



Universidade Estadual de Maringá
Centro de Ciências Exatas
Departamento de Física

Trabalho de Conclusão de Curso

Desenvolvimento da Teoria da Relatividade Restrita e Incorporação na Mecânica Quântica

Acadêmico: João Victor Balieiro da Silva

Orientador: Prof. Dr. Miguel Jorge Bernabé Ferreira

Coorientadora: Profa. Dra. Hatsumi Mukai

Maringá, 7 de maio de 2021



Universidade Estadual de Maringá
Centro de Ciências Exatas
Departamento de Física

Trabalho de Conclusão de Curso

Desenvolvimento da Teoria da Relatividade Restrita e Incorporação na Mecânica Quântica

Trabalho de conclusão de curso apresentado ao Departamento de Física, da Universidade Estadual de Maringá, como parte dos requisitos para obtenção do título de bacharel em Física

Acadêmico: João Victor Balieiro da Silva
Orientador: Prof. Dr. Miguel Jorge Bernabé Ferreira
Coorientadora: Profa. Dra. Hatsumi Mukai

Maringá, 30 de julho de 2020

Sumário

Resumo	5
Introdução	6
1 Motivação Histórica da Teoria da Relatividade	8
1.1 Princípio da Relatividade na Mecânica	8
1.2 Princípio da Relatividade no Eletromagnetismo	10
1.3 Transformações de Galileu e Equações de Maxwell	11
2 Transformações de Lorentz	12
2.1 Dedução das transformações	12
2.2 Consequências cinemáticas	16
2.3 Transformações de Lorentz entre referenciais com velocidade relativa numa direção arbitrária sem rotação relativa	18
2.4 Geometria da Relatividade Restrita	19
2.4.1 Distância na mecânica newtoniana	19
2.4.2 Distância na relatividade	20
2.5 Matriz de transformação e transformação geral de Lorentz	21
3 Grupo de Lorentz	24
3.1 Grupos	24
3.2 Matrizes de Lorentz e estrutura de grupo	25
3.3 Propriedades do grupo de Lorentz	26
3.4 Interpretação geométrica do grupo de Lorentz	29
3.4.1 Espaço euclidiano 2D	29
3.4.2 Espaço de Minkowski (1+1)D	31
4 Dinâmica relativística em quatro dimensões	34
4.1 Quadri vetores	34
4.1.1 Quadrivelocidade	34
4.1.2 Quadriaceleração	35

4.1.3	Quadrimento	36
4.2	Equação de movimento	36
4.3	Energia relativística e partículas sem massa	39
4.4	Formulação lagrangiana e hamiltoniana	41
5	Incorporação na Mecânica Quântica	44
5.1	Equação de Schrödinger	44
5.1.1	Interpretação estatística	45
5.2	Equação de Klein-Gordon	45
5.2.1	Interpretação Estatística	46
5.3	Equação de Dirac	47
5.3.1	Interpretação estatística	51
6	Considerações finais	53
A	Álgebra tensorial	55
A.1	Componentes de um vetor	55
A.2	Diática e transformação de coordenadas	55
A.3	Contração	56
A.4	Tensor métrico	57

Resumo

Este trabalho teve como objetivo fazer uma revisão da teoria da Relatividade Restrita na forma de um compilado geral desta teoria. A motivação para este trabalho foi que, inicialmente, a intenção era fazer um trabalho sobre Teorias Clássicas de Campos, porém, para tal feito, viu-se que era necessário um conhecimento maior sobre a Teoria da Relatividade Restrita por parte do autor, por esse motivo, decidimos trabalhar mais profundamente neste tema com o objetivo de explorar as Teorias Clássicas de Campos no futuro. Neste trabalho, partimos do desenvolvimento histórico da Relatividade Restrita, cujo ponto de partida é a imposição do princípio da relatividade sobre as equações de Maxwell. Tal imposição culmina em transformações das coordenadas especiais e do tempo chamadas de transformações de Lorentz. Exploramos também as principais consequências cinemáticas de tais transformações, a mudança na geometria do espaço que elas nos revelam e suas propriedades como a estrutura de Grupo que elas formam. Assumindo que as transformações de Lorentz são as verdadeiras transformações de coordenadas entre referenciais inerciais, foi apresentada a versão relativística para a mecânica newtoniana explorando algumas de suas consequências. Por fim, foram apresentadas versões relativísticas da equação de Schrödinger conhecidas como equação de Klein-Gordon e equação de Dirac, também apresentando algumas de suas principais consequências.

Introdução

A teoria do Eletromagnetismo de Maxwell, quando desenvolvida, foi calcada na ideia da existência de um meio material chamado de éter que seria o mediador das interações elétricas e magnéticas e o meio que suportaria as ondas eletromagnéticas. Assim como ondas na água, as propriedades das ondas eletromagnéticas e sua velocidade de propagação deveriam ser medidas no referencial de repouso do éter. Dessa forma, no final do séc. XIX e início de XX, os físicos estavam preocupados em medir a velocidade com que a Terra se movia através do éter. O espantoso é que todos os experimentos concluíram que não havia movimento relativo entre o éter e a Terra.

Com base nos experimentos realizados na época, foi postulado por Henri Poincaré [7] que as leis do eletromagnetismo, assim como as leis da mecânica newtoniana, deveriam ter a mesma forma em qualquer referencial inercial. Desse postulado surgem as chamadas transformações de Lorentz que são as transformações de coordenadas que tornam o eletromagnetismo invariante sobre uma mudança de referencial.

No Capítulo 1 apresentamos a razão histórica que levou à necessidade do desenvolvimento da Teoria da Relatividade Restrita, mais especificamente a discordância entre as transformações de Galileu e o Eletromagnetismo de Maxwell.

No Capítulo 2 exploramos as principais propriedades e consequências das transformações de Lorentz. Em particular, mostramos que referenciais que se movimentam com velocidade relativa constante, podem observar o mesmo evento acontecerem em instantes diferentes, medem intervalos de tempos diferentes e comprimentos diferentes. A partir das transformações de Lorentz, mostramos que na Teoria da Relatividade Restrita a geometria do espaço é diferente da geometria usual da mecânica newtoniana e a métrica gerada pelas transformações de Lorentz só tem sentido completo quando consideramos o tempo e o espaço conjuntamente, o chamado espaço-tempo.

Uma outra importante consequência, explorada no Capítulo 3, é que as transformações de Lorentz formam uma estrutura algébrica de teoria de Grupo e, com isso, é dado uma interpretação geométrica dessas transformações.

Conhecendo as transformações de Lorentz e suas propriedades, no Capítulo 4, apresenta-se por meio de um postulado que a mecânica também deve ser invariante sobre essas transformações e exploramos as principais consequências desse postulado estudando a mecânica clássica relativística no formalismo newtoniano e nos formalismos lagrangiano e hamiltoniano.

Por fim, no Capítulo 5, apresentamos um exemplo de como a Teoria da Relatividade Restrita pode ser incorporada à Mecânica Quântica, obtendo-se assim a equação de Dirac, que descreve o comportamento de campos quânticos relativísticos e que tem como uma de

suas consequências a interpretação de criação/destruição de partículas e anti-partículas. Finalizando com a apresentação no Apêndice [A](#) sobre álgebra tensorial, que é a ferramenta utilizada no desenvolvimento deste trabalho, e na sequência as referências bibliográficas, na ordem de alfabética.

Capítulo 1

Motivação Histórica da Teoria da Relatividade

Neste primeiro capítulo veremos a como a incompatibilidade das transformações de Galileu, já muito bem estabelecida para a mecânica newtoniana, com a Teoria eletromagnética de Maxwell trouxe a necessidade do surgimento da Teoria de Relatividade Restrita.

1.1 Princípio da Relatividade na Mecânica

A mecânica clássica formulada por Newton é definida sob uma classe especial de referenciais chamados de referenciais inerciais. Para definir estes referenciais, Newton define o conceito físico de espaço absoluto. O espaço absoluto não tem relação alguma com qualquer ente externo, é fixo e imutável. A princípio, apenas neste referencial as equações da mecânica são válidas.

Newton pode mostrar, usando o que conhecemos hoje por transformações de Galileu, que referenciais que se movem com velocidade constante em relação ao espaço absoluto também seriam igualmente bons para descrever os fenômenos mecânicos, ou seja, a forma de suas equações é a mesma em qualquer destes referenciais, chamados inerciais. Dessa forma, as leis de Newton são válidas para todos os referenciais inerciais.

Uma visualização das transformações de Galileu é dada a seguir. Suponha um referencial inercial S e um referencial S' que, visto de S , se move com uma velocidade constante $\vec{v} = v\hat{x}$. Assim, S' é também inercial. Suponha que observadores em S e S' tenham em mãos instrumentos capazes de medir comprimentos e intervalos de tempo, como réguas e relógios. Vamos impor que, o instante de tempo em que as origens de S e S' coincidem ocorre quando os tempos medidos em cada referencial t e t' marcam ambos 0 e que estes relógios estejam sincronizados antes de o observador em S' começar a sua viagem. Este esquema é mostrado na Figura 1.1. Para estes sistemas de referenciais, as relações entre as medidas de posição e tempo entre os observadores em S e S' , dadas pela mecânica newtoniana, as já mencionadas transformações de Galileu, são dadas por:

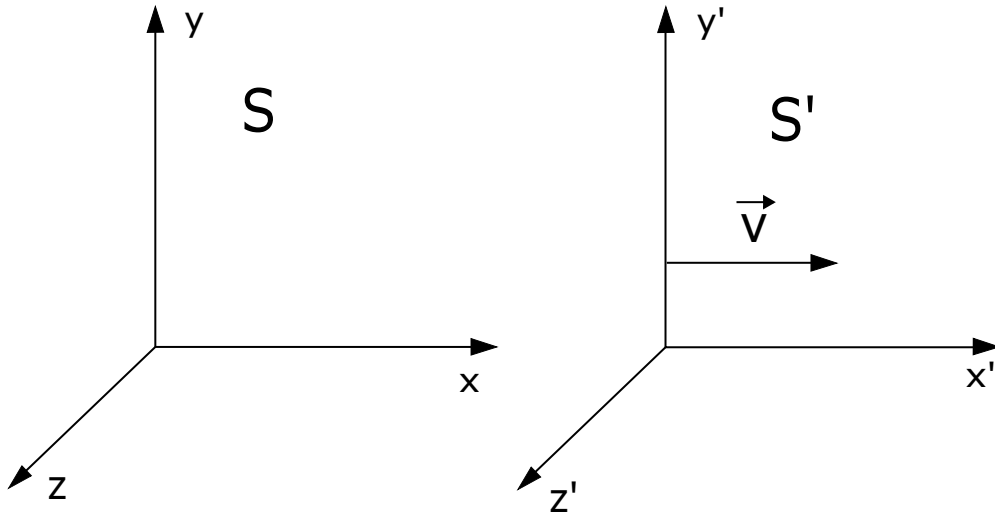


Figura 1.1: Referencial S' se movendo com velocidade \vec{v} paralela ao eixo x em relação ao referencial S

$$\begin{cases} x' = x - vt; \\ y' = y; \\ z' = z \\ t' = t. \end{cases} \quad (1.1)$$

As equações (1.1) expressam o fato de que o tempo, na mecânica newtoniana, deve passar da mesma maneira para ambos os observadores. Para garantir esta propriedade do tempo Newton, em seu livro, define o conceito físico de tempo absoluto. Novamente, assim como o espaço absoluto, o tempo absoluto também não sofre influencia de nenhum ente externo, é imutável e flui uniformemente.

Mostraremos agora que o princípio da relatividade é válido para a mecânica newtoniana quando usamos as transformações de Galileu (1.1). Suponha que um observador em repouso no referencial S' vê uma partícula de massa m constante em movimento no espaço sob a ação de uma força $\vec{F}' = F'_x \hat{x} + F'_y \hat{y} + F'_z \hat{z}$ como na figura 1.2.

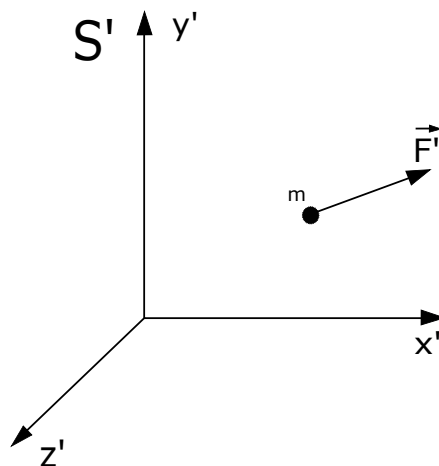


Figura 1.2: Partícula de massa m vista do referencial S' sob ação a força \vec{F}'

A segunda lei de Newton no referencial S' em cada componente espacial é dada por:

$$\begin{cases} \ddot{x}' = F'_x/m \\ \ddot{y}' = F'_y/m \\ \ddot{z}' = F'_z/m \end{cases}, \quad (1.2)$$

em que o índice " significa derivada segunda em relação ao tempo t' (ou t já que neste caso eles são iguais). Substituindo as equações (1.1) em (1.2) obtemos

$$\begin{cases} \ddot{x} = F'_x/m = F_x/m \\ \ddot{y} = F'_y/m = F_y/m \\ \ddot{z} = F'_z/m = F_z/m \end{cases}, \quad (1.3)$$

em que F_x, F_y, F_z são as componentes da força \vec{F}' quando medidas por um observador no referencial S.

De (1.3) obtemos que $F'_x = F_x, F'_y = F_y$ e $F'_z = F_z$. Além disso podemos ver que a forma das equações de movimento da partícula é a mesma nos dois referenciais mostrando que o princípio da relatividade é válido para a mecânica newtoniana quando as transformações de coordenadas utilizadas são as transformações de Galileu (1.1). Uma das consequências deste fato, já muito bem observado por Galileu em seu tempo, é que deve ser impossível medir a velocidade de um referencial inercial usando apenas experimentos que envolvam movimento de corpos (experimento mecânico). Para uma discussão mais aprofundada sobre o desenvolvimento histórico do princípio da relatividade para a mecânica ver [7].

1.2 Princípio da Relatividade no Eletromagnetismo

Em meados do séc. XIX muitos físicos acreditavam que a luz, por ser um fenômeno ondulatório, deveria se propagar por um meio material que permeava todo o espaço, denominado de éter luminífero. Assim, considerando que a Terra deve se mover através deste meio, deveria ser possível, usando algum experimento óptico, determinar a velocidade da Terra em relação ao éter.

Maxwell, quando sintetizou as equações do eletromagnetismo para descrever os fenômenos eletromagnéticos, fez isto baseado na ideia fundamental da existência do éter, meio que seria responsável por intermediar as interações. Em seus estudos, ele chegou a uma equação de onda para os campos elétrico e magnético que se propagam na mesma velocidade que a luz. Assim, ele imaginou que o éter seria o meio material que suportaria estas ondas. Dessa forma, se imaginava que tais equações só seriam válidas em um referencial em repouso em relação ao éter e que, devido ao suposto movimento da Terra em relação ao éter, poderia ser possível medir o movimento relativo da Terra em relação ao éter baseando-se em experimentos eletromagnéticos ou ópticos.

O experimento mais preciso da época foi o bem conhecido experimento de Michelson-Morley. Ele tinha o objetivo de determinar a diferença de velocidade entre dois feixes de luz que caminham em direções perpendiculares entre si vindos da mesma fonte usando um interferômetro. O resultado do experimento foi que os feixes tinham a mesma velocidade

e isso não dependia da direção que o interferômetro apontasse e nem da época do ano (já que se poderia argumentar que quando a Terra translada em torno do Sol ela pode estar em repouso com relação ao éter numa determinada época e em movimento em outra época). Este resultado deixou a comunidade científica em um grande impasse.

Algumas hipóteses foram criadas para tentar explicar este experimento mas que seriam apenas uma solução *ad hoc*. Por exemplo, Lorentz supôs que objetos que se movem paralelamente ao éter sofrem uma pequena contração compensando o caminho que um feixe andaria a mais que o outro no experimento, assim, explicando o resultado inesperado do experimento. Porém, nada garante de ante mão que isso devesse realmente acontecer.

A explicação mais aceita nos dias de hoje é de que o resultado do experimento nos mostra que as equações de Maxwell não devem valer apenas no referencial do éter, mas sim em qualquer referencial inercial que se move com velocidade constante em relação ao éter, ou seja, não importa em que referencial inercial estejamos, as equações de Maxwell devem ser as mesmas. Assim, é imposto que o princípio da relatividade não seja válido apenas para a mecânica, mas também para o eletromagnetismo.

1.3 Transformações de Galileu e Equações de Maxwell

Aqui surge um problema. Ao usarmos as transformações de Galileu nota-se que as equações de Maxwell não são as mesmas para qualquer referencial inercial. Vejamos, por exemplo, a equação de onda eletromagnética sem fontes. Suponha que as equações de Maxwell são válidas no referencial S, mostrado na Figura 1.1. A equação de onda sem fontes para o campo elétrico tem a seguinte forma:

$$\nabla^2 \vec{E} = \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2}, \quad (1.4)$$

supondo que o princípio da relatividade seja válido para o eletromagnetismo, no referencial inercial S' devemos obter uma equação de onda da mesma forma, já que as próprias equações de Maxwell tem a mesma forma:

$$\nabla'^2 \vec{E}' = \frac{1}{c^2} \frac{\partial^2 \vec{E}'}{\partial t'^2}, \quad (1.5)$$

em que os índices (') significam que estamos definindo as grandezas no referencial S'. Devemos destacar aqui uma importante conclusão: apenas impondo o princípio da relatividade para o eletromagnetismo, chegamos à conclusão de que a velocidade das ondas eletromagnéticas devem ser as mesmas em qualquer referencial. Ou seja, concluímos que o segundo postulado de Einstein nada mais é do que uma consequência do primeiro.

