

Universidade Federal do Rio de Janeiro

Uma introdução às singularidades em Relatividade Geral

Hugo Tremonte de Carvalho

Rio de Janeiro

Fevereiro de 2013

Universidade Federal do Rio de Janeiro

Uma introdução às singularidades em Relatividade Geral

Hugo Tremonte de Carvalho

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Orientadores: Henrique Bursztyn e Alejandro Cabrera.

Rio de Janeiro
2013

Universidade Federal do Rio de Janeiro

Uma introdução às singularidades em Relatividade Geral

Hugo Tremonte de Carvalho

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Aprovada por:

Presidente, Prof. Alejandro Cabrera (IM-UFRJ)

Prof. Henrique Bursztyn (IMPA)

Prof. Heudson Tosta Mirandola (IM-UFRJ)

Prof. Marcus Venícius Cougo Pinto (IF-UFRJ)

Prof. Maurício Ortiz Calvão (IF-UFRJ)

Prof. Alexei Mailybaev (suplente - IMPA)

Rio de Janeiro
Fevereiro de 2013

C331in

Carvalho, Hugo Tremonte de.

Uma introdução às singularidades em Relatividade Geral/ Hugo Tremonte de Carvalho. – Rio de Janeiro: UFRJ/ IM, 2013.

viii, 122 f. : il. ; 29 cm.

Orientadores: Alejandro Cabrera e Henrique Bursztyn.

Dissertação (Mestrado) — UFRJ/ IM/ Programa de Pós-graduação em Matemática Aplicada, 2013.

Referências Bibliográficas: p. 119-122 .

1. Relatividade Geral. 2. Relatividade Restrita
3. Teoremas de Singularidade. 4. Stephen Hawking

I. Cabrera, Alejandro (Orient.) II. Bursztyn, Henrique (Orient.)
III. Universidade Federal do Rio de Janeiro,
Instituto de Matemática. IV. Título.

CDD519

CDD510 7

Agradecimentos

Essa é uma modesta lista das pessoas que foram importantes para mim ao longo do tempo. Certamente não vou esquecer ninguém, mas se eu for citar todos os motivos dos agradecimentos, isso daria uma nova dissertação! Além disso, não há uma ordem de importância nos agradecimentos. Não se ofenda se você ficou um pouco pra baixo na lista!

Aos meus pais, por mesmo sem entenderem esses símbolos estranhos que estudei ao longo dos últimos anos, sempre me deram todo o apoio e carinho.

Ao Professor Felipe Acker, por ter criado o curso de Matemática Aplicada, um “curso para as pessoas desajustadas”, em suas próprias palavras, um lugar onde aprendi muito e fiz ótimos amigos e por ter me ensinado a pensar matematicamente.

Aos Professores Fábio Ramos e César Niche, por serem não somente ótimos professores mas também nossos amigos e “conselheiros espirituais” em vários momentos difíceis.

Ao Professor Nilson, por ter sido o melhor professor que tive na faculdade, opinião essa que compartilho com vários colegas.

Ao Professor Alfonso Romero, que ministrou um curso sobre Geometria Lorentziana na Escola de Verão de Geometria e Topologia no início de 2011 em São Carlos, que caiu como uma luva para mim, e por inúmeras sugestões para este trabalho.

Aos meus orientadores Henrique Bursztyń e Alejandro Cabrera, por terem me dado toda a liberdade para fazer esse trabalho.

Aos integrantes da minha banca, pelas inúmeras sugestões para esse trabalho. Espero daqui a algum tempo publicá-lo como livro.

Ao Gregory Naber, que escreveu o livro que permitiu que essa dissertação exista [Liv:Nab88].

Aos meus (mais do que) amigos Renato e Krystal e às suas famílias, que conheci no colégio há 9 anos, por estarem sempre presente na minha vida, mesmo quando morei longe, e me obrigarem a tirar férias eventualmente! Sem vocês eu teria mais tempo para me dedicar a esse trabalho, mas certamente o caminho até aqui não seria tão divertido! E sem dúvida eu seria uma pessoa muito mais ranzinza, chata e paranóica do que sou hoje.

Ao meu amigo Júnior, que completa minha banda com o Renato e a Krystal. Apesar de você esquecer ensaios, não prestar atenção no que a gente fala fazendo com que a gente precise repetir a mesma coisa 5 vezes e morar um pouquinho longe, gosto de você!

Aos meus amigos Raphael e Christiano. Estar com vocês é diversão garantida!

À minha amiga Carol, pelas diversas conversas por e-mail e, depois que passamos a morar perto, pessoalmente. Você me fez refletir sobre várias coisas que me tornaram uma pessoa melhor.

Ao meu (também mais do que) amigo Claudio e à sua família, por ter me acolhido quando eu mais precisava (aturar minhas maluquices por 2 anos não deve ser nada fácil!) e por inúmeras discussões sobre Matemática, Música, Filosofia, a Vida, o Universo e Tudo o Mais.

Ao meu amigo Filipe, integrante do Trio Ternura comigo e com o Claudio e nosso companheiro de aventuras. Agradeço por nunca ter me chamado para um certo passeio de carro em Vicente de Carvalho em um dia chuvoso.

Aos meus colegas de turma Lucas e Renata, meus primeiros companheiros nessa jornada através da Matemática.

Aos meus amigos Nicolau e Braga, pelas inúmeras viagens no 485 quando a gente passava mal de tanto rir.

Ao Yuri Saporito, por ceder os arquivos .tex da sua dissertação, que aos poucos foram se modificando até chegar na minha. Sem isso acho que eu jamais saberia como formatar esse texto.

Às minhas orientadoras de Iniciação Científica Glória Braz (na Bioquímica) e Marianty Ionel (na Matemática).

Ao Wallace e o Markus, meus futuros colegas no LPS, por me darem todo o apoio na minha transição para a Engenharia Elétrica no Doutorado.

Aos professores Luiz Wagner e Flávio Ávila, por ter aceitado me orientar no doutorado.

Aos meus amigos da ABC-116 Gabriel Victorino, Gabriel Sanfins, Raphael Lourenço, Daniel, Gabriel Castor, Bruno Santiago (aka Mestre), Trem (aka Felipe Senra), Frodo (aka João Paulo), Roberto Velho, Cecília, Paloma, Guilherme, Aloízio, que dividiram comigo esse espaço ao longo dos anos.

Ao Donald Knuth e Leslie Lamport pela criação do \TeX e \LaTeX , respectivamente.

Aos mestres-cervejeiros do mundo todo (especialmente os da Bélgica) que mudaram a minha vida depois que descobri a magia da boa cerveja.

Ao Willis Carrier, inventor do ar-condicionado, aparelho imprescindível para um morador do Rio de Janeiro.

A Jimmy Wales e Larry Sanger, criadores da Wikipedia, página que me ajudou muito durante esse trabalho.

Aos servidores de Internet que têm como sede Niue, Samoa Ocidental, Rússia e Equador e a seus mantenedores.

Aos meus grandes companheiros de estudos Johann Sebastian Bach, Robert de Visèe, Marin Marais, Agustín Barrios, Sylvius Leopold Weiß, Fernando Sor, John Dowland, Gaspar Sanz, Luyz de Narváez, Wolfgang Amadeus Mozart, Issac Albéniz, Joaquin Rodrigo, Paulinho Nogueira, e seus intérpretes Fábio Zanon, Julian Bream, David Russel, Andrés Segóvia, Hopkinson Smith, Paul O'Dette, Rolf Lislevand, José Miguel Moreno, Lutz Kirchhof, Göran Söllscher, Jakob Lindberg, Hespèrion XX, Jordi Savall, Hille Perl, Lee Santana, João Carlos Martins, Viktoria Mullova, Mstislav Rostropovich, Jacqueline du Pré e muitos, muitos outros compositores e intérpretes que escuto sem cansar.

Aos meus professores de violão Ramón e Leo Rugero.

Ao Cristiano e aos funcionários da Secretaria da Pós-Graduação do IM-UFRJ, por serem excepcionalmente eficientes na hora de resolver nossos problemas burocráticos.

Ao Walter, diretor financeiro do IM-UFRJ, que nos ajudou muito a fazer melhorias no nosso espaço.

Ao Nicola, responsável por matar nossa sede há mais de 8 anos.

Aos funcionários da limpeza do IM-UFRJ, por deixarem nosso ambiente de trabalho sempre limpo.

Peço infinitas desculpas se esqueci alguém. Tomei todo o cuidado do mundo para que isso não aconteça e dei mais atenção aos Agradecimentos do que a muitas partes da Dissertação.

Hugo Tremonte de Carvalho
Fevereiro de 2013

Resumo

Uma introdução às singularidades em Relatividade Geral

Hugo Tremonte de Carvalho

Resumo da dissertação de Mestrado submetida ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Resumo: Essa dissertação tem como objetivo expor uma introdução ao estudo de singularidades em Relatividade Geral. Para tal fim é apresentada brevemente a teoria da Relatividade Restrita, dando atenção a aspectos históricos. Após isso, são expostos resultados básicos sobre a teoria de Causalidade em Relatividade Geral, e por fim prova-se um Teorema de Stephen Hawking sobre a existência de singularidades em espaços-tempo satisfazendo condições fisicamente razoáveis.

Palavras-chave. Relatividade Geral, Relatividade Restrita, Teoremas de Singularidade, Stephen Hawking.

**Rio de Janeiro
Fevereiro de 2013**

Abstract

An Introduction to Singularities in General Relativity

Hugo Tremonte de Carvalho

Abstract da dissertação de Mestrado submetida ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Abstract: The present work intends to expose an introduction to the study of singularities in General Relativity. For this, is briefly presented the Special Relativity theory, with an emphasis in historical aspects. After that, basic results in Causality theory in General Relativity are exposed, and finally a Theorem of Stephen Hawking about the existence of singularities in physically reasonably space-times is proved.

Keywords. General Relativity, Special Relativity, Singularity Theorems, Stephen Hawking.

**Rio de Janeiro
Fevereiro 2013**

Sumário

I	RELATIVIDADE RESTRITA	9
1	Introdução Histórica - Da Antiguidade até a Teoria da Relatividade Restrita	11
1.1	A primeira idéia de espaço-tempo - Da Antiguidade até Galileu	11
1.2	O Princípio da Relatividade de Galileu e sua influência na estrutura do espaço-tempo . . .	12
1.3	As Leis de Newton e suas influências na estrutura do espaço-tempo	13
1.4	A Lei da Gravitação Universal	14
1.5	O Princípio da Equivalência	17
2	A Teoria da Relatividade Restrita	19
2.1	As Equações de Maxwell	19
2.2	O experimento de Michelson-Morley	21
2.3	A dilatação temporal	23
2.4	Princípios Básicos	25
2.5	Transformações de Lorentz	31
2.6	A dilatação temporal - II	34
2.7	A transformação de Lorentz padrão	34
2.8	A contração de FitzGerald-Lorentz	36
2.9	Adição de velocidades	37
2.10	Simultaneidade	38
2.11	O paradoxo dos gêmeos	39
3	A Geometria do Espaço de Minkowski	41
3.1	Motivação	41
3.2	A Métrica	42
3.3	Bases ortonormais e o grupo de isometrias de \mathcal{M}	43
3.4	Vetores especiais têm propriedades especiais	49
3.4.1	Vetores do tipo luz	49
3.4.2	Vetores do tipo tempo	51
3.4.3	Vetores do tipo espaço	59
3.5	Um pouco de Mecânica Relativística	60
3.5.1	Partículas Materiais	60
3.6	O Tensor de Energia-Momento	61
II	RELATIVIDADE GERAL	65
4	A Teoria da Relatividade Geral - Os fundamentos	67
4.1	Introdução	67
4.2	Os Fundamentos	69
4.3	Três exemplos úteis	71
4.3.1	O Espaço-tempo de Minkowski	71

4.3.2	O espaço-tempo de deSitter	71
4.3.3	O espaço-tempo de Einstein-deSitter	72
4.4	Causalidade	73
4.4.1	Orientação no tempo	73
4.4.2	Condição cronológica	75
4.4.3	Estabilidade causal	80
4.4.4	Hiperbolicidade global	82
4.5	A Equação de Einstein	84
4.5.1	Alguns conceitos de Geometria (semi)-Riemanniana	84
4.5.2	O Tensor de Energia-Momento e as Equações de Einstein	85
5	Teoremas de Singularidade	87
5.1	Motivação: Porque singularidades são importantes em Relatividade Geral?	87
5.2	Mas o que é uma singularidade?	88
5.3	Completude Geodésica	88
5.4	O Teorema de Laplace	90
5.5	Um Teorema de S. Hawking	92
5.5.1	Preliminares matemáticas	92
5.5.2	O Teorema	96
5.5.3	Prova do Lema 5.17	100
5.6	Considerações Finais	116

Lista de Figuras

1.1	<i>Segunda Lei de Kepler.</i>	15
2.1	<i>Referenciais em movimento relativo.</i>	20
2.2	<i>Experimento de Michelson-Morley.</i>	22
2.3	<i>Relógios no foguete.</i>	24
2.4	<i>Simultaneidade na Mecânica Clássica.</i>	26
2.5	<i>Cálculo de velocidades na Mecânica Clássica.</i>	27
2.6	<i>Simultaneidade na Relatividade Restrita.</i>	27
2.7	<i>Em Relatividade Restrita, a simultaneidade não é um conceito que independe do observador.</i>	28
2.8	<i>Colocando coordenadas no espaço-tempo.</i>	29
2.9	<i>A distância entre dois pontos é independente do tempo.</i>	29
3.1	<i>Intervalo no espaço-tempo.</i>	42
3.2	<i>Vetores tipo tempo, luz e espaço.</i>	45
3.3	<i>Não-transitividade da relação $<$.</i>	49
3.4	<i>Cone de luz.</i>	50
3.5	<i>Gêmeos viajando.</i>	55
4.1	<i>Espaço-tempo de Sitter.</i>	71
4.2	<i>Espaço-tempo de Einstein-de Sitter.</i>	73
4.3	<i>Variiedade lorentziana não orientável no tempo.</i>	74
4.4	<i>p antecede q cronologicamente.</i>	76
4.5	<i>Espaço tempo no qual $p \in I^+(p)$.</i>	77
4.6	<i>Falha na causalidade.</i>	78
4.7	<i>Prova da Proposição 4.18.</i>	79
4.8	<i>Espaço-tempo no qual existem curvas fechadas do tipo luz.</i>	80
4.9	<i>Informações não se propagam sobre curvas assim.</i>	83
5.1	<i>Teorema de Laplace.</i>	91
5.2	<i>Varição de uma curva γ.</i>	93
5.3	<i>Varição Γ_i da geodésica γ.</i>	98
5.4	<i>$H_{p,k}$.</i>	101
5.5	<i>$N_T^\pm(p)$.</i>	102
5.6	<i>Lema de Gauß.</i>	102
5.7	<i>Ilustração da prova da Proposição 4.15.</i>	104
5.8	<i>Suavização do levantamento da viagem em T_xM.</i>	105
5.9	<i>$J^-(p)$.</i>	106
5.10	<i>$J^+(S)$.</i>	107
5.11	<i>$J^-(p) \cap J^+(S)$.</i>	107
5.12	<i>Prova do Lema 5.29.</i>	109
5.13	<i>L não é contínua.</i>	112

5.14 Conjuntos U_i, V_i e L_i	113
5.15 Prova do Lema 5.31.	114

Prefácio

Antes de tudo, vale ressaltar que esse trabalho quase deixou de existir em Setembro de 2011, quando cientistas do CERN anunciaram a descoberta de neutrinos que viajam mais rápido que a luz [Art:CERN11]. Por sorte, isso não passou de um mal-entendido (veja <http://www.nature.com/news/2011/110922/full/news.2011.554.html>) e você pode continuar lendo esse texto sem medo das teorias terem perdido o seu significado físico!

A Teoria da Relatividade Restrita nasceu em 1905 com Albert Einstein, no artigo “Zur Elektrodynamik bewegter Körper” [Art:Ein05]. Apesar de tal Teoria ser hoje em dia conhecida como a generalização da Mecânica Clássica para velocidades próximas à da luz, ela não nasceu de considerações mecânicas, mas sim de certas inconsistências entre o Princípio da Relatividade de Galileu e o Eletromagnetismo. Apesar de seus paradoxos aparentes e aspectos pouco intuitivos, é uma das Teorias físicas mais testadas experimentalmente hoje em dia em aceleradores de partículas. Além disso os modernos aparelhos de GPS usam Relatividade para realizar certas correções nos cálculos, tornando a localização ainda mais precisa. Seu formalismo matemático é bem simples e elegante, porém não é possível a introdução de campos gravitacionais sem gerar inconsistências.

Somente dez anos depois, Einstein conseguiu modificar sua Teoria de modo a considerar efeitos gravitacionais [Art:Ein15]. A nova Teoria, agora chamada de Relatividade Geral, é bem mais complexa matematicamente, tendo que apelar para a recém-nascida Geometria Riemanniana para sua formulação. Agora corpos com massa influenciam na geometria do espaço-tempo de maneira não-trivial, através das equações de Einstein. Uma solução para tais equações é uma variedade munida de uma métrica Lorentziana, na qual estão embutidos conceitos cronológicos e geométricos sobre o espaço-tempo. Dessa forma, é natural a busca por soluções, e essa empreitada se iniciou logo após a publicação do artigo de 1915 [Art:Ein15]. Salvo raras exceções de significado físico duvidoso (porém não menos interessante do ponto de vista matemático), todas as soluções modelavam espaços-tempo nos quais havia buracos-negros [Art:Sch16] [Art:Rei16] [Art:Nor18] [Art:Ker63] [Art:New65].

A pergunta natural é se toda solução das equações de Einstein que modelam espaços tempo fisicamente razoáveis apresentam algum tipo de comportamento anômalo como os exemplos acima. A primeira resposta veio em 1955, com Raychaudhuri [Art:Ray55]. Ele demonstrou que um espaço-tempo que tenha somente uma “nuvem” de partículas que não estão em rotação em algum momento colapsará em um buraco-negro. Porém, esse não é um caso fisicamente razoável. Respostas mais gerais para essa pergunta foram dadas por Stephen Hawking em 1966 e 1967 [Art:Haw66A] [Art:Haw66B] [Art:Haw67].

O objetivo desse trabalho é demonstrar um desses Teoremas, que afirma que espaços-tempo em expansão em que a gravidade exerce um efeito atrativo, hipóteses fisicamente comprovadas, além de satisfazer certas condições cronológicas também razoáveis, nasceram de uma singularidade, conceito a ser definido precisamente no Capítulo 5. Isso é interpretado como a existência do Big-Bang. Sua demonstração foi baseada em diversas referências, entre elas [Liv:Haw75] [Liv:Onl83] [Liv:Pen05] [Liv:Pen87] [Liv:Wal84] [Liv:Wal92] [Liv:Woo07] [Liv:Nab88], sendo o último o mais importante.

No Capítulo 1 é apresentada uma perspectiva histórica da noção de espaço, começando na Grécia Antiga. No Capítulo 2 discute-se brevemente a Teoria da Relatividade Restrita. Apesar de tal Capítulo não ser necessário do ponto de vista rigoroso para o entendimento do resto do trabalho, nele são apresentadas idéias que fazem da Relatividade uma Teoria muito mais natural. O Capítulo 3 tem algumas redundâncias com o anterior, mas seu objetivo é tornar rigoroso, porém sem perder a intuição, vários conceitos básicos da Relatividade Restrita para sua posterior generalização. No Capítulo 4 introduz-se a Relatividade Geral e discutem-se aspectos cronológicos dos espaços-tempo, preparando o terreno para os Teoremas de Singularidade. Finalmente, no Capítulo 5 é discutido o conceito de singularidade e prova-se um dos Teoremas de Hawking. Sua prova é razoavelmente simples, porém lança mão de um Lema, cuja prova é bastante técnica e é apresentada na Seção 5.5.3.

As referências bibliográficas são citadas de um jeito não usual. Vejamos um exemplo. A referência [Liv:Nab88] deve ser interpretada da seguinte forma: “Liv” indica que é citado um livro. As outras possibilidades são “Art”(artigo), “Cls”(obra clássica) ou “Int”(página da Internet). “Nab” é uma abreviação do nome do autor, Gregory Naber nesse caso. Por fim, “88” indica o ano de publicação do trabalho. Quando o autor publicou dois ou mais trabalhos no mesmo ano, distinguimos como “xxA” e “xxB”. No caso da publicação ser anterior a 1900, a data é escrita por extenso para não haver ambiguidade.

HOW DO I LOVE YOU, THESIS? LET ME COUNT THE WAYS...



WWW.PHDCOMICS.COM

Parte I

RELATIVIDADE RESTRITA

Capítulo 1

Introdução Histórica - Da Antiguidade até a Teoria da Relatividade Restrita

Desde a Antiguidade o conceito de espaço é uma questão recorrente nas discussões filosóficas e científicas. É natural de se esperar essa recorrência, já que estamos falando de nossa “casa”, o lugar em que todas nossas ações acontecem. Ao longo do tempo várias hipóteses sobre tal conceito foram propostas. Por exemplo, durante muito tempo acreditou-se que a Terra era plana, até que notou-se que quando um barco veleja até o horizonte ele “some” aos poucos: primeiro o casco e em seguida as velas. Caso a Terra fosse de fato plana, não veríamos o barco desaparecer dessa forma¹. Já na Grécia antiga estimou-se o raio da Terra com precisão admirável. Não é meu objetivo aqui fazer uma descrição extensiva das diversas teorias sobre espaço que existiram na história. Para isso, veja o excelente texto de Max Jammer² [Liv:Jam10a]. Esse trabalho tem como foco a teoria da Relatividade Geral, proposta por Einstein³ em 1915 [Art:Ein15]. Para que a introdução das idéias de tal teoria sejam mais naturais é necessário uma perspectiva histórica. Esse capítulo tem como objetivo expor alguns fatos históricos importantes e exibir a evolução de certas idéias físicas que têm influência direta no conceito de espaço até chegarmos na Teoria da Relatividade Restrita. Faremos então uma pausa na exposição história para ver alguns conceitos de tal teoria, essenciais para a sua posterior generalização. Não seguirei exatamente a ordem cronológica e não apresentarei os conceitos na linguagem que era usada na época, mas sim os traduzirei em termos matemáticos modernos. Também não é meu objetivo nesse primeiro capítulo ser extremamente preciso nas definições rigorosas dos conceitos. Tal rigor tornaria a discussão obscura e não deixaria transparecer a essência filosófica da evolução do conceito de espaço-tempo. Esse capítulo é amplamente baseado nos excelentes textos [Liv:Pen05] e [Liv:Fey08].

1.1 A primeira idéia de espaço-tempo - Da Antiguidade até Galileu

A primeira idéia que surge ao pensarmos em espaço é que vivemos em um espaço euclidiano tridimensional e que o tempo flui continuamente de acordo com um parâmetro. Essa idéia pode ser traduzida matematicamente da seguinte forma: vivemos em \mathbb{R}^3 e o tempo está em \mathbb{R} . Dessa forma, o espaço-tempo é dado pelo produto cartesiano $\mathbb{R} \times \mathbb{R}^3$. Com esse modelo, espaço e tempo são entidades totalmente distintas e uma não influencia na outra. Isso tem consequências extremamente importantes. Vejamos algumas delas.

¹Como o alcance da nossa visão é finito, veríamos o barco ficando cada vez menor, até que de tão pequeno não o veríamos mais. Ou veríamos o barco sumindo todo de uma só vez, e os navegadores teriam o trágico destino de serem engolidos pelo monstro que supostamente habitava o Universo depois das fronteiras da Terra. Como em muitas vezes os navegadores iam e voltaram, essa segunda hipótese não é nada razoável.

²Max Jammer, Berlin 13/04/1915 - Jerusalem 18/12/2010

³Albert Einstein, Ulm 14/03/1879 - Princeton 18/04/1955

1.2. O PRINCÍPIO DA RELATIVIDADE DE GALILEU E SUA INFLUÊNCIA NA ESTRUTURA DO ESPAÇO-TEMPO

Primeiramente, façamos a seguinte analogia: o espaço \mathbb{R}^3 é a tela de um cinema e o movimento dos objetos, as ações que acontecem, é o filme sendo projetado na tela. É isso que essa estrutura de espaço-tempo quer dizer. O espaço é um ente absoluto, sem nenhuma interação não-trivial com o tempo. Repare que se fixarmos um ponto na tela, à medida que o filme passa faz sentido dizer que tal ponto se manteve no mesmo lugar. Analogamente, nessa estrutura de espaço-tempo faz sentido dizer que um certo ponto em \mathbb{R}^3 é o mesmo após um certo instante de tempo. Em outras palavras, a noção de “estar parado” é um conceito absoluto, não depende do observador.

É importante introduzir o conceito de *evento* pois ele será usado durante todo o trabalho. Nesse caso, dizemos que um *evento* é um ponto de $\mathbb{R} \times \mathbb{R}^3$. A princípio isso pode parecer uma definição sem graça, pois estamos somente dando um nome diferente para um objeto bem conhecido, porém o nome “ponto no espaço-tempo” é somente um termo geométrico sem significado físico. Essa definição idealiza o conceito de algo que acontece em algum ponto do espaço com uma duração de tempo muito pequena, como uma explosão ou uma colisão entre partículas.

Note que existe também uma noção de simultaneidade absoluta. Dizemos que dois eventos (t_x, x_1, x_2, x_3) e (t_y, y_1, y_2, y_3) são *simultâneos* se $t_x = t_y$. Temos também a noção de *intervalo de tempo* e *distância entre eventos simultâneos* bem definidos: o primeiro é dado por $|t_x - t_y|$ e o segundo por $|(x_1, x_2, x_3) - (y_1, y_2, y_3)|$, onde usamos o mesmo símbolo $|\cdot|$ para denotar a distância em \mathbb{R} e \mathbb{R}^3 . Todos os habitantes desse espaço-tempo concordam em relação a essas quantidades.

Vale ressaltar que a primeira teoria axiomática de espaço surge na Grécia Antiga com Aristóteles⁴ em seu livro *Physica*[Cls:Ari], uma série de tratados onde se discutem os conceitos filosóficos das ciências naturais.

1.2 O Princípio da Relatividade de Galileu e sua influência na estrutura do espaço-tempo

Em 1632, Galileu⁵ introduz o chamado *Princípio da Relatividade* [Cls:Gal1632], cujo conteúdo é o seguinte: “As leis do movimento são as mesmas em qualquer referencial que se move com velocidade constante.” Galileu introduz esse princípio fazendo uma analogia com o movimento de um barco. Façamos algo semelhante: imagine que você está em um avião que se move em velocidade uniforme e em linha reta, em relação à Terra, por exemplo. Se você não olhar pela janela, você não sentirá velocidade alguma. É como se você estivesse parado. Isso quer dizer que não existe um observador universal privilegiado, em relação ao qual podem ser medidas todas as velocidades absolutas. Se estamos dentro de um objeto movendo-se com velocidade constante, não podemos realizar nenhum experimento que nos permita dizer qual é a nossa velocidade.

Com essa observação de Galileu, a noção anterior de espaço-tempo não é mais adequada, pois nela existe um observador privilegiado, por exemplo, alguém parado no ponto $0 \in \mathbb{R}^3$ e em relação a ele podemos medir velocidades absolutas. Todos os observadores concordarão com essas medições, já que as distâncias temporais e espaciais são conceitos absolutos. Como consequência, existe a noção universal de “estar parado”. Precisamos então de um novo conceito de espaço, um no qual não exista uma noção universal de “estar parado” nem um observador universal privilegiado. Se analisarmos certos fatos do cotidiano mais atentamente, veremos que essa noção realmente não existe. Para ver isso, fixe um ponto na superfície da Terra e espere por 1 minuto. Podemos dizer onde está esse ponto agora? Podemos pensar em considerar a rotação da Terra em torno de seu próprio eixo e calcular onde tal ponto estaria. Mas isso não resolve, pois não estamos considerando o movimento de translação da Terra. Mesmo se levarmos tal movimento em conta nos nossos cálculos estaríamos esquecendo do movimento de rotação

⁴Aristóteles, *Stageira* 384 a.C. - *Euboea* 322 a.C.

⁵Galileu Galilei - *Pisa* 15/02/1564 - *Florença* 08/01/1642

da Via Láctea em torno do seu centro, seu movimento de translação em relação ao *cluster* no qual ela se localiza, e assim por diante. Porém, se considerarmos um certo observador particular, podemos dizer que um ponto se manteve parado em relação àquele observador durante um certo intervalo de tempo. Por exemplo, agora enquanto estou digitando esse trabalho posso dizer que meu computador se manteve parado em relação a mim durante um pequeno intervalo de tempo, digamos, uma hora⁶. É por isso que esse princípio leva tal nome: não faz mais sentido em falar de noções universais, mas sim somente relativas a observadores. Agora não temos mais *um* observador universal privilegiado como anteriormente, mas sim uma *classe* deles, os chamados *observadores inerciais*: aqueles em relação aos quais as leis do movimento sempre são as mesmas, observadores que se movem em linha reta e com velocidade constante uns em relação aos outros.

Assim, nossa noção de espaço-tempo toma outra forma. Como não temos mais uma noção global de “estar parado”, podemos dizer que para cada instante de tempo t temos um \mathbb{R}_t^3 distinto, que não se relaciona de maneira natural com \mathbb{R}_s^3 , se $t \neq s$. Em termos matemáticos modernos, nosso espaço-tempo torna-se um fibrado vetorial cuja base é \mathbb{R} e fibra é \mathbb{R}^3 . Note que $\mathbb{R} \times \mathbb{R}^3$ também é um fibrado com essas características, porém ele é *trivial*, ou seja, podemos escrevê-lo como o produto do espaço-base \mathbb{R} pela fibra \mathbb{R}^3 . Agora estamos admitindo fibrados mais gerais, que podem eventualmente não serem triviais. Por enquanto não sabemos quase nada sobre a estrutura desse fibrado. Na próxima seção vamos inferir algumas conclusões a partir das Leis de Newton.

