

# TEORIA DA RELATIVIDADE RESTRITA

Júlio César Fabris e Antônio Brasil Batista

Departamento de Física

Universidade Federal do Espírito Santo

Goiabeiras - Vitória - CEP29060-900

Espírito Santo

2017

# 1 Mecânica e eletromagnetismo

A Física Clássica se alicerça em duas teorias fundamentais: a mecânica newtoniana e o eletromagnetismo de Maxwell. Se nos limitamos aos fenômenos envolvendo corpos macroscópicos, não excessivamente massivos <sup>1</sup>, que se deslocam com velocidades ordinárias, significativamente inferiores à velocidade da luz, a mecânica newtoniana fornece resultados em acordo com a experiência e a observação. Isto se aplica em particular ao movimento de objetos sobre a superfície terrestre assim como para o movimento dos planetas. Por sua vez, o eletromagnetismo é uma teoria igualmente bem comprovada, que teve um dos seus sucessos mais impressionantes na identificação da luz como onda eletromagnética e na consequente predição da existência de ondas eletromagnéticas em frequências fora da faixa correspondente à luz visível. Faremos a seguir uma breve revisão destas duas teorias, enfatizando sobretudo os aspectos que serão relevantes para a posterior construção da teoria da relatividade restrita.

## 2 A mecânica newtoniana

### 2.1 As leis de Newton

A mecânica newtoniana baseia-se nas três leis de Newton e, conseqüentemente, está fortemente alicerçada sobre a noção de referencial inercial. Podemos enunciar as três leis da seguinte forma:

1. Todo corpo permanece em repouso ou em estado de movimento retilíneo uniforme se nenhuma força age sobre ele ou se a soma vetorial de todas as forças que agem sobre o corpo é nula;
2. A ação de uma força sobre um corpo resultará na aceleração deste corpo, a razão entre os módulos da força e da aceleração sendo igual à massa inercial do corpo;
3. A toda ação de um corpo  $A$  sobre um corpo  $B$  corresponde uma reação

---

<sup>1</sup>A condição "não excessivamente massivo" pode ser expressa estabelecendo que a grandeza adimensional  $\frac{GM}{c^2}$  satisfaz a condição  $\frac{GM}{c^2} \ll 1$ .

de mesma intensidade, e de sentido contrário, que agirá sobre o corpo que provoca a ação, neste caso o próprio corpo  $A$ .

A estas leis, podemos acrescentar a lei da Gravitação Universal:

1. Dois corpos se atraem proporcionalmente ao inverso do quadrado da distância que os separa e em proporção direta ao produto de suas massas gravitacionais.

No enunciado destas leis, introduzimos as noções de massas gravitacional e inercial, a primeira estando relacionada ao fato que existe uma interação gravitacional (constituindo assim uma "carga gravitacional": quanto maior a massa gravitacional, maior a força gravitacional que será exercida), a segunda se referindo à resistência apresentada por um corpo quando sobre ele age uma força (quanto maior a massa inercial de um corpo, menor é a aceleração transmitida a este corpo por uma dada força). A experiência mostra que as duas massas são iguais com uma precisão de  $10^{-12}$ [?]. Na teoria da relatividade geral, que será abordada no final deste livro, as duas massas são consideradas exatamente iguais.

A primeira lei de Newton define os "bons" referenciais para os quais as demais leis da Física devem ser escritas. Estes bons referenciais são denominados de "inerciais". Referencial inercial é aquele para o qual a primeira lei de Newton se aplica: neste referencial, se a força resultante sobre um corpo é nula, então a velocidade deste corpo é constante. A segunda lei, por outro lado, diz qual será o efeito de uma força medido por um observador situado naquele "bom" referencial, o referencial "inercial". Neste sentido a primeira lei é pré-requisito indispensável para a segunda lei: primeiro se define quem observa (aquele para quem, na ausência de uma força, um corpo deve ficar em repouso ou em movimento retilíneo uniforme), para depois se dizer qual é a consequência da ação de uma força (ela provoca aceleração na proporção inversa de sua massa inercial). A terceira lei está relacionada à conservação do momento linear. A terceira lei determina também quando uma dada força é devida a uma interação física: à sua *ação*, podemos determinar qual é a *reação*. Obviamente, podemos tratar um fenômeno físico através de um referencial não inercial. Mas, neste caso, a aceleração observada não estará forçosamente relacionada às interações

físicas: forças fictícias, ditas inerciais, que não verificam a terceira lei (visto que, para elas, é impossível determinar qual é a reação <sup>2</sup>).

Para tratarmos quantitativamente estas leis, definimos um referencial (inercial), e denotamos a distância de um dado corpo à origem deste referencial pelo vetor  $\vec{r}$ . A velocidade deste corpo é então definida como a taxa de variação da posição com o tempo  $\vec{v} = \frac{d}{dt}\vec{r}$  e a aceleração é definida como a taxa de variação da velocidade com o tempo, ou ainda como a segunda derivada da posição em relação ao tempo,  $\vec{a} = \frac{d}{dt}\vec{v} = \frac{d^2}{dt^2}\vec{r}$ . Desta maneira, a segunda lei se escreve como

$$\vec{F} = m_I \vec{a} = m_I \frac{d}{dt}\vec{v} = m_I \frac{d^2}{dt^2}\vec{r} \quad , \quad (1)$$

onde  $m_I$  designa a massa inercial. Se definirmos uma nova quantidade auxiliar denominada de *momento linear*, expressa como

$$\vec{p} = m_I \vec{v} \quad , \quad (2)$$

então a segunda lei de Newton assume a forma

$$\vec{F} = \frac{d}{dt}\vec{p} \quad . \quad (3)$$

Observamos que o termo "quantidade auxiliar" empregado acima corresponde a uma grandeza sem sentido físico intrínseco, podendo seu valor, por exemplo, depender do referencial onde é medido. Desta forma, uma grandeza auxiliar se distingue de grandezas - como força e aceleração - que têm o mesmo valor em todos os referenciais inerciais, assumindo conseqüentemente a condição de invariantes da teoria.

As definições (1,3) coincidem quando a massa de um corpo é constante. Entretanto, elas diferem quando a massa do corpo é variável, como no problema do foguete cuja massa varia em consequência da ejeção de gases. A definição fundamental da segunda lei de Newton é dada por (3).

A independência da primeira e da segunda leis de Newton, no sentido que a primeira não determina a segunda e nem é um caso particular dela, se reflete

---

<sup>2</sup>Mach propôs de atribuir um sentido físico às forças inerciais ao considerar que a reação a este tipo de força ocorre sobre a distribuição global de matéria no Universo. Isto se enquadra na tentativa de Mach de interpretar a inércia como a interação entre um corpo e toda a matéria restante no Universo [?]. Este conceito foi retomado posteriormente por Feynmann e Wheeler, no âmbito da teoria quântica, através do conceito de absorvedores [?].

no fato que, admitida a primeira lei, seria possível supor que o efeito da força poderia ser diferente do produto simples entre a massa e a aceleração. Um exemplo disto é a teoria MOND (acrônimo para *Modified Newtonian Dynamics*), proposta recentemente [?] visando explicar a anomalia observada na curva de rotação das galáxias espirais - ver apêndice 1 para uma descrição do conteúdo e das motivações para esta teoria. Na teoria MOND, a segunda lei de Newton é substituída pela expressão  $\vec{F} = m\mu(a/a_0)\vec{a}$ , onde  $a_0$  é uma aceleração dita crítica, e  $\mu(x)$  é uma função tal que  $\mu(x) \sim 1$  se  $x \gg 1$  e  $\mu(x) \sim x$  se  $x \ll 1$ . Neste caso, a primeira lei permanece inalterada mas a relação entre força e aceleração é diferente da expressão correspondente da teoria newtoniana.

## 2.2 As transformações de Galileu

Dado um referencial inercial  $S$ , qualquer outro referencial  $S'$  que se desloca em relação a  $S$  com velocidade constante será também inercial. A forma da segunda lei de Newton é preservada quando passamos do referencial inercial  $S$  para este outro referencial inercial  $S'$ . Designemos a velocidade (constante) de  $S'$ , medida por um observador em  $S$ , como  $\vec{V}$ . Esta velocidade relativa, a título de simplificação, será considerada como orientada ao longo do eixo  $x$ . Para simplificar, supomos que no tempo  $t = 0$  as duas origens coincidam, assim como os respectivos eixos. Neste caso as relações entre as coordenadas  $x, y, z$  e  $x', y'$  e  $z'$  são,

$$x' = x - Vt \quad , \quad (4)$$

$$y' = y \quad , \quad (5)$$

$$z' = z \quad . \quad (6)$$

Observamos que estas transformações, conhecidas como transformações de Galileu, pressupõem um tempo universal:

$$t = t' \quad . \quad (7)$$

Temos assim:

$$v'_x = \frac{dx'}{dt} = v_x - V \quad ; \quad a'_x = \frac{dv'_x}{dt} = a_x \quad ; \quad (8)$$

$$v'_y = \frac{dy'}{dt} = v_y \quad ; \quad a'_y = \frac{dv'_y}{dt} = a_y \quad ; \quad (9)$$

$$v'_z = \frac{dz'}{dt} = v_z \quad ; \quad a'_z = \frac{dv'_z}{dt} = a_z \quad . \quad (10)$$

Como consequência temos que,

$$\vec{F}' = m_I \vec{a}' = m_I \vec{a} = \vec{F} \quad . \quad (11)$$

Desta forma, a força medida no referencial  $S'$ , assim como a aceleração sofrida pelo corpo, é a mesma que a medida no referencial  $S$ : as leis de Newton são portanto invariantes em relação às transformações de Galileu. Estamos supondo a massa inercial do corpo como sendo uma propriedade intrínseca do corpo, logo independente do observador. Note-se também que, frequentemente, utilizaremos os termos *referencial* e *observador* como equivalentes: todo observador pressupõe um referencial e vice-versa.

Como a lei da Gravitação Universal está baseada na posição relativa entre dois corpos, sua expressão também não é afetada por esta mudança de sistema de referência. De fato, na mecânica, todas as leis de força são baseadas seja na posição relativa, seja na velocidade relativa entre dois corpos (como a força de Lorentz entre duas cargas em movimento), e permanecem, em consequência, invariantes pelas transformações de Galileu.

## 2.3 Leis de conservação

Quando uma força é tal que a integral do produto interno desta força com o vetor deslocamento do corpo sobre o qual ela age, ao longo de um percurso fechado  $l$ , é nula,

$$\oint \vec{F} \cdot d\vec{r} = 0 \quad (12)$$

então esta força é dita "conservativa". Isto implica que o trabalho realizado por esta força quando o corpo se desloca de uma posição inicial  $\vec{r}_i$  para uma posição final  $\vec{r}_f$  independe do caminho empregado neste trajeto. Neste caso, a força pode ser obtida como gradiente de um campo escalar:  $\vec{F} = -\nabla\phi(\vec{r})$ . De fato, usando o teorema de Stokes (ver apêndice), temos

$$\oint \vec{F} \cdot d\vec{r} = \int_S \nabla \times \vec{F} \cdot d\vec{S} = 0 \quad , \quad (13)$$

onde  $S$  é a superfície cuja borda é definida pela curva fechada  $l$ . Como esta expressão é válida para qualquer curva fechada, o integrando deverá ser identicamente nulo para qualquer superfície  $S$ . Isto pode ser obtido se  $\vec{F}$  for o gradiente de uma função escalar. Logo, a expressão (13) implica que, para uma força conservativa,

$$\vec{F} = -\nabla\phi \quad , \quad (14)$$

onde o sinal negativo foi introduzido por comodidade futura. Este campo escalar  $\phi(\vec{r})$  é denominado de energia potencial, sendo uma grandeza auxiliar que, justamente por ser um escalar, facilita em geral o cálculo de certas quantidades físicas. Vamos, para o que se segue, suprimir o índice  $I$  que é utilizado para indicar a massa inercial. Este índice será restabelecido quando necessário. Para um corpo sob ação de uma força conservativa, cujo potencial automaticamente não pode depender explicitamente do tempo, a energia mecânica, definida como

$$E = \frac{1}{2}mv^2 + \phi(\vec{r}) \quad , \quad (15)$$

se conserva.

De fato, podemos, no caso de uma força conservativa, escrever a segunda lei de Newton como

$$m \frac{d}{dt} \vec{v} = -\nabla\phi \quad . \quad (16)$$

Tomando o produto interno com  $\vec{v}$ , utilizando a definição da velocidade como derivada da posição e a noção de derivada implícita, encontramos,

$$\begin{aligned} m\vec{v} \cdot \frac{d}{dt} \vec{v} &= -\vec{v} \cdot \nabla\phi = -\sum_{i=1}^3 \frac{dx_i}{dt} \cdot \frac{\partial\phi}{\partial x_i} \\ \frac{1}{2}m \frac{d}{dt} v^2 &= -\frac{d}{dt} \phi \\ \frac{d}{dt} \left( \frac{1}{2}mv^2 + \phi \right) &= 0 \\ \frac{1}{2}mv^2 + \phi = E &= \text{constante} \quad . \end{aligned}$$

Nesta dedução, utilizamos o fato que o vetor  $\vec{r}$  indica a posição da partícula, sendo uma função do tempo:  $\vec{r} = \vec{r}(t)$ . Além do mais, a energia potencial, por estar relacionada a uma força conservativa, não depende explicitamente do tempo, mas apenas da posição da partícula. Chamamos  $E$  de *energia mecânica*

da *partícula*, uma quantidade que se conserva quando a força que age sobre esta partícula é conservativa.

Por outro lado, a conservação do momento linear de um sistema de partículas que interagem entre si pode ser obtido da terceira lei de Newton. Para simplificar, consideremos um sistema de duas partículas. A terceira lei de Newton diz que

$$\begin{aligned}\vec{F}_{21} = -\vec{F}_{12} &\Rightarrow \frac{d}{dt}\vec{p}_1 = -\frac{d}{dt}\vec{p}_2 \\ \frac{d}{dt}(\vec{p}_1 + \vec{p}_2) = 0 &\Rightarrow \vec{p}_1 + \vec{p}_2 = \vec{P} = \text{constante} \quad .\end{aligned}$$

Nestas expressões foi usado o fato que  $\vec{F}_{ij}$ , força que a partícula  $i$  exerce sobre a partícula  $j$ , é igual à variação do momento linear da partícula  $j$ . A generalização da expressão para a conservação do momento linear para um sistema fechado de  $n$  partículas (sistema fechado é aquele sobre o qual não existem forças externas ao sistema agindo sobre seus constituintes) é proposta nos exercícios.

As leis de conservação da energia e do momento linear representam alguns dos principais alicerces da mecânica newtoniana. Essas leis são complementadas pela a lei da conservação do momento angular, que é verificada quando a força de interação entre qualquer par de partículas está na direção da reta que as une. O momento angular de uma dada partícula se define como

$$\vec{L} = m\vec{r} \times \vec{v} = \vec{r} \times \vec{p} \quad , \quad (17)$$

onde  $\vec{r}$  é o vetor posição da partícula e  $\vec{v}$  sua velocidade. Para demonstrar que, sob a condição explicitada acima, o momento angular se conserva, vamos considerar um sistema de duas partículas, cujas quantidades correspondentes serão designadas pelos subscritos 1 e 2. A terceira lei de Newton é sempre satisfeita, evidentemente. O momento angular total deste sistema de duas partículas é dado por

$$\vec{L}_T = \vec{L}_1 + \vec{L}_2 = \vec{r}_1 \times \vec{p}_1 + \vec{r}_2 \times \vec{p}_2 \quad . \quad (18)$$

A taxa de variação do momento angular total é

$$\begin{aligned}\frac{d}{dt}\vec{L}_T &= \vec{v}_1 \times \vec{p}_1 + \vec{r}_1 \times \frac{d}{dt}\vec{p}_1 + \vec{v}_2 \times \vec{p}_2 + \vec{r}_2 \times \frac{d}{dt}\vec{p}_2 \\ &= \vec{r}_1 \times \vec{F}_{21} + \vec{r}_2 \times \vec{F}_{12}\end{aligned}$$



$$\begin{aligned}
&= (\vec{r}_1 - \vec{r}_2) \times \vec{F}_{21} \\
&= 0 \quad .
\end{aligned}$$

Nesta dedução utilizamos que  $\vec{F}_{21} = -\vec{F}_{12}$  (terceira lei de Newton), onde  $\vec{r}_1 - \vec{r}_2$  é o vetor de posição da primeira partícula em relação à segunda e, finalmente, que a força está orientada ao longo desta posição relativa. Logo, o momento angular do sistema de partículas se conserva quando não existem forças externas agindo sobre o sistema, e as forças entre as partículas do sistema estão orientadas na direção do vetor posição relativa entre as partículas. No caso onde a força depende da velocidade relativa entre as partículas, a dedução da conservação do momento angular é mais sutil. Uma situação onde a força depende da velocidade ocorre no eletromagnetismo, como por exemplo na força de Lorentz. Neste caso, é preciso levar em conta também o momento transportado pelo campo eletromagnético. Esta questão será analisada mais tarde.

A existência de leis conservação reflete a presença de simetrias no sistema físico. A conservação da energia, por exemplo, está relacionada à invariância das leis da física por translações temporais: se as condições do sistema são as mesmas, pouco importa se fazemos uma dada experiência hoje ou amanhã, o resultado será o mesmo. Por outro lado, a conservação do momento linear reflete a invariância dos sistemas físicos por translações espaciais: mantendo-se as mesmas condições no sistema, o resultado de uma experiência é independente de onde ela é feita. Por fim, a conservação do momento angular está relacionada com a invariância do sistema por rotação. Demonstrações mais gerais e mais rigorosas das leis de conservação podem ser encontradas nas referências [?, ?, ?].

## 3 O eletromagnetismo

### 3.1 As equações de Maxwell

A teoria eletromagnética é representada matematicamente pelas equações de Maxwell que, na forma integral, são escritas como se segue:

$$\oint \vec{E} \cdot d\vec{S} = \frac{q}{\epsilon_0} \quad , \quad (19)$$

$$\oint \vec{B} \cdot d\vec{S} = 0 \quad , \quad (20)$$

$$\oint \vec{E} \cdot d\vec{l} = -\frac{d}{dt} \int B \cdot d\vec{S} \quad , \quad (21)$$

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 i + \frac{d}{dt} \int E \cdot d\vec{S} \quad , \quad (22)$$

onde  $d\vec{S}$  e  $d\vec{l}$  representam os elementos de integração sobre uma superfície e uma linha, respectivamente. Nestas equações  $\vec{E}$  e  $\vec{B}$  são, respectivamente, os campos elétrico e magnético,  $q$  e  $i$  são a carga e a corrente que criam aqueles campos, enquanto  $\epsilon_0$  e  $\mu_0$  são as constantes de permissividade elétrica e permeabilidade magnética do vácuo.

A equação (19) é a lei de Gauss (ou equivalentemente, lei de Coulomb) para o campo elétrico; a equação (20) estabelece a inexistência de monopolos (cargas) magnéticos; a equação (21) é a lei de Faraday, que indica que a variação no fluxo de campo magnético cria campo elétrico; por fim, a equação (22) é a lei de Ampère, suplementada pelo termo de corrente de deslocamento, que estabelece que correntes elétricas, assim como a variação no fluxo de campo elétrico, criam campo magnético.

Os teoremas de Gauss e Stokes estabelecem que, para um dado campo vetorial  $\vec{A}$ , as seguintes relações são verdadeiras:

$$\oint A \cdot d\vec{S} = \int_V \nabla \cdot \vec{A} dV \quad , \quad (23)$$

$$\oint A \cdot d\vec{l} = \int_S \nabla \times \vec{A} \cdot d\vec{S} \quad . \quad (24)$$

Sendo  $\rho$  a densidade de carga e  $\vec{j}$  a densidade corrente, a carga total contida no interior de uma superfície fechada  $S$  e a corrente  $i$  que atravessa uma dada superfície delimitada por um percurso fechado  $l$ , são expressas como,

$$q = \int_V \rho dV \quad , \quad (25)$$

$$i = \int_S \vec{j} \cdot d\vec{S} \quad . \quad (26)$$

Lembramos que a densidade de corrente  $\vec{j}$  se escreve como  $\vec{j} = \rho \vec{v}$ , onde  $\vec{v}$  é o campo de velocidade das cargas. De posse destas informações, podemos re-escrever as equações de Maxwell na sua forma diferencial:

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad ; \quad (27)$$

$$\nabla \cdot \vec{B} = 0 ; \quad (28)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} ; \quad (29)$$

$$\nabla \times \vec{B} = \mu_0 \vec{j} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} . \quad (30)$$

As equações (28,29) são chamadas de equações homogêneas de Maxwell porque, nelas, não aparecem os termos de fontes, representados pelas densidades de carga e corrente. Já as equações (27,30) contêm estes termos de fontes e por isto são chamadas de equações de Maxwell não homogêneas.

As quantidades  $\rho$  e  $\vec{j}$  estão ligadas por uma equação de conservação:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0 \quad (31)$$

que pode ser obtida a partir das equações de Maxwell tomando-se a derivada parcial em relação ao tempo de (27) e combinando com o divergente de (30).

### 3.2 Ondas eletromagnéticas

A partir das equações de Maxwell, nós podemos deduzir uma equação de onda. Podemos escrever a equação de onda para o campo elétrico tomando-se o rotacional de (29) e utilizando-se a relação,

$$\nabla \times \nabla \times \vec{E} = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = \frac{\nabla \rho}{\epsilon_0} - \nabla^2 \vec{E} . \quad (32)$$

Temos assim, empregando ainda a derivada temporal de (30),

$$\nabla^2 \vec{E} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} = \mu_0 \frac{\partial \vec{j}}{\partial t} + \frac{\nabla \rho}{\epsilon_0} . \quad (33)$$

Utilizando a lei de Ohm,  $\vec{j} = \sigma \vec{E}$ , onde  $\sigma$  é a condutividade do meio, esta equação pode ser re-escrita como

$$\nabla^2 \vec{E} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} = \mu_0 \sigma \frac{\partial \vec{E}}{\partial t} + \frac{\nabla \rho}{\epsilon_0} . \quad (34)$$

Da mesma forma, tomando-se o rotacional de (30) e a derivada em relação ao tempo de (29), nós obtemos a seguinte equação para o campo  $\vec{B}$ :

$$\nabla^2 \vec{B} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} = -\mu_0 \nabla \times \vec{j} . \quad (35)$$

Esta equação pode ser re-escrita, usando novamente a lei de Ohm e agora também a lei de Faraday, como

$$\nabla^2 \vec{B} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} = \mu_0 \sigma \frac{\partial \vec{B}}{\partial t} \quad . \quad (36)$$

Os termos contendo primeiras derivadas representam um efeito dissipativo devido às propriedades do meio.

No vácuo, estas duas equações se escrevem:

$$\square \vec{E} = 0 \quad , \quad (37)$$

$$\square \vec{B} = 0 \quad , \quad (38)$$

onde definimos o operador D'Alambertiano

$$\square = \frac{\partial^2}{\partial t^2} - \frac{1}{\mu_0 \epsilon_0} \nabla^2 \quad . \quad (39)$$

Estas expressões representam a equação de uma onda que se propaga com a velocidade

$$c^2 = \frac{1}{\mu_0 \epsilon_0} \quad . \quad (40)$$

As constantes  $\mu_0$  e  $\epsilon_0$  têm seus valores medidos experimentalmente<sup>3</sup>:  $\mu_0 = 1,26 \times 10^{-6} \frac{H}{m}$  e  $\epsilon_0 = 8,85 \times 10^{-12} \frac{F}{m}$ . Substituindo os valores obtidos na expressão acima, nós obtemos,

$$c \sim 3 \times 10^{10} \frac{cm}{s} \quad . \quad (41)$$

Esta é a velocidade de propagação de uma onda eletromagnética, que se revela igual ao valor conhecido para a velocidade de propagação da luz. Este resultado permitiu identificar a luz como sendo uma onda eletromagnética, e foi um dos maiores sucessos da teoria Eletromagnética de Maxwell.

A existência das ondas eletromagnéticas revela que o conceito de *campo* possui uma realidade, não sendo apenas uma definição formal. Os campos *existem* e a onda eletromagnética é a consubstanciação deste fato: a onda eletromagnética reflete o fato que os campos se induzem reciprocamente, o campo elétrico variável criando campo magnético e vice-versa. A realidade da onda

---

<sup>3</sup>As unidades  $H$  (henry) e  $F$  (farad) se definem como  $H = \frac{J}{A}$  e  $F = \frac{A^2 s^2}{J}$  onde  $J$  e  $A$  designam as unidades de energia (joule) e corrente (ampère), respectivamente.

eletromagnética fica evidente devido ao fato que a onda transporta momento e energia, o que é expresso pelo vetor de Poynting [?, ?],

$$\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B} \quad , \quad (42)$$

que indica a quantidade de energia por unidade de tempo e unidade de área transportada pela onda eletromagnética. O fato que o eletromagnetismo de Maxwell faz uso dos conceitos de campo, atribuindo uma realidade a estes novos objetos, faz dele a primeira teoria de campo coerentemente formulada. O programa de *teoria de campos* será amplamente explorado, a partir do século XX, tanto a nível clássico quanto quântico, sendo um dos fundamentos da física moderna.

### 3.3 Os potenciais eletromagnéticos

Outra formulação do eletromagnetismo, equivalente à descrita acima, pode ser obtida expressando os campos em termos de potenciais. A partir de (28), podemos escrever o campo magnético como o rotacional de um vetor:

$$\vec{B} = \nabla \times \vec{A} \quad . \quad (43)$$

Substituindo esta relação em (29), obtemos

$$\nabla \times \left( \vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0 \quad , \quad (44)$$

o que nos leva a escrever a quantidade entre parênteses como o gradiente de um campo escalar, resultando na expressão,

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \quad . \quad (45)$$

Podemos substituir agora estas expressões para  $\vec{E}$  e  $\vec{B}$  nas equações (27,30). Obtemos,

$$-\nabla^2\phi - \frac{\partial}{\partial t} \nabla \cdot \vec{A} = \frac{\rho}{\epsilon_0} \quad , \quad (46)$$

$$\nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A} = \mu_0 \vec{j} + \mu_0 \epsilon_0 \left[ -\frac{\partial}{\partial t} \nabla\phi - \frac{\partial^2}{\partial t^2} \vec{A} \right] \quad . \quad (47)$$

Estas são as equações de Maxwell não homogêneas escritas em termos dos potenciais. Lembre-se que as equações de Maxwell homogêneas conduzem às definições dos potenciais. As equações escritas em termos dos potenciais parecem

mais complicadas. Mas, isto é apenas aparente, como veremos a seguir. Na verdade, os potenciais  $\phi$  e  $\vec{A}$  permitem simplificar a obtenção dos campos  $\vec{E}$  e  $\vec{B}$  para uma dada distribuição de cargas e correntes.

### 3.4 As transformações de calibre

Os potenciais escalar e vetor  $\phi$  e  $\vec{A}$  não são únicos. As expressões para os campos  $\vec{E}$  e  $\vec{B}$  permanecem inalteradas se fizermos as redefinições

$$\vec{A}' = \vec{A} + \nabla\chi, \quad (48)$$

$$\phi' = \phi - \frac{\partial}{\partial t}\chi, \quad (49)$$

onde  $\chi$  é uma função arbitrária da posição e do tempo,  $\chi = \chi(\vec{r}, t)$ . Isto é consequência do fato que os campos  $\vec{E}$  e  $\vec{B}$  são as quantidades físicas mensuráveis enquanto que os potenciais  $\phi$  e  $\vec{A}$  são grandezas auxiliares não tendo, a nível puramente clássico, realidade física<sup>4</sup>. As transformações (465,466) são denominadas transformações de calibre, e deixam invariante as equações de Maxwell, visto que  $\vec{E}$  e  $\vec{B}$  permanecem os mesmos. Observe-se que tratam-se de transformações distintas daquelas que envolvem passagem de um referencial inercial a outro, que implicam mudanças de coordenadas: as transformações de calibre referem-se aos potenciais e não aos referenciais.

Utilizando a invariância resultante das transformações (465,466) podemos impor uma condição sobre os potenciais, visto que eles são definidos a menos de uma função arbitrária  $\chi$ . Esta função arbitrária pode ser escolhida de maneira a eliminar uma das componentes do conjunto dos potenciais  $\phi$  e  $\vec{A}$ , ou de forma a estabelecer uma relação entre elas. Uma condição bastante útil é a denominada condição de Lorentz, que se escreve <sup>5</sup>

$$\nabla \cdot \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \phi = 0. \quad (50)$$

---

<sup>4</sup>Isto é inteiramente análogo aos potenciais a partir dos quais se derivam as leis de força, quando conservativas, na Mecânica newtoniana. Entretanto, este carácter puramente auxiliar dos potenciais eletromagnéticos não é preservado a nível quântico, como indica o efeito Aharonov-Bohm, ver por exemplo J.J. Sakurai, **Modern quantum mechanics**, Addison-Wesley, Nova Iorque (1994).

<sup>5</sup>O fator  $c$  nesta expressão é introduzido de maneira a torná-la homogênea dimensionalmente, visto que  $[A]/T = [\phi]/L \rightarrow [A] = (T/L)[\phi]$ , onde os colchetes designam a dimensão das quantidades em questão.

Podemos entender a possibilidade de se impor esta condição considerando inicialmente o caso onde (50) não é satisfeita. Utilizando as transformações (465,466), obtemos

$$\nabla \cdot \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \phi = \nabla \cdot \vec{A}' + \frac{1}{c^2} \frac{\partial \phi'}{\partial t} + \nabla^2 \chi - \frac{1}{c^2} \frac{\partial^2 \chi}{\partial t^2}. \quad (51)$$

A função  $\chi$  é arbitrária. Logo, podemos escolhê-la de tal forma que

$$\nabla \cdot \vec{A}' + \frac{1}{c^2} \frac{\partial \phi'}{\partial t} = -\nabla^2 \chi + \frac{1}{c^2} \frac{\partial^2 \chi}{\partial t^2} \quad (52)$$

obtendo (50).

Usando a condição de Lorentz, as equações (46,47) se escrevem,

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \phi - \nabla^2 \phi = \frac{\rho}{\epsilon_0}, \quad (53)$$

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} - \nabla^2 \vec{A} = \mu_0 \vec{j}. \quad (54)$$

No vácuo, essas equações se reduzem a,

$$\square \phi = 0, \quad (55)$$

$$\square \vec{A} = 0. \quad (56)$$

Essas são as equações de onda para os potenciais  $\phi$  e  $\vec{A}$ .

A imposição de um calibre deve em princípio fixar os potenciais  $\vec{A}$  e  $\phi$ . No caso do calibre de Lorentz expresso pela equação (50), e obtido da forma descrita acima, os potenciais não estão, entretanto, ainda completamente fixados: existe uma liberdade residual na definição dos calibres, representada pelas transformações

$$A \rightarrow A + \nabla \chi, \quad \phi \rightarrow \phi - \frac{\partial \chi}{\partial t}, \quad (57)$$

que conecta diferentes conjuntos de potenciais  $(\phi, \vec{A})$  que satisfazem (50), desde que

$$\square \chi = 0, \quad (58)$$

como pode ser verificado a partir de (52). Observe que, agora,  $\chi$ , ao contrário do caso anterior, não é uma função completamente arbitrária, mas deve satisfazer a equação de onda homogênea (58). Uma transformação do tipo (57), com o campo  $\chi$  satisfazendo a condição (58), não apenas mantém invariante as equações de Maxwell, mas também preserva a condição de Lorentz (50). Isto

está relacionado com o fato que as ondas eletromagnéticas são ondas transversais com duas direções de polarização possíveis, contendo portanto dois graus de liberdade. Observe, assim, que a fixação do calibre de Lorentz implica em uma família de potenciais, conectadas pelas transformações (57) suplementadas pela condição (58).

A invariância por transformações de calibre implica mais do que a liberdade para se impor condições sobre os potenciais. Na verdade, é necessário utilizar esta propriedade de invariância pela redefinição de potenciais, do contrário o sistema conterà graus de liberdade espúrios, no sentido de serem não físicos. A escolha de um calibre (não necessariamente o de Lorentz) é uma necessidade ditada pelo fato que se tem mais variáveis do que equações, visto que a equação da continuidade, por exemplo, implica uma relação entre as equações de Maxwell. Do ponto de vista matemático, é preciso impor uma condição de calibre para que as soluções sejam únicas.