Analisando o que foi exposto acima chegamos a conclusão de que este resultado contradiz as transformações de Galileu, pois como a onda tem a mesma velocidade nos dois referenciais, ela não respeita a lei de soma de velocidades das transformações de Galileu. Portanto, o principal desafio dos físicos era o de encontrar uma lei de transformação que mantenha as equações de Maxwell invariantes entre dois referenciais inerciais e, portanto, que respeite o princípio da relatividade. Veremos isto no próximo capítulo.

Capítulo 2

Transformações de Lorentz

Uma notação importante a ser definida a partir deste ponto é a notação de soma de Einstein. Definimos que quando um mesmo índice aparecer em uma expressão, uma soma naquele índice estará implícita. Por exemplo, o produto escalar entre dois vetores em \mathbb{R}^3 no espaço euclidiano pode ser escrito na seguinte forma:

$$\vec{A} \cdot \vec{B} = A_1 B_1 + A_2 B_2 + A_3 B_3 = A_i B_i, \quad (2.1)$$

em que A_1, A_2, A_3 são as componentes do vetor \vec{A} e B_1, B_2, B_3 são as componentes do vetor \vec{B} .

2.1 Dedução das transformações

A dedução das transformações de Lorentz aqui apresentadas estão baseadas na referencia [9].

Vamos supor uma configuração entre dois referenciais inerciais como na Figura 1.1. Definiremos os vetores auxiliares de quatro componentes da seguinte forma $w_i = (x, y, z, t)$ e $w'_i = (x', y', z', t')$ ¹ em que x', y', z', t' são as medidas feitas por um observador em S' das coordenadas espaciais e do instante de tempo em que um evento ocorre e x, y, z, t são as medidas feitas por um observador em S das coordenadas espaciais e do instante de tempo em que o mesmo evento ocorre. Assim, a primeira pergunta que nos vem a mente é, que tipo de transformação é possível entre as medidas feitas de S para S'?

Essa transformação deve ser linear, pois usando o argumento de homogeneidade do espaço, se a transformação não fosse linear, medidas de comprimentos e intervalos de tempos dependeriam das posições onde elas fossem medidas.

Como exemplo, tome $x' = x^2$. Um comprimento medido por x' seria dado por $\Delta x' = x'_2 - x'_1 = x_2^2 - x_1^2$. Dessa forma, suponha uma barra de comprimento unitário em S. Digamos que suas extremidades estejam em $x_2 = 2$ e $x_1 = 1$, assim, $\Delta x' = 3$. Porém, digamos que as extremidades estejam em $x_2 = 5$ e $x_1 = 4$, assim, $\Delta x' = 9$. Ou seja, este comprimento dependeria da posição onde ele foi medido e isto não pode ser verdade devi

¹Neste capítulo ainda não estamos utilizando a notação usual de quadri vetor $x^\mu = (ct, x, y, z)$, pois o propósito destas seções iniciais é fazer uma construção intuitiva da Teoria da Relatividade Restrita sem o uso de ferramentas matemáticas mais sofisticadas.

a homogeneidade do espaço. Portanto, vamos supor a transformação linear mais geral possível dada por:

$$w'_i = \Lambda_i^j w_j, \quad (2.2)$$

ou em notação matricial

$$\begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} = \begin{pmatrix} \Lambda_1^1 & \Lambda_1^2 & \Lambda_1^3 & \Lambda_1^4 \\ \Lambda_2^1 & \Lambda_2^2 & \Lambda_2^3 & \Lambda_2^4 \\ \Lambda_3^1 & \Lambda_3^2 & \Lambda_3^3 & \Lambda_3^4 \\ \Lambda_4^1 & \Lambda_4^2 & \Lambda_4^3 & \Lambda_4^4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \quad (2.3)$$

com Λ_i^j sendo os coeficientes da matriz de transformação. Lembrando que seus coeficientes não devem depender das posições, porém, podem depender da velocidade relativa entre os referenciais eventualmente.

Para deduzirmos todos os coeficientes da matriz de transformação usaremos alguns argumentos geométricos, de simetria, homogeneidade do espaço e por fim usaremos o princípio da relatividade.

Da Figura 1.1, um ponto de coordenadas $y = 0$ e $z = 0$ está sobre o eixo x . Porém, os eixos x' e x coincidem, isso significa que o ponto considerado também está sobre o eixo x' , conseqüentemente, $y' = 0$ e $z' = 0$. Como conclusão:

$$\text{Se } y = z = 0 \quad \Rightarrow \quad y' = z' = 0. \quad (2.4)$$

Usando este resultado na equação (2.2) obtemos que $\Lambda_2^1 = \Lambda_2^4 = 0$ e $\Lambda_3^1 = \Lambda_3^4 = 0$. Assim, por hora, as transformações para y' e z' ficam escritas como:

$$y' = \Lambda_2^2 y + \Lambda_2^3 z \quad (2.5)$$

$$z' = \Lambda_3^2 y + \Lambda_3^3 z, \quad (2.6)$$

dependendo única e exclusivamente de y e z .

Outro fato geométrico que deve ser analisado é que se um ponto estiver situado em $z = 0$, ele pertence ao plano xy . Porém, este plano é equivalente ao plano $x'y'$, o que significa que ele tem coordenada $z' = 0$. O mesmo raciocínio é válido os planos xz e $x'z'$. Como consequência, obtemos que $\Lambda_2^3 = \Lambda_3^2 = 0$, resultando que:

$$y' = \Lambda_2^2 y \quad (2.7)$$

$$z' = \Lambda_3^3 z. \quad (2.8)$$

Suponha agora que de S é vista uma barra de certo comprimento, em repouso no referencial S' e que está paralela ao eixo y' como mostra a Figura 2.1.

Agora suponha que a mesma barra é colocada em S paralela ao eixo y , mas queremos medi-la vista de S'. Como não há distinção entre as duas situações (pois para S, S' é quem está se movendo mas, para S', é S quem se move), as medidas devem ser exatamente iguais. Usando o mesmo raciocínio para comprimentos paralelos aos eixos z e z' e usando

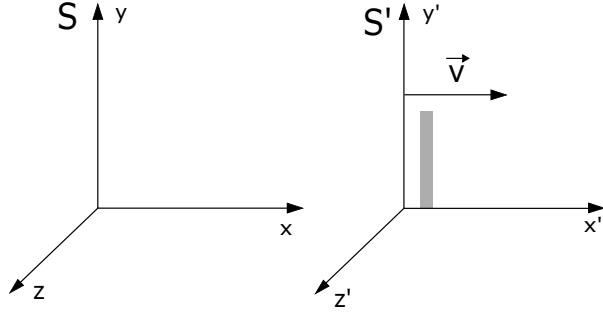


Figura 2.1: Barra em repouso no referencial S' paralela ao eixo y'

as equações (2.7) e (2.8) obtemos:

$$\Lambda_2^2 = \frac{1}{\Lambda_2^2} \Rightarrow \Lambda_2^2 = 1 \Rightarrow y' = y \quad (2.9)$$

$$\Lambda_3^3 = \frac{1}{\Lambda_3^3} \Rightarrow \Lambda_3^3 = 1 \Rightarrow z' = z, \quad (2.10)$$

assim, já obtivemos as transformações das coordenadas y' e z' . Para a transformação de t' , temos:

$$t' = \Lambda_4^1 x + \Lambda_4^2 y + \Lambda_4^3 z + \Lambda_4^4 t, \quad (2.11)$$

note que, se dois relógios sincronizados estivessem situados em (x, y, z) e em $(x, -y, -z)$, ou seja, distribuídos simetricamente, ao serem vistos de S' pareceriam dessincronizados o que não deve ser verdade devido a homogeneidade do espaço. Assim, t' não deve depender de y e z nos dando que $\Lambda_4^2 = \Lambda_4^3 = 0$. Dessa forma, por hora temos:

$$t' = \Lambda_4^1 x + \Lambda_4^4 t. \quad (2.12)$$

Para a transformação de x' temos que:

$$x' = \Lambda_1^1 x + \Lambda_1^2 y + \Lambda_1^3 z + \Lambda_1^4 t. \quad (2.13)$$

Tomemos um ponto em S' de coordenada $x' = 0$. Visto de S, este ponto se move com velocidade v no sentido positivo de x . Assim, a componente x do vetor posição deste ponto visto de S deve ser $x = vt$. Portanto, concluímos que se $x = vt$, então, $x' = 0$. Usando este resultado na equação (2.13) ficamos com:

$$0 = (\Lambda_1^1 v + \Lambda_1^4)t + \Lambda_1^2 y + \Lambda_1^3 z, \quad (2.14)$$

disto, obtemos que

$$\Lambda_1^4 = -\Lambda_1^1 v \quad (2.15)$$

$$\Lambda_1^2 = 0 \quad (2.16)$$

$$\Lambda_1^3 = 0, \quad (2.17)$$

portanto,

$$x' = \Lambda_1^1(x - vt), \quad (2.18)$$

dessa forma, temos o seguinte conjunto de equações para as transformações:

$$\begin{cases} x' = \Lambda_1^1(x - vt) \\ y' = y \\ z' = z \\ t' = \Lambda_4^1 x + \Lambda_4^4 t \end{cases} . \quad (2.19)$$

Vamos usar agora o princípio da constância da velocidade da luz derivado do princípio da relatividade² mostrado no Capítulo 1 para determinar o valor das constantes que ainda restam. Suponha que no instante $t = t' = 0$, instante em que as origens dos sistemas S e S' coincidem, uma onda eletromagnética esférica é produzida na origem dos dois referenciais. Como ela deve ter a mesma velocidade em ambos os referenciais, as equações para a frente de onda em cada referencial são:

$$x^2 + y^2 + z^2 = c^2 t^2 \quad (2.20)$$

$$x'^2 + y'^2 + z'^2 = c^2 t'^2 , \quad (2.21)$$

substituindo as leis de transformações (2.19) na equação (2.21) obtemos:

$$[(\Lambda_1^1)^2 - c^2(\Lambda_4^1)^2] x^2 - 2 [(\Lambda_1^1)^2 v + c^2 \Lambda_4^1 \Lambda_4^4] xt + y^2 + z^2 = [(\Lambda_4^4)^2 c^2 - (\Lambda_1^1)^2 v^2] t^2 . \quad (2.22)$$

Comparando a equação (2.22) com a equação (2.20) obtemos o seguinte sistema de equações para os Λ 's:

$$\begin{cases} (\Lambda_4^4)^2 c^2 - (\Lambda_1^1)^2 v^2 = c^2 \\ (\Lambda_1^1)^2 v + c^2 \Lambda_4^1 \Lambda_4^4 = 0 \\ (\Lambda_1^1)^2 - c^2 (\Lambda_4^1)^2 = 1 \end{cases} , \quad (2.23)$$

cuja solução é:

$$\Lambda_1^1 = \Lambda_4^4 = \frac{1}{\sqrt{1 - v^2/c^2}}; \quad \Lambda_4^1 = -\frac{v}{c^2} \frac{1}{\sqrt{1 - v^2/c^2}} . \quad (2.24)$$

Definindo o parâmetro γ como

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} , \quad (2.25)$$

chegamos às equações das transformações de Lorentz:

$$\begin{cases} x' = \gamma(x - vt) \\ y' = y \\ z' = z \\ t' = \gamma(t - vx/c^2) \end{cases} . \quad (2.26)$$

²Lembre-se, este princípio foi postulado devido aos resultados dos experimentos da época que não conseguiram medir a velocidade relativa entre a Terra e o éter. Dessa forma, uma consequência da imposição de que as equações de Maxwell são as mesmas em qualquer referencial inercial é que a velocidade da luz no vácuo é mesma para todos estes referenciais.

Da definição do parâmetro γ podemos observar que, como ele deve ser um número real, $v < c$. Outra propriedade é que, pela definição, pode-se concluir que $\gamma \geq 1$, já que o denominador é maior do que 0 e menor do que 1. Note que no limite $c \rightarrow \infty$ retornamos às Transformações de Galileu (equação (1.1)).

2.2 Consequências cinemáticas

As principais consequências cinemáticas já podem ser derivadas diretamente das transformações de Lorentz (equações (2.26)). A primeira delas é que comprimentos que estão perpendiculares à direção definida pela velocidade entre os referenciais não são afetados, pois essas medidas são iguais nos dois referenciais.

Tomemos agora o intervalo de tempo entre dois eventos³ medidos no referencial S' submetidos às transformações de Lorentz:

$$\Delta t' = t'_2 - t'_1 = \gamma(t_2 - t_1) - \frac{\gamma v}{c^2}(x_2 - x_1) = \gamma \Delta t - \frac{\gamma v}{c^2} \Delta x. \quad (2.27)$$

Suponha que observado de S os eventos medidos são simultâneos⁴, ou seja, $\Delta t = 0$. Assim, $\Delta t' = -\frac{\gamma v}{c^2} \Delta x$, o que significa que se dois eventos são simultâneos em um determinado referencial, em geral, eles não são simultâneos em um referencial que se move em relação ao primeiro. Este fenômeno é conhecido como relatividade da simultaneidade.

Agora, vamos supor que visto de S os eventos não são simultâneos mas, são medidos numa mesma posição⁵, ou seja, $\Delta x = 0$. Assim, $\Delta t' = \gamma \Delta t$, o que significa que o intervalo de tempo medido em S' sofre uma dilatação em relação ao intervalo medido em S, visto que $\gamma \geq 1$. Este fenômeno é conhecido como dilatação temporal.

Quando dois eventos são medidos no mesmo lugar em um determinado referencial, este é chamado de próprio e o intervalo de tempo medido neste referencial é dito tempo próprio. Podemos, então, concluir que o intervalo de tempo medido em um referencial próprio é o menor possível.

Tomemos agora, o intervalo espacial na direção x' entre dois eventos medidos no referencial S':

$$\Delta x' = x'_2 - x'_1 = \gamma(x_2 - x_1) - \gamma v(t_2 - t_1) = \gamma \Delta x - \gamma v \Delta t. \quad (2.28)$$

Suponha agora que em S' há uma barra em repouso de comprimento d' , medido por um observador em S', posicionada na horizontal na direção de x' como mostrado na figura 2.2.

Para medirmos o comprimento dessa barra em S, devemos medir o tempo que a barra leva para atravessar um determinado ponto x de S. Com este procedimento em mente, o comprimento medido em S é $d = -v \Delta t$.

³Eventos são definidos como um acontecimento físico que ocorre em um instante de tempo e em uma posição bem definidos.

⁴Por exemplo, a queda simultânea de dois raios no solo.

⁵Por exemplo, duas palmas consecutivas.

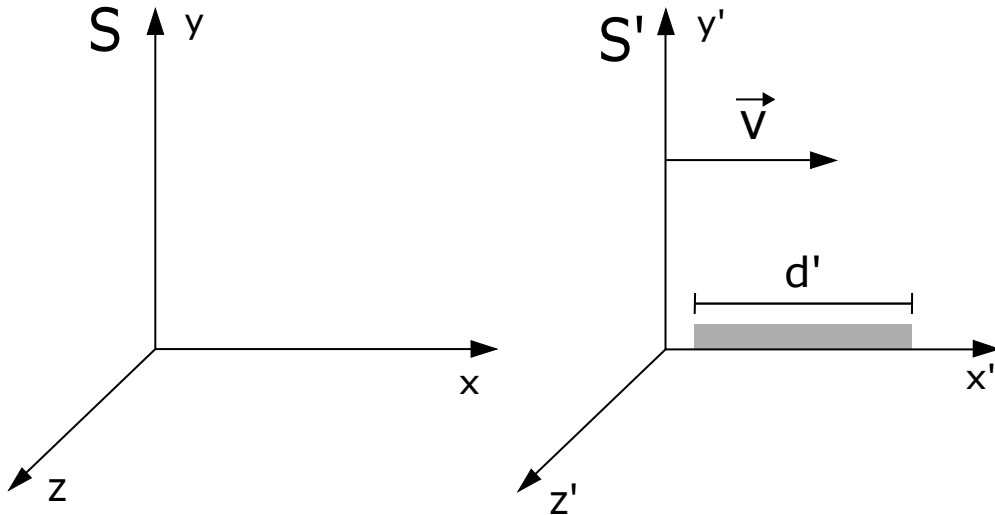


Figura 2.2: Barra de comprimento d' medido por um observador em repouso em S' posicionada na horizontal sobre o eixo x'

Observe que, nesta situação, S é o referencial próprio pois, ele irá medir o intervalo de tempo entre dois eventos que ocorrem na mesma posição, portanto $\Delta x = 0$. Assim, da equação (2.28), temos que, $d' = -\gamma v \Delta t = \gamma d$, ou seja, $d = d'/\gamma$. Portanto, há uma contração na medida da barra quando ela é vista em movimento em um determinado referencial em relação à medida dela quando vista em repouso. Este fenômeno é conhecido como contração espacial.