1.3 As Leis de Newton e suas influências na estrutura do espaço-tempo

Newton⁷ é considerado um dos cientistas mais influentes de todos os tempos. Sua obra principal, *Philosophiae Naturalis Principia Mathematica* [Cls:New1687], é também considerada uma das obras mais influentes da ciência⁸, pois nela são apresentadas as 3 Leis de Newton e a Lei da Gravitação Universal, as bases da mecânica clássica. A Lei da Gravitação Universal será nosso assunto em breve. Por enquanto, vamos olhar somente para as Leis de Newton, publicadas originalmente em Latim na seguinte forma:

- *Lex I: Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare.*
- *Lex II: Mutationem motus proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimitur.*
- *Lex III: Actioni contrariam semper et aequalem esse reactionem: sive corporum duorum actiones in se mutuo semper esse aequales et in partes contrarias dirigi.*

Em Português e em termos modernos, temos:

- 1a. Lei: “O movimento de um corpo que não está sob a ação de nenhuma força é em linha reta e com velocidade uniforme (chamado de *movimento inercial*)”;

⁶Eu poderia começar aqui uma longa discussão sobre o que podem significar os termos “pequena porção do espaço” e “pequeno intervalo de tempo”. Conceitos como *pequeno*, *grande*, *longe*, *perto*, *rápido* são ambíguos e variam de caso para caso. No nosso exemplo, quando eu olho para meu quarto durante uma hora isso pode ser considerado minúsculo em relação a escala do Universo. Mas pode ser também uma imensidão e uma eternidade quando olhamos na escala sub-atômica.

⁷Sir Isaac Newton, Woolsthorpe-by-Colsterworth 04/01/1643 - Londres 31/03/1727

⁸A obra considerada como mais influente de toda a história da ciência são os Elementos, de Euclides (Euclides de Alexandria, ±300 a.C - desconhecido). Composto de 13 livros e escrito por volta de 300 a.C., é o primeiro tratamento axiomático de uma teoria e foi usado como referência no estudo de geometria durante mais de 20 séculos! O próprio Newton e outros grandes cientistas estudaram geometria pelo tratado de Euclides. Uma curiosidade é que essa obra só tem menos edições publicadas do que a Bíblia, que chega ao número impressionante de 1000 edições. Existe, inclusive, uma versão em português dos Elementos [Liv:Euc09].

1.4. A LEI DA GRAVITAÇÃO UNIVERSAL

- 2a. Lei: “A força exercida em um corpo é diretamente proporcional a sua aceleração, essa constante de proporcionalidade sendo a massa do corpo. Mais precisamente, $F = ma$ ”;
- 3a. Lei: “A força que um corpo A faz em um corpo B é igual em magnitude e contrária em direção à força que B faz em A”.

Começamos nossa análise com a 1a. Lei. No seu enunciado, está escrito o termo “linha reta”. Isso quer dizer que devemos saber dizer o que é uma linha reta no espaço-tempo. Temos diversas maneiras equivalentes de fazer isso. Duas possibilidades são considerar o espaço-tempo como um fibrado vetorial como acima e munido de uma conexão livre de torção e curvatura⁹ ou considerá-lo como um espaço afim de dimensão 4. É melhor pensarmos com conexão, pois será mais fácil na hora que precisarmos generalizar. Tendo então uma conexão, podemos considerar as geodésicas dela. Se considerarmos somente as geodésicas que são seções (pelo menos locais) do fibrado, teremos, rigorosamente, os movimentos inerciais enunciados na 1a. Lei.

A 2a. Lei versa sobre força e aceleração. Podemos dizer que as seções do fibrado que representa o espaço-tempo que não são geodésias são os *movimentos acelerados*. Mas que sorte! Podemos calcular a aceleração de uma curva usando a conexão. Juntanto essa informação com a 2a. Lei, podemos dizer que a conexão nos dá, de certa forma, a força agindo sobre o corpo em um movimento não-inercial¹⁰.

A 3a. Lei supõe, implicitamente, uma hipótese extremamente profunda sobre o comportamento das forças. Repare que para que seu enunciado faça sentido a força que A faz em B deve ser sentida imediatamente por B. Caso contrário, não poderíamos falar que a força que B faz em A é contrária em direção, pois em um instante seguinte não teríamos como dizer qual era a posição de A no momento que a força foi aplicada. Dessa forma, o enunciado da 3a. Lei supõe que as forças são sentidas imediatamente.

1.4 A Lei da Gravitação Universal

A Astronomia é uma das ciências mais antigas. Desde tempos remotos, as civilizações estudavam os movimentos dos planetas e outros objetos que conseguiam observar no céu. Tão importante era a importância dos objetos celestes para os antigos que seus monumentos eram construídos de acordo com as posições de estrelas, como as Pirâmides do Egito e o Stonehenge, na Inglaterra. Existem registros astronômicos Babilônicos de aproximadamente 1200 a.C. Os antigos já sabiam que os planetas giravam em torno do Sol, fato redescoberto por Copérnico¹¹ muito mais tarde. Porém, não se sabia *como* era esse movimento. No início do século XV existiam grandes debates sobre se eles realmente giravam em torno do Sol ou não. Tycho Brahe¹² teve uma idéia diferente: no lugar de prosseguir com argumentos filosóficos ele resolveu fazer diversas medições precisas da posição dos planetas no céu em seu observatório na ilha de Hven, perto de Copenhagen. Com isso ele esperava obter novas informações e a partir delas estabelecer outro ponto de vista sobre o movimento dos planetas. Brahe fez inúmeras medições da posição de Marte ao longo de vários anos. Certo tempo depois, Kepler¹³ estudou esses dados e a partir deles deduziu três leis, hoje conhecidas como *Leis de Kepler*, sobre o movimento planetário.

A primeira conclusão de Kepler foi que os planetas giravam em torno do Sol não em uma órbita circular, mas sim elíptica, onde o Sol está em um dos focos. A segunda conclusão foi que a velocidade dos planetas não é constante, mas que varia de acordo com a sua distância ao Sol: se o planeta estiver mais perto, maior é sua velocidade e se estiver mais longe, menor é a velocidade. Além disso, essa variação de

⁹Por motivos técnicos que não vem ao caso no momento.

¹⁰Já temos aqui algum protótipo da relação entre a geometria de um espaço e os movimentos não acelerados, analogia deveras importante na formulação de Teoria da Relatividade Geral. Veja o Capítulo 4 para mais detalhes, ou tenha paciência para chegar até lá. O segundo caminho certamente é mais emocionante.

¹¹Nicolaus Copernicus, Torun 19/02/1473 - Frombork 24/05/1543.

¹²Tyge Ottesen Brahe, Scania 14/12/1546 - Praga 24/10/1601

¹³Johannes Kepler, Weil der Stadt 27/12/1571 - Regensburg 15/11/1630.

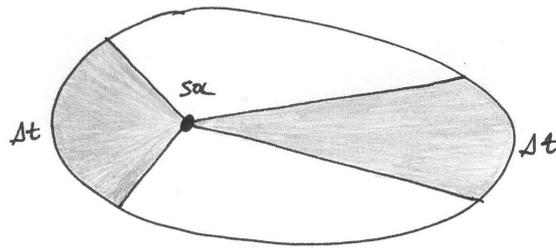


Figura 1.1: Segunda Lei de Kepler.

velocidade satisfaz uma propriedade bem interessante: suponha que sejam feitas diversas observações da posição de um planeta com um certo intervalo de tempo, digamos, Δt . Então o vetor que vai do Sol até o planeta varrerá a mesma área, em cada intervalo de tempo Δt , como na Figura abaixo:

As duas observações acima foram publicadas em 1609 [Cls:Kep1609] e são as duas primeiras Leis de Kepler. Elas podem ser enunciadas como:

- 1a. Lei: “Cada planeta se move ao redor do Sol em uma elipse, com o Sol em um dos focos”.
- 2a. Lei: “O vetor que vai do Sol ao planeta varre áreas iguais em intervalos de tempos iguais”.

10 anos depois, Kepler descobriu e publicou a terceira lei [Cls:Kep1619]. Ela diz que:

- 3a. Lei: “O quadrado do período de um planeta é proporcional ao cubo do semi-eixo maior da elipse que é sua órbita”,

onde o *período* de um planeta é o tempo que ele leva para dar uma volta completa ao redor do Sol.

Com essas leis, sabemos *como* os planetas orbitam o Sol, porém ainda não sabemos o que causa esse movimento. Antigamente se acreditava que havia anjos batendo as asas e empurrando a Terra! Como temos em mãos as Leis de Newton, podemos chegar em uma conclusão mais razoável. Elas nos dizem que um objeto que não está sob ação de nenhuma força se move em linha reta e com velocidade constante¹⁴ e que a única forma de mudar a direção do movimento de um objeto é aplicando nele uma força. Em outras palavras, se um objeto acelera, então uma força foi aplicada na direção da aceleração; por outro lado, se a direção do movimento muda, então uma força deve ter sido aplicada lateralmente no objeto. Essas considerações nos levam a acreditar que não é uma força tangencial que mantém os planetas em movimento, mas sim uma força radial, pois só há mudança de direção, não de velocidade. A partir dessas idéias, Newton supôs que o próprio Sol deveria ser a fonte dessa força. Após uma análise cuidadosa das Leis de Kepler, Newton observou que quanto mais afastado do Sol está o planeta menor é a força que age sobre ele e que essa taxa de decaimento é dada pelo quadrado da distância do planeta ao Sol. Porém, essas observações são feitas somente para planetas orbitando em torno do Sol. Apesar disso, Newton propôs que quaisquer dois corpos que têm massa se atraem e enunciou a famosa lei:

$$F = G \frac{m_1 m_2}{r^2},$$

onde m_1 e m_2 são as massas dos corpos em questão e G é uma constante hoje conhecida como a *constante gravitacional*. Essa generalização não é tão gratuita assim, tendo em vista que já se conhecia que outros

¹⁴Novamente, note que essa Lei não diz *porque* tal movimento acontece. Ela apenas nos diz *como* que ele acontece. Vale ressaltar que isso acontece com mais frequência do que se imagina: inúmeras vezes na Física sabemos dizer *como* um certo fenômeno se comporta, mas não sabemos dizer sua natureza, ou *porque* ele se comporta dessa forma. Apenas deduzimos as equações por argumentos físicos ou matemáticos e verificamos que elas são válidas, realizando experimentos.

planetas tinham suas próprias luas. Nada mais justo que os planetas sejam a fonte da força que as mantém girando ao seu redor!

A relação entre a Lei da Gravitação Universal e as Leis de Kepler é incrível. Em qualquer livro de Mecânica Clássica é possível encontrar uma dedução das Leis de Kepler a partir da Lei da Gravitação Universal. Veja por exemplo o excelente livro de Mecânica do Spivak [Liv:Sp10]. É possível deduzir também a Lei da Gravitação Universal a partir das Leis de Kepler. Isso é mais complicado e pode ser encontrado em [Liv:Tun07].

Agora temos uma nova Lei em mãos. Para sabermos se ela é útil, devemos conseguir extrair dela mais informação do que já temos. Ela também não pode “falhar” em nenhum teste: se encontrarmos um caso no qual ela seja falsa, toda essa teoria está errada e precisa ser corrigida. Isso quase aconteceu quando Rømer¹⁵ estudava as luas de Júpiter. Apesar dos seus estudos serem anteriores à Lei da Gravitação Universal, eles poderiam ser usados para contestá-la. Observando os eclipses de tais luas, em particular de Io, por ser de mais fácil visualização estando na Terra, ele notou que o intervalo de tempo entre dois eclipses simultâneos de Io variavam de acordo com a posição da Terra em relação a Júpiter. Se a Terra estivesse mais perto, o tempo era menor, e se estivesse mais longe, era maior. Isso poderia não estar de acordo com a Lei da Gravitação Universal, pois a força gravitacional exercida por Júpiter sobre Io é muito maior do que a da Terra, implicando que na fórmula da força sobre Io o único termo relevante vem da influência de Júpiter. De onde viria então esse fenômeno observado? Rømer notou então que a razão para esse desvio era muito simples: a luz leva mais tempo para chegar em nós quando Júpiter está longe da Terra do que quando está perto, e com isso obteve uma primeira estimativa da velocidade da luz.

Como foi dito acima, a Lei da Gravitação Universal diz que quaisquer dois corpos que têm massa se atraem. Portanto, ao estudar o movimento dos planetas ao redor do Sol não devemos levar em conta apenas a atração do Sol sobre os planetas, mas também a força entre os planetas. Isso faz com que a órbita de um planeta não seja uma elipse perfeita. Porém, como o Sol é muito mais massivo que qualquer planeta do Sistema Solar, podemos, em certos casos, aproximar a órbita por uma elipse. Tentativas foram feitas para analisar os movimentos de Júpiter, Saturno e Urano com base na Lei da Gravitação Universal. Levou-se em consideração os efeitos gravitacionais de cada um no outro para verificar-se se as irregularidades nas suas órbitas podiam ser respondidas somente com tal Lei. Para Júpiter e Saturno, tudo funcionou. Porém, os dados de Urano ainda não batiam com as previsões. Isso podia ser o fim da nossa Lei. Mas, felizmente, Adams¹⁶ e Le Verrier¹⁷, na Inglaterra e na França respectivamente, supuseram de forma independente que deveria existir outro planeta, que estaria causando tais desvios na órbita de Urano. Eles calcularam onde tal planeta deveria estar para causar esses desvios e enviaram cartas aos seus observatórios, indicando a posição do céu no qual esse planeta deveria ser encontrado. Felizmente de novo, a mensagem deles não foi ignorada e foi encontrado Netuno. Vale ressaltar outra participação de Galileu na nossa história: seus desenhos datados de 28 de Dezembro de 1612 e 27 de Janeiro de 1613 mostram que ele “descobriu” Netuno, porém, o classificou como uma estrela.

Temos então evidências de que a Lei da Gravitação Universal é válida para o Sistema Solar. Temos também evidências da existência da gravidade em escalas maiores de distância, como os sistemas binários (sistemas com duas estrelas, uma orbitando a outra), as galáxias, os aglomerados de galáxias e outras estruturas muito grandes encontradas no Universo [Liv:Ast08]. Note que ainda não falamos o *porquê* da gravidade, mas somente como ela se comporta em pequenas distâncias. Isso ainda é um problema em aberto na Física.

¹⁵Ole Christensen Rømer, Århus 25/09/1644 - Copenhagen 19/09/1710

¹⁶John Couch Adams, Laneast 05/06/1819 - Cambridgeshire 21/01/1892

¹⁷Urbain Le Verrier, Saint-Lô 11/03/1811 - Paris 23/09/1877

1.5 O Princípio da Equivalência

O Princípio da Equivalência é uma observação profunda sobre a natureza da gravidade. Provavelmente esse foi o primeiro passo no caminho de uma compreensão mais clara sobre seus mecanismos, a idéia essencial sendo de Galileu, apesar de existirem precursores, como Stevin¹⁸ em 1586 e outros mais cedo, como Ioannes Philoponus¹⁹ nos séculos V e VI. É um fato bem conhecido no folclore científico o experimento que Galileu supostamente realizou na Torre de Pisa: ele deixou cair dois objetos de massas diferentes e notou que eles caíam da mesma forma²⁰. Não se sabe ao certo se ele de fato realizou esse experimento, mas a idéia de que, se não levarmos em consideração os efeitos de atrito do ar, a gravidade age de forma igual em corpos de massa distinta é extremamente profunda. Como temos a 2a. Lei de Newton e a Lei da Gravitação Universal em mãos, podemos estudar essa observação de Galileu de uma forma mais precisa. Suponha que um corpo de massa m é atraído por um de massa M . A 2a. Lei nos diz que $F = ma$, enquanto que a Lei da Gravitação Universal nos diz que $F = GmM/r^2$. Podemos então igualar as duas equações e obter que

$$\frac{GmM}{r^2} = ma \Rightarrow a = \frac{GM}{r^2}.$$

Em palavras, um corpo de massa M gera um certo campo gravitacional que faz uma força sobre um corpo de massa m . Essa força causa uma aceleração em m que é *independente* de m . Isso é exatamente a “demonstração” do resultado obtido por Galileu em seu suposto experimento. Repare que isso é quase uma conspiração da natureza: as duas Leis usadas dependem da mesma forma da massa m .

É uma pena que esse raciocínio elegante tenha uma pequena falha. Nós deixamos passar um pequeno detalhe. Na 2a. Lei de Newton, a massa m que aparece é a chamada *massa inercial*, que, essencialmente, mede a dificuldade de se colocar um corpo em movimento. A massa que aparece na Lei da Gravitação tem dois papéis. Quando consideramos a massa m sob ação do campo gravitacional gerado pela massa M , o número m representa agora a capacidade do objeto sentir tal campo e é chamada de *massa gravitacional passiva*. Quando olhamos para a massa M no campo gerado por m , o número m representa agora a capacidade do objeto de produzir campo gravitacional e é chamada, naturalmente, de *massa gravitacional ativa*²¹. A 3a. Lei de Newton nos diz que as massas gravitacionais ativas e passivas são iguais e ambas são então chamadas de *massa gravitacional*. Porém, a princípio, nada nos garante que as massas gravitacional e inercial são iguais. Caso não o fossem, não poderíamos deduzir a igualdade acima. Vamos então fazer o processo inverso: como podemos verificar experimentalmente que o movimento de um corpo em um campo gravitacional não depende da natureza do campo, mas sim somente de sua posição e velocidade inicial, tomamos isso como um princípio:

- Princípio da Equivalência: “O movimento de uma massa puntiforme em um campo gravitacional depende somente de sua posição e velocidade inicial e é independente da natureza da massa”.

Com isso, podemos provar que as massas inercial e gravitacional são proporcionais. De fato, sejam m_1^i e m_2^i as massas inerciais de dois objetos, m_1^g e m_2^g suas respectivas massas gravitacionais. Se um corpo de massa m_0 produz um campo gravitacional, temos que

$$\text{força na massa 1 devido ao campo gerado pela massa 0} = F_1 = G \frac{m_0 m_1^g}{r^2},$$

$$\text{força na massa 2 devido ao campo gerado pela massa 0} = F_2 = G \frac{m_0 m_2^g}{r^2}.$$

¹⁸Simon Stevin, 1548/1549 - 1620. Pouco se conhece sobre sua vida.

¹⁹Ioannes Philoponus, 490 - 570. Também se conhece pouco sobre sua vida, mas bastante se conhece sobre sua obra.

²⁰Aristóteles tinha uma visão diferente. Ele acreditava que corpos mais massivos caíam mais rápido.

²¹Para uma discussão sobre a evolução histórica do conceito de massa veja [Liv:Jam10b].

1.5. O PRINCÍPIO DA EQUIVALÊNCIA

Da 2a. Lei de Newton, temos que $F_j = m_j^i a_j$, $j = 1, 2$. Se ambas as massas 1 e 2 estão à mesma distância r da massa 0, pelo Princípio da Equivalência, $a_1 = a_2$ e, portanto,

$$\frac{F_1}{m_1^i} = \frac{F_2}{m_2^i} \Rightarrow G \frac{m_0 m_1^g}{m_1^i r^2} = G \frac{m_0 m_2^g}{m_2^i r^2} \Rightarrow \frac{m_1^g}{m_1^i} = \frac{m_2^g}{m_2^i}.$$

Se a boa lógica de Philoponus o levou a realizar algum experimento como o suposto experimento de Galileu na Torre de Pisa, tal experimento não tinha uma grande precisão. Ao longo da história vários experimentos têm sido realizados para verificar o Princípio da Equivalência. Em 1999, usando balança de torção, tal princípio verificou-se com uma precisão de 5×10^{-13} ! Veja [Art:Bae99]. Experimentos ainda mais precisos são planejados para o futuro.

Podemos então reescrever a igualdade ingênua acima como

$$a = K \frac{GM}{r^2},$$

onde K é uma constante universal de proporcionalidade entre as massas gravitacionais e inerciais. Escolhendo unidades adequadas, podemos tomar $K = 1$, mas não vamos nos preocupar com esse detalhe aqui. Para todos os efeitos, vamos simplesmente tomar $K = 1$. O significado dessa igualdade é muito profundo. Imagine que você está em queda livre em um planeta de massa M . Então você sentirá uma aceleração dada pela fórmula acima. Imagine agora que você está em uma nave espacial que está acelerando com a mesma aceleração dada por essa fórmula. Se você estiver de olhos fechados, não poderá distinguir qual é a natureza dessa aceleração. Ela pode ser proveniente tanto da atração gravitacional de um planeta de massa M quanto do movimento da nave espacial.

Como a aceleração está relacionada com a conexão e, portanto, com a geometria, isso nos sugere que efeitos gravitacionais, por serem da mesma natureza que aceleração, também estão relacionados com a geometria. Poderíamos continuar essa discussão envolvendo o Princípio da Equivalência e estudar como ele modifica a estrutura do espaço-tempo. Seria uma discussão muito elegante, porém, nos levaria quase diretamente à Teoria de Relatividade Geral de Einstein. Vamos um pouco mais devagar e falar um pouco sobre a Teoria da Relatividade Restrita.

Capítulo 2

A Teoria da Relatividade Restrita

Nesse capítulo vamos discutir alguns conceitos básicos da Teoria da Relatividade Restrita, apresentando também os fatos históricos que levaram ao seu surgimento. Não seguiremos o caminho puramente histórico pois ele sempre é o mais duro, porém, minha abordagem será mais ou menos construtiva, introduzindo os axiomas da teoria após entender o porquê deles. Este capítulo foi baseado em [Liv:Fey08], [Liv:Nab88], [Liv:Whe92], [Liv:Woo07] e [Liv:Wal92]. Algumas partes desse capítulo e do próximo podem parecer semelhantes, porém são feitas de pontos de vista distintos. Nesse capítulo tornamos rigoroso alguns aspectos da Relatividade Restrita, ao passo que no capítulo seguinte veremos isso de um ponto de vista geométrico. Porém, as idéias físicas discutidas aqui são essenciais para um bom entendimento futuro.

2.1 As Equações de Maxwell

Apesar de toda essa discussão sobre teorias da gravitação, Leis de Newton, Kepler, Princípio de Equivalência, e outras coisas fortemente ligadas ao conceito de espaço, a Teoria da Relatividade Restrita não nasceu de nenhum argumento desse tipo, mas sim de considerações sobre fenômenos eletromagnéticos. Em 1861 e 1862 Maxwell¹ publicou uma série de quatro artigos na qual ele enunciou as famosas *Leis de Maxwell do eletromagnetismo* [Cls:Max1862]. Em termos modernos, elas são enunciadas como

$$\begin{aligned}\nabla \cdot E &= \frac{\rho}{\epsilon_0} \\ \nabla \cdot B &= 0 \\ \nabla \times E &= -\frac{\partial B}{\partial t} \\ \nabla \times B &= \mu_0 J + \mu_0 \epsilon_0 \frac{\partial E}{\partial t},\end{aligned}$$

onde E é o campo elétrico, B é o campo magnético, ρ é a densidade de carga, J é a densidade de corrente, μ_0 é a constante magnética ou constante de permeabilidade do espaço e ϵ_0 é a constante elétrica ou a permissividade do espaço. Antes do advento do Cálculo Vetorial não existiam os símbolos ∇ , nem $\nabla \times$. Dessa forma, as equações como postuladas por Maxwell eram em número muito maior e muito mais feias!

No vácuo, ou seja, na ausência de cargas e correntes elétricas, elas tomam a seguinte forma:

¹James Clerk Maxwell, Edinburgh 13/06/1831 - Cambridge 05/11/1879

$$\begin{aligned}\nabla \cdot E &= 0 \\ \nabla \cdot B &= 0 \\ \nabla \times E &= -\frac{\partial B}{\partial t} \\ \nabla \times B &= \mu_0 \varepsilon_0 \frac{\partial E}{\partial t}.\end{aligned}$$

No capítulo anterior, foram enunciados o Princípio de Relatividade de Galileu e as Leis de Newton. Como as Leis de Newton são equações que descrevem completamente a dinâmica dos corpos, devemos ser capazes de enunciarmos o Princípio da Relatividade em termos das Leis de Newton. Para fazer isso, considere dois sistemas de coordenadas, um se movendo em relação ao outro com uma certa velocidade u na direção x , como mostrado na figura abaixo:

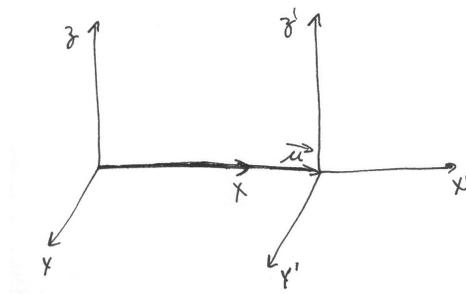


Figura 2.1: Referenciais em movimento relativo.

Se x, y, z são as coordenadas do sistema “parado” e x', y', z' as coordenadas do sistema “em movimento”², podemos relacioná-las pelas transformações

$$\begin{aligned}x' &= x - ut \\ y' &= y \\ z' &= z \\ t' &= t.\end{aligned}$$

Portanto, ao fazermos essa transformação na 2a. Lei de Newton, obtemos a mesma equação. É dessa forma que o fato das leis de movimentos serem as mesmas em qualquer referencial inercial se expressa matematicamente. Deveríamos esperar que o mesmo acontece com as Equações de Maxwell. Porém, ao fazermos tal transformação nessas equações notamos que elas *não* tomam a mesma forma. A princípio pensou-se que as Equações de Maxwell estavam erradas, pois elas eram muito recentes, enquanto que as Leis de Newton tinham mais de 200 anos e já tinham sido verificadas experimentalmente diversas vezes. Tentou-se então modificar as Equações de Maxwell de modo que ela fosse invariante pela transformação acima³, porém, os fenômenos eletromagnéticos previstos pelas novas equações eram incompatíveis com os experimentos. Portanto se as leis do Eletromagnetismo são diferentes quando analisadas em um objeto em movimento e em um objeto parado poderíamos usar essa divergência para determinar a velocidade absoluta de um objeto em movimento.

²Note que usei as palavras *parado* e *em movimento* em um sentido intuitivo. O Princípio da Relatividade não nos permite dizer se existe um estado de repouso absoluto, em relação ao qual os outros referenciais inerciais estão em movimento. Para ser mais preciso, eu deveria dizer parado e em movimento em relação ao sistema de coordenadas x, y, z .

³Um exemplo notável foi a Teoria da Emissão, proposta por W. Ritz (Walther Ritz, Sion 22/02/1878 - Göttingen 07/07/1909) em 1908.

Tal fato contraria o Princípio da Relatividade de Galileu. Porém, modificar o Eletromagnetismo para que ele se “ajuste” à Mecânica Clássica não foi bem sucedido. O que teve que ser feito então é abandonar certas noções clássicas que já estavam bem enraizadas na ciência. Uma delas foi aceitar o fato de que a velocidade da luz é constante e independe do observador inercial. Claro que isso é bastante não-intuitivo, mas Einstein certamente tinha bons motivos para acreditar que isso deveria ser verdade. Mas, a lei de adição de velocidades de observadores inerciais $v'' = v + v'$ é incompatível com essa nova Teoria. Isso nos dá indícios de que essa pequena alteração trará diferenças profundas na nova Teoria, pois mesmo a mais simples das fórmulas deverá ser alterada. Uma maneira de nos convenceremos de que a velocidade da luz é constante é o experimento de Michaelson-Morley, descrito abaixo. Não se sabe se Einstein tinha conhecimento desse resultado quando ele propôs a Teoria da Relatividade Restrita.

2.2 O experimento de Michelson-Morley

Michelson⁴ e Morley⁵, dois cientistas americanos, Michelson sendo o primeiro americano a ganhar o Prêmio Nobel de Física, em 1907, realizaram em 1887 o experimento mais famoso para se determinar a velocidade da Terra em relação ao *éter luminífero*, uma substância que acreditava-se na época que preenchia todo o universo e era responsável pela propagação de ondas eletromagnéticas. Acreditava-se também que poderia-se considerar o éter como um referencial inercial absoluto, em relação ao qual podem-se medir velocidades absolutas. Vou antecipar o final da história falando que o resultado foi negativo, fato que esperou 18 anos até o surgimento da Teoria da Relatividade Restrita para poder ser explicado. O experimento foi realizado diversas vezes por eles com grau de precisão cada vez maior. Vamos descrevê-lo brevemente aqui.

O aparelho é essencialmente o esquema mostrado na Figura 2.2. Ele consiste de uma fonte de luz A , uma lâmina de vidro parcialmente coberta de prata B e dois espelhos C e E , tudo montado sobre uma base rígida. Os espelhos são colocados a distâncias iguais L em relação a B . A lâmina de vidro B divide em dois um feixe de luz que parte de A , que partem em direções perpendiculares aos espelhos, onde são refletidos de volta a B . Ao chegarem em B , os dois feixes são recombinados como dois feixes superpostos, D e F . Se o tempo decorrido para a luz ir de B a E e voltar for o mesmo que o tempo de ir de B a C e voltar, os feixes D e F estarão em fase e reforçarão um ao outro. Caso o tempo seja diferente, eles estarão ligeiramente fora de fase, resultando em uma interferência. Se o aparelho estiver “em repouso” no éter os tempos deveriam ser exatamente iguais, mas caso ele esteja se movendo para a direita com uma velocidade u deveria haver uma diferença nos tempos. Vejamos por quê.