## 4 O eletromagnetismo e as transformações de Galileu

Vimos na seção (2.2) que as equações resultantes das leis de Newton são invariantes pelas transformações de Galileu. Uma vez que estas transformações expressam como mudam as relações entre as grandezas cinemáticas quando se passa de um referencial inercial a outro, é imprescindível analisar as implicações das transformações de Galileu para as equações de Maxwell. Uma transformação geral de coordenadas, passando do sistema  $x, y, z, t$  para o sistema  $x', y', z', t'$  se escreve,

$$x' = x'(x, y, z, t), y' = y'(x, y, z, t), z' = z'(x, y, z, t), t' = t'(x, y, z, t) \quad (59)$$

onde incluímos a transformação para a coordenada temporal de maneira a preparar o terreno para uma classe de transformações mais geral que as de Galileu. Sob uma transformação do tipo (59), as derivadas parciais se transformam como,

$$\frac{\partial}{\partial x} = \frac{\partial x'}{\partial x} \frac{\partial}{\partial x'} + \frac{\partial y'}{\partial x} \frac{\partial}{\partial y'} + \frac{\partial z'}{\partial x} \frac{\partial}{\partial z'} + \frac{\partial t'}{\partial x} \frac{\partial}{\partial t'} \quad , \quad (60)$$



$$\frac{\partial}{\partial y} = \frac{\partial x'}{\partial y} \frac{\partial}{\partial x'} + \frac{\partial y'}{\partial y} \frac{\partial}{\partial y'} + \frac{\partial z'}{\partial y} \frac{\partial}{\partial z'} + \frac{\partial t'}{\partial y} \frac{\partial}{\partial t'} , \quad (61)$$

$$\frac{\partial}{\partial z} = \frac{\partial x'}{\partial z} \frac{\partial}{\partial x'} + \frac{\partial y'}{\partial z} \frac{\partial}{\partial y'} + \frac{\partial z'}{\partial z} \frac{\partial}{\partial z'} + \frac{\partial t'}{\partial z} \frac{\partial}{\partial t'} , \quad (62)$$

$$\frac{\partial}{\partial t} = \frac{\partial x'}{\partial t} \frac{\partial}{\partial x'} + \frac{\partial y'}{\partial t} \frac{\partial}{\partial y'} + \frac{\partial z'}{\partial t} \frac{\partial}{\partial z'} + \frac{\partial t'}{\partial t} \frac{\partial}{\partial t'} . \quad (63)$$

Considerando agora o caso particular das transformações de Galileu (4, 5,6,7), obtemos

$$\nabla = \nabla' , \quad \frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - \vec{V} \cdot \nabla . \quad (64)$$

Observe que a derivada parcial em relação ao tempo não é invariante pelas transformações de Galileu. Isto se deve ao fato que estas transformações envolvem explicitamente o tempo. Mas, é imediato observar que a derivada total em relação ao tempo permanece invariante, como se deveria esperar:

$$\frac{d}{dt} = \frac{d}{dt'} . \quad (65)$$

De fato, a derivada total em relação ao tempo pode ser escrita como

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \vec{v} \cdot \nabla . \quad (66)$$

Usando as transformações acima e empregando a lei da adição das velocidades, obtemos (65).

Em relação aos termos de fonte, densidades de carga e corrente, as transformações de Galileu conduzem a

$$\rho' = \rho , \quad \vec{j}' = \rho \vec{v} = \rho(\vec{v}' + \vec{V}) = \vec{j}' + \epsilon_0 \vec{V} \nabla \cdot \vec{E} , \quad (67)$$

onde utilizamos o fato que  $\rho$  é uma função escalar e que  $\nabla \cdot \vec{E} = \rho/\epsilon_0$ . Aplicando-se estas expressões nas equações de Maxwell, obtemos

$$\nabla' \cdot \vec{E} = \frac{\rho}{\epsilon_0} , \quad (68)$$

$$\nabla' \cdot \vec{B} = 0 , \quad (69)$$

$$\nabla' \times \vec{E} = -\frac{\partial \vec{B}}{\partial t'} + (\vec{V} \cdot \nabla') \vec{B} , \quad (70)$$

$$\nabla' \times \vec{B} = \mu_0 \vec{j}' + \mu_0 \epsilon_0 \left[ \frac{\partial \vec{E}}{\partial t'} + \vec{V} \nabla \cdot \vec{E} - (\vec{V} \cdot \nabla') \vec{E} \right] . \quad (71)$$

Obviamente, as "novas" equações de Maxwell diferem das equações originais.

Entretanto, temos ainda a liberdade de redefinir os campos  $\vec{E}$  e  $\vec{B}$ . Afinal, um observador em repouso em relação a uma carga elétrica identifica unicamente a existência de um campo elétrico, ao passo que um observador em movimento em relação a esta mesma carga identifica a existência de campos elétrico e magnético. Portanto, as transformações de Galileu, aplicadas a um sistema de cargas e correntes, devem fazer "aparecer" ou "desaparecer" componentes dos campos elétrico ou magnético. Em vista disto, é natural admitir que as expressões que relacionam os campos elétrico e magnético medidos por diferentes observadores inerciais dependam da velocidade relativa entre estes observadores. Mas, como determinar as relações conectando os campos em diferentes referenciais? Lembramos inicialmente que as dimensões dos campos elétrico e magnético no sistema SI se relacionam como  $[E] = \frac{L}{T}[B]$ , onde  $L$  e  $T$  designam as dimensões espaciais e temporais respectivamente. Podemos então, por argumentos dimensionais e em função da situação descrita anteriormente sobre a medida dos campos criados por uma carga elétrica feita por dois observadores inerciais diferentes, escrever os campos elétrico e magnético no referencial  $S$  em termos dos campos medidos no referencial  $S'$  como

$$\vec{E} = \vec{E}' + a\vec{V} \times \vec{B}' \quad , \quad \vec{B} = \vec{B}' - \frac{b}{c^2}\vec{V} \times \vec{E}' \quad , \quad (72)$$

onde  $a$  e  $b$  são constantes adimensionais a serem determinadas ulteriormente. Estas seriam as relações mais simples que visam reobter as equações de Maxwell reescritas inteiramente em termos das novas coordenadas  $x', y', z'$  e dos novos campos  $\vec{E}'$  e  $\vec{B}'$ .

Utilizamos, agora, as relações vetoriais,

$$\nabla \times (\vec{V} \times \vec{A}) = -(\vec{V} \cdot \nabla)\vec{A} + \vec{V}(\nabla \cdot \vec{A}) - \vec{A}(\nabla \cdot \vec{V}) + (\vec{A} \cdot \nabla)\vec{V} \quad , \quad (73)$$

$$\nabla \cdot (\vec{V} \times \vec{A}) = -\vec{V} \cdot (\nabla \times \vec{A}) + \vec{A} \cdot (\nabla \times \vec{V}) \quad . \quad (74)$$

Aplicadas aos campos  $\vec{E}$  e  $\vec{B}$ , e lembrando que  $\vec{V}$  é constante, aquelas relações nos fornecem,

$$\nabla \cdot \vec{E} = \nabla \cdot \vec{E}' - a\vec{V} \cdot (\nabla \times \vec{B}') \quad , \quad (75)$$

$$\nabla \cdot \vec{B} = \nabla \cdot \vec{B}' + \frac{b}{c^2}\vec{V} \cdot (\nabla \times \vec{E}') \quad , \quad (76)$$

$$\nabla \times \vec{E} = \nabla \times \vec{E}' - a\left[(\vec{V} \cdot \nabla)\vec{B}' - \vec{V}\nabla \cdot \vec{B}'\right] \quad , \quad (77)$$

$$\nabla \times \vec{B} = \nabla \times \vec{B}' + \frac{b}{c^2} \left[ (\vec{V} \cdot \nabla) \vec{E}' - \vec{V} \nabla \cdot \vec{E}' \right] , \quad (78)$$

Para que as equações de Maxwell que envolvem divergências dos campos  $\vec{E}$  e  $\vec{B}$  permaneçam invariantes, é necessário impor  $a = b = 0$ . No entanto, com estas escolhas, as demais equações de Maxwell, aquelas que envolvem os rotacionais dos campos, são alteradas pelas transformações de Galileu. Assim, uma inspeção direta de todas estas transformações mostra que não é possível recuperar a forma original das equações de Maxwell para nenhum valor das constantes  $a$  e  $b$ .

Isto pode ser visto de uma forma mais direta, utilizando-se a equação de onda para o potencial  $\phi$ . Temos,

$$\square \phi = \frac{\partial^2 \phi}{\partial t^2} - c^2 \nabla^2 \phi = 0 . \quad (79)$$

Aplicando-se o fato que  $\nabla = \nabla'$  e que  $\frac{\partial^2}{\partial t'^2} = \frac{\partial^2}{\partial t^2} - 2\vec{V} \cdot \nabla' \frac{\partial}{\partial t'} + (\vec{V} \cdot \nabla')^2$ , obtemos,

$$\frac{\partial^2}{\partial t'^2} \phi - 2\vec{V} \cdot \nabla' \frac{\partial \phi}{\partial t'} + (\vec{V} \cdot \nabla')^2 \phi - c^2 \nabla'^2 \phi = 0 , \quad (80)$$

que já não representa uma onda se propagando com velocidade  $c$ . De fato, se uma onda plana tal que

$$\phi(\vec{x}, t) = A e^{i(\omega t - \vec{k} \cdot \vec{x})} , \quad (81)$$

onde  $A$  é uma constante, é solução de (79), com  $\omega = |\vec{k}|$ , por outro lado a solução de (80) exige que

$$\omega = |\vec{k}|(c - \vec{V} \cdot \vec{n}) , \quad (82)$$

onde  $\vec{n}$  é o vetor unitário que indica a direção de propagação da onda eletromagnética. Isto está em acordo com o fato que, na cinemática galileana, diferentes observadores medem diferentes velocidades, a diferença entre as velocidades sendo determinada pela velocidade relativa dos referenciais. Mas, como a equação da onda (79) é consequência das equações de Maxwell, isto quer dizer que diferentes observadores escrevem as equações de Maxwell de diferentes maneiras. Assim, o eletromagnetismo não é invariante pelas transformações de Galileu, ao contrário do que acontece com a mecânica Newtonina.

A não invariância das equações de Maxwell com relação às transformações de Galileu pode ser resolvida se considerarmos que estas equações são válidas em

um referencial bem definido. Isto é o que acontece com as ondas mecânicas, um exemplo sendo as ondas sonoras: a equação de onda correspondente é escrita no referencial de repouso do meio onde se propaga esta onda, no caso o ar; em outros referenciais em movimento em relação a este meio, a equação de propagação da onda sonora assume uma forma bem diferente da equação usual da onda. Isto implica, por exemplo, que um observador que se aproxima de uma fonte em repouso neste meio não é uma situação equivalente a uma fonte que se aproxima, com a mesma velocidade, de um observador em repouso em relação ao meio (isto será discutido com maior detalhe na seção 3.4); a existência de um meio onde uma onda se propaga implica em um referencial privilegiado para os fenômenos relativos a esta onda. No caso da luz, postulou-se que este referencial seria o do éter, meio onde as ondas eletromagnéticas se propagariam e onde elas teriam velocidade  $c$ . Mas, fazendo experiências com sinais luminosos, seria em princípio possível detectar este meio, como é o caso com as ondas sonoras. Esta experiência foi feita por A.A. Michelson e E.W. Morley, em torno de 1888, e repetida diversas vezes depois, cada vez com maior refinamento técnico, e seu resultado foi negativo: não houve indícios da existência do éter.

## 5 A experiência de Michelson-Morley

A teoria eletromagnética de Maxwell prevê a existência de ondas eletromagnéticas. A luz é, tão somente, onda eletromagnética em uma dada faixa de comprimento de onda. Durante o século XIX, admitia-se que a luz se propagava no éter, que é um referencial bem particular. Esta hipótese era conforme ao paradigma das ondas mecânicas. Se é assim, é natural assumir que a Terra se desloca em relação a este meio. Uma indicação que a Terra se desloca em relação ao éter é o fenômeno da aberração da luz. O fenômeno da aberração da luz refere-se ao fato que a direção da velocidade da luz emitida por uma estrela, medida por um observador na Terra, muda durante o período de tempo em que a Terra descreve sua órbita em torno do Sol. Se existe um meio onde a luz se propaga, no caso o éter, este fenômeno implica que a Terra se desloca em relação ao éter: o movimento aparente da estrela devido à velocidade de translação da Terra só pode ser compreendido se a luz se propaga com velocidade fixa  $c$  no meio inte-

restelar que seria inteiramente preenchido por este meio material indetectável<sup>6</sup>. O fenômeno de aberração da luz será discutido mais em detalhe no próximo capítulo<sup>7</sup>. Se Terra se desloca em relação ao éter, onde a luz teria velocidade  $c$ , seria possível medir os efeitos devido a este deslocamento, fazendo interferir feixes de luz que se propagam paralela e perpendicularmente à direção do deslocamento da Terra em relação ao éter. É um efeito similar ao da propagação da onda sonora na direção paralela ou perpendicular do vento, que é o movimento do meio onde se propaga a onda sonora, o ar.

A experiência de Michelson-Morley consistia precisamente em detectar o "vento" do éter. Trata-se de uma experiência de interferometria. Envia-se um feixe luminoso sobre um espelho semi-transparente. O feixe se divide em dois, que se propagam em direções perpendiculares. À distâncias  $L_1$  e  $L_2$  do espelho semi-transparente se encontram espelhos totalmente refletores. Os feixes são assim enviados novamente ao primeiro espelho, onde eles se interferem.

Considerando que uma das direções está orientada no sentido do deslocamento da Terra, e a outra perpendicularmente, que a Terra se desloca com velocidade  $V$  em relação ao éter, onde a luz tem velocidade  $c$ , usando a lei galileiana de adição das velocidades, os tempos gastos na ida e volta ao espelho semi-transparente, nas direções paralela ( $T_1$ ) e perpendicular ao movimento de translação da Terra ( $T_2$ ), são

$$T_1 = \frac{L_1}{c - V} + \frac{L_1}{c + V} = \frac{2L_1}{c(1 - \frac{V^2}{c^2})} \quad , \quad (83)$$

$$T_2 = \frac{2L_2}{c\sqrt{1 - \frac{V^2}{c^2}}} \quad . \quad (84)$$

Estes tempos são diferentes mesmo se as distâncias  $L_1$  e  $L_2$  forem iguais. Mas, para evitar problemas de medidas, como os relacionados com o conhecimento

<sup>6</sup>A não ser eventualmente pelos fenômenos relacionados à propagação da luz.

<sup>7</sup>Entretanto, vale assinalar que a posição das estrelas apresenta um movimento aparente, visto da Terra, devido a dois fenômenos distintos: a paralaxe e a aberração. Ambos serão discutidos no próximo capítulo. Podemos, no entanto, antecipar que os dois fenômenos podem ser distinguidos uma vez que o movimento aparente das estrelas devido à paralaxe depende da *posição* da Terra ao longo de sua órbita, ao passo que o movimento aparente devido à aberração depende da *direção da velocidade* da Terra ao longo de sua órbita. Além do mais, somente as estrelas mais próximas apresentam uma paralaxe apreciável, enquanto todas as estrelas apresentam um deslocamento aparente devido à aberração.

impreciso de  $L_1$  e  $L_2$ , é mais simples considerar  $L_1 \neq L_2$ , e repetir a experiência depois de girar todo o dispositivo de um ângulo  $\theta = \frac{\pi}{2}$ . Se a velocidade da luz depende da direção, devido ao deslocamento da Terra em relação ao éter, haverá um deslocamento das franjas de interferência quando for feita esta rotação.

Esse deslocamento das franjas não foi observado. Michelson e Morley repetiram e aperfeiçoaram esta experiência por mais de dez anos. A conclusão imediata é que o "vento" do éter não existe. Isto compromete a existência do éter e de um referencial privilegiado para as ondas eletromagnéticas. Observe-se, entretanto, que a experiência de Michelson e Morley não implica necessariamente que a luz tem uma velocidade constante que é independente do observador. Na verdade, as transformações de Lorentz, a serem introduzidas posteriormente, foram inicialmente propostas visando explicar o resultado negativo desta experiência: haveria uma contração dos corpos na direção do movimento, como sugerido inicialmente por Fitzgerald e Lorentz, o que explicaria a ausência de alteração das franjas de interferência quando o aparato era girado.

As transformações de Lorentz, formuladas originalmente para explicar o resultado negativo das experiências de Michelson-Morley, foram obtidas conjunta e independentemente por Lorentz e Poincaré antes do surgimento da teoria da Relatividade Restrita.

## 6 Os dois axiomas da relatividade restrita

Existe uma classe de "bons" observadores em Física: são os observadores localizados em referenciais inerciais. A situação descrita anteriormente mostra que as leis de Newton são válidas para todos os referenciais inerciais, se as medidas feitas em diferentes referenciais são conectadas pelas transformações de Galileu. Entretanto, o eletromagnetismo não possui esta invariância, e se quisermos manter as transformações de Galileu, temos que supor um referencial privilegiado onde as equações de Maxwell sejam válidas. A experiência de Michelson-Morley mostra, por outro lado, que não existe tal referencial privilegiado. Isto cria um impasse: se aceitarmos que as leis da Física são as mesmas em todos os referenciais inerciais, então ou as transformações de Galileu têm que ser modificadas, conduzindo também a uma modificação da mecânica Newtoniana, ou o eletro-

magnetismo deve ser modificado de forma a se adaptar às transformações de Galileu.

Este impasse é resolvido na Relatividade Restrita através de dois axiomas fundamentais:

1. Todos os observadores localizados em referenciais inerciais escrevem as leis da Física da mesma forma;
2. A velocidade da luz,  $c$ , é uma constante universal, tendo o mesmo valor em todos os referenciais inerciais.

O primeiro axioma põe todas as leis da Física no mesmo nível, do ponto de vista da invariância por uma transformação de um referencial inercial a outro. Mas, a mecânica Newtoniana não obedece às mesmas leis de transformação que o eletromagnetismo. O segundo axioma tem como consequência que a equação para as ondas eletromagnéticas guarda a mesma forma em todos os referenciais inerciais. Assim, o segundo axioma implica que as equações de Maxwell são invariantes em relação à passagem de um referencial inercial a outro e que, conseqüentemente, as transformações de Galileu não constituem a boa forma de traduzir as medidas de um referencial para outro. Como corolário, temos que a mecânica Newtoniana tem que ser modificada. É necessário, então, encontrar as boas leis de transformação, e adaptar a mecânica Newtoniana a elas.

O primeiro axioma seria, de uma certa forma, compartilhado por um físico newtoniano. Entretanto, o segundo axioma representa a ruptura com a física newtoniana visto que ele estabelece que a velocidade da luz não depende de quem a mede. Este axioma conduz a resultados inesperados. Estes resultados serão discutidos detalhadamente mais tarde. No momento, analisemos o que significa definir um sistema de referência. É preciso, em primeiro lugar, se ter uma medida de distância e uma medida de tempo. Utilizando-se o segundo postulado da Relatividade Restrita, pode-se estabelecer medidas de distâncias e temporais utilizando-se feixes luminosos. A distância de um certo ponto à origem pode ser obtida, colocando-se neste ponto um espelho refletor: envia-se um feixe luminoso, que retorna ao ponto de partida. Multiplica-se o tempo de ida e volta do feixe por  $c$  e divide-se depois por dois. Isto nos fornece a distância deste ponto. A medida de tempo, por outro lado, implica em que em

cada ponto do referencial deve-se ter um relógio e o conjunto destes relógios deve estar sincronizado. Para obter-se a sincronização, envia-se um feixe luminoso a partir da origem, no instante que por convenção se considera  $t = 0$ . Quando o feixe passar por um relógio a uma distância  $x$  da origem, ele deverá indicar o tempo  $t = \frac{x}{c}$ .

Um dos problemas que decorre dos axiomas da Relatividade Restrita é que não se pode utilizar os dados de um referencial  $S$  para se estabelecer medidas de distância e tempo em um outro referencial  $S'$  em movimento em relação a  $S$ . Suponhamos um referencial  $S$  onde identificamos dois pontos  $A$  e  $B$  equidistantes da origem, possuindo relógios sincronizados. Consideramos agora pontos  $A'$  e  $B'$ , também equidistantes da origem no referencial  $S'$ . Se o observador em  $S'$  utilizar o sinal luminoso que o observador em  $S$  usou para sincronizar seus relógios, ele poderá sincronizar os tempos de  $A'$  e  $B'$ , mas encontrará que os tempos de  $A$  e  $B$  não estarão sincronizados, já que a velocidade da luz é a mesma nos dois referenciais, mas para o observador em repouso em  $S'$ , os pontos  $A$  e  $B$  se deslocam durante o processo de sincronização. Em consequência, relógios sincronizados em um referencial, não o serão em outro referencial.

Para entendermos melhor esta questão, precisamos primeiro achar as leis de transformações que resultam dos dois axiomas da Relatividade Restrita.

## 7 As transformações de Lorentz

Nosso objetivo agora é estabelecer como se traduzem os dados físicos de um referencial inercial  $S$  para outro referencial inercial  $S'$ , quando os axiomas da Relatividade Restrita são respeitados. Procuremos inicialmente utilizar a forma geral de se passar das coordenadas  $(\vec{x}, t)$  medidas no referencial  $S$  para as coordenadas  $(\vec{x}', t')$  no referencial  $S'$ . A velocidade relativa, constante, entre os dois referenciais é  $\vec{V}$ .

Suponhamos inicialmente que o observador em repouso em  $S$  siga o movimento de um relógio que se desloca em relação a ele com velocidade constante<sup>8</sup>. O tempo medido por este relógio é  $\tau$ . Devido à homogeneidade do espaço,

---

<sup>8</sup>Esta seção segue de perto a apresentação do texto **Introduction to special relativity**, de Wolfgang Rindler.



incrementos iguais em  $\tau$  implicam incrementos iguais nas coordenadas  $(\vec{x}, t)$ .

Logo,

$$\frac{dx^\mu}{d\tau} = \text{constante} \quad , \quad \frac{d^2x^\mu}{d\tau^2} = 0 \quad , \quad (85)$$

onde  $\mu = 0, 1, 2, 3$ , sendo que  $x^0 = ct$ ,  $x^1 = x$ ,  $x^2 = y$  e  $x^3 = z$ . Isto deve ser também verdade para o observador no referencial  $S'$ : incrementos iguais em  $\tau$  implicam incrementos iguais em  $(\vec{x}', t')$ . Mas,

$$\frac{dx^{\mu'}}{d\tau} = \sum_{\nu} \frac{\partial x^{\mu'}}{\partial x^\nu} \frac{dx^\nu}{d\tau} = \text{constante} \quad , \quad (86)$$

$$\frac{d^2x^{\mu'}}{d\tau^2} = \sum_{\nu, \lambda} \frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\lambda} \frac{dx^\nu}{d\tau} \frac{dx^\lambda}{d\tau} + \sum_{\nu} \frac{\partial x^{\mu'}}{\partial x^\nu} \frac{d^2x^\nu}{d\tau^2} = 0 \quad . \quad (87)$$

Nós obtemos assim as condições de transformação do sistema de coordenadas  $x^\mu$  para o sistema  $x^{\mu'}$ :

$$\frac{\partial x^{\mu'}}{\partial x^\nu} = \text{constante} \quad ; \quad (88)$$

$$\frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\lambda} = 0. \quad (89)$$

Logo, a transformação de  $x^\mu$  para  $x^{\mu'}$  tem a forma,

$$x^{\mu'} = \sum_{\nu} A^{\mu'}_{\nu} x^\nu + B^{\mu'} \quad , \quad (90)$$

onde  $A^{\mu'}_{\nu}$  e  $B^{\mu'}$  são constantes. Suponhamos agora que o relógio esteja em repouso na origem do sistema  $S$ . Assim,  $x^i = 0$  e diferenciando (90), temos

$$dt' = A^0_{\mu'} dt \quad , \quad (91)$$

$$dx^{i'} = A^{i'}_{\mu'} dt \quad , \quad (92)$$

conduzindo a

$$\frac{dx^{i'}}{dt'} = \frac{A^{i'}_{\mu'}}{A^0_{\mu'}} = \text{constante} \quad . \quad (93)$$

Isto implica que a primeira lei de Newton é válida: um relógio em repouso em  $S$  terá velocidade constante em  $S'$ .

Nós podemos usar a isotropia do espaço (nada se altera se girarmos os eixos coordenadas), de forma que o eixo  $x$  seja paralelo ao eixo  $x'$ . Da mesma forma, nós podemos fazer com que os planos definidos pelas coordenadas  $y = 0$  e  $y' = 0$ , assim como o definido por  $z = 0$  e  $z' = 0$ , sejam os mesmos. A homogeneidade

do tempo nos permite que façamos  $t = 0$  coincidir com  $t' = 0$ . Se o referencial  $S'$  se desloca em relação a  $S$  com velocidade constante  $\vec{V}$  que, por simplicidade e usando as propriedades de simetria descritas acima, escolhemos como orientada ao longo do eixo  $x$ , então a origem de  $S'$  terá coordenada  $x = Vt$ , medida em  $S$ . Além do mais, suponhamos que as equações de transformação de coordenadas sejam invariantes em relação às seguintes operações:

$$t, x, y, z \quad , \quad t', x', y', z' \quad \rightarrow \quad t', -x', y', -z' \quad , \quad t, -x, y, -z \quad ; \quad (94)$$

$$t, x, y, z \quad , \quad t', x', y', z' \quad \rightarrow \quad t', -x', -y', z' \quad , \quad t, -x, -y, z \quad . \quad (95)$$

Estas operações equivalem a inverter os papéis de  $S$  e  $S'$ .

Nós podemos agora buscar as transformações que satisfazem as condições acima. Uma vez que a transformação deve ser linear, temos,

$$y' = Ax + By + Dt + E \quad . \quad (96)$$

Mas,  $y = 0$  implica  $y' = 0$ . Assim,  $A = D = E = 0$ . Considerando a primeira inversão acima temos,

$$y = By' = B^2y \rightarrow B = \pm 1 \quad . \quad (97)$$

Mas, no limite  $V \rightarrow 0$ , a transformação deve se reduzir à identidade. Assim, temos  $B = 1$ . O mesmo raciocínio se aplica à coordenada  $z$ . Logo,

$$y' = y \quad , \quad (98)$$

$$z' = z \quad . \quad (99)$$

Consideremos agora a coordenada  $x$ . Utilizando novamente a condição de linearidade, temos

$$x' = \Gamma x + Fy + Gz + Ht \quad . \quad (100)$$

Nós já vimos que  $x' = 0$  implica  $x = Vt$ . Assim,  $F = G = 0$  e  $H = -\Gamma V$ , resultando em,

$$x' = \Gamma(x - Vt) \quad . \quad (101)$$

A inversão  $xy$  conduz a

$$x = \Gamma(x' + Vt') \quad . \quad (102)$$

Para calcularmos o parâmetro  $\Gamma$  usamos o segundo axioma da Relatividade Restrita. Como a velocidade da luz é a mesma em todos os referenciais inerciais, se uma frente de onda é emitida no momento  $t = t' = 0$ , nós temos,

$$ct' = \Gamma(c - V)t \quad , \quad (103)$$

$$ct = \Gamma(c + V)t' \quad . \quad (104)$$

Temos então,

$$\begin{aligned} c^2 tt' &= \Gamma^2 tt' (c^2 - V^2) \quad , \\ \Gamma &= \frac{c}{\sqrt{c^2 - V^2}} \quad , \\ \Gamma &= \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} \quad . \end{aligned} \quad (105)$$

Escolhemos o sinal positivo da raiz de maneira a se ter a transformação identidade quando  $V \rightarrow 0$ .

A lei de transformação para  $t'$  decorre das precedentes:

$$\begin{aligned} x' &= \Gamma(x - Vt) \quad , \\ x &= \Gamma(x' + Vt') \quad , \\ x &= \Gamma \left[ \Gamma(x - Vt) + Vt' \right] \quad , \\ \Gamma Vt' &= x(1 - \Gamma^2) + \Gamma^2 Vt \quad , \\ \Gamma Vt' &= \Gamma^2 \left( Vt - \frac{V^2}{c^2} x \right) \quad , \\ t' &= \Gamma \left( t - \frac{V}{c^2} x \right) \quad , \end{aligned} \quad (106)$$

$$t = \Gamma \left( t' + \frac{V}{c^2} x' \right) \quad . \quad (107)$$

Desta forma, nós temos o seguinte conjunto de transformações entre as coordenadas do referencial  $S$  e as do referencial  $S'$ :

$$t' = \Gamma \left( t - \frac{V}{c^2} x \right) \quad ; \quad t = \Gamma \left( t' + \frac{V}{c^2} x' \right) \quad ; \quad (108)$$

$$x' = \Gamma(x - Vt) \quad ; \quad x = \Gamma(x' + Vt') \quad ; \quad (109)$$

$$y' = y \quad , \quad z' = z \quad ; \quad y = y' \quad , \quad z = z' \quad . \quad (110)$$

Estas expressões são denominadas *Transformações de Lorentz*.

Nós podemos mostrar que, se a seção espacial de  $S$  é euclidiana, a de  $S'$  também o será. A distância percorrida por um raio luminoso, em  $S$ , é dada por

$$c^2 dt^2 = dx^2 + dy^2 + dz^2 \quad . \quad (111)$$

Aplicando as transformações de Lorentz, temos,

$$\begin{aligned} c^2 \left[ \Gamma^2 (dt' + \frac{V}{c^2} dx')^2 \right] &= \Gamma^2 (dx' + V dt')^2 + dy'^2 + dz'^2 \quad , \\ dt'^2 (c^2 \Gamma^2 - \Gamma^2 V^2) &= dx'^2 (\Gamma^2 - \frac{V^2}{c^2} \Gamma^2) + dy'^2 + dz'^2 \quad , \\ c^2 dt'^2 &= dx'^2 + dy'^2 + dz'^2 \quad . \end{aligned} \quad (112)$$

Isto nos assegura que o espaço tridimensional observado por  $S'$  permanece euclidiano.

## 8 Propriedades das transformações de Lorentz

Vamos analisar algumas consequências das transformações deduzidas anteriormente. Um conceito importante que será utilizado sucessivamente aqui é o de evento. Um evento é um fenômeno que ocorre em um ponto do espaço, em um dado momento. Assim, a um evento  $P$  nós associamos quatro coordenadas, uma temporal e três espaciais:  $P = P(t, \vec{x})$ .

### 8.1 O conceito de simultaneidade

Consideremos dois eventos no referencial  $S$  tendo coordenadas,

$$P_1 = (t_1, x_1) \quad , \quad (113)$$

$$P_2 = (t_2, x_2) \quad . \quad (114)$$

Suponhamos agora que  $t_1 = t_2$ : os eventos são simultâneos no referencial  $S$ . Consideremos um referencial  $S'$  que se desloca em relação a  $S$  com velocidade constante  $V$ , ao longo do eixo  $x$ . As coordenadas temporais destes dois eventos medidas no referencial  $S'$  são,

$$t_1' = \Gamma(t_1 - \frac{V}{c^2} x_1) \quad , \quad (115)$$

$$t_2' = \Gamma(t_2 - \frac{V}{c^2} x_2) \quad . \quad (116)$$

Como  $t_1 = t_2$ , temos,

$$\Delta t' = t_2' - t_1' = \Gamma \frac{V}{c^2} (x_1 - x_2) \quad . \quad (117)$$

Visto do referencial  $S'$ , estes dois eventos não são simultâneos. Eles só o serão se  $x_1 = x_2$ . Mas, neste caso, eles constituiriam um único evento, visto que  $P_1$  ocorreria no mesmo lugar e no mesmo momento que  $P_2$ . Observe-se que, no referencial  $S'$ , o evento  $P_1$  pode ocorrer antes ou depois do evento  $P_2$ , dependendo do sinal de  $\Delta x = x_1 - x_2$ . Isto é possível, como veremos mais tarde, porque se os eventos são simultâneos em um dado referencial, necessariamente eles não têm conexão causal. Desta forma, a ordem temporal com que eles são vistos em um outro referencial é irrelevante.

O fato que nem todos os observadores concordam sobre quais fenômenos são simultâneos é um das chaves para compreender muitos dos aparentes paradoxos da relatividade restrita. A noção de simultaneidade na teoria newtoniana é absoluta: todos os observadores concordam sobre quais são os fenômenos que ocorrem ao mesmo tempo. Mas, este não é mais o caso quando se substitui as transformações de Galileu pelas transformações de Lorentz. Como veremos depois, esta "relativização" do conceito de simultaneidade está intimamente ligada à noção de causalidade em relatividade restrita.

## 8.2 A contração das distâncias

Consideremos uma régua de comprimento  $L$  em repouso em um referencial  $S$ . Essa régua terá seu comprimento medido em um referencial  $S'$  que se desloca com velocidade  $V$  em relação a  $S$ . Para fazer esta medida, o observador em  $S'$  deverá assinalar, simultaneamente, as posições das duas extremidades da régua. Utilizando as transformações de Lorentz, e considerando que as medidas são feitas no mesmo tempo  $t'_1 = t'_2$ , obtemos,

$$\begin{aligned} x_2 &= \Gamma(x'_2 + Vt'_2) \quad , \\ x_1 &= \Gamma(x'_1 + Vt'_1) \quad , \\ \Delta x &= x_2 - x_1 = \Gamma \Delta x' \quad . \end{aligned} \quad (118)$$

Visto que  $\Delta x = L$ , escrevendo o comprimento da régua no referencial  $S'$  como  $\Delta x' = L'$ , obtemos,

$$L' = \frac{L}{\Gamma} = \left(1 - \frac{V^2}{c^2}\right)^{1/2} L \quad . \quad (119)$$

Assim, o comprimento da régua medido em  $S'$  é menor que o medido em  $S$ . Isto caracteriza o que se denomina de contração das distâncias. Notemos que a régua terá comprimento máximo no referencial onde ela está em repouso.