O leitor pode estar se perguntando se essas consequências cinemáticas altamente contra intuitivas ocorrem na natureza. A resposta é sim e um exemplo de prova experimental da veracidade das transformações de Lorentz é a detecção de partículas instáveis chamadas de Múons na superfície da Terra.

Os Múons são criados quando partículas vindas dos Sol colidem com as moléculas da atmosfera da Terra à uma altura de aproximadamente 60 Km da superfície. O tempo de meia vida⁶ dessa partícula, medido no referencial laboratório em que ela está em repouso, é por volta de $1,5 \mu s$. Dessa forma, caso os Múons viajassem a velocidade da luz até a superfície da Terra, o que não é verdade, se passariam aproximadamente 130 meias vidas até o caminho à superfície. Logo, não deveria-se encontrar a quantidade de mésons detectadas.

A única explicação possível para esta anomalia é que, devido as altas velocidades dos Mésons, o tempo de meia vida deles é dilatado para um referencial em repouso na superfície da Terra enquanto eles percorrem toda a atmosfera (60 Km) e, para um referencial que caminha junto com o Méson, o tempo de meia vida é o mesmo ($1,5 \mu s$) porém há uma contração no comprimento da atmosfera neste referencial. Dessa forma, a quantidade de méson observadas na superfície da Terra só pode ser explicada se admitirmos que as transformações de Lorentz estão corretas e, portanto, os fenômenos de dilatação temporal e contração espacial sejam reais.

Veremos agora como se relacionam as medidas de velocidade de um corpo em movimento entre dois referenciais inerciais em movimento relativo como na Figura 1.1.

⁶Tempo de meia vida é o tempo que uma determinada amostra leva para que metade dela se desintegre.

Suponha que de S' observamos um objeto em movimento. Sua velocidade sobre o eixo x' é $u'_x = \Delta x' / \Delta t'$. Assim:

$$\begin{aligned} u'_x &= \frac{\Delta x'}{\Delta t'} \\ u'_x &= \frac{\gamma \Delta x - \gamma v \Delta t}{\gamma \Delta t - \frac{\gamma v}{c^2} \Delta x} \\ u'_x &= \frac{\gamma \Delta x (1 - v \frac{\Delta t}{\Delta x})}{\gamma \Delta t (1 - \frac{v}{c^2} \frac{\Delta x}{\Delta t})}, \end{aligned}$$

mas, $\Delta x / \Delta t$ é a velocidade u_x deste objeto visto de S. Dessa forma:

$$u'_x = \frac{u_x - v}{1 - u_x v / c^2}. \quad (2.29)$$

A transformação inversa é obtida apenas trocando o sinal de v , já que para S, S' se move com velocidade v no sentido positivo de x e para S', S se move com velocidade v no sentido negativo de x' .

$$u_x = \frac{u'_x + v}{1 + u'_x v / c^2}. \quad (2.30)$$

Esta é a lei de transformação de velocidades relativística que retorna ao caso galileano quando $c \rightarrow \infty$. Faremos um exemplo usando esta lei de transformação para entender a sua diferença em relação a lei de soma de velocidades das transformações de Galileu.

Suponha que de S' é vista uma partícula viajando com velocidade $0,6c$ e que S' se mova com velocidade $0,8c$ relativamente a S. De acordo com a cinemática de Galileu a velocidade da partícula vista de S seria $u_x = u'_x + v = 0,6c + 0,8c = 1,4c$. Neste cenário, a partícula viajaria mais rápido do que a luz vista de S.

Usando a transformação de velocidades relativística para a situação descrita acima, o valor que obtemos é $u_x \approx 0,94c$, valor abaixo da velocidade da luz. Pode-se mostrar que isso é sempre verdade para a transformação relativística (equação (2.30)).

2.3 Transformações de Lorentz entre referenciais com velocidade relativa numa direção arbitrária sem rotação relativa

Consideremos agora que o referencial S' se move não mais na direção de x , mas em uma direção arbitrária definida pelo vetor velocidade \vec{v} . Usaremos argumentos para a dedução dessas Transformações de Lorentz baseados em [6].

Escrevamos o vetor posição \vec{r}' :

$$\vec{r}' = x' \hat{x} + y' \hat{y} + z' \hat{z},$$

substituindo as equações (2.26) obtemos:

$$\vec{r}' = (\gamma x - \gamma vt)\hat{x} + y\hat{y} + z\hat{z},$$

que pode ser escrito na forma:

$$\vec{r}' = \vec{r} + x(\gamma - 1)\hat{x} - \gamma vt\hat{x}, \quad (2.31)$$

em que $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$. Sabemos que neste caso $\vec{v} = v\hat{x}$, então:

$$\hat{x} = \frac{\vec{v}}{v}$$

$$x = \frac{\vec{v} \cdot \vec{r}}{v},$$

substituindo esses valores na equação (2.31) e na parte temporal de (2.26) e definindo o fator $\vec{\beta} = \vec{v}/c$, obtemos

$$\begin{cases} \vec{r}' = \vec{r} + (\gamma - 1)\frac{\vec{\beta} \cdot \vec{r}}{\beta^2}\vec{\beta} - \gamma c\vec{\beta}t \\ t' = \gamma\left(t - \frac{\vec{\beta} \cdot \vec{r}}{c}\right) \end{cases}. \quad (2.32)$$

Observe que as transformações (2.32) não tornam as equações (2.26) inválidas, pois podemos simplesmente fazer uma rotação nos eixos de S e S' de tal forma que os eixos x e x' coincidam com a direção de \vec{v} . Para isso, podemos simplesmente fazer $\vec{\beta} = \frac{v}{c}\hat{x}$ retornando as transformações (2.26).

2.4 Geometria da Relatividade Restrita

Como vimos nas seções anteriores, espaço e tempo não são grandezas absolutas, ou seja, não tem o mesmo valor em referenciais que se movem um em relação ao outro. Dessa forma, o conceito de distância espacial euclidiana e distância temporal dada na mecânica de Galileu perde seu sentido no contexto da Relatividade Restrita. Fica então a pergunta, como definir uma distância entre dois eventos? A resposta vem de uma combinação entre o espaço e o tempo.

2.4.1 Distância na mecânica newtoniana

Para a mecânica newtoniana, que utiliza o espaço Euclidiano, a distância d' entre dois pontos distintos do espaço de coordenadas (x'_1, y'_1, z'_1) e (x'_2, y'_2, z'_2) em um determinado referencial S' é dada pelo teorema de Pitágoras:

$$d' = \sqrt{(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2}. \quad (2.33)$$

Suponha agora que estejamos em um referencial S que se move com relação a S' com uma velocidade \vec{v} como mostra a Figura 1.1. Usando as transformações de Galileu (equações (1.1)), concluímos que a distância d medida entre estes mesmos dois pontos, mas vistos de S é exatamente igual a d' . Podemos, então, concluir que a distância ente

dois pontos na mecânica newtoniana é um invariante, isto é, tem o mesmo valor independente do referencial escolhido, ou seja, a partir da definição de distância e usando as transformações de Galileu determinou-se um invariante, o comprimento espacial.

Sabemos também que na mecânica newtoniana o tempo passa igualmente para todos os referenciais. Dessa forma, o intervalo de tempo $\Delta t = \Delta t'$ entre dois eventos é um invariante. Então, se somarmos d' e $\Delta t'$ formando uma quantidade $r' = d' + \Delta t'$ também o resultado é um invariante.

2.4.2 Distância na relatividade

Para a relatividade restrita iremos fazer o caminho contrário, vamos encontrar uma quantidade invariante sob as transformações de Lorentz e, com ela, definir a distância entre dois eventos⁷ no contexto desta teoria.

Suponha que estejamos em um referencial S' e observamos dois eventos, de coordenadas espaciais (x', y', z') e coordenada temporal t' e outro evento de coordenadas espaciais infinitesimalmente próximas $(x' + dx', y' + dy', z' + dz')$ e temporal $t' + dt'$. A diferença entre as coordenadas espaciais é o vetor (dx', dy', dz') e a diferença entre os instantes de tempo vale dt' . Agora, tomemos a seguinte quantidade:

$$ds'^2 = c^2 dt'^2 - dx'^2 - dy'^2 - dz'^2, \quad (2.34)$$

usando as transformações de Lorentz (equações (2.26)), obtemos

$$\begin{aligned} dx' &= \gamma(dx - vdt) \\ dy' &= dy \\ dz' &= dz \\ dt' &= \gamma\left(dt - \frac{v}{c^2}dx\right), \end{aligned}$$

substituindo esses valores na equação (2.34) obtemos

$$ds'^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = ds^2. \quad (2.35)$$

Podemos observar então que esta quantidade é um invariante sobre as transformações de Lorentz (2.26). Usaremos, então, ds^2 como a definição de distância entre dois eventos no contexto da Relatividade Restrita. Note que, se $ds^2 = ds'^2 = 0$ obtemos as equações (2.20) e (2.21), que representam a equação de movimento de um pulso luminoso que anda com velocidade c em qualquer referencial. Portanto, esta definição é coerente com o princípio da relatividade, informação que usaremos mais adiante.

Veja que nem espaço e nem tempo sozinhos são invariantes, apenas uma particular combinação entre eles. Isso sugere que espaço e tempo estão conectados e não podem ser tratados de forma separada, mas sim, como componentes de um ente físico que possui um significado concreto, o espaço-tempo.

⁷Neste trabalho tratamos um ponto no espaço-tempo como sendo equivalente as coordenadas espaciais e temporal em que um evento ocorre.

Podemos então considerar que “moramos” em um espaço de 4 dimensões em que uma delas é temporal e as outras três são espaciais e a distância entre dois eventos neste espaço é representada pela equação (2.34). Neste espaço, definiremos as coordenadas espaço-temporais de um evento, chamado de quadrivetor posição, por (x^0, x^1, x^2, x^3) cujas componentes x^μ (aqui usaremos as componentes contravariantes do vetor, ver o Apêndice A), são $x^0 = ct$, $x^1 = x$, $x^2 = y$ e $x^3 = z$ (perceba a mudança de notação em relação a seção 2.1)⁸. Assim, a distância neste espaço fica definida por:

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2, \quad (2.36)$$

ou ainda, na notação de soma de Einstein:

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \quad (2.37)$$

com μ e ν variando de 0 a 3. A partir de agora, quando usarmos índices gregos eles sempre irão variar de 0 a 3 e quando usarmos índices latinos, eles irão variar de 1 a 3.

O objeto $\eta_{\mu\nu}$, na equação (2.37) é chamado de tensor métrico pois define a noção de distância no espaço. No caso da Relatividade Restrita, quando usamos as coordenadas oblíquas quadridimensionais ct, x, y, z , este tensor, representado na forma de uma matriz, tem a forma:

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (2.38)$$

Algumas vezes é também comum representar a matriz η por $\eta = \text{diag}(1, -1, -1, -1)$.

Este espaço quadridimensional munido desta métrica é conhecido como espaço de Minkowski e o tensor métrico (equação (2.38)), define a chamada métrica de Minkowski.

Algumas propriedades importantes de serem notadas desta matriz é que $\det(\eta) = -1$ e $\eta^2 = \eta \cdot \eta = \mathbb{I}$ em que \mathbb{I} é a matriz identidade, ou seja, a inversa de η é ela mesma. Na notação de tensores, esta última propriedade fica:

$$\eta_{\mu\beta} \eta^{\beta\nu} = \delta_\mu^\nu, \quad (2.39)$$

em que δ_μ^ν é a função delta de Kronecker que representa os coeficientes da matriz identidade em quatro dimensões que vale 1 se $\mu = \nu$ e 0 se $\mu \neq \nu$.

2.5 Matriz de transformação e transformação geral de Lorentz

Usando a notação de quadrivetor, podemos escrever as transformações de Lorentz como sendo matrizes Λ , cujas componentes são Λ_ν^μ , pois basta notar que podemos escrever as leis de transformação como:

⁸Fizemos essa mudança de notação, pois achamos que este é o momento certo de introduzir a matemática dos tensores que nos ajudará no desenvolvimento da Teoria da Relatividade Restrita.

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}. \quad (2.40)$$

Para o caso das transformações onde a velocidade relativa entre os referenciais está na direção de x , dada pela equação (2.26), esta matriz fica:

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.41)$$

Já para o caso das transformações com velocidade em uma direção arbitrária, dadas pelas equações (2.32), as componentes da matriz de transformação, extensa demais para ser escrita explicitamente, são:

$$\Lambda_0^0 = \gamma \quad (2.42)$$

$$\Lambda_i^0 = -\gamma\beta_i \quad (2.43)$$

$$\Lambda_0^i = -\gamma\beta^i \quad (2.44)$$

$$\Lambda_j^i = \delta_j^i + (\gamma - 1) \frac{\beta^i \beta_j}{(\beta)^2} \quad (2.45)$$

Dados estes exemplos vem a pergunta: qual é a transformação de Lorentz mais geral possível? Pois, a princípio, podemos primeiro fazer uma rotação relativa entre os referenciais e depois movê-los em uma direção arbitrária. Esta é uma situação candidata a respeitar o princípio da relatividade e, assim, gerar uma transformação de Lorentz. A resposta vem do invariante ds^2 , como comentado anteriormente, ds^2 , por si só, já respeita o princípio da relatividade. Assim, definimos que as transformações de Lorentz mais gerais possíveis são transformações de coordenadas que mantêm ds^2 invariante. Segue então a seguinte definição:

Definição 1: uma transformação que mantém a quantidade ds^2 invariante, definida pela equação (2.34), é chamada de Transformação de Lorentz.

Vejamos qual a equação que essa transformação deve satisfazer.

$$ds'^2 = ds^2 \quad (2.46)$$

$$\begin{aligned} \eta_{\mu\nu} dx'^{\mu} dx'^{\nu} &= \eta_{\mu\nu} dx^{\mu} dx^{\nu} \\ \eta_{\mu\nu} \Lambda_{\alpha}^{\mu} dx^{\alpha} \Lambda_{\beta}^{\nu} dx^{\beta} &= \eta_{\alpha\beta} dx^{\alpha} dx^{\beta} \\ \Lambda_{\alpha}^{\mu} \eta_{\mu\nu} \Lambda_{\beta}^{\nu} &= \eta_{\alpha\beta}, \end{aligned} \quad (2.47)$$

usando notação matricial temos:

$$\Lambda^T \eta \Lambda = \eta, \quad (2.48)$$

em que Λ^T é a matriz transposta de Λ . Multiplicando ambos os lados da equação (2.48) a esquerda por G e a direita pela inversa de Λ , denotada por Λ^{-1} , e usando as propriedades já mencionadas de G , obtemos:

$$\eta \Lambda^T \eta = \Lambda^{-1}. \quad (2.49)$$

Outra manipulação possível é multiplicar a equação (2.48) a direita por G obtendo:

$$\Lambda^T \eta \Lambda \eta = \mathbb{I}, \quad (2.50)$$

portanto,

$$\Lambda^T \eta = (\Lambda \eta)^{-1}, \quad (2.51)$$

então,

$$\Lambda \eta \Lambda^T \eta = \mathbb{I}, \quad (2.52)$$

assim,

$$\Lambda \eta \Lambda^T = \eta. \quad (2.53)$$

Todas as equações: (2.47), (2.48), (2.49) e (2.53) são equivalentes e representam a equação que a matriz de transformação Λ deve respeitar para ser uma transformação de Lorentz e, portanto, respeitar o princípio da relatividade. Uma outra forma de interpretar essas equações é que a matriz Λ é aquela que deixa a métrica do espaço invariante por uma transformação de isometria.

A equação (2.49) nos mostra também que a matriz Λ é pseudo-ortogonal por ter uma correção na relação entre a sua transposta e sua inversa (matrizes ortogonais são aquelas em que a sua inversa é igual a sua transposta).

Usando a interpretação de tensor para $\eta_{\mu\nu}$ (Apêndice A), a equação (2.47) nos mostra que o tensor métrico é invariante por uma transformação entre referenciais inerciais, ou seja, ele tem a mesma forma em qualquer referencial inercial. Este fato é muito importante e será usado no Capítulo 4.

Capítulo 3

Grupo de Lorentz

Neste capítulo iremos estudar algumas propriedades das transformações de Lorentz baseadas na análise de [6]. Desejamos aqui mostrar que as transformações de Lorentz formam um grupo. O uso de teorias de Grupos neste trabalho é importante para dar uma nova interpretação para as transformações de Lorentz. Mostramos que elas podem ser vistas como rotações em um espaço hiperbólico que deixam a métrica invariante. Mas, antes de provar esta propriedade, apresentemos a definição de grupos.