Primeiro calculamos o tempo necessário para luz ir de B a E e voltar. Digamos que o tempo de ida seja t_1 e o tempo de retorno seja t_2 . Agora, enquanto a luz está a caminho de B até E , o aparelho se desloca de uma distância ut_1 , então a luz precisa percorrer uma distância $L + ut_1$, à velocidade c . Essa distância é dada também por ct_1 . Temos então que

$$ct_1 = L + ut_1 \Rightarrow t_1 = \frac{L}{c - u}.$$

De forma análoga, o tempo t_2 pode ser calculado. Durante esse tempo, a lâmina B avança uma distância ut_2 , de forma que a distância de retorno da luz é $L - ut_2$. Portanto,

$$ct_2 = L - ut_2 \Rightarrow t_2 = \frac{L}{c + u}.$$

Dessa forma, o tempo total é

$$t_1 + t_2 = \frac{2Lc}{c^2 - u^2}.$$

⁴Albert Abraham Michelson, Strzelno 19/12/1852 - Pasadena 09/05/1931

⁵Edward Williams Morley, Newark 29/01/1838 - West Hartford 24/02/1923

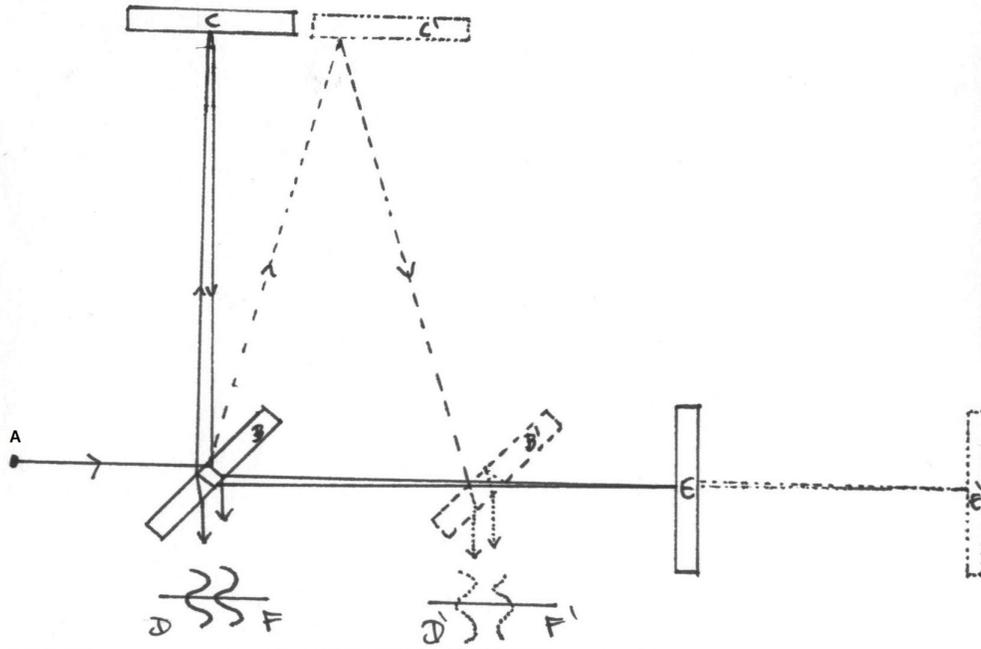


Figura 2.2: Experimento de Michelson-Morley.

Vamos escrever essa igualdade como

$$t_1 + t_2 = \frac{2L/c}{1 - u^2/c^2}, \quad (2.1)$$

pois isso vai ser útil mais a frente.

Calculamos agora o tempo t_3 que a luz leva para ir de B até o espelho C. Como antes, o espelho C move-se para a direita de uma distância ut_3 , até a posição C' . Ao mesmo tempo, a luz percorre uma distância ct_3 ao longo da hipotenusa de um triângulo retângulo, que é BC' . Temos então que

$$(ct_3)^2 = L^2 + (ut_3)^2 \Rightarrow L^2 = c^2t_3^2 - u^2t_3^2 = (c^2 - u^2)t_3^2 \Rightarrow t_3 = \frac{L}{\sqrt{c^2 - u^2}}.$$

Para a viagem de volta de C' até B' a distância é a mesma, portanto, o tempo é também o mesmo. Dessa forma, o tempo total é dado por

$$2t_3 = \frac{2L}{\sqrt{c^2 - u^2}} = \frac{2L/c}{\sqrt{1 - u^2/c^2}}. \quad (2.2)$$

Podemos agora comparar os tempos gastos pelos dois feixes de luz. Nas fórmulas 2.1 e 2.2 os numeradores são idênticos e representam o tempo que decorreria se o aparelho estivesse em repouso. Nos denominadores, o termo u^2/c^2 será pequeno, a não ser que u seja comparável a c . Os denominadores representam as modificações nos tempos causadas pelo movimento do aparelho. Note que essas modificações não são iguais: o tempo para ir até C e voltar é um pouco menor que o tempo para ir até E e voltar, devido ao movimento do aparelho em relação ao éter. Tudo que temos que fazer é medir essa diferença com precisão.

Ao realizarem o experimento, Michelson e Morley orientaram o aparelho de modo que a linha BE estivesse quase paralela ao movimento da Terra em sua órbita. A velocidade orbital da Terra em relação ao

Sol é de aproximadamente $29.000m/s$, e qualquer “corrente do éter” deveria ser pelo menos desse tanto em algum instante do dia ou da noite em algum período do ano. Apesar da diferença de fase entre os feixes D e F ser de difícil detecção, o aparelho que foi usado tinha grande precisão e, como foi dito acima, o experimento foi realizado diversas vezes com grau de precisão cada vez maior. Ainda assim, o resultado foi nulo. A velocidade da Terra através do éter não pôde ser detectada.

Analisando a estrutura do experimento, ele nos permitiria medir variações na velocidade da luz, a medida que a Terra se desloca em relação ao éter. O resultado negativo do experimento nos dá indícios de que a velocidade da luz é constante e não depende do observador. Claro que isso é muito pouco para fazer essa afirmação, mas existem outros experimentos para comprovar esse fato [Art:Bre77].

O que está errado com nosso experimento? A primeira idéia para explicar esse resultado foi de FitzGerald⁶ e H. Lorentz⁷ em 1889 e 1892, respectivamente. Eles sugeriram que os corpos materiais se contraem quando se movem e que essa contração é apenas na direção do movimento. Mais ainda: se L_0 é o comprimento em repouso, então quando ele se mover a uma velocidade u paralela ao seu comprimento, o novo comprimento será dado por

$$L_{//} = L_0 \sqrt{1 - u^2/c^2}.$$

Quando levamos em conta essa contração no experimento de Michelson-Morley, temos que a distância BC permanece inalterada e, portanto, a Equação 2.2 não se altera. Por outro lado, a distância de B até E torna-se $L \sqrt{1 - u^2/c^2}$. A equação 2.1 torna-se então

$$t_1 + t_2 = \frac{(2L/c) \sqrt{1 - u^2/c^2}}{1 - u^2/c^2} = \frac{2L/c}{\sqrt{1 - u^2/c^2}}.$$

Temos então que $t_1 + t_2 = 2t_3$, o que justifica o resultado negativo do experimento.

É difícil aceitar a sugestão de FitzGerald-Lorentz, devido ao seu caráter artificial e a ela ter sido inventada com o propósito de explicar a dificuldade. Mas se olharmos com cuidado, ela não é tão absurda assim: as forças que agem na escala microscópica de um corpo rígido são majoritariamente eletromagnéticas, forças essas que parecem ser afetadas por movimentos retilíneos em velocidade uniforme, já que as Equações de Maxwell não são invariantes por transformações de Galileu. Então não seria tão absurdo dizer que tais forças seriam afetadas pelo movimento através do éter. Por outro lado, a natureza parecia estar conspirando contra os cientistas introduzindo algum fenômeno novo que impeça a medição de u , pois foram realizados muitos outros experimentos e todos davam um resultado nulo. Por fim, Poincaré⁸ observou que *uma conspiração total é por si própria uma lei da natureza* e propôs que não é possível descobrir o movimento no éter por meio de nenhum experimento.

2.3 A dilatação temporal

Do mesmo jeito que vimos que a medição de comprimentos muda de acordo com a velocidade, podemos deduzir que os intervalos de tempo também mudam. Esse é um fenômeno *muito* estranho e devemos discuti-lo com cuidado. Imagine que realizamos o experimento de Michelson-Morley de outra forma. Suponha que o aparelho está em repouso em um foguete movendo-se a uma velocidade u e temos dois observadores: um em repouso e outro dentro do foguete. Para o observador que está viajando, o tempo que a luz leva para ir de B até C e voltar é $2L/c$. Porém, para o outro observador é $(2L/c) / \sqrt{1 - u^2/c^2}$, como na Equação 2.2. Repare que usamos aqui o fato de que ambos observadores veem a luz com a mesma velocidade c . Caso tivéssemos trocado a luz por outra coisa sem essa propriedade veríamos o

⁶George Francis FitzGerald, Dublin 03/08/1851 - Dublin 22/02/1901

⁷Hendrik Antoon Lorentz, Arnhem 18/07/1853 - Haarlem 04/02/1928

⁸Jules Henri Poincaré, Nancy 29/04/1854 - Paris 17/07/1912.

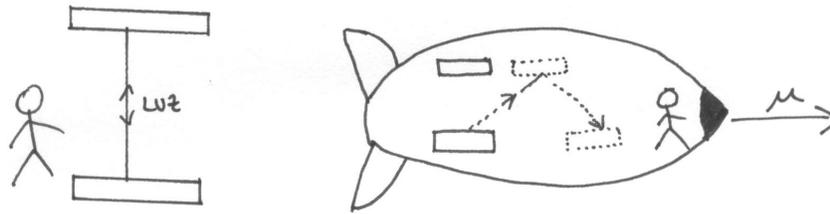


Figura 2.3: Relógios no foguete.

mesmo tempo. Essa discrepância entre os tempos quer dizer que o observador externo vê o que se passa no foguete de uma forma mais devagar. Todas as ações do homem do foguete parecem mais lentas do que o normal para o homem fora do foguete. Porém, o observador viajante não nota isso. Para ele o tempo decorre normalmente.

Para entender melhor esse fenômeno, vamos criar um relógio e ver como é seu mecanismo em movimento. Vamos considerar o relógio mais simples que podemos. Ele é composto por uma régua graduada com um espelho em cada uma das extremidades. Quando iniciamos um sinal luminoso entre os espelhos, a luz vai indo e voltando, e cada vez que a luz é refletida consideramos isso como o “tique-taque” de um relógio comum. Construímos dois relógios idênticos e os sincronizamos⁹ para iniciarem juntos. Dessa forma, eles sempre concordarão, pois eles têm o mesmo comprimento e a luz sempre viaja na mesma velocidade. Um desses relógios fica em um foguete com um observador e ele o posiciona na direção perpendicular ao movimento, de modo que a contração de FitzGerald-Lorentz não altere o tamanho da régua. O outro relógio fica com um observador fora do foguete (veja a Figura 2.3). Note que o observador do foguete não pode ver nenhum fenômeno estranho no relógio, caso contrário, ele notaria que está viajando, contradizendo o Princípio da Relatividade. Agora, quando o observador externo olha o relógio do foguete ele vê a luz fazendo um caminho em ziguezague, já que a barra está se movimentando lateralmente o tempo todo. Portanto, a luz de fato leva mais tempo para ir de uma extremidade a outra do relógio em movimento do que no relógio parado, em relação ao observador fora do foguete. Note também que quanto maior for u mais devagar parece que o tempo passa.

Tudo bem. Fizemos essa dedução para *esse* tipo de relógio em especial. Alguém poderia afirmar que esse fenômeno é exclusivo desse relógio e que ele não mede o tempo direito, certo? Não! Considere agora dois relógios quaisquer idênticos. Eles podem ser feitos de rodas e engrenagens, baseados na desintegração atômica, ou de qualquer outro jeito a sua escolha. Sincronize-os de modo que marquem o mesmo tempo e funcionem em sincronia com nosso primeiro relógio. Um desses relógios é levado no foguete, junto com o primeiro. Caso nesse novo relógio não se verifique a dilatação temporal, o observador dentro do foguete saberá que ele está se movendo e poderá, através da discrepância entre os relógios, calcular sua velocidade, contradizendo o Princípio da Relatividade.

Portanto, independente do mecanismo do relógio, nele se verificará a dilatação temporal. Mas se *todos* os relógios em movimento funcionam mais lentamente podemos dizer que o próprio tempo parece passar mais devagar. Todos os fenômenos dentro do foguete, a pulsação do observador, seus processos de pensamento, o tempo que ele leva para se mover, crescer e envelhecer parecem mais lentos para o outro observador.

Uma evidência experimental da dilatação temporal é fornecida pelos mésons μ [Art:Eas91], também conhecidos como múons. Elas são partículas que se desintegram espontaneamente após uma vida média de $2,2 \times 10^{-6}$ segundos. Elas atingem a Terra em raios cósmicos e também podem ser produzidos artificialmente em laboratório. Alguns deles se desintegram na atmosfera, mas o restante só se desintegra

⁹Como fazemos isso? Discutiremos isso em breve.

após encontrar um pedaço de matéria e parar. Os múons viajam em velocidades distintas, alguns deles quase na velocidade da luz. Apesar disso, seu pouco tempo de vida não os permitiria percorrer os 10 km da atmosfera terrestre. Porém eles são encontrados naturalmente em laboratório. Isso é possível por causa da dilatação temporal. Do ponto de vista deles, eles vivem apenas aproximadamente $2\mu s$, mas do nosso ponto de vista eles vivem consideravelmente mais: o suficiente para que eles possam atingir a Terra.

Portanto, podemos verificar experimentalmente o fenômeno da dilatação temporal. E quanto à contração espacial? Que experimento podemos realizar para mostrar que de fato os corpos se contraem quando se movem em relação ao éter? Não podemos usar métodos como o usado por Michelson e Morley, pois a contração é o fenômeno que explica o resultado negativo de tais tipos de experimento; não podemos medir o comprimento do corpo com uma régua pois a régua também estará contraída ao se mover. Então a idéia de FitzGerald-Lorentz depende de uma “distância absoluta”, uma distância medida em relação ao éter. Somente distâncias medidas em relação ao éter estão corretas. Porém nenhum experimento pode detectar o éter! Então como vamos medir distâncias absolutas?

Einstein foi o primeiro a entender essa dificuldade e notar que a solução não está em modificar as Equações de Maxwell, mas sim em re-examinar a noção de distância e, por consequência, de simultaneidade. É necessário darmos *definições operacionais* para os conceitos: definições que nos permitam medir tais conceitos na prática. Isso significa que, por exemplo, se não é possível descrever um experimento para detectar o éter, então ele não tem lugar na teoria. O mesmo vale para medições de tempo e distância. Em particular, não é possível falar de distância enquanto não se decida o que é simultaneidade.

2.4 Princípios Básicos

Primeiramente é bom saber porque do nome Relatividade *Restrita*. Como o nome indica, ela só vale em uma família restrita de situações: aquelas nas quais as medições são feitas em relação a observadores são inerciais e não existe campo gravitacional. O caso geral é, logicamente, tratado na Teoria da Relatividade Geral. Como falei no começo do capítulo, não vou seguir o caminho histórico e vou tentar ser rigoroso matematicamente, porém, sem perder a intuição física. Para uma discussão de um ponto de vista mais físico, veja [Liv:Syn72]. Veja também o artigo original do Einstein [Art:Ein05]. Existe uma tradução para o Inglês bem acessível em [Liv:Lor63].

A Teoria da Relatividade Restrita é construída com base em essencialmente dois princípios básicos:

- *Princípio da Relatividade de Galileu*: Nenhum experimento pode medir a velocidade absoluta de um observador inercial. De maneira equivalente, as leis do movimento são as mesmas em qualquer referencial que se move com velocidade constante;
- *Universalidade da velocidade da luz*: A velocidade da luz relativa a qualquer observador não-acelerado é $c = 3 \times 10^8 m/s$, independente do movimento da fonte da luz relativo ao observador. Mais precisamente, se dois observadores não-acelerados medem a velocidade do mesmo raio de luz cada um encontrará o mesmo valor c , independente do movimento de um relativo ao outro.

Considere um observador inercial O como na Mecânica Clássica, ou seja, se movendo sem aceleração. Suponha que O carregue consigo um relógio¹⁰. Ao se mover O descreve uma curva no espaço-tempo, chamada de sua *trajetória* ou *história*. Sobre sua trajetória O sabe assinalar coordenadas: as coordenadas espaciais são $(0, 0, 0)$ e o tempo é a marcação do relógio que O carrega. E para os pontos que estão fora de sua trajetória? Como O dá coordenadas a esses pontos? Em particular, como O julga se dois eventos

¹⁰Note que não é absurda essa imagem de um observador inercial carregar um relógio, pois um simples átomo pode agir como um relógio. Essa é a base teórica dos *relógios atômicos*. Tais relógios, idealizados por Lord Kelvin (William Thomson, 1o. Barão Kelvin, Belfast 26/06/1824 - Largs 17/12/1907) em 1879 [Liv:Kel79] são os instrumentos de medição de tempo mais precisos que conhecemos hoje em dia.

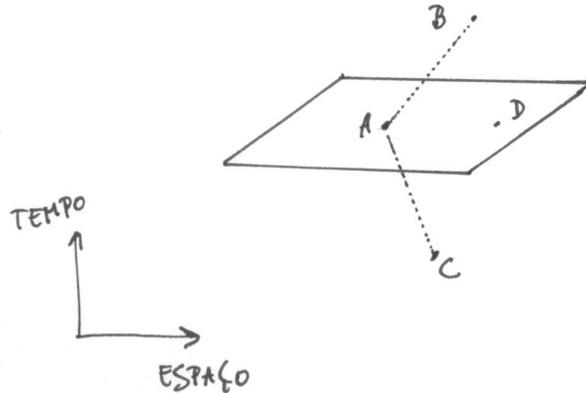


Figura 2.4: Simultaneidade na Mecânica Clássica.

são simultâneos? Na Mecânica Clássica, procedíamos da seguinte forma: como não existe, por hipótese, nenhum limite de quão rápido um observador pode se mover, podemos dizer que dois eventos A e D são simultâneos se não é possível que nenhum observador inercial esteja presente nos dois eventos, ou seja, que sua trajetória contenha os dois eventos. Veja a Figura 2.4.

Note que em nenhum momento usamos observações feitas pelo observador O . Isso quer dizer que a noção de simultaneidade na Mecânica Clássica é absoluta, ou seja, não depende do observador. Dessa forma, conseguimos medir velocidades com facilidade, como está ilustrado na Figura 2.5. Se O e \tilde{O} são observadores inerciais, O consegue medir a velocidade de \tilde{O} medindo a posição de \tilde{O} em instantes distintos. No tempo t_1 , O faz uso da geometria Euclidiana da superfície de simultaneidade para calcular a distância x_1 até \tilde{O} . Em um outro tempo t_2 , O faz o mesmo procedimento e calcula a distância x_2 de \tilde{O} e calcula $v = (x_2 - x_1)/(t_2 - t_1)$. Esse procedimento envolve a noção de simultaneidade acima, pois está implícito que O está medindo a posição de \tilde{O} na superfície de simultaneidade definida por O (que é absoluta e não depende de O) nos instantes de tempo t_1 e t_2 .

Porém, essa noção de simultaneidade não pode ser transportada para a Relatividade Restrita. A universalidade da velocidade da luz nos garante que não podemos nos mover tão rápidos quanto quisermos¹¹. De fato, caso fosse possível nos movermos mais rápido que a luz não veríamos a luz com velocidade c . O que devemos fazer é utilizar esse princípio para definir novamente a noção de simultaneidade. Para isso, como na Figura 2.6, considere dois eventos A e B e seja O um observador inercial. Suponha, por simplicidade, que O passa por A . O quer julgar se A e B são simultâneos. Suponha que um sinal luminoso é emitido de B em todas as direções. Então existe um evento D sobre a trajetória de O no qual tal sinal luminoso encontra sua trajetória. Analogamente, existe um evento C sobre a trajetória de O de modo que se um sinal luminoso é emitido em todas as direções em C , então tal sinal chega ao evento B .

Seja t_1 o tempo medido por O em D subtraído do tempo medido por O em A . Analogamente, seja t_2 o tempo medido por O em A subtraído do tempo medido por O em C . Note que t_1 ou t_2 , mas não ambos, pode ser negativo. Definimos $\Delta t := (t_1 - t_2)/2$ como o *intervalo de tempo* entre A e B e $\Delta x := c(t_1 + t_2)/2$ como a *distância espacial* entre A e B . Portanto, podemos dizer que A e B são simultâneos se $\Delta t = 0$. Se \tilde{O} é outro observador inercial, podemos agora calcular a velocidade relativa de \tilde{O} em relação a O usando o mesmo procedimento acima: nas superfícies de simultaneidades definidas por O nos instantes de tempo t_1 e t_2 calculamos as distâncias x_1 e x_2 de \tilde{O} para O e calculamos $v = (x_2 - x_1)/(t_2 - t_1)$. Note que

¹¹Os *táchions* são partículas hipotéticas que viajariam mais rápido do que a luz. Para uma discussão sobre as consequências disso e sua possível existência veja [Int:Wik] e [Int:Bae].

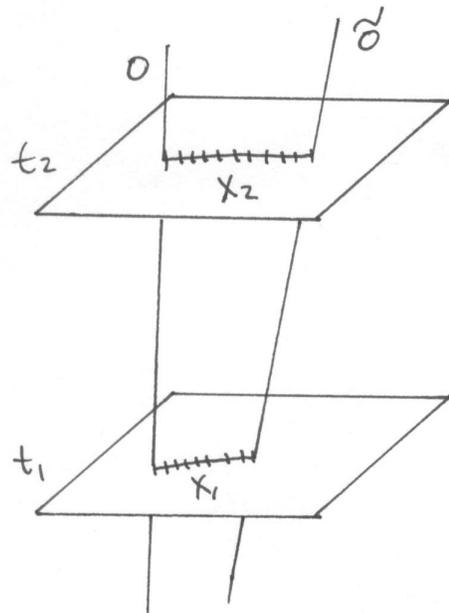


Figura 2.5: Cálculo de velocidades na Mecânica Clássica.

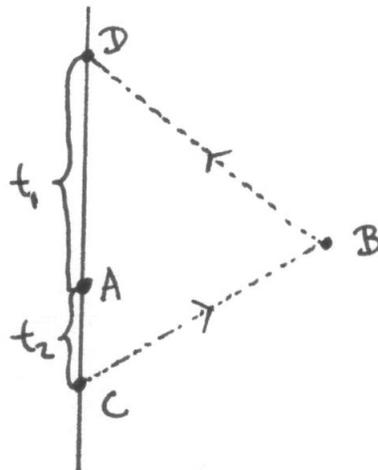


Figura 2.6: Simultaneidade na Relatividade Restrita.

2.4. PRINCÍPIOS BÁSICOS

agora a noção de simultaneidade não é mais absoluta mas ainda estamos supondo que as superfícies de simultaneidade são munidas de uma geometria Euclidiana. Se \tilde{O} é um outro observador inercial que também passa por A e sua velocidade relativa em relação a O é diferente de zero, então O julga que A e E são simultâneos, ao passo que \tilde{O} julga que A e F são simultâneos, como na Figura 2.7.

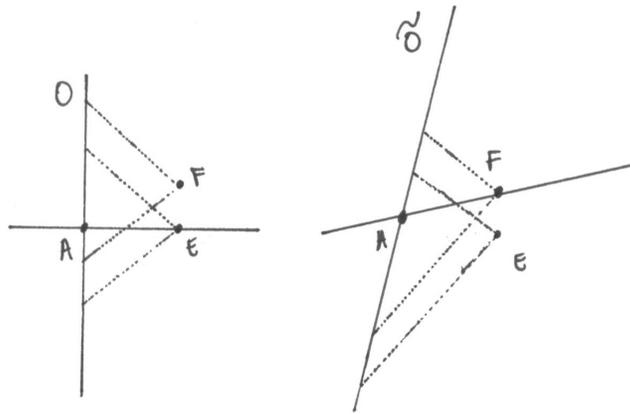


Figura 2.7: Em Relatividade Restrita, a simultaneidade não é um conceito que independe do observador.

Com esse método O consegue assinalar coordenadas a todos os eventos do espaço-tempo. Dado um evento B , existe um evento A sobre a trajetória de O simultâneo com B . Associamos a B o tempo marcado pelo relógio de O em A e a coordenada espacial dada pela direção e distância de A para B .

Se você não gostou desse método, podemos introduzir coordenadas de outra forma. Considere O um observador inercial. Seja \tilde{O} um observador inercial em repouso em relação a O . Dessa forma, \tilde{O} é também um observador inercial. Suponha que ambos carreguem relógios idênticos. Queremos sincronizá-los. Para isso, considere outro observador inercial em repouso em relação a O e \tilde{O} mas situado no ponto médio entre eles. De um evento A sobre a trajetória desse observador parte um sinal luminoso em todas as direções. Quando esse sinal alcançar as trajetórias de O e \tilde{O} dizemos que ambos relógios marcam o tempo 0. Com isso O associa a cada evento na trajetória de \tilde{O} o tempo marcado por seu relógio nesse evento e como coordenada espacial a direção e distância entre tal evento e o evento sobre a trajetória de O no instante que seu relógio marcar o mesmo tempo. Dado então um evento arbitrário no espaço-tempo, considere um observador inercial em repouso em relação a O que passe por tal evento e associe coordenadas a ele da mesma forma como fizemos acima.

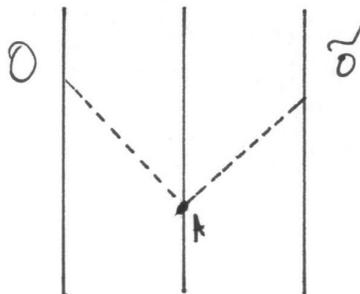


Figura 2.8: Colocando coordenadas no espaço-tempo.

Note que as coordenadas introduzidas por esses dois métodos tem algumas propriedades em comum. Primeiro, para todo observador inercial, a geometria do espaço obtido fixando-se o tempo é Euclidiana, por hipótese. Segundo, fixado um observador inercial, a distância entre dois pontos x e y é independente do tempo. Mais precisamente, se O é um observador inercial e (t_1, x) e (t_1, y) são dois pontos na superfície de simultaneidade definida por O em um instante de tempo t_1 então a distância entre (t_2, x) e (t_2, y) é igual a distância entre (t_1, x) e (t_1, y)

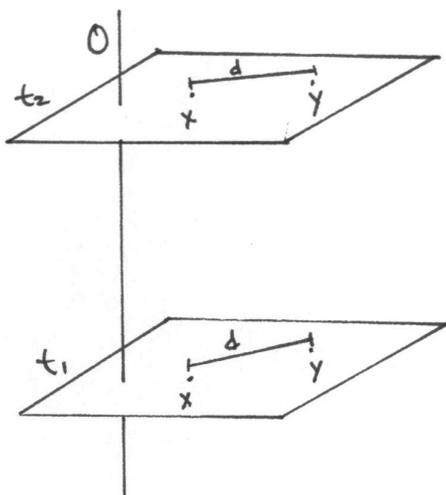


Figura 2.9: A distância entre dois pontos é independente do tempo.

Como sabemos associar coordenadas para qualquer evento do espaço-tempo pelos métodos acima, podemos, para simplificar o raciocínio, imaginar que em cada ponto do espaço existe um relógio, todos idênticos entre si e que estão sincronizados. Essa sincronização é dada pelo seguinte método: o observador inercial O , situado na origem de seu sistema de coordenadas, envia um sinal luminoso em todas as direções. Nesse instante seu relógio começa a contar o tempo a partir do zero. Dado um ponto x , tal sinal alcançará x após um tempo $\text{dist}(x, 0)/c$. Nesse instante, o relógio situado em x começa a contar o tempo a partir de tal valor. Note que como estamos pedindo que os relógios sejam idênticos, está

implícito que eles medem o tempo com a mesma taxa, ou seja, nenhum marca o tempo “mais rápido” do que outro. Note também que isso só diz respeito aos relógios relativos ao observador O . É possível que observadores distintos vejam os relógios um do outro medido o tempo com taxas distintas.

Resumindo, fixado um observador, introduzimos um sistema de coordenadas no espaço-tempo com as seguintes propriedades:

- A geometria do espaço obtido fixando-se t é Euclidiana;
- A distância entre dois pontos x e y é independente do tempo t ;
- Os relógios que estão em cada ponto marcando o tempo estão sincronizados e marcam o tempo com a mesma taxa.

Um sistema de coordenadas satisfazendo essas propriedades é dito um sistema de *coordenadas inerciais*¹². Nos convencemos por argumentos puramente físicos que existe um (e portanto, infinitos) sistemas de coordenadas inerciais, a saber, os sistemas dados por observadores inerciais usando qualquer um dos métodos acima. Para concluir, façamos um breve comentário sobre o que é um observador O fazer uma *observação*. O observar um certo evento NÃO é o evento que acontece quando ele toma consciência do acontecimento do primeiro evento. Por exemplo, O observar uma explosão não é o evento que acontece quando ele vê a explosão acontecendo. Fazer uma observação é o ato de associar coordenadas espaciais e temporais a um certo evento. Usando o exemplo acima, O observa a explosão quando ele diz que, relativo ao seu sistema de coordenadas, ela aconteceu no ponto x no tempo t .

Agora podemos compreender melhor as definições rigorosas dos objetos em questão. Em primeiro lugar, definimos o espaço-tempo como sendo uma variedade diferenciável de dimensão¹³ 4. Pela discussão acima, pedimos que o espaço-tempo admita um sistema global de coordenadas inerciais. Não tomamos o próprio \mathbb{R}^4 como espaço-tempo pois, como vimos no Capítulo 1, a presença da origem contradiz o Princípio de Relatividade de Galileu. Porém, após a introdução de coordenadas inerciais, faremos todas as contas como se estivéssemos em \mathbb{R}^4 . Não podemos também separar as palavras espaço e tempo pois não sabemos como essas entidades se relacionam. Um *evento* é um ponto do espaço-tempo. A definição de *partícula* também é importada da Mecânica Clássica e consideramos também a luz como uma partícula, chamada de *fóton*. Note que aqui os termos *partícula* e *fóton* nada têm a ver com os conceitos da Mecânica Quântica, sendo somente uma idealização de objetos com tamanho tão pequeno quanto se queira. Dependendo do contexto, podemos considerar certos objetos como uma partícula, como por exemplo, um navio atravessando o oceano Atlântico em relação ao oceano. Porém no momento que o navio chega no porto, não podemos mais considerá-lo como uma partícula em relação ao porto. Vamos supor que cada partícula carregue consigo um relógio que mede o tempo τ dos eventos em sua história. τ é dito o *tempo próprio* da partícula. Um *observador* é um sistema de coordenadas no espaço-tempo. Um observador é dito *inercial* se tal sistema de coordenadas é inercial. Apesar dessa definição, não há nada de errado em pensar que um observador é algo que se move como uma partícula livre, conceito a ser definido abaixo, que carrega um relógio consigo, onde ele marca seu tempo próprio τ . Denotamos as coordenadas fixadas por um observador inercial como (x^0, x^1, x^2, x^3) . É conveniente fazermos $x^0 = c\tau$, de modo que todas as coordenadas terão a mesma unidade de medida.