Essa contração é, de fato, um efeito de "perspectiva" do observador que vê a régua em movimento. Consideremos um exemplo. Um observador está em repouso em um referencial  $S''$ , onde as distâncias são graduadas. Duas réguas de mesmo comprimento  $L$  se deslocam no referencial  $S$  com velocidades iguais mas opostas. As extremidades destas réguas são designadas por  $PQ$  e por  $P'Q'$ . O ponto  $P(Q)$  encontra o ponto  $P'(Q')$  no mesmo momento em  $S$ . Se nós considerarmos agora estes eventos vistos do referencial  $S'$  da régua  $PQ$ , o encontro de  $PP'$  ocorre depois do encontro  $QQ'$ , já que, neste referencial, a régua  $P'Q'$  parecerá menor que a régua  $PQ$ ; além disto, visto a partir do referencial  $S''$  da régua  $P'Q'$ , o encontro  $QQ'$  ocorre depois que o de  $PP'$ , pelo mesmo motivo. Vê-se aqui um exemplo da relatividade da simultaneidade.

A contração das distâncias pode conduzir a "falsos" paradoxos. Consideremos o seguinte caso. Um carro de  $20m$  de comprimento, em seu próprio referencial, se aproxima de uma garagem que, também no seu referencial próprio, mede  $10m$ , com velocidade  $V$ . Pode o carro entrar na garagem? Suponhamos  $V = 0,866c$ , de maneira que  $\Gamma = 2$ . No referencial da garagem, o carro terá comprimento igual a  $10m$ , e assim ele poderá entrar inteiramente na garagem. Mas no referencial do carro, a garagem medirá  $5m$ , e aparentemente o carro não caberia nela. Como podemos conciliar isto com o princípio da simetria entre os referenciais?

Vamos nos pôr agora no referencial do carro. A frente do carro encontra o fundo da garagem. A parte traseira do carro ainda não recebeu o sinal deste evento. Ela o receberá num tempo  $t = \frac{20}{c}$  depois. Mas, durante este tempo, a parte traseira terá se deslocado uma distância igual a  $x = tV = 17,32m$ . Como a distância entre a porta da garagem e a parte traseira do carro era de  $15m$ , todo o carro terá atravessado a entrada da garagem quando o observador na parte

traseira receber o sinal que a parte dianteira chegou no fundo da garagem: eventos que eram simultâneos no referencial da garagem, correspondendo às partes dianteira e traseira no interior da garagem, não o são no referencial do carro. A solução deste paradoxo se encontra na propagação finita dos sinais, em especial a da luz.

### 8.3 A dilatação do tempo

Consideremos, como antes, dois referenciais  $S$  e  $S'$  com velocidade relativa  $V$ . Suponhamos que no referencial  $S$  haja um relógio em repouso. Neste referencial se escoa um intervalo de tempo  $T = \Delta t$  medido por este relógio. Qual é o intervalo de tempo  $T' = \Delta t'$  correspondente que medirá um observador em repouso no referencial  $S'$ ? Utilizamos novamente as transformações de Lorentz, relacionando os tempos  $t_1, t_2$  com  $t'_1, t'_2$  que marcam o início e o fim deste intervalo de tempo em cada referencial:

$$t'_2 = \Gamma(t_2 - \frac{V}{c^2}x_2) \quad , \quad (120)$$

$$t'_1 = \Gamma(t_1 - \frac{V}{c^2}x_1) \quad . \quad (121)$$

Como o relógio está em repouso no referencial  $S$ , temos  $x_1 = x_2$ . Logo,

$$\Delta t' = \Gamma \Delta t \quad \rightarrow \quad T' = \Gamma T \quad . \quad (122)$$

O intervalo de tempo, que em  $S$  é igual a  $T$ , será dilatado para um observador em  $S'$  por um fator  $\Gamma$ . Se  $T$  representa um segundo no relógio no referencial  $S$ , este segundo parecerá mais longo para um observador em  $S'$ , que concluirá que o relógio em repouso em  $S$  está se atrasando em relação ao seu próprio relógio.

Devido a equivalência dos referenciais inerciais, o observador em  $S$  dirá, pelas mesmas razões expostas acima, que o relógio em repouso no referencial  $S'$  está se atrasando. A dilatação do tempo é um puro efeito de movimento relativo de referenciais inerciais, e evidentemente não pode conduzir a uma constatação absoluta, visto que isto implicaria que os dois relógios, em  $S$  e em  $S'$ , deveriam ter seus indicadores de tempo comparado duas vezes no mesmo ponto do espaço. Mas isto implicaria que um dos relógios deveria "retornar", revertendo seu movimento, e isto conduz a uma aceleração de um dos referenciais, que deixaria

portanto de ser inercial. Este é o caso do paradoxo dos gêmeos, que novamente é um "falso" paradoxo: dois irmãos gêmeos se separam, um ficando na Terra e o outro empreendendo uma viagem interestelar com velocidade próxima à da luz. Quando o gêmeo "astronauta" retorna, ele se revela mais jovem que o seu irmão que ficou na Terra. A solução para este problema está no fato que o "retorno" implica na passagem de um referencial inercial para um não inercial, e toda a análise acima não se aplica mais. Nós trataremos deste paradoxo com mais detalhe posteriormente.

## 8.4 Transformações de Lorentz na forma diferencial

Escrevamos inicialmente as transformações de Lorentz sob a forma de diferenças finitas:

$$\Delta t' = \Gamma(\Delta t - \frac{V}{c^2}\Delta x) \quad ; \quad \Delta x' = \Gamma(\Delta x - V\Delta t) \quad ; \quad (123)$$

$$\Delta y' = \Delta y \quad , \quad \Delta z' = \Delta z \quad . \quad (124)$$

No limite em que o intervalo de tempo e, consequentemente os intervalos de espaço, tende a zero, obtém-se as seguintes relações diferenciais:

$$dt' = \Gamma(dt - \frac{V}{c^2}dx) \quad ; \quad dx' = \Gamma(dx - Vdt) \quad ; \quad (125)$$

$$dy' = dy \quad , \quad dz' = dz \quad ; \quad dy = dy' \quad , \quad dz = dz' \quad . \quad (126)$$

Essa forma de se escrever é possível uma vez que a velocidade  $V$ , e consequentemente o fator  $\Gamma$ , é constante.

## 8.5 Existe uma velocidade limite

Consideremos de novo os referenciais  $S$  e  $S'$  que têm uma velocidade relativa igual a  $V$ . Se  $V$  for maior que  $c$ , a velocidade da luz, o fator  $\Gamma$  torna-se imaginário. Isto mostra que, para que o formalismo da Relatividade Restrita seja coerente, contendo unicamente observáveis reais, passíveis de serem medidos, nenhum objeto pode ter velocidade superior à da luz, já que a qualquer objeto podemos associar um referencial. Além do mais, nenhum objeto pode ter a velocidade igual à da luz, já que isto implica uma divergência do fator  $\Gamma$ . A velocidade da luz é uma velocidade limite na natureza, permitida unicamente



aos objetos não massivos. Tais objetos, como o fóton e o gráviton <sup>9</sup>, mediadores das interações eletromagnéticas e gravitacional, por serem não massivos, não podem ter a eles associado um referencial no sentido usual, e suas velocidades serão, como mostraremos mais tarde, invariantes, e sempre iguais a  $c$ .

Além disto, a existência de uma velocidade superior à da luz conduziria a uma violação da causalidade. Consideremos, por exemplo, dois eventos  $P_1 = (t_1, x_1)$  e  $P_2 = (t_2, x_2)$  vistos a partir do referencial  $S$ , tal que  $\Delta t = t_2 - t_1 > 0$  e  $\Delta x = x_2 - x_1 > 0$ . Assumamos que o evento  $P_1$  é causa do evento  $P_2$ . No referencial  $S'$ , temos,

$$\Delta t' = \Gamma(\Delta t - \frac{V}{c^2}\Delta x) = \Gamma\Delta t(1 - \frac{V}{c^2}U) \quad (127)$$

onde

$$U = \frac{\Delta x}{\Delta t} \quad (128)$$

é a velocidade com que o sinal emitido por  $P_1$  se desloca até atingir  $P_2$ . Se  $U$  for maior que  $c$  de forma que  $UV > c^2$ , então  $\Delta t' < 0$ , e a ordem de acontecimento dos dois eventos visto de  $S'$  seria invertida: o efeito aconteceria antes da causa, violando a causalidade. Observe-se que, fixando  $U > c$ , sempre será possível encontrar um referencial  $S'$  tal que a condição  $UV > c^2$  será satisfeita. A única forma de se preservar o ordenamento temporal de eventos que têm conexão causal é impor a impossibilidade de se ter velocidades superiores a  $c$ .

## 9 O espaço-tempo de Minkowski

As transformações de Lorentz combinam de forma não trivial as coordenadas espaciais e temporais. Desta forma, a visão habitual da mecânica newtoniana de um espaço euclidiano, onde se definem os diferentes referenciais, dotado de um tempo universal, não afetado pela mudança de referenciais, não é mais adequada. É preciso agora uma concepção *espaço-temporal*, devido ao fato que cada referencial define seu tempo, da mesma maneira que define suas coordenadas espaciais. A estrutura criada pelas transformações de Lorentz requer um contínuo quadri-dimensional, com três coordenadas espaciais e uma coordenada temporal, que se "deforma" quando se passa de um referencial a outro.

---

<sup>9</sup>Ao contrário do fóton, o gráviton ainda não foi detectado experimentalmente

Podemos construir um diagrama espaço-temporal, a quatro dimensões, e nele representar os efeitos das transformações de Lorentz. De forma a simplificar esta representação, vamos considerar um espaço-tempo bidimensional, com coordenadas  $(ct, x) = (x^0, x^1)$ . Definimos a nova coordenada  $x^0 = ct$  de maneira a se ter homogeneidade dimensional. O eixo  $x^0$  é definido por  $x^1 = 0$  ( $t = 0$ ), e o eixo  $x^1$  por  $x^0 = 0$ . Um feixe de luz neste diagrama tem uma trajetória retilínea, fazendo um ângulo de  $\frac{\pi}{4}$  com o semi-eixo positivo ou negativo de  $x^1$ , uma vez que para este feixe,  $x = x^1 = ct = x^0$ . As retas que fazem este ângulo com o eixo  $x^1$  definem o que é denominado de cone de luz. O interior do cone de luz é a região causalmente conectada com a origem.

De fato, qualquer evento  $P = (x^0, x^1)$  localizado no interior do cone de luz tem conexão causal com o evento da origem  $P_0 = (0, 0)$ , pois um sinal emitido a partir da  $P_0$ , com velocidade inferior à da luz, pode atingir o evento  $P$ , se  $x^0 > 0$ ; ou ainda, um sinal emitido de  $P$  com velocidade inferior à da luz pode atingir  $P_0$ , se  $x^0 < 0$ . Desta forma, a parte superior do cone de luz é constituída de eventos que sofreram a ação de  $P_0$ , ao passo que a parte inferior do cone de luz é constituída de eventos que agiram sobre  $P_0$ . Por este motivo, a parte superior do cone de luz é chamada de futuro, e a parte inferior de passado. As noções de passado e futuro se referem ao evento da origem,  $P_0$ . Todos os eventos localizados fora do cone de luz não têm conexão causal com a origem, já que um sinal que os conecte com a origem deve necessariamente se propagar com velocidade superior à da luz. Notemos, *en passant*, que um evento  $P_1$  dentro do cone de luz de  $P_0$  pode ter conexão causal com um evento  $P_2$  fora do cone de luz de  $P_0$ ; para tanto, basta que  $P_1$  esteja dentro do cone de luz de  $P_2$ .

Temos pelas Transformações de Lorentz,

$$x^{0'} = \Gamma(x^0 - \frac{V}{c}x^1) \quad , \quad (129)$$

$$x^{1'} = \Gamma(x^1 - \frac{V}{c}x^0) \quad . \quad (130)$$

Podemos construir os novos eixos  $(x^{0'}, x^{1'})$ . Para definirmos o eixo  $x^{0'}$ , fazemos  $x^{1'} = 0$ , obtendo  $\frac{x^1}{x^0} = \frac{V}{c}$ . Da mesma forma, para obtermos o eixo  $x^{1'}$ , fazemos  $x^{0'} = 0$ , obtendo  $\frac{x^0}{x^1} = \frac{V}{c}$ . Os ângulos entre os novos eixos  $(x^{0'}, x^{1'})$  e os antigos eixos  $(x^0, x^1)$  são dados por  $\tan \theta = \frac{V}{c}$ . Observe-se que, visto do referencial  $S$ , os eixos do referencial  $S'$  têm a curiosa forma de uma cunha. Isto se deve à natureza

hiperbólica do espaço-tempo de Minkowski, como será discutido mais tarde. Dito de outra maneira, a forma de cunha é consequência do fato que a velocidade da luz é a mesma nos dois referenciais  $S$  e  $S'$ : ela deve ser representada pela bissetriz dos eixos  $(x^0, x^1)$  e  $(x^{0'}, x^{1'})$  simultaneamente.

A partir das transformações de Lorentz, é fácil verificar (veja o exercício 5) que

$$(x^0)^2 - (x^1)^2 = (x^{0'})^2 - (x^{1'})^2 . \quad (131)$$

Logo, a hipérbole  $x^2 - c^2 t^2 = (x^1)^2 - (x^0)^2 = cte$  permanece invariante quando aplicamos uma transformação de Lorentz. Dependendo do sinal da constante nessa expressão, a hipérbole corta seja os eixos  $x^0, x^{1'}$  ou  $x^0, x^{0'}$ , definindo unidades de medidas de espaço ou de tempo.

## 10 Transformações de velocidades

O observador no referencial  $S$  define como as componentes da velocidade de um objeto que se desloca no espaço como sendo

$$v_x = \frac{dx}{dt} , \quad v_y = \frac{dy}{dt} , \quad v_z = \frac{dz}{dt} , \quad (132)$$

enquanto o observador no referencial  $S'$  define as componentes da velocidade deste mesmo objeto como sendo

$$v'_x = \frac{dx'}{dt'} , \quad v'_y = \frac{dy'}{dt'} , \quad v'_z = \frac{dz'}{dt'} , \quad (133)$$

Podemos relacionar as componentes da velocidade deste objeto medidas em  $S'$  com as medidas em  $S$  através das transformações de Lorentz. De fato, temos

$$\begin{aligned} v'_x &= \frac{dx'}{dt'} = \frac{d[\Gamma(x - Vt)]}{d[\Gamma(t - \frac{V}{c^2}x)]} \\ &= \frac{dx - Vdt}{dt - \frac{V}{c^2}dx} = \frac{v_x - V}{1 - \frac{Vv_x}{c^2}} ; \end{aligned} \quad (134)$$

$$\begin{aligned} v'_y &= \frac{dy'}{dt'} = \frac{dy}{d[\Gamma(t - \frac{V}{c^2}x)]} \\ &= \frac{dy}{\Gamma(dt - \frac{V}{c^2}dx)} = \frac{v_y}{\Gamma(1 - \frac{Vv_x}{c^2})} ; \end{aligned} \quad (135)$$

$$v'_z = \frac{dz'}{dt'} = \frac{dz}{d[\Gamma(t - \frac{V}{c^2}x)]}$$

$$= \frac{dz}{\Gamma(dt - \frac{V}{c^2}dx)} = \frac{v_z}{\Gamma(1 - \frac{Vv_x}{c^2})} . \quad (136)$$

Ao fazermos esta dedução nós utilizamos o fato que  $V$ , a velocidade relativa entre os dois referenciais, e consequentemente  $\Gamma$ , é constante.

Podemos então resumir as leis de transformação ligando as componentes da velocidade do objeto medidas em  $S$  e em  $S'$  pelas expressões:

$$v'_x = \frac{v_x - V}{1 - \frac{Vv_x}{c^2}} ; \quad (137)$$

$$v'_y = \frac{v_y}{\Gamma(1 - \frac{Vv_x}{c^2})} ; \quad (138)$$

$$v'_z = \frac{v_z}{\Gamma(1 - \frac{Vv_x}{c^2})} . \quad (139)$$

Observe-se que, muito embora o movimento relativo entre os dois referenciais seja ao longo do eixo  $x$ , as componentes da velocidade ao longo dos eixos  $y$  e  $z$  são afetadas. Isto se deve ao efeito de dilatação do tempo.

Uma consequência notável das transformações (137),(138),(139) refere-se à medida da velocidade da luz. Se no referencial  $S$  um feixe de luz se propaga com velocidade  $c$ , ao longo do eixo positivo dos  $x$  (logo  $v_x = c$ ), este mesmo feixe será visto no referencial  $S'$  com velocidade

$$v'_x = \frac{c - V}{1 - \frac{Vc}{c^2}} = \frac{c - V}{1 - \frac{V}{c}} = c . \quad (140)$$

Assim, os dois referenciais concordam que a velocidade da luz é igual a  $c$ .

As transformações (137),(138),(139) se reduzem às galileanas quando a velocidade relativa entre os referenciais é muito pequena em relação à velocidade da luz. Neste caso, podemos desprezar todos os termos envolvendo  $\frac{V}{c}$ , e assumir  $\Gamma \approx 1$ , e reencontramos as transformações usuais da cinemática newtoniana.

## 11 Comparando a cinemática lorentziana com a galileana

Em situações onde a velocidade dos corpos envolvidos no problema é muito pequena em relação à velocidade da luz, nós podemos utilizar a fórmula de adição das velocidades galilenianas em vez das transformações (137),(138),(139): a diferença entre os resultados fornecidos por essas diferentes expressões para a adição

das velocidades é desprezível. Entretanto, quando as velocidades dos corpos se tornam comparáveis com a da luz, então inevitavelmente deveremos usar as transformações lorentzianas. Exemplifiquemos isto com dois casos concretos.

Consideremos dois corpos cujas velocidades, medidas em um referencial  $S$ , têm o mesmo valor em módulo, a mesma direção mas têm sentidos contrários. Identifiquemos a direção destas velocidades como sendo o eixo  $x$ . Em um primeiro caso, consideremos que o módulo da velocidade destes corpos é  $|v_1| = |v_2| = 72 \frac{km}{h} = 20 \frac{m}{s}$ . Velocidades como estas aparecem normalmente em uma rodovia. Qual é a velocidade com que o observador localizado em um dos corpos vê o outro?

Suponhamos que o corpo 1 se mova no sentido positivo dos  $x$ . Se identificarmos o referencial  $S'$  como um referencial co-móvel a este corpo, teremos, utilizando a lei das transformações das velocidades galileanas,

$$v' = v_2 - V = 2v_2 = 144 \frac{km}{h} = 40 \frac{m}{s} . \quad (141)$$

Utilizando a expressão lorentziana (137), temos

$$v'_1 = \frac{v_1 - V}{1 - \frac{Vv_1}{c^2}} = \frac{2v_1}{1 + \frac{(v_1)^2}{c^2}} \quad (142)$$

Logo, as duas formas de calcular a velocidade vista pelo observador no referencial  $S'$  co-móvel com corpo 1 diferem por um termo  $\frac{(v_1)^2}{c^2} = \frac{1,6 \times 10^3}{9 \times 10^{16}} \approx 1,5 \times 10^{-13}$ . Realizando uma expansão em série em (142), obtemos que a diferença no resultado final em relação a (141) é  $\Delta v \approx 6 \times 10^{-12} \frac{m}{s}$ . Obviamente, tal erro no cálculo da velocidade está acima da precisão de praticamente todos os instrumentos de medidas normalmente utilizados. Podemos, portanto, usar a fórmula galileniana para adição das velocidades com segurança.

Consideremos agora que as velocidades sejam  $|v_1| = |v_2| = 270.000 \frac{km}{s} = 0,9c$ . Velocidades como estas são usualmente obtidas em aceleradores de partículas quando se deseja realizar uma colisão entre nêutrons, prótons ou elétrons. Se usarmos a relação de velocidades galileana, teremos,

$$v' = v_2 - V = 2v_2 = 540.000 \frac{km}{s} = 1,8c . \quad (143)$$

A partícula 2 verá a partícula 1 com uma velocidade bastante superior à da luz.

Por outro lado, empregando (137), obtemos,

$$v'_1 = \frac{v_1 - V}{1 - \frac{Vv_1}{c^2}} = \frac{2v_1}{1 + \frac{(v_1)^2}{c^2}} = \frac{1,8c}{1 + 0,81} = \frac{1,8}{1,81}c \approx 0,994c \quad . \quad (144)$$

Assim, a velocidade da partícula 2 medida referencial da partícula 1 será maior que a velocidade medida no referencial  $S$ , mas ainda inferior à velocidade da luz. O resultado difere substancialmente do obtido usando as transformações de velocidade galileanas.

A razão entre a velocidade do objeto  $v$  e a da luz  $c$  é uma boa referência para decidirmos quando deveremos utilizar as expressões lorentzianas ou quando podemos nos contentar com as expressões galileanas. Para o movimento da Terra em torno do sol, temos que a velocidade orbital é  $v \approx 30 \frac{km}{s}$ , e  $\frac{v}{c} = 10^{-4}$ ; as expressões não relativistas podem ser utilizadas com segurança. Entretanto, para um elétron em um acelerador de partícula, temos  $v \approx 0,9c$ , e  $\frac{v}{c} \approx 0,9$ : forçosamente deveremos utilizar as expressões relativistas.

## 12 O Efeito de arrasto

Um dos problemas clássicos da ótica refere-se à medida da velocidade da luz em um meio em movimento, em um tubo com água corrente por exemplo. Se a velocidade da luz no vácuo é  $c$ , e se a velocidade da água é  $V$ , a medida da velocidade da luz neste meio indica um resultado aproximado de

$$v \approx v' + V\left(1 - \frac{1}{n^2}\right) \quad , \quad (145)$$

onde  $n$  é o índice de refração da água e  $v' = \frac{c}{n}$ . Este resultado é obtido através do estudo da interferência de um feixe de luz, que é dividido em dois através de um espelho semi-transparente, de tal forma que um dos feixes resultantes se propaga neste meio, no sentido da corrente de água, enquanto que o outro que se propaga no sentido contrário ao da corrente: a diferença final de fase, quando os feixes voltam a se encontrar está relacionada com a diferença nas velocidades de propagação; a medida das franjas de interferência dos feixes permite deduzir a diferença de fase entre eles, que por sua vez fornece a diferença de caminho ótico. Esta diferença de caminho ótico está relacionada com a diferença das

velocidades dos dois feixes na água corrente. A experiência foi feita por Fizeau no início do século XIX.

Do ponto de vista clássico, o resultado (145) revela-se misterioso. No entanto, ele é facilmente explicado utilizando-se a cinemática relativista mas exigindo, ao contrário um complexo mecanismo envolvendo o conceito do éter quando se emprega a mecânica newtoniana. No que se segue, empregaremos diretamente a lei de transformação das velocidades (137).

Consideremos um feixe de luz que se propaga ao longo de um tubo com água corrente. Qual é a velocidade deste feixe de luz medido no referencial do laboratório (que podemos tomar também como sendo o referencial do tubo)? A velocidade da água no tubo é  $V$ . Esta é, por conseguinte, a velocidade relativa entre os referenciais da água e do laboratório. Se no vácuo a luz tem velocidade  $c$ , na água (em movimento ou em repouso) ela será  $v' = \frac{c}{n}$  onde  $n$  é o índice de refração da água. Uma dedução rigorosa desta relação envolve o problema de interação da radiação com a matéria, estando além dos nossos objetivos aqui. Vamos assumir que essa é uma relação empírica.

Utilizando as transformações de velocidade inversas, obtemos,

$$v = \frac{v' + V}{1 + \frac{Vv'}{c^2}} \quad . \quad (146)$$

Consideremos que  $V$  é muito pequeno em relação a  $c$ . Podemos fazer uma expansão em série, em termos de  $\frac{V}{c}$ , obtendo,

$$v \approx (v' + V)\left(1 - \frac{Vv'}{c^2}\right) \approx v' + V - \frac{Vv'^2}{c^2} \quad . \quad (147)$$

Como  $v' = \frac{c}{n}$ , temos então,

$$v \approx v' + V\left(1 - \frac{1}{n^2}\right) \quad . \quad (148)$$

Se usássemos as transformações de Galileu, em conjunção com a teoria do éter, e considerássemos que o meio onde se propagam as ondas eletromagnéticas não é arrastado pela água, encontraríamos  $v = v'$ ; se, por outro lado considerássemos que a água arrasta integralmente o éter, teríamos  $v = v' + V$ . O resultado acima indica, do ponto de vista da teoria do éter, que temos um "arrasto" parcial do éter pela água. A expressão (148) foi obtida no início experimentalmente, e foi

interpretada, através da teoria do éter, como uma indicação que o éter é parcialmente arrastado pela água em movimento. Entretanto, o uso das transformações de velocidades deduzidas a partir das transformações de Lorentz, indica que se trata na verdade de um efeito relativista, a expressão experimental constituindo o primeiro termo em uma expansão  $\frac{V}{c}$ .

## 13 A Aberração da luz

Quando a Terra descreve a translação em torno do Sol, a posição aparente das estrelas mais próximas muda devido a dois efeitos. O primeiro é o efeito de paralaxe: a posição aparente de estrela próxima muda, em relação às estrelas mais distantes, devido à mudança de posição da Terra ao longo de sua trajetória em torno do sol; o segundo é o efeito de aberração: como a Terra possui uma velocidade de translação, a direção da luz vinda da estrela sofre uma mudança, ganhando um ângulo de inclinação que muda devido à mudança na direção da velocidade orbital da Terra. Tanto a paralaxe quanto a aberração fazem a posição aparente da estrela se alterar, descrevendo uma elipse ao longo da translação da Terra. Entretanto, os dois efeitos podem ser distinguidos, devido às suas características: a paralaxe se deve à mudança de posição da Terra, enquanto a aberração se deve à mudança da velocidade de translação da Terra, o que implica que a elipse devido à paralaxe tem uma orientação distinta da devido à aberração. Além disto, o efeito de paralaxe ocorre apenas para estrelas extremamente próximas, enquanto o efeito de aberração ocorre mesmo para estrelas distantes.

Usando as expressões para a transformação de velocidade, nós podemos deduzir o ângulo de aberração, quer dizer, o ângulo de posição aparente da estrela em função do ângulo sob o qual ela seria visto caso a Terra não estivesse em movimento em relação ao Sol. Para tanto, assumamos que o referencial do Sol é inercial (referencial  $S$ ) e que um observador nele em repouso veria a estrela sob um ângulo  $\alpha$  no plano  $xy$ ; por outro lado, o observador na Terra, que se desloca com velocidade  $V$  em relação ao Sol (referencial  $S'$ ), ao longo do eixo  $x$ , vê a estrela sob um ângulo  $\alpha'$ .

Neste caso, a velocidade da luz se decompõe nos referenciais  $S$  e  $S'$  como se



segue:

$$v_x = -c \cos \alpha \quad , \quad v_y = -c \sin \alpha \quad ; \quad (149)$$

$$v'_x = -c \cos \alpha' \quad , \quad v'_y = -c \sin \alpha' \quad . \quad (150)$$

Usando a transformação relativista das velocidades, temos:

$$v'_x = \frac{v_x + V}{1 + \frac{v_x V}{c^2}} \quad \Rightarrow \quad \cos \alpha' = \frac{\cos \alpha + \frac{V}{c}}{1 + \frac{\cos \alpha V}{c}} \quad , \quad (151)$$

$$v'_y = \frac{v_y}{\Gamma(1 + \frac{v_x V}{c^2})} \quad \Rightarrow \quad \sin \alpha' = \frac{\sin \alpha}{\Gamma(1 + \frac{\cos \alpha V}{c})} \quad . \quad (152)$$

Para colocarmos estas expressões numa forma mais conveniente, devemos fazer algumas manipulações trigonométricas. Observamos primeiro que,

$$\frac{\sin \alpha'}{1 + \cos \alpha'} = \frac{2 \sin \frac{\alpha'}{2} \cos \frac{\alpha'}{2}}{2 \cos \frac{\alpha'}{2}} = \tan \frac{\alpha'}{2} \quad . \quad (153)$$

Usando as relações (151,152), temos:

$$\begin{aligned} \tan \frac{\alpha'}{2} &= \frac{\sin \alpha}{\Gamma[(1 + \frac{V}{c}) \cos \alpha + 1 + \frac{V}{c}]} = \frac{1}{\Gamma(1 + \frac{V}{c})} \frac{\sin \alpha}{1 + \cos \alpha} \\ &= \frac{1}{\Gamma(1 + \frac{V}{c})} \tan \frac{\alpha}{2} \quad . \end{aligned} \quad (154)$$

Utilizando a definição de  $\Gamma$ , obtemos finalmente a expressão relacionando o ângulo  $\alpha'$  com  $\alpha$ :

$$\tan \frac{\alpha'}{2} = \left( \sqrt{\frac{c - V}{c + V}} \right) \tan \frac{\alpha}{2} \quad . \quad (155)$$

Empregando agora uma expansão em série de Taylor, obtemos o limite newtoniano desta expressão:

$$\tan \frac{\alpha'}{2} = \left( 1 - \frac{V}{c} \right) \tan \frac{\alpha}{2} \quad . \quad (156)$$

## 14 Efeito Doppler

A frequência de uma onda que um observador detecta depende do estado do movimento da fonte e do observador. No caso de ondas que necessitam de um meio para se propagarem, a situação correspondendo à fonte estar em movimento e o observador em repouso não é equivalente a uma fonte em repouso

e um observador em movimento. Quando a onda (como por exemplo no caso do som) exige um meio para se propagar, existe um referencial em relação ao qual o estado de movimento é definido: o referencial do meio onde se propaga a onda. A fonte estar em movimento em relação ou o observador correspondem a situações físicas distintas. Nós vamos agora analisar esta questão, para o caso das ondas sonoras, em detalhe.

As ondas sonoras exigem um meio para se propagar. Vamos considerar conseqüentemente a seguinte situação: a fonte sonora se propaga com velocidade  $v_f$  em relação ao solo, o observador se desloca com velocidade  $v_0$  em relação ao solo, e o meio se desloca com velocidade  $v_m$  também em relação ao solo. A velocidade da onda em relação ao meio é  $v_s$ . Nós utilizaremos as leis de transformação galileanas, visto que as velocidades envolvidas são pequenas em relação à da luz.

No instante  $t = 0$  a fonte emite uma frente de onda. No tempo  $t = T$ ,  $T$  sendo o período, ela emite uma segunda frente de onda. Assim, a separação entre as duas frentes de onda é:

$$\lambda' = (v_s - v_m - v_f)T \quad . \quad (157)$$

O observador recebe a primeira frente de onda no instante  $t = t_1$  e a segunda frente de onda no instante  $t = t_2$ . Para ele, a velocidade da onda é  $v_s - v_m - v_0$ , e o período  $T' = t_2 - t_1$ . Tem-se então:

$$T' = \frac{\lambda'}{v_s - v_m - v_0} \quad . \quad (158)$$

Desta relação extraímos,

$$T' = \frac{(v_s - v_m - v_f)T}{v_s - v_m - v_0} \quad \Rightarrow \quad \nu' = \frac{v_s - v_m - v_0}{v_s - v_m - v_f} \nu \quad . \quad (159)$$

Esta expressão mostra a relação entre a frequência da fonte e a frequência medida pelo observador. Pode-se notar que constituem situações inteiramente distintas o observador estar em movimento em relação ao meio ou a fonte estar em movimento em relação ao meio. Ao contrário do que poderia se esperar, a relação entre a frequência medida e a emitida não depende unicamente da velocidade relativa. Isto ocorre porque neste caso existe um meio onde a onda se propaga, o que define um referencial privilegiado para o estudo das ondas sonoras.

A luz por sua vez não exige um meio para se propagar. Vamos analisar então o efeito Doppler para ondas luminosas; nesta análise empregaremos as transformações de Lorentz, já que se trata de um fenômeno essencialmente relativista. Consideremos um observador na origem de um referencial  $S$ , e uma partícula que se desloca com velocidade  $\vec{v}$  arbitrária. Esta velocidade pode ser decomposta em componentes radial e tangencial:  $\vec{v} = \vec{v}_r + \vec{v}_t$ . A partícula emite uma onda luminosa de frequência  $\nu = \frac{1}{T}$ , onde  $T$  é o período, comprimento de onda  $\lambda$ , medidos no referencial próprio da partícula. Pela dilatação do tempo, o tempo entre a emissão de duas frentes de onda consecutivas, medido no referencial  $S$ , será

$$T' = T\Gamma \quad , \quad \Gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad . \quad (160)$$

Por outro lado o comprimento de onda medido será de

$$\lambda' = cT' + T'v_r = (c + v_r)T\Gamma = \frac{(c + v_r)}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{\lambda}{c} = \frac{(1 + \frac{v_r}{c})}{\sqrt{1 - \frac{v^2}{c^2}}} \lambda \quad . \quad (161)$$

A relação entre a frequência medida por um observador no referencial da partícula,  $\nu$  e a frequência  $\nu'$  medida no referencial  $S$  é:

$$\nu' = \nu \frac{\sqrt{1 - \frac{v^2}{c^2}}}{1 + \frac{v_r}{c}} \quad . \quad (162)$$

A comparação entre as equações (159) e (162) revela, como esperado, diferenças essenciais:

1. A expressão relativista depende apenas da velocidade relativa, ao passo que a das ondas sonoras depende do movimento de cada um dos elementos em relação ao meio onde a onda se propaga. Isto é consequência do fato que as ondas luminosas não exigem um meio para se propagarem;
2. No limite de baixas velocidades (velocidades do meio, da fonte e do observador pequenas em relação à velocidade de propagação da onda no meio ou no vácuo para o caso da luz), as duas expressões coincidem e só dependem da velocidade relativa entre fonte e observador;

3. Na equação (162) existe o efeito Doppler mesmo quando a fonte não está se afastando do observador, como ocorre quando a fonte descreve um movimento circular uniforme em torno do observador. Esse é o assim denominado "efeito Doppler transversal" e se deve unicamente à dilatação temporal, constituindo-se um efeito puramente relativista.