3.1 Grupos

Suponha um conjunto \mathfrak{Y} qualquer não vazio. Suponha também que exista uma operação simbolizada por \bullet que opera entre dois elementos quaisquer deste conjunto tal que o resultado desta operação também esteja no conjunto (esse tipo de operação é chamada de binária). Ou seja, se $a, b \in \mathfrak{Y} \Rightarrow a \bullet b = c \in \mathfrak{Y}$. Além disso, suponha que esta operação seja associativa, isto é, para todo $a, b, c \in \mathfrak{Y}$ temos que

$$a \bullet (b \bullet c) = (a \bullet b) \bullet c. \quad (3.1)$$

Suponha também que exista um elemento neutro dentro deste conjunto, denotado por e , que quando operado sobre qualquer outro elemento do conjunto o resultado é este mesmo elemento, ou seja

$$e \bullet a = a \bullet e = a. \quad (3.2)$$

Finalmente, tome um elemento a que pertence ao conjunto e suponha que exista um elemento chamado de inverso e denotado por a^{-1} que quando operado com a resulte no elemento neutro, isto é

$$a \bullet a^{-1} = a^{-1} \bullet a = e. \quad (3.3)$$

Se essas condições forem satisfeitas, dizemos que o conjunto \mathfrak{Y} munido da operação binária \bullet é um grupo. Caso a operação seja comutativa, isto é, $a \bullet b = b \bullet a$, para todo $a, b \in \mathfrak{Y}$ dizemos que o grupo é um grupo abeliano. Caso seja não comutativo o grupo é denominado de não abeliano.

Podemos mostrar que se e é elemento neutro do grupo \mathfrak{Y} então ele é único. Para isso, suponha que haja um outro elemento \bar{e} que respeita as propriedades de elemento

neutro do grupo. Então,

$$\begin{cases} e \bullet a = a \\ \bar{e} \bullet a = a \end{cases},$$

para todo $a \in \mathfrak{M}$, assim,

$$\begin{aligned} e \bullet \bar{e} &= e \\ e \bullet \bar{e} &= \bar{e}, \end{aligned}$$

então,

$$e = \bar{e}. \tag{3.4}$$

Podemos mostrar que o elemento inverso também é único. Suponha que haja um outro elemento inverso de a dado por \bar{a}^{-1} . Então,

$$\begin{aligned} a \bullet a^{-1} &= e \\ a \bullet \bar{a}^{-1} &= e, \end{aligned}$$

portanto,

$$\begin{aligned} a \bullet a^{-1} &= a \bullet \bar{a}^{-1} \\ a^{-1} \bullet (a \bullet a^{-1}) &= a^{-1} \bullet (a \bullet \bar{a}^{-1}) \\ (a^{-1} \bullet a) \bullet a^{-1} &= (a^{-1} \bullet a) \bullet \bar{a}^{-1} \\ e \bullet a^{-1} &= e \bullet \bar{a}^{-1} \\ a^{-1} &= \bar{a}^{-1}. \end{aligned}$$

Um exemplo de um conjunto que tem a estrutura de grupo é o conjunto das matrizes reais $n \times n$ inversíveis sobre o produto usual das matrizes. Pelo produto ser o usual das matrizes, isso já nos garante que o conjunto é fechado e o produto é associativo. O elemento neutro é a matriz identidade \mathbb{I} e o elemento inverso, por definição está no grupo e pode ser calculado. Nas próximas seções iremos omitir o símbolo \bullet para não carregar a notação. Assim, quando dois elementos aparecerem um ao lado do outro o produto estará implícito.

O estudo de teoria de grupos é muito vasto e rico com grandes aplicações na Física, principalmente em mecânica quântica. Aqui só foram apresentados a sua definição e simples propriedades que são o suficiente para o estudo das transformações de Lorentz. Para leitores que se interessarem em ver com mais profundidade a teoria de grupos recomendamos ver [5].

3.2 Matrizes de Lorentz e estrutura de grupo

Tomaremos nesta seção a definição do conjunto de todas as transformações de Lorentz dada por $\Lambda G \Lambda^T = G$, equação (2.53). Desejamos mostrar que elas formam uma estrutura de grupo com a operação binária sendo o produto usual de matrizes. Assim, trivialmente, já mostramos que a operação é associativa, pois o produto entre matrizes o é.

Provaremos agora que a operação é fechada. Sejam Λ e $\bar{\Lambda}$ dois elementos que pertençam ao conjunto das transformações de Lorentz. Partindo da definição e sabendo que $\Lambda^T \bar{\Lambda}^T = (\bar{\Lambda} \Lambda)^T$, temos que

$$\begin{aligned}\Lambda \eta \Lambda^T &= \eta \\ \bar{\Lambda} \Lambda \eta \Lambda^T \bar{\Lambda}^T &= \bar{\Lambda} \eta \bar{\Lambda}^T \\ \bar{\Lambda} \Lambda \eta (\bar{\Lambda} \Lambda)^T &= \eta,\end{aligned}$$

portanto, se vê que o produto $\bar{\Lambda} \Lambda$ pertence ao conjunto, pois respeita a definição dada pela equação (2.53).

Mostraremos agora a existência do elemento neutro e . Um candidato natural é a matriz identidade \mathbb{I} já que $\mathbb{I} \Lambda = \Lambda \mathbb{I} = \Lambda$, então trivialmente

$$\mathbb{I} \eta \mathbb{I}^T = \eta \mathbb{I} = \eta, \quad (3.5)$$

assim, \mathbb{I} pertence ao conjunto, pois respeita a definição (equação (2.53)).

Vejam agora se o elemento inverso Λ^{-1} , que satisfaz $\Lambda \Lambda^{-1} = \Lambda^{-1} \Lambda = \mathbb{I}$, pertence ao conjunto. Partindo da definição do conjunto e sabendo que $\eta = \eta^{-1}$ e que $(\Lambda^{-1})^T = (\Lambda^T)^{-1}$, temos que

$$\begin{aligned}\Lambda \eta \Lambda^T &= \eta \\ (\Lambda \eta \Lambda^T)^{-1} &= \eta^{-1} \\ \Lambda^{-1} \eta^{-1} (\Lambda^T)^{-1} &= \eta^{-1} \\ \Lambda^{-1} \eta (\Lambda^T)^{-1} &= \eta \\ \Lambda^{-1} \eta (\Lambda^{-1})^T &= \eta,\end{aligned}$$

dessa forma, vemos que Λ^{-1} respeita a definição (equação (2.53)), portanto, pertence ao conjunto. Logo, o conjunto das transformações de Lorentz munido da operação usual de matrizes forma um grupo que será denominado de \mathfrak{L} .

3.3 Propriedades do grupo de Lorentz

Da definição do grupo de Lorentz (equação (2.53)) e sabendo que $\det(\Lambda^T) = \det(\Lambda)$ temos que:

$$\begin{aligned}\Lambda \eta \Lambda^T &= \eta \\ \det(\Lambda \eta \Lambda^T) &= \det(\eta) \\ \det(\Lambda) \det(\eta) \det(\Lambda^T) &= \det(\eta) \\ \det(\Lambda) (-1) \det(\Lambda) &= -1 \\ \det^2(\Lambda) &= 1 \\ \det(\Lambda) &= \pm 1,\end{aligned}$$

quando $\det(\Lambda) = +1$ a transformação de Lorentz é dita própria, caso contrário, ela é dita imprópria.

Um exemplo de transformação de Lorentz de determinante $+1$ é dada pela equação (2.26). Um exemplo de transformação de Lorentz de determinante -1 é a inversão de eixos conhecida como transformação de paridade que tem a seguinte forma[11]:

$$\begin{cases} x' = -x \\ y' = -y \\ z' = -z \\ t' = t \end{cases} . \quad (3.6)$$

Usando a notação de tensor para a definição de Λ , equação (2.47), e usando o caso particular quando $\alpha = \beta = 0$, temos

$$\begin{aligned} \Lambda_0^\mu \eta_{\mu\nu} \Lambda_0^\nu &= \eta_{00} \\ (\Lambda_0^0)^2 - (\Lambda_0^1)^2 - (\Lambda_0^2)^2 - (\Lambda_0^3)^2 &= 1 \\ (\Lambda_0^0)^2 &= 1 + (\Lambda_0^1)^2 + (\Lambda_0^2)^2 + (\Lambda_0^3)^2, \end{aligned} \quad (3.7)$$

dessa forma,

$$\begin{aligned} (\Lambda_0^0)^2 &\geq 1 \\ \Lambda_0^0 &\geq 1 \quad \text{ou} \quad \Lambda_0^0 \leq -1. \end{aligned}$$

Se $\Lambda_0^0 \geq 1$ a transformação de Lorentz é dita ortócrona e ela preserva o sentido do tempo. Caso $\Lambda_0^0 \leq -1$, a transformação de Lorentz é dita não-ortócrona e ela inverte o sentido do tempo.

Quando a transformação é própria e ortócrona ela é chamada de restrita. Para simplificar os tipos de transformação usaremos a seguinte notação:

$$\begin{aligned} \mathfrak{L}_+^\uparrow : \quad & \det(\Lambda) = +1 \quad \text{e} \quad \Lambda_0^0 \geq +1 \\ \mathfrak{L}_+^\downarrow : \quad & \det(\Lambda) = +1 \quad \text{e} \quad \Lambda_0^0 \leq -1 \\ \mathfrak{L}_-^\uparrow : \quad & \det(\Lambda) = -1 \quad \text{e} \quad \Lambda_0^0 \geq +1 \\ \mathfrak{L}_-^\downarrow : \quad & \det(\Lambda) = -1 \quad \text{e} \quad \Lambda_0^0 \leq -1 \end{aligned}$$

Que são subconjuntos do grupo de Lorentz. Apenas \mathfrak{L}_+^\uparrow forma um subgrupo do grupo de Lorentz chamado de grupo de Lorentz restrito. Para mostrar este fato basta notar que, pelas propriedades deste subconjunto, o elemento neutro, que é a própria matriz identidade, está no subconjunto, ou seja, $e = I \in \mathfrak{L}_+^\uparrow$.

O elemento inverso Λ^{-1} também pertence ao subconjunto pois,

$$\begin{aligned} \Lambda \Lambda^{-1} &= \mathbb{I} \\ \det(\Lambda \Lambda^{-1}) &= \det(\mathbb{I}) \\ \det(\Lambda) \det(\Lambda^{-1}) &= \det(\mathbb{I}) \\ 1 \det(\Lambda^{-1}) &= 1 \\ \det(\Lambda^{-1}) &= 1, \end{aligned}$$

da definição da inversa temos que

$$\begin{aligned} \Lambda_\alpha^0 (\Lambda^{-1})_0^\alpha &= \delta_0^0 = 1 \\ \Lambda_0^0 (\Lambda^{-1})_0^0 + \Lambda_1^0 (\Lambda^{-1})_0^1 + \Lambda_2^0 (\Lambda^{-1})_0^2 + \Lambda_3^0 (\Lambda^{-1})_0^3 &= 1. \end{aligned}$$

Vamos definir os seguintes vetores:

$$\vec{x} = (\Lambda_1^0, \Lambda_2^0, \Lambda_3^0) \quad (3.8)$$

$$\vec{y} = ((\Lambda^{-1})_0^1, (\Lambda^{-1})_0^2, (\Lambda^{-1})_0^3), \quad (3.9)$$

resultando em,

$$\Lambda_0^0 (\Lambda^{-1})_0^0 = 1 - \vec{x} \cdot \vec{y}, \quad (3.10)$$

da equação (3.7) e dos vetores definidos nas equações (3.8) e (3.9), temos que

$$\begin{aligned} (\Lambda_0^0)^2 &= 1 + x^2 \\ [(\Lambda^{-1})_0^0]^2 &= 1 + y^2. \end{aligned}$$

Como já se sabe que $\Lambda_0^0 \geq 1$, resta que $\Lambda_0^0 = \sqrt{1 + x^2}$. Já para $(\Lambda^{-1})_0^0$ não temos certeza. Suponha que seja negativo, assim da equação (3.10)

$$\begin{aligned} -\sqrt{1 + x^2} \sqrt{1 + y^2} &= 1 - \vec{x} \cdot \vec{y} \\ \vec{x} \cdot \vec{y} - 1 &= \sqrt{1 + x^2} \sqrt{1 + y^2}, \end{aligned}$$

porém,

$$\begin{aligned} \sqrt{1 + x^2} \sqrt{1 + y^2} &= \sqrt{1 + y^2 + x^2 + x^2 y^2} \\ &= \sqrt{(1 + xy)^2 + (x - y)^2} \\ &\geq 1 + xy, \end{aligned}$$

assim,

$$\vec{x} \cdot \vec{y} - 1 \geq 1 + xy \quad (3.11)$$

$$\vec{x} \cdot \vec{y} \geq 2 + xy. \quad (3.12)$$

O que é falso. Portanto, $(\Lambda^{-1})_0^0$ deve ser positivo. Dessa forma,

$$(\Lambda^{-1})_0^0 = \sqrt{1 + b^2} \geq 1 \quad (3.13)$$

$$(\Lambda^{-1})_0^0 \geq 1. \quad (3.14)$$

Mostrando que a inversa pertence ao subconjunto por respeitar as propriedades de \mathfrak{L}_+^\dagger .

Falta provar que a operação binária respeita as propriedades do subgrupo \mathfrak{L}_+^\dagger . Suponha $\Lambda, \bar{\Lambda} \in \mathfrak{L}_+^\dagger$. Assim,

$$\begin{aligned} \bar{\bar{\Lambda}} &= \Lambda \bar{\Lambda} \\ \det(\bar{\bar{\Lambda}}) &= \det(\Lambda \bar{\Lambda}) \\ \det(\bar{\bar{\Lambda}}) &= \det(\Lambda) \det(\bar{\Lambda}) \\ \det(\bar{\bar{\Lambda}}) &= 1 \times 1 = 1. \end{aligned}$$

O termo $\bar{\bar{\Lambda}}_0^0$ é igual à

$$\bar{\bar{\Lambda}}_0^0 = \bar{\Lambda}_0^0 \Lambda_0^0 + \bar{\Lambda}_1^0 \Lambda_1^0 + \bar{\Lambda}_2^0 \Lambda_2^0 + \bar{\Lambda}_3^0 \Lambda_3^0. \quad (3.15)$$

Vamos definir os vetores

$$\begin{aligned}\vec{a} &= (\bar{\Lambda}_1^0, \bar{\Lambda}_2^0, \bar{\Lambda}_3^0) \\ \vec{b} &= (\Lambda_0^1, \Lambda_0^2, \Lambda_0^3),\end{aligned}$$

assim, a equação (3.15) fica:

$$\bar{\bar{\Lambda}}_0^0 = \bar{\Lambda}_0^0 \Lambda_0^0 + \vec{a} \cdot \vec{b}, \quad (3.16)$$

mas, de (3.7)

$$\begin{aligned}(\Lambda_0^0)^2 &= 1 + a^2 \\ (\bar{\Lambda}_0^0)^2 &= 1 + b^2,\end{aligned}$$

portanto, a equação (3.16) pode ser escrita na forma:

$$\bar{\bar{\Lambda}}_0^0 = \sqrt{1 + a^2} \sqrt{1 + b^2} + \vec{a} \cdot \vec{b}, \quad (3.17)$$

porém,

$$\begin{aligned}\sqrt{1 + a^2} \sqrt{1 + b^2} &= \sqrt{1 + b^2 + a^2 + a^2 b^2} \\ &= \sqrt{(1 + ab)^2 + (a - b)^2} \\ &\geq 1 + ab,\end{aligned}$$

consequentemente,

$$\bar{\bar{\Lambda}}_0^0 \geq 1 + ab + \vec{a} \cdot \vec{b}.$$

Como $\vec{a} \cdot \vec{b} \geq -ab$, chegamos que

$$\bar{\bar{\Lambda}}_0^0 \geq 1, \quad (3.18)$$

concluindo então que o subconjunto \mathfrak{L}_+^\uparrow é um subgrupo do grupo de Lorentz \mathfrak{L} .

3.4 Interpretação geométrica do grupo de Lorentz

As interpretações físicas das transformações de Lorentz foram dadas nos capítulos anteriores. Nosso objetivo aqui é dar uma interpretação geométrica para tais transformações. Primeiramente estudaremos o espaço euclidiano bidimensional e as transformações de isometria que mantêm sua métrica invariante. Posteriormente, analisaremos o espaço de Minkowski bidimensional fazendo uma analogia com o euclidiano [8].

3.4.1 Espaço euclidiano 2D

Sabemos que neste espaço a métrica é dada pela matriz identidade. Dessa forma, a transformação de isometria por uma matriz \mathbb{S} , a ser descoberta, é:

$$\mathbb{S}^T \mathbb{I} \mathbb{S} = \mathbb{I}. \quad (3.19)$$

Multiplicando a equação (3.19) a direita pela inversa de \mathbb{S} obtemos:

$$\mathbb{S}^T = \mathbb{S}^{-1}, \quad (3.20)$$

portanto, a matriz \mathbb{S} é ortogonal. Além disso, calculando o determinante da equação (3.19) obtemos:

$$\det(\mathbb{S}) = \pm 1. \quad (3.21)$$

Escolhendo o valor $+1$ por simplicidade, ficamos com o conjunto das matrizes de ordem 2, ortogonais e de determinante $+1$ conhecido como grupo $SO(2)$ ¹. Vamos analisar agora, qual a forma mais geral dessas matrizes. Uma matriz arbitrária 2×2 é da forma:

$$\mathbb{S} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (3.22)$$

A equação, em função dos elementos da matriz, que expressa o determinante é:

$$ad - bc = 1. \quad (3.23)$$

Por sua ortogonalidade, ou pela equação (3.19), temos que:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (3.24)$$

assim, obtêm-se as seguintes equações:

$$a^2 + b^2 = 1 \quad (3.25)$$

$$ac - bd = 0 \quad (3.26)$$

$$c^2 + d^2 = 1. \quad (3.27)$$

Manipulando as equações (3.23), (3.25), (3.26), (3.27) obtemos as seguintes relações:

$$a = d \quad (3.28)$$

$$c = -b, \quad (3.29)$$

dessa forma, nossa matriz contém apenas 2 termos independentes a e b . Portanto a matriz fica na forma:

$$\mathbb{S} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \quad (3.30)$$

¹Lê-se S do inglês *Special* significa que o determinante das matrizes do grupo é igual a $+1$, O que significa que as matrizes são ortogonais e 2 significa que são matrizes de ordem 2.