Queremos definir o que é uma *partícula livre*. Como na Mecânica Clássica, queremos que seja uma partícula fora da influência de outros corpos, de modo que se mova em linha reta e com velocidade uniforme relativamente a observadores inerciais. Dizemos então que uma partícula é *livre* se, em um sistema de coordenadas inerciais sua trajetória é dada pela equação

¹²Repare que isso não é uma definição circular. Usamos o que queríamos definir para extrair propriedades e usamos tais propriedades para definir o conceito desejado.

¹³Veja só que ironia! Se $n \neq 4$, existe apenas uma estrutura diferencial em \mathbb{R}^n , a menos de difeomorfismos. Em \mathbb{R}^4 , justamente o espaço que serve de modelo para nosso espaço-tempo, existem não-enumeráveis estruturas diferenciais não equivalentes [Art:Tau87]. Ou seja, existe uma quantidade não enumerável de variedades homeomorfas porém não difeomorfas a \mathbb{R}^4 . Será que isso é importante para a Relatividade? Há quem diga que sim. Veja [Liv:Mal07].

$$x = V\tau + T, \quad (2.3)$$

onde τ é o tempo próprio da partícula, $x = (x^0, x^1, x^2, x^3)$, e $V = (V^0, V^1, V^2, V^3)$ e $T = (T^0, T^1, T^2, T^3)$ são constantes satisfazendo

$$(V^1)^2 + (V^2)^2 + (V^3)^2 < (V^0)^2. \quad (2.4)$$

Para tornar precisa a idéia de que a velocidade da luz independe do observador, dizemos que dois eventos A e B estão na trajetória de um fóton se e somente se suas coordenadas $x = (x^0, x^1, x^2, x^3)$ e $y = (y^0, y^1, y^2, y^3)$ são relacionadas por

$$(x^0 - y^0)^2 - (x^1 - y^1)^2 - (x^2 - y^2)^2 - (x^3 - y^3)^2 = 0. \quad (2.5)$$

Reescrevendo a equação 2.5 temos que $(x^0 - y^0)^2 = (x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2$, ou seja, a distância espacial de A até B é c vezes o intervalo de tempo entre A e B . Para x fixo, 2.5 define o *cone de luz* do evento A , representando a união de todas as trajetórias de fótons que passam por A . Finalmente, a desigualdade 2.4 é uma forma precisa de dizer que não podemos nos mover mais rápido do que a luz, pois esse é justamente o critério para que tenhamos o vetor V “dentro” do cone de luz.

2.5 Transformações de Lorentz

Nosso objetivo agora é encontrar uma relação entre coordenadas inerciais de diferentes observadores. Vamos fixar algumas notações.

Notação (Coeficientes métricos¹⁴). As quantidades g_{ij} e g^{ij} são definidas como

$$\begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} g^{00} & g^{01} & g^{02} & g^{03} \\ g^{10} & g^{11} & g^{12} & g^{13} \\ g^{20} & g^{21} & g^{22} & g^{23} \\ g^{30} & g^{31} & g^{32} & g^{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Os índices i e j sempre percorrerão o conjunto $\{0, 1, 2, 3\}$.

Observação. Agora você pode se perguntar: “Porquê esse cara deu dois nomes para a mesma coisa?!” O que acontece é o seguinte: como foi observado na nota de rodapé acima, g será o tensor da métrica no espaço de Minkowski. Sabemos da Geometria Diferencial que um tensor métrico é representado, em cada ponto, por uma matriz. No caso particular da Relatividade Restrita, onde o espaço-tempo é um espaço vetorial, podemos definir a métrica a partir de uma única matriz. No nosso caso, escolhemos a matriz g para tal fim. Em Geometria, é comum a notação g^{ij} para denotar as coordenadas da inversa de g_{ij} . Nesse caso particular a matriz e sua inversa coincidem. Por isso esses símbolos foram definidos da mesma forma. Nada dessa discussão feita sobre métricas será usada nesse capítulo, só precisaremos dessas quantidades g_{ij} e g^{ij} .

¹⁴Esse nome não é por acaso. Veremos mais a frente que tal expressão será o tensor da métrica no espaço de Minkowski (Hermann Minkowski, Alekostas 22/06/1864 - Göttingen 12/01/1909.).

2.5. TRANSFORMAÇÕES DE LORENTZ

Notação (Notação de Einstein). Se um índice aparece como um índice superior e inferior, então está implícito que há uma soma nesse índice¹⁵. Ou seja,

$$a_i b^i = \sum_{i=0}^3 a_i b^i.$$

Com essa notação, a Equação 2.5 escreve-se como

$$g_{ij}(x^i - y^i)(x^j - y^j) = 0.$$

Sejam O e \tilde{O} dois observadores inerciais com coordenadas inerciais x^i e \tilde{x}^i , respectivamente. Suponha que cada observador ocupe a origem das suas coordenadas espaciais. Queremos encontrar a relação entre as coordenadas x^i e \tilde{x}^i . Para isso, escrevemos $x^i(t) = x^i(\tilde{x}^0(t), \tilde{x}^1(t), \tilde{x}^2(t), \tilde{x}^3(t))$. Seja P uma partícula livre. Pela equação 2.3,

$$\frac{d^2 x^i}{d\tau^2} = \frac{d^2 \tilde{x}^i}{d\tau^2} = 0,$$

onde τ é o tempo próprio da partícula. Por outro lado,

$$\frac{d^2 x^i}{d\tau^2} = \frac{\partial^2 x^i}{\partial \tilde{x}^k \partial \tilde{x}^j} \frac{d\tilde{x}^j}{d\tau} \frac{d\tilde{x}^k}{d\tau} + \frac{\partial x^i}{\partial \tilde{x}^j} \frac{d^2 \tilde{x}^j}{d\tau^2} = \frac{\partial^2 x^i}{\partial \tilde{x}^k \partial \tilde{x}^j} \frac{d\tilde{x}^j}{d\tau} \frac{d\tilde{x}^k}{d\tau}.$$

Como isso vale para qualquer partícula livre temos que

$$\frac{\partial^2 x^i}{\partial \tilde{x}^k \partial \tilde{x}^j} = 0.$$

Dessa forma, a dependência de x em \tilde{x} é linear e podemos escrever que

$$x^i = K_j^i \tilde{x}^j + T^i, \quad (2.6)$$

onde K_j^i e T^i são constantes.

Suponha agora que A e B são eventos que estão na trajetória de um fóton. Escrevendo suas coordenadas em relação a O e \tilde{O} , respectivamente, como x^i e y^i , pela equação 2.5,

$$g_{ij}(x^i - y^i)(x^j - y^j) = 0 = g_{ij}(\tilde{x}^i - \tilde{y}^i)(\tilde{x}^j - \tilde{y}^j). \quad (2.7)$$

Substituindo 2.6 na primeira igualdade de 2.7

$$g_{ij}(x^i - y^i)(x^j - y^j) = g_{ij}K_k^i K_l^j (\tilde{x}^k - \tilde{y}^k)(\tilde{x}^l - \tilde{y}^l) = g_{kl}(\tilde{x}^k - \tilde{y}^k)(\tilde{x}^l - \tilde{y}^l) = 0$$

Pelas duas últimas igualdades, temos que $g_{ij}K_k^i K_l^j$ é proporcional a g_{kl} . Podemos então escrever $K_j^i = \xi L_j^i$, onde a constante ξ é escolhida de modo que valha

$$g_{ij}L_k^i L_l^j = g_{kl}. \quad (2.8)$$

Note que essa equação nos diz que $L^t g L = g$, onde $g = (g_{ij})$ e $L = (L_j^i)$.

¹⁵Existe no folclore científico uma história curiosa sobre essa notação. Uma secretária de Einstein, na hora de redigir um artigo, esqueceu de colocar o símbolo de somatório e desde então isto se chama notação de Einstein. Infelizmente essa história não é verdade. Em [Art:Ein16], quando essa notação aparece pela primeira vez, Einstein avisa que a está introduzindo.

Sejam τ e $\tilde{\tau}$ os tempos próprios dos observadores O e \tilde{O} , respectivamente. Suas trajetórias se escrevem em seus referenciais como

$$\begin{aligned} (O) \quad x^0 &= c\tau, \quad x^1 = x^2 = x^3 = 0; \\ (\tilde{O}) \quad \tilde{x}^0 &= c\tilde{\tau}, \quad \tilde{x}^1 = \tilde{x}^2 = \tilde{x}^3 = 0. \end{aligned}$$

Nas coordenadas de O , usando a equação 2.6, o segundo conjunto de equações torna-se

$$x^i = K_j^i \tilde{x}^j + T^i = \xi L_j^i \tilde{x}^j + T^i = \xi L_0^i \tilde{x}^0 + T^i = \xi L_0^i c\tilde{\tau} + T^i. \quad (2.9)$$

Portanto, a velocidade de \tilde{O} em relação a O é dada por

$$\vec{v} = \left(\frac{dx^1}{d\tilde{\tau}}, \frac{dx^2}{d\tilde{\tau}}, \frac{dx^3}{d\tilde{\tau}} \right) \frac{d\tilde{\tau}}{d\tau} = (L_0^1, L_0^2, L_0^3) \frac{c}{L_0^0},$$

e portanto, $v = \sqrt{(L_0^1)^2 + (L_0^2)^2 + (L_0^3)^2} \frac{c}{L_0^0}$. Tomando $k = j = 0$ em 2.8 temos que $(L_0^1)^2 + (L_0^2)^2 + (L_0^3)^2 = (L_0^0)^2 - 1$. Dessa forma, $v = c \sqrt{1 - (L_0^0)^{-2}}$ e $L_0^0 = 1 / \sqrt{1 - v^2/c^2}$. Vamos chamar essa última constante de $\gamma(v)$.

Reciprocamente, podemos escrever as coordenadas de \tilde{O} em função de x^i como $\tilde{x}^i = \xi^{-1} M_j^i (x^j - T^j)$, onde M é a matriz inversa de L . Mais explicitamente, $M = gL^t g$ ou

$$M_j^i = g^{ik} g_{jl} L_k^l. \quad (2.10)$$

Dois identidades importantes que seguem de 2.8 são

$$g^{im} g_{kl} L_m^l L_j^k = \delta_j^i \quad (2.11)$$

e

$$g^{kl} L_k^i L_l^j = g^{ij}. \quad (2.12)$$

A primeira diz nada além de $M = L^{-1}$. A segunda diz que $LgL^t = g$. Usando a fórmula para M temos que $M_0^0 = L_0^0$, o que mostra que a velocidade de O em relação a \tilde{O} , em módulo, também é dada por v . Para determinarmos ξ , note que a relação entre x^0 e $\tilde{\tau}$ ao longo da trajetória de \tilde{O} é dada por

$$x^0 = \frac{c\xi\tilde{\tau}}{\sqrt{1 - v^2/c^2}} + \text{constante}$$

e, reciprocamente, a relação entre \tilde{x}^0 e τ ao longo da trajetória de O é

$$\tilde{x}^0 = \frac{1}{\xi} \frac{c\tau}{\sqrt{1 - v^2/c^2}} + \text{constante}.$$

Mas o princípio da relatividade diz que somente movimento relativo pode ser detectado, ou seja, nada deve se alterar se trocarmos O e \tilde{O} . Devemos ter então que $\xi = \pm 1$. O sinal de ξ é fixado pedindo-se que todos os observadores inerciais concordem na direção do tempo. Para isso, tomamos também raízes positivas nas duas equações acima.

Resumindo, dois sistemas inerciais de coordenadas estão relacionados por uma transformação da forma

$$x^i = L_j^i \tilde{x}^j + T^i,$$

onde $g_{ij}L_k^iL_l^j = g_{kl}$ e $L_0^0 > 0$. Consideraremos somente o caso $T^i = 0$, que é quando não há um deslocamento da origem do espaço-tempo. O conjunto de todas as matrizes L_j^i satisfazendo 2.8 forma um grupo, chamado de *grupo de Lorentz*, e seus elementos são chamados naturalmente de *transformações de Lorentz*. Note que não necessariamente as transformações desse grupo devem satisfazer $L_0^0 > 0$. Vamos pedir isso mais adiante para que observadores distintos concordem na orientação do tempo. Note também que $\det L = \pm 1$. As transformações de Lorentz satisfazendo $\det L = 1$ e $L_0^0 > 0$ terão uma importância fundamental no próximo capítulo.

2.6 A dilatação temporal - II

Uma consequência direta das transformações de Lorentz é podermos deduzir rigorosamente o efeito da dilatação temporal. Isso segue diretamente da equação

$$x^0 = \frac{c\tilde{\tau}}{\sqrt{1 - v^2/c^2}}.$$

Como $x^0 = c\tau$, onde τ é o tempo próprio do observador O , temos que

$$\tau = \frac{\tilde{\tau}}{\sqrt{1 - v^2/c^2}}.$$

Essa equação diz que, para o observador O , o relógio de \tilde{O} marca o tempo mais devagar por um fator de $\sqrt{1 - v^2/c^2}$.

2.7 A transformação de Lorentz padrão

Para entender melhor o significado das Transformações de Lorentz, vamos considerar um caso particular. Suponha que O e \tilde{O} são dois observadores inerciais cujas origens coincidem em um certo evento A . Suponha que a velocidade relativa de \tilde{O} em relação a O seja dada por $(v, 0, 0)$ e, portanto, a velocidade relativa de O em relação a \tilde{O} é dada por $(-v, 0, 0)$. Isso fixa a direção do eixo x^1 , porém deixa livre os eixos x^2 e x^3 . Vimos acima que os dois sistemas de coordenadas estão relacionados por

$$x^i = L_k^i \tilde{x}^k \quad \text{e} \quad \tilde{x}^i = M_k^i x^k.$$

Vamos descobrir qual é a cara da matriz L . Para isso, note que

- $\vec{v} = (v, 0, 0) \Rightarrow (L_0^1, L_0^2, L_0^3) = \gamma(v)(v/c, 0, 0)$
- Equação 2.10 $\Rightarrow (M_0^1, M_0^2, M_0^3) = -(L_1^0, L_2^0, L_3^0)$
- $-\vec{v} = (-v, 0, 0) \Rightarrow (M_0^1, M_0^2, M_0^3) = \gamma(v)(-v/c, 0, 0) \Rightarrow -(L_1^0, L_2^0, L_3^0) = \gamma(v)(-v/c, 0, 0)$
- $k = 0$ e $l = 1$ em 2.8 $\Rightarrow L_1^1 = \gamma(v)$
- $k = l = 1$ em 2.8 $\Rightarrow L_1^2 = L_1^3 = 0$
- $k = l = 1$ em 2.12 $\Rightarrow L_2^1 = L_3^1 = 0$

- $k = l = 2$ em 2.8 $\Rightarrow (L_2^2)^2 + (L_3^2)^2 = 1$
- $k = l = 3$ em 2.8 $\Rightarrow (L_2^3)^2 + (L_3^3)^2 = 1$
- $i = 2$ e $j = 3$ em 2.11 $\Rightarrow L_2^2 L_3^2 + L_2^3 L_3^3 = 0$

Isso nos permite escrever L como

$$L = \begin{pmatrix} L_0^0 & L_1^0 & L_2^0 & L_3^0 \\ L_0^1 & L_1^1 & L_2^1 & L_3^1 \\ L_0^2 & L_1^2 & L_2^2 & L_3^2 \\ L_0^3 & L_1^3 & L_2^3 & L_3^3 \end{pmatrix} = \begin{pmatrix} \gamma(v) & \gamma(v)v/c & 0 & 0 \\ \gamma(v)v/c & \gamma(v) & 0 & 0 \\ 0 & 0 & \cos(\theta) & \sin(\theta) \\ 0 & 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix},$$

para algum ângulo θ . Fazendo uma rotação adequada nos eixos x^2 e x^3 , podemos tomar $\theta = 0$. Dessa forma, com $(x^0, x^1, x^2, x^3) = (ct, x, y, z)$ e $(\tilde{x}^0, \tilde{x}^1, \tilde{x}^2, \tilde{x}^3) = (c\tilde{t}, \tilde{x}, \tilde{y}, \tilde{z})$ temos que

$$\begin{aligned} x &= \gamma(v)(\tilde{x} + v\tilde{t}) = \frac{\tilde{x} + v\tilde{t}}{\sqrt{1 - v^2/c^2}} \\ y &= \tilde{y} \\ z &= \tilde{z} \\ t &= \gamma(v)(\tilde{t} + v\tilde{x}/c^2) = \frac{\tilde{t} + v\tilde{x}/c^2}{\sqrt{1 - v^2/c^2}}. \end{aligned}$$

Essa é dita a *transformação de Lorentz padrão*. A princípio poderíamos pensar que essa classe de transformações é menos geral que as transformações de Lorentz arbitrárias, de modo que existam fenômenos que só podem ser deduzidos usando a forma geral da transformação. Felizmente temos o seguinte resultado:

Teorema 2.1. *Se L é uma transformação de Lorentz satisfazendo $\det L = 1$ e $L_0^0 > 0$, então existem matrizes ortogonais 3×3 H e \tilde{H} de modo que*

$$\begin{pmatrix} 1 & 0 \\ 0 & H \end{pmatrix} L \begin{pmatrix} 1 & 0 \\ 0 & \tilde{H} \end{pmatrix}$$

é uma transformação de Lorentz padrão.

A prova pode ser encontrada em [Liv:Nai64]. Não a reproduziremos aqui pois isso foge do escopo desse texto.

Note que, pelo Princípio da Relatividade, não há nenhum referencial privilegiado. Dessa forma, a inversa de L , que chamamos de M e que escreve as coordenadas de \tilde{O} em função das coordenadas de O , é obtida tomando-se o negativo de v :

$$L = \begin{pmatrix} \gamma(-v) & -\gamma(-v)v/c & 0 & 0 \\ -\gamma(-v)v/c & \gamma(-v) & 0 & 0 \\ 0 & 0 & \cos(\theta) & \sin(\theta) \\ 0 & 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Fazendo $\theta = 0$, temos

$$\begin{aligned} \tilde{x} &= \gamma(v)(x - vt) = \frac{x - vt}{\sqrt{1 - v^2/c^2}} \\ \tilde{y} &= y \\ \tilde{z} &= z \\ \tilde{t} &= \gamma(v)(t - vx/c^2) = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}}. \end{aligned}$$

Tomando os limites $c \rightarrow \infty$ na transformação de Lorentz padrão acima temos que

$$\begin{aligned}\tilde{x} &= x - vt \\ \tilde{y} &= y \\ \tilde{z} &= z \\ \tilde{t} &= t,\end{aligned}$$

ou seja, no caso em que não há limite máximo de velocidade para os objetos (em particular, na Mecânica Clássica), a transformação de Lorentz se reduz a uma transformação de Galileu. Esse também é o caso se $v \ll c$, de modo que $v^2/c^2 \ll 1$, ou seja, em baixas velocidades as transformações de Lorentz são aproximadas por transformações de Galileu.

2.8 A contração de FitzGerald-Lorentz

Agora podemos tornar rigorosa a dedução do efeito da contração espacial. Para isso considere uma barra rígida sobre o eixo x^1 em repouso em relação ao observador \tilde{O} . Suponha, por simplicidade, que um dos extremos da barra está na origem das coordenadas espaciais de \tilde{O} . Então as trajetórias W_1 e W_2 das duas pontas da barra são dadas nas coordenadas de \tilde{O} como

$$\begin{aligned}(W_1) \quad \tilde{x}^0 &= c\tilde{\tau}, \quad \tilde{x}^1 = \tilde{x}^2 = \tilde{x}^3 = 0; \\ (W_2) \quad \tilde{x}^0 &= c\tilde{\tau}, \quad \tilde{x}^1 = D, \quad \tilde{x}^2 = \tilde{x}^3 = 0,\end{aligned}$$

onde $\tilde{\tau}$ é o tempo próprio ao longo de W_1 e W_2 e D é o comprimento da barra, como medido por \tilde{O} . Se supusermos que as coordenadas de um outro observador O se relacione com as coordenadas de \tilde{O} por meio de uma transformação de Lorentz padrão, então nas coordenadas de O temos que

$$\begin{aligned}(W_1) \quad x^0 &= c\gamma(v)\tilde{\tau}, \quad x^1 = \gamma(v)v\tilde{\tau}, \quad x^2 = x^3 = 0; \\ (W_2) \quad x^0 &= c\gamma(v)(\tilde{\tau} + vD/c^2), \quad x^1 = \gamma(v)(D + v\tilde{\tau}), \quad x^2 = x^3 = 0.\end{aligned}$$

Queremos saber qual é o comprimento da barra para O , ou seja, queremos calcular a distância espacial em relação a O entre dois eventos simultâneos ocorrendo em W_1 e W_2 . Para isso, devemos eliminar a variável $\tilde{\tau}$, escrevendo-a em função do tempo próprio de O , denotado por t . Como $ct = x^0 = c\gamma(v)\tilde{\tau}$ no ponto W_1 , temos que $\tilde{\tau} = t/\gamma(v)$. Dessa forma, temos que

$$(W_1) \quad x^0 = ct, \quad x^1 = vt, \quad x^2 = x^3 = 0.$$

Analogamente, em W_2 temos que $ct = x^0 = c\gamma(v)(\tilde{\tau} + vD/c^2) \Rightarrow \tilde{\tau} = t/\gamma(v) - vD/c^2$. Substituindo na equação para W_2 , temos que

$$(W_2) \quad x^0 = ct, \quad x^1 = vt + D\sqrt{1 - v^2/c^2}, \quad x^2 = x^3 = 0.$$

Com isso, o comprimento da barra para O é dado por $D\sqrt{1 - v^2/c^2}$, que é menor que D .

Podemos mostrar também que a contração acontece somente na direção do movimento, fazendo o mesmo raciocínio acima com a barra posicionada em algum eixo ortogonal ao eixo x^1 . Com isso, o resultado negativo no experimento de Michelson-Morley está explicado, sem precisar recorrer a métodos heurísticos para introduzir a contração.

2.9 Adição de velocidades

Vimos que adição de velocidades para observadores distintos não pode mais ser dada por $v'' = v + v'$, pois essa fórmula não respeita o limite de c para a velocidade dos objetos. Para deduzirmos uma nova fórmula, sejam O , \tilde{O} e \bar{O} observadores inerciais com coordenadas x^i , \tilde{x}^i e \bar{x}^i , respectivamente. Suas coordenadas são relacionadas por

$$x^i = L_j^i \tilde{x}^j \quad \text{e} \quad \tilde{x}^i = L_j^i \bar{x}^j,$$

onde L e L' são transformações de Lorentz padrão dadas por

$$L = \begin{pmatrix} \gamma(v) & \gamma(v)v/c & 0 & 0 \\ \gamma(v)v/c & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad L' = \begin{pmatrix} \gamma(v') & \gamma(v')v'/c & 0 & 0 \\ \gamma(v')v'/c & \gamma(v') & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

onde v é a velocidade de \tilde{O} em relação a O e v' é a velocidade de \bar{O} em relação a \tilde{O} . Para saber a velocidade de \bar{O} em relação a O , combinamos as transformações de coordenadas acima para obtermos que $x^i = L_j^i L_k^j \bar{x}^k$. Mas a matriz LL' é dada por

$$LL' = \begin{pmatrix} \gamma(v)\gamma(v')(1 + vv'/c^2) & \gamma(v)\gamma(v')(v + v')/c & 0 & 0 \\ \gamma(v)\gamma(v')(v + v')/c & \gamma(v)\gamma(v')(1 + vv'/c^2) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Se chamarmos de v'' a velocidade de \bar{O} em relação a O e igualarmos o termo $(LL')_0^0$ a $\gamma(v'')$, concluímos que

$$v'' = \frac{c^2(v + v')}{c^2 + vv'}, \quad (2.13)$$

de modo que a matriz LL' torna-se

$$LL' = \begin{pmatrix} \gamma(v'') & \gamma(v'')v''/c & 0 & 0 \\ \gamma(v'')v''/c & \gamma(v'') & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

É conveniente buscarmos por um parâmetro θ que esteja relacionado com a velocidade de modo que se θ é a velocidade de \tilde{O} em relação a O e θ' representa a velocidade de \bar{O} em relação a \tilde{O} , então $\theta'' := \theta + \theta'$ representa a velocidade de \bar{O} em relação a O . Já que θ está relacionado com velocidades relativas, é razoável pedirmos que v seja uma função bijetiva de θ . Como vimos acima, temos que $v'' = c^2(v + v')/(c^2 + vv')$. Se escrevermos $v = f(\theta)$, $v' = f(\theta')$ e $v'' = f(\theta + \theta')$, temos que f deve satisfazer a seguinte equação funcional

$$f(\theta + \theta') = \frac{c^2(f(\theta) + f(\theta'))}{c^2 + f(\theta)f(\theta')}. \quad (2.14)$$

Isso nos sugere que devemos ter $v = f(\theta) = c \tanh(\theta)$, ou $\theta = \tanh^{-1}(v/c)$. Note que $\tanh^{-1}(\cdot/c) : (-c, c) \rightarrow \mathbb{R}$ é um difeomorfismo suave com a propriedade que $v \rightarrow \pm c$ implica $\tanh^{-1}(v/c) \rightarrow \pm\infty$, de modo que o θ correspondente a velocidade da luz é $+\infty$, o que é razoável com a adição de velocidades que estamos procurando.

Para determinar unicamente f , note que se v e v' são suficientemente pequenos a Fórmula 2.13 pode ser aproximada por $v'' \approx v + v'$. Portanto, para baixas velocidades v e θ devem ser aproximadamente iguais, ou seja,

$$\lim_{\theta \rightarrow 0} \frac{f(\theta)/c}{\theta} = 1 \quad (2.15)$$

Temos então o seguinte resultado:

Proposição 2.2. *Existe uma única função diferenciável $f(\theta)$ em \mathbb{R} satisfazendo 2.14 e 2.15.*

Demonstração. A existência é imediata, já que vimos acima que $f(\theta) = \tanh(\theta)$ satisfaz 2.14. Verifica-se facilmente também que tal f satisfaz 2.15. Para provar a unicidade, usamos as duas fórmulas para calcular $f'(\theta) = \lim_{\theta' \rightarrow 0} (f(\theta + \theta') - f(\theta))/(\theta')$, mostrando que f satisfaz a EDO $f' = 1 - f^2$ com a condição inicial $f(0) = 0$. Pelo Teorema de Existência e Unicidade (veja [Liv:Cod84]), temos o resultado desejado. \square

Dessa forma, podemos escrever uma transformação de Lorentz padrão da seguinte forma

$$L = \begin{pmatrix} \cosh(\theta) & \sinh(\theta) & 0 & 0 \\ \sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

o que nos permite interpretá-la como uma “rotação” na coordenada temporal e na primeira coordenada espacial.

2.10 Simultaneidade

Lembremos de como a coordenada temporal é transformada segundo uma transformação de Lorentz padrão:

$$t = \frac{\tilde{t} + v\tilde{x}/c^2}{\sqrt{1 - v^2/c^2}}.$$

Olhando cuidadosamente para essa equação, podemos concluir que eventos que ocorrem em dois lugares diferentes ao mesmo tempo como vistos por um observador não ocorrem ao mesmo tempo em relação a outro observador. De fato, se um evento ocorre no ponto $(\tilde{x}_1, 0, 0)$ no tempo \tilde{t}_0 e outro no ponto $(\tilde{x}_2, 0, 0)$ no mesmo tempo \tilde{t}_0 em relação a \tilde{O} , seus tempos correspondentes em relação a O são

$$t_1 = \frac{\tilde{t}_0 + v\tilde{x}_1/c^2}{\sqrt{1 - v^2/c^2}} \quad \text{e} \quad t_2 = \frac{\tilde{t}_0 + v\tilde{x}_2/c^2}{\sqrt{1 - v^2/c^2}},$$

de modo que a diferença entre esses tempos é dada por

$$t_2 - t_1 = \frac{v(\tilde{x}_2 - \tilde{x}_1)/c^2}{\sqrt{1 - v^2/c^2}}.$$

Para tornar a idéia um pouco mais clara, considere o seguinte experimento.

Suponha que um homem que está se movendo em um foguete colocou um relógio em cada uma das duas extremidades do foguete e quer ter certeza se os relógios estão sincronizados. Uma maneira fácil de se fazer isso é localizar o ponto médio entre os dois relógios e então enviar um sinal luminoso nas duas direções. Esses sinais terão a mesma velocidade e chegarão nos dois relógios ao mesmo tempo, segundo esse observador. Usando esse procedimento, ele terá certeza que os relógios estão sincronizados. Caso

contrário, ele saberia que está viajando, contradizendo o Princípio de Relatividade de Galileu. Considere agora um observador parado na Terra. Será que para ele os dois relógios estão sincronizados? Para ele, uma vez que o foguete está indo para a frente, o relógio na extremidade dianteira está se afastando do sinal luminoso, de modo que a luz tem que percorrer mais para alcançá-lo. Analogamente, o relógio na extremidade traseira está se aproximando, fazendo com que a luz o alcance primeiro. Portanto, para o observador no foguete a chegada dos sinais luminosos em ambos os relógios são eventos simultâneos e os relógios estão sincronizados, enquanto que para o observador na Terra não.