## 15 Transformação das acelerações

Através das transformações de Lorentz é possível obter as coordenadas  $(t', x', y', z')$  de um evento  $P$  em um referencial  $S'$  a partir das coordenadas  $(t, x, y, z)$  deste mesmo evento medidas no referencial  $S$ . Empregando as transformações (137, 138, 139), obtemos as componentes da velocidade de um corpo no referencial  $S'$  a partir das componentes da velocidade deste mesmo corpo, medidas no referencial  $S$ . Agora, nós estabeleceremos as relações entre as componentes da aceleração medidas em  $S$  e em  $S'$ .

O procedimento é inteiramente análogo à relação entre as velocidades. No referencial  $S'$ , temos

$$a'_x = \frac{dv'_x}{dt'} \quad , \quad a'_y = \frac{dv'_y}{dt'} \quad , \quad a'_z = \frac{dv'_z}{dt'} \quad . \quad (163)$$

Como no caso das transformações das velocidades, empregamos novamente as transformações de Lorentz, agora combinadas com (137, 138, 139). Lembrando-se que  $V$ , a velocidade relativa entre os referenciais  $S$  e  $S'$  é constante, tem-se

$$\begin{aligned} a'_x &= \frac{dv'_x}{dt'} = \frac{1}{dt'} \left\{ \frac{dv_x}{1 - \frac{v_x V}{c^2}} + \frac{v_x - V}{(1 - \frac{v_x V}{c^2})^2} \frac{dv_x V}{c^2} \right\} = \\ &= \frac{1}{\Gamma(dt - \frac{V}{c^2} dx)} \left\{ \frac{dv_x}{1 - \frac{v_x V}{c^2}} + \frac{v_x - V}{(1 - \frac{v_x V}{c^2})^2} \frac{dv_x V}{c^2} \right\} = \\ &= \frac{1}{\Gamma(1 - \frac{v_x V}{c^2})^3} \left\{ a_x (1 - \frac{v_x V}{c^2}) + (v_x - V) \frac{a_x V}{c^2} \right\} = \\ &= \frac{a_x}{\Gamma^3(1 - \frac{v_x V}{c^2})^2} \quad . \end{aligned}$$

Procedemos de forma similar para as componentes  $y$  e  $z$  das acelerações:

$$\frac{dv'_y}{dt'} = \frac{\frac{dv_y}{\Gamma(1 - \frac{v_x V}{c^2})} + \frac{v_y \frac{dv_x V}{c^2}}{\Gamma(1 - \frac{v_x V}{c^2})^2}}{\Gamma(dt - \frac{V}{c^2} dx)} = \frac{\frac{a_y}{\Gamma(1 - \frac{v_x V}{c^2})} + \frac{v_y \frac{a_x V}{c^2}}{\Gamma(1 - \frac{v_x V}{c^2})^2}}{\gamma(1 - \frac{V v_x}{c^2})}$$

$$= \frac{1}{\Gamma^2(1 - \frac{v_x V}{c^2})^3} \left( a_y \left( 1 - \frac{v_x V}{c^2} \right) + v_y \frac{a_x V}{c^2} \right) .$$

Existe uma expressão semelhante para a relação entre  $a'_z$  e  $a_z$ . Desta forma, as transformações da aceleração são,

$$a'_x = \frac{a_x}{\Gamma^3(1 - \frac{v_x V}{c^2})^3} , \quad (164)$$

$$a'_y = \frac{1}{\Gamma^2(1 - \frac{v_x V}{c^2})^3} \left[ a_y \left( 1 - \frac{v_x V}{c^2} \right) + v_y \frac{a_x V}{c^2} \right] , \quad (165)$$

$$a'_z = \frac{1}{\Gamma^2(1 - \frac{v_x V}{c^2})^3} \left[ a_z \left( 1 - \frac{v_x V}{c^2} \right) + v_z \frac{a_x V}{c^2} \right] . \quad (166)$$

As transformações (164,165,166) têm algumas propriedades importantes. Em primeiro lugar, elas se reduzem às transformações galileanas correspondentes ( $a'_x = a_x$ ,  $a'_y = a_y$ ,  $a'_z = a_z$ ) obtidas a partir da condição  $c \rightarrow \infty$  ( o que significa considerar velocidades muito inferiores à da luz ). Entretanto, é fundamental observar que elas implicam que as expressões para a aceleração não se mantêm invariantes quando se passa de um referencial  $S$  para um referencial  $S'$ . Consequentemente, as expressões usuais de força não preservam sua forma. Isto era de se esperar, visto que a força medida em um referencial  $S$  era idêntica à força medida em um referencial  $S'$  quando se utilizava as transformações de Galileu. Como nós estamos agora empregando as transformações de Lorentz, nós teremos que modificar as relações dinâmicas entre uma força e as quantidades cinemáticas. Em outras palavras, teremos que reformular a Mecânica.

Por outro lado, observemos que, se em um referencial  $S$  só existe aceleração na direção  $x$ , no referencial  $S'$  pode-se ter acelerações nas direções  $y$  e  $z$ , caso as velocidades  $v_y$  e  $v_z$  não sejam nulas. Isto é resultado dos últimos termos das equações (165,166). Esses termos são usualmente denominados de acelerações transversais. Sua origem reside no fenômeno da dilatação do tempo. De fato, mesmo que o movimento relativo entre os dois referenciais seja ao longo do eixo  $x$ , as velocidades medidas ao longo dos eixos  $y$  e  $z$  não são as mesmas nos dois referenciais, e a própria dedução das expressões (138,139) mostra que isto se deve à não invariância do tempo quando se passa de um referencial a outro. Novamente, isto origina acelerações transversais, no sentido que a aceleração de um corpo ao longo do eixo  $x$  provoca o aparecimento, no referencial  $S'$ , de

de acelerações nas direções perpendiculares, a menos que a velocidade do corpo nestas direções perpendiculares seja nula.

Uma vez que mostramos que a forma para as acelerações não se preserva quando se passa do referencial  $S$  para o referencial  $S'$  e se utiliza as transformações de Lorentz para se realizar esta passagem, é inevitável se perguntar o que permanece idêntico nesta passagem. Para ser preciso, é necessário distinguir o que permanece invariante, quer dizer, o que preserva o seu valor numérico sob aquela transformação, e o que é covariante, quer dizer, o que preserva sua relação formal. Mas, para identificar o que é invariante e o que é covariante, é preciso reformular tudo o que foi feito até agora, expressando todas as quantidades como grandezas definidas em um espaço-tempo a quatro dimensões (três espaciais e uma temporal), dotado de uma estrutura hiperbólica, quer dizer, o espaço-tempo de Minkowski. Além do mais, em função do fato que estaremos definindo uma nova estrutura geométrica, será preciso desenvolver a noção de tensor, como generalização dos conceitos usuais de escalar e vetor. É o que faremos a seguir.

## 16 O espaço-tempo de Minkowski

Um feixe de luz se propaga de tal forma que a relação entre a variação no tempo e a nas coordenadas espaciais é dada por

$$c^2(\Delta t)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \quad . \quad (167)$$

Outra forma de escrever esta expressão é

$$c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 = 0 \quad . \quad (168)$$

Se um objeto se propaga com uma velocidade inferior à da luz, temos

$$c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 > 0 \quad , \quad (169)$$

já que, no mesmo tempo que a luz percorre uma certa distância, este objeto obviamente percorreria uma distância menor. Por outro lado, se fosse possível que um objeto viajasse com velocidade superior a da luz, teríamos,

$$c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 < 0 \quad . \quad (170)$$

Desta forma, podemos definir a quantidade

$$c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 = (\Delta s)^2 \quad , \quad (171)$$

sendo que  $(\Delta s)^2$  pode ser maior, menor ou igual a zero. O fato de  $\Delta s^2$  poder ser negativo é algo, para um geômetra euclidiano, chocante, uma vez que isto implica em  $\Delta s$  imaginário. Mas, é preciso ter em mente que estamos falando de uma "distância" espaço-temporal, e que não devemos nos prendermos conceitos oriundos da geometria euclidiana, como o que diz que a distância entre dois pontos deve ser sempre positiva: no caso euclidiano usual, esta distância é puramente espacial, e a positividade se impõe. Tal imposição (e mesmo a imposição que o valor seja real) não ocorre no caso de uma distância espaço-temporal.

Nós definimos o elemento de distância espaço-temporal entre dois eventos  $E_1$  e  $E_2$  com coordenadas  $(t_1, x_1, y_1, z_1)$  e  $(t_2, x_2, y_2, z_2)$ , medidas em um referencial  $S$ , como

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \quad . \quad (172)$$

Escrita na forma diferencial, (172) torna-se,

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad . \quad (173)$$

As transformações de Lorentz deixam invariante (172). A demonstração é simples: considere-se os mesmos eventos  $E_1$  e  $E_2$  medidos no referencial  $S'$  que se desloca com velocidade constante  $V$  em relação a  $S$ . Suas coordenadas neste referencial são  $(t'_1, x'_1, y'_1, z'_1)$  e  $(t'_2, x'_2, y'_2, z'_2)$ . Logo,

$$\Delta t = \Gamma(\Delta t' + \frac{V}{c^2} \Delta x') \quad , \quad (174)$$

$$\Delta x = \Gamma(\Delta x' + V \Delta t') \quad , \quad (175)$$

$$y = y' \quad , \quad z = z' \quad . \quad (176)$$

Inserindo (174,175,176) em (172), obtemos,

$$\begin{aligned} \Delta s^2 &= \Gamma^2 \left( c^2 \Delta t'^2 + \frac{V^2}{c^2} \Delta x'^2 + 2 \frac{V}{\Delta} \Delta t' \Delta x' - \Delta x'^2 - V^2 \Delta t'^2 - 2V \Delta t' \Delta x' \right) \\ &- \Delta y'^2 - \Delta z'^2 \\ &= c^2 \Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2 = \Delta s'^2 \quad , \end{aligned}$$

onde usamos explicitamente que  $\Gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ . Quando  $\Delta s^2$  é zero, temos a propagação de um raio luminoso: os dois eventos  $E_1$  e  $E_2$  representam a passagem de uma frente de onda luminosa no ponto  $P_1 = (x_1, y_1, z_1)$  no tempo  $t_1$  e a passagem desta mesma frente luminosa no ponto  $P_2 = (x_2, y_2, z_2)$ , no tempo  $t_2$ . Por outro lado, se  $\Delta s^2 > 0$ , isto significa que um sinal que se propaga com velocidade inferior a da luz pode ser emitido de  $P_1$  no tempo  $t_1$  e alcançar  $P_2$  no tempo  $t_2$ . Finalmente, se  $\Delta s^2 < 0$ <sup>10</sup>, para que um sinal saia de  $P_1$  no tempo  $t_1$  e alcance  $P_2$  no tempo  $t_2$ , sua velocidade teria que ser necessariamente superior à da luz. Observe que a estrutura causal descrita acima é preservada pelas transformações de Lorentz.

Podemos representar os eventos descritos acima em um diagrama do espaço-tempo definido por um sistema de coordenadas contendo  $t$ ,  $x$ ,  $y$  e  $z$ . Entretanto, a representação gráfica de um tal sistema quadridimensional se revela impossível. Por esta razão, e para fins de representação gráfica, nós suprimimos as coordenadas  $y$  e  $z$ , e construímos um diagrama contendo unicamente  $t$  e  $x$ . Se além disto, ao invés de considerarmos a coordenada  $t$ , nós empregarmos  $ct$ , de maneira a ter as mesmas dimensões que as demais coordenadas espaciais, a trajetória do raio luminoso será representada no diagrama  $(ct, x)$  pela equação  $ct = \pm x$ , ou seja, por retas de inclinação igual a  $\pm \frac{\pi}{4} \text{ rad}$ . O sinal positivo refere-se a um raio luminoso que se propaga na direção positiva dos  $x$  e o sinal negativo refere-se à propagação de um raio luminoso na direção negativa dos  $x$ . Se considerarmos a coordenada  $y$ , a propagação dos raios luminosos define um cone usualmente denominado *cone de luz*; incluindo as coordenadas  $y$  e  $z$  temos um hipercone, um cone com superfície espacial tridimensional, em um espaço-tempo a quatro dimensões.

Retornemos à expressão (173). Tomemos agora um objeto que se propaga com velocidade inferior à da luz. Neste caso, em um mesmo intervalo de tempo referente à propagação de um raio luminoso, a distância espacial percorrida por este objeto é menor que a percorrida pela luz. Logo, dois eventos associados à

---

<sup>10</sup>Neste caso,  $\Delta s$  torna-se imaginário. Mas, como já foi enfatizado anteriormente, isto não representa nenhum problema pois  $\Delta s$  seria uma "distância" no espaço-tempo, não tendo uma contrapartida com a noção usual de distância tal como empregada no sentido puramente espacial



propagação deste objeto têm  $ds^2 > 0$ . Se associarmos a origem a um dos eventos (digamos  $E_1$ ), o outro, que denominamos  $E_2$ , estará localizado dentro do cone de luz. Se o evento  $E_2$  ocorrer antes do  $E_1$ , ele estará no interior da parte inferior do cone de luz; se ele ocorrer depois, ele estará no interior da parte superior do cone de luz. Assim, todos os fenômenos que estão no interior da parte superior do cone de luz podem ter sido influenciados pelo evento  $E_1$  na origem, no sentido que um observador em  $E_1$  pode enviar um sinal que se propaga com velocidade inferior à da luz, que será recebido por um observador em  $E_2$ . Inversamente, se  $E_2$  estiver localizado na parte inferior do cone, um sinal que se propaga com velocidade inferior a da luz, emitido por um observador situado neste ponto do espaço-tempo, alcançará  $E_1$ . Em função disto, a parte superior do cone de luz é denominado de *futuro* do evento  $E_1$ , que está localizado na origem, no sentido que ela reúne todos os eventos que receberão sinais subluminais enviados por  $E_1$ . Por outro lado,  $E_1$  terá recebido sinais também subluminais enviados por todos os eventos que estão na parte inferior do cone de luz. Por esta razão, a parte inferior do cone de luz é denominada de *passado* do evento da origem.

Consideremos agora dois eventos tais que  $\Delta s^2 < 0$ . Chamemos um dos eventos novamente de  $E_1$ , localizado na origem, e outro de  $E_3$ . Neste caso, para que um sinal enviado por  $E_1$  alcance  $E_3$ , seria necessário que ele tivesse uma velocidade superior à da luz. Desta forma, uma vez que não existe propagação de nenhum sinal com velocidade superior à da luz<sup>11</sup>, os eventos que estão localizados na parte exterior do cone de luz não têm nenhuma relação causal com o que ocorre na origem. Observamos que esta distinção entre eventos futuros, passados e eventos sem conexão causal, se referem ao evento  $E_1$  localizado na origem: dois eventos podem estar no futuro do evento  $E_1$ , e terem consequentemente conexões causais com ele, e não terem nenhuma relação causal entre si. Por outro lado, o evento  $E_2$ , que está no interior do cone de luz de  $E_1$ , tendo consequentemente uma relação causal com  $E_1$ , pode ter conexão causal com o evento  $E_3$ , que por sua vez não tem relação causal com  $E_1$ . Cada ponto do espaço-tempo possui seu próprio cone de luz, com passado e futuro próprios.

---

<sup>11</sup>Observe que se algum objeto ou sinal se deslocar com velocidade superior a  $c$ , poderíamos associar a ele um referencial inercial para o qual o fator  $\Gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$  seria imaginário e toda estrutura lorentziana empregada até aqui perderia todo o sentido

Pares de eventos tais que  $\Delta s^2 > 0$  são chamados *tipo tempo*, visto que o termo temporal predomina sobre o espacial. Se  $\Delta s^2 = 0$ , os eventos são chamados tipo luz, já que eles se localizam sobre o cone de luz. Finalmente, se  $\Delta s^2 < 0$ , os eventos são denominados *tipo espaço*, já que o termo espacial predomina sobre o temporal. Dois eventos tipo tempo jamais poderão ser feitos simultâneos por uma transformação de Lorentz (uma vez que isto implicaria em anular o  $\Delta t$ , e conseqüentemente o  $\Delta s$  se tornaria negativo, contrariando a condição de ser tipo tempo), mas poderão ser transportados, num dado referencial, à uma mesma posição espacial; dois eventos tipo espaço, poderão ser feitos simultâneos num dado referencial, mas jamais poderão ter suas posições espaciais coincidentes. Por fim, dois eventos tipo luz em um dado referencial, continuarão o sendo em qualquer outro referencial inercial.

## 17 Transformações de Lorentz e rotações no espaço-tempo de Minkowski

É importante notar que o espaço-tempo de Minkowski possui uma estrutura hiperbólica, devido à diferença de sinal entre a coordenada temporal e as coordenadas espaciais na expressão do  $\Delta s^2$ . As transformações de Lorentz podem ser consideradas como rotações neste espaço-tempo hiperbólico, cuja estrutura geométrica é não-euclídeana. Vamos considerar inicialmente a rotação de um sistema de coordenadas  $S$  em um espaço euclídeano bi-dimensional onde se definem os eixos  $(x, y)$ . Definimos um novo sistema de coordenadas  $S'$ , com eixos  $(x', y')$ ,  $x'$  e  $y'$  fazendo um ângulo  $\theta$  com os eixos  $x$  e  $y$  originais. Consideremos um ponto  $P$ , situado a uma distância  $r$  da origem, que é comum a  $S$  e a  $S'$ . O ângulo entre o vetor  $\vec{r}$ , que liga a origem ao ponto  $P$ , e o eixo  $x$  é  $\phi$ ; por outro lado, o ângulo entre  $\vec{r}$  e o eixo  $x'$  é  $\phi - \theta$ . Desta forma, nós podemos calcular as coordenadas  $x'_0$  e  $y'_0$  do ponto  $P$  em  $S'$  e relacioná-las com as coordenadas  $x_0$  e  $y_0$  em  $S$ . Lembrando que  $x_0 = r \cos \phi$  e  $y_0 = r \sin \phi$ , temos

$$\begin{aligned} x'_0 &= r \cos(\phi - \theta) = r \left( \cos \phi \cos \theta + \sin \phi \sin \theta \right) \\ &= x_0 \cos \theta + y_0 \sin \theta \quad , \end{aligned} \tag{177}$$

$$\begin{aligned}
y'_0 &= r \sin(\phi - \theta) = r \left( \sin \phi \cos \theta - \sin \theta \cos \phi \right) \\
&= y_0 \cos \theta - x_0 \sin \theta \quad .
\end{aligned} \tag{178}$$

Esta transformação pode ser representada por uma forma matricial,

$$x^{i'}_0 = \sum_1^2 R^i_j x^j_0 \quad , \tag{179}$$

onde

$$R^i_j = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad . \tag{180}$$

A matriz  $R^i_j$  é uma representação do grupo  $SO(2)$ , grupo de rotações espaciais a duas dimensões. De fato, todos os requisitos necessários para que esta matriz represente um grupo são satisfeitos: a transformação unitária é um caso particular; existe uma transformação inversa; o produto de duas rotações espaciais, uma de ângulo  $\theta_1$ , outra de um ângulo  $\theta_2$ , é ainda uma rotação espacial de um ângulo  $\theta_1 + \theta_2$ . Note que o fato da matriz transformação ser de determinante unitário é fundamental para que se verifiquem estas propriedades de grupo. Além do mais, o elemento de distância do ponto  $P$  à origem é preservado pela rotação:

$$r^2 = x_0^2 + y_0^2 = x_0'^2 + y_0'^2 \quad . \tag{181}$$

Na nomenclatura deste grupo,  $O$  indica que o grupo é ortogonal. Quer dizer, definindo a transposta  $R_i^{jT} = R^j_i$ , temos que  $R^i_j R^j_k = \delta^i_k$ : o produto da matriz por sua transposta é igual à matriz unitária. Por outro lado, o termo  $S$  indica um grupo *especial*, quer dizer, de determinante igual a 1.

Consideremos agora as transformações de Lorentz, levando em conta apenas as coordenadas  $x, t$ . Temos,

$$t' = \Gamma \left( t - \frac{V}{c^2} x \right) \quad , \tag{182}$$

$$x' = \Gamma \left( x - \frac{V}{t} \right) \quad . \tag{183}$$

Como já foi dito anteriormente, o "elemento de distância espaço-temporal" é preservado por esta transformação. Se definirmos  $\Gamma = \cosh \theta$  e  $\Gamma \frac{V}{c} = \sinh \theta$ , de tal forma que a velocidade relativa dos referenciais seja dada por

$$V = c \tanh \theta \quad , \tag{184}$$

e escrevermos  $x^0 = ct$ ,  $x^1 = x$ , obtemos,

$$x^{0'} = \cosh \theta x^0 - \sinh \theta x^1 \quad , \quad (185)$$

$$x^{1'} = \cosh \theta x^1 - \sinh \theta x^0 \quad . \quad (186)$$

Um evento  $E$  constitui um ponto no diagrama espaço-temporal. As coordenadas  $(x^0, x^1)$  deste evento, são determinadas através da expressão,

$$x^0 = \Delta s \cosh \phi \quad , \quad x^1 = \Delta s \sinh \phi \quad , \quad (187)$$

onde  $\phi$  é o ângulo entre a reta que liga a origem ao evento  $E$  e o eixo  $x^1$ . Esta expressão é condizente com o fato que o elemento de distância no espaço-tempo, em termos das coordenadas  $x^0$  e  $x^1$  é dado por

$$\Delta s^2 = \Delta x^{02} - \Delta x^{12} \quad . \quad (188)$$

Este fato determina a natureza particular da geometria do espaço-tempo: as projeções do elemento de distância espaço-temporal em cada eixo, são feitas através de funções hiperbólicas, e não trigonométricas. Isto é consequência direta da forma de calcular este elemento de distância (188), onde os termos espacial e temporal entram com sinais opostos. Observe que o fato do sinal de  $x^0$  ser positivo e o de  $x^1$  ser negativo, não desempenha um papel essencial; o fundamental é que eles entram na expressão do  $\Delta s^2$  com sinais opostos.

Observamos então que a passagem do referencial  $S$  para o referencial  $S'$  corresponde a uma rotação hiperbólica dos eixos, como indicam as relações (174,175). Como representar graficamente esta transformação? Notemos que se o eixo  $x^{1'}$  faz um ângulo  $\theta$  com o eixo  $x^1$ , as projeções do  $\Delta s$ , que deve ser o mesmo nos dois referenciais, são dadas por,

$$\begin{aligned} x^{0'} &= \Delta s \cosh(\phi - \theta) = \Delta s(\cosh \phi \cosh \theta - \sinh \phi \sinh \theta) \\ &= x^0 \cosh \theta - x^1 \sinh \theta \quad , \end{aligned} \quad (189)$$

$$\begin{aligned} x^{1'} &= \Delta s \sinh(\phi - \theta) = \Delta s(\sinh \phi \cosh \theta - \cosh \phi \sinh \theta) \\ &= x^1 \cosh \theta - x^0 \sinh \theta \quad , \end{aligned} \quad (190)$$

onde usamos as decomposições de  $\Delta s$  nos eixos  $x^0$  e  $x^1$  (187), e reobtivemos as transformações de Lorentz (185,186). Surge agora a questão de como represen-

tar os eixos  $x^{0'}$  e  $x^{1'}$  em relação aos eixos  $x^0$  e  $x^1$  já construídos. Para isso, recordemos o segundo axioma da Relatividade Restrita: a velocidade da luz é a mesma para os dois referenciais e tal que  $\Delta s = \Delta s' = 0$ . No eixo  $(x^0, x^1)$ , isto implicava em que a propagação da luz era representada por uma reta fazendo um ângulo de  $\frac{\pi}{4}$  com o eixo  $x^1$  (ou, equivalentemente,  $x^0$ ). A forma de "girar" os eixos é aparentemente muito exótica, mas observe que é a única maneira que garante que o observador em  $S'$  veja a luz se propagando com velocidade  $c$ , e que conseqüentemente o raio luminoso seja a bissetriz dos novos eixos. Note-se também que estão sendo representados os eixos  $x^{0'}$  e  $x^{1'}$  tais como vistos pelo observador em  $S$ . Neste último caso, a partir de (184), isto requer  $\theta \rightarrow \infty$ , coerentemente com o fato que na geometria hiperbólica  $-\infty < \theta < +\infty$ .

A necessidade de girar os eixos  $(ct, x)$  da forma indicada também pode ser entendida observando que ela preserva a estrutura causal: eventos que estão conectados causalmente no referencial  $S$  continuam o sendo no referencial  $S'$ . Isto não aconteceria se os eixos fossem rodados da mesma maneira que no caso euclidiano. Além do mais, no limite em que a velocidade relativa entre os referenciais tende para a velocidade da luz, os novos eixos  $(ct', x')$  passam a coincidir com o cone de luz, como é de se esperar.

A forma como os pontos no espaço-tempo são decompostos nos eixos respectivos, pode ser entendida através da definição de um ângulo complexo  $\phi = i\theta$ . Definindo agora os eixos onde  $x^0 = ict$ , temos as relações para a decomposição do ponto  $E_0$ ,

$$x^0 = \Delta s \cos \phi \quad , \quad x^1 = \Delta s \sin \phi \quad . \quad (191)$$

### 17.1 Visão geométrica da dilatação do tempo e da contração das distâncias

A rotação hiperbólica descrita anteriormente mantém invariante tanto a propagação dos raios luminosos, quanto as hipérboles definidas nesse espaço tempo. De fato, a equação

$$\Delta s^2 = \Delta x_0^2 - \Delta x_1^2 \quad (192)$$

define uma hipérbole quando  $\Delta s^2$  é uma dada constante, positiva ou negativa, que é a mesma para o referencial  $S$  e para o referencial  $S'$ , devido à invariância de

Lorentz. Retornemos ao elemento de distância espaço-temporal de dois eventos escritos nos referenciais  $S$  e  $S'$ :

$$\Delta s^2 = \Delta x_0^2 - \Delta x_1^2 = \Delta x'_0{}^2 - \Delta x'_1{}^2 . \quad (193)$$

Se o evento é do tipo luz, de tal forma que  $\Delta s^2 = 0$ , temos

$$\frac{\Delta x_1}{\Delta x_0} = \frac{\Delta x'_1}{\Delta x'_1} = \pm 1 , \quad (194)$$

Por outro lado, se o evento é do tipo tempo,  $\Delta s^2 > 0$ , e a hipérbole definida pela relação (193) corta os eixos  $x_0$  e  $x'_0$ ; se o evento é do tipo espaço, então a hipérbole resultante cortará os eixos  $x_1$  e  $x'_1$ . Essas hipérboles, que se mantêm invariantes pelas transformações de Lorentz uma vez que  $\Delta s^2 = \Delta s'^2$ , nos permitem entender de uma forma mais geométrica os fenômenos de dilatação temporal e contração das distâncias.

Consideremos inicialmente o caso da dilatação temporal. Os observadores nos referenciais  $S$  e  $S'$  querem usar o mesmo padrão para a unidade de tempo, que eles denominarão *segundo*. Para tanto, eles utilizam eventos que se localizam sobre o eixo temporal (o que equivale a utilizar relógios em repouso no seu próprio referencial), e que correspondam a fenômenos físicos idênticos (por exemplo, um certo múltiplo da frequência de uma dada radiação eletromagnética). Assim, eles utilizarão o mesmo  $\Delta s^2 > 0$ . Isto equivale a utilizar uma hipérbole definida por este  $\Delta s^2 > 0$ , e que corta o eixo  $x_0$  no ponto  $A$  e o eixo  $x'_0$  no ponto  $A'$ . Esses pontos definem a unidade de tempo nos referenciais  $S$  e  $S'$ , usando um mesmo padrão. Entretanto, observamos que a projecção de  $A'$  sobre o eixo  $x_0$ , nos fornece,

$$T = T' \cosh \phi = \Gamma T' , \quad (195)$$

O *segundo* definido no referencial  $S'$  equivale a um tempo igual ao segundo do referencial  $S$  multiplicado por um fator  $\Gamma > 1$ : o observador do referencial  $S$  vê o relógio do referencial  $S'$  se atrasando em relação ao seu próprio relógio, muito embora eles sejam construídos com o mesmo princípio físico. Isto constitui a dilatação temporal. Visto do referencial  $S'$ , o diagrama seria o mesmo, trocando os termos com linha pelos sem linha, e vice-versa. Assim, usando este mesmo padrão, o observador de  $S'$  veria o relógio de  $S$  se atrasando. Observe que este fenômeno se deve a uma perspectiva de referencial, isto é, como o observador

em  $S$  vê os eixos (em princípio idênticos) de  $S'$ : ele os vê inclinados em relação aos seus próprios eixos.

Par o caso da contração espacial, nós devemos tomar duas hipérboles, cada uma correspondendo a uma das extremidades da régua padrão. O comprimento da régua em  $S'$  é dada pela distância  $L_0 = x'_2 - x'_1$ . Entretanto, esta distância deve ser projetada em uma reta a tempo constante no referencial  $S$ , já que uma medida de distância entre dois pontos deverá ser feita assinalando a posição destes dois pontos simultaneamente, para que realmente isto corresponda a uma verdadeira distância medida em  $S$ . Para analisar qual é esta distância, devemos tomar pontos das duas hipérboles que correspondam a tempos idênticos no referencial  $S$  e projetar no eixo  $x'_1$ . Desta forma, encontramos que o comprimento da régua medida em  $S$  é, em relação ao comprimento dela medida em  $S'$ ,

$$L = \frac{L_0}{\cosh \theta} \quad . \quad (196)$$

Isto corresponde à contração das distâncias, e novamente é um efeito de perspectiva de como um referencial vê o outro.

## 17.2 Vetores Contravariantes e Covariantes

Quantidades físicas diferentes se transformam diferentemente quando da passagem de um sistema de coordenadas a outro (ou de um referencial a outro). Consideremos uma transformação de um sistema de coordenadas  $x^{\mu'}$  para  $x^\mu$ . Se a transformação é inversível, a matriz jacobiana desta transformação, definida como  $\Delta^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^\nu}$ , possui determinante não nulo. Uma transformação de coordenadas deste tipo pode ser de natureza passiva ou ativa. No caso da transformação passiva, nós mudamos unicamente as coordenadas permanecendo no mesmo referencial; na transformação ativa, nós passamos para um outro referencial inercial. A passagem do sistema  $x, y, z$  cartesiano para o sistema esférico  $r, \theta, \phi$  é uma transformação passiva. As transformações de Lorentz, assim como as de Galileu, são ativas.

Consideremos por exemplo, agora, a diferencial  $dx^{\mu'}$ , que são as componentes do vetor  $d\vec{r}$ . Quando passamos do sistema  $x^{\mu'}$  para o sistema  $x^\mu$ , temos então

que estas componentes se transformam como

$$dx^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\nu} dx^\nu = \Delta^{\mu'}_{\nu} dx^\nu \quad . \quad (197)$$

Por outro lado, se considerarmos as componentes do operador vetorial nabla  $\vec{\nabla}$ , dadas por  $\frac{\partial}{\partial x^{\mu'}}$ , temos

$$\frac{\partial}{\partial x^{\mu'}} = \frac{\partial x^\nu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\nu} = \Delta^\nu_{\mu'} \frac{\partial}{\partial x^\nu} \quad . \quad (198)$$

Note-se que em (198) surge a inversa da matriz jacobiana ao contrário do que ocorre em (197). As duas grandezas se transformam diferentemente na passagem de um sistema de coordenadas a outro. Em uma linguagem mais tradicional  $d\vec{r}$  é chamado de vetor, enquanto  $\vec{\nabla}$  é denominado de co-vetor.

Afim de simplificar a notação, escreveremos daqui por diante

$$\Delta^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^\nu} \quad , \quad (199)$$

designando a matriz Jacobiana, e

$$\Delta^\mu_{\nu'} = \frac{\partial x^\mu}{\partial x^{\nu'}} \quad , \quad (200)$$

designando sua inversa. Destas definições resulta a propriedade,

$$\Delta^{\mu'}_{\nu} \Delta^\nu_{\lambda'} = \delta^{\mu'}_{\lambda'} \quad (201)$$

sendo  $\delta^\mu_{\nu}$  o delta de Kronecker, que é igual a um quando  $\mu = \nu$  e igual a 0 quando  $\mu \neq \nu$ .

Uma quantidade vetorial cujas as componentes  $A^\mu$  se transformam como

$$A^{\mu'} = \Delta^{\mu'}_{\nu} A^\nu \quad , \quad (202)$$

quando se passa do sistema de coordenadas  $x^{\mu'}$  para o sistema de coordenadas  $x^\nu$  é dita contravariante. Por outro lado uma quantidade vetorial (na verdade, co-vetorial), de componentes  $B_\mu$  que se transformam como

$$B_{\mu'} = \Delta^\nu_{\mu'} B_\nu \quad (203)$$

é dita covariante.