A equação (3.25), nos permite interpretar esses elementos como sendo os pontos sobre uma circunferência de raio 1. Portanto, podemos parametrizar os valores de a e b , sem perda de generalidade, da seguinte forma:

$$a = \cos \theta \tag{3.31}$$

$$b = \sin \theta, \tag{3.32}$$

em que $0 < \theta < 2\pi$. Dessa forma, a matriz pode ser escrita como:

$$\mathbb{S} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \tag{3.33}$$

Esta é exatamente a matriz de rotação do espaço euclidiano bidimensional. Ou seja, uma rotação por um ângulo θ nos eixos das coordenadas mantém a métrica invariante. Com esse interpretação em mãos faremos uma analogia com rotações que preservam a métrica porém no contexto do espaço hiperbólico de Minkowski (1+1)D.

3.4.2 Espaço de Minkowski (1+1)D

Neste espaço, a métrica é dada pela matriz (2.38), porém em apenas duas dimensões ela é da forma²:

$$G = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.34}$$

assim, a transformação de isometria, já vista na equação (2.48), é

$$\Lambda^T G \Lambda = G. \tag{3.35}$$

Vamos novamente supor uma matriz com elementos genéricos e descobrir quais as restrições sobre eles:

$$\Lambda = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{3.36}$$

Do Capítulo 3, já discutimos que o determinante de Λ pode ser ± 1 . Tomemos, como no caso anterior, o valor $+1$. Assim, da definição da matriz Λ , obtemos

$$ad - bc = 1. \tag{3.37}$$

Da relação de isometria, usando termos genéricos para essa matriz, ficamos com:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.38}$$

²Aqui denotamos a matriz que representa a métrica por G para não confundir com a métrica no espaço-tempo de quatro dimensões.

dessa forma, chegamos as seguintes equações:

$$a^2 - c^2 = 1 \quad (3.39)$$

$$ab - cd = 0 \quad (3.40)$$

$$b^2 - d^2 = -1. \quad (3.41)$$

Por meio das equações (3.37), (3.39), (3.40), (3.41) obtemos:

$$b = c \quad (3.42)$$

$$a = d. \quad (3.43)$$

A matriz fica, então

$$\Lambda = \begin{pmatrix} a & c \\ c & a \end{pmatrix}. \quad (3.44)$$

A equação (3.39), nos permite interpretar esses elementos como sendo os pontos sobre uma hipérbole. Portanto podemos reparametrizar estes coeficientes da seguinte forma:

$$a = \cosh \xi \quad (3.45)$$

$$c = \sinh \xi, \quad (3.46)$$

em que ξ pode ser qualquer número real. Dessa forma, a matriz pode ser escrita como:

$$\Lambda = \begin{pmatrix} \cosh \xi & \sinh \xi \\ \sinh \xi & \cosh \xi \end{pmatrix}. \quad (3.47)$$

Em analogia com a matriz de rotação anterior, podemos interpretar essa matriz como sendo uma rotação no espaço hiperbólico de Minkowski. Uma visualização destas rotações pode ser vista na Figura 3.1 em que as interpretamos como se um ponto no espaço-tempo estivesse deslizando sobre uma determinada hipérbole.

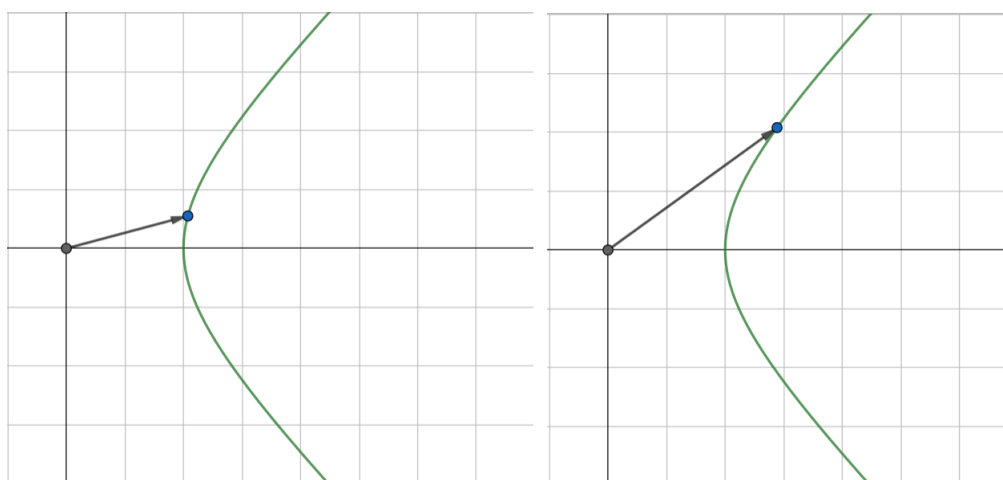


Figura 3.1: Ponto deslizando sobre uma hipérbole

Outra visualização possível das rotações hiperbólicas pode ser vista na Figura 3.2 que mostra como objetos e malhas quadriculadas são distorcidos quando vistos de um referencial que se move com velocidade próximas a da luz. Note que as suas distorções se parecem com hipérboles em total acordo com nossa análise.

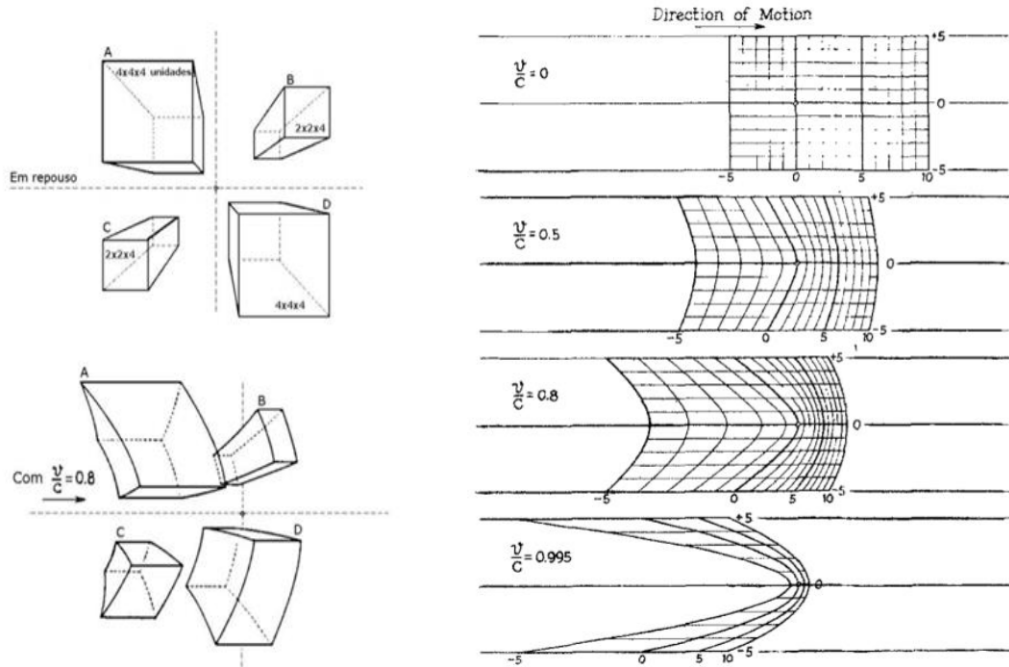


Figura 3.2: Objetos e malhas quadriculadas sendo distorcidos quando vistos de um referencial que se move com velocidades próximas a da luz. Imagem retirada da referencia [3].

Capítulo 4

Dinâmica relativística em quatro dimensões

Neste capítulo, iremos abandonar definitivamente as transformações de Galileu e mostraremos as consequências sob a mecânica ao se admitir as transformações de Lorentz das medidas de espaço e tempo entre referenciais distintos e que a geometria do espaço não é mais euclidiana, mas sim, uma métrica de Minkowski representada pelo tensor métrico (2.38). Iremos escrever a equação de movimento na forma covariante, pois a equação se torna a mesma em qualquer referencial inercial devida a propriedade das transformações dos tensores (Apêndice A).

4.1 Quadrivetores

Assim como na equação (2.40) definiremos como um quadrivetor a quantidade física de quatro componentes $V^\mu = (V^0, V^1, V^2, V^3)$ que se transforma de um referencial para outro por uma transformação de Lorentz.

$$V'^\mu = \Lambda^\mu_\nu V^\nu, \quad (4.1)$$

ou seja, essa quantidade se transforma da mesma forma que as coordenadas. Assim, definimos que um escalar é uma quantidade que é invariante sobre uma transformação de Lorentz. Como exemplo podemos citar a carga de uma partícula, a velocidade da luz no vácuo, entre outros [6].

Como dito anteriormente, queremos definir uma nova formulação para a segunda lei de Newton, porém levando em consideração as transformações de Lorentz. Para isso, precisamos definir as quantidades físicas como velocidade, momento, aceleração e força.

4.1.1 Quadrivelocidade

Suponha que em um referencial inercial S estejamos observando uma partícula se movendo. Poderíamos definir sua velocidade como dx^μ/dt , porém, dt não é um escalar (seu valor depende do referencial observado), assim, devemos usar uma medida de tempo que seja um escalar. Podemos então usar o tempo próprio $d\tau$, ou seja, o tempo no referencial onde

a partícula encontra-se em repouso. Definimos então a quadrivelocidade como

$$U^\mu = \frac{dx^\mu}{d\tau}. \quad (4.2)$$

Pelo intervalo invariante podemos obter uma relação entre essas duas medidas de tempo ($d\tau$ e dt):

$$\begin{aligned} ds'^2 &= ds^2 \\ c^2 d\tau^2 &= c^2 dt^2 - \vec{r} \cdot \vec{r} \\ c^2 d\tau^2 &= c^2 dt^2 - \vec{v} \cdot \vec{v} dt^2 \\ c^2 d\tau^2 &= c^2 \left(1 - \frac{v^2}{c^2}\right) dt^2 \\ \gamma d\tau &= dt. \end{aligned} \quad (4.3)$$

Usando a regra da cadeia, a quadrivelocidade fica:

$$\begin{aligned} U^\mu &= \frac{dx^\mu}{d\tau} \\ U^\mu &= \frac{dx^\mu}{dt} \frac{dt}{d\tau} \\ U^\mu &= \gamma \frac{dx^\mu}{dt}. \end{aligned} \quad (4.4)$$

Suas componentes são $U^\mu = (\gamma c, \gamma \dot{x}, \gamma \dot{y}, \gamma \dot{z}) = (\gamma c, \gamma \vec{v})$. Uma consequência desta definição é que o quadrado da quadrivelocidade é uma constante. Para realizar este cálculo, precisamos definir quais são as componentes covariantes de U (Apêndice A), elas são dadas por $U_\mu = g_{\mu\nu} U^\nu$.

Como discutido na seção 2.5, o tensor métrico tem a mesma forma para todos os referenciais inerciais e é dado por (2.38). Portanto, $U_\mu = (\gamma c, -\gamma \vec{v})$ e o produto $U^\mu U_\mu = U_\mu U^\mu = c^2$.

4.1.2 Quadriaceleração

De maneira análoga definiremos a quadriaceleração como sendo $A^\mu = dU^\mu/d\tau = d^2 x^\mu/d\tau^2$. Usando a regra da cadeia ficamos com $A^\mu = \gamma dU^\mu/dt$. Em termos das componentes a quadriaceleração é dada por:

$$A^\mu = \gamma \left(c \frac{d\gamma}{dt}, \vec{v} \frac{d\gamma}{dt} + \gamma \frac{d\vec{v}}{dt} \right),$$

mas,

$$\frac{d\gamma}{dt} = \gamma^3 \frac{\vec{v}}{c^2} \cdot \frac{d\vec{v}}{dt},$$

fazendo $\vec{a} = d\vec{v}/dt$ a quadriaceleração fica:

$$A^\mu = \left(\gamma^4 \frac{\vec{v} \cdot \vec{a}}{c}, \gamma^4 \frac{\vec{v} \cdot \vec{a}}{c^2} \vec{v} + \gamma^2 \vec{a} \right). \quad (4.5)$$

Uma notável propriedade é que $A^\mu U_\mu = 0$. Ou seja, no espaço de Minkowski os quadrivetores aceleração e velocidade são ortogonais.

4.1.3 Quadrimomento

O quadrimomento é simplesmente definido como $P^\mu = mU^\mu$. Dessa forma, suas componentes são $P^\mu = (\gamma mc, \gamma m\vec{v}) = (\gamma mc, \vec{p})$. Em que $\vec{p} = \gamma m\vec{v}$ é a componente espacial do momento relativístico. Aqui, também obtemos que o módulo do quadrimomento é uma constante e vale $P^\mu P_\mu = m^2 c^2$.

4.2 Equação de movimento

Por consistência a equação de movimento proposta será:

$$\frac{dP^\mu}{d\tau} = \mathcal{F}^\mu, \quad (4.6)$$

em que \mathcal{F}^μ é a quadriforça ou força de Minkowski. Usando novamente a regra da cadeia obtemos:

$$\begin{aligned} \frac{dP^\mu}{d\tau} &= \frac{dP^\mu}{dt} \frac{dt}{d\tau} \\ \gamma \frac{dP^\mu}{dt} &= \mathcal{F}^\mu, \end{aligned} \quad (4.7)$$

do módulo do quadrimomento, derivando ambos os lados em relação a τ , temos que

$$\begin{aligned} P^\mu P_\mu &= m^2 c^2 \\ 2 \frac{dP^\mu}{d\tau} P_\mu &= 0 \\ \mathcal{F}^\mu P_\mu &= 0, \end{aligned} \quad (4.8)$$

dessa forma, \mathcal{F}^μ deve satisfazer as equações (4.6) e (4.8) simultaneamente para ser uma definição consistente.

Para determinar a quadriforça, devemos relacioná-la com a força tridimensional \vec{F} . Para isso, vamos supor que a parte espacial da equação de movimento continua válida com \vec{p} sendo o momento relativístico.

$$\frac{d\vec{p}}{dt} = \frac{d}{dt}(\gamma m\vec{v}) = \vec{F}, \quad (4.9)$$

assim, usando a parte espacial da equação (4.7) obtemos

$$\mathcal{F}^i = \gamma \frac{d\vec{p}}{dt} = \gamma \vec{F}, \quad (4.10)$$

portanto, $\mathcal{F} = (\mathcal{F}^0, \gamma \vec{F})$. Para descobrir a parte temporal, basta usar a equação (4.8). Depois de uma certa álgebra, chega-se que

$$\mathcal{F}^0 = \frac{\gamma}{c} \vec{F} \cdot \vec{v}, \quad (4.11)$$

portanto, nosso quadrivetor força fica:

$$\mathcal{F}^\mu = \left(\frac{\gamma}{c} \vec{F} \cdot \vec{v}, \gamma \vec{F} \right), \quad (4.12)$$

assim, dada uma força espacial \vec{F}^1 basta substituí-la na equação (4.11) e resolve a equação para encontrar como o corpo se move pelo espaço.

Para o caso simples em que a massa é constante no tempo a equação de movimento é dada por:

$$\mathcal{F}^\mu = mA^\mu \quad (4.13)$$

$$\left(\vec{F} \cdot \vec{v}, \vec{F} \right) = \left(m\gamma^3 \vec{a} \cdot \vec{v}, m\gamma^3 \frac{\vec{a} \cdot \vec{v}}{c^2} \vec{v} + m\gamma \vec{a} \right). \quad (4.14)$$

Veremos com um exemplo como este sistema se comporta. Suponha um sistema unidimensional ($y = 0$ e $z = 0$ para todo tempo t) e que sobre uma partícula de massa m constante atue uma força constante $\vec{F} = F\hat{x}$. Substituindo essas informações na equação (4.14) e seja v sua velocidade sobre o eixo x a a sua aceleração sobre o eixo x obtemos:

$$F = \frac{ma}{\sqrt{1 - v^2/c^2}}. \quad (4.15)$$

Como $a = dv/dt$ podemos resolver a equação (4.15) para a velocidade v em função do tempo t encontrando

$$v = \frac{Fct}{\sqrt{m^2c^2 + F^2t^2}}. \quad (4.16)$$

Da equação (4.16) nota-se que quando $t \rightarrow \infty$ a velocidade $v \rightarrow c$, ou seja, não importa por quanto tempo apliquemos uma força sobre a partícula, ela nunca irá ultrapassar a velocidade da luz. A figura 4.1 mostra o comparativo entre a velocidade relativística (em roxo) dada pela equação (4.1) e a velocidade não relativística (em vermelho) que cresce linearmente com o tempo quando uma força constante atua sobre a partícula². Note que para baixas velocidades as duas funções se confundem, o que já era esperado.