2.11 O paradoxo dos gêmeos

Na Teoria da Relatividade Restrita temos vários fenômenos inesperados e estranhos, como a dilatação temporal e a contração espacial. Nossa primeira reação é pensar que nossa teoria está errada e procurar por uma correção. Porém, a Relatividade Restrita é uma das teorias físicas mais testadas experimentalmente e suas previsões são confirmadas diariamente com altíssimo grau de precisão. O fato de algumas de suas conclusões serem estranhas não quer dizer que a teoria está errada. O problema é que esses fenômenos são perceptíveis somente em velocidades próximas à velocidade da luz, e não estamos acostumados a presenciar essas velocidades no nosso dia-a-dia. Justamente pelo fato dessas conclusões serem contrárias a nossas experiências é muito comum interpretá-las erroneamente, levando a certos paradoxos. Note que a Teoria não é inconsistente nem errada do ponto de vista físico (pelo menos por enquanto!). O que chamamos aqui de paradoxos são somente interpretações erradas de alguns conceitos básicos, e eles podem ser resolvidos olhando mais atentamente para o que está acontecendo.

Certamente o paradoxo dos gêmeos é o mais famoso da Relatividade Restrita. Sua resolução é bem fácil, mas é um bom começo para vermos como que as coisas podem dar errado quando as interpretamos sem tomar os devidos cuidados.

Paradoxo 2.3 (dos gêmeos). *Pedro tem um irmão gêmeo, o Bob. Apesar deles serem gêmeos idênticos, Bob é muito mais aventureiro que Pedro e um dia resolve ir viajar. Ele convidou Pedro, porém ele preferiu ficar em casa. Bob pegou então seu foguete e partiu da Terra em linha reta com velocidade constante, bem próxima à da luz, durante um certo tempo. Quando ele cansou de viajar, deu meia volta e voltou para a Terra na mesma velocidade. No seu referencial, Pedro ficou parado enquanto Bob viajou. Então pelo fenômeno da dilatação temporal Bob deve estar mais novo quando eles se reencontram. Porém no referencial de Bob, quem viajou foi Pedro! Isso quer dizer que no reencontro quem estará mais novo é o Pedro.*

O problema aqui é a escolha dos referenciais. O referencial de Pedro é inercial, ao passo que o de Bob não é, pois ele não se move com velocidade constante em relação a um referencial inercial. No momento que Bob resolveu retornar a Terra, ele teve que desacelerar, mudar de direção e acelerar novamente para retomar à velocidade de cruzeiro. O fenômeno da dilatação temporal só vale quando consideramos corpos em movimento com velocidade constante em relação a referenciais inerciais. Ou seja, a conclusão que Bob estará mais novo quando eles se reencontrarem é verdadeira, porém a outra não é. Existe uma variação desse paradoxo no caso em que uma das dimensões do espaço é compactificada, possibilitando que Bob chegue de volta a Terra ainda sendo um referencial inercial [Art:Fal07]. Concluímos então que a resolução desse paradoxo é somente uma interpretação correta de quando vale o fenômeno da dilatação temporal, notando que a simetria aparente entre os dois observadores não é de fato uma simetria boa para aplicar os conceitos relativísticos.

Capítulo 3

A Geometria do Espaço de Minkowski

3.1 Motivação

Vamos retomar a discussão feita no início da Seção 2.4. Lembremos que na Mecânica Clássica temos uma noção de simultaneidade universal. Se quisermos estudar a geometria do espaço-tempo, devemos buscar por propriedades que sejam independentes do observador. Quais são essas propriedades no caso clássico? Da discussão feita na Seção 2.4, segue que o intervalo de tempo entre dois eventos e a distância espacial entre eventos simultâneos independem do observador. Além disso, as trajetórias dos observadores inerciais nos dá uma noção natural de “linha reta” no espaço-tempo, conectando quaisquer dois eventos não-simultâneos. Quais são as quantidades intrínsecas, independentes do observador, que temos em mãos na Relatividade Restrita? Vimos que o intervalo de tempo entre dois eventos depende fortemente do observador, e não podemos sequer pensar em falar que a distância entre eventos simultâneos é absoluta, pois nem a noção de simultaneidade é absoluta. Porém, não estamos perdidos. Lembremos que dados dois eventos A e B no espaço-tempo e um observador inercial O que passa por A , tomamos eventos C e D sobre a trajetória de O de modo que um sinal luminoso enviado de B alcança D e um sinal luminoso enviado de C alcança A (veja a Figura 2.6). Definimos t_1 como o tempo medido por O em D subtraído do tempo medido por O em A e t_2 como o tempo medido por O em A subtraído do tempo medido por O em C . Definimos também $\Delta t := (t_1 - t_2)/2$ como o *intervalo de tempo* entre A e B e $\Delta x := c(x_1 + x_2)/2$ como a *distância espacial* entre A e B . Mostra-se que a quantidade $-t_1 t_2$ não depende do observador. Para uma demonstração muito bonita desse fato usando argumentos físicos, veja o Apêndice de [Liv:Wal92]¹. Mais adiante, tornaremos essa idéia rigorosa e provaremos essa invariância matematicamente. Lógico que esse sinal negativo não tem nada de especial e a quantidade $t_1 t_2$ também é conservada. Veremos em breve porque esse sinal será conveniente. Para deixar mais claro o significado de tal quantidade, notemos que

$$-t_1 t_2 = \left(\frac{t_1 - t_2}{2}\right)^2 - \left(\frac{t_1 + t_2}{2}\right)^2 = (\Delta t)^2 - \left(\frac{\Delta x}{c}\right)^2. \quad (3.1)$$

Vamos chamar essa quantidade de *intervalo no espaço-tempo* ou simplesmente *intervalo* entre dois eventos. Vejamos algumas informações que esse número pode nos dar. Lembre que observamos que t_1 ou t_2 , mas não ambos, pode ser negativo e que prometi dar um significado a esses sinais.

¹Apesar de demonstrações usando argumentos físicos não serem rigorosas do ponto de vista matemático, sou obrigado a confessar que sua beleza e elegância é sem igual. Existe, inclusive, um livro [Liv:Mar09] que usa argumentos físicos para provar vários teoremas, entre eles o teorema de Gauß-Bonnet!

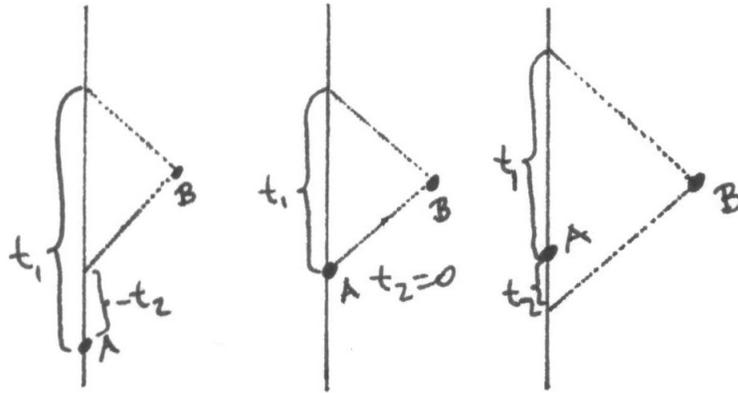


Figura 3.1: Intervalo no espaço-tempo.

Se $-t_1 t_2$ é negativo, então ambos t_1 e t_2 são positivos. Isso quer dizer não é possível que um observador esteja presente em A e B , mas existe um observador que vê esses eventos como simultâneos. Dizemos nesse caso que A e B são *especialmente relacionados*. Para um observador que veja A e B como simultâneos, o intervalo entre eles é dado por $-(\Delta x/c)^2$. Se $-t_1 t_2$ é positivo, então ou t_1 ou t_2 , mas não ambos, é negativo. Nesse caso é possível que um observador esteja presente em ambos eventos e dizemos que eles são *temporalmente relacionados*. Para tal observador, o intervalo entre os eventos é dado por $(\Delta t)^2$. Finalmente, se $-t_1 t_2$ é zero, então ou t_1 ou t_2 é zero. Nesse caso, é possível enviar um sinal luminoso de A para B ou de B para A , dependendo de qual termo é zero. É claro que isso não está rigoroso, mas deve servir como motivação para as definições principais desse capítulo.

A razão para colocarmos aquele sinal negativo na definição do intervalo é puramente estética. Como vimos acima, se é possível chegar de um evento A em um evento B , então o intervalo entre eles é positivo. Caso isso não seja possível, o intervalo é negativo. Além disso, enunciaremos vários teoremas da forma “se $g(x, y) > 0$ então alguma quantidade é positiva”. Com outras convenções, esses teoremas podem tomar a forma “se $g(x, y) > 0$ então alguma quantidade é negativa”. Os teoremas enunciados dessa forma ficam difíceis de lembrar posteriormente. Isso é uma mera convenção e os resultados independem dessa escolha. Cada livro tem uma convenção própria.

Sendo bem entendidas essas idéias físicas, podemos partir para as definições rigorosas. Tudo que afirmamos acima será então provado como propriedades geométricas do espaço-tempo.

3.2 A Métrica

A discussão acima nos motiva então a definir o *espaço de Minkowski* como sendo um espaço vetorial de dimensão 4 (três para o espaço e uma para o tempo) munido de uma certa estrutura geométrica. Para isso, precisaremos de alguns resultados básicos de Álgebra Linear.

Definição 3.1. Seja V um espaço vetorial de dimensão finita. Um *produto interno* em V é uma aplicação $g : V \times V \rightarrow \mathbb{R}$ bilinear (ou seja, linear em cada variável), simétrica ($g(x, y) = g(y, x)$, para todos $x, y \in V$) e não-degenerada (se $g(x, y) = 0$ para todo $y \in V$ então $x = 0$).

Teorema 3.2. Seja $g : V \times V \rightarrow \mathbb{R}$ um produto interno sobre um espaço vetorial de dimensão finita. Então existe uma base $\{e_1, \dots, e_n\}$, onde n é a dimensão de V , tal que se $x = x^i e_i$ e $y = y^j e_j$, então

$$g(x, y) = x^1 y^1 + \dots + x^{n-r} y^{n-r} - x^{n-r+1} y^{n-r+1} - \dots - x^n y^n.$$

A princípio o número r pode depender da base $\{e_1, \dots, e_n\}$, mas prova-se que para qualquer base que tenha a propriedade de g se escrever como acima, o número r é sempre o mesmo. Damos a ele então o nome de *índice* do produto interno g .

Dessa forma, definimos o espaço de Minkowski, denotado por \mathcal{M} , como um espaço vetorial real de dimensão 4 munido de um produto interno g de índice 3, ou seja, existe uma base $\{e_0, e_1, e_2, e_3\}$ de \mathcal{M} com a propriedade que se $x = x^i e_i$ e $y = y^i e_i$ então

$$g(x, y) = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3.$$

Chamaremos g da *métrica* em \mathcal{M} , apesar de tal espaço não ser um espaço métrico, pois g não é positiva definida (ou seja, $g(x, x) > 0$, se $x \neq 0$). Note que na notação introduzida no capítulo anterior, podemos escrever $g(x, y) = g_{ij} x^i y^j$.

3.3 Bases ortonormais e o grupo de isometrias de \mathcal{M}

A definição de g prioriza uma classe de bases. Note que uma base dessa classe satisfaz $g(e_i, e_j) = g_{ij}$. Dizemos que qualquer base para \mathcal{M} satisfazendo essa propriedade é dita uma *base ortonormal*. Existe uma relação estreita entre observadores inerciais e bases ortonormais de \mathcal{M} . A saber, suponha que um observador inercial introduz coordenadas x^i usando qualquer um dos métodos descritos na seção 2.4. Definindo os vetores $e_0 = (1, 0, 0, 0)$, $e_1 = (0, 1, 0, 0)$, $e_2 = (0, 0, 1, 0)$, $e_3 = (0, 0, 0, 1)$ escritos nas coordenadas x^i , $\{e_i\}$ é uma base ortonormal para \mathcal{M} . Reciprocamente, dada uma base ortonormal $\{e_i\}$ para \mathcal{M} considere um observador inercial que no tempo τ , segundo seu relógio, está situado no ponto $(\tau, 0, 0, 0)$. Então com essas coordenadas ele é um observador inercial. Portanto, se você não gostou da discussão física desenvolvida no capítulo anterior, você pode ignorá-la completamente e considerar somente bases ortonormais em \mathcal{M} .

A princípio, teríamos problemas com unidades, pois g mistura quantidades medidas em unidades de espaço e tempo. Porém, sanamos esse problema no capítulo anterior, quando definimos a coordenada x^0 como sendo $c\tau$, onde τ é o tempo marcado pelo relógio de um observador que introduz essas coordenadas (tempo próprio). Dessa forma, todas as quantidades envolvidas na definição de g têm a mesma unidade, a saber, unidade de espaço. Parece uma idéia estranha medir espaço e tempo com as mesmas unidades, mas esse é um artifício ao qual recorreremos diariamente. Sempre falamos “eu moro perto do trabalho, só 10 minutos caminhando” ou “é muito rápido chegar lá, são só 100 metros de distância”. Na Relatividade Restrita é semelhante. No nosso caso, estamos medindo o tempo em metros. 1 metro de tempo quer dizer o tempo que a luz leva para percorrer 1 metro. Poderíamos ter escolhido medir tempo e espaço em anos. 1 ano de espaço seria o espaço que a luz percorre em 1 ano, ou seja, 1 ano-luz, unidade de medida muito usada para representar distâncias astronômicas.

Sejam agora $\{e_i\}$ e $\{\tilde{e}_i\}$ duas bases ortonormais de \mathcal{M} . Sabemos que existe uma única transformação linear $L : \mathcal{M} \rightarrow \mathcal{M}$ satisfazendo $L(e_i) = \tilde{e}_i$. Queremos estudar qual é a transformação L . Dizemos que uma transformação L *preserva produto interno* ou é *ortogonal* se $g(Lx, Ly) = g(x, y)$, para todos $x, y \in \mathcal{M}$. Temos o seguinte resultado:

Lema 3.3. *Seja $L : \mathcal{M} \rightarrow \mathcal{M}$ uma transformação linear. São equivalentes:*

1. L é ortogonal;
2. $g(Lx, Lx) = g(x, x)$, para todo $x \in \mathcal{M}$;

3. L leva bases ortonormais em bases ortonormais.

Demonstração. $\underline{3} \Rightarrow \underline{1}$: Sejam $\{e_i\}$ e $\{Le_i\}$ bases ortonormais de \mathcal{M} . Escrevendo $x = x^i e_i$ e $y = y^j e_j$ temos que

$$g(x, y) = x^i y^j g(e_i, e_j) = g_{ij} x^i y^j.$$

Por outro lado, temos que

$$g(Lx, Ly) = g(x^i Le_i, y^j Le_j) = x^i y^j g(Le_i, Le_j) = x^i y^j g_{ij} = g(x, y).$$

$\underline{2} \Rightarrow \underline{1}$: Dados $x, y \in \mathcal{M}$, olhemos para a quantidade $g(L(x+y), L(x+y)) - g(L(x-y), L(x-y))$ de duas formas distintas. Por um lado, temos que

$$g(L(x+y), L(x+y)) - g(L(x-y), L(x-y)) = 4g(Lx, Ly).$$

Por outro lado, usando a hipótese, temos que

$$g(L(x+y), L(x+y)) - g(L(x-y), L(x-y)) = g(x+y, x+y) - g(x-y, x-y) = 4g(x, y).$$

Portanto, $g(Lx, Ly) = g(x, y)$.

$\underline{1} \Rightarrow \underline{3}$: Seja $\{e_i\}$ uma base ortonormal de \mathcal{M} . Queremos provar que $\{Le_i\}$ também é uma base ortonormal de \mathcal{M} . Note que a hipótese implica que

$$g(Le_i, Le_j) = g(e_i, e_j) = g_{ij}.$$

Dessa forma, $\{Le_i\}$ é uma base ortonormal de \mathcal{M} . □

Consideremos, então, uma transformação ortogonal L e $\{e_i\}$ uma base para \mathcal{M} . Pelo Lema acima, temos que $\{Le_i = \tilde{e}_i\}$ também é uma base ortonormal de \mathcal{M} . Podemos escrever então cada e_i como combinação linear dos \tilde{e}_j :

$$e_i = L_i^0 \tilde{e}_0 + L_i^1 \tilde{e}_1 + L_i^2 \tilde{e}_2 + L_i^3 \tilde{e}_3 = L_i^k \tilde{e}_k.$$

Usando que $\{e_i\}$ é base ortonormal de \mathcal{M} , temos que

$$g_{ij} = g(e_i, e_j) = g(L_i^k \tilde{e}_k, L_j^l \tilde{e}_l) = L_i^k L_j^l g(\tilde{e}_k, \tilde{e}_l) = L_i^k L_j^l g_{kl}.$$

Definimos a matriz associada a transformação L como

$$L = \begin{pmatrix} L_0^0 & L_1^0 & L_2^0 & L_3^0 \\ L_0^1 & L_1^1 & L_2^1 & L_3^1 \\ L_0^2 & L_1^2 & L_2^2 & L_3^2 \\ L_0^3 & L_1^3 & L_2^3 & L_3^3 \end{pmatrix}.$$

Dessa forma, as entradas da matriz L satisfazem a mesma condição 2.8 satisfeita pela mudança de coordenadas entre observadores inerciais. Isso é condizente (e esperado) da nossa associação entre observadores inerciais e bases ortonormais. Note que as colunas da matriz L são as coordenadas dos vetores $\{e_i\}$ na nova base $\{\tilde{e}_i\}$, ou seja, essa é a matriz de mudança de base de $\{e_i\}$ para $\{\tilde{e}_i\}$. Portanto, se $x = x^i e_i = \tilde{x}^i \tilde{e}_i$, então suas coordenadas nas duas bases se relacionam da seguinte forma:

$$\begin{pmatrix} \tilde{x}^0 \\ \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix} = \begin{pmatrix} L_0^0 & L_1^0 & L_2^0 & L_3^0 \\ L_0^1 & L_1^1 & L_2^1 & L_3^1 \\ L_0^2 & L_1^2 & L_2^2 & L_3^2 \\ L_0^3 & L_1^3 & L_2^3 & L_3^3 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Já vimos que a condição 2.8 é equivalente a $L^t g L = g$ e também a $L g L^t = g$. Dessa forma, o grupo de isometrias de \mathcal{M} está contido no grupo de todas as matrizes satisfazendo tais condições, que chamamos no capítulo anterior de *grupo de Lorentz*. Reciprocamente, dada uma transformação L satisfazendo $L^t g L = g$, temos que ela preserva a quantidade g :

$$g(x, y) = g_{ij} x^i y^j = g_{ij} L_k^i \tilde{x}^k L_l^j \tilde{y}^l = g_{ij} L_k^i L_l^j \tilde{x}^k \tilde{y}^l = g_{kl} \tilde{x}^k \tilde{y}^l = g(\tilde{x}, \tilde{y}).$$

Com isso, concluímos que o grupo de Lorentz coincide com o grupo de isometrias de \mathcal{M} . Lembrando da analogia que fizemos entre bases ortonormais e observadores inerciais, acabamos de provar que a quantidade g independe do observador inercial.

A discussão feita acima sobre as quantidades t_1 e t_2 nos motiva a fazer as seguintes definições.

Definição 3.4. Um vetor $x \in \mathcal{M}$ é dito *tipo tempo*, *tipo espaço* ou *tipo luz* se $g(x, x) > 0$, $g(x, x) < 0$ ou $g(x, x) = 0$, respectivamente. Dizemos também que $x, y \in \mathcal{M}$ são *ortogonais* se $g(x, y) = 0$.

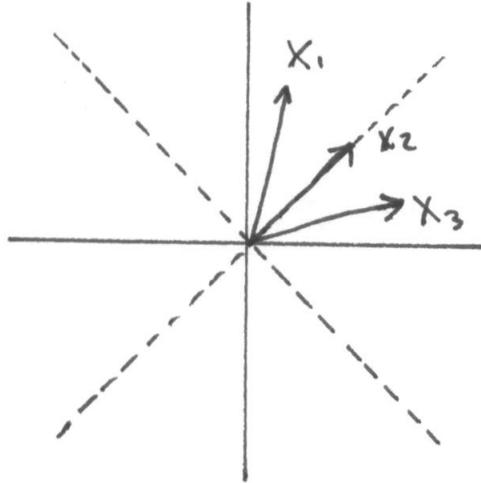


Figura 3.2: Vetores tipo tempo, luz e espaço.

Vimos no capítulo anterior que se L é uma transformação de Lorentz, então $\det(L) = \pm 1$. Dizemos que L é *própria* se $\det(L) = 1$ e *imprópria* caso contrário. Lembremos que não podemos ter $L_0^0 = 0$. De fato, fazendo $k = l = 0$ em 2.8 temos que $(L_0^0)^2 = 1 + (L_0^1)^2 + (L_0^2)^2 + (L_0^3)^2$. Dizemos que L é *ortocrônica*² se $L_0^0 > 0$ e *não-ortocrônica* se $L_0^0 < 0$. Para vermos o porquê desse nome, provemos um resultado auxiliar.

Teorema 3.5. Sejam x tipo tempo, y tipo tempo ou luz, $\{e_i\}$ base ortonormal de \mathcal{M} com $x = x^i e_i$ e $y = y^i e_i$. Então $x^0 y^0 > 0$ (e nesse caso, $g(x, y) > 0$) ou $x^0 y^0 < 0$ (e nesse caso, $g(x, y) < 0$).

²Dos prefixos gregos orthos (correto) e chronos (tempo).

Demonstração. Por hipótese, temos que

$$g(x, x) = x^0x^0 - x^1x^1 - x^2x^2 - x^3x^3 = x^0x^0 - \sum_{i=1}^3 x^i x^i > 0$$

$$g(y, y) = y^0y^0 - y^1y^1 - y^2y^2 - y^3y^3 = y^0y^0 - \sum_{i=1}^3 y^i y^i \geq 0,$$

onde a soma é sobre os índices 1, 2 e 3.³

Dessa forma, temos $x^0x^0 > \sum_{i=1}^3 x^i x^i$ e $y^0y^0 \geq \sum_{i=1}^3 y^i y^i$ e, portanto, $x^0x^0y^0y^0 > \sum_{i=1}^3 x^i x^i y^i y^i$. Tomando a raiz quadrada dos dois lados, obtemos que

$$|x^0y^0| > \left(\sum_{i=1}^3 x^i x^i y^i y^i \right)^{1/2}. \quad (3.2)$$

Note agora que, para todo $t \in \mathbb{R}$ vale que

$$0 \leq (ty^1 + x^1)^2 + (ty^2 + x^2)^2 + (ty^3 + x^3)^2 = t^2 \left(\sum_{i=1}^3 y^i y^i \right) + 2t \left(\sum_{i=1}^3 x^i y^i \right) + \left(\sum_{i=1}^3 x^i x^i \right).$$

Portanto, a equação do segundo grau não tem raízes reais distintas, ou seja, $4\left(\sum_{i=1}^3 x^i y^i\right)^2 - 4\left(\sum_{i=1}^3 y^i y^i\right)\left(\sum_{i=1}^3 x^i x^i\right) \leq 0$ e, portanto,

$$\left| \sum_{i=1}^3 x^i y^i \right| \leq \left(\sum_{i=1}^3 x^i x^i y^i y^i \right)^{1/2}. \quad (3.3)$$

Juntando o resultado das equações 3.2 e 3.3, temos que

$$|x^0y^0| > \left| \sum_{i=1}^3 x^i y^i \right|,$$

e, portanto, $x^0y^0 \neq 0$ e $g(x, y) \neq 0$. Caso $x^0y^0 > 0$ temos que $x^0y^0 > \left| \sum_{i=1}^3 x^i y^i \right| \geq \sum_{i=1}^3 x^i y^i$ e, portanto, $g(x, y) > 0$. Caso tenhamos $x^0y^0 < 0$ trocamos y por $-y$ que a primeira condição será satisfeita. Dessa forma teremos que $g(x, -y) > 0$, ou seja, $g(x, y) < 0$. \square

Corolário 3.6. *Se um vetor não nulo em \mathcal{M} é ortogonal a um vetor tipo tempo então ele deve ser tipo espaço.*

Demonstração. Seja $y \neq 0$ ortogonal a x tipo tempo. Suponha que y não seja do tipo espaço. Então ele é tipo luz ou tempo. Estando nas condições do Teorema 3.5, temos que $g(x, y) > 0$ ou $g(x, y) < 0$, contradição. \square

Esse Teorema diz que para vetores x e y como na hipótese, o produto x^0y^0 tem o mesmo sinal em qualquer base ortonormal. Isso quer dizer que se os números têm o mesmo sinal (ou seja, ambos os vetores apontam “para cima” ou “para baixo”) então isso acontecerá em qualquer base ortonormal e

³Alguns livros usam a convenção de letras latinas para índices tomando os valores 0, 1, 2 e 3 e letras gregas para 1, 2 e 3 ou vice-versa. Não seguirei aqui essa convenção, pois acho ela confusa quando é necessário comparar textos. Já basta nos preocuparmos com a diferença no sinal da métrica. Nesse texto, a notação de Einstein indica a soma nos índices de 0 até 4. Quando somarmos em outros índices, usaremos o símbolo de somatório.

se os sinais são opostos (ou seja, um vetor aponta “para cima” e outro aponta “para baixo”) então isso acontecerá em qualquer base ortonormal. Repare que isso não quer dizer que separadamente eles terão o mesmo sinal (seria o caso de que se um vetor aponta para cima ou para baixo em uma certa base ortonormal então ele o faz em qualquer base ortonormal. Isso claramente está errado). É possível que em uma base ortonormal um vetor do tipo tempo ou luz *aponte para o futuro* (ou seja, $x^0 > 0$) e em outra *aponte para o passado*⁴ (isto é, $x^0 < 0$). As bases nas quais o vetor sempre aponta para a mesma direção temporal são justamente aquelas obtidas através de mudanças de coordenadas dadas pelas transformações de Lorentz ortocrônicas, como veremos no próximo Teorema.

Teorema 3.7. *Seja L uma transformação de Lorentz e $\{e_i\}$ uma base ortonormal para \mathcal{M} . Então são equivalentes:*

1. L é ortocrônica;
2. L preserva orientação no tempo de todos os vetores do tipo luz diferentes de zero, ou seja, para todo $0 \neq x = x^i e_i$ do tipo luz os números x^0 e $\tilde{x}^0 = L_1^0 x^1$ têm o mesmo sinal;
3. L preserva orientação no tempo de todos os vetores do tipo tempo diferentes de zero.

Demonstração. Seja $0 \neq x = x^i e_i$ um vetor tipo tempo ou luz. Pela desigualdade de Cauchy⁵-Schwarz⁶-Bunyakovskii⁷ em \mathbb{R}^3 temos que

$$(L_1^0 x^1 + L_2^0 x^2 + L_3^0 x^3)^2 \leq ((L_1^0)^2 + (L_2^0)^2 + (L_3^0)^2)((x^1)^2 + (x^2)^2 + (x^3)^2). \quad (3.4)$$

Agora, fazendo $i = j = 0$ em 2.12 temos que

$$1 = g^{00} = (L_0^0)^2 - (L_1^0)^2 - (L_2^0)^2 - (L_3^0)^2 \Rightarrow (L_0^0)^2 > (L_1^0)^2 + (L_2^0)^2 + (L_3^0)^2. \quad (3.5)$$

Como x é tipo tempo ou luz, $g(x, x) \geq 0$ e, portanto, $(x^0)^2 \geq (x^1)^2 + (x^2)^2 + (x^3)^2$. Juntamente a equação 3.4 com as duas desigualdades obtidas e com a hipótese de $x \neq 0$, concluímos que

$$(L_1^0 x^1 + L_2^0 x^2 + L_3^0 x^3)^2 < (L_0^0 x^0)^2,$$

que podemos escrever como

$$(L_1^0 x^1 + L_2^0 x^2 + L_3^0 x^3)^2 - (L_0^0 x^0)^2 < 0 \Rightarrow (L_1^0 x^1 + L_2^0 x^2 + L_3^0 x^3 - L_0^0 x^0)(L_1^0 x^1 + L_2^0 x^2 + L_3^0 x^3 + L_0^0 x^0) < 0. \quad (3.6)$$

Defina $y \in \mathcal{M}$ como $y = L_0^0 e_0 + L_1^0 e_1 + L_2^0 e_2 + L_3^0 e_3$. Pela desigualdade 3.5, y é tipo tempo. Com isso, podemos escrever 3.6 como

$$(-g(x, y))\tilde{x}^0 < 0 \Rightarrow g(x, y)\tilde{x}^0 > 0.$$

Dessa forma, os números $g(x, y)$ e \tilde{x}^0 têm o mesmo sinal. Mostremos que $L_0^0 > 0$ se e somente se x^0 e \tilde{x}^0 têm o mesmo sinal. De fato, se $L_0^0 > 0$ e $x^0 > 0$, então $x^0 y^0 = x^0 L_0^0 > 0$ e, pelo Teorema acima, $g(x, y) > 0$, implicando em $\tilde{x}^0 > 0$. Analogamente, se $x^0 < 0$, então $\tilde{x}^0 < 0$. Para provar que se x^0 e \tilde{x}^0 têm o mesmo sinal então $L_0^0 > 0$, suponha que $L_0^0 < 0$. Então por um argumento similar ao feito acima concluímos que x^0 e \tilde{x}^0 têm sinais opostos. \square

⁴Apontar para o passado e para o futuro é um jeito mais simpático de dizer aponta para cima ou para baixo.

⁵Augustin-Louis Cauchy, Paris 21/08/1789 - Sceaux 23/05/1857.

⁶Karl Hermann Amand Schwarz, Hermsdorf 25/01/1843 - Berlin 30/11/1921.

⁷Viktor Yakovlevich Bunyakovskii, Bar 16/12/1804 - São Petersburgo 12/12/1889.

Esse teorema dá uma explicação do nome que demos para essa classe de transformações, pois elas preservam orientação temporal de vetores do tipo luz ou tipo tempo. Note que agora temos um conjunto de bases ortonormais mais conveniente. Fixada uma base ortonormal arbitrária, dizemos que uma outra base é *admissível* se a transformação que leva uma base na outra é uma transformação de Lorentz ortocrônica. Fisicamente, isso quer dizer que, fixada uma orientação temporal no espaço-tempo, todos os observadores inerciais devem estar de acordo com essa orientação.