Podemos definir o produto de um tensor covariante com um contravariante, resultando em uma grandeza escalar:

$$A^\mu B_\mu = C \quad . \quad (204)$$

Podemos mostrar que  $C$  é de fato um escalar estudando o seu comportamento por uma transformação de coordenadas:

$$\begin{aligned} C' &= A^{\mu'} B_{\mu'} = \Delta^{\mu'}_\nu A^\nu \Delta^\lambda_{\mu'} B_\lambda = \\ &= \Delta^{\mu'}_\nu \Delta^\lambda_{\mu'} A^\nu B_\lambda = \delta^\lambda_\nu A^\nu B_\lambda = A^\nu B_\nu = C \quad . \end{aligned} \quad (205)$$

A função  $C$  não sente a transformação de coordenadas, sendo efetivamente um escalar.

## 18 Interpretação geométrica dos vetores covariantes e contravariantes

Quando lidamos com a geometria euclidiana normal, sobretudo quando se empregam coordenadas cartesianas, a distinção entre vetores contravariante e covariantes (ou, vetores e co-vetores) é, muito frequentemente, irrelevante. Entretanto, podemos ter uma noção do sentido desta distinção considerando, em um espaço euclidiano ordinário, um sistema de coordenadas não convencional. De fato, considere o sistema de coordenadas cartesiano normal, a duas dimensões, com os eixos  $x$  e  $y$  perpendiculares entre si. Considere agora um vetor neste espaço e sua projeção naqueles eixos. A projeção do vetor no eixo  $x$  (o mesmo poderá ser dito da projeção no eixo  $y$ ) pode ser definido de duas formas equivalentes para este caso: traça-se uma linha da extremidade do vetor ao eixo  $x$  que seja perpendicular a este eixo; ou traça-se uma linha da extremidade do vetor ao eixo  $x$  que seja paralela ao eixo  $y$ . O resultado é o mesmo nos dois casos.

Por outro lado considere agora um novo sistema de coordenadas  $(x', y')$  em que os respectivos eixos não sejam perpendiculares, mas tal que o eixo  $x'$  faça um ângulo  $\theta$  com o antigo eixo  $x$ , enquanto o eixo  $y'$  coincide com o eixo  $y$ . Para o novo sistema de coordenadas  $(x', y')$ , os dois procedimentos acima conduzem a resultados diferentes. Chamaremos os sistemas  $(x, y)$  e  $(x', y')$  como sistemas ortogonal e não ortogonal, respectivamente.

Um vetor  $\vec{r}$  se escreve no sistema de coordenadas  $(x, y)$  e  $(x', y')$ , respectivamente, como

$$\vec{r} = x\vec{i} + y\vec{j} = x'\vec{i}' + y'\vec{j}' \quad , \quad (206)$$

onde  $(\vec{i}, \vec{j})$  e  $(\vec{i}', \vec{j}')$  são os vetores unitários nos sistemas ortogonal e não ortogonal, respectivamente. Temos

$$\vec{i} \cdot \vec{i} = \cos \theta \quad , \quad \vec{i} \cdot \vec{j}' = 0 \quad , \quad \vec{j} \cdot \vec{i} = \sin \theta \quad , \quad \vec{j} \cdot \vec{j}' = 1 \quad . \quad (207)$$

Suponhamos agora que o ângulo do vetor  $\vec{r}$  com o eixo dos  $x$  seja  $\phi$ . Neste caso, se projetarmos este vetor segundo uma paralela ao eixo dos  $x(y)$ , obtemos

$$x = r \cos \phi \quad , \quad y = r \sin \phi \quad . \quad (208)$$

Encontremos agora a projeção nos novos eixos  $(x', y')$ . Temos, por definição para um vetor, e utilizando as decomposições (206) e as relações (207)

$$\vec{r} \cdot \vec{i}' = x' + y' \sin \theta = x \cos \theta + y \sin \theta \quad , \quad (209)$$

$$\vec{r} \cdot \vec{j}' = y' + x' \sin \theta = y \quad . \quad (210)$$

Podemos expressar então as novas coordenadas em termos das antigas e vice-versa:

$$x' = \frac{x}{\cos \theta} \quad , \quad y' = -x \tan \theta + y \quad ; \quad (211)$$

$$x = x' \cos \theta \quad , \quad y = x' \sin \theta + y' \quad . \quad (212)$$

Estas transformações podem ser escritas como,

$$x^{\lambda'} = \Delta_{\nu}^{\lambda'} x^{\nu} \quad , \quad x^{\nu} = \Delta_{\lambda'}^{\nu} x^{\lambda} \quad , \quad (213)$$

sendo que  $\lambda, \nu = 1, 2$ , e

$$\Delta_{\nu}^{\lambda'} = \frac{\partial x^{\lambda'}}{\partial x^{\nu}} \quad , \quad \Delta_{\lambda'}^{\nu} = \frac{\partial x^{\nu}}{\partial x^{\lambda'}} \quad , \quad (214)$$

são a matriz transformação do sistema de coordenadas  $(x', y')$  para o sistema de coordenada  $(x, y)$  e sua inversa. Estas transformações podem ser expressas na forma matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\cos \theta} & 0 \\ -\tan \theta & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad , \quad (215)$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad . \quad (216)$$

O que discutiu-se até agora foi a transformação das componentes de um genuíno vetor quando se passa de um sistema de coordenadas ortogonal para um sistema de coordenadas não ortogonal. O que garante que estamos lidando com vetores usuais é a decomposição (206), que faz uso explícito da regra do paralelograma para a composição de vetores. Assim, os vetores se representam como matrizes colunas. O que ocorre se considerarmos agora uma decomposição onde as componentes da quantidade  $\vec{r}$  sejam obtidas através de uma perpendicular ao eixo desejado? Obviamente esta quantidade  $\vec{r}$  (para a qual usamos a mesma notação que a empregada para os vetores usuais) não poderá ser decomposta como em (206). Por outro lado obtemos agora,

$$x = r \cos \phi \quad , \quad y = r \sin \phi \quad , \quad (217)$$

$$x' = r \cos(\phi - \theta) \quad , \quad y' = r \sin \phi \quad . \quad (218)$$

Logo,

$$x' = x \cos \theta + y \sin \theta \quad , \quad y' = y \quad , \quad (219)$$

$$x = \frac{x'}{\cos \theta} - y \tan \theta \quad , \quad y = y' \quad . \quad (220)$$

Do ponto de vista matricial, esta transformação pode ser escrita como,

$$\begin{pmatrix} x' & y' \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 1 \end{pmatrix} \quad , \quad (221)$$

$$\begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} \frac{1}{\cos \theta} & 0 \\ -\tan \theta & 1 \end{pmatrix} \quad . \quad (222)$$

Há duas observações importantes nas transformações (221,222) acima: elas têm a forma inversa em relação às transformações vetoriais (215,216); as projeções perpendiculares aos eixos implicam em representar o "vetor" como uma matriz coluna. Essas quantidades são, na verdade, co-vetores ou, em uma linguagem mais moderna, vetores covariantes. Esses vetores são duais aos vetores usuais. Isto quer dizer, que os co-vetores (vetores) agem sobre os vetores (co-vetores) de maneira a resultar em um escalar, quer dizer, uma quantidade que não é afetada pela transformação de coordenadas. De fato, escrevendo  $(x_v, y_v)$  e  $(x_c, y_c)$  como as componentes de um vetor e de um co-vetor respectivamente, temos

$$\begin{pmatrix} x'_c & y'_c \end{pmatrix} \begin{pmatrix} x_v \\ y_v \end{pmatrix} = \begin{pmatrix} x_c & y_c \end{pmatrix} \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\cos \theta} & 0 \\ -\tan \theta & 1 \end{pmatrix} \begin{pmatrix} x_v \\ y_v \end{pmatrix} \quad ,$$

$$\rightarrow x'_c x'_v + y'_c y_v = x_c x_v + y_c y_v \quad . \quad (223)$$

Isto implica que o produto de um vetor por um co-vetor é um escalar, uma quantidade que permanece invariante por transformação de coordenadas.

Um espaço dado sempre possui o seu dual. Os elementos do espaço agindo sobre os elementos do espaço dual, e vice-versa, resultam em um escalar. Os elementos do espaço vetorial dual, se transformam de maneira inversa aos elementos do espaço vetorial original. A existência de um espaço dual a um dado espaço é condição necessária para que possamos definir um produto interno, onde dois elementos "vetoriais" resultam em um escalar. Na verdade, no produto escalar, um vetor de um dado espaço age sobre o vetor do seu espaço dual, como explicitado acima.

Por fim, é importante ressaltar que, aparentemente, os vetores do espaço dual, os que denominamos co-vetores, parecem não obedecer à regra do paralelograma para a adição de vetores, como indica a decomposição nos eixos coordenadas. Entretanto, é preciso ressaltar que a base de co-vetores unitários não é a base de vetores unitários usuais, nem eles se decompõem da mesma forma. De fato esta base deve ser escrita como  $\vec{\omega}_1, \vec{\omega}_2$ , onde a ação desses elementos sobre a base vetorial usual  $\vec{i}$  e  $\vec{j}$  resulta em um número. Para um co-vetor, portanto,

$$\vec{A} = A_1 \vec{\omega}_1 + A_2 \vec{\omega}_2 \quad . \quad (224)$$

## 19 Generalização da noção de vetores: tensores

Um tensor é a generalização das noções de escalar e vetor (ou co-vetor). Como foi visto anteriormente, um vetor agindo sobre um co-vetor, resulta em um número. O vetor está definido em um certo espaço, e o co-vetor no seu espaço dual. Como dito acima, é este fato que torna possível definir o produto interno, que sempre envolve um vetor e um elemento do seu espaço dual. Quanto ao tensor, nós podemos defini-lo como sendo uma aplicação multilinear de vetores e co-vetores na reta real. Um tensor se define também pela forma como suas componentes se transformam por uma mudança de sistema de coordenadas.

Simbolicamente podemos representar a ação de um tensor sobre vetores e co-vetores como

$$\mathbf{T}(\mathbf{V}_1, \dots, \mathbf{V}_n, \mathbf{C}_1, \dots, \mathbf{C}_m) = R \quad , \quad (225)$$

onde os  $\mathbf{V}_i$  indicam vetores e os  $\mathbf{C}_j$  indicam co-vetores, enquanto  $R$  é um número. Em termos de componentes, em um dado sistema de coordenadas, temos então,

$$T_{\mu_1 \dots \mu_n}{}^{\nu_1 \dots \nu_m} V^{\mu_1} \dots V^{\mu_n} C_{\nu_1} \dots C_{\nu_m} = R \quad . \quad (226)$$

No exemplo acima, temos um tensor  $\mathbf{T}$  de ordem  $n + m$ ,  $n$  vezes covariante e  $m$  contravariante. A ordem dos índices covariantes e contravariantes é importante, e no exemplo em questão, os  $n$  primeiros índices são covariantes, enquanto que os  $m$  últimos índices são contravariantes. No caso geral, a ordem do tensor, sua natureza contravariante, covariante ou mista, e a posição dos índices dependem do tipo específico de tensor com o qual lidamos. Exemplos específicos serão apresentados mais tarde.

Um escalar é um tensor de ordem zero: permanece invariante por transformações de coordenadas. Um vetor é um tensor contravariante de ordem 1; um co-vetor é um tensor covariante também de ordem 1. Veremos a seguir tensores mais gerais, de ordem superior, tais como o tensor métrico e o tensor eletromagnético, ambos de ordem 2. Na verdade, o conceito de tensor já é familiar em outras situações. Por exemplo, o estudo de um sólido deformável é feito com a ajuda do tensor de elasticidade. Neste caso, o conceito de uma aplicação multilinear, como definida acima, é bastante importante. Afinal, uma tensão aplicada em um direção  $x$ , pode provocar deformações tanto na direção  $x$ , quanto nas direções  $y$  e  $z$  perpendiculares. Tal relação não pode ser expressa simplesmente através de vetores, mas se torna factível através de um tensor, no caso, um tensor de segunda ordem. De fato, podemos definir a deformação sofrida por um sólido  $d^i$  sob a aplicação de uma tensão  $t^i$ , com a ajuda do tensor de elasticidade  $E^i_j$ :

$$d^i = E^i_j t^j \quad . \quad (227)$$

Assim, para o caso da aplicação da tensão na direção  $x$ , o que implica que a única componente não nula da tensão é  $t^x$ , temos para as deformações,

$$d^x = E^x_x d^x \quad , \quad d^y = E^y_x d^x \quad , \quad d^z = E^z_x d^x \quad . \quad (228)$$

Em geral, o tensor de segunda ordem pode ser representado por uma matriz  $n \times n$ , da mesma forma que um vetor pode ser representado por uma matriz coluna e um co-vetor por uma matriz linha. Veremos mais tarde, no entanto, que nem toda matriz  $n \times n$  é um tensor de segunda ordem: o que define efetivamente um tensor é a maneira como ele se transforma sob uma mudança de coordenadas.

No caso do tensor de elasticidade, se a matriz  $E^i_j$  é diagonal, o meio é linear. Se além do mais, os elementos da diagonal são iguais, então a relação tensorial se reduz a uma relação puramente escalar. De fato, se  $E^i_j = E\delta^i_j$  então

$$d^i = E\delta^i_j t^j \quad \rightarrow \quad d^i = E t^i \quad . \quad (229)$$

Neste caso, o meio é dito "isotrópico". A polarização de um meio dielétrico sob a ação de um campo magnético externo também é um tensor de segunda ordem. Para meios isotrópicos, o tensor de polarização se reduz a um escalar, pelos mesmos motivos que o tensor de elasticidade. Mas, o caso geral deve incluir anisotropias.

Para o caso do tensor  $\mathbf{T}$  definido acima, sob a ação de uma mudança de um sistema de coordenadas, suas componentes se transformam como

$$T_{\mu'_1 \dots \mu'_n}^{\nu'_1 \dots \nu'_m} = \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \dots \frac{\partial x^{\mu_n}}{\partial x^{\mu'_n}} \frac{\partial x^{\nu'_1}}{\partial x^{\nu_1}} \dots \frac{\partial x^{\nu'_m}}{\partial x^{\nu_m}} T_{\mu_1 \dots \mu_n}^{\nu_1 \dots \nu_m} \quad . \quad (230)$$

Lembrando que um vetor e um co-vetor se transformam respectivamente como

$$V^{\mu'_i} = \frac{\partial x^{\mu_i}}{\partial x^{\mu'_i}} V^{\mu_i} \quad , \quad C_{\nu'_i} = \frac{\partial x^{\nu_i}}{\partial x^{\nu'_i}} C_{\nu_i} \quad , \quad (231)$$

e que

$$\frac{\partial x^{\mu'_i}}{\partial x^{\mu_i}} \frac{\partial x^{\mu_i}}{\partial x^{\mu_j}} = \delta^{\mu'_i}_{\mu_j} \quad , \quad \frac{\partial x^{\nu'_i}}{\partial x^{\nu_i}} \frac{\partial x^{\nu_i}}{\partial x^{\nu'_j}} = \delta^{\nu'_i}_{\nu'_j} \quad , \quad (232)$$

onde a convenção da soma sobre índices repetidos e a regra da cadeia da derivação foram usadas. Assim,

$$\begin{aligned} T_{\mu'_1 \dots \mu'_n}^{\nu'_1 \dots \nu'_m} V^{\mu'_1} \dots V^{\mu'_n} C_{\nu'_1} \dots C_{\nu'_m} &= \\ T_{\mu_1 \dots \mu_n}^{\nu_1 \dots \nu_m} V^{\mu_1} \dots V^{\mu_n} C_{\nu_1} \dots C_{\nu_m} &= R \quad , \end{aligned} \quad (233)$$

de maneira que a aplicação tensorial multilinear explicitada acima, resulta efetivamente em um número, um escalar.

## 20 Operações com tensores

Os tensores podem ser puramente covariantes, contravariantes ou mistos. Por exemplo, podemos definir um tensor covariante de segunda ordem cujas componentes são  $A_{\mu\nu}$ , ou um tensor contravariante de segunda ordem com componentes  $B^{\mu\nu}$ , ou ainda um tensor misto de segunda ordem com componentes  $C^\mu_\nu$ . Tensores de ordem superior devem ser definidos da mesma forma. É importante ressaltar que a posição dos índices é crucial: por exemplo, em princípio as componentes de um tensor covariante de segunda ordem  $A_{\mu\nu}$  são diferentes das componentes  $A_{\nu\mu}$ . A natureza dos tensores (covariante, contravariante ou misto), assim como sua ordem (seu número de índices, falando de uma forma simplificada) depende do contexto em que eles aparecem, determinado pela maneira como eles se transformam sob uma mudança do sistema de coordenadas. O vetor deslocamento  $d\vec{r}$ , como já foi visto anteriormente, se transforma como um tensor contravariante de primeira ordem; o operador diferencial  $\vec{\nabla}$  se transforma como um tensor covariante de primeira ordem. Veremos posteriormente como tensores de ordem superior aparecem e como sua ordem e sua natureza são definidos.

A primeira operação com tensores é a sua multiplicação direta, que podemos também denominar de produto tensorial. Consideremos, para sermos específicos, dois tensores de ordens diferentes, que notemos abstratamente como **A** e **B** (não há nenhuma menção aos tensores apresentados acima com as mesmas letras). Digamos que **A** é um tensor misto de segunda ordem de componentes  $A^\mu_\nu$ , enquanto **B** é um tensor também misto, de terceira ordem, com componentes  $B_\lambda^\gamma{}_\rho$ . O produto tensorial entre estes dois tensores, resulta em um tensor misto de quinta ordem, que denominamos **C**:

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} \quad \rightarrow \quad C^\mu{}_\nu{}_\lambda{}^\gamma{}_\rho = A^\mu{}_\nu B_\lambda^\gamma{}_\rho \quad . \quad (234)$$

Podemos verificar se **C** é um tensor misto de quinta ordem verificando como ele se comporta por uma transformação de coordenadas. Uma vez que **A** e **B** foram definidos como tensores, ao passarmos de um sistema de coordenadas  $x^{\mu'}$  para um sistema de coordenadas  $x^\alpha$ , temos,

$$A^{\mu'}{}_{\nu'} = \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\nu'}} A^\alpha{}_\beta \quad , \quad (235)$$

$$B_{\lambda' \gamma' \rho'} = \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^{\gamma'}}{\partial x^\beta} \frac{\partial x^\sigma}{\partial x^{\rho'}} B_{\alpha \beta \sigma} . \quad (236)$$

Desta forma, temos que

$$\begin{aligned} C^{\mu' \nu' \lambda' \gamma' \rho'} &= A^{\mu' \nu'} B_{\lambda' \gamma' \rho'} \\ &= \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\nu'}} A^\alpha_\beta \frac{\partial x^\delta}{\partial x^{\lambda'}} \frac{\partial x^{\gamma'}}{\partial x^\omega} \frac{\partial x^\sigma}{\partial x^{\rho'}} B_{\delta \omega \sigma} \\ &= \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\nu'}} \frac{\partial x^\delta}{\partial x^{\lambda'}} \frac{\partial x^{\gamma'}}{\partial x^\omega} \frac{\partial x^\sigma}{\partial x^{\rho'}} A^\alpha_\beta B_{\delta \omega \sigma} \\ &= \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\nu'}} \frac{\partial x^\delta}{\partial x^{\lambda'}} \frac{\partial x^{\gamma'}}{\partial x^\omega} \frac{\partial x^\sigma}{\partial x^{\rho'}} C^\alpha_{\beta \delta \omega \sigma} . \end{aligned} \quad (237)$$

Logo,  $\mathbf{C}$  é um tensor misto de quinta ordem duplamente contravariante e triplamente covariante. Novamente, é preciso enfatizar que a posição dos índices é crucial. Observemos também que  $\mathbf{A} \otimes \mathbf{B}$  é, em princípio, diferente de  $\mathbf{B} \otimes \mathbf{A}$ , muito embora ambos resultem em tensores mistos de quinta ordem, duplamente contravariantes e triplamente covariantes. A posição dos índices das componentes é, no entanto, diferente.

Uma outra operação com tensores é a contração de índices no interior de um mesmo tensor. Consideremos o tensor  $\mathbf{C}$  do exemplo acima cujas componentes são escritas no sistema de coordenadas  $x^{\mu'}$ . Somemos nos índices  $\nu'$  e  $\gamma'$ , por exemplo. Temos então,

$$C^{\mu' \nu' \lambda' \gamma' \rho'} = C^{\mu' \lambda' \rho'} . \quad (238)$$

A contração dos índices  $\nu'$  e  $\gamma'$  resultou em um tensor misto de terceira ordem, uma vez contravariante e duas vezes covariante. A operação de contração de índices no seio de um tensor sempre reduz duas vezes sua ordem, suprimindo um índice contravariante e um covariante. A demonstração disto se faz usando a regra da cadeia para as derivadas:

$$\begin{aligned} C^{\mu' \nu' \lambda' \gamma' \rho'} &= \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\nu'}} \frac{\partial x^\delta}{\partial x^{\lambda'}} \frac{\partial x^{\nu'}}{\partial x^\omega} \frac{\partial x^\sigma}{\partial x^{\rho'}} C^\alpha_{\beta \delta \omega \sigma} \\ &= \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\delta}{\partial x^{\lambda'}} \frac{\partial x^\sigma}{\partial x^{\rho'}} \delta^\beta_\omega C^\alpha_{\beta \delta \sigma} \\ &= \frac{\partial x^{\mu'}}{\partial x^\alpha} \frac{\partial x^\delta}{\partial x^{\lambda'}} \frac{\partial x^\sigma}{\partial x^{\rho'}} C^\alpha_{\delta \sigma} . \end{aligned} \quad (239)$$

Logo,  $C^{\mu' \nu' \lambda' \gamma' \rho'}$  se transforma como um tensor de terceira ordem, uma vez



contravariante e duas vezes covariante, que podemos escrever como

$$C^{\mu'}{}_{\nu'}{}_{\lambda'}{}^{\nu'}{}_{\rho'} = C^{\mu'}{}_{\lambda'}{}_{\rho'}{}^{\nu'}{}_{\nu'} \quad . \quad (240)$$

A contração de índices só pode ser feita, evidentemente, entre índices covariantes e contravariantes.

Um tensor de segunda ordem pode ser decomposto em uma parte simétrica e uma parte anti-simétrica. De fato, consideremos um tensor covariante de segundo ordem  $A_{\mu\nu}$  como exemplo. Podemos escrever

$$\begin{aligned} A_{\mu\nu} &= A_{\mu\nu} + \frac{1}{2}A_{\nu\mu} - \frac{1}{2}A_{\nu\mu} \\ &= \frac{1}{2}(A_{\mu\nu} + A_{\nu\mu}) + \frac{1}{2}(A_{\mu\nu} - A_{\nu\mu}) \\ &= A_{(\mu\nu)} + A_{[\mu\nu]} \quad , \end{aligned} \quad (241)$$

onde definimos

$$A_{(\mu\nu)} = \frac{1}{2}(A_{\mu\nu} + A_{\nu\mu}) \quad , \quad (242)$$

$$A_{[\mu\nu]} = \frac{1}{2}(A_{\mu\nu} - A_{\nu\mu}) \quad . \quad (243)$$

Para um tensor puramente simétrico, temos  $A_{\mu\nu} = A_{\nu\mu}$ , enquanto para um tensor puramente anti-simétrico,  $A_{\mu\nu} = -A_{\nu\mu}$ .

Podemos também gerar um novo tensor, a partir de um tensor dado, derivando-o. Mas, para tanto, será preciso generalizar a noção de derivada covariante.

## 21 A noção de métrica

Os resultados descritos acima parecem indicar que só podemos fazer o produto entre dois vetores, obtendo uma função escalar, quando um deles é covariante e o outro é contravariante. Isto traria problemas para obter a norma de um dado vetor, definida como a raiz quadrada do produto interno de um vetor por ele próprio. Entretanto, podemos de fato definir o produto interno de um vetor covariante por ele próprio, assim como podemos definir o produto interno de um vetor contravariante por outro também contravariante. Isto é feito com ajuda da *métrica*. A métrica é uma quantidade tensorial que permite estabelecer como

se calcula a distância infinitesimal entre dois pontos em uma dada estrutura geométrica.

A distância infinitesimal entre dois pontos no espaço euclidiano a três dimensões pode ser escrita, utilizando coordenadas cartesianas, como

$$ds^2 = dx^2 + dy^2 + dz^2 = g_{\mu\nu} dx^\mu dx^\nu \quad , \quad (244)$$

onde  $\mu, \nu = 1, 2, 3$  e onde introduzimos o termo  $g_{\mu\nu}$ , que neste caso é tal que

$$\begin{aligned} g_{\mu\nu} = \delta_{\mu\nu} &= 1 \quad \text{se } \mu = \nu \quad , \\ &= 0 \quad \text{se } \mu \neq \nu \quad . \end{aligned}$$

O termo  $g_{\mu\nu}$  como apresentado acima é denominado de *métrica do espaço euclidiano a três dimensões*: é a quantidade que diz como calcular a distância entre dois pontos infinitesimalmente próximos no espaço euclidiano a três dimensões. No caso acima, temos a expressão para a métrica euclidiana quando se emprega o sistema de coordenadas cartesianas. É claro que este intervalo não deve mudar quando fazemos uma transformação de coordenadas, visto que ela é uma propriedade intrínseca da geometria em questão; no nosso exemplo acima, a geometria euclidiana. Assim, se fizermos uma transformação de coordenadas do tipo

$$x^\mu = x^\mu(x^{\lambda'}) \quad , \quad x^{\lambda'} = x^{\lambda'}(x^\mu) \quad , \quad (245)$$

o elemento de distância entre dois pontos se transforma da seguinte maneira:

$$ds^2 = g_{\mu\nu} \frac{\partial x^\mu}{\partial x^{\lambda'}} \frac{\partial x^\nu}{\partial x^{\gamma'}} dx^{\lambda'} dx^{\gamma'} = g_{\lambda'\gamma'} dx^{\lambda'} dx^{\gamma'} \quad , \quad (246)$$

onde

$$g_{\lambda'\gamma'} = g_{\mu\nu} \frac{\partial x^\mu}{\partial x^{\lambda'}} \frac{\partial x^\nu}{\partial x^{\gamma'}} \quad . \quad (247)$$

Isto nos mostra que a quantidade  $g_{\mu\nu}$  se transforma duplamente como uma quantidade covariante. Dizemos então que  $g_{\mu\nu}$  é um tensor covariante de segunda ordem.

Dada uma métrica  $g_{\mu\nu}$ , podemos obter a inversa, que será um tensor contravariante de segunda ordem. Esta inversa é obtida estabelecendo que o produto da métrica por sua inversa é a identidade:

$$g_{\mu\lambda} g^{\lambda\nu} = \delta_\mu^\nu \quad , \quad (248)$$

onde  $\delta_\mu^\nu$  é a função delta de Kronecker definida anteriormente.

Podemos exemplificar o que foi exposto anteriormente ainda no âmbito da geometria euclideana. Consideremos a transformação de um sistema de coordenadas cartesianas para um sistema de coordenadas esféricas, definida por

$$x = r \sin \theta \cos \phi \quad , \quad y = r \sin \theta \sin \phi \quad , \quad z = r \cos \theta \quad , \quad (249)$$

$$r = \sqrt{x^2 + y^2 + z^2} \quad , \quad \tan \theta = \frac{\sqrt{x^2 + y^2}}{z} \quad , \quad \tan \phi = \frac{y}{x} \quad . \quad (250)$$

Essa transformação e sua inversa são bem definidas, à exceção quando  $x = y = z = 0$ , ela é degenerada (esse mesmo ponto pode ser representado por uma infinidade de coordenadas  $r, \theta, \phi$ ). Neste caso, o elemento infinitesimal de distância se transforma como,

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \quad . \quad (251)$$

Neste caso revela-se conveniente representar a métrica como uma matrix:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad ; \quad g_{\lambda'\gamma'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad . \quad (252)$$

Consequentemente, as matrizes inversas se escrevem como,

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad ; \quad g^{\lambda'\gamma'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & r^{-2} \sin^{-2} \theta \end{pmatrix} \quad . \quad (253)$$

Consideremos agora o caso do espaço-tempo quadridimensional de Minkowski. Neste caso, o elemento invariante que permite calcular a distância entre dois eventos (pontos no espaço-tempo) é dado por

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu \quad . \quad (254)$$

Observe que neste caso a métrica foi escrita como  $g_{\mu\nu} = \eta_{\mu\nu}$ . Isto é feito para seguir a tradição segundo a qual a métrica no espaço-tempo de Minkowski se simboliza por  $\eta_{\mu\nu}$ . Em coordenadas esféricas  $(ct, r, \theta, \phi)$ , a métrica de Minkowski se torna,

$$ds^2 = c^2 dt^2 - dr^2 - r^2(d\theta^2 + \sin^2 \theta d\phi^2) = \eta_{\mu'\nu'} dx^{\mu'} dx^{\nu'} \quad . \quad (255)$$

A métrica  $\eta_{\mu\nu}$  e sua inversa  $\eta^{\mu\nu}$  se escrevem, em coordenadas cartesianas e esféricas, como

$$\eta_{\mu\nu} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ; \quad (256)$$

$$\eta_{\lambda'\gamma'} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix} . \quad (257)$$

$$\eta^{\mu\nu} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ; \quad (258)$$

$$\eta^{\lambda'\gamma'} = \begin{pmatrix} c^{-2} & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^{-2} & 0 \\ 0 & 0 & 0 & -r^{-2} \sin^{-2} \theta \end{pmatrix} . \quad (259)$$

A contração da métrica, na sua forma covariante, com um vetor contravariante, resulta em um vetor covariante. Podemos demonstrar isto facilmente, verificando como esta quantidade se transforma quando passamos de um sistema de coordenadas a outro. Consideremos então a quantidade  $B_\mu = g_{\mu\nu} B^\nu$ . Reescrevamos esta quantidade no sistema de coordenadas  $x^{\lambda'}$ :

$$\begin{aligned} B_\mu = g_{\mu\nu} B^\nu &= g_{\lambda'\gamma'} \frac{\partial x^{\lambda'}}{\partial x^\mu} \frac{\partial x^{\gamma'}}{\partial x^\nu} B^{\rho'} \frac{\partial x^\nu}{\partial x^{\rho'}} = \\ &= g_{\lambda'\gamma'} \frac{\partial x^{\lambda'}}{\partial x^\mu} \delta_{\rho'}^{\gamma'} B^{\rho'} = \frac{\partial x^{\lambda'}}{\partial x^\mu} B_{\gamma'} \end{aligned} \quad (260)$$

o que mostra que o vetor  $B_\mu$  construído a partir do vetor contravariante  $B^\nu$ , com ajuda da métrica, se comporta efetivamente como um vetor covariante. Da mesma forma, utilizando a métrica inversa  $g^{\lambda\gamma}$  podemos obter um vetor contravariante a partir de um vetor covariante.

Desta forma, o produto interno de dois vetores covariantes (ou contravariantes) pode ser feito com ajuda da métrica. De fato, dado um vetor contravariante  $A^\mu$ , podemos construir sua versão covariante utilizando a métrica:

$$A^\mu \Rightarrow A_\mu = g_{\mu\nu} A^\nu \quad . \quad (261)$$

Assim, a norma do vetor pode ser calculada:

$$A^\mu A_\mu = g_{\mu\nu} A^\mu A^\nu \quad . \quad (262)$$

No caso da métrica de Minkowski, em coordenadas cartesianas, temos

$$A^\nu A_\nu = \eta_{\mu\nu} A^\mu A^\nu = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 \quad . \quad (263)$$

Observe que no espaço-tempo de Minkowski, a norma de um vetor contravariante não é definida positiva.

Da mesma forma, usando a métrica inversa  $\eta^{\mu\nu}$ , podemos associar um vetor contravariante a um vetor covariante:

$$B_\mu \Rightarrow B^\mu = \eta^{\mu\nu} B_\nu \quad . \quad (264)$$

A norma de um vetor covariante se escreve então como

$$B^\mu B_\mu = \eta^{\mu\nu} B_\mu B_\nu \quad . \quad (265)$$

No espaço-tempo de Minkowski, utilizando coordenadas cartesianas, a norma se lê

$$B^\mu B_\mu = \eta_{\mu\nu} B^\mu B^\nu = (B_0)^2 - (B_1)^2 - (B_2)^2 - (B_3)^2 \quad . \quad (266)$$

Como no caso anterior, ela não é definida positiva.

Note-se que se estamos, por outro lado realizando a contração de um vetor contravariante  $A^\mu$  com um covariante  $B_\nu$ , obtemos, ainda no espaço-tempo de Minkowski,

$$A^\mu B_\mu = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3 \quad . \quad (267)$$

Por outro lado, para contrair um vetor, por exemplo, contravariante  $A^\mu$  com outro vetor contravariante  $C^\nu$ , teremos necessidade de introduzir a métrica convertendo uma das quantidades em um vetor covariante (qual delas, não é relevante

para este propósito):

$$A^\mu C_\nu = A^\mu \eta_{\mu\nu} C^\nu = \eta_{\mu\nu} A^\mu C^\nu = A_\mu C^\mu = A^0 C^0 - A^1 C^1 - A^2 C^2 - A^3 C^3 \quad . \quad (268)$$

A partir das considerações acima, pode-se dar conta porque a distinção entre vetores covariantes e contravariantes não é muito enfatizada quando se desenvolve uma teoria no espaço euclidiano usual. Utilizando-se o sistema de coordenadas cartesianas os valores numéricos das componentes de um vetor contravariante  $A^\mu$  são iguais aos valores numéricos das componentes covariantes correspondentes  $A_\mu = g_{\mu\nu} A^\nu$  devido à forma da métrica e sua inversa neste sistema de coordenadas (252,253). Entretanto, esta distinção já se torna relevante, mesmo no caso euclidiano, quando se emprega outros sistemas de coordenadas, mesmo que ortogonais, como o esférico, cilíndrico, etc.