Integrando a equação (4.16) podemos encontrar a função da posição da partícula em função do tempo, já que $v = dx/dt$, dada por:

$$x = \frac{c}{F} \left(\sqrt{F^2t^2 + m^2c^2} - mc \right), \quad (4.17)$$

¹Por exemplo uma força elástica $\vec{F} = -kx$ ou uma força de arraste $\vec{F} = -b\dot{x}$.

²Neste gráfico usamos todas as constantes iguais a 1 para simplificar a análise. Dessa forma, para o caso relativístico $v \rightarrow 1$ quando $t \rightarrow \infty$

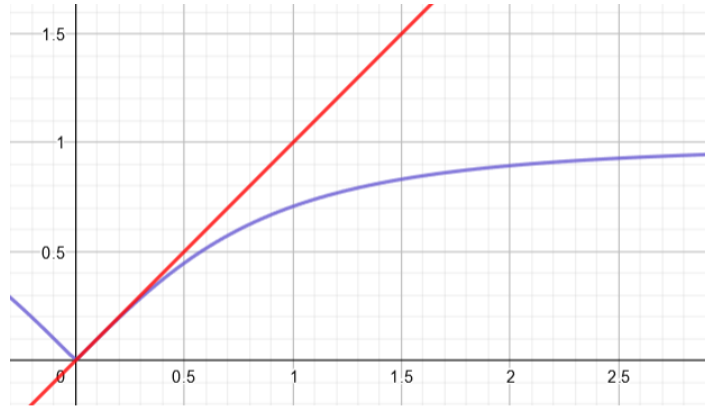


Figura 4.1: Gráfico comparando as velocidade em função do tempo para os caso relativísticos e não relativísticos quando uma força constante atua sobre uma partícula

que pode ser manipulada para ficar na forma de uma equação de hipérbole:

$$\frac{\left(x + \frac{mc^2}{F}\right)^2}{\left(\frac{mc^2}{F}\right)^2} - \frac{t^2}{\left(\frac{mc}{F}\right)^2} = 1. \quad (4.18)$$

A figura 4.2 mostra o comparativo³ entre a função horária da posição relativística (em azul) dada pela equação (4.2) e a função horária da posição não relativística (em laranja) que tem a forma $x = Ft^2/2m$. Note novamente que as funções se confundem para velocidades baixas (tempo t pequeno) como esperado.

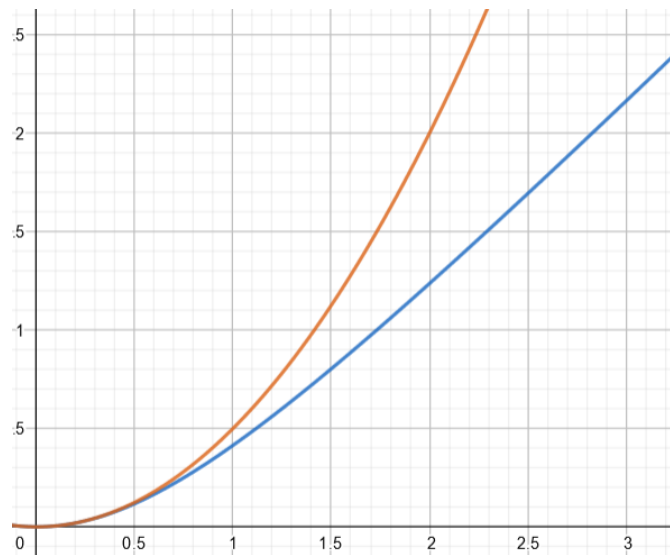


Figura 4.2: Caption

³Novamente com todas as constantes sendo colocadas iguais a 1 para a plotagem do gráfico.

4.3 Energia relativística e partículas sem massa

Por definição, a parte espacial de \mathcal{F}^μ corresponde as equações de movimento relativísticas. Mas, qual o significado da parte temporal? Para entendermos o seu significado partiremos da equação (4.12).

$$\begin{aligned}\gamma \frac{dP^0}{dt} &= \mathcal{F}^0 \\ \gamma \frac{d}{dt}(\gamma mc) &= \frac{\gamma}{c} \vec{F} \cdot \vec{v} \\ \frac{d}{dt}(\gamma mc^2) &= \vec{F} \cdot \vec{v},\end{aligned}\tag{4.19}$$

Para entendermos o significado da equação (4.19) relembremos alguns conceitos de mecânica newtoniana. Um trabalho infinitesimal de uma força \vec{F} é dado por:

$$\begin{aligned}dW &= \vec{F} \cdot d\vec{x} \\ dW &= \vec{F} \cdot \vec{v} dt,\end{aligned}$$

assim,

$$\frac{dW}{dt} = \vec{F} \cdot \vec{v} = P,$$

em que P é chamada de potência fornecida pela força \vec{F} . Por outro lado, na mecânica newtoniana $\vec{F} = m\vec{a} = m d\vec{v}/dt$ quando a massa da partícula é constante. Então o trabalho infinitesimal é dado por:

$$\begin{aligned}dW &= m \frac{d\vec{v}}{dt} \cdot \vec{v} dt \\ dW &= m\vec{v} \cdot d\vec{v}\end{aligned}$$

Definindo a energia de partícula livre newtoniana ou energia cinética newtoniana como

$$K = \frac{mv^2}{2} = \frac{m\vec{v} \cdot \vec{v}}{2},$$

obtemos que

$$\begin{aligned}dK &= m\vec{v} \cdot d\vec{v} \\ \frac{dK}{dt} &= P = \vec{F} \cdot \vec{v}.\end{aligned}$$

Podemos então interpretar que o trabalho infinitesimal de uma força newtoniana \vec{F} causa uma variação infinitesimal na energia de partícula livre newtoniana K . Dessa forma, a equação (4.19) nos mostra que, no contexto da Relatividade Restrita, a potencia

fornecida por um uma força \vec{F} causa a variação de um novo tipo de energia de partícula livre denotada por E dada por:

$$E = \gamma mc^2, \quad (4.20)$$

portanto, no contexto da relatividade restrita, um trabalho realizado por uma força causa uma variação desta nova energia e não uma variação sobre a energia cinética newtoniana $mv^2/2$.

Ao expandir E em série de Taylor obtemos:

$$E = mc^2 + \frac{1}{2}mv^2 + \dots, \quad (4.21)$$

assim, a energia relativística, a baixas velocidades, recai na energia cinética clássica com a adição do termo mc^2 . Este termo é chamado de energia de repouso pois, quando a velocidade é zero a energia resultante é este próprio fator. Dessa forma, podemos considerar que há uma energia associada a massa das partículas e que, a princípio, podemos converter matéria em energia.

Em teorias quânticas modernas costuma-se associar as partículas elementares uma energia e não uma massa. Este fato vem da energia da repouso, já que ela nos diz que matéria e energia não manifestações diferentes do mesmo ente físico.

Da equação (4.20) podemos reescrever a expressão do quadrimomento como:

$$P^\mu = \left(\frac{E}{c}, \vec{p} \right), \quad (4.22)$$

portanto, a energia pode ser interpretada como a componente na direção do tempo do quadrimomento. Uma outra relação entre momento e energia, chamada de relação pitagórica, é:

$$E^2 = p^2 c^2 + m^2 c^4, \quad (4.23)$$

quando $m = 0$, partículas sem massa, obtemos:

$$E = |\vec{p}| c, \quad (4.24)$$

além disso:

$$\begin{aligned} |\vec{p}| c &= \gamma mc^2 \\ \gamma m |\vec{v}| c &= \gamma mc^2 \\ |\vec{v}| &= c, \end{aligned} \quad (4.25)$$

portanto, no contexto da relatividade restrita, é possível a existência de partículas sem massa, mas que possuam energia e momento. Porém, obrigatoriamente, essas partículas devem viajar na velocidade da luz. Um exemplo de tal partícula é o fóton.

4.4 Formulação lagrangiana e hamiltoniana

Nesta sessão mostraremos como podemos utilizar as formulações lagrangiana e hamiltoniana no contexto na mecânica relativística. Essas formulações são muito úteis por serem mais simples de ser aplicadas na prática, por exporem as simetrias do sistema estudado de forma mais clara, o que não é visto facilmente no formalismo de forças e por facilitarem o estudo de quantidades conservadas do sistema.

Suponha que a força tridimensional \vec{F} seja conservativa, ou seja, tenha uma função energia potencial $V(\vec{r})$ associada e que ela seja determinada por $\vec{F} = -\vec{\nabla}V(\vec{r})$. Assim,

$$\frac{d}{dt} \left(\frac{m\vec{v}}{\sqrt{1 - v^2/c^2}} \right) = -\vec{\nabla}V(\vec{r}), \quad (4.26)$$

note que,

$$\frac{m\dot{x}}{\sqrt{1 - (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)/c^2}} = \frac{\partial}{\partial \dot{x}} \left(-mc^2 \sqrt{1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2}} - V \right), \quad (4.27)$$

já que V depende apenas das posições generalizadas. Substituindo a equação (4.27) em (4.26) e fazendo manipulações que não alteram o resultado podemos escrever que:

$$\frac{d}{dt} \left\{ \frac{\partial}{\partial \dot{x}} \left[-mc^2 \sqrt{1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2}} - V \right] \right\} - \frac{\partial}{\partial x} \left[-mc^2 \sqrt{1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2}} - V \right] = 0, \quad (4.28)$$

cujo resultado vale para todas posições e velocidades generalizadas. A equação (4.28) tem a mesma estrutura que as equações de Euler-Lagrange:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^k} \right) - \frac{\partial L}{\partial q^k} = 0, \quad (4.29)$$

o que nos leva a definir uma função lagrangiana para a relatividade restrita dada por:

$$L(\vec{r}, \dot{x}, \dot{y}, \dot{z}) = -mc^2 \sqrt{1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2}} - V(\vec{r}) = -\frac{mc^2}{\gamma} - V(\vec{r}). \quad (4.30)$$

Podemos definir a função lagrangiana outra forma. Tome um quadrivetor no espaço tempo de tamanho infinitesimal:

$$d\vec{x} = dx^\mu \vec{e}_\mu, \quad (4.31)$$

em que \vec{e}_μ é uma base do espaço.

Seu módulo é dado por:

$$|d\vec{x}| = \sqrt{dx^\mu \vec{e}_\mu \cdot dx^\nu \vec{e}_\nu} = \sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} dt. \quad (4.32)$$

Integrando ambos os lados encontramos o comprimento total S percorrido pela partícula entre os tempo t_1 e t_2 :

$$S = \int_{t_1}^{t_2} \sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} dt. \quad (4.33)$$

Fisicamente uma partícula que não está sujeita a um potencial (partícula livre), segue sempre o caminho mais curto entre um ponto e outro. Assim, poderíamos usar a equação (4.33) para definir a parte livre da função lagrangiana. Contudo, o objeto dentro da integral não tem dimensão de energia sendo necessária a multiplicação deste termo por $-mc$ para que se cumpra tal objetivo. O termo $-mc$ é escolhido de tal forma que obtemos as equações de movimento já conhecidas. Portanto a parte livre da função lagrangiana L_o fica:

$$L_o = -mc\sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} \quad (4.34)$$

Assim, as equações de movimento dependem da forma de $g_{\mu\nu}$. Na relatividade Geral o tensor métrico é tratado com um objeto dinâmico que depende da distribuição de matéria e energia no espaço. Ou seja, a interação gravitacional é interpretada como a distorção no espaço-tempo causada por uma dada distribuição de matéria e energia. Assim, outros corpos sentem esta curvatura e seguem curvas geodésicas⁴ neste espaço curvo. Se não houver outro tipo de interação, a lagrangiana livre é a própria função lagrangiana. Por exemplo, a parte livre da lagrangiana gerada por uma partícula pontual de massa M é dada por:

$$L = -mc^2\sqrt{\mathcal{L}}, \quad (4.35)$$

em que

$$\mathcal{L} = 1 - \frac{r_s}{r} - \frac{1}{1 - \frac{r_s}{r}} \frac{\dot{r}^2}{c^2} - \frac{r^2 \dot{\varphi}^2}{c^2} - \frac{r^2 \sin^2(\varphi) \dot{\theta}^2}{c^2}, \quad (4.36)$$

com $r_s = 2GM/c^2$ é chamado de raio de Schwarzschild.

Para o caso do espaço de Minkowski a parte livre da lagrangiana fica:

$$L_o = -mc\sqrt{\left(\frac{dx^0}{dt}\right)^2 - \left(\frac{dx^1}{dt}\right)^2 - \left(\frac{dx^2}{dt}\right)^2 - \left(\frac{dx^3}{dt}\right)^2}, \quad (4.37)$$

$$L_o = -mc\sqrt{c^2 - (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)}, \quad (4.38)$$

$$L_o = -mc^2\sqrt{1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2}} = -\frac{mc^2}{\gamma}. \quad (4.39)$$

Adicionando o termo de interação obtemos:

$$L = -\frac{mc^2}{\gamma} - V(\vec{r}). \quad (4.40)$$

⁴Curvas geodésicas são as curvas de menor comprimento em um dado espaço.

Definida a função lagrangiana, basta fazermos uma transformada de Legendre para obter a função hamiltoniana:

$$\mathcal{H}(\vec{r}, \dot{x}, \dot{y}, \dot{z}) = \dot{q}^k \frac{\partial L}{\partial \dot{q}^k} - L = \gamma mc^2 + V(\vec{r}). \quad (4.41)$$

Aqui pode-se ver claramente que a função \mathcal{H} é a energia total do sistema sempre que V não depende do tempo e como a função L não depende explicitamente do tempo, a energia é uma quantidade conservada.

Capítulo 5

Incorporação na Mecânica Quântica

Neste capítulo veremos a incorporação da Relatividade Restrita na Mecânica Quântica que culminaram nas conhecidas equação de Dirac (ED) e equação de Klein-Gordon (EKG). Não entraremos em muitos detalhes sobre as soluções de tais equações mas veremos suas principais implicações. Focaremos nas suas derivações e nos princípios que tornam estas equações as versões relativísticas da equação de Schrödinger (ES). Além disso veremos com se comportam as densidades de probabilidade de cada teoria para que possamos compara-las [10].

5.1 Equação de Schrödinger

A ES pode ser escrita em termo dos operadores energia e hamiltoniano:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V\Psi \quad (5.1)$$

$$\hat{E}\Psi = \hat{H}\Psi, \quad (5.2)$$

em que $\Psi = \Psi(\vec{r}, t)$ é uma função escalar. Essa equação expressa a conservação de energia do sistema e pode ser derivada deste princípio apenas fazendo a substituição da energia total pelo operador energia e a soma da energia cinética e potencial pelo operador hamiltoniano.

A ES trouxe enormes avanços para a física e para a sociedade. Ela prevê alguns fenômenos que eram observados em laboratório¹ e alguns novos constatados apenas mais tarde². Tal equação proporcionou o desenvolvimento de diversas tecnologias, como a eletrônica moderna, que contribuíram muito para nossa sociedade.

¹Como por exemplo o espectro de emissão dos átomos e a discretização da energia em sistemas confinados.

²Como por exemplo o tunelamento quântico.

5.1.1 Interpretação estatística

Para o caso da partícula livre, basta fazermos a função energia potencial V ser igual a zero.

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi. \quad (5.3)$$

Multiplicando a equação (5.3) pelo complexo conjugado de Ψ , obtemos

$$i\hbar \Psi^* \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Psi^* \nabla^2 \Psi. \quad (5.4)$$

Escrevendo o complexo conjugado da equação (5.3) e multiplicando por Ψ obtemos,

$$i\hbar \Psi \frac{\partial \Psi^*}{\partial t} = \frac{\hbar^2}{2m} \Psi \nabla^2 \Psi^*. \quad (5.5)$$

Somando as equações (5.4) e (5.5) obtemos,

$$\begin{aligned} i\hbar \left(\Psi^* \frac{\partial \Psi}{\partial t} + \Psi \frac{\partial \Psi^*}{\partial t} \right) &= \frac{\hbar^2}{2m} (\Psi \nabla^2 \Psi^* - \Psi^* \nabla^2 \Psi) \\ \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} &= 0, \end{aligned} \quad (5.6)$$

em que,

$$\rho = \Psi \Psi^*, \quad (5.7)$$

é interpretado como a densidade de probabilidade e

$$\vec{j} = \frac{i\hbar}{2m} (\Psi \vec{\nabla} \Psi^* - \Psi^* \vec{\nabla} \Psi), \quad (5.8)$$

é interpretado como a densidade de corrente de probabilidade.

A função ρ pode ser interpretado dessa forma pois tem um valor sempre positivo. Veremos que essa propriedade será violada na equação de Klein-Gordon (EKG) mas não na equação de Dirac (ED). Além disso, a equação (5.6) nos dá uma lei de conservação do fluxo de probabilidade.