Outra classe importante de transformações de Lorentz são as *rotações*. Tais transformações são da seguinte forma:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & [L_j^i] & & \\ 0 & & & \end{pmatrix},$$

onde $[L_j^i], i, j = 1, 2, 3$ é uma matriz ortogonal com determinante 1. Essas transformações representam rotações dos eixos coordenados espaciais em um dado referencial. A seguinte proposição caracteriza as rotações.

Lema 3.8. *Seja L uma transformação de Lorentz própria e ortocrônica. Então são equivalentes:*

1. L é uma rotação;
2. $L_0^1 = L_0^2 = L_0^3 = 0$;
3. $L_1^0 = L_2^0 = L_3^0 = 0$;
4. $L_0^0 = 1$.

Demonstração. Fazendo $k = l = 0$ em 2.8 temos

$$(L_0^0)^2 - (L_0^1)^2 - (L_0^2)^2 - (L_0^3)^2 = 1.$$

Analogamente, $i = j = 0$ em 2.12 implica que

$$(L_0^0)^2 - (L_1^0)^2 - (L_2^0)^2 - (L_3^0)^2 = 1.$$

A equivalência entre 2, 3 e 4 segue imediatamente das igualdades acima e da hipótese que L é ortocrônica. Como uma rotação claramente satisfaz 2, 3 e 4, resta mostrar que se L satisfaz qualquer uma dessas condições, então L é uma rotação. De fato, se L satisfaz essas condições, então $\det(L) = 1$. A Fórmula 2.10 para a inversa de L nos diz que

$$L^{-1} = \begin{pmatrix} L_0^0 & -L_0^1 & -L_0^2 & -L_0^3 \\ -L_1^0 & L_1^1 & L_1^2 & L_1^3 \\ -L_2^0 & L_2^1 & L_2^2 & L_2^3 \\ -L_3^0 & L_3^1 & L_3^2 & L_3^3 \end{pmatrix},$$

que no nosso caso se reduz a

$$L^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & L_1^1 & L_1^2 & L_1^3 \\ 0 & L_2^1 & L_2^2 & L_2^3 \\ 0 & L_3^1 & L_3^2 & L_3^3 \end{pmatrix},$$

provando que $[L_j^i], i, j = 1, 2, 3$ é uma matriz ortogonal □

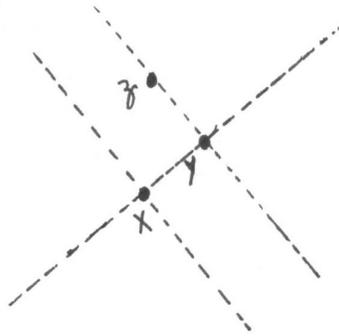


Figura 3.3: Não-transitividade da relação $<$.

Apesar de não utilizarmos esse resultado no trabalho, as transformações de Lorentz, transformações de Lorentz próprias, transformações de Lorentz ortocrônicas, transformações de Poincaré (que são transformações de Lorentz transladadas) têm estrutura de grupo. O estudo desses grupos e suas representações é de extrema importância em Teoria Quântica de Campos. Veja [Liv:Zee10].

3.4 Vetores especiais têm propriedades especiais

Nessa seção estudaremos propriedades específicas das três classes de vetores que definimos. Começemos pelos vetores tipo luz.

3.4.1 Vetores do tipo luz

Dados dois eventos distintos x e x_0 , considere o *vetor deslocamento* $x - x_0$ de x_0 para x . Se tal vetor é do tipo luz, então, relativo a qualquer base ortonormal, temos que

$$(x^0 - x_0^0)^2 - (x^1 - x_0^1)^2 - (x^2 - x_0^2)^2 - (x^3 - x_0^3)^2 = 0.$$

Não podemos ter $x^0 - x_0^0 = 0$ e tal número, relativo a qualquer sistema de coordenadas admissível, sempre terá o mesmo sinal. Lembrando a definição feita mais acima, dizemos que o vetor $x - x_0$ está *direcionado para o futuro* se $x^0 - x_0^0 > 0$ e *direcionado para o passado* se $x^0 - x_0^0 < 0$. Definimos então uma relação binária em \mathcal{M} da seguinte forma:

$$x_0 < x \text{ se e somente se } x - x_0 \text{ é tipo luz e direcionado para o futuro.}$$

A interpretação física dessa ordem é que fenômenos eletromagnéticos iniciados em x_0 podem influenciar x . Em particular, um raio de luz pode ser enviado de x_0 para x . Note que essa relação não é transitiva. De fato, tomando $\{e_i\}$ uma base admissível para \mathcal{M} e $x = (0, 0, 0, 0)$, $y = (1, 1, 0, 0)$ e $z = (2, 0, 0, 0)$ escritos nessa base, verifica-se que $x < y$, $y < z$ porém $x \not< z$. Veja a Figura 3.3. Veremos mais adiante em quais circunstâncias vale a transitividade.

Dados x e x_0 eventos tais que $g(x - x_0, x - x_0) = 0$, definimos o *raio de luz* através de x_0 e x como

$$R_{x_0, x} = \{x_0 + t(x - x_0), t \in \mathbb{R}\}.$$

Definimos também o *cone de luz* com vértice em x_0 como

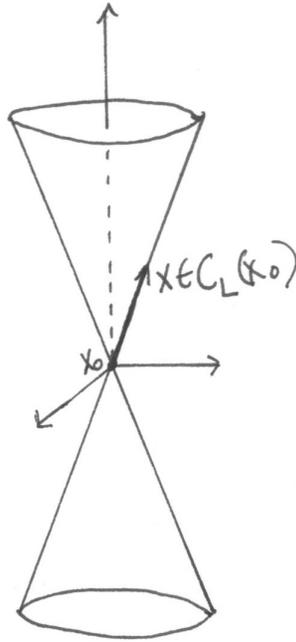


Figura 3.4: Cone de luz.

$$C_L(x_0) = \{x \in \mathcal{M} \mid g(x - x_0, x - x_0) = 0\}.$$

Uma propriedade importante que vetores tipo luz possuem é provada no seguinte lema.

Lema 3.9. *Dois vetores não-nulos do tipo luz $x, y \in \mathcal{M}$ são ortogonais se e somente se são paralelos.*

Demonstração. Claramente se x e y são paralelos então são ortogonais. Provemos a volta. Fixe $\{e_i\}$ base ortonormal de \mathcal{M} . Por hipótese, $x = x^i e_i$ e $y = y^j e_j$ são ortogonais. Suponha, primeiramente, que x^0 e y^0 têm o mesmo sinal. Portanto:

$$0 = g(x, y) = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3 \Rightarrow |x^0| |y^0| = x^0 y^0 = (x^1 y^1 + x^2 y^2 + x^3 y^3). \quad (3.7)$$

Como x e y são tipo luz, temos que:

$$\begin{aligned} 0 = g(x, x) &= (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \Rightarrow (x^0)^2 = (x^1)^2 + (x^2)^2 + (x^3)^2. \\ 0 = g(y, y) &= (y^0)^2 - (y^1)^2 - (y^2)^2 - (y^3)^2 \Rightarrow (y^0)^2 = (y^1)^2 + (y^2)^2 + (y^3)^2. \end{aligned}$$

Juntando com a Equação 3.7, temos:

$$(x^1 y^1 + x^2 y^2 + x^3 y^3) = ((x^1)^2 + (x^2)^2 + (x^3)^2)^{1/2} ((y^1)^2 + (y^2)^2 + (y^3)^2)^{1/2}.$$

Portanto, existe $\lambda \in \mathbb{R}$ tal que $(x^1, x^2, x^3) = \lambda(y^1, y^2, y^3)$, pela desigualdade de Cauchy-Schwarz-Bunyakovskii em \mathbb{R}^3 . Basta provarmos que $x^0 = \lambda y^0$. Novamente pela Equação 3.7 e por y ser tipo luz e não-nulo temos:

$$x^0 y^0 = (x^1 y^1 + x^2 y^2 + x^3 y^3) = \lambda((y^1)^2 + (y^2)^2 + (y^3)^2) = \lambda(y^0)^2 \Rightarrow x^0 = \lambda(y^0).$$

Caso x^0 e y^0 tenham sinais opostos, troque y por $-y$. Dessa forma x^0 e $-y^0$ têm o mesmo sinal, ainda valem as hipóteses do lema e concluímos que x e $-y$ são paralelos. \square

3.4.2 Vetores do tipo tempo

Sejam agora dois eventos x e x_0 tais que $g(x - x_0, x - x_0) > 0$, ou seja, o vetor deslocamento é tipo tempo. Nosso objetivo é mostrar que existe um observador inercial que vê esses dois eventos acontecendo no mesmo lugar. Primeiramente fixe $\{e_i\}$ uma base ortonormal para \mathcal{M} . Escreva $x = x^i e_i$ e $x_0 = x_0^j e_j$. Matematicamente, queremos mostrar que existe uma base admissível $\{\bar{e}_i\}$ na qual $\bar{x}^i - \bar{x}_0^i = 0$, $i = 1, 2, 3$, onde $x = \bar{x}^i \bar{e}_i$ e $x_0 = \bar{x}_0^j \bar{e}_j$.

Isso é razoável, pois como $g(x - x_0, x - x_0) > 0$, então

$$\frac{\sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2}}{(x^0 - x_0^0)} < 1.$$

Se lembrarmos da escolha de unidade que fizemos na definição da métrica, temos que a velocidade da luz é 1. Dessa forma, é de se esperar que um observador que se move com velocidade $\frac{\sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2}}{(x^0 - x_0^0)}$ paralelo à reta que liga (x_0^1, x_0^2, x_0^3) a (x^1, x^2, x^3) na superfície de simultaneidade definida pela observador inercial referente a base $\{e_i\}$ verá os eventos acontecendo no mesmo lugar.

De fato, seja $v = \frac{\sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2}}{(x^0 - x_0^0)}$ e sejam d^1, d^2 e d^3 os co-senos diretores da reta orientada que liga (x_0^1, x_0^2, x_0^3) a (x^1, x^2, x^3) . Cada d^i é dado por $d^i = (x^i - x_0^i) / \sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2}$. Seja L a transformação de Lorentz cuja primeira linha é dada por $L_0^0 = \gamma(v) = 1 / \sqrt{1 - v^2}$ e $L_i^0 = -v\gamma(v)d^i$, para $i = 1, 2, 3$.⁸ Calculemos a quantidade $\bar{x}^0 - \bar{x}_0^0$. Note que podemos escrever $\gamma(v)$ como $(x^0 - x_0^0) / \sqrt{g(x - x_0, x - x_0)}$.

$$\begin{aligned} \bar{x}^0 - \bar{x}_0^0 &= L_0^0(x^0 - x_0^0) + L_1^0(x^1 - x_0^1) + L_2^0(x^2 - x_0^2) + L_3^0(x^3 - x_0^3) = \\ &= \gamma(v)(x^0 - x_0^0) - v\gamma(v)(d^1(x^1 - x_0^1) + d^2(x^2 - x_0^2) + d^3(x^3 - x_0^3)) = \\ &= \gamma(v)(x^0 - x_0^0) - v\gamma(v)\sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2} = \\ &= \frac{(x^0 - x_0^0)^2}{\sqrt{g(x - x_0, x - x_0)}} - \frac{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2}{\sqrt{g(x - x_0, x - x_0)}} = \\ &= \sqrt{g(x - x_0, x - x_0)}. \end{aligned}$$

Dessa forma, temos que $g(x - x_0, x - x_0) = (\bar{x}^0 - \bar{x}_0^0)^2$. Por outro lado, $g(x - x_0, x - x_0) = (\bar{x}^0 - \bar{x}_0^0)^2 - (\bar{x}^1 - \bar{x}_0^1)^2 - (\bar{x}^2 - \bar{x}_0^2)^2 - (\bar{x}^3 - \bar{x}_0^3)^2$. Com isso, concluímos que $\bar{x}^1 - \bar{x}_0^1 = \bar{x}^2 - \bar{x}_0^2 = \bar{x}^3 - \bar{x}_0^3 = 0$.

Isso nos motiva a seguinte definição: Para dois eventos x e x_0 tais que o vetor deslocamento $x - x_0$ é do tipo tempo, definimos a *duração* $\tau(x - x_0)$ de $x - x_0$ como

$$\tau(x - x_0) = \sqrt{g(x - x_0, x - x_0)}.$$

⁸Já veremos que tal transformação sempre existe.

Acabamos de ver que essa quantidade deve ser interpretada como o intervalo de tempo entre x e x_0 em relação a um observador que vê ambos eventos ocorrendo no mesmo ponto do espaço.

Note que uma possibilidade para a transformação de Lorentz utilizada acima é dada por

$$\begin{pmatrix} \gamma(v) & -v\gamma(v)d^1 & -v\gamma(v)d^2 & -v\gamma(v)d^3 \\ -v\gamma(v)d^1 & (\gamma(v)-1)d^1d^1+1 & (\gamma(v)-1)d^1d^2 & (\gamma(v)-1)d^1d^3 \\ -v\gamma(v)d^2 & (\gamma(v)-1)d^2d^1 & (\gamma(v)-1)d^2d^2+1 & (\gamma(v)-1)d^2d^3 \\ -v\gamma(v)d^3 & (\gamma(v)-1)d^3d^1 & (\gamma(v)-1)d^3d^2 & (\gamma(v)-1)d^3d^3+1 \end{pmatrix}.$$

Com esse argumento, podemos provar que para qualquer reta da forma $\{x_0 + t(x - x_0), t \in \mathbb{R}\}$, onde $x - x_0$ é tipo tempo, existe uma base admissível na qual todos os eventos dessa reta ocorrem no mesmo ponto espacial. Essa reta pode ser interpretada como um assistente de um observador inercial, que está em repouso em relação a este. Uma reta como essa acima é chamada de *reta tipo tempo*. Se uma reta tipo tempo passa pela origem, podemos provar ainda mais: ela pode ser identificada com o conjunto de eventos na história de algum observador admissível, ou seja, com o eixo x_0 de alguma base admissível para \mathcal{M} . Para provarmos isso, tome um vetor \hat{e}_0 sobre a reta, que chamaremos de T , tal que $g(\hat{e}_0, \hat{e}_0) = 1$ e seja $[\hat{e}_0]$ o espaço gerado por tal vetor. Como conjunto, temos que $T = [\hat{e}_0]$. Seja $[\hat{e}_0]^\perp$ o complemento ortogonal de $[\hat{e}_0]$. Afirmamos que $\mathcal{M} = [\hat{e}_0] \oplus [\hat{e}_0]^\perp$. De fato, tome $x \in \mathcal{M}$ e considere o vetor $v = x - g(x, \hat{e}_0)\hat{e}_0$. Tal vetor está em $[\hat{e}_0]^\perp$, pois $g(v, \hat{e}_0) = g(x, \hat{e}_0) - g(x, \hat{e}_0) = 0$. Como podemos escrever $x = v + g(x, \hat{e}_0)\hat{e}_0$, temos que $\mathcal{M} = [\hat{e}_0] + [\hat{e}_0]^\perp$. A soma é direta pois $[\hat{e}_0] \cap [\hat{e}_0]^\perp = \{0\}$. Como todo vetor de $[\hat{e}_0]$ é tipo tempo, todo vetor de $[\hat{e}_0]^\perp$ é tipo espaço, pelo Corolário 3.6. Dessa forma, a restrição de g a $[\hat{e}_0]^\perp$ é negativo definido, implicando assim na existência de uma base $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ satisfazendo $g(\hat{e}_i, \hat{e}_j) = -\delta_i^j$. Com isso, $\{\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3\}$ é uma base ortonormal de \mathcal{M} , não necessariamente admissível. Para resolver isso, note que as colunas da matriz de mudança de base consiste nas coordenadas dos vetores $\{\hat{e}_i\}$ escritos na base $\{e_i\}$. Mudando alguns \hat{e}_i por $-\hat{e}_i$, se necessário, podemos converter tal base em uma base admissível.

A quantidade $\tau(x - x_0)$ é a menor cota inferior para o intervalo de tempo entre os eventos x_0 e x :

Proposição 3.10. *Se $x - x_0$ é tipo tempo e s é um número real arbitrário não-negativo, então existe um referencial admissível no qual a separação espacial entre x_0 e x é s . Além disso, a separação temporal entre x_0 e x pode assumir qualquer valor maior ou igual que $\tau(x - x_0)$.*

Demonstração. Como vimos acima, seja $\{e_i\}$ um referencial admissível no qual $x = x^i e_i$, $x_0 = x_0^j e_j$, $x^1 - x_0^1 = x^2 - x_0^2 = x^3 - x_0^3 = 0$ e $\tau(x - x_0) = x^0 - x_0^0$. Considere a seguinte transformação de Lorentz especial, como no capítulo anterior:

$$L = \begin{pmatrix} \cosh(\theta) & \sinh(\theta) & 0 & 0 \\ \sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Temos então que

$$L(x_0 - x) = \begin{pmatrix} \cosh(\theta) & \sinh(\theta) & 0 & 0 \\ \sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 - x_0^0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \cosh(\theta)(x^0 - x_0^0) \\ \sinh(\theta)(x^0 - x_0^0) \\ 0 \\ 0 \end{pmatrix}.$$

Dessa forma, a separação temporal e espacial nesse novo referencial são, respectivamente, $\cosh(\theta)(x^0 - x_0^0)$ e $\sinh(\theta)(x^0 - x_0^0)$. Como $\min \cosh$ é 1 quando $\theta = 0$, o menor valor possível para a separação temporal é $\tau(x - x_0) = x^0 - x_0^0$. Por outro lado, \sinh assume todos os valores reais, implicando que $\sinh(\theta)(x^0 - x_0^0)$ também o faz. \square

Se $x - x_0$ é tipo tempo, então $(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2 < (x^0 - x_0^0)^2$, ou seja, $x - x_0$ está "dentro" do cone de luz em x_0 . Definimos o cone de tempo em x_0 como $C_T(x_0) = \{x \in \mathcal{M} \mid g(x - x_0) > 0\}$. Para um vetor $x \in C_T(x_0)$ temos $x^0 - x_0^0 \neq 0$, e o sinal de tal número independe do observador admissível, como vimos no Teorema 3.7. Como fizemos para vetores do tipo luz, dizemos que $x - x_0$ é direcionado para o futuro se $x^0 - x_0^0 > 0$ e é direcionado para o passado se $x^0 - x_0^0 < 0$. Definimos agora outra relação binária em \mathcal{M} da seguinte forma:

$x_0 \ll x$ se e somente se $x - x_0$ é tipo tempo e direcionado para o futuro.

Com isso, podemos dividir o cone de tempo em x_0 em duas partes: o cone de tempo futuro $C_T^+(x_0) = \{x \in \mathcal{M} \mid x_0 \ll x\}$ e o cone de tempo passado $C_T^-(x_0) = \{x \in \mathcal{M} \mid x \ll x_0\}$.

Nosso objetivo agora é provar análogos da desigualdade triangular e da desigualdade de Cauchy-Schwarz-Bunyakovskii para o espaço de Minkowski. Para isso, precisamos provar dois resultados auxiliares.

Lema 3.11. *A soma de finitos vetores do tipo tempo ou luz direcionados para o futuro é tipo tempo e direcionada para o futuro, exceto quanto todos os vetores são tipo luz e paralelos; nesse caso a soma é tipo luz e direcionada para o futuro.*

Demonstração. Claramente a soma de vetores direcionados para o futuro é direcionada para o futuro. Usando indução, basta provarmos o lema para soma de dois vetores. O caso geral segue de combinações dos casos especiais abaixo.

- x_1, x_2 tipo tempo e direcionados para o futuro: $g(x_1 + x_2, x_1 + x_2) = g(x_1, x_1) + 2g(x_1, x_2) + g(x_2, x_2)$. Cada $g(x_i, x_i)$ é positivo. Como ambos apontam para o futuro, temos que cada x_i^0 é positivo, portanto $x_1^0 x_2^0$ também é positivo. Pelo Teorema 3.5 temos que $g(x_1, x_2)$, e portanto $g(x_1 + x_2, x_1 + x_2)$, são positivos.
- x_1 tipo luz, x_2 tipo tempo e ambos direcionados para o futuro: $g(x_1 + x_2, x_1 + x_2) = g(x_1, x_1) + 2g(x_1, x_2) + g(x_2, x_2)$. O termo $g(x_1, x_1)$ é zero pois x_1 é tipo luz e $g(x_2, x_2)$ é positivo pois x_2 é tipo tempo. Para mostrar que o termo $g(x_1, x_2)$ é positivo usamos o mesmo argumento acima, pois ainda estamos nas hipóteses do Teorema 3.5.
- x_1, x_2 tipo luz e direcionados para o futuro: Note que $g(x_1 + x_2, x_1 + x_2) = 2g(x_1, x_2)$, que é zero se e somente se x_1 e x_2 são paralelos, pelo Lema 3.9. Dessa forma, $x_1 + x_2$ é tipo luz se e somente se x_1 e x_2 são paralelos. Suponha agora que x_1 e x_2 não são paralelos. Defina $y_n \in \mathcal{M}$ da seguinte forma: Seja $\{e_i\}$ uma base admissível para \mathcal{M} , na qual $x_1 = x_1^i e_i$ e $x_2 = x_2^j e_j$. Seja $y_n = (x_1^0 + 1/n)e_0 + x_1^1 e_1 + x_1^2 e_2 + x_1^3 e_3$, para $n \in \mathbb{N}$. Note que cada y_n é direcionado para o futuro e tipo tempo, pois $g(y_n, y_n) = (x_1^0 + 1/n)^2 + (x_1^1)^2 + (x_1^2)^2 + (x_1^3)^2 = 2x_1^0/n + 1/n^2 > 0$. Como $x_2^0 y_n^0 = x_2^0(x_1^0 + 1/n) > 0$, pelo Teorema 3.5 temos que $g(x_2, y_n) > 0$. Mas $g(x_2, y_n) = g(x_2, x_1 + e_0/n) = g(x_2, x_1) + x_2^0/n$. Portanto, $g(x_2, x_1) > -1/nx_2^0$ para todo $n \in \mathbb{N}$. Dessa forma, temos que $g(x_2, x_1) \geq 0$. Mas $g(x_2, x_1) \neq 0$, pois os vetores não são paralelos. Portanto, concluímos que $g(x_2, x_1) > 0$, logo, $x_1 + x_2$ é tipo tempo. □

Considerando $-x_i$ no Lema acima, podemos trocar "direcionado para o futuro" para "direcionado para o passado" para obtermos o resultado análogo.

O Lema acima nos permite mostrar que a relação \ll é transitiva, ao passo que $<$ não o é. De fato, se $x \ll y$ e $y \ll z$, então $y - x$ e $z - y$ são tipo tempo e direcionados para o futuro. Dessa forma, $z - x = (z - y) + (y - x)$ é tipo tempo e direcionado para o futuro. Porém, se $x < y$ e $y < z$ então $y - x$ e $z - y$ são tipo luz e direcionados para o futuro, e $z - x$ é tipo luz se e somente se $z - y$ e $y - x$ são paralelos. A partir desse Lema podemos provar também o análogo ao Teorema 3.5 para vetores tipo luz não nulos e não paralelos.

Teorema 3.12. *Sejam x e y vetores tipo luz não nulos e não paralelos. Então x e y têm a mesma orientação no tempo se e somente se $g(x, y) > 0$.*

Demonstração. Se x e y têm a mesma orientação no tempo, então $x + y$ é tipo tempo e $g(x + y, x + y) = 2g(x, y) > 0$. Para provar o outro lado, provemos a contrapositiva. Se x e y têm orientação oposta no tempo, então x e $-y$ têm a mesma orientação e, portanto, $g(x, -y) > 0$. Dessa forma, $g(x, y) < 0$. \square

A interpretação física da relação \ll é que, se $x \ll y$ então um fenômeno material iniciado em x pode influenciar y . As duas relações \ll e $<$ são chamadas de *relações causais no espaço-tempo*. Pode-se provar o seguinte resultado (demonstração disponível em [Liv:Nab92]):

Teorema. *Para x e y distintos em \mathcal{M} vale que:*

- $x < y$ se e somente se : $x \not\ll y$ e ($y \ll z$ implica $x \ll z$);
- $x \ll y$ se e somente se : $x \not< y$ e ($x < z < y$, para algum $z \in \mathcal{M}$).

Uma bijeção $F : \mathcal{M} \rightarrow \mathcal{M}$ é dita um *automorfismo causal* se preserva a relação $<$, ou seja, $x < y$ se e somente se $F(x) < F(y)$, para todos $x, y \in \mathcal{M}$. Como podemos escrever uma relação em função da outra, F preserva $<$ se e somente se preserva \ll , bastando assim considerar somente uma relação. Translações, multiplicações por escalares não nulos e transformações de Lorentz ortocrônicas e próprias são exemplos de automorfismos causais. Deveríamos imaginar que existem muitos outros, pois a definição de tal classe de funções não nos dá nenhuma pista de que F deve ser linear, sequer contínua! Porém, Zeeman⁹ provou em [Art:Zee64] que todo automorfismo causal é uma composição desses três exemplos. Para uma prova desse fato, veja [Liv:Nab92]. Provemos então os dois resultados desejados.

Teorema 3.13 (Desigualdade de Cauchy-Schwarz-Bunyakovskii invertida). *Se x e y são vetores tipo tempo em \mathcal{M} então $g(x, y)^2 \geq g(x, x)g(y, y)$, com igualdade se e somente se x e y são linearmente dependentes.*

Demonstração. Considere o vetor $u = ax - by$, onde $a = g(x, y)$ e $b = g(x, x)$. Note que $g(u, x) = g(ax - by, x) = ag(x, x) - bg(x, y) = 0$. Como x é tipo tempo, então u é tipo espaço ou nulo, pelo Corolário 3.6. Dessa forma, temos que $0 \geq g(u, u) = g(ax - by, ax - by) = a^2g(x, x) - 2abg(x, y) + b^2g(y, y)$, com igualdade se e somente se $u = 0$. Com isso, $2abg(x, y) \geq a^2g(x, x) + b^2g(y, y)$, ou seja,

$$\begin{aligned} 2g(x, y)^2g(x, x) &\geq g(x, y)^2g(x, x) + g(x, x)^2g(y, y) \\ \Leftrightarrow 2g(x, y)^2 &\geq g(x, y)^2 + g(x, x)g(y, y) \\ \Leftrightarrow g(x, y)^2 &\geq g(x, x)g(y, y), \end{aligned}$$

com igualdade se e somente se $u = 0$. Mas $u = 0$ implica que $ax = by$. Como $a = g(x, y) \neq 0$ pelo Corolário 3.6 e $b = g(x, x) \neq 0$ pois x é tipo tempo, temos que x e y são linearmente dependentes. \square

Teorema 3.14 (Desigualdade Triangular invertida). *Se x e y são vetores do tipo tempo com a mesma orientação no tempo, então $\tau(x + y) \geq \tau(x) + \tau(y)$, com igualdade se e somente se x e y são linearmente dependentes.*

Demonstração. Pela desigualdade de Cauchy-Schwarz-Bunyakovskii invertida temos que $g(x, y)^2 \geq g(x, x)g(y, y)$, ou seja, $|g(x, y)| \geq \sqrt{g(x, x)}\sqrt{g(y, y)}$, pois ambos $g(x, x)$ e $g(y, y)$ são positivos. Como x e y têm a mesma orientação no tempo, $g(x, y) \geq 0$, pelo Teorema 3.5. Dessa forma, $g(x, y) \geq \sqrt{g(x, x)}\sqrt{g(y, y)}$.

Note agora que

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

⁹Sir Erik Christopher Zeeman, Japão 04/02/1925

o que implica que $\sqrt{g(x+y, x+y)} \geq \sqrt{g(x, x)} + \sqrt{g(y, y)}$, ou seja, $\tau(x+y) \geq \tau(x) + \tau(y)$.

Se vale igualdade, podemos fazer o velho truque de ler a sequência lógica acima ao contrário e concluir que $g(x, y) = \sqrt{g(x, x)}\sqrt{g(y, y)}$ e, pela desigualdade de Cauchy-Schwarz-Bunyakovskii invertida temos que x e y são linearmente dependentes. \square

O paradoxo dos gêmeos, discutido no final do Capítulo 2, pode ser visto como uma consequência direta da desigualdade triangular invertida. De fato, suponha que os gêmeos Pedro e Bob viajam de acordo com o diagrama abaixo, isto é, Pedro chega até a ponta do vetor x ficando parado e Bob chega ao mesmo ponto primeiro “percorrendo” o vetor y_1 e logo em seguida o vetor y_2 . Note que $x = y_1 + y_2$. Mas pela desigualdade triangular invertida, temos que $\tau(x) \geq \tau(y_1) + \tau(y_2)$. Lembrando a interpretação que demos para τ , concluímos que, em relação ao irmão que representa um observador inercial (Pedro), o tempo próprio de Bob foi maior do que o seu.

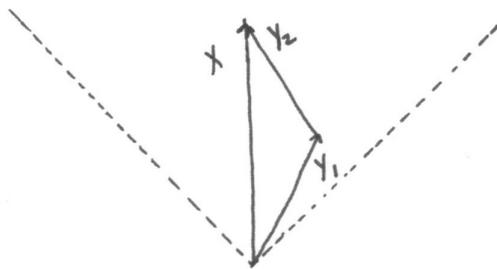


Figura 3.5: Gêmeos viajando.

Note que tanto a desigualdade de Cauchy-Schwarz-Bunyakovskii quanto a desigualdade triangular se estendem para uma soma finita de vetores do tipo tempo, pelo Lema 3.11.

Vamos generalizar alguns desses conceitos e resultados para curvas no espaço-tempo que representem movimento de partículas não necessariamente livres. Começemos com algumas definições.