As estruturas geométricas não se esgotam com o espaço euclidiano e o espaço-tempo de Minkowski. Existem infinitas geometrias possíveis, com estruturas próprias, classificadas pela forma como se calcula a distância entre dois pontos infinitesimalmente próximos. Estas geometrias, que apresentam em geral uma curvatura intrínseca, se denominam *geometrias riemaniannas*. No caso onde elas apresentam uma coordenada temporal, como no caso minkowskiano, elas são denominadas normalmente como *pseudo-riemaniannas*. Muitas vezes, utiliza-se a expressão *geometria riemanianna* para designar em geral estruturas espaciais ou espaço-temporais dotadas de uma métrica.

## 22 Métrica sobre a esfera bi-dimensional

Podemos exemplificar o significado da métrica estudando o caso da esfera. O objetivo é definir a métrica sobre a esfera, isto é, como se calcula a distância infinitesimal entre dois pontos sobre esta superfície. A esfera está imersa no espaço euclidiano a três dimensões. Em coordenadas cartesianas, a distância entre dois pontos no espaço tri-dimensional é dada pela expressão (244). Se re-expressarmos este elemento de distância em coordenadas esféricas, temos então a expressão (251). Como queremos nos restringir à superfície da esfera, o raio é

constante e igual a  $R$ , resultando na métrica bi-dimensional,

$$ds^2 = R^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad . \quad (269)$$

Em que a métrica (269) difere da métrica de um plano cartesiano? Em princípio poder-se-ia realizar uma transformação para o sistema de coordenadas cartesiano bi-dimensional. Mas, a situação é um pouco mais sutil. O que se tem a fazer, é projetar os pontos da superfície da esfera sobre o plano. Isto implica em realizar uma projeção estereográfica, mapear a esfera no plano. Definamos, para este fim, as coordenadas no plano,

$$x = r \cos \alpha \quad , \quad y = r \sin \alpha \quad , \quad (270)$$

$$r = \sqrt{x^2 + y^2} \quad , \quad \alpha = \arctan \frac{y}{x} \quad . \quad (271)$$

Temos que  $\beta = \frac{\pi}{2} - \frac{\theta}{2}$ . Isto resulta em

$$d \cos \left( \frac{\pi}{2} - \frac{\theta}{2} \right) = d \sin \frac{\theta}{2} = 2R \quad , \quad (272)$$

$$d \sin \left( \frac{\pi}{2} - \frac{\theta}{2} \right) = d \cos \frac{\theta}{2} = r \quad . \quad (273)$$

Podemos agora relacionar as coordenadas  $(\theta, \phi)$  sobre a esfera com as coordenadas  $(r, \alpha)$  sobre o plano:

$$r = 2R \cot \frac{\theta}{2} \quad , \quad \alpha = \phi \quad . \quad (274)$$

A partir destas transformações, obtém-se,

$$d\theta = -\frac{4R dr}{4R^2 + r^2} \quad , \quad \sin \theta = \frac{4Rr}{4R^2 + r^2} \quad , \quad d\phi = d\alpha \quad . \quad (275)$$

Assim, a métrica da esfera (269) adquire a forma,

$$ds^2 = \frac{16R^4}{(4R^2 + r^2)^2} (dr^2 + r^2 d\alpha^2) \quad . \quad (276)$$

Ou ainda, utilizando as coordenadas cartesianas  $(x, y)$  sobre o plano,

$$ds^2 = \frac{1}{\left(1 + \frac{x^2 + y^2}{4R^2}\right)^2} (dx^2 + dy^2) \quad . \quad (277)$$

A expressão da métrica sobre a esfera expressa em coordenadas cartesianas do plano (277) mostra que a geometria da esfera é bastante diferente da geometria do plano, como deveríamos esperar. A diferença consiste fundamentalmente

no fator comum em (277) que multiplica a métrica do plano euclidiano. Por ser um fator comum, ele é denominado de *fator conforme*. As componentes da métrica da esfera neste sistema de coordenadas se escrevem então,

$$g_{xx} = g_{yy} = \frac{1}{\left(1 + \frac{x^2 + y^2}{4R^2}\right)^2}, \quad g_{xy} = g_{yx} = 0. \quad (278)$$

De posse da métrica de uma superfície esférica, podemos calcular, por exemplo, a distância entre dois pontos sobre a superfície, a área de uma região delimitada por duas latitudes (duas curvas com  $\theta$  constante), etc. O espaço definido pelos elementos métricos (278) é um espaço a curvatura constante positiva, com raio de curvatura  $k = \frac{1}{R}$ . Este é um exemplo de uma geometria não euclidiana, com o espaço dotado de uma curvatura intrínseca.

## 23 A derivada covariante

A derivada ordinária de um tensor não é, em geral, um tensor. Isto pode ser visto de maneira simples. Considere-se, para ser específico, um vetor contravariante  $A^{\mu'}$ , cujas componentes são escritas em um sistema de coordenadas  $x^{\lambda'}$ , e considere-se a derivada ordinária destas quantidades utilizando este mesmo sistema de coordenadas. Agora, realizemos uma transformação para um novo sistema de coordenadas  $x^\gamma$ , tal que  $x^{\lambda'} = x^{\lambda'}(x^\gamma)$ . Tem-se então

$$\begin{aligned} \frac{\partial A^{\mu'}}{\partial x^{\nu'}} &= \frac{\partial x^\rho}{\partial x^{\nu'}} \frac{\partial}{\partial x^\rho} \left( \frac{\partial x^{\mu'}}{\partial x^\sigma} A^\sigma \right) \\ &= \frac{\partial x^\rho}{\partial x^{\nu'}} \frac{\partial x^{\mu'}}{\partial x^\sigma} \frac{\partial A^\sigma}{\partial x^\rho} + \frac{\partial x^\rho}{\partial x^{\nu'}} \frac{\partial^2 x^{\mu'}}{\partial x^\rho \partial x^\sigma} A^\sigma. \end{aligned} \quad (279)$$

A quantidade inicial, a derivada ordinária do tensor contravariante de primeira ordem, não se transformou como um tensor devido a presença do segundo termo em (279). O mesmo ocorre para qualquer outro tensor, de qualquer natureza e qualquer ordem, excetuando o caso do tensor de ordem zero (um escalar) cuja a derivada ordinária resulta em um tensor covariante de primeira ordem. A razão é aue os vetores da base mudam de direção de um ponto a outro, e isto tem que ser levado em conta ao se comparar dois vetores em pontos distintos do espaço.

Entretanto, é fundamental que possamos construir quantidades que derivadas se comportam como tensores. A razão é simples: as leis físicas, assim



como muitas quantidades matemáticas, são expressas como derivadas de certas grandezas; se elas não se transformarem tensorialmente sob uma mudança de sistema de coordenadas, isto implicará que as leis físicas não guardarão sua forma quando da passagem de um sistema de coordenadas a outro: a invariância das leis físicas quando da passagem de um sistema inercial a outro não estaria assegurada. Para tanto, será preciso generalizar a noção de derivada, introduzindo a *derivada covariante*.

Existem muitas formas de se definir uma derivada covariante, todas obviamente equivalentes. Introduziremos a noção de derivada covariante impondo que a derivada covariante das componentes da métrica seja nula. Isto na verdade não é obrigatório e existem estruturas onde a derivada covariante da métrica não é nula. Na verdade, ao impormos que a derivada covariante da métrica é nula, estamos supondo que a métrica guarda sua forma sobre toda a variedade, quer dizer, a geometria é puramente definida pela métrica. Isto define os espaços riemaniannos. Existem outras estruturas, como as dos espaços dotados de torção, os espaços de Weyl, onde isto não é verdade. Mas, visto que estamos interessados em espaços riemaniannos, nos deteremos neste tipo de estrutura.

Denotemos a derivada covariante como

$$D_\mu \equiv \frac{D}{Dx^\mu} \quad . \quad (280)$$

A ação da derivada covariante sob a métrica resulta em um valor nulo. A derivação covariante é a derivação ordinária mais a ação de uma conexão. Uma conexão leva em conta o fato que ao derivarmos, estamos subtraindo a quantidade em dois pontos diferentes da variedade e, neste caso, devemos levar em conta não unicamente a variação intrínseca da quantidade, mas também o fato que os eixos do sistema de coordenadas mudam quando passamos de um ponto a outro, notadamente se a variedade é curva (muito embora esta condição não seja necessária, já que a noção de derivada covariante é necessária mesmo em um espaço plano quando se, por exemplo, coordenadas curvilíneas). Definimos então a derivada covariante sob um dado tensor, digamos  $A^\mu{}_\nu$ , como

$$D_\rho A^\mu{}_\nu = \partial_\rho A^\mu{}_\nu + \Gamma^\mu_{\rho\sigma} A^\sigma{}_\nu - \Gamma^\sigma_{\rho\nu} A^\mu{}_\sigma \quad . \quad (281)$$

Note-se que a ação da conexão sobre índices contravariantes leva a um sinal

positivo, equanto sua ação sobre índices covariantes leva a um sinal negativo.

Qual é a forma da conexão? Para tanto, impomos que a derivada covariante da métrica seja nula:

$$D_\mu g_{\nu\lambda} = \partial_\mu g_{\nu\lambda} - \Gamma_{\mu\nu}^\rho g_{\rho\lambda} - \Gamma_{\mu\lambda}^\rho g_{\nu\rho} = 0 \quad . \quad (282)$$

Para descobrirmos a forma da conexão re-escrevamos a expressão acima realizando uma permutação cíclica dos índices:

$$D_\mu g_{\nu\lambda} = \partial_\mu g_{\nu\lambda} - \Gamma_{\mu\nu}^\rho g_{\rho\lambda} - \Gamma_{\mu\lambda}^\rho g_{\nu\rho} = 0 \quad ; \quad (283)$$

$$D_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - \Gamma_{\lambda\mu}^\rho g_{\rho\nu} - \Gamma_{\lambda\nu}^\rho g_{\mu\rho} = 0 \quad ; \quad (284)$$

$$D_\nu g_{\lambda\mu} = \partial_\nu g_{\lambda\mu} - \Gamma_{\nu\lambda}^\rho g_{\rho\mu} - \Gamma_{\nu\mu}^\rho g_{\lambda\rho} = 0 \quad . \quad (285)$$

Ao subtrair (284,285) de (283), e considerando que a conexão é simétrica nos índices inferiores, obtém-se

$$\partial_\mu g_{\nu\lambda} - \partial_\lambda g_{\mu\nu} - \partial_\nu g_{\lambda\mu} = 2\Gamma_{\lambda\nu}^\rho g_{\mu\rho} \quad . \quad (286)$$

Multiplicando pela métrica inversa  $g^{\mu\sigma}$ , utilizando o fato que  $g^{\mu\sigma} g_{\mu\rho} = \delta_\rho^\sigma$ , obtemos finalmente para a conexão,

$$\Gamma_{\mu\nu}^\rho = \frac{1}{2} g^{\rho\sigma} \left( \partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right) \quad . \quad (287)$$

Sob uma transformação de coordenadas,

$$g_{\mu'\nu'} = \frac{\partial x^\rho}{\partial x^{\mu'}} \frac{\partial x^\sigma}{\partial x^{\nu'}} g_{\rho\sigma} \quad , \quad g^{\mu'\nu'} = \frac{\partial x^{\mu'}}{\partial x^\rho} \frac{\partial x^{\nu'}}{\partial x^\sigma} g^{\rho\sigma} \quad . \quad (288)$$

Logo, a conexão se transforma como,

$$\begin{aligned} \Gamma_{\nu'\lambda'}^{\mu'} &= \frac{1}{2} g^{\mu'\gamma'} \left( \partial_{\nu'} g_{\gamma'\lambda'} + \partial_{\lambda'} g_{\gamma'\nu'} - \partial_{\gamma'} g_{\nu'\lambda'} \right) \\ &= \frac{1}{2} \frac{\partial x^{\mu'}}{\partial x^\rho} \frac{\partial x^{\gamma'}}{\partial x^\sigma} g^{\rho\sigma} \left[ \frac{\partial x^\alpha}{\partial x^{\nu'}} \partial_\alpha \left( \frac{\partial x^\beta}{\partial x^{\gamma'}} \frac{\partial \omega}{\partial x^{\lambda'}} g_{\beta\omega} \right) + \frac{\partial x^\omega}{\partial x^{\lambda'}} \partial_\omega \left( \frac{\partial x^\beta}{\partial x^{\gamma'}} \frac{\partial \alpha}{\partial x^{\nu'}} g_{\beta\alpha} \right) \right. \\ &\quad \left. - \frac{\partial x^\beta}{\partial x^{\gamma'}} \partial_\beta \left( \frac{\partial x^\alpha}{\partial x^{\nu'}} \frac{\partial \omega}{\partial x^{\lambda'}} g_{\alpha\omega} \right) \right] \\ &= \frac{\partial x^{\mu'}}{\partial x^\rho} \frac{\partial x^\alpha}{\partial x^{\nu'}} \frac{\partial x^\omega}{\partial x^{\lambda'}} \Gamma_{\alpha\omega}^\rho + \frac{\partial^2 x^\rho}{\partial x^{\nu'} \partial x^{\lambda'}} \frac{\partial x^{\mu'}}{\partial x^\rho} \quad . \end{aligned} \quad (289)$$

A presença do último termo em (289) mostra que a conexão, conhecida neste caso como símbolo de Christoffel, não é um tensor.

Podemos analisar agora o comportamento da derivada covariante sob uma mudança de sistema de coordenadas. Temos,

$$\begin{aligned}
D_{\mu'} A^{\nu'} &= \partial_{\mu'} A^{\nu'} + \Gamma_{\mu' \rho'}^{\nu'} A^{\rho'} \\
&= \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\sigma} \frac{\partial A^\sigma}{\partial x^\alpha} + \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\sigma} A^\sigma \\
&+ \frac{\partial x^{\nu'}}{\partial x^\beta} \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^\omega}{\partial x^{\rho'}} \Gamma_{\alpha\omega}^\beta \frac{\partial x^{\rho'}}{\partial x^\kappa} A^\kappa \\
&+ \frac{\partial^2 x^\alpha}{\partial x^{\mu'} \partial x^{\rho'}} \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial x^{\rho'}}{\partial x^\sigma} A^\sigma \\
&= \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\sigma} \frac{\partial A^\sigma}{\partial x^\alpha} + \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\sigma} A^\sigma \\
&+ \frac{\partial x^{\nu'}}{\partial x^\beta} \frac{\partial x^\alpha}{\partial x^{\mu'}} \Gamma_{\alpha\kappa}^\beta A^\kappa \\
&+ \left[ \frac{\partial}{\partial x^\sigma} \left( \frac{\partial x^\alpha}{\partial x^{\mu'}} \right) \right] \frac{\partial x^{\nu'}}{\partial x^\alpha} A^\sigma .
\end{aligned} \tag{290}$$

O último termo pode ser re-arranjado conforme se segue:

$$\begin{aligned}
\left[ \frac{\partial}{\partial x^\sigma} \left( \frac{\partial x^\alpha}{\partial x^{\mu'}} \right) \right] \frac{\partial x^{\nu'}}{\partial x^\alpha} &= \frac{\partial}{\partial x^\sigma} \left[ \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\alpha} \right] - \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\sigma} \\
&= \frac{\partial}{\partial x^\sigma} \left[ \frac{\partial x^{\nu'}}{\partial x^{\mu'}} \right] - \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\sigma} \\
&= \frac{\partial}{\partial x^\sigma} \left[ \delta_{\mu'}^{\nu'} \right] - \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\sigma} \\
&= - \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\sigma} .
\end{aligned} \tag{291}$$

Assim, o terceiro termo de (290) é anulado pelo quarto termo. Temos então

$$D_{\mu'} A^{\nu'} = \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\sigma} \left( \frac{\partial A^\sigma}{\partial x^\alpha} + \Gamma_{\alpha\kappa}^\sigma A^\kappa \right) = \frac{\partial x^\alpha}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\sigma} D_\alpha A^\sigma . \tag{292}$$

Logo, a derivada covariante se transforma como um tensor. O mesmo pode ser demonstrado para a derivada covariante de um vetor covariante.

## 24 Os operadores Laplaceano e D'Alambertiano

A equação de propagação de uma onda para uma dada quantidade  $\phi$  se escreve como

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial^2 \phi}{\partial y^2} - \frac{\partial^2 \phi}{\partial z^2} = 0 . \tag{293}$$

Ou ainda como

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = 0 \quad , \quad (294)$$

onde

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad , \quad (295)$$

é o operador Laplaceano em coordenadas cartesianas.

A equação da onda pode ser escrita de uma forma mais econômica como

$$\square \phi = 0 \quad , \quad (296)$$

onde

$$\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \quad , \quad (297)$$

é o operador D'Alambertiano em coordenadas cartesianas. As expressões (295,297) representam, na verdade, os operadores Laplaceano e D'Alambertiano no espaço euclidiano e minkowskiano, respectivamente, em coordenadas cartesianas. Em particular, esses operadores podem ser escritos como

$$\nabla^2 = \delta^{ij} \partial_i \partial_j \quad , \quad \square = \eta^{\mu\nu} \partial_\mu \partial_\nu \quad , \quad (298)$$

onde  $i, j = 1, 2, 3$  e  $\mu, \nu = 0, 1, 2, 3$ .

Em uma geometria qualquer definida por uma métrica  $g_{\mu\nu}$ , o operador D'Alambertiano se escreve como

$$\square = g^{\mu\nu} D_\mu D_\nu = g^{\mu\nu} \left( \partial_\mu \partial_\nu - \Gamma_{\mu\nu}^\rho \partial_\rho \right) \quad . \quad (299)$$

Nesta expressão, foram utilizadas as derivadas covariantes, de maneira que este operador seja covariante. Podemos pensar que o operador Laplaceano é o caso particular do operador D'Alambertiano quando a métrica é a do espaço euclidiano em três dimensões. Esta expressão covariante permite, em particular, calcular o operador Laplaceano em qualquer sistema de coordenadas de uma maneira muito simples.

De fato, considere-se como exemplo o operador Laplaceano no espaço euclidiano em coordenadas esféricas. Primeiro, devemos obter a métrica neste sistema de coordenadas:

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad . \quad (300)$$

Assim, os coeficientes métricos não nulos são:

$$g_{rr} = 1 \quad , \quad g_{\theta\theta} = r^2 \quad , \quad g_{\phi\phi} = r^2 \sin^2 \theta \quad , \quad (301)$$

$$g^{rr} = 1 \quad , \quad g^{\theta\theta} = r^{-2} \quad , \quad g^{\phi\phi} = r^{-2} \sin^{-2} \theta \quad . \quad (302)$$

Os símbolos de Christoffel não nulos são:

$$\Gamma_{\theta\theta}^r = -r \quad , \quad \Gamma_{\phi\phi}^r = -r \sin^2 \theta \quad , \quad (303)$$

$$\Gamma_{\phi\phi}^\theta = -\cos \theta \sin \theta \quad , \quad \Gamma_{\theta\phi}^\phi = \cot \theta \quad (304)$$

Assim, temos para o operador Laplaceano, lembrando que estamos neste caso no espaço euclideano a três dimensões:

$$\begin{aligned} \nabla^2 &= g^{\mu\nu} \left( \partial_\mu \partial_\nu - \Gamma_{\mu\nu}^\rho \partial_\rho \right) \\ &= \partial_r^2 + \frac{1}{r^2} \partial_\theta^2 + \frac{1}{r^2 \sin^2 \theta} \partial_\phi^2 + \frac{2}{r^2} \partial_r + \frac{1}{r^2} \cot \theta \partial_\theta \\ &= \frac{1}{r^2} \partial_r \left( r^2 \partial_r \right) + \frac{1}{r^2} \left[ \frac{1}{\sin \theta} \partial_\theta \left( \sin \theta \partial_\theta \right) + \frac{1}{\sin^2 \theta} \partial_\phi^2 \right] \quad . \quad (305) \end{aligned}$$

Obtivemos assim a expressão conhecida para o operador Laplaceano em coordenadas esféricas simplesmente através da noção de derivada covariante. Da mesma forma, podemos determinar a forma do operador Laplaceano (ou D'Alambertiano) em qualquer outro sistema de coordenadas, partindo apenas da distância infinitesimal entre dois pontos naquele sistema de coordenadas.

## 25 A cinemática relativista

A cinemática busca descrever o movimento de uma partícula através de sua posição, velocidade e aceleração. Na mecânica newtoniana tais quantidades têm três componentes, já que a mecânica newtoniana é formulada em um espaço euclideano tri-dimensional, ao qual se acrescenta o tempo como parâmetro universal. Entretanto, a relatividade restrita indica que vivemos em um espaço-tempo quadri-dimensional. Portanto, as quantidades correspondentes devem ter quatro componentes, uma delas associada às coordenadas temporais. Será preciso, portanto, generalizar as expressões usuais newtonianas para as quantidades cinemáticas.

No que diz respeito à posição, esta generalização é direta, e nós já a abordamos anteriormente. O quadri-vetor (vetor com quatro componentes) posição é, em coordenadas espaciais cartesianas,

$$x^\mu = (ct, x, y, z) \quad . \quad (306)$$

A velocidade é a variação da posição com o tempo. Em mecânica newtoniana esta definição não traz grandes dificuldades, uma vez que a estrutura geométrica newtoniana é o espaço euclidiano tri-dimensional ao qual se acrescenta o tempo como parâmetro universal. A velocidade, portanto, é a derivada do vetor posição em relação a este tempo universal, independente do observador. Em relatividade restrita, por outro lado cada observador mede um tempo particular ao seu referencial.

É necessário para definirmos a quadri-velocidade um parâmetro que seja, como no caso newtoniano, independente do observador. Este parâmetro é fornecido pela distância infinitesimal entre dois pontos no espaço tempo,

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad . \quad (307)$$

Como já foi visto anteriormente, este elemento de linha tem o mesmo valor para todos observadores inerciais. Por outro lado, o tempo próprio  $\tau$  é o tempo medido por um observador em repouso em relação á partícula. Para este observador, temos então

$$ds^2 = c^2 d\tau^2 \quad . \quad (308)$$

Consequentemente, o tempo próprio também é um invariante, pois ele difere de  $ds^2$  pelo fator constante  $c$ . Isto significa que, se dois observadores localizados em referenciais diferentes, descrevem o movimento de uma dada partícula, eles discordarão sobre a medida de tempo respectivas, mas concordarão sobre qual é a medida de tempo do observador em repouso em relação a partícula, quer dizer, o tempo próprio da partícula.

Tomaremos então o tempo próprio como sendo o parâmetro universal, uma vez que ele é um invariante. Assim, a quadri-velocidade se define como

$$u^\mu = \frac{dx^\mu}{d\tau} \quad . \quad (309)$$

A quadri-velocidade assim definida é um vetor ou, dito de outra forma, um tensor contravariante de primeira ordem. Podemos facilmente relacionar a quadri-velocidade, que é um tensor no espaço-tempo a quatro dimensões, com as grandezas usuais tri-dimensionais. De fato consideremos novamente a distância infinitesimal  $ds$ . Temos então,

$$\begin{aligned} ds^2 &= c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad , \\ d\tau^2 &= dt^2 - \frac{dx^2 + dy^2 + dz^2}{c^2} \quad , \\ d\tau^2 &= dt^2 \left( 1 - \frac{v^2}{c^2} \right) \quad , \\ d\tau &= dt \sqrt{1 - \frac{v^2}{c^2}} = \gamma(v) dt \quad , \quad \gamma(v) = \sqrt{1 - \frac{v^2}{c^2}} \quad . \end{aligned} \quad (310)$$

Desta forma temos

$$\frac{d\tau}{dt} = \gamma(v) \quad . \quad (311)$$

Assim

$$\begin{aligned} u^\mu &= \frac{dt}{d\tau} \frac{dx^\mu}{dt} = \gamma(v) \left( c, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) \\ &= \gamma(v) (c, v_x, v_y, v_z) = \gamma(v) (c, \vec{v}) \quad . \end{aligned} \quad (312)$$

Esta é a expressão para a quadri-velocidade expressa em termos da velocidade tri-dimensional  $\vec{v}$  medida em um referencial inercial específico. Observe que, neste caso, escrevemos o fator de Lorentz como  $\gamma(v)$  e não como  $\Gamma(V)$  pois queremos distinguir os fatores de Lorentz que correspondem à passagem de um referencial inercial arbitrário para o referencial próprio da partícula ( $\gamma(v)$ ) do caso onde se processa a passagem de um referencial inercial arbitrário a outro igualmente arbitrário ( $\Gamma(V)$ ). Um dos motivos para isto é que o referencial próprio pode ser apenas localmente inercial, como no caso onde a partícula está sendo acelerada: é preciso neste caso definir uma sequência de referenciais inerciais, cada um correspondendo ao referencial próprio da partícula apenas em um intervalo infinitesimal de tempo.

O quadri-vetor velocidade  $u^\mu$  é do tipo tempo. Há duas formas de mostrar isto. Uma delas seria calcular explicitamente a contração  $u^\mu u_\mu = \eta_{\mu\nu} u^\mu u^\nu$  utilizando (312) e a expressão da métrica minkowskiana  $\eta_{\mu\nu}$ . A outra faz uso da definição (309) e da noção de elemento de linha  $ds^2$  e do tempo próprio

$d\tau$ . Vamos considerar esta segunda possibilidade, que é bem mais econômica. Utilizando então a definição (309), temos

$$\begin{aligned} u^\mu u_\mu &= \eta_{\mu\nu} u^\mu u^\nu = \eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \\ &= \eta_{\mu\nu} \frac{dx^\mu dx^\nu}{d\tau^2} = \frac{ds^2}{d\tau^2} \quad , \end{aligned} \quad (313)$$

onde utilizou-se o fato que  $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ . Lembrando agora que  $ds^2 = c^2 d\tau^2$ , temos finalmente

$$u^\mu u_\mu = c^2 \quad . \quad (314)$$

Logo, a norma do quadri-vetor velocidade é igual ao quadrado da velocidade da luz, sendo portanto positiva definida. Assim, o quadri-vetor velocidade é do tipo tempo. Isto está ligado à noção de causalidade: uma partícula só pode seguir uma trajetória no espaço-tempo de Minkowski tal que a relação causal esteja assegurada.

A generalização quadri-dimensional da aceleração segue os mesmos passos. A aceleração é a derivada da velocidade. Novamente, para que esta afirmação faça sentido no espaço-tempo quadri-dimensional, a derivada deve ser em relação a um parâmetro universal, no caso o tempo próprio medido pela partícula da qual se calcula a aceleração. Temos então,

$$a^\mu = \frac{du^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2} \quad . \quad (315)$$

Podemos também expressar a quadri-aceleração em termos das grandezas tri-dimensionais usuais, como a posição, a velocidade e a aceleração medidas em um referencial específico. Utilizando novamente a relação entre o tempo próprio da partícula  $\tau$  e o tempo coordenado medido em um referencial inercial  $t$  (311), tem-se

$$\begin{aligned} a^\mu &= \gamma(v) \frac{du^\mu}{dt} = \gamma(v) \frac{d}{dt} (\gamma(v)c, \gamma(v)\vec{v}) \\ &= \gamma(v) (\dot{\gamma}(v)c, \dot{\gamma}(v)\vec{v} + \gamma(v)\dot{\vec{v}}) \quad . \end{aligned} \quad (316)$$

Temos, por outro lado,

$$\dot{\gamma}(v) = \frac{1}{c^2} \frac{\vec{v} \cdot \dot{\vec{v}}}{\left(1 - \frac{v^2}{c^2}\right)^{3/2}} = \gamma^3(v) \frac{\vec{v} \cdot \dot{\vec{v}}}{c^2} \quad . \quad (317)$$



Assim, a quadri-aceleração expressa em termos das quantidades tri-dimensionais, medidas em um referencial particular, assume a forma

$$a^\mu = \left( \gamma^4(v) \frac{\vec{v} \cdot \dot{\vec{v}}}{c}, \gamma^4 \vec{v} \cdot \dot{\vec{v}} \frac{\vec{v}}{c^2} + \gamma^2 \vec{a} \right) \quad , \quad (318)$$

sendo  $\vec{a}$  a aceleração tri-dimensional da partícula medida em um referencial  $S$  específico.

O quadri-vetor aceleração  $a^\mu$  é do tipo espaço. Novamente, pode-se demonstrar isto de duas formas: utilizando a definição da quadri-aceleração (??) em termos das grandezas tri-dimensionais, ou mostrando que a quadri-aceleração é ortogonal a um vetor tipo tempo, como a quadri-velocidade. Seguiremos este último procedimento por ser mais econômico e por nos revelar uma relação entre os quadri-vetores velocidade e aceleração que será muito útil posteriormente.

Considere-se a norma da quadri-velocidade,

$$u^\mu u_\mu = c^2 \quad . \quad (319)$$

Diferenciando em relação ao tempo próprio e observando que o lado direito de (319) é constante, obtém-se

$$u_\mu \frac{du^\mu}{d\tau} = 0 \quad \Rightarrow \quad \eta_{\mu\nu} u^\mu \frac{du^\nu}{d\tau} = \eta_{\mu\nu} u^\mu a^\nu = u_\mu a^\mu = 0 \quad . \quad (320)$$

Uma vez que os dois quadri-vetores são ortogonais entre si e que o quadri-vetor velocidade é do tipo tempo, o quadri-vetor aceleração é forçosamente do tipo espaço.

Ao exprimimos os quadri-vetores velocidade e aceleração de uma forma manifestamente tensorial temos que, dadas suas componentes em um sistema de referência, podemos calculá-las em outro sistema de referência. De fato, as componentes destas duas quantidades, expressas em um referencial  $S$  e um referencial  $S'$ , se relacionam como

$$u^{\mu'} = \Lambda^{\mu'}_\nu u^\nu \quad , \quad (321)$$

$$a^{\mu'} = \Lambda^{\mu'}_\nu a^\nu \quad . \quad (322)$$

Utilizando estas relações, podemos expressar, por exemplo, as quantidades tri-dimensionais (velocidade, aceleração, etc.) medidas em um referencial em termos das mesmas quantidades em outro referencial.

## 26 O momento e a lei de força relativista

Na mecânica newtoniana, a segunda lei indica qual é o efeito de uma força  $\vec{F}$  sobre uma partícula:  $\vec{F} = m\vec{a}$ . Como vimos anteriormente, utilizando a definição de  $\vec{a}$  como segunda derivada da posição em relação ao tempo, obtemos que esta expressão é invariante pelas transformações de Galileu. Se focalizarmos a atenção nas transformações de Lorentz, temos que substituir todas as grandezas tridimensionais por grandezas quadridimensionais definidas no espaço de Minkowski. Essas grandezas devem ser tensores, se transformando como tais quando de uma mudança de um referencial a outro. Construir relações tensoriais é a única forma de garantir que teremos expressões que manterão sua forma por transformações de Lorentz; além disto, devido à natureza das transformações de Lorentz, que são definidas em um espaço-tempo a quatro dimensões, tais expressões são forçosamente quadri-dimensionais. Desta forma, a generalização mínima (podemos sempre pensar em algumas generalizações mais complicadas) da segunda lei de tal forma a incorporar estes princípios relativistas é:

$$F^\mu = m_0 a^\mu = m_0 \frac{du^\mu}{d\tau} = m_0 \frac{d^2 x^\mu}{d\tau^2} \quad , \quad (323)$$

onde  $a^\mu$ ,  $u^\mu$ ,  $x^\mu$  e  $\tau$  são a quadri-aceleração, a quadri-velocidade, a quadri-posição e o tempo próprio, respectivamente. O termo  $m_0$  é a massa de repouso da partícula; o seu significado ficará mais claro posteriormente.

A quadri-aceleração é um tensor de primeira ordem. Como  $m_0$  é um escalar,  $F^\mu$  é também um tensor de primeira ordem. (Lembramos que o tempo próprio, sendo proporcional ao elemento  $ds$  ( $ds^2 = c^2 d\tau^2$ ), mantém-se invariante por transformações de Lorentz). Este fato nos permite relacionar as componentes de  $F^\mu$  com as expressões usuais newtonianas de força. E mais: nos permitirá interpretar fisicamente o que significa a componente temporal desta quantidade  $F^0$  que, aparentemente, e apenas aparentemente, não tem equivalente newtoniano.