5.2 Equação de Klein-Gordon

Uma maneira de incorporar a Relatividade Restrita na Mecânica Quântica é por meio da EKG. Ela parte da expressão da energia de uma partícula livre relativística:

$$E^2 = p^2 c^2 + m^2 c^4, \quad (5.9)$$

fazendo a substituição dos operadores de energia e momento definidos em (5.1) e (5.2) e aplicando sobre uma função $\Psi(\vec{r}, t)$, obtemos

$$-\hbar^2 \frac{\partial^2 \Psi}{\partial t^2} = (-\hbar^2 c^2 \nabla^2 + m^2 c^4) \Psi,$$

que pode ser escrita da seguinte forma

$$\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} - \nabla^2 \Psi + \left(\frac{mc}{\hbar} \right)^2 \Psi = 0, \quad (5.10)$$

esta é a EKG. A fim de obtermos uma versão covariante da EKG usaremos o operador de onda ou D'Alembertiano definido por:

$$\square^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 = \frac{\partial^2}{\partial x_0^2} - \nabla^2 = \partial^\mu \partial_\mu,$$

assim a versão covariante da EKG é dada por:

$$\left[\square^2 + \left(\frac{mc}{\hbar} \right)^2 \right] \Psi = 0. \quad (5.11)$$

Pelo princípio da relatividade exposto no Capítulo 1 tal equação deve ser a mesma em qualquer referencial inercial. Assim a EKG deve ser invariante sobre transformações de Lorentz. A função de onda Ψ é trivialmente invariante por se tratar de um escalar, assim como o termo mc^2/\hbar . Precisamos então mostrar que o D'Alembertiano é um invariante relativístico.

$$\begin{aligned} \bar{\partial}^\mu \bar{\partial}_\mu &= \frac{\partial \bar{x}^\mu}{\partial x^\nu} \frac{\partial x^\sigma}{\partial \bar{x}^\mu} \partial^\nu \partial_\sigma \\ &= \frac{\partial x^\sigma}{\partial x^\nu} \partial^\nu \partial_\sigma \\ &= \delta_\nu^\sigma \partial^\nu \partial_\sigma \\ &= \partial^\nu \partial_\nu, \end{aligned}$$

assim, mostramos que o operador D'Alembertiano é invariante por qualquer transformações de coordenadas o que inclui as transformações de Lorentz.

O fato de a função Ψ na EKG ser uma função escalar faz com que esta equação seja aplicável apenas para partículas que não possuem *spin* chamadas de Bósons (exemplos são o méson π , o fóton, bóson de Higgs e o glúon), pois para partículas com esta propriedade intrínseca deveríamos obter pelo menos duas funções de onda distintas associadas aos *spin up* e *spin down*.

5.2.1 Interpretação Estatística

Multiplicando a equação (5.10) por $-i\Psi^*$ obtemos

$$-i\Psi^* \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} + i\Psi^* \nabla^2 \Psi - i\Psi^* \left(\frac{mc}{\hbar} \right)^2 \Psi = 0. \quad (5.12)$$

Tomando o complexo conjugado da equação (5.10) e multiplicando por $i\Psi$ obtemos

$$i\Psi \frac{1}{c^2} \frac{\partial^2 \Psi^*}{\partial t^2} - i\Psi \nabla^2 \Psi^* + i\Psi \left(\frac{mc}{\hbar} \right)^2 \Psi^* = 0. \quad (5.13)$$

Somando as equações (5.12) e (5.13) e rearranjando os termos chegamos que

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0, \quad (5.14)$$

em que

$$\rho = \frac{i\hbar}{2mc^2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t} \right),$$

e

$$\vec{j} = \frac{i\hbar}{2m} (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi).$$

O problema é que ρ não pode ser interpretado como uma densidade de probabilidade por não ser sempre uma função positiva. Essencialmente este problema vem do fato de que a EKG é de segunda ordem no tempo, o que faz com que apareçam as primeiras derivadas temporais na função densidade de probabilidade. Tal problema foi posteriormente corrigido com a segunda quantização advinda da Teoria Quântica de Campos (TQC). Além disso, soluções da EKG preveem estados com energias negativas. Assim, a EKG foi considerada sem sentido na época em que foi proposta. Porém, as soluções com energias negativas foram associadas as antipartículas, que foram descobertas posteriormente [2]. Ademais, quando calculamos os espectros de energia para partículas bosônicas³, usando a EKG, obtemos o valor correto obtido experimentalmente. Portanto, concluímos que a EKG é de extrema utilidade na física, contudo, só pode ser usada no estudo de bósons, mesmo que tal restrição não apareça explicitamente em sua dedução.

5.3 Equação de Dirac

Na tentativa de obter uma versão relativística da ES que fosse consistente com a interpretação probabilística, Dirac chegou em uma equação conhecida como Equação de Dirac (ED) que é válida apenas para férmions⁴.

A equação proposta por Dirac parte da expressão relativística da energia (equação (4.23)), mas escrita da seguinte forma:

$$E = \sqrt{p^2 c^2 + m^2 c^4}, \quad (5.15)$$

porém, não podemos simplesmente substituir o operador momento nesta expressão pois ficaríamos com a raiz quadrada de um operador que, quando expandida em série de Taylor, resultaria em uma equação diferencial de ordem infinita. Devemos então linearizar a expressão da raiz quadrada. Dirac supôs uma linearização da seguinte forma:

$$\sqrt{p^2 c^2 + m^2 c^4} = -i\hbar c \alpha^k \frac{\partial}{\partial x^k} + \beta mc^2, \quad (5.16)$$

em que estamos usando a convenção de soma ($k = 1, 2, 3$), sendo \hbar a constante de Planck dividida por 2π e com α^k e β parâmetros a serem determinados. Substituindo a equação (5.16) na equação (5.15) obtemos

$$i\hbar \frac{\partial \Psi}{\partial t} = (-i\hbar c \alpha^k \partial_k + \beta mc^2) \Psi, \quad (5.17)$$

³Partículas que não possuem *spin*.

⁴Partículas com *spin*.

em que $\partial_k = \partial/\partial x^k$. Esta é a Equação de Dirac.

Para obter os valores dos parâmetros α^k e β , Dirac impôs que quando aplicamos o operador energia duas vezes sobre Ψ devemos obter a EKG. Dessa forma ficamos com:

$$\begin{aligned} -\hbar^2 \frac{\partial^2 \Psi}{\partial t^2} &= (-i\hbar c \alpha^i \partial_i + \beta m c^2)(-i\hbar c \alpha^j \partial_j + \beta m c^2)\Psi \\ -\hbar^2 \frac{\partial^2 \Psi}{\partial t^2} &= -\hbar^2 c^2 \alpha^i \alpha^j \partial_i \partial_j \Psi - i\hbar m c^3 (\alpha^i \beta + \beta \alpha^i) \partial_i \Psi + \beta^2 m^2 c^4 \Psi, \end{aligned}$$

que pode ser escrita como

$$-\hbar^2 \frac{\partial^2 \Psi}{\partial t^2} = -\hbar^2 c^2 \frac{(\alpha^i \alpha^j + \alpha^j \alpha^i)}{2} \partial_i \partial_j \Psi - i\hbar m c^3 (\alpha^i \beta + \beta \alpha^i) \partial_i \Psi + \beta^2 m^2 c^4 \Psi,$$

multiplicando a equação por $-1/\hbar^2 c^2$ obtemos

$$\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} - \frac{(\alpha^i \alpha^j + \alpha^j \alpha^i)}{2} \partial_i \partial_j \Psi + \beta^2 \left(\frac{mc}{\hbar}\right)^2 \Psi = i \frac{mc}{\hbar} (\alpha^i \beta + \beta \alpha^i) \partial_i \Psi, \quad (5.18)$$

comparando as equações (5.18) e (5.10) obtemos as seguintes condições para α^i e β :

$$\alpha^i \alpha^j + \alpha^j \alpha^i = 2 \delta^{ij} \quad (5.19)$$

$$(\alpha^i)^2 = (\beta)^2 = 1 \quad (5.20)$$

$$\alpha^i \beta + \beta \alpha^i = 0. \quad (5.21)$$

Essas condições mostram que α^i e β não podem ser números. Note que, caso eles fossem números, de (5.21) obtemos que $\alpha^i = 0$ ou $\beta = 0$, porém essas duas soluções não podem satisfazer a condição (5.20). Então α^i e β devem ser matrizes de ordem N e Ψ deve ser um vetor de N componentes

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \vdots \\ \Psi_N \end{pmatrix}_{N \times 1}. \quad (5.22)$$

O análogo ao complexo conjugado da função de onda neste caso é o conjugado hermitiano:

$$\Psi^\dagger = (\Psi_1^* \quad \Psi_2^* \quad \dots \quad \Psi_N^*)_{1 \times N}. \quad (5.23)$$

Devemos então corrigir, por conta dos argumentos citados acima, as condições sobre α^i e β para:

$$\alpha^i \alpha^j + \alpha^j \alpha^i = 2 \delta^{ij} \mathbb{I}_{N \times N} \quad (5.24)$$

$$(\alpha^i)^2 = (\beta)^2 = \mathbb{I}_{N \times N} \quad (5.25)$$

$$\alpha^i \beta + \beta \alpha^i = \mathbb{O}_{N \times N}, \quad (5.26)$$

em que $\mathbb{I}_{N \times N}$ é a matriz identidade $N \times N$ e $\mathbb{O}_{N \times N}$ é a matriz nula $N \times N$. Dadas essas propriedades, vamos encontrar qual a forma dessas matrizes. Para isso apresenta-se primeiramente as consequências dessas propriedades:

1) os autovalores das matrizes α^i e β são ± 1 .

Demonstração 1:

Considere \vec{v} um autovetor de β com autovalor q , isto é

$$\begin{aligned}\beta\vec{v} &= q\vec{v} \\ \beta\beta\vec{v} &= q\beta\vec{v} \\ \vec{v} &= q^2\vec{v} \\ q &= \pm 1.\end{aligned}$$

O mesmo vale para α^i , como queríamos demonstrar.

2) o traço das matrizes α^i e β é igual a zero.

Demonstração 2:

Da condição (5.26) temos

$$\begin{aligned}\alpha^i\beta &= -\beta\alpha^i \\ \alpha^i\beta\beta &= -\beta\alpha^i\beta \\ \alpha^i &= -\beta\alpha^i\beta.\end{aligned}$$

Tirando o traço desta equação

$$\begin{aligned}tr(\alpha^i) &= -tr(\beta\alpha^i\beta) \\ tr(\alpha^i) &= -tr(\alpha^i\beta\beta) \\ tr(\alpha^i) &= -tr(\alpha^i) \\ tr(\alpha^i) &= 0.\end{aligned}$$

De forma análoga $tr(\beta) = 0$, como queríamos demonstrar.

Com essas duas consequências em mãos podemos inferir que a ordem das matrizes α^i e β é par.

Demonstração 3:

Nesta demonstração explicitaremos quando houver uma soma para não haver confusão de índices.

Seja \hat{e}_k o autovetor do operador β e q_k seu autovalor, isto é

$$\beta\hat{e}_k = q_k\hat{e}_k,$$

Seja também os vetores \vec{u} e \vec{v} escritos na base dos autovalores de β ,

$$\begin{aligned}\vec{u} &= \sum_i u_i \hat{e}_i \\ \vec{v} &= \sum_j v_j \hat{e}_j,\end{aligned}$$

de tal forma que

$$\vec{v} = \beta \vec{u}. \quad (5.27)$$

Multiplicando a equação (5.27) a esquerda por \hat{e}_k^\dagger e sabendo que os autovetores \hat{e}_i são ortogonais por virem de uma equação de autovalor de um operador hermitiano obtemos

$$v_k = u_i \hat{e}_k^\dagger \beta \hat{e}_i.$$

Assim, podemos interpretar que o fator $\hat{e}_k^\dagger \beta \hat{e}_i$ é o elemento ki da matriz β que leva a componente i do vetor \vec{u} na componente k do vetor \vec{v} . Aplicando β em \hat{e}_i obtemos

$$\hat{e}_k^\dagger \beta \hat{e}_i = \beta_{ki} = q_i \delta_{ki},$$

ou seja, β é uma matriz diagonal cujos elementos da diagonal principal são seus autovalores. Mas eles só podem ser ± 1 e quando os somamos o resultado deve ser zero, ou seja, devemos ter a mesma quantidade de entradas $+1$ e -1 , portanto a ordem da matriz β deve ser par. Como as outras matrizes devem ter a mesma ordem de β , então elas também tem ordem par, como queríamos demonstrar.

Poderíamos supor que a ordem das matrizes α^i e β é 2. Porém é impossível encontrar 4 matrizes diferentes que respeitem as propriedades (5.24)-(5.26). Nesta condição, as únicas matrizes que contemplam estas propriedades são as 3 matrizes de Pauli:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (5.28)$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (5.29)$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (5.30)$$

logo a menor ordem possível para as matrizes α^i e β é 4. Uma das escolhas possíveis para essas matrizes, conhecida como representação padrão, é:

$$\alpha^i = \begin{pmatrix} \mathbb{O} & \sigma_i \\ \sigma_i & \mathbb{O} \end{pmatrix} \quad (5.31)$$

$$\beta = \begin{pmatrix} \mathbb{I} & \mathbb{O} \\ \mathbb{O} & -\mathbb{I} \end{pmatrix}, \quad (5.32)$$

em que as entradas das matrizes α^i e β são matrizes nulas (\mathbb{O}), ou matrizes identidade (\mathbb{I}) ou matrizes de Pauli, todas de ordem 2. Porém, há outras escolhas possíveis.

Uma outra consequência imediata da fixação da ordem das matrizes é que a função de onda Ψ se torna um vetor de quatro componentes e não mais uma função escalar, como havíamos suposto inicialmente. Quando resolvemos a ED explicitamente nota-se que as duas primeiras componentes estão associadas aos *spins up* e *down* da partícula e as duas últimas estão associadas aos *spins up* e *down* da sua respectiva anti-partícula. Portanto, a noção de *spins* aparece naturalmente na teoria desenvolvida por Dirac, em contraste com a de Schrödinger, em que tal propriedade das partículas deve ser colocada a mão.

Vamos escrever a ED (equação (5.17)) em sua forma covariante. Multiplicando à esquerda a equação (5.17) por β/c obtemos

$$\begin{aligned} -\frac{i\hbar}{c}\beta\partial_t\Psi - i\hbar\beta\alpha^i\partial_i\Psi + mc\Psi &= 0 \\ i\hbar\beta\partial_0\Psi + i\hbar\beta\alpha^i\partial_i\Psi - mc\Psi &= 0, \end{aligned}$$

definindo as matrizes de Dirac como

$$\gamma^0 = \beta \quad (5.33)$$

$$\gamma^i = \beta\alpha^i, \quad (5.34)$$

podemos escrever a ED (equação (5.17)) na forma covariante:

$$\left(i\gamma^\mu\partial_\mu - \frac{mc}{\hbar}\mathbb{I}\right)\Psi = 0. \quad (5.35)$$

Vale ressaltar as propriedades das matrizes de Dirac:

$$(\gamma^0)^\dagger = \gamma^0 \quad (5.36)$$

$$(\gamma^0)^2 = \mathbb{I} \quad (5.37)$$

$$(\gamma^i)^\dagger = -\gamma^i \quad (5.38)$$

$$(\gamma^i)^2 = -\mathbb{I}, \quad (5.39)$$

além disso elas satisfazem uma álgebra de Clifford, isto é:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{I}, \quad (5.40)$$

em que $\eta^{\mu\nu}$ é a métrica de Minkowski. Este último fato revela que a função Ψ não é um vetor qualquer, mas sim, um objeto matemático conhecido como espinor⁵.

5.3.1 Interpretação estatística

Veremos agora se a ED é compatível com a interpretação probabilística. Multiplicando a equação (5.17) a esquerda por Ψ^\dagger obtemos

$$i\hbar\Psi^\dagger\frac{\partial\Psi}{\partial t} = -i\hbar c\Psi^\dagger\alpha^i\partial_i\Psi + mc^2\Psi^\dagger\beta\Psi. \quad (5.41)$$

Aplicando o conjugado hermitiano na equação (5.17) e multiplicando a direita por $-\Psi$ obtemos

⁵Espinoros ou campos espinoriais são grandezas que podem ser associadas ao espaço euclidiano ou ao de Minkowski que se transformam de uma forma específica quando o espaço é rotacionado infinitesimalmente, ou por uma rotação infinitesimal (no caso do espaço euclidiano), ou por uma transformação de Lorentz infinitesimal (no caso do espaço de Minkowski). Porém, ao compormos um sucessão de rotações infinitesimais para compor uma rotação geral, a transformação do espinor depende da ordem das rotações infinitesimais ao contrário dos vetores e tensores. Por exemplo, um espinor se transforma no seu inverso quando ocorre uma rotação de 360° no espaço euclidiano.[12]

$$i\hbar \frac{\partial \Psi^\dagger}{\partial t} \Psi = -i\hbar c (\partial_i \Psi^\dagger) \alpha^i \Psi - mc^2 \Psi^\dagger \beta \Psi,$$

porém, como o operador hamiltoniano é hermitiano, α^i e β também devem ser. Assim, chegamos que:

$$i\hbar \frac{\partial \Psi^\dagger}{\partial t} \Psi = -i\hbar c (\partial_i \Psi^\dagger) \alpha^i \Psi - mc^2 \Psi^\dagger \beta \Psi, \quad (5.42)$$

somando as equações (5.41) e (5.42) e dividindo por $i\hbar$ obtemos

$$\frac{\partial}{\partial t} (\Psi^\dagger \Psi) + \partial_i (c \Psi^\dagger \alpha^i \Psi) = 0, \quad (5.43)$$

ou seja, $\Psi^\dagger \Psi$ pode ser interpretada como a densidade de probabilidade e $c \Psi^\dagger \alpha^i \Psi$ é a componente i do vetor densidade de corrente de probabilidade.