Seja $I \subset \mathbb{R}$ um intervalo. Uma função $\alpha : I \rightarrow \mathcal{M}$ é dita uma *curva* em \mathcal{M} . Dado um sistema de coordenadas admissível $\{e_i\}$, podemos escrever $\alpha(t) = (x^0(t), x^1(t), x^2(t), x^3(t))$. Quando as funções $x^i(t)$ são suaves, dizemos que a curva é *suave*. Note que essa definição independe da escolha de coordenadas admissíveis. α é dita *tipo tempo* se $\alpha'(t)$ é um vetor tipo tempo, para todo $t \in I$, e *direcionada para o futuro* se $\alpha'(t)$ o é, para todo $t \in I$. Uma curva suave tipo tempo e direcionada para o futuro em \mathcal{M} é chamada de *história* ou *trajetória de uma partícula material*. Note que essa definição quer dizer que a partícula não viaja mais rápido que a velocidade da luz e, além disso, sempre vai para o futuro, o que é razoável de se esperar. Se α tem a forma $\alpha(t) = x_0 + t(x - x_0)$, onde $x_0 \ll x$, então α é a história de uma partícula *livre*. Isso representa uma versão relativística da Primeira Lei de Newton: partículas livres se movimentam em linha reta. Se tivermos que $x_0 < x$, teremos a história de um fóton. Se α é a trajetória de uma partícula material, definimos o *comprimento* ou *tempo próprio* de α como

$$L(\alpha) = \int_a^b \sqrt{g(\alpha'(t), \alpha'(t))} dt = \int_a^b \sqrt{g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt,$$

em um sistema de coordenadas admissível. Mostra-se facilmente que $L(\alpha)$ não depende da parametrização nem do sistema de coordenadas escolhido. Note que só podemos definir essa quantidade para curvas tipo tempo. Essa quantidade deve ser interpretada como o *intervalo de tempo entre $\alpha(a)$ e $\alpha(b)$ como medido*

pelo relógio que a partícula carrega consigo. Para entender isso, vamos precisar do seguinte resultado auxiliar, que provaremos logo em seguida mas que segue como corolário de um resultado muito mais geral, a ser provado no Capítulo 5:

Teorema 3.15. *Sejam $p, q \in \mathcal{M}$. Então $p \ll q$ se e somente se existe uma curva suave $\alpha : [a, b] \rightarrow \mathcal{M}$ direcionada para o futuro e tipo tempo tal que $\alpha(a) = p, \alpha(b) = q$.*

É fácil provar que se $p \ll q$, então tal curva existe. Basta tomar o segmento de reta ligando p a q . Vamos particionar o intervalo $[a, b]$ em subintervalos da forma $a = t_0 < t_1 < \dots < t_{n-1} < t_n = b$. Pelo Teorema acima, $\alpha(a) = \alpha(t_0) \ll \alpha(t_1) \ll \dots \ll \alpha(t_{n-1}) \ll \alpha(t_n) = \alpha(b)$. Dessa forma, cada vetor $v_i = \alpha(t_i) - \alpha(t_{i-1})$ é tipo tempo e direcionado para o futuro. $\tau(v_i)$ é interpretado como o tempo entre $\alpha(t_{i-1})$ e $\alpha(t_i)$, medido por um observador admissível que está presente em ambos eventos. Se a partícula material que a curva α representa tem velocidade constante entre os eventos $\alpha(t_{i-1})$ e $\alpha(t_i)$, então $\tau(v_i)$ mede o intervalo de tempo entre esses eventos como medido pelo relógio carregado pela partícula. Em qualquer referencial admissível podemos escrever que

$$\tau(v_i) = \sqrt{g(\alpha(t_i) - \alpha(t_{i-1}), \alpha(t_i) - \alpha(t_{i-1}))} = \sqrt{g_{kl} \Delta x_i^k \Delta x_i^l} = \sqrt{g_{kl} \frac{\Delta x_i^k}{\Delta t} \frac{\Delta x_i^l}{\Delta t} \Delta t},$$

onde $\Delta x_i^k = v_i^k = x^k(t_i) - x^k(t_{i-1})$ e $\Delta t = t_i - t_{i-1}$. Pela continuidade de α , podemos escolher os Δt suficientemente pequenos de modo que a velocidade da partícula seja aproximadamente constante nos intervalos $[t_{i-1}, t_i]$. Dessa forma, $\tau(v_i)$ será uma boa aproximação para o intervalo de tempo entre $\alpha(t_{i-1})$ e $\alpha(t_i)$ como medido pela partícula. Dessa forma, a soma

$$\sum_{i=1}^n \sqrt{g_{kl} \frac{\Delta x_i^k}{\Delta t} \frac{\Delta x_i^l}{\Delta t} \Delta t} \quad (3.8)$$

é uma boa aproximação para o intervalo de tempo entre $\alpha(a)$ e $\alpha(b)$ como medido pela partícula. No limite $\max \Delta t \rightarrow 0$, a soma converge para a integral na definição de $L(\alpha)$. Provemos, então, o Teorema. Para isso, precisaremos do seguinte Lema:

Lema 3.16. *Seja $\alpha : (a, b) \rightarrow \mathcal{M}$ uma curva suave, tipo tempo e direcionada para o futuro. Então, dado $t_0 \in (a, b)$, existe $\varepsilon > 0$ tal que:*

- i. $(t_0 - \varepsilon, t_0 + \varepsilon) \subset (a, b)$;
- ii. $\alpha(t) \in C_T^-(\alpha(t_0))$, se $t \in (t_0 - \varepsilon, t_0)$;
- iii. $\alpha(t) \in C_T^+(\alpha(t_0))$, se $t \in (t_0, t_0 + \varepsilon)$.

Demonstração. Seja $\{e_i\}$ uma base admissível para \mathcal{M} e escreva $\alpha(t) = x^i(t)e_i$. Claramente para ε suficientemente pequeno a condição (i) é satisfeita. Provemos que vale (ii), pois o argumento para (iii) é similar. Suponha que não existe $\varepsilon > 0$ de modo que valha (ii), ou seja, para todo $\varepsilon > 0$ existe $t \in (t_0, t_0 + \varepsilon)$ tal que $\alpha(t) \notin C_T^+(\alpha(t_0))$. Podemos então construir uma sequência decrescente $\{t_n\}$, convergindo para t_0 tal que $\alpha(t_n) \notin C_T^+(\alpha(t_0))$, para todo $n \in \mathbb{N}$. Pela definição de $C_T^+(p)$, temos duas possibilidades:

- a. $\alpha(t_n) - \alpha(t_0)$ não é tipo tempo, para todos $n \in \mathbb{N}$;
- b. $\alpha(t_n) - \alpha(t_0)$ é tipo tempo mas não é direcionado para o futuro, para todos $n \in \mathbb{N}$.

Provemos que essas duas possibilidades são impossíveis. De fato, se vale (a), então $\alpha(t_n) - \alpha(t_0)$ é tipo luz ou espaço, ou seja, $g(\alpha(t_n) - \alpha(t_0), \alpha(t_n) - \alpha(t_0)) \leq 0$, para todos $n \in \mathbb{N}$. Como $t_n - t_0$ é sempre positivo, podemos escrever que $g\left(\frac{\alpha(t_n) - \alpha(t_0)}{t_n - t_0}, \frac{\alpha(t_n) - \alpha(t_0)}{t_n - t_0}\right) \leq 0$. Passando o limite $n \rightarrow \infty$, temos que $g(\alpha'(t_0), \alpha'(t_0)) \leq 0$,

contradição, pois $\alpha'(t_0)$ é, por hipótese, tipo tempo.

Para ver que (b) é impossível, note que $\alpha(t_n) - \alpha(t_0)$ é direcionado para o passado e tipo tempo, enquanto que $\alpha'(t_0)$ é direcionado para o futuro e tipo tempo. Pelo Teorema 3.5, temos que $g(\alpha(t_n) - \alpha(t_0), \alpha'(t_0)) < 0$, para todos $n \in \mathbb{N}$. Novamente como $t_n - t_0$ é sempre positivo, podemos escrever $g\left(\frac{\alpha(t_n) - \alpha(t_0)}{t_n - t_0}, \alpha'(t_0)\right) < 0$, para todos $n \in \mathbb{N}$. Passando ao limite, concluímos que $g(\alpha'(t_0), \alpha'(t_0)) \leq 0$, contradição. \square

Agora podemos provar o Teorema.

Demonstração (do Teorema 3.15). Como já foi observado, se $p \ll q$, claramente a curva existe. Basta tomarmos uma reta que liga esses dois pontos. Provemos então a outra parte. Tome $\alpha : [a, b] \rightarrow \mathcal{M}$ suave, direcionada para o futuro e tipo tempo tal que $\alpha(a) = p$, $\alpha(b) = q$. Considere uma extensão de α a um intervalo (A, B) contendo $[a, b]$ de modo que a extensão ainda seja direcionada para o futuro e tipo tempo. Tal extensão sempre existe pois α é suave e $C_T(p)$ é aberto em \mathcal{M} . Pelo Lema anterior, existe $\varepsilon_1 > 0$ tal que

- $(a, a + \varepsilon_1) \subset (A, B)$;
- $\alpha(t)$ está em $C_T^+(p)$, para $t \in (a, a + \varepsilon_1)$.

Seja $t_0 = \sup(a + \varepsilon_1)$. Provemos que $t_0 = B$. Suponha que $A < t_0 < B$. Novamente pelo Lema, existe $\varepsilon > 0$ tal que

- $(t_0 - \varepsilon, t_0 + \varepsilon) \subset (A, B)$;
- $\alpha(t) \in C_T^-(\alpha(t_0))$ para $t \in (t_0 - \varepsilon, t_0)$;
- $\alpha(t) \in C_T^+(\alpha(t_0))$ para $t \in (t_0, t_0 + \varepsilon)$.

Temos 3 possibilidades para $\alpha(t_0)$:

- i. $\alpha(t_0) \in C_T^+(p)$;
- ii. $\alpha(t_0) \in C_L(p)$;
- iii. $\alpha(t_0) \in (C_T^+(p) \cup C_L(p))^c$.

Analisemos cada caso separadamente. Se vale (i.), temos que, para todo $t \in (t_0, t_0 + \varepsilon)$ vale que $\alpha(t) - \alpha(t_0)$ e $\alpha(t_0) - p$ são tipo tempo e direcionado para o futuro, pois $\alpha(t) \in C_T^+(\alpha(t_0))$ e $\alpha(t_0) \in C_T^+(p)$. Portanto, $\alpha(t) - p = (\alpha(t) - \alpha(t_0)) + (\alpha(t_0) - p)$ é tipo tempo e direcionado para o futuro, para todo $t \in (t_0, t_0 + \varepsilon)$. Assim, $\alpha(t) \in C_T^+(p)$, para todo $t \in (t_0, t_0 + \varepsilon)$, contradição com a definição de t_0 .

Se vale (ii.), então $C_T^-(\alpha(t_0)) \cap C_T^+(p) = \emptyset$. Por outro lado, para todo $t \in (t_0 - \varepsilon, t_0)$ temos que $\alpha(t) \in C_T^-(\alpha(t_0))$, pelo lema, e $\alpha(t) \in C_T^+(p)$, por definição de t_0 .

Finalmente, se vale (iii.), como $(C_T^+(p) \cup C_L(p))^c$ é aberto e α é contínua, temos que $\alpha(t) \in (C_T^+(p) \cup C_L(p))^c$, para alguns $t \in (t_0 - \varepsilon, t_0)$. Porém, para tais t deveríamos ter que $\alpha(t) \in C_T^+(p)$, pela definição de t_0 . \square

Podemos também definir um análogo à parametrização por comprimento de arco para curvas tipo tempo. Vejamos porque isso é interessante. Suponha que o domínio de α contenha o 0. Defina a *função comprimento* ou *função tempo próprio* de α como

$$\tau = \tau(t) = \int_0^t \sqrt{g(\alpha'(s), \alpha'(s))} ds.$$

Dessa forma, τ é suave e positiva, pois α é tipo tempo. Sua inversa $t = h(\tau)$ está então bem definida e satisfaz $dh/d\tau = (d\tau/dt)^{-1} = \sqrt{g(\alpha'(t), \alpha'(t))}^{-1} > 0$. Podemos então considerar uma reparametrização de α , dada por $\alpha(\tau) = \alpha(h(\tau)) = (x^0(\tau), x^1(\tau), x^2(\tau), x^3(\tau))$, em um sistema de coordenadas admissível. Essa reparametrização é dita a *parametrização por tempo próprio*. Essa parametrização goza da propriedade que $g(\alpha'(\tau), \alpha'(\tau)) = 1$, como pode ser facilmente verificado.

Façamos um último comentário sobre partículas materiais. Considere uma curva $\alpha : [a, b] \rightarrow \mathcal{M}$ tipo tempo e direcionada para o futuro. Sejam $p = \alpha(a)$ e $q = \alpha(b)$. Vimos que a quantidade $L(\alpha)$ pode ser arbitrariamente aproximada pelas somas 3.8, que podemos escrever como

$$\sum_{i=1}^n \tau(v_i) = \sum_{i=1}^n \tau(\alpha(t_i) - \alpha(t_{i-1})) = \tau(\alpha(t_1) - \alpha(t_0)) + \tau(\alpha(t_2) - \alpha(t_1)) + \cdots + \tau(\alpha(t_n) - \alpha(t_{n-1})).$$

Mas a desigualdade triangular invertida nos diz que essa soma é no máximo

$$\tau(\alpha(t_1) - \alpha(t_0) + \alpha(t_2) - \alpha(t_1) + \cdots + \alpha(t_n) - \alpha(t_{n-1})) = \tau(\alpha(t_n) - \alpha(t_0)) = \tau(q - p).$$

Portanto, no limite $\max \Delta t \rightarrow 0$, deveríamos imaginar que $\tau(q - p) \geq L(\alpha)$, ou seja, o tempo entre p e q é maior para o observador que está parado (em relação a um sistema de coordenadas admissível) que passa pelos dois eventos do que para qualquer outro observador que também passa pelos dois eventos. Isso quer dizer que “relógios que se movimentam contam o tempo mais devagar”. Vimos no capítulo anterior uma explicação razoável para esse efeito. Provemos então esse fato.

Teorema 3.17. *Seja $\alpha : [a, b] \rightarrow \mathcal{M}$ uma curva tipo tempo e direcionada para o futuro. Sejam $p = \alpha(a)$ e $q = \alpha(b)$. Então $L(\alpha) \leq \tau(q - p)$, com igualdade se e somente se α é a trajetória de uma partícula material livre.*

Demonstração. Pelo Teorema 3.15 temos que $p \ll q$. Considere $\{e_i\}$ uma base admissível para \mathcal{M} na qual $p = x_p^0 e_0 + x_p^1 e_1 + x_p^2 e_2 + x_p^3 e_3$, $q = x_q^0 e_0 + x_q^1 e_1 + x_q^2 e_2 + x_q^3 e_3$ e $\alpha(t) = (x^0(t), x^1(t), x^2(t), x^3(t))$. Dessa forma, temos que $\tau(q - p) = \tau((x_q^0 - x_p^0)e_0) = \sqrt{g((x_q^0 - x_p^0)e_0, (x_q^0 - x_p^0)e_0)} = |x_q^0 - x_p^0| = x_q^0 - x_p^0$, pois $q - p$ é direcionado para o futuro.

Considere a função suave $t \mapsto x^0(t)$. Como α é tipo direcionada para o futuro, $(x^0)'(t) > 0$, para todo t . Dessa forma, podemos escrever $t = t(x^0)$ e usar x^0 como parâmetro para α . Nessa parametrização, a curva se escreve como $\alpha(x^0) = (x^0, x^1(x^0), x^2(x^0), x^3(x^0))$. Temos então que:

$$\begin{aligned} L(\alpha) &= \int_{x_p^0}^{x_q^0} \sqrt{g(\alpha'(x^0), \alpha'(x^0))} dx^0 = \int_{x_p^0}^{x_q^0} \sqrt{1 - \left[\left(\frac{dx^1}{dx^0} \right)^2 + \left(\frac{dx^2}{dx^0} \right)^2 + \left(\frac{dx^3}{dx^0} \right)^2 \right]} dx^0 \\ &\leq \int_{x_p^0}^{x_q^0} dx^0 = x_q^0 - x_p^0 = \tau(q - p). \end{aligned}$$

A igualdade vale se e somente se $dx^i/dx^0 = 0$, para $i = 1, 2, 3$. Mas isso quer dizer que x^i é constante, para $i = 1, 2, 3$. Com isso, $\alpha(x^0) = (x^0, x^1, x^2, x^3)$, que é a equação de uma reta. \square

Note que essa é uma versão mais geral do Paradoxo dos Gêmeos no espaço de Minkowski. Tal Teorema nos diz que se quisermos sair do evento $\alpha(a)$ e chegarmos até o evento $\alpha(b)$ (supondo que isso seja possível), então o caminho que leva “mais tempo” é a linha reta que conecta os dois eventos. Dessa forma, Bob, o irmão que viajou no foguete, percorreu uma trajetória que leva “menos tempo”. Note que essas relações de “mais tempo” e “menos tempo” são em relação a algum observador inercial, que por conveniência é escolhido como sendo Pedro, o irmão que ficou na Terra. A lição que tiramos disso é: relaxe se quiser viver mais!

3.4.3 Vetores do tipo espaço

Para concluir o estudo das classes especiais de vetores em \mathcal{M} , estudemos brevemente os vetores tipo espaço. Considere dois eventos x_0 e x de modo que $x_0 - x$ seja tipo espaço. Relativo a qualquer referencial admissível, $(x^0 - x_0^0)^2 < (x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2$, ou seja, $x - x_0$ está *fora* do cone de luz em x_0 . Dessa forma, não existe nenhum observador que passe por ambos eventos (para isso, ele precisaria viajar mais rápido que a velocidade da luz). Por outro lado, podemos provar que existe um observador que vê tais eventos como simultâneos.

De fato, seja $\{e_i\}$ uma base admissível para \mathcal{M} e escreva $x = x^i e_i$ e $x_0 = x_0^j e_j$. Sejam $v = (x^0 - x_0^0) / \sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2} < 1$ e $d^i = (x^i - x_0^i) / \sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2}$ (co-senos diretores da reta orientada ligando x_0 a x , como fizemos no começo da seção sobre vetores do tipo tempo). Considere L uma transformação de Lorentz tal que $L_0^0 = \gamma(v)$ e $L_i^0 = -v\gamma(v)d^i$, para $i = 1, 2, 3$. Lembre que já exibimos uma transformação de Lorentz com essas propriedades. Como $\{\bar{e}_i = L e_i\}$ também é uma base ortonormal de \mathcal{M} , escreva $x = \bar{x}^i \bar{e}_i$ e $x_0 = \bar{x}_0^j \bar{e}_j$. Provemos que $\bar{x}^0 - \bar{x}_0^0 = 0$. De fato, usando que $\gamma(v)$ pode ser escrito como $\sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2} / \sqrt{-g(x - x_0, x - x_0)}$, temos que:

$$\begin{aligned} \bar{x}^0 - \bar{x}_0^0 &= L_0^0(x^0 - x_0^0) + L_1^0(x^1 - x_0^1) + L_2^0(x^2 - x_0^2) + L_3^0(x^3 - x_0^3) = \\ &= \gamma(v)(x^0 - x_0^0) - v\gamma(v)(d^1(x^1 - x_0^1) + d^2(x^2 - x_0^2) + d^3(x^3 - x_0^3)) = \\ &= \gamma(v)(x^0 - x_0^0) - v\gamma(v) \sqrt{(x^1 - x_0^1)^2 + (x^2 - x_0^2)^2 + (x^3 - x_0^3)^2} = \\ &= \text{algumas contas...} = \\ &= 0. \end{aligned}$$

Escrevendo $g(x - x_0, x - x_0)$ nessa base, temos que

$$g(x - x_0, x - x_0) = -(\bar{x}^1 - \bar{x}_0^1)^2 - (\bar{x}^2 - \bar{x}_0^2)^2 - (\bar{x}^3 - \bar{x}_0^3)^2.$$

Dessa forma, $\sqrt{-g(x - x_0, x - x_0)}$ é a distância espacial entre x e x_0 medida por um observador que vê esses eventos como simultâneos. Denotamos essa quantidade por $S(x - x_0)$ e a chamamos de *separação espacial entre x e x_0* .

Podemos provar também um análogo à proposição 3.10 para vetores tipo espaço.

Proposição 3.18. *Se x_0 e x são eventos tais que $x - x_0$ é tipo espaço e s é um número real arbitrário, então existe um referencial admissível no qual a separação temporal entre x_0 e x é s . Em particular, observadores admissíveis não concordam sequer na ordem temporal de x_0 e x . Além disso, $S(x - x_0)$ é uma cota inferior para a distância espacial entre x e x_0 para observadores que vêem tais eventos como simultâneos.*

Demonstração. Comece com um referencial admissível $\{e_i\}$ no qual $x = x^i e_i$, $x_0 = x_0^j e_j$ e $x^0 - x_0^0 = 0$. Considere a transformação de Lorentz padrão dada por

$$L = \begin{pmatrix} \cosh(\theta) & \sinh(\theta) & 0 & 0 \\ \sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Temos então que

$$L(x_0 - x) = \begin{pmatrix} \cosh(\theta) & \sinh(\theta) & 0 & 0 \\ \sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ x^1 - x_0^1 \\ x^2 - x_0^2 \\ x^3 - x_0^3 \end{pmatrix} = \begin{pmatrix} \sinh(\theta)(x^1 - x_0^1) \\ \cosh(\theta)(x^1 - x_0^1) \\ x^2 - x_0^2 \\ x^3 - x_0^3 \end{pmatrix}.$$

Como \sinh toma qualquer valor real, a separação temporal também o faz. Como $\min \cosh = 1$ é atingido quando $\theta = 0$, temos que a distância entre x e x_0 nesse referencial, que vê tais eventos como simultâneos, é dada exatamente por $S(x - x_0)$. \square

3.5 Um pouco de Mecânica Relativística

Vimos na seção anterior, onde discutimos os vetores do tipo tempo, que a história de uma partícula material é dada por uma curva $\alpha : I \rightarrow \mathcal{M}$ suave, tipo tempo e direcionada para o futuro. Porém, uma partícula tem uma informação a mais, que é crucial para seu estudo do ponto de vista físico. Tal informação é a *massa*. Mas como fazemos para medir massa na Relatividade Restrita? Lembremos que no caso newtoniano podemos colidir a partícula em questão com uma partícula teste, de massa previamente conhecida ou fixada como uma unidade de massa, aplicar as leis de conservação de massa, momento e, em certos casos, de energia, e obteremos a massa de nossa partícula. Porém, não sabemos se tais leis valem no caso relativístico. Para resolver esse problema, lembremos que, em baixas velocidades uma transformação de Lorentz é bem aproximada por uma transformação de Galileu. Isso nos diz que em baixas velocidades a mecânica newtoniana é uma boa aproximação para a mecânica relativística. O procedimento a ser adotado agora então é colidir várias vezes a partícula em questão com uma partícula teste, só que com velocidades de colisão cada vez menores. Quanto menor a velocidade, melhor é a aproximação newtoniana à Relatividade. Dessa forma, tomando o limite dessa sequência obtida, obtemos um número que podemos chamar de *massa* da partícula. Do ponto de vista matemático, a massa pode ser definida somente como um número real maior ou igual a zero. É claro que muito da essência física é perdida com essa definição simplificada e a discussão física é necessária para que se entenda esse conceito de maneira adequada.

3.5.1 Partículas Materiais

Uma *partícula material* então é dada por um par (α, m) , onde α é a história da partícula e m é sua massa. É conveniente parametrizarmos α por seu comprimento de arco τ , de modo que tenhamos $g(\alpha'(\tau), \alpha'(\tau)) = 1$. O vetor $U(\tau) = \alpha'(\tau)$ é dito a *4-velocidade* da partícula e $P(\tau) = mU(\tau)$ é dito seu *4-momento*. Outro vetor importante é a *4-aceleração*, definida por $A(\tau) = \alpha''(\tau)$. Note que

$$g(\alpha'(\tau), \alpha'(\tau)) = 1 \Rightarrow 0 = \frac{d}{dt} g(\alpha'(\tau), \alpha'(\tau)) = g(\alpha''(\tau), \alpha'(\tau)) = g(U(\tau), A(\tau)).$$

Assim, pelo Corolário 3.6, temos que A ou é do tipo espaço ou é zero.

Uma outra parametrização adequada (e mais razoável de ser feita por um observador) é através do parâmetro x^0 . Vimos que isso é possível na seção anterior, aonde discutimos os vetores do tipo tempo. Dessa forma, se $\alpha(x^0) = (x^0, x^1(x^0), x^2(x^0), x^3(x^0))$ em algum sistema de coordenadas admissível, então

$$\alpha'(x^0) = \left(1, \frac{dx^1}{dx^0}, \frac{dx^2}{dx^0}, \frac{dx^3}{dx^0} \right) (x^0),$$

e portanto,

$$g(\alpha'(x^0), \alpha'(x^0)) = 1 - \left(\frac{dx^1}{dx^0} \right)^2 - \left(\frac{dx^2}{dx^0} \right)^2 - \left(\frac{dx^3}{dx^0} \right)^2 = 1 - v^2(x^0).$$

Dessa forma, o tempo próprio de α nesse sistema é dado por

$$\tau = \tau(x^0) = \int_0^{x^0} \sqrt{1 - v^2(s)} ds.$$

Mais ainda, a 4-velocidade da partícula é dada por

$$U = \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau} \right) = \left(1, \frac{dx^1}{dx^0}, \frac{dx^2}{dx^0}, \frac{dx^3}{dx^0} \right) \frac{dx^0}{d\tau} = \frac{1}{\sqrt{1 - v^2}} \left(1, \frac{dx^1}{dx^0}, \frac{dx^2}{dx^0}, \frac{dx^3}{dx^0} \right),$$

e seu 4-momento é dado por

$$P = \frac{m}{\sqrt{1 - v^2}} \left(1, \frac{dx^1}{dx^0}, \frac{dx^2}{dx^0}, \frac{dx^3}{dx^0} \right).$$

Simplificamos essas expressões definindo $\gamma = 1/\sqrt{1 - v^2}$, $v^i = dx^i/dx^0$ e $\vec{v} = (v^1, v^2, v^3)$, pois assim temos que

$$U = \gamma(1, \vec{v}) \text{ e } P = m\gamma(1, \vec{v}).$$

Denote por p^i a i -ésima componente do vetor P . Temos que, para $i = 1, 2, 3$,

$$p^i = m\gamma v^i = \frac{m}{\sqrt{1 - v^2}} v^i = mv^i + \frac{1}{2}mv^i v^2 + \dots,$$

e

$$p^0 = m\gamma = \frac{m}{\sqrt{1 - v^2}} = m + \frac{1}{2}mv^2 + \dots$$

Essas duas expressões têm termos que nos são bastante familiares. Na primeira, se v é suficientemente pequeno, p^i reduz-se à i -ésima componente do vetor momento newtoniano da partícula. Na segunda equação, reconhecemos o termo $1/2mv^2$ como a energia cinética newtoniana da partícula. Por essa razão, chamamos p^0 de *energia relativística total* de (α, m) .

A massa inercial na mecânica newtoniana é vista como uma medida de resistência a acelerações. Se pensarmos no termo $m/\sqrt{1 - v^2}$ como uma espécie de “massa relativística”, essa resistência se torna ilimitada, a medida que v se aproxima de 1. Isso quer dizer que, teoricamente, não há como passar da velocidade da luz. Para mais curiosidades físicas assim veja o excelente site [Int:Bae].

Note também que quando $v = 0$, a segunda equação se reduz a $p^0 = m$. Se estivéssemos usando unidades convencionais para a velocidade, encontraríamos que $p^0 = mc^2$. Como chamamos p^0 de energia relativística total, essa quantidade merece ser chamada também de E e temos a famosa fórmula $E = mc^2$. Ela nos diz que, na Relatividade, massa e energia são conceitos equivalentes. Para uma discussão mais profunda sobre isso, veja [Liv:Fey08].

3.6 O Tensor de Energia-Momento

Considere, informalmente, um “feixe” de partículas materiais, cada um com a mesma massa inercial m e histórias paralelas (ou seja, todas viajam na mesma direção com a mesma velocidade relativa a um referencial admissível). Cada partícula possui a mesma 4-velocidade, denotada por U , e portanto, o mesmo 4-momento P . Porém, essas informações não são suficientes para determinar unicamente todas as características do feixe, pois nem U nem P contém nenhuma informação sobre, por exemplo, a quantidade de partículas por unidade de volume ou a densidade de energia. Vejamos que um objeto matemático que codifica essas informações não pode ser um vetor de \mathcal{M} . Para isso, considere o referencial admissível no qual o feixe esteja em repouso (vimos que tal referencial sempre existe). Nesse referencial, a energia de cada partícula é m , e chamemos de n o número de partículas por unidade de volume nesse

mesmo referencial. Dessa forma, a energia por unidade de volume é dada por mn . Denotaremos essa quantidade por ρ . Considere agora um outro referencial que se move com velocidade v em relação ao referencial inicial. Lembremos que vimos no Capítulo 2 que a matriz que representa essa mudança de coordenadas é dada por

$$L = \begin{pmatrix} \gamma(v) & \gamma(v)v & 0 & 0 \\ \gamma(v)v & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

onde $\gamma(v) = 1/\sqrt{1-v^2}$. Lembremos também que isso implica que os comprimentos são contraídos por um fator de $\gamma(v)$ e, por consequência, volumes também são contraídos pelo mesmo fator. Dessa forma, uma unidade de volume no primeiro referencial torna-se uma fração de $\gamma(v)$ dessa quantidade.

Assim, no novo referencial, temos que o número de partículas por unidade de volume, denotado por \tilde{n} , é dado por $n/\sqrt{1-v^2}$. E como as partículas estão em movimento no novo referencial, temos que sua nova energia, denotada por \tilde{m} , é dada por $m/\sqrt{1-v^2}$, como vimos no início da seção. Dessa forma, a energia por unidade de volume no novo referencial é dada por $\tilde{n}\tilde{m} = mn/(1-v^2) = \rho/(1-v^2)$. Essa transformação envolve o quadrado de $\gamma(v)$, enquanto que toda transformação de Lorentz padrão depende linearmente de $\gamma(v)$. Assim, um objeto matemático que contenha a informação ρ não pode ser um vetor, pois vetores se transformam somente com uma transformação de Lorentz. O objeto matemático que codifica informações relativas a várias partículas é o *tensor de energia-momento*, a ser definido abaixo.

Em um referencial admissível (x^0, x^1, x^2, x^3) , defina o tensor T como

$$T(dx^i, dx^j) := \text{fluxo da } i\text{-ésima componente do 4-momento através da superfície de } x^j \text{ constante.}$$

Vejamos quem são suas componentes. Considere primeiro a componente T^{00} . Ela é, por definição, o fluxo de p^0 (ou seja, energia), através da superfície $t = \text{constante}$. Isso é somente a densidade de energia:

$$T^{00} = \text{densidade de energia.}$$

De maneira similar,

$$T^{0i} = \text{fluxo de energia através da superfície } x^i = \text{constante;}$$

$$T^{i0} = \text{densidade de } p^i;$$

$$T^{ij} = \text{fluxo de } p^i \text{ através da superfície } x^j = \text{constante.}$$

Como para qualquer tensor, basta conhecermos as componentes de T em um sistema de coordenadas que o conhecemos em todos. No exemplo que introduz a seção, se considerarmos o referencial no qual as partículas estão em repouso, temos que

$$T^{00} = \rho = mn$$

$$T^{0i} = T^{i0} = T^{ij} = 0.$$

Verifica-se facilmente que o tensor $mnU \otimes U$ tem exatamente essas componentes nesse referencial. Dessa forma, conseguimos deduzir qual é a forma do tensor de energia-momento de um feixe de partículas. Note que ele é simétrico. Isso não é uma característica exclusiva do nosso tensor. Na verdade, deve-se provar que isso é de fato um tensor, pois nossa definição de $T(dx^i, dx^j)$ não garante isso de imediato. Como não vamos usar nesse texto muito a fundo propriedades do tensor de energia-momento, não provaremos isso aqui. Veja [Liv:Sch09] para uma prova desses fatos usando argumentos físicos.

Note que o T definido como acima é um $(2, 0)$ -tensor. É conveniente modificá-lo para que ele se torne um $(0, 2)$ -tensor. Fazemos isso através de uma técnica muito usada em Geometria Riemanniana, que é usar

a métrica para transformar um tensor do tipo (r, s) em um do tipo $(r - 1, s + 1)$, ou vice-versa. No nosso caso particular, sejam T^{ij} as componentes de T em algum referencial admissível (e_i) . Dado um vetor V , o escrevemos em coordenadas como $V = V^i e_i$. Definimos a 1-forma associada a V , denotada por V^b , em coordenadas como

$$V_i^b = V_i = g_{ij} V^j.$$

Definimos então o $(0, 2)$ -tensor associado a T como

$$T^{bb}(V, W) = T(V^b, W^b) = T^{ij} V_i W_j = T^{ij} g_{ik} V^k g_{jl} W^l.$$

Dessa forma, as coordenadas desse novo tensor no referencial admissível em questão são dadas por

$$T_{ij}^{bb} = g_{ki} g_{lj} T^{kl}.$$

Para não carregar a notação, denotaremos os dois tensores por T , sempre ficando claro no contexto se o estaremos enxergando como um $(0, 2)$ -tensor ou como um $(2, 0)$ -tensor.

Note que no caso particular de um feixe de partículas, podemos escrever

$$T(V, W) = T(V^b, W^b) = mnU \otimes U(V^b, W^b) = mn g(U, V) g(U, W).$$

Em um sistema de coordenadas admissível (e_i) que se relaciona com o referencial no qual o feixe está em repouso por uma transformação de Lorentz padrão com parâmetro v , temos

$$T(e_0, e_0) = mn g(U, e_0) g(U, e_0) = \frac{mn}{1 - v^2},$$

que é exatamente a densidade de energia como vista pelo referencial admissível (e_i) . No nosso caso particular esse número é sempre positivo. Pela sua interpretação física, ele deveria ser positivo em todos os casos.

Em geral, definimos um *tensor de energia-momento* em \mathcal{M} como sendo um $(0, 2)$ -tensor simétrico T que satisfaz $T(V, V) \geq 0$, para todo V do tipo tempo. Essa condição é chamada de *condição fraca de energia*. Note que, pela discussão acima, para todo vetor unitário do tipo tempo V , a quantidade $T(V, V)$ representa a densidade de energia medida em um referencial admissível que tenha $V = e_0$. Note também que se T satisfaz a condição fraca de energia, por continuidade, a desigualdade vale também para todo V do tipo luz.

Dessa forma, o tensor de energia-momento é a maneira matemática adequada de se descrever a distribuição de matéria em \mathcal{M} . No Capítulo 4 vamos levar esse conceito para espaços-tempos mais gerais. A distribuição de matéria na Relatividade Geral é extremamente importante, pois ela se relaciona com a geometria do espaço-tempo de uma maneira não-trivial através das *equações de Einstein*.

Parte II

RELATIVIDADE GERAL

Capítulo 4

A Teoria da Relatividade Geral - Os fundamentos

4.1 Introdução

A Teoria da Relatividade Restrita nasceu por considerações acerca do Eletromagnetismo. Nada mais natural então do que se tentar introduzir conceitos eletromagnéticos na Teoria de Einstein. O procedimento feito é o seguinte¹: se define um campo eletromagnético como um certo tensor em M e se deduz as equações que caracterizam o movimento de uma partícula carregada sob ação desse campo. Seria natural proceder de maneira análoga em relação a efeitos gravitacionais. De fato essa abordagem foi tentada por muitos, inclusive pelo próprio Einstein, nos primeiros anos da Relatividade Restrita (para mais detalhes sobre essas tentativas, veja [Liv:MTW73]). Porém, as previsões da nova teoria não eram condizentes com os resultados experimentais. Citando Gregory Naber em [Liv:Nab88], *“Einstein logo deixou de lado a tarefa de formular teorias cada vez mais refinadas na esperança de acomodar os dados observacionais e prontamente buscou por uma razão física que explicasse a falha dessas idéias aparentemente naturais. Como sempre, a resposta estava lá para todos, mas só Einstein a viu”*. Para que isso fique claro, devemos retomar a discussão feita no final do primeiro capítulo sobre o Princípio da Equivalência. Lembremos o que ele nos diz:

- Princípio da Equivalência: “A trajetória de uma massa pontual em um campo gravitacional depende somente de sua posição e velocidade inicial e é independente da natureza da massa”.

A resposta para nosso impasse está toda contida nesse princípio. Note que, na abordagem adotada com campos eletromagnéticos, ele é tratado como uma entidade externa ao espaço-tempo, algo a mais, que não modifica sua estrutura, mas sim somente a trajetória de uma certa classe de partículas, a saber, as partículas carregadas. O Princípio da Equivalência nos diz que os campos gravitacionais não são dessa natureza, pois eles modificam a trajetória de *qualquer* partícula que tenha massa, e de maneira independente da natureza dessa massa. É quase como se a trajetória de partículas sob a ação de um campo gravitacional, dadas sua posição e velocidade inicial, fossem trajetórias “naturais” a serem seguidas no espaço-tempo. Além disso, lembremos também que concluímos o Capítulo 1 notando que o Princípio da Equivalência implica que gravidade está intimamente relacionada com aceleração, e, como este conceito é proveniente, em casos gerais, de uma conexão, isso nos indica que gravidade e conexão estão intimamente ligadas (páginas 15 e 16). Nessa hora já devemos imaginar que o espaço-tempo da Relatividade Geral é uma variedade diferenciável munida de uma conexão, que será responsável por descrever o campo gravitacional nele presente. Vejamos mais calmamente porquê esse é o contexto matemático adequado para nossa nova teoria.

¹Não faremos isso aqui, mas isso pode ser encontrado em qualquer texto de Relatividade, por exemplo, [Liv:Woo07].

Façamos como no começo do estudo da Teoria da Relatividade Restrita: comecemos com um conjunto M , a princípio sem estrutura adicional, cujos pontos chamamos de eventos, e queremos deduzir quais propriedades adicionais tal conjunto deve ter para que ele seja digno de receber o nome de *espaço-tempo*. Newton nos ensinou que a intensidade de campos gravitacionais decai com o quadrado da distância de sua fonte. Isso nos sugere que suficientemente longe de fontes de campos gravitacionais, M deve ser “parecida” como M . Porém, Einstein nos diz mais: ele propõe que *todo* evento, independente de sua proximidade com fontes de campos gravitacionais, possui uma vizinhança suficientemente pequena na qual o espaço-tempo se parece com M . Para entender isso, façamos o seguinte experimento mental: imagine que em um elevador encontram-se um observador e diversos objetos. Por infortúnio, o cabo do elevador se soltou e agora ele encontra-se em queda livre. Pelo Princípio da Relatividade, todos os corpos presentes dentro do elevador devem reagir da mesma forma, e portanto, devem manter-se em repouso um em relação ao outro durante a queda. Dessa forma, se o observador levantar um objeto do chão e o soltar, ele continuará nessa mesma posição. É muito pouco provável que um observador em tal situação esteja preocupado com princípios físicos ou que alguém se coloque em tal situação propositalmente em prol da ciência. Porém todos já vimos imagens na televisão de astronautas brincando com objetos dentro de suas cápsulas espaciais. Essa situação é análoga àquela proposta acima pois estando os motores dos foguetes desligados, os astronautas estão somente sob o efeito da gravidade. Então nesse pequeno intervalo de tempo e espaço no qual o elevador está em queda livre ou os astronautas estão orbitando nosso planeta podemos dizer que o espaço-tempo se comporta aproximadamente como M . Os objetos nessas condições se movem de acordo com a Lei da Inércia. Um observador nessas condições pode introduzir coordenadas, pelo menos localmente, no espaço-tempo através de algum dos métodos discutidos no início do Capítulo 2 e tornar-se assim um observador admissível.

Um observador nessas condições observa, ao menos localmente, as ações se passando como se estivesse em M . Em tal referencial vale a Lei da Inércia e o observador vê objetos livres se movendo em linha reta. Então tudo nos indica que esse é o referencial ideal que devemos adotar no nosso novo espaço-tempo. Sendo esse referencial tão especial, devemos dar a ele um nome. Chamaremos-o de *referencial em queda livre*. Note que esse nome é bem adequado tendo em vista as duas situações que usamos para introduzi-lo. No caso do elevador o observador está claramente em queda livre. No caso do astronauta também, já que ele está somente sobre a ação da gravidade da Terra. A diferença é que a trajetória em queda livre do astronauta é mais segura do que a trajetória do observador no elevador.

Repare que tudo isso vale localmente. O que quer dizer o termo “localmente”? Bem, na Física podemos pensar da seguinte forma: lembremos-nos do experimento do elevador. Os objetos estão em queda livre, porém, o campo gravitacional da Terra os atrai para seu centro, implicando em um pequeno movimento relativo entre os corpos. Se tivermos equipamentos suficientemente precisos, podemos medir esses desvios, fazendo com que tal observador não esteja mais em queda livre. Dessa forma, a palavra “local” depende de quão preciso queremos que seja nosso experimento: dada a precisão desejada, pode-se encontrar distâncias e tempos suficientemente pequenos de modo que o erro do experimento esteja dentro da precisão.

É natural pensarmos que quanto menor for a vizinhança do evento, mais ela se parece com M . Dessa forma, tomando uma vizinhança infinitamente pequena, teríamos que o espaço-tempo é idêntico a M . Do ponto de vista matemático esse argumento está totalmente não-rigoroso, mas podemos torna-lo preciso com facilidade. Faremos isso a partir de então.

4.2 Os Fundamentos

A Teoria da Relatividade Geral lança mão da Geometria Riemanniana, introduzida por Bernhard Riemann² em 1854 [Art:Rie1854]³. Porém, no começo do século XX a Geometria Diferencial era muito diferente do que conhecemos hoje em dia, devido a falta de noções fundamentais para esta, como espaço topológico e fibrados vetoriais⁴. Por isso que a Teoria da Relatividade Geral foi considerada extremamente complexa, do ponto de vista matemático, quando ela surgiu. Demorou um certo tempo até que ela fosse bem aceita pelos físicos. Felizmente, hoje em dia a Geometria Diferencial é uma teoria bem conhecida e com bases sólidas, possibilitando-nos então formular a Relatividade Geral de maneira rigorosa, e daí extrair Teoremas que têm importantes implicações físicas.

A partir de agora usaremos a linguagem de Geometria Diferencial e Riemanniana. Muitas afirmações não serão provadas, por já serem bem conhecidas na literatura. Apesar disso, eu as enunciarei e indicarei uma referência para uma demonstração.

Bem, a discussão na seção anterior nos sugere postular que o espaço-tempo é uma variedade diferenciável M de dimensão 4. Nossa observação de que infinitamente perto de um dado ponto M é idêntica a M se traduz em termos modernos dizendo-se que o espaço tangente em todo ponto de M é munido de um produto interno de índice 3. Dessa forma, M possui uma estrutura adicional, conhecida como *métrica lorentziana*.

As observações de Einstein sobre referenciais em queda livre nos levam então ao seguinte princípio:

- Princípio da Equivalência (segunda versão): “Todo experimento que não envolva gravidade realizado em um referencial em queda livre apresenta o mesmo resultado se realizado em M .”

Em outras palavras, esse Princípio nos diz que as leis da Física que não envolvem efeitos gravitacionais são as mesmas em M e em qualquer referencial em queda livre.

A partir de agora vamos fazer uma série de definições análogas às feitas no caso da Relatividade Restrita.

Definição 4.1. Uma *métrica lorentziana* em uma variedade diferenciável M de dimensão n é um tensor métrico suave, simétrico e não degenerado, denotado por g , tal que em cada espaço tangente g tem índice $n - 1$.

Repare que o nome *métrica* é um abuso de nomenclatura pois, em contraste com o caso Riemanniano, não podemos definir uma métrica a partir desse tensor.

Definição 4.2. Um *espaço-tempo* é uma variedade M de dimensão 4 munida de uma métrica lorentziana.

Denotaremos um espaço-tempo pelo par (M, g) , quando houver necessidade de se deixar claro quem é a métrica em questão. Uma variedade de dimensão arbitrária munida de uma métrica lorentziana é dita uma *variedade lorentziana*. Todas as definições aqui valem para variedades lorentzianas em geral, porém nossa intuição sempre deve estar voltada para aquelas de dimensão 4. Quando escrevermos somente M está implícito que M é uma variedade lorentziana ou um espaço-tempo.

²Georg Friedrich Bernhard Riemann, Breselenz 17/09/1826 - Selasca 20/07/1866.

³Em [Liv:Spi99], volume II, encontra-se uma tradução para o inglês de tal obra. Vale a pena ler o trabalho de Riemann. É impressionante ver como ele já conhecia naquela época todos os fundamentos da Geometria Riemanniana, só não tinha teoria suficiente para torná-lo rigoroso. O legal dessa tradução do Spivak é que logo em seguida ele escreve em termos modernos o que Riemann nos ensinou. Vale ressaltar que no mesmo livro encontra-se também uma tradução do trabalho de Gauß (Johann Carl Friedrich Gauß - Braunschweig 30/04/1777 - Göttingen 23/02/1855) sobre geometria das superfícies em \mathbb{R}^3 . Na verdade, vale a pena ler qualquer livro do Spivak. Ele escreve de um jeito muito diferente, muitas vezes melhor do que vemos nos outros textos, além de trazer exercícios não-mecânicos e muito bem elaborados.

⁴Na verdade, a noção de espaço topológico surge na mesma época, com Riesz e Hausdorff. Porém não se tinha a definição precisa do que é uma variedade diferenciável. Em 1913, Hermann Weyl deu uma primeira definição de superfícies de Riemann [Liv:Wei09]. A definição de variedade usando atlas foi dada por Hassler Whitney em 1936 [Art:Whi36].

Definição 4.3. Seja (M, g) uma variedade lorentziana. Dizemos que $v \in T_p M$ é do tipo *tempo, espaço* ou *luz* se a quantidade $g(v, v)$ é positiva, negativa ou zero, respectivamente.

Definição 4.4. Dada uma curva diferenciável $\alpha : I \rightarrow M$, dizemos que ela é do tipo *tempo, espaço* ou *luz* se $\alpha'(t)$ o é, para todo $t \in I$.

Notação. Denotaremos um espaço-tempo arbitrário por M , enquanto que reservamos o símbolo \mathcal{M} para o espaço de Minkowski.

Analogamente ao caso riemanniano, toda métrica lorentziana admite uma única conexão compatível, no seguinte sentido:

Teorema 4.5. *Seja (M, g) uma variedade lorentziana. Então existe uma única conexão linear ∇ livre de torção e compatível com a métrica, ou seja, o tensor de torção dado por $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$ é nulo e para todos campos de vetores X, Y, Z em M vale que*

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Uma demonstração pode ser encontrada em qualquer texto de Geometria Riemanniana, e sua demonstração funciona perfeitamente no caso lorentziano ou semi-riemanniano em geral (onde g é um tensor métrico suave, simétrico, não degenerado e seu índice é arbitrário). Um excelente texto é [Liv:Lee97]. Lá também podem ser encontradas diversas equivalências para se descrever a compatibilidade entre a conexão e a métrica.

Estando nossa variedade munida de uma conexão, temos muitas ferramentas em mãos para trabalhar, como a aplicação exponencial, geodésicas, tensor de curvatura, entre outras. As introduziremos na hora certa. Por enquanto usaremos somente a aplicação exponencial. Lembremos que mais acima falamos que “localmente” podemos dizer que M é “parecida” \mathcal{M} . Agora podemos tornar isso rigoroso, usando *coordenadas normais*. O Teorema abaixo vale para variedades munidas com uma conexão linear em geral, mas o enunciaremos somente no nosso caso particular.

Teorema 4.6. *Seja (M, g) uma variedade lorentziana e $p \in M$. Então existe uma vizinhança estrelada U de $0 \in T_p M$ e vizinhança V de p em M tal que:*

1. $\exp_p : U \rightarrow V$ é um difeomorfismo (dizemos que V é uma vizinhança normal de p em M);
2. V é geodesicamente convexa (ou convexa, por simplicidade), ou seja, é vizinhança normal de todos os seus pontos.

Uma prova pode ser encontrada em [Liv:Onl83]. A razão para uma vizinhança que satisfaz a segunda propriedade ser chamada de geodesicamente convexa é o seguinte fato, cuja demonstração se encontra no mesmo texto:

Proposição 4.7. *Seja $U \subset M$ um aberto convexo. Então para quaisquer par de pontos $p, q \in U$ existe uma única geodésica $\alpha : [0, 1] \rightarrow M$ contida em U tal que $\alpha(0) = p$ e $\alpha(1) = q$.*

Observação. Note que dois pontos quaisquer em uma variedade podem não ser conectados por nenhuma geodésica ou por infinitas. Um exemplo para o primeiro caso é o plano sem a origem e dois pontos da forma x e $-x$. Já o segundo pode ser ilustrado com pontos antípodas sobre a esfera.

Sabendo da existência de vizinhanças normais em torno de todo ponto de M , podemos olhar a aplicação exponencial como uma carta local. As coordenadas introduzidas dessa forma são chamadas de *coordenadas normais*.

4.3 Três exemplos úteis

Quando estudamos Geometria Riemanniana nossas intuições baseiam-se essencialmente em três exemplos: o espaço euclidiano, o plano hiperbólico e a esfera. Tais espaços são especiais principalmente por terem curvatura escalar constante nula, negativa e positiva, respectivamente. Agora apresentaremos três modelos de espaço-tempo que serão bastante úteis para ilustrar os conceitos introduzidos.

4.3.1 O Espaço-tempo de Minkowski

Esse espaço-tempo já é bem conhecido nosso e o estudamos a fundo no Capítulo 3. Lembemos que o denotamos por \mathcal{M} e ele é um espaço vetorial real de dimensão 4 munido de um produto interno de índice 3.

4.3.2 O espaço-tempo de deSitter

É nosso segundo exemplo e o denotamos pela letra \mathcal{D} . Considere a métrica $\tilde{g}(x, y) = -x^0 y^0 + x^1 y^1 + x^2 y^2 + x^3 y^3 + x^4 y^4$ em \mathbb{R}^5 . Sua forma quadrática associada é dada por

$$Q(x) = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2.$$

Claramente Q é uma função suave e $1 \in \mathbb{R}$ é valor regular de Q . Definimos então $\mathcal{D} = Q^{-1}(1)$. Dessa forma \mathcal{D} é um hiperbolóide de dimensão 4 em \mathbb{R}^5 .

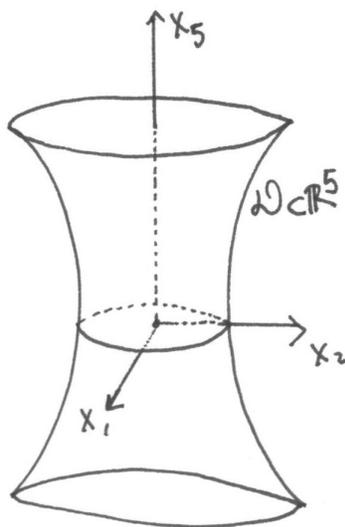


Figura 4.1: Espaço-tempo deSitter.

Bem, um hiperbolóide como acima em \mathbb{R}^3 é difeomorfo a $\mathbb{S}^1 \times \mathbb{R}$. Vale o análogo nesse caso.

Proposição 4.8. \mathcal{D} é difeomorfo a $\mathbb{S}^3 \times \mathbb{R}$.

Demonstração. Vamos enxergar o conjunto $\mathbb{S}^3 \times \mathbb{R}$ como

$$\{(y^1, y^2, y^3, y^4, t) \mid \sum (y^i)^2 = 1 \text{ e } -\infty < t < \infty\}.$$

Defina $F : \mathbb{S}^3 \times \mathbb{R} \rightarrow \mathbb{R}^5$ como

$$F(y^1, y^2, y^3, y^4, t) = (t, (1+t^2)^{1/2}y^1, (1+t^2)^{1/2}y^2, (1+t^2)^{1/2}y^3, (1+t^2)^{1/2}y^4).$$

F é claramente suave e verificamos que

$$Q(F(y^1, y^2, y^3, y^4, t)) = -t^2 + (1+t^2) \left(\sum (y^i)^2 \right) = 1.$$

Assim, $F(\mathbb{S}^3 \times \mathbb{R}) \subseteq \mathcal{D}$.

F é também bijetiva, pois a função $G : \mathcal{D} \rightarrow \mathbb{S}^3 \times \mathbb{R}$ dada por

$$G(x^0, x^1, x^2, x^3, x^4) = (x^0, (1+(x^0)^2)^{1/2}x^1, (1+(x^0)^2)^{1/2}x^2, (1+(x^0)^2)^{1/2}x^3, (1+(x^0)^2)^{1/2}x^4)$$

é uma inversa para F . Como G também é suave, concluímos que F é um difeomorfismo. \square

Seja $D = \{(u^0, u^1, u^2, u^3) \in \mathbb{R}^4 \mid -\infty < u^0 < \infty, -\pi < u^1, u^3 < \pi \text{ e } 0 < u^2 < \pi\}$. Defina $\varphi : D \rightarrow \mathbb{R}^5$ como $\varphi(u^0, u^1, u^2, u^3) = (x^0, x^1, x^2, x^3, x^4)$, onde

$$\begin{aligned} x^0 &= \sinh(u^0) \\ x^1 &= \cosh(u^0) \cos(u^1) \\ x^2 &= \cosh(u^0) \sin(u^1) \cos(u^2) \\ x^3 &= \cosh(u^0) \sin(u^1) \sin(u^2) \cos(u^3) \\ x^4 &= \cosh(u^0) \sin(u^1) \sin(u^2) \sin(u^3). \end{aligned}$$

Uma verificação tediosa nos garante que φ é uma parametrização para \mathcal{D} . Vamos definir uma métrica lorentziana g em \mathcal{D} restringindo a métrica \tilde{g} de \mathbb{R}^5 a cada espaço tangente e \mathcal{D} e tomando o sinal negativo. Outra verificação mais tediosa ainda nos diz que, relativo à parametrização φ , as componentes da métrica são dadas por:

$$\begin{aligned} g_{00} &= 1 \\ g_{11} &= -\cosh^2(u^3) \\ g_{22} &= -\cosh^2(u^3) \sin^2(u^0) \\ g_{33} &= -\cosh^2(u^3) \sin^2(u^0) \sin^2(u^1) \\ g_{ij} &= 0, \text{ se } i \neq j. \end{aligned}$$

Isso mostra que (\mathcal{D}, g) é de fato um espaço-tempo. Claro que a parametrização φ não cobre todo \mathcal{D} , porém, se a considerarmos no domínio $\{(u^0, u^1, u^2, u^3) \in \mathbb{R}^4 \mid -\infty < u^0 < \infty, -\pi \leq u^1, u^3 \leq \pi \text{ e } 0 \leq u^2 \leq \pi\}$ toda a superfície será coberta, com o preço da aplicação deixar de ser uma parametrização. Mas isso não é um problema grave, já que se considerarmos a mesma função porém com u^1, u^2 e u^3 restitos a intervalos suficientemente pequenos em torno de $-\pi, 0$ e π , respectivamente, teremos um atlas para \mathcal{D} . Repare que essa escolha de parametrização e seus conjuntos de definição não são misteriosos, pois as coordenadas u^1, u^2 e u^3 são responsáveis por parametrizar cada “fatia” com x^0 fixo de \mathcal{D} (ou seja, uma esfera). Dessa forma, para que cubramos \mathcal{D} por inteiro, devemos ter que todas essas “fatias” sejam cobertas por parametrizações, ou seja, precisamos ter essencialmente um atlas para a esfera.

4.3.3 O espaço-tempo de Einstein-deSitter

Esse modelo é bem simples, pois é um sub-conjunto aberto de \mathbb{R}^4 com uma métrica diferente da métrica de Minkowski usual. Considere $\mathcal{E} = (0, \infty) \times \mathbb{R}^3$ e defina $g(p) = g(u^0, u^1, u^2, u^3)$ como

$$g_{00} = 1$$

$$g_{ii} = -(u^0)^{4/3}, \text{ se } i = j = 1, 2, 3$$

$$g_{ij} = 0, \text{ se } i \neq j.$$

Dessa forma, dados $v, w \in T_p\mathcal{E}$, onde $p = (u^0, u^1, u^2, u^3)$, temos que $g(v, w) = v^0w^0 - (u^0)^{4/3}(v^1w^1 + v^2w^2 + v^3w^3)$. Observe, também, que o cone de luz em $T_p\mathcal{E}$ é dado por $\{v = (v^0, v^1, v^2, v^3) \in T_p\mathcal{E} \mid (u^0)^{4/3}((v^1)^2 + (v^2)^2 + (v^3)^2) = (v^0)^2\}$, ou seja, os cones de luz ficam mais “finos” a medida que p fica “mais alto”.

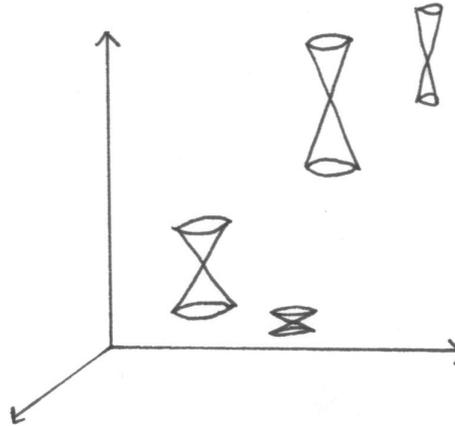


Figura 4.2: Espaço-tempo de Einstein-deSitter.

O espaço-tempo de deSitter⁵ representa um universo não realista, com densidade de energia negativa. Outra possível interpretação é como um universo sem matéria satisfazendo as Equações de Einstein com constante cosmológica não-nula. Já espaço-tempo de Einstein-deSitter é, depois do espaço de Minkowski, o mais simples dos modelos cosmológicos em Relatividade que tem algum significado físico. Ele representa, em uma primeira aproximação, um universo que é ocupado por uma “nuvem” de galáxias homogênea e isotrópica.

4.4 Causalidade

4.4.1 Orientação no tempo

Analogamente o que fizemos para o espaço de Minkowski, vamos estudar relações causais em M , ou seja, quando que um certo evento acontece “antes” ou “depois” de algum outro evento. Motivado pelo Teorema 3.5, fazemos a seguinte definição:

Definição 4.9. Seja (M, g) uma variedade lorentziana e $v, w \in T_pM$, ambos do tipo tempo. Dizemos que eles têm a mesma orientação se $g(v, w) > 0$.

Queremos dizer, da mesma forma que fizemos no caso da Relatividade Restrita, quando que um vetor aponta para o futuro ou para o passado. Mais ainda: como consideraremos partículas e seus respectivos vetores tangentes, queremos dizer quando que um campo de vetores aponta para o futuro ou para o passado. Note que não podemos adotar o procedimento feito no caso restrito, pois ele só valeria em um

⁵Willem deSitter - Sneek 06/05/1872 - Leiden 20/11/1934.

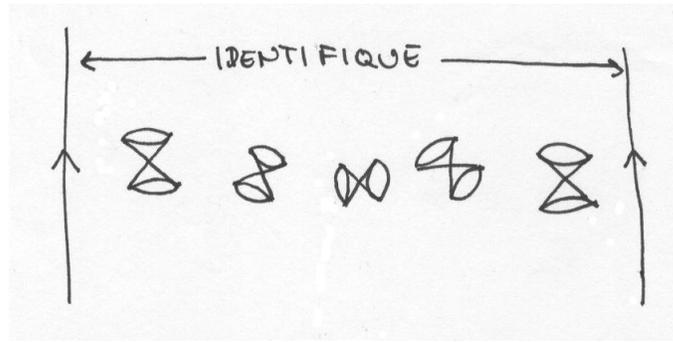


Figura 4.3: Variedade lorentziana não orientável no tempo.

espaço tangente em particular. E mesmo se o repetíssemos para todos os espaços tangentes