De fato, a quadri-força  $F^\mu$ , sendo um tensor, visto que é construído a partir de quantidades tensoriais, se comporta, por uma transformação do sistema de coordenadas  $x^\mu$  para o sistema de coordenadas  $x^{\mu'}$ , como

$$F^{\mu'} = \Lambda^{\mu'}_\nu F^\nu \quad , \quad (324)$$

onde  $\Lambda_{\nu}^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^{\nu}}$  é a matrix de transformação do sistema de coordenadas  $x^{\nu}$  ao sistema de coordenadas  $x^{\mu'}$ . No caso específico das transformações de Lorentz, isto conduz a:

$$F^{0'} = \Gamma \left( F^0 - \frac{V}{c} F^1 \right) ; \quad (325)$$

$$F^{1'} = \Gamma \left( F^1 - \frac{V}{c} F^0 \right) ; \quad (326)$$

$$F^{2'} = F^2 ; \quad (327)$$

$$F^{3'} = F^3 . \quad (328)$$

Nós conhecemos as leis de força macroscópicas e que obedecem às relações newtonianas com grande precisão. Como relacionar essas leis de força ordinárias com a expressão aparentemente mais abstrata dada por (470)? Para vermos isto, retomemos a expressão da quadri-aceleração  $a^{\mu}$ :

$$a^{\mu} = \frac{du^{\mu}}{d\tau} = \frac{du^{\mu}}{dt} \frac{dt}{d\tau} . \quad (329)$$

Mas,

$$\frac{dt}{d\tau} = \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} . \quad (330)$$

Desta forma,

$$F^{\mu} = m_0 \gamma (f^0, \vec{f}) , \quad (331)$$

onde  $\vec{f} = \frac{d}{dt} \gamma \vec{v}$  é a componente espacial da quadri-força medida em um dado referencial  $S$ . Desta forma, a relação da quadri-força com a lei de força ordinária se torna

$$F^i = \gamma f^i , \quad (332)$$

sendo  $f^i$  a expressão tridimensional de uma força medida em um referencial específico: para o caso de uma força restauradora,  $f^i = -kx^i$ ; para o caso da força gravitacional,  $f^i = -G \frac{Mm}{r^3} x^i$ .

Necessitamos entretanto, interpretar  $F^0$ . Lembremos que o quadri-vetor aceleração é do tipo espaço, ao passo que o quadri-vetor velocidade é do tipo tempo. Logo,  $a^{\mu} v_{\mu} = 0$ . Assim,

$$F^{\mu} u_{\mu} = 0 \rightarrow \gamma^2 \left( f^0 c - f^i v_i \right) = 0 \rightarrow f^0 = \frac{1}{c} f^i \cdot v_i . \quad (333)$$

Desta forma, a componente temporal da quadri-força está ligada à potência transmitida pela força  $\vec{f}$ .

O resultado precedente permite reescrever as leis de transformação das componentes da força relativista:

$$f'_1 = \frac{f_1 - \frac{V\vec{f}\cdot\vec{v}}{c^2}}{1 - \frac{v_1 V}{c^2}} , \quad (334)$$

$$f'_2 = \frac{f_2}{\Gamma(1 - \frac{v_1 V}{c^2})} , \quad (335)$$

$$f'_3 = \frac{f_3}{\Gamma(1 - \frac{v_1 V}{c^2})} . \quad (336)$$

## 27 O momento linear

Na mecânica newtoniana, o momento se escreve como  $\vec{p} = m\vec{v}$ . A generalização quadridimensional é imediata. Nós escrevemos o momento em quatro-dimensões como

$$p^\mu = m_0 u^\mu = m_0 \gamma(c, \vec{v}) . \quad (337)$$

Novamente, aparece  $m_0$ , a massa de repouso da partícula;  $v^\mu$  é, como antes, a quadri-velocidade. Como  $v^\mu$  é um tensor de primeira ordem e  $m_0$  é um escalar,  $p^\mu$  será também um tensor de primeira ordem e se transforma como tal quando passamos de um sistema de coordenadas  $x^\mu$  para um sistema de coordenadas  $x^{\mu'}$ :

$$p^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\nu} p^\nu . \quad (338)$$

No caso específico da transformação de Lorentz, isto conduz às seguintes expressões conectando as componentes do quadri-momento no novo sistema de coordenadas em relações no antigo sistema de coordenadas:

$$p^{0'} = \Gamma \left( p^0 - \frac{V}{c} p^1 \right) , \quad (339)$$

$$p^{1'} = \Gamma \left( p^1 - \frac{V}{c} p^0 \right) , \quad (340)$$

$$p^{2'} = p^2 , \quad p^{3'} = p^3 . \quad (341)$$

Como o quadri-momento é diretamente proporcional à quadri-velocidade, ele também é do tipo tempo.

Como no caso da definição da quadri-força, devemos interpretar a componente temporal do quadri-momento. Para fazermos isto, consideremos novamente a quadri-força e escrevemo-la como derivada do quadri-momento em relação ao tempo próprio:

$$F^\mu = \frac{dp^\mu}{d\tau} \quad . \quad (342)$$

Esta relação pode ainda ser reescrita como,

$$F^\mu = \frac{dp^\mu}{dt} \frac{dt}{d\tau} \quad . \quad (343)$$

Mas,  $\frac{dt}{d\tau} = \gamma$ , de maneira que temos,

$$F^\mu = \gamma \frac{dp^\mu}{dt} \quad . \quad (344)$$

Recordando as expressões da quadri-força em relação às grandezas tridimensionais usuais, nós obtemos,

$$f^0 = \frac{dp^0}{dt} \quad , \quad f^i = \frac{dp^i}{dt} \quad . \quad (345)$$

Como  $f^0 = \frac{1}{c} \frac{dE}{dt}$ , então  $p^0 = \frac{E}{c}$ : a componente temporal do quadri-momento está relacionada com a energia da partícula. Além disto, observamos que

$$p^0 = m_0 \gamma c = \frac{E}{c} \rightarrow E = \gamma m_0 c^2 \quad . \quad (346)$$

Notemos que quando a partícula está em repouso  $\gamma = 1$  e ainda assim ela possui uma energia dada por

$$E_0 = m_0 c^2 \quad , \quad (347)$$

denominada energia de repouso da partícula. Essa energia de repouso não tem análogo newtoniano

Por outro lado, nós temos,

$$f^i = \frac{dp^i}{dt} = m_0 \frac{d}{dt} \gamma v^i \quad . \quad (348)$$

Esta expressão nos permite escrever a relação entre força e momento relativista projetada no espaço tridimensional:

$$\vec{f} = m_0 \left( \dot{\gamma} \vec{v} + \gamma \vec{a} \right) \quad . \quad (349)$$

O resultado obtido anteriormente para o quadri-momento linear poderia, de certa forma, ser antecipado por razões puramente conceituais. Já foi dito anteriormente que as leis de conservação estão relacionadas à invariâncias da teoria em relação a certos tipos de transformações. O fato de um sistema ser invariante por uma translação espacial (pouco importa fazer uma experiência em Vitória ou em Cachoeiro, desde que os sistemas sejam idênticos) conduz à conservação do momento; o sistema físico é também invariante por transformações temporais, o resultado de experiências feitas sobre sistemas idênticos sendo o mesmo independentemente se a experiência foi feita hoje ou ontem: esta simetria conduz à conservação da energia. Como agora estamos trabalhando em um contínuo espaço-temporal, as conservações em separado do momento (invariância espacial) e da energia (invariância temporal) se unificam em uma única invariância espaço-temporal, que é expressa pela conservação, para um sistema fechado, do quadri-momento:

$$p_f^\mu = p_i^\mu \quad , \quad (350)$$

onde os índices  $i$  e  $f$  designam momentos inicial e final, respectivamente. A equação (350) expressa simultaneamente as leis de conservação da energia e do momento linear tridimensional, que se apresentam como leis separadas na mecânica newtoniana. Essa relação de conservação será amplamente usada ulteriormente.

Da expressão do momento (337), e lembrando que  $v^\mu v_\mu = c^2$ , nós obtemos,

$$p^\mu p_\mu = m_0^2 c^2 = \gamma^2 \left( \frac{E^2}{c^2} - \vec{p}^2 \right) \quad . \quad (351)$$

Definindo,

$$\bar{E} = \gamma E \quad , \quad \bar{p} = \gamma p \quad , \quad (352)$$

obtemos a expressão,

$$\bar{E} = \sqrt{m_0^2 c^4 + \bar{p}^2 c^2} \quad , \quad (353)$$

que relaciona a energia da partícula com sua massa de repouso e momento linear.

## 28 Problemas envolvendo leis de força

Vamos analisar agora em detalhe alguns problemas envolvendo o cálculo das trajetórias de uma partícula para alguns casos particulares.

### 28.1 Força constante

Vamos inicialmente considerar o exemplo mais simples: uma partícula inicialmente em repouso que sofre a ação de uma força constante, orientada segundo o eixo dos  $x$ , de tal forma que, no referencial do laboratório, ela assume a forma  $\vec{f} = f_x \vec{i} = f \vec{i}$ ,  $f$  sendo uma constante. Neste caso, a quadri-força se escreve,

$$F = \gamma \left( \frac{\vec{f} \cdot \vec{v}}{c}, f, 0, 0 \right) . \quad (354)$$

Considerando que o momento é dado por

$$P = \gamma m_0 (c, \vec{v}) , \quad (355)$$

e que  $F = \frac{dP}{d\tau} = \gamma \frac{dP}{dt}$ , obtemos,

$$\frac{fv}{c} = m_0 c \frac{d\gamma}{dt} , \quad (356)$$

$$f = m_0 \left( \frac{d\gamma}{dt} v + \gamma \frac{dv}{dt} \right) . \quad (357)$$

Uma vez que  $f$  é constante, temos

$$\gamma v = \frac{f}{m_0} t = \omega_0 t , \quad (358)$$

onde fizemos  $\omega_0 = \frac{f}{m_0}$ . A solução para a velocidade é,

$$v = \frac{\omega_0 t}{\sqrt{1 + \frac{\omega_0^2 t^2}{c^2}}} , \quad (359)$$

ao passo que a solução para a distância percorrida é,

$$x = \frac{c^2}{\omega_0} \left( \sqrt{1 + \frac{\omega_0^2 t^2}{c^2}} - 1 \right) . \quad (360)$$

No limite em que a velocidade da luz tende a infinito, obtemos as relações newtonianas usuais,

$$v \sim \omega_0 t , \quad x = \frac{\omega_0 t^2}{2} . \quad (361)$$

Por outro lado, podemos observar que a velocidade jamais ultrapassará o valor limite  $c$ :

$$t \rightarrow \infty \Rightarrow v \rightarrow c \quad . \quad (362)$$

Finalmente, é interessante estabelecer uma relação entre o tempo medido no laboratório e o tempo próprio medido por um relógio fixo na partícula. Temos,

$$\tau = \int_0^t \sqrt{1 - \frac{v^2}{c^2}} dt = \frac{c}{\omega_0} \sinh^{-1} \frac{\omega_0 t}{c} \quad . \quad (363)$$

No limite  $t \rightarrow \infty$ , temos a relação

$$\tau \sim \frac{c}{\omega_0} \ln \frac{2\omega_0 t}{c} \quad , \quad (364)$$

o que mostra que o tempo próprio cresce muito mais lentamente que o tempo do laboratório.

## 28.2 Força constante ortogonal à velocidade inicial

Consideremos agora outro problema clássico da mecânica newtoniana: uma partícula com velocidade inicial  $v_0$  orientada ao longo do eixo  $x$  que entra em uma região onde é submetida a uma força  $\vec{F}$  orientada ao longo do eixo  $y$ . Na mecânica newtoniana, trata-se de um problema trivial, visto que existirá aceleração apenas ao longo do eixo  $y$ , a velocidade na direção  $x$  permanecendo constante. Em relatividade restrita, entretanto, sabemos que existirá também aceleração ao longo do eixo  $x$  como efeito da velocidade não nula nesta direção.

Vamos escrever as relações que nos interessam. Temos,

$$F^\mu = m_0 a^\mu \rightarrow \frac{\vec{f}}{m_0} = \gamma^3 \frac{\vec{v}\vec{a}}{c^2} + \gamma \vec{a} \rightarrow \frac{\vec{f}}{m_0} = \frac{d}{dt} \gamma \vec{v} \quad . \quad (365)$$

Obtemos então,

$$\gamma \vec{v} = \frac{\vec{f}}{m_0} t + \vec{v}_0 \quad . \quad (366)$$

Isto nos conduz a duas relações:

$$\gamma v_x = \gamma_0 v_0 \quad , \quad \gamma v_y = \frac{f}{m_0} t \quad . \quad (367)$$

Elevando ao quadrado as duas expressões e adicionando-as, obtemos

$$\gamma^2 v^2 = \gamma_0^2 v_0^2 + \omega_0^2 t^2 \quad , \quad \omega_0^2 = \frac{f^2}{m_0^2} \quad . \quad (368)$$



Esta última equação, nos permite obter  $\gamma$  em termos de  $t$ :

$$\gamma = \sqrt{1 + \frac{\gamma_0 v_0^2 + \omega_0^2 t^2}{c^2}} . \quad (369)$$

Inserindo (369) nas das equações apresentadas em (367), obtemos

$$v_x = \frac{\gamma_0 v_0}{\sqrt{1 + \frac{\gamma_0 v_0^2 + \omega_0^2 t^2}{c^2}}} , \quad (370)$$

$$v_y = \frac{\omega_0 t}{\sqrt{1 + \frac{\gamma_0 v_0^2 + \omega_0^2 t^2}{c^2}}} . \quad (371)$$

Observemos que no limite  $t \rightarrow \infty$ , temos  $v_x \rightarrow 0$  e  $v_y \rightarrow c$ , o que quer dizer que a partícula se desloca com velocidade praticamente igual à da luz na direção  $y$ , em relação ao que a velocidade na direção  $x$  se torna desprezível.

Estas relações podem ainda serem integradas, fornecendo  $x$  e  $y$  em função de  $t$ :

$$x(t) = \frac{v_0 c}{\omega_0} \sinh^{-1} \left( \frac{\omega_0 t}{\sqrt{c^2 + v_0^2}} \right) , \quad (372)$$

$$y(t) = \frac{c^2}{\omega_0} \left( \sqrt{1 + \frac{\gamma_0 v_0^2 + \omega_0^2 t^2}{c^2}} - \sqrt{1 + \frac{\gamma_0 v_0^2}{c^2}} \right) , \quad (373)$$

o que nos fornece a relação implícita,

$$x = \frac{v_0 c}{\omega_0} \sinh^{-1} \left( y^2 - 1 \right) . \quad (374)$$

Esta relação pode ser comparada com sua equivalente newtoniana,

$$x = \quad (375)$$

Note-se que no limite em que  $\frac{v}{c} \rightarrow 0$ , as duas relações coincidem.

### 28.3 Partícula carregada em movimento sob a ação de um campo magnético uniforme

Suponhamos uma partícula de massa de repouso  $m_0$  e carga  $q$  que se move com velocidade  $\vec{v}$  perpendicular a um campo magnético estático  $\vec{B}$ . Suponhamos então que

$$\vec{v} = v_x \vec{i} + v_y \vec{j} , \quad \vec{B} = B \vec{k} . \quad (376)$$

A lei de força se escreve

$$\vec{f} = q\vec{v} \times \vec{B} = q(-v_x\vec{j} + v_y\vec{i})B \quad . \quad (377)$$

Usando as relações relativistas deduzidas anteriormente, temos

$$m_0 \frac{d}{dt} \gamma(v) \vec{v} = q(-v_x\vec{j} + v_y\vec{i})B \quad .. \quad (378)$$

Por outro lado, devido à ortogonalidade entre  $\vec{v}$  e  $\vec{B}$  em qualquer instante de tempo, temos que

$$\vec{f} \cdot \vec{v} = \frac{dE}{dt} = \frac{d}{dt} m_0 \gamma(v) c^2 = 0 \quad . \quad (379)$$

Logo,  $\gamma(v)$  = constante, implicando que o módulo da velocidade permanece constante.

Desta forma, podemos então escrever

$$\dot{v}_x = \omega^2 v_y \quad , \quad \dot{v}_y = -\omega^2 v_x \quad , \quad (380)$$

onde  $\omega^2 = \frac{qB}{m_0 \gamma(v)}$ . Estas equações podem ser reduzidas a uma única equação:

$$\ddot{v}_i + \omega^2 v_i = 0 \quad , \quad i = x, y \quad . \quad (381)$$

A solução então se escreve como,

$$v_x = v_{x0} \cos(\omega t) \quad , \quad v_y = v_{y0} \cos(\omega t) \quad , \quad (382)$$

onde a fase foi escolhida de forma a satisfazer as condições iniciais. Integrando novamente para obter  $x$  e  $y$ , temos

$$x(t) = \frac{v_{x0}}{\omega} \sin(\omega t) \quad , \quad y(t) = \frac{v_{y0}}{\omega} \sin(\omega t) \quad (383)$$

onde fixamos que a partícula no tempo  $t = 0$  se encontrava na origem. Desta forma, a partícula descreve um círculo de raio

$$r = \sqrt{x^2 + y^2} = \frac{v_0}{\omega} \quad , \quad v_0 = \sqrt{v_{x0}^2 + v_{y0}^2} \quad . \quad (384)$$

A expressão newtoniana correspondente a este caso é idêntica à obtida anteriormente, com a única diferença que  $\omega_n = \frac{qB}{m_0}$ . O parâmetro  $\omega$  é conhecido como *frequência de ciclotron*. Assim, a frequência de ciclotron relativista é menor que a newtoniana, devido à presença do fator  $\gamma(v)$  no denominador, e o raio do círculo descrito pela partícula é maior.

## 28.4 O movimento da órbita dos planetas. Precessão das órbitas elípticas

Inicialmente vamos estabelecer o contexto do problema através da análise newtoniana. Neste caso, temos que

$$\vec{F} = m_0 \vec{a} \quad \rightarrow \quad \vec{a} = -G \frac{M}{r^2} \hat{r} \quad , \quad (385)$$

já que  $\vec{F} = -G \frac{M m_0}{r^2} \hat{r}$ . A aceleração escreve-se, em coordenadas esféricas,

$$\vec{a} = (\ddot{r} - r\dot{\theta}^2) \hat{r} + (r\ddot{\theta} + 2\dot{r}\dot{\theta}) \hat{\theta} \quad . \quad (386)$$

Inserindo esta relação em (385), obtém-se duas equações:

$$\ddot{r} - r\dot{\theta}^2 = -G \frac{M}{r^2} \quad , \quad (387)$$

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 \quad . \quad (388)$$

A última equação pode ser integrada, fornecendo

$$r^2 \dot{\theta} = l = \text{constante} \quad (389)$$

o que implica que o momento angular é conservado, fato bem conhecido para o caso da força central. Observe-se que a quantidade  $l$  é, na verdade, o momento angular por unidade de massa.

Reinserindo este último resultado na primeira equação, obtém-se

$$\ddot{r} - \frac{l^2}{r^3} = -G \frac{M}{r^2} \quad . \quad (390)$$

Agora, torna-se interessante realizar uma mudança de variável, de  $t$  para  $\theta$ . Desta forma, escrevemos

$$\dot{r} = r' \frac{l}{r^2} \quad , \quad (391)$$

$$\ddot{r} = \frac{d}{dt} \left( r' \frac{l}{r^2} \right) = \frac{l^2}{r^2} \left( \frac{r''}{r^2} - 2 \frac{r'^2}{r^3} \right) \quad . \quad (392)$$

Assim, a equação (390) assume a forma:

$$\frac{r''}{r^2} - 2 \frac{r'^2}{r^3} - \frac{1}{r} = -G \frac{M}{l^2} \quad . \quad (393)$$

Ou ainda:

$$\left(\frac{1}{r}\right)'' + \frac{1}{r} = G\frac{M}{l^2} \quad . \quad (394)$$

Definindo  $u = \frac{1}{r}$ , esta equação admite uma forma simples:

$$u'' + u = G\frac{M}{l^2} \quad . \quad (395)$$

Tem-se agora uma equação diferencial de segunda ordem não homogênea, do tipo oscilador harmônico. Neste caso, temos que procurar a solução geral da equação homogênea correspondente e uma solução particular da não homogênea. O resultado final escreve-se

$$u = A \cos \theta + G\frac{M}{l^2} \quad , \quad (396)$$

que, em termos de  $r$ , resulta em

$$r = \frac{1}{A \cos \theta + G\frac{M}{l^2}} \quad . \quad (397)$$

A solução (397) corresponde a uma órbita parabólica, hiperbólica ou elíptica, dependendo se  $A$  é igual, menor, ou maior que  $G\frac{M}{l^2}$ . Utilizando a expressão da energia  $E = \frac{1}{2}m_0v^2 + G\frac{Mm_0}{r}$ , que é uma constante para a solução acima (o que era de se esperar pois a força central é conservativa), esses casos correspondem a energia zero, positiva e negativa, respectivamente.

Para o caso de energia negativa, ( $A > G\frac{M}{l^2}$ ) temos uma órbita elíptica: a  $\theta = 0$  o raio da órbita é mínimo (periélio), crescendo até atingir o valor máximo em  $\theta = \pi$  (afélio), e decrescendo novamente até o valor mínimo em  $\theta = 2\phi$ . Logo, o periélio ocorre sempre no mesmo lugar, ao fim de uma volta completa. Entretanto, não é isto que se observa, e os planetas descrevem órbitas elípticas cujo periélio acusa um pequeno avanço a cada volta completa. O caso mais notório é o de Mercúrio, cujo periélio avança a uma taxa de  $42''$  de arco por século. Uma anomalia pequena, mas detectável, em relação à solução newtoniana apresentada anteriormente.

A anomalia observada nas órbitas elípticas dos planetas, devido ao avanço do periélio, pode ser em princípio explicada através da perturbação que a órbita de um dado planeta sofre devido à presença dos demais planetas. Afinal, o problema tratado acima é um típico problema de dois corpos, mas na verdade a

órbita de um planeta no sistema solar é um problema de  $N$  corpos<sup>12</sup>, já que um planeta descreve sua trajetória sob a ação não apenas do sol, mas também dos demais oito planetas<sup>13</sup>. Isto é feito através de técnicas perturbativas. Entretanto, mesmo tal análise, mais completa e detalhada, consegue explicar apenas parcialmente o avanço do periélio dos planetas. Para o caso mais extremo, que é o de Mercúrio, o avanço do periélio é de aproximadamente  $550''$  de arco por século; mas somente cerca de  $510''$  de arco por século podem ser explicados como devidos aos efeitos dos outros planetas; algo como  $40''$  de arco por século resistem a qualquer explicação utilizando a mecânica newtoniana<sup>14</sup>. É um efeito pequeno, mas perfeitamente detectável e que está além da margem de erro das técnicas perturbativas empregadas.

Consideremos agora o caso relativista. Temos as seguintes equações:

$$\vec{f} \cdot \vec{v} = m_0 \dot{\gamma}(v) c^2, \quad (398)$$

$$\vec{f} = m_0 (\dot{\gamma} \vec{v} + \gamma \vec{a}). \quad (399)$$

Como antes, vamos supor,

$$\vec{f} = -G \frac{M m_0}{r^2} \hat{r}. \quad (400)$$

Rigorosamente falando, estamos introduzindo um elemento estranho à teoria da relatividade: a lei de força gravitacional newtoniana não obedece aos postulados relativistas, já que ela supõe uma interação instantânea. No entanto, vamos ignorar este fato pelo momento, considerando que o problema que estamos tratando é uma primeira aproximação ao problema real. A solução mais completa de problemas envolvendo forças gravitacionais implica considerar a teoria da Relatividade Geral, que se encontra além da análise à qual nos propomos aqui.

---

<sup>12</sup>O problema de dois corpos é perfeitamente resolúvel em mecânica newtoniana. Quando um dado problema envolve mais que dois corpos, e em princípio não admite solução exata, ele é denominado *problema de N corpos*

<sup>13</sup>Para sermos mais precisos deveríamos levar também em conta a presença dos satélites no caso de alguns planetas, dos asteróides, etc.

<sup>14</sup>Foi aventada a possibilidade que este efeito se deveria ao achatamento do sol: se o sol for achatado, ele terá um momento quadropolar que gera uma precessão do periélio. Mas os parâmetros de achatamento do sol conhecidos não são capazes de gerar um efeito da magnitude observada

Retornaremos à esta questão no capítulo dedicado à teoria da Relatividade Geral.

## 29 Conservação do momento

Na mecânica newtoniana, as leis de conservação da energia e do momento linear são independentes. A conservação da energia se verifica quando não existem forças dissipativas agindo no sistema; as forças envolvidas são conservativas, o que significa que elas podem ser deduzidas de um potencial. A conservação do momento, por outro lado, está ligada à terceira lei de Newton, como foi visto no primeiro capítulo, e não é verificada caso haja forças externas ao sistema. Mas, existe uma forma mais elegante de derivar estas leis de conservação: elas estão associadas às simetrias do sistema. A lei de conservação da energia corresponde à invariância da estrutura do sistema com relação às translações temporais, enquanto a lei de conservação do momento está associada à invariância com relação às translações espaciais. A invariância temporal é quebrada quando existem forças dissipativas, visto que a evolução do sistema se torna irreversível, enquanto a invariância espacial não se verifica mais quando há forças externas, visto que diferentes posições espaciais não mais se equivalem devido à presença de um agente externo ao sistema.

A teoria da Relatividade Restrita unifica, de uma certa forma, tempo e espaço em uma única entidade, que é o espaço-tempo. Em consequência, as duas leis de conservação citadas acima, se reduzem a uma única lei de conservação. Dito de outra forma, uma vez que energia e momento linear se expressam em termos, agora, de um único tensor, o quadrivetor momento, existe agora uma única lei de conservação associada às translações espaço-temporais. Consideremos, por exemplo, um conjunto de partículas, cada uma dotada de seu quadrivetor momento. Podemos somar estes quadrivetores, para obter o quadrivetor total do sistema:

$$P\left(\frac{E}{c}, \vec{p}\right) = \sum_{k=1}^N P_k\left(\frac{E_k}{c}, \vec{p}_k\right). \quad (401)$$

A conservação do quadrimomento significa que a soma total dos quadrimomentos individuais deve ser a mesma em qualquer momento da evolução do sistema,

na ausência de forças dissipativas e de agentes externos ao sistema:

$$P^i\left(\frac{E_i}{c}, \vec{p}^i\right) = P^f\left(\frac{E_f}{c}, \vec{p}^f\right), \quad (402)$$

onde os superescritos  $i$  e  $f$  indicam "inicial" e "final". Uma vez que, em uma igualdade de quadrivetores, cada componente deve se igualar separadamente, temos,

$$E^i = E^f, \quad \vec{p}^i = \vec{p}^f, \quad (403)$$

onde

$$E^i = \sum_{k=1}^N E_k^i, \quad E^f = \sum_{k=1}^N E_k^f, \quad (404)$$

$$\vec{p}^i = \sum_{k=1}^N \vec{p}_k^i, \quad \vec{p}^f = \sum_{k=1}^N \vec{p}_k^f, \quad (405)$$

sendo  $E_k^i$  ( $E_k^f$ ) a energia inicial (final) da  $k$ -ésima partícula. A mesma notação se aplica para os momentos tri-dimensionais.

A lei de conservação do quadrivetor momento é particularmente importante nos problemas de colisão elásticos. Vamos ver a seguir alguns exemplos.

## 29.1 Choque entre duas partículas

Duas partículas, a primeira de massa  $m$  e a segunda de massa  $M$  colidem elasticamente. Suponhamos a situação em que a primeira partícula se move com velocidade  $\vec{V} = V\vec{i}$ , enquanto a segunda partícula se encontra em repouso na origem. Tratemos primeiramente o problema do ponto de vista Newtoniano. Ainda, por simplicidade, consideremos o caso puramente uni-dimensional. Neste caso, temos as leis de conservação da energia e do momento:

$$\frac{1}{2}mV^2 = \frac{1}{2}mv_1^2 + \frac{1}{2}Mv_2^2, \quad (406)$$

$$mV = mv_1 + Mv_2. \quad (407)$$

Uma manipulação algébrica simples, conduz aos seguintes resultados para  $v_1$  e  $v_2$ :

$$v_1 = \frac{m - M}{m + M}V, \quad (408)$$

$$v_2 = \frac{2m}{m + M}V. \quad (409)$$

Os casos limites são bem determinados: se  $m \gg M$ , a partícula 1 continua seu movimento como se a partícula 2 não existisse, e a partícula 2 adquire uma velocidade que é duas vezes a da primeira partícula; se  $M \gg m$ , a segunda partícula permanece parada depois da colisão enquanto a primeira retorna com a mesma velocidade inicial em módulo; se as duas partículas têm a mesma massa, depois da colisão a primeira partícula fica em repouso na origem, enquanto a segunda partícula passa a ter a velocidade que antes era a da primeira partícula.

Ao analisar, do ponto de vista relativista, o choque entre duas partículas, vamos primeiro considerar o caso em que as massas são iguais. A primeira partícula tem velocidade  $\vec{V} = V\vec{1}$  enquanto a segunda partícula está em repouso na origem. O problema torna-se mais simples se analisado no referencial do centro de massa. Logo, vamos analisar a colisão no referencial  $S'$ , que se desloca em relação a  $S$  com velocidade  $U$  orientada ao longo do eixo positivo dos  $x$ . Neste referencial, as velocidades das partículas são,

$$v'_1 = \frac{v_1 - U}{1 - \frac{v_1 U}{c^2}} = \frac{V - U}{1 - \frac{VU}{c^2}}, \quad (410)$$

$$v'_2 = \frac{v_2 - U}{1 - \frac{v_2 U}{c^2}} = -U, \quad (411)$$

onde fixamos  $v_1 = V$  e  $v_2 = 0$ . Impondo  $v'_2 = -v'_1$ , de tal maneira que  $S'$  corresponda ao referencial do centro de massa, obtemos

$$U = \frac{c^2}{V} \left\{ 1 - \sqrt{1 - \frac{V^2}{c^2}} \right\}. \quad (412)$$

Observe que, no limite de baixas velocidades, obtemos  $U = \frac{V}{2}$ , em acordo com o resultado Newtoniano. Aliás o sinal em frente à raiz quadrada de (412) foi escolhido de maneira a reproduzir este limite.

Podemos generalizar esta análise para o caso em que o choque não seja uni-dimensional, mas se passe no plano  $(x, y)$ . Suponhamos que a partícula 1 saia com velocidade  $\vec{v}_1$  fazendo um ângulo  $\theta$  com o semi-eixo positivo dos  $x$  e a partícula 2 saia com velocidade  $\vec{v}_2$  fazendo um ângulo  $\phi$  com o semi-eixo negativo dos  $x$ . Neste caso, as conservações da energia e do momento tridimensional fornecem as seguintes equações:

$$2m\gamma(v)c^2 = m\gamma(v_1)c^2 + m\gamma(v_2)c^2, \quad (413)$$



$$0 = m\gamma(v_1)v_1 \cos \theta - m\gamma(v_2)v_2 \cos \phi, \quad (414)$$

$$0 = m\gamma(v_1)v_1 \sin \theta - m\gamma(v_2)v_2 \sin \phi. \quad (415)$$

As duas últimas equações podem serem combinadas, resultando em

$$\gamma_1^2 v_1^2 = \gamma_2^2 v_2^2, \quad (416)$$

onde, por simplicidade, chamamos  $\gamma(v_i) = \gamma_i$ . Usando a relação,

$$\gamma_i^2 \frac{v_i^2}{c^2} = \gamma_i^2 \left(1 - \frac{1}{\gamma_i^2}\right), \quad (417)$$

encontramos  $\gamma_1 = \gamma_2$ . Usando a primeira equação, tem-se que  $\gamma_1 = \gamma_2 = \gamma(v)$ , o que implica  $v_1 = v_2 = v$ . Logo,  $\theta = \phi$ .

Podemos agora achar os ângulos de espalhamento no referencial original  $S$ . Para isto consideremos a transformação de velocidades inversa:

$$\bar{v}_{1x} = \frac{v_{1x} + U}{1 + \frac{v_{1x}U}{c^2}}, \quad (418)$$

$$\bar{v}_{1y} = \frac{v_{1y}}{\Gamma(U) \left[1 + \frac{v_{1x}U}{c^2}\right]}, \quad (419)$$

$$\bar{v}_{2x} = \frac{v_{2x} + U}{1 + \frac{v_{2x}U}{c^2}}, \quad (420)$$

$$\bar{v}_{2y} = \frac{v_{2y}}{\Gamma(U) \left[1 + \frac{v_{2x}U}{c^2}\right]}. \quad (421)$$

Desegnemos  $\bar{\theta}$  e  $\bar{\phi}$  os ângulos de espalhamento (sempre com o eixo  $x$ ) no referencial  $S$ . Eles serão dados, a partir das expressões acima, por,

$$\tan \bar{\theta} = \frac{1}{\Gamma(U)} \frac{v \sin \theta}{v \cos \theta + U}, \quad (422)$$

$$\tan \bar{\phi} = \frac{1}{\Gamma(U)} \frac{-v \sin \phi}{-v \cos \phi + U}. \quad (423)$$

Lembrando que  $U = v$ , e que  $\phi = \theta + \pi$ , obtemos,

$$\tan \bar{\theta} = \frac{1}{\Gamma(v)} \frac{\sin \theta}{\cos \theta + 1}, \quad (424)$$

$$\tan \bar{\phi} = \frac{1}{\Gamma(v)} \frac{\sin \theta}{-v \cos \theta + U}. \quad (425)$$

As velocidades, em módulo, são:

## 30 As Equações de Maxwell

As equações de Maxwell relacionam os campos elétrico e magnético entre si, e com as fontes, cargas e correntes. Elas se escrevem da seguinte forma:

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} ; \quad (426)$$

$$\nabla \cdot \vec{B} = 0 ; \quad (427)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} ; \quad (428)$$

$$\nabla \times \vec{B} = \mu_0 \vec{j} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} . \quad (429)$$

A relação de compatibilidade entre elas (por exemplo, a divergência de (30) combinada com a derivada parcial em relação ao tempo de (426)) nos conduz à equação da continuidade:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0 . \quad (430)$$

As grandezas fundamentais são os vetores campo elétrico e magnético e, no que diz respeito às fontes, a densidade de carga e o vetor densidade de corrente. Estes últimos possuem quatro componentes no total e podem ser representados por um quadri-vetor. Se levarmos em conta a covariância pelas transformações de Lorentz, deveremos escrever o vetor quadri-corrente como

$$J^\mu = \rho_0 v^\mu , \quad (431)$$

onde  $v^\mu = \gamma(c, \vec{v})$ . Logo,

$$J^\mu = \gamma(c\rho, \vec{j}) . \quad (432)$$

O fator  $\gamma$  revela que a densidade de carga, e consequentemente a densidade de corrente, são afetadas quando se passa de um referencial inercial a outro, devido à contração das distâncias.

Os vetores campo elétrico e magnético têm no total 6 componentes. Elas não podem ser acomodadas em um quadri-vetor, i.e., em um tensor de primeira ordem, como a densidade e a corrente. Um tensor de segunda ordem, carregando dois índices, contém em princípio 16 componentes. Se ele for simétrico, ele terá dez componentes, e se ele for anti-simétrico, ele terá 6 componentes. Somos conduzidos assim a pensar que os campos elétrico e magnético podem ser acomodados em um tensor anti-simétrico de segunda ordem. A questão consiste em

como construir este tensor anti-simétrico de segunda ordem, identificando suas componentes com as componentes dos campos elétrico e magnético de modo a satisfazer as equações de Maxwell escritas anteriormente.

Como já vimos, as equações de Maxwell homogêneas, (28,29) permitem escrever os campos elétrico e magnético em termos de grandezas auxiliares, os potenciais vetor e escalar:

$$\vec{B} = \nabla \times \vec{A} \quad , \quad (433)$$

$$\vec{E} = -\nabla\phi - \frac{\partial\vec{A}}{\partial t} \quad . \quad (434)$$

Os campos elétrico e magnético assim obtidos são invariantes pelas transformações de calibre,

$$\vec{A} \rightarrow \vec{A} + \nabla\chi \quad , \quad (435)$$

$$\phi \rightarrow \phi - \frac{1}{c} \frac{\partial\chi}{\partial t} \quad . \quad (436)$$

Podemos partir de uma observação simples: os potenciais eletromagnéticos são um vetor e um escalar, totalizando quatro componentes. A partir deles podemos pensar em construir um quadri-vetor. Vamos considerar então o seguinte quadri-vetor,

$$A^\mu = (\phi, c\vec{A}) \quad . \quad (437)$$

O fator  $c$  foi introduzido por razões dimensionais. Observemos que os campos são obtidos a partir dos potenciais por processo de derivação. Como desejamos identificar os campos como componentes de um tensor anti-simétrico de segunda ordem, e que seja construído a partir dos potenciais, temos uma escolha natural para a expressão que relaciona os potenciais ao tensor anti-simétrico de segunda ordem, que denotamos  $F_{\mu\nu}$ :

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \quad (438)$$

Observe-se que o tensor  $F_{\mu\nu}$  foi definido covariante, enquanto o quadri-vetor potencial  $A^\mu$  foi definido contravariante. Logo, nas expressões acima conectando o tensor  $F_{\mu\nu}$  aos potenciais  $A^\mu$ , deveremos considerar que  $A_\mu = (\phi, -c\vec{A})$ . Obviamente, temos que

$$F_{\mu\nu} = -F_{\nu\mu} \quad . \quad (439)$$

O nosso problema agora consiste em determinar as componentes deste tensor de segunda ordem. Temos

$$F_{0i} = \partial_0 A_i - \partial_i A_0 = -\left(\frac{\partial \vec{A}}{\partial t}\right)_i - (\nabla \phi)_i = E_i \quad , \quad (440)$$

$$F_{12} = \partial_1 A_2 - \partial_2 A_1 = \frac{\partial A_1}{\partial x_2} - \frac{\partial A_2}{\partial x_1} = (\nabla \times \vec{A})_3 = B_3 \quad , \quad (441)$$

$$F_{23} = \partial_2 A_3 - \partial_3 A_2 = \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} = (\nabla \times \vec{A})_1 = B_1 \quad , \quad (442)$$

$$F_{31} = \partial_3 A_1 - \partial_1 A_3 = \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} = (\nabla \times \vec{A})_2 = B_2 \quad . \quad (443)$$

Devido à anti-simetria do tensor  $F_{\mu\nu}$ , temos  $F_{i0} = -E_i$ ,  $F_{21} = -B_3$ ,  $F_{32} = -B_1$ ,  $F_{13} = -B_2$ . Logo, o tensor  $F_{\mu\nu}$  pode ser representado pela matriz,

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix} \quad (444)$$

Esse tensor contém as componentes dos campos elétrico e magnético.

A forma contravariante do tensor  $F_{\mu\nu}$  (que daqui por diante denominado tensor eletromagnético ou tensor de Maxwell) dada por  $F^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}F_{\rho\sigma}$ , escreve-se

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix} \quad (445)$$

As equações de Maxwell envolvem derivadas dos campos que, em duas delas, estão conectadas com os termos de fonte. Desta forma, podemos supor que as duas equações de Maxwell não-homogêneas podem ser escritas como

$$F_{,\mu}^{\mu\nu} = kJ^\nu \quad , \quad (446)$$

onde a vírgula denota derivada com relação às coordenadas  $x_\mu$ , e  $k$  é uma constante a ser determinada.

Podemos desenvolver esta expressão para os casos  $\nu = 0$  e  $\nu = i$ :

- $\nu = 0$ :

$$\begin{aligned}\partial_0 F^{00} + \partial_i F^{i0} &= kJ^0 \quad , \\ \partial_i F^{i0} &= k\rho c \quad .\end{aligned}\tag{447}$$

A comparação com a primeira das equações de Maxwell e o fato que  $F^{i0} = E_i$  permitem escrever  $k = \frac{1}{c\epsilon_0}$ .

- $\nu = i$ :

$$\begin{aligned}\partial_0 F^{0i} + \partial_j F^{ji} &= kJ^i \quad , \\ \frac{\partial E_i}{c\partial t} + \frac{\partial F^{ij}}{\partial x^j} &= k\vec{j} \quad .\end{aligned}\tag{448}$$

Desenvolvendo esta expressão, observamos que a quarta equação de Maxwell é reproduzida se usarmos  $F^{12} = -cB_3$ ,  $F^{31} = -cB_2$  e  $F^{23} = -cB_1$ .

Pode-se, a partir do tensor  $F^{\mu\nu}$ , obter outro tensor de segunda ordem anti-simétrico, utilizando-se o tensor completamente anti-simétrico de Levi-Civita a quatro dimensões:

$${}^*F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \quad ,\tag{449}$$

onde  $\epsilon^{0123} = 1$  e qualquer permutação de dois índices conduz a uma mudança de sinal.  ${}^*F^{\mu\nu}$  é conhecido como *tensor eletromagnético dual*. Cálculo direto nos permite escrever as matrizes associadas a  ${}^*F^{\mu\nu}$  e a  ${}^*F_{\mu\nu}$ :

$${}^*F^{\mu\nu} = \begin{pmatrix} 0 & cB_x & cB_y & cB_z \\ -cB_x & 0 & -E_z & E_y \\ -cB_y & E_z & 0 & -E_x \\ -cB_z & -E_y & E_x & 0 \end{pmatrix} \quad ,\tag{450}$$

$${}^*F_{\mu\nu} = \begin{pmatrix} 0 & -cB_x & -cB_y & -cB_z \\ cB_x & 0 & -E_z & E_y \\ cB_y & E_z & 0 & -E_x \\ cB_z & -E_y & E_x & 0 \end{pmatrix} \quad .\tag{451}$$

A partir das expressões acima, podemos observar que a passagem do tensor eletromagnético  $F^{\mu\nu}$  ao seu dual implica nas substituições  $\vec{E} \rightarrow -c\vec{B}$  e  $c\vec{B} \rightarrow \vec{E}$ .

As duas equações de Maxwell homogêneas podem ser obtidas de

$$\partial_\mu {}^*F^{\mu\nu} = 0 \quad ,\tag{452}$$

como pode ser verificado usando os mesmos passos que no caso anterior das equações com fonte. Assim, as equações de Maxwell passam a ser escritas como,

$$\partial_\mu F^{\mu\nu} = \frac{J^\nu}{\epsilon_0 c} \quad , \quad (453)$$

$$\partial_\mu {}^*F^{\mu\nu} = 0 \quad . \quad (454)$$

Derivando a equação (453) em relação à coordenada  $\nu$ , e lembrando que  $\partial_\mu \partial_\nu F^{\mu\nu} = 0$  devido à anti-simetria do tensor eletromagnético, obtemos

$$\partial_\mu J^\mu = 0 \quad . \quad (455)$$

A equação (455) expressa de forma covariante a equação da continuidade.

Desta forma, conseguimos construir um tensor eletromagnético, anti-simétrico, de segunda ordem, que resume as equações de Maxwell em duas equações tensoriais, consequentemente covariantes pelas transformações de Lorentz, que são (453,454). Além disto, a própria construção do tensor eletromagnético mostra como ele se relaciona com o potencial  $A^\mu$  dado por (437). As fontes são dadas pelo quadri-vetor (432).

Resta-nos agora tornar covariante as transformações de calibre (435,436). Isto pode ser obtido escrevendo-se

$$A'_\mu = A_\mu - \partial_\mu \chi \quad . \quad (456)$$

O calibre de Lorentz assume a forma

$$\partial_\mu A^\mu = \frac{1}{c} \frac{\partial \phi}{\partial t} + c \nabla \cdot \vec{A} = 0 \quad . \quad (457)$$

Desta forma, todas as relações fundamentais do eletromagnetismo de Maxwell estão escritas em forma tensorial, manifestamente covariante.

## 31 Leis de transformação para os campos elétrico e magnético

Na passagem de um referencial inercial  $S$  para um referencial  $S'$ , que se move em relação a  $S$  com velocidade constante  $V$  que consideramos, por simplicidade, paralela ao eixo dos  $x$ , os campos elétrico e magnético devem ser expressos de forma

diferente em um referencial e no outro. O fato que o tensor eletromagnético  $F_{\mu\nu}$  obedece à leis de transformação usuais de um tensor, nos dirá exatamente como devemos escrever os campos  $\vec{E}'$  e  $\vec{B}'$ , medidos no referencial  $S'$ , em função dos campos  $\vec{E}$ ,  $\vec{B}$  e da velocidade relativa entre os dois referenciais  $V$ .

É óbvio que os campos nos referenciais  $S$  e  $S'$  não devem ter as mesmas expressões. Por exemplo, uma carga em repouso no referencial  $S$  cria apenas um campo elétrico; no referencial  $S'$  esta carga terá uma velocidade  $-V$  e, além do campo elétrico, ela criará também campo magnético. Mas, precisamos estabelecer exatamente como isto ocorre, não apenas para esta situação, relativamente trivial, como para situações mais gerais. A covariância das equações de Maxwell, em particular do tensor eletromagnético, pelas transformações de Lorentz nos conduzirá nesta tarefa.

O tensor eletromagnético  $F'_{\mu\nu}$  no referencial  $S'$  se escreve em termos do tensor eletromagnético  $F_{\mu\nu}$  no referencial  $S$  como

$$F'_{\mu\nu} = \Lambda_{\mu'}^{\rho} \Lambda_{\nu'}^{\sigma} F_{\rho\sigma} \quad , \quad (458)$$

onde, como antes,  $\Lambda_{\mu'}^{\rho}$  representa a matriz da transformação de Lorentz. Podemos então calcular as diversas componentes do tensor  $F'_{\mu\nu}$ . Para a componente  $F'_{01}$  temos

$$\begin{aligned} F'_{01} &= E'_x = \Lambda_0^{\rho} \Lambda_1^{\sigma} F_{\rho\sigma} = \Lambda_0^0 \Lambda_1^1 F_{01} + \Lambda_0^1 \Lambda_1^0 F_{10} \\ &= \Gamma^2 E_x - \frac{V^2}{c^2} \Gamma E_x = E_x \quad , \end{aligned} \quad (459)$$

onde  $\Gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$  é o fator de Lorentz. As componentes  $F'_{02} = E'_y$  e  $F'_{03} = E'_z$  têm comportamento similar. Assim, apresentamos apenas o cálculo para uma delas:

$$\begin{aligned} F'_{02} &= E'_y = \Lambda_0^{\rho} \Lambda_2^{\sigma} F_{\rho\sigma} = \Lambda_0^0 \Lambda_2^2 F_{02} + \Lambda_0^2 \Lambda_2^0 F_{20} + \Lambda_0^1 \Lambda_2^2 F_{12} \\ &= \Gamma E_y + \frac{V}{c} \Gamma c B_z = \Gamma (E_y + V B_z) \quad . \end{aligned} \quad (460)$$

Repetindo o cálculo para a componente  $F'_{03}$ , obtém-se

$$E'_z = \Gamma (E_z - V B_y) \quad . \quad (461)$$

As componentes do campo magnético se calculam de forma similar. A componente  $F'_{32} = cB'_x$  se apresenta como

$$F'_{32} = cB'_x = \Lambda_{3'}^\rho \Lambda_{2'}^\sigma F_{\rho\sigma} = \Lambda_{3'}^3 \Lambda_{2'}^2 F_{32} = cB_x \quad . \quad (462)$$

Para a componente  $F'_{13} = cB'_y$  obtém-se

$$\begin{aligned} F'_{13} &= cB'_y = \Lambda_{1'}^\rho \Lambda_{3'}^\sigma F_{\rho\sigma} = \Lambda_{1'}^0 \Lambda_{3'}^3 F_{03} + \Lambda_{1'}^1 \Lambda_{3'}^3 F_{13} \\ &= -\frac{V}{c} \Gamma E_z + \Gamma cB_y = \Gamma \left( cB_y - \frac{V}{c} E_z \right) \quad , \end{aligned} \quad (463)$$

enquanto para a componente  $F'_{21}$  se escreve

$$cB'_y = \Gamma \left( cB_y + \frac{V}{c} E_z \right) \quad . \quad (464)$$

Resumindo, as expressões para as transformações dos campos elétrico e magnético se escrevem como

$$E'_x = E_x \quad , \quad B'_x = B_x \quad , \quad (465)$$

$$E'_y = \Gamma(E_y + V B_z) \quad , \quad B'_y = \Gamma \left( B_y - \frac{V}{c^2} E_z \right) \quad , \quad (466)$$

$$E'_z = \Gamma(E_z - V B_x) \quad , \quad B'_z = \Gamma \left( B_z + \frac{V}{c^2} E_x \right) \quad . \quad (467)$$

Estas expressões podem ser reescritas de uma forma mais compacta lembrando-se que a velocidade  $V$  entre os referenciais  $S$  e  $S'$  está orientada ao longo do eixo  $x$ . Uma inspeção direta das expressões (465,466,467) revela que, vetorialmente, elas podem ser escritas como

$$\vec{E}'_{\parallel} = \vec{E}_{\parallel} \quad , \quad \vec{B}'_{\parallel} = \vec{B}_{\parallel} \quad , \quad (468)$$

$$\vec{E}'_{\perp} = \Gamma(\vec{E}_{\perp} - \vec{V} \times \vec{B}) \quad , \quad \vec{B}'_{\perp} = (\vec{B}_{\perp} + \frac{\vec{V}}{c^2} \times \vec{E}) \quad , \quad (469)$$

onde os símbolos  $\parallel$  e  $\perp$  denotam as direções paralela e perpendicular à velocidade relativa entre os dois referenciais  $\vec{V}$ .

## 32 Aplicação: a força de Lorentz

De posse das leis de transformação para o campo elétrico e magnético, e utilizando também as leis de transformação para a quadri-força, podemos fazer



uma importante aplicação: deduzir a lei de força de Lorentz a partir da lei de força elétrica devido a um campo  $\vec{E}$  que age sobre uma carga. Esta aplicação é particularmente interessante pois normalmente, vista como uma teoria tri-dimensional, e ignorando a sua natureza relativista, o eletromagnetismo seria obrigado a considerar a força de Lorentz como uma expressão empírica. Considerando a estrutura relativista do eletromagnetismo, a força de Lorentz aparece como consequência da força que age sobre uma carga em repouso devido a sua interação com um campo elétrico.

Para verificar isto, considere uma carga  $q$  sobre a qual age um campo elétrico  $\vec{E}$  em um dado referencial  $S$  onde esta carga esteja em repouso:

$$\vec{f} = q\vec{E} \quad . \quad (470)$$

As expressões que relacionam as componentes da força  $\vec{f}$  agindo sobre um dado corpo medidas em  $S$  com as componentes da força  $\vec{f}'$  medida em  $S'$  são

$$\vec{f}_{\parallel} = \frac{\vec{f}'_{\parallel} + \frac{\vec{V}\vec{f}' \cdot \vec{v}'}{c^2}}{1 + \frac{\vec{v}' \cdot \vec{V}}{c^2}} \quad , \quad (471)$$

$$\vec{f}_{\perp} = \frac{\vec{f}'_{\perp}}{\Gamma(1 + \frac{\vec{v}' \cdot \vec{V}}{c^2})} \quad , \quad (472)$$

onde  $\vec{v}'$  é a velocidade deste corpo medida em  $S'$  e  $\vec{V}$  é a velocidade relativa entre os dois referenciais. Como antes, a força é separada em suas componentes paralela e perpendicular à velocidade relativa  $\vec{V}$ . Como a carga se encontrava em repouso em  $S$ ,  $\vec{v}' = -\vec{V}$ , o que reduz as expressões acima para

$$\vec{f}_{\parallel} = \frac{\vec{f}'_{\parallel} - \frac{\vec{V}\vec{f}' \cdot \vec{V}}{c^2}}{1 - \frac{\vec{V} \cdot \vec{V}}{c^2}} = \frac{\vec{f}'_{\parallel} - \frac{\vec{f}' V^2}{c^2}}{1 - \frac{V^2}{c^2}} = \vec{f}'_{\parallel} \quad , \quad (473)$$

$$\vec{f}_{\perp} = \frac{\vec{f}'_{\perp}}{\Gamma(1 - \frac{V^2}{c^2})} = \Gamma \vec{f}'_{\perp} \quad , \quad (474)$$

onde usamos o fato que  $\vec{V}\vec{f}' \cdot \vec{V} = \vec{f}'_{\parallel} V^2$  devido à própria definição de  $\vec{f}'_{\parallel}$  como sendo a componente de  $\vec{f}'$  paralela a  $\vec{V}$ .

A tarefa agora consiste em inserir (473,474) na expressão (470) e utilizar as transformações (468,484) inversas. Para a componente paralela encontra-se

$$\vec{f}_{\parallel} = \vec{f}'_{\parallel} = q\vec{E}_{\parallel} = q\vec{E}'_{\parallel} \quad . \quad (475)$$

Para a componente perpendicular, tem-se

$$\vec{f}_\perp = \Gamma \vec{f}'_\perp = q \vec{E}_\perp = q \Gamma (\vec{E}'_\perp + \vec{V} \times \vec{B}') \Rightarrow \vec{f}'_\perp = q (\vec{E}'_\perp + \vec{V} \times \vec{B}') \quad . \quad (476)$$

Utilizando agora as duas expressões acima, obtém-se

$$\vec{f}'_\parallel + \vec{f}'_\perp = \vec{f}' = q \{ \vec{E}'_\parallel + \vec{E}'_\perp + \vec{V} \times \vec{B}' \} = q \{ \vec{E}' + \vec{V} \times \vec{B}' \} \quad . \quad (477)$$

Assim, um observador no referencial  $S'$  verá a carga  $q$  que se move com velocidade  $\vec{V}$  em uma região onde existem campos elétrico e magnético  $\vec{E}'$  e  $\vec{B}'$  sendo submetida à força de Lorentz.

A dedução da lei de força que age sobre uma carga  $q$  em movimento submetida a campos elétrico e magnético, a força de Lorentz, é um exemplo notável de como a exploração da natureza intrinsecamente relativista do eletromagnetismo pode lançar luz sobre expressões que, vistas do ponto de vista tridimensional e não relativista, devem ser obtidas a partir da experiência, empiricamente.

### 33 Invariantes eletromagnéticos

Como já foi discutido no capítulo 3, escalares construídos a partir das quantidades físicas fundamentais desempenham um papel importante em Relatividade Restrita. Um dos aspectos essenciais da Relatividade Restrita reside na covariância, isto é, na invariância das leis fundamentais por transformações de Lorentz. Entretanto, ao resolver um problema específico, escolhe-se um referencial particular. Na análise dos resultados obtidos neste referencial é fundamental distinguir entre o que é específico ao referencial empregado, e o que independe do referencial. Uma forma de realizar esta distinção é através do estudo dos escalares (logo, invariantes por mudança de referencial) construídos a partir das quantidades fisicamente relevantes. Veremos uma aplicação concreta disto para o caso do eletromagnetismo.

As quantidades fundamentais do eletromagnetismo expressas na sua formulação manifestamente covariante a quatro dimensões são o tensor eletromagnético  $F_{\mu\nu}$  e o termo de fonte representado pela quadri-corrente  $j^\mu$ . Com o termo de fonte, é possível formar o escalar  $j^2 = j^\mu j_\mu$ . Usando a expressão explícita para

a quadri-corrente (431), pode-se determinar  $j^2$ :

$$j^2 = j^\mu j_\mu = j^0{}^2 - \sum_{i=1}^3 j^i{}^2 = \gamma^2(c^2 \rho_0^2 - v^2 \rho_0^2) = c^2 \rho_0^2 \quad . \quad (478)$$

Esta expressão apenas nos diz que  $j^2$  está relacionada com a densidade de carga no referencial próprio do sistema.

Mais relevantes são as quantidades que podem ser construídas a partir do tensor eletromagnético  $F_{\mu\nu}$ . Na verdade, como foi visto anteriormente, o formalismo manifestamente covariante do eletromagnetismo emprega também o tensor eletromagnético dual  $*F_{\mu\nu}$ . Pode-se construir, em princípio, três quantidades escalares com o tensor eletromagnético e o seu dual:

$$I_1 = F^{\mu\nu} F_{\mu\nu} \quad , \quad I_2 = *F^{\mu\nu} F_{\mu\nu} \quad , \quad I_3 = *F^{\mu\nu} *F_{\mu\nu} \quad . \quad (479)$$

Usando a definição do tensor dual  $*F^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$ , e o fato que  $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu\nu\lambda\gamma} = \delta_\lambda^\rho \delta_\gamma^\sigma - \delta_\gamma^\rho \delta_\lambda^\sigma$ , pode-se mostrar que  $I_3 = 2I_1$ . Assim, os únicos invariantes realmente independentes são  $I_1$  e  $I_2$ .

Podemos calcular os invariantes simplesmente através da multiplicação das matrizes que representam o tensor eletromagnético e o seu dual. Por exemplo, construindo a matrix  $M_\nu^\mu = F^{\mu\rho} F_{\rho\nu}$ , tem-se que  $I_1 = M_\mu^\mu = Tr(M_\nu^\mu)$ . Assim,

$$\begin{aligned} I_1 &= Tr \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix} \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix} \\ &= 2(E^2 - c^2 B^2) \quad . \end{aligned} \quad (480)$$

Da mesma forma,  $I_2 = Tr(N^{\mu\nu})$ , sendo  $N_\nu^\mu = *F^{\mu\rho} F_{\rho\nu}$ . O cálculo explícito de  $I_2$  conduz a

$$\begin{aligned} I_2 &= Tr \begin{pmatrix} 0 & cB_x & cB_y & cB_z \\ -cB_x & 0 & -E_z & E_y \\ -cB_y & E_z & 0 & -E_x \\ -cB_z & -E_y & E_x & 0 \end{pmatrix} \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix} \\ &= -4c \vec{E} \cdot \vec{B} \quad . \end{aligned} \quad (481)$$

Qual é o significado e a utilidade destes invariantes? O valor numérico destes invariantes é o mesmo em qualquer sistema de referência, exatamente porque tratam-se de *invariantes*. Sabe-se que em alguns referenciais uma configuração de fontes conduz à existência apenas de campo elétrico, ao passo que uma outra configuração implica a existência apenas de campo magnético. Um exemplo do primeiro caso é uma carga  $q$  em repouso, enquanto um exemplo do segundo caso é uma corrente atravessando um condutor que é globalmente neutro eletricamente. Em outros referenciais, estas configurações levarão a uma combinação de campos elétrico e magnético, devido às leis de transformação de campos escritas anteriormente. Inversamente, os valores dos invariantes  $I_1$  e  $I_2$  permitem saber se existe, para uma dada configuração de campos elétrico e magnético, um referencial onde só existirá o campo elétrico ou o campo magnético. Suponhamos, por exemplo, que  $I_1 > 0$  e  $I_2 = 0$ . Neste caso, sempre será possível encontrar um referencial onde existirá apenas campo elétrico. Por outro lado, se  $I_1 < 0$  e  $I_2 = 0$ , existe um referencial onde apenas ocorrerá a presença de campo magnético. No entanto, se  $I_2 \neq 0$ , em todos os referenciais haverá campos elétrico e magnético, independentemente do valor de  $I_1$ . O caso particular onde  $I_1 = I_2 = 0$  corresponde a uma onda eletromagnética: isto implica  $E = cB$  e  $\vec{E} \perp \vec{B}$ , o que, de fato, caracteriza uma onda eletromagnética.

Neste sentido, por exemplo, o conhecimento dos invariantes, obtido a partir dos valores dos campos elétrico e magnético, permite saber se existe uma carga líquida diferente de zero (sempre haverá campo elétrico), ou se a carga total é nula (é possível só se ter campo magnético).

### 34 Exemplos de campos elétrico e magnético

O cálculo dos campos elétrico e magnético para uma dada distribuição de fontes não é, normalmente, uma tarefa simples. No eletromagnetismo clássico, técnicas matemáticas apuradas são empregadas para se obter respostas a problemas às vezes aparentemente simples. Em situações estáticas, a determinação dos campos elétrico e magnético se passa da forma habitual, aparentemente sem nenhuma consideração relativista. A palavra "aparentemente" foi usada porque, no fundo as equações de Maxwell são essencialmente relativistas. En-

tretanto, esta natureza relativista se manifesta de forma mais clara em situações dinâmicas, onde a configuração de campos muda com o tempo. Neste caso, se se quer manter a estrutura relativista, é preciso ter em mente que os sinais, inclusive a mudança do valor do campo em um dado ponto a uma distância  $\vec{r}$  de uma carga em movimento, se propagam com a velocidade da luz.

Mesmo que o emprego da relatividade restrita possa conduzir a situações extremamente complexas no estudo de problemas de eletrodinâmica, o fato que as equações relativistas em geral, e as do eletromagnetismo em particular, são covariantes, pode frequentemente simplificar a tarefa de determinar as configurações de campos: calcula-se o campo em um dado referencial, onde o problema adquire uma forma mais simples, e depois transpõe-se o resultado para o referencial em que se pretende conhecer aquela configuração. O fato de que sabemos como expressar os campos elétrico e magnético em um referencial  $S$  em função destes mesmos campos medidos no referencial  $S'$  é, obviamente, crucial para que possamos assim proceder. Daremos a seguir dois exemplos deste procedimento, um envolvendo essencialmente o campo elétrico e outro o campo magnético.

### 34.1 O campo de uma carga elétrica em movimento

Considere-se uma carga elétrica pontual  $q$  que se move com velocidade constante  $\vec{v}$  em um dado referencial  $S$ . Já que as equações são invariantes por uma rotação dos eixos coordenados espaciais, usaremos este fato para fixar que a partícula se move retilineamente ao longo do eixo  $x$ :  $\vec{v} = v\vec{i}$ . Normalmente, este problema apresenta uma dificuldade não trivial: a carga em movimento criará campos elétrico e magnético em um dado ponto  $\vec{r}$ , cujo valor em um dado tempo  $t$  é devido à posição da carga em um momento  $t' = t - \frac{r}{c}$  anterior.

Entretanto, este problema pode ser tratado de uma forma relativamente simples se ele é inicialmente analisado em um referencial  $S'$  onde a carga está em repouso, sendo depois os resultados transpostos para o referencial  $S$  através das leis de transformações para os campo elétrico e magnético determinadas anteriormente.

No referencial  $S'$ , que se move em relação a  $S$  com velocidade  $\vec{V} = \vec{v} = v\vec{i}$ , a carga  $q$  encontra-se em repouso. Podemos supor sem perda de generalidade

que a carga  $q$  encontra-se na origem de  $S'$ . Neste caso os campos elétrico e magnético em um dado ponto distante  $\vec{r}'$  da carga, assumem a forma

$$\vec{E}' = \frac{1}{4\pi\epsilon_0} \frac{q}{r'^3} \vec{r}' \quad , \quad \vec{B}' = 0 \quad . \quad (482)$$

Utilizando as leis de transformação para os campos elétrico e magnético (468,484), podemos escrever

$$\vec{E}_x = \vec{E}'_x \quad , \quad E_y = \Gamma E'_y \quad , \quad E_z = \Gamma E'_z \quad , \quad (483)$$

$$B_x = 0 \quad , \quad B_y = \Gamma(V) \frac{V}{c^2} E'_z \quad , \quad B_z = -\Gamma(V) \frac{V}{c^2} E'_y \quad . \quad (484)$$

Deve-se observar que

$$r'^2 = x'^2 + y'^2 + z'^2 = \Gamma(V)^2 (x - Vt)^2 + y^2 + z^2 \quad . \quad (485)$$

Desta forma, podemos escrever as expressões finais para os campos elétrico e magnético para uma carga em movimento uniforme:

$$E_x = \frac{1}{4\pi\epsilon_0} \Gamma(V) \frac{q(x - Vt)}{[\Gamma(V)^2 (x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (486)$$

$$E_y = \frac{1}{4\pi\epsilon_0} \Gamma(V) \frac{qy}{[\Gamma(V)^2 (x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (487)$$

$$E_z = \frac{1}{4\pi\epsilon_0} \Gamma(V) \frac{qz}{[\Gamma(V)^2 (x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (488)$$

$$B_x = 0 \quad , \quad (489)$$

$$B_y = \frac{1}{4\pi\epsilon_0} \Gamma(V) \frac{V}{c^2} \frac{qy}{[\Gamma(V)^2 (x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (490)$$

$$B_z = -\frac{1}{4\pi\epsilon_0} \Gamma(V) \frac{V}{c^2} \frac{qy}{[\Gamma(V)^2 (x - Vt)^2 + y^2 + z^2]^{3/2}} \quad . \quad (491)$$

No limite não-relativista, tem-se que  $\Gamma(V) \sim 1$ , o que implica que as relações acima tomam a forma

$$E_x \approx \frac{1}{4\pi\epsilon_0} \frac{q(x - Vt)}{[(x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (492)$$

$$E_y \approx \frac{1}{4\pi\epsilon_0} \frac{qy}{[(x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (493)$$

$$E_z \approx \frac{1}{4\pi\epsilon_0} \frac{qz}{[(x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (494)$$

$$B_x = 0 \quad , \quad (495)$$

$$B_y \approx \frac{\mu_0}{4\pi} \frac{qVy}{[(x - Vt)^2 + y^2 + z^2]^{3/2}} \quad , \quad (496)$$

$$B_z \approx -\frac{\mu_0}{4\pi} \frac{qVy}{[(x - Vt)^2 + y^2 + z^2]^{3/2}} \quad . \quad (497)$$

As expressões para o campo magnético tomam a forma conhecida para o caso de uma corrente criada por uma carga pontual, com a propagação do campo sendo considerada como praticamente instantânea, como era de se esperar.

### 34.2 O campo criado por uma corrente uniforme em um fio retilíneo eletricamente neutro

Vamos considerar agora um fio retilíneo percorrido por uma corrente uniforme  $i$ . O fio é eletricamente neutro. Neste caso tem-se um referencial  $S$  onde a densidade de cargas positivas e negativas é a mesma; conseqüentemente, o campo elétrico neste referencial é nulo. Tem-se apenas o campo magnético que pode ser obtido aplicando-se diretamente a lei de Ampère:

$$\vec{B} = \frac{\mu_0 i}{4\pi r} \hat{\theta}, \quad (498)$$

onde  $\hat{\theta}$  é o vetor unitário associado à coordenada  $\theta$ .

No referencial  $S$ , as expressões para os campos elétrico e magnético são as seguintes:

$$\vec{E}_{\parallel} = \vec{E}_{\perp} = 0 \quad , \quad (499)$$

$$\vec{B}_{\parallel} = 0 \quad , \quad \vec{B}_{\perp} = \frac{\mu_0 i}{4\pi r} \hat{\theta} \quad , \quad (500)$$

onde os termos paralelo e perpendicular referem-se à direção de propagação da corrente.