Assim como a EKG, a ED prevê soluções com energias negativas. Dessa forma, como as partículas sempre tendem a um estado de menor energia, elas deveriam decair infinitamente nos níveis de energia negativos. Para contornar este problema, Dirac supôs que estes estados de energias negativas já estariam ocupados e, portanto, as partículas não mais decairiam dos estados de energia positiva. Estes estados ocupados ficaram conhecidos como “mar de Dirac”.

Essa interpretação de Dirac gerou muitas controvérsias na época, porém uma de suas consequências é que um fóton poderia excitar uma das partículas nos estados negativos de energia até um estado positivo deixando um “buraco”. Este estado desocupado se comporta exatamente como a partícula que foi excitada, porém com carga oposta. Assim, Dirac interpretou que sua equação estava prevendo novas partículas associadas as que já conhecíamos chamadas de antipartículas. Três anos mais tarde o pósitron, antipartícula do elétron, foi descoberta por Carl David Anderson em experimentos com raios cósmicos [4].

Capítulo 6

Considerações finais

Neste trabalho fizemos uma revisão da Teoria da Relatividade Restrita com o objetivo de fundamentar melhor as bases dessa teoria a fim de iniciar estudos em teorias mais avançadas da física onde o conhecimento profundo de tal teoria é extremamente necessário, como Teorias de Campos. Iniciamos nosso trabalho com a motivação histórica que levou ao surgimento da teoria da Relatividade Restrita. Vimos que, o fato da teoria eletromagnética de Maxwell ser incompatível com as transformações de Galileu e que os experimentos da época indicavam que tal teoria deveria ser a mesma em qualquer referencial inercial, levaram ao surgimento das transformações de Lorentz que mudaram a forma como vemos o mundo para sempre.

Tais transformações nos mostram que medidas de distância e intervalos de tempo dependem do referencial que observamos, ou seja, são grandezas relativas. Elas também mostram que nem espaço e nem tempo tem uma realidade física sozinhos, apenas uma combinação dos dois é que realmente faz sentido físico, o chamado espaço-tempo. Vimos também que não é qualquer combinação de espaço e tempo que faça sentido, apenas uma particular combinação entre eles, chamado de espaço de Minkowski, que tem realidade física. Tal espaço tem uma métrica diferente da euclidiana que estamos acostumados em nosso cotidiano.

Em seguida, fizemos um estudo mais profundo sobre a estrutura de Grupo que as transformações de Lorentz possuem a fim de fazer uma interpretação geométrica de tais transformações e vimos que podemos interpretar tais transformações como rotações hiperbólicas em analogia com as rotações no espaço euclidiano. Além disso, a estrutura de grupo é de extrema utilidade no estudo de simetrias em teorias físicas mais avançadas como Teorias Quânticas de Campos.

Tomando o espaço de Minkowski como a verdadeira estrutura de nosso mundo, foi mostrada a versão relativística da mecânica newtoniana chamada de mecânica relativística. Nesse contexto, mostramos que a velocidade no espaço-tempo de qualquer partícula deve ser igual a velocidade da luz e, portanto, podemos interpretar que os efeitos de dilatação temporal e contração espacial são devidos a este fato. Vimos também um exemplo de uma força constante agindo sobre uma partícula pontual e mostramos que não importa por quanto tempo aceleremos uma partícula, sua velocidade espacial nunca ultrapassa a velocidade da luz e também mostramos que em baixas velocidades as dinâmicas newtoniana e relativística se confundem, como esperado.

Mostramos que a parte temporal das equações de movimento relativísticas estão associadas ao teorema do trabalho e energia e que revelam uma nova forma de energia de partícula livre e ainda o surgimento da energia de repouso, associada a massa das partículas. Vimos também que partículas sem massa podem ter quantidades dinâmicas bem definidas, porém devem sempre ter uma velocidade espacial igual a c . Além disso, foram mostradas as formulações lagrangiana e hamiltoniana da dinâmica relativística, pois tais formalismos são muito úteis na prática e no estudos de simetrias e quantidades conservadas.

Por fim, mostramos as duas principais versões relativísticas da equação de Schrödinger (ES), conhecidas como equação de Klein-Gordon (EKG) e equação de Dirac (ED). Vimos que apesar da EKG ter problemas com a interpretação estatística, que posteriormente foi corrigida com a segunda quantização, ela é essencial no estudos dos bósons (partículas sem *spin*) por calcular com precisão o espectro de energia de sistemas com tais partículas. Além disso, vimos que os estados de energia negativas foram associados as antipartículas, descobertas posteriormente. Já a ED, é coerente com a interpretação estatística e prevê a existência dos *spins* das partículas de forma natural, ao contrário da ES, em que tal propriedade da matéria deve ser posta a mão, ou seja, a ED serve para descrevermos os férmions. Além disso, vimos que a interpretação do “mar de Dirac” para os estados de energia negativa, mesmo que controversa em sua época, trouxe a interpretação das antipartículas vindas da excitação por um fótons de partículas nos estados preenchidos de energia negativa.

Apêndice A

Álgebra tensorial

Neste capítulo faremos uma introdução a matemática dos tensores, já que estes foram usadas durante este trabalho baseados em [1]. Os resultados aqui expostos valem para um espaço com dimensão n qualquer e por isso serão usados índices latinos, os quais assumem os valores $1, 2, \dots, n$, em contraste com os do texto principal, que se aplicam apenas para o caso quadridimensional.

A.1 Componentes de um vetor

Primeiramente faremos a distinção entre as componentes covariantes e contravariantes de um vetor. Sua definição é bem simples: As componentes contravariantes v^i de um vetor \vec{v} são determinadas pelas combinações lineares dos vetores da base do sistema de coordenadas. Já as componentes covariantes v_i de um vetor \vec{v} são determinadas fazendo-se a projeção do vetor sobre seus respectivos vetores de base. Ambas as representações definem igualmente bem um vetor. Note que em coordenadas cartesianas as duas componentes são idênticas, mas, em geral, não são.

A.2 Diática e transformação de coordenadas

Para definir tensores podemos, preliminarmente, dizer que eles são o resultado de uma operação conhecida como diática. Ela é feita da seguinte forma: multiplique uma componente qualquer de um vetor por uma componente qualquer de outro vetor. O resultado desta operação forma uma componente de um tensor similar a uma matriz [1]. Dessa forma se chamarmos este objeto de T e a equação que representa suas componentes é

$$T_{ij} = A_i B_j. \quad (\text{A.1})$$

Na realidade, a forma mais precisa de definir um tensor é olhando para a transformação de suas componentes quando fazemos uma mudança de base. Assim, suponha que temos as equações que relacionam as componentes contravariantes, \bar{x}^i , de um sistema de coordenadas \bar{S} em função das componentes x^i contravariantes do antigo sistema de coordenadas S . De maneira geral elas são dadas por:

$$\bar{x}^i = \bar{x}^i(x^1, x^2, \dots, x^n), \quad (\text{A.2})$$

calculando o diferencial de \bar{x}^i obtemos:

$$d\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^j} dx^j. \quad (\text{A.3})$$

A equação (A.3) define como as componentes contravariantes de um vetor se transformam de um sistema de coordenadas para outro. Assim qualquer grandeza que se transforma como:

$$\bar{V}^i = \frac{\partial \bar{x}^i}{\partial x^j} V^j, \quad (\text{A.4})$$

é chamada de vetor contravariante.

Vamos agora definir a lei de transformação das componentes covariantes de um vetor. Tome uma função escalar Φ . Por definição, um escalar é uma quantidade que não depende do sistema de coordenadas, ou seja, é invariante sobre transformações de coordenadas. Assim, $\Phi(\bar{x}^i) = \Phi(x^j)$. Porém, as derivadas espaciais de Φ são vetores (podemos interpretar como sendo as componentes do gradiente de Φ). Calculando as derivadas em relação ao “novo” sistema de coordenadas e usando a regra da cadeia obtemos:

$$\begin{aligned} \frac{\partial \Phi}{\partial \bar{x}^k} &= \frac{\partial \Phi}{\partial x^m} \frac{\partial x^m}{\partial \bar{x}^k} \\ \frac{\partial \Phi}{\partial \bar{x}^k} &= \frac{\partial \Phi}{\partial x^m} \frac{\partial x^m}{\partial \bar{x}^k}. \end{aligned} \quad (\text{A.5})$$

A equação (A.5) define a transformação das componentes covariantes de um vetor, ou seja, qualquer quantidade que se transforme seguindo:

$$\bar{V}_i = \frac{\partial x^j}{\partial \bar{x}^i} V_j \quad (\text{A.6})$$

é chamado de vetor covariante. Dessa forma, como um tensor pode ser definido como uma diádica, a lei de transformação para tensores é dada por:

$$\bar{T}_j^i = \bar{A}^i \bar{B}_j = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^n}{\partial \bar{x}^j} T_n^m \quad (\text{A.7})$$

Este é um exemplo de um tensor misto de segunda ordem (há dois índices livres), que tem um índice contravariante i e um covariante j . As transformações de outros tipos de tensores (ordem maiores e posições dos índices), são definidas de forma análoga.

A.3 Contração

A operação de contração é definida quando em uma expressão que envolve a multiplicação entre componentes de tensores e vetores há pelo menos um par de índices mudos, ou seja, esse índice se repete e o resultado final não depende do índice somado. O importante a se

notar é que, necessariamente, estes índices devem ser um co e o outro contravariante, caso contrário a quantidade resultante não é um tensor. Mostremos isto com um exemplo.

Considere um vetor contravariante A^i construído a partir da operação $A^i = T^{ij}V_j$. Dessa forma, A^i deve respeitar a lei de transformação (A.3) quando feita uma mudança de coordenadas. Calculemos tal transformação:

$$\begin{aligned}\bar{A}^i &= \bar{T}^{ij}\bar{V}_j \\ \bar{A}^i &= \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} T^{kl} \frac{\partial x^m}{\partial \bar{x}^j} V_m,\end{aligned}$$

porém,

$$\frac{\partial x^m}{\partial \bar{x}^j} \frac{\partial \bar{x}^j}{\partial x^l} = \frac{\partial x^m}{\partial x^l} = \delta_l^m, \quad (\text{A.8})$$

portanto,

$$\bar{A}^i = \frac{\partial \bar{x}^i}{\partial x^k} \delta_l^m T^{kl} V_m = \frac{\partial \bar{x}^i}{\partial x^k} T^{km} V_m = \frac{\partial \bar{x}^i}{\partial x^k} A^k, \quad (\text{A.9})$$

assim, concluímos que A^i é um vetor contravariante. Se considerarmos que $A^i = T^{ij}V_j$, o cancelamento das derivadas parciais mostradas em (A.8) não ocorreria e A^i não respeitaria a lei de transformação (A.3), ou seja, não seria um vetor contravariante (tensor de primeira ordem contravariante).

Com isso em mente a representação de um vetor \vec{v} , dada em função das suas componentes co e contravariantes deve ser:

$$\begin{aligned}\vec{v} &= v^i \vec{e}_i \\ \vec{v} &= v_i \vec{e}^i.\end{aligned}$$

Usando o que foi usado nesta seção, mostraremos que a função Delta de Kronecker deve ser um tensor misto de segunda ordem, isto é, um índice co e outro contravariante para que tenha as mesmas propriedades em todos os sistemas de coordenadas. Vejamos o porque:

$$\bar{\delta}_j^i = \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial x^l}{\partial \bar{x}^j} \delta_l^k = \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial x^k}{\partial \bar{x}^j} = \frac{\partial \bar{x}^i}{\partial \bar{x}^j}, \quad (\text{A.10})$$

assim, concluímos que a função Delta de Kronecker deve ser um tensor misto de segunda ordem.

A.4 Tensor métrico

Para calcular distâncias em um sistema de coordenadas arbitrário podemos, primeiramente, calcular a distância infinitesimal para integrar posteriormente. Assim, definimos um escalar ds^2 que define a distância em tal espaço.

$$d\vec{s} \cdot d\vec{s} = ds^2 = (dx^i dx^j) (\vec{e}_i \cdot \vec{e}_j), \quad (\text{A.11})$$

em que \vec{e}_i e \vec{e}_j são os vetores da base do sistema de coordenadas.

Este produto escalar entre os vetores da base é chamado de tensor métrico g_{ij} . Dessa forma, o cálculo da distância infinitesimal ds entre dois pontos fica da seguinte forma:

$$ds^2 = g_{ij} dx^i dx^j . \quad (\text{A.12})$$

Desta relação, intercalando os índices i e j e sabendo que este objeto é um escalar obtemos $ds'^2 = g_{ji} dx^j dx^i = ds^2$. Ou seja, g_{ij} é um tensor simétrico. Fato que já poderia ter sido constatado facilmente pela definição de g_{ij} e sabendo que o produto escalar é comutativo.

Um exemplo simples é o tensor métrico no sistema de coordenadas cartesiano onde ele é numericamente igual a função Delta de Kronecker δ_j^i . Assim, usando a definição de tensor covariante de segunda ordem, podemos calcular o tensor métrico em qualquer sistema de coordenadas desde que conheçamos a função que relaciona um sistema arbitrário de coordenadas com o sistema cartesiano:

$$\bar{g}_{kl} = \frac{\partial x^i}{\partial \bar{x}^k} \frac{\partial x^i}{\partial \bar{x}^l} . \quad (\text{A.13})$$

Vamos definir agora o tensor recíproco como sendo a matriz inversa de g (tensor métrico representado na forma de matriz):

$$g^{ij} g_{jk} = \delta_j^i , \quad (\text{A.14})$$

em que os índices do tensor recíproco devem ser ambos contravariantes para haver uma igualdade nas posições dos índices de ambos os lados da equação (A.14). Portanto ele deve ser um tensor contravariante de segunda ordem. O tensor recíproco também pode ser definido por $g^{ij} = \vec{e}^i \cdot \vec{e}^j$.

Uma propriedade muito importante do tensor métrico é a de transformar componentes covariantes em componentes contravariantes e vice versa, vejamos como. Da definição de produto escalar e da equação (A.12) podemos escrever que:

$$\begin{aligned} ds^2 &= dx_i dx^i = g_{ij} dx^i dx^j \\ dx_i &= g_{ij} dx^j , \end{aligned} \quad (\text{A.15})$$

ou seja, o tensor métrico g_{ij} é um operador que transforma a componente contravariante de um vetor em sua componente covariante (abaixamento de índices). Multiplicando a esquerda ambos os lados de (A.15) por g^{ki} ficamos com:

$$\begin{aligned} g^{ki} dx_i &= g^{ki} g_{ij} dx^j \\ g^{ki} dx_i &= \delta_j^k dx^j \\ dx^k &= g^{ki} dx_i , \end{aligned} \quad (\text{A.16})$$

ou seja, o tensor recíproco transforma a componente covariante de um vetor em sua componente contravariante (levantamento de índices).

Bibliografia

- [1] Balboa, Medrano. *Campos Vetoriais, tensores e aplicações na Física*. RANSTEC, 1994.
- [2] Bueno, Caio. “Equação de Klein-Gordon”. Em: (2014).
- [3] Capiberibe, Ayni R. *O Princípio da Relatividade: Física (Lições)*. ALRISHA, 2020.
- [4] Ferreira Neto, Osvaldo et al. “Explorando a Equação de Dirac: um passeio por Isolantes Topológicos e Férmions de Majorana”. Em: (2019).
- [5] Gordon, James e Martin, Liebeck. *Representations and characters of groups*. Cambridge University Press, 2001.
- [6] Lemos, Nivaldo A. *Mecânica analítica*. Editora Livraria da Física, 2007.
- [7] MARTINS, R de A. *A origem histórica da Relatividade Especial*. São Paulo: Livraria da Física, 2015.
- [8] Neto, Jose Abdalla Helayël. *Aula 2 (Parte 1) - $D=2$, Lie e Clifford; Espinores; Quatérnions; Bilineares espinoriais - Helayel - YouTube*. https://www.youtube.com/watch?v=7unUiidTAHs&list=PLjEqTgJ_GMqrzsXr66jhqgluas4_EfdCS&index=5. (Acessado em 11/09/2020). 2017.
- [9] Resnick, Robert e Watanabe, Shigeo. *Introdução à relatividade especial*. Editora da Universidade de São Paulo, 1971.
- [10] Rocha, Roldão da. *Aula 10, Mecânica Quântica III, UFABC: Equação de Klein-Gordon e derivação da eq. de Dirac - YouTube*. <https://www.youtube.com/watch?v=1TiVk6vpZXg>. (Acessado em 11/09/2020). 2020.
- [11] Tavares, Marcio e Aroucha, Sarah. *Mecânica Quântica Relativística: Uma introdução ao Campo de Klein-Gordon*. Novas Edições Acadêmicas, 2020.
- [12] Wikipedia. *Espinor*. https://es.wikipedia.org/wiki/Espinor#cite_note-2. (Acessado em 06/05/2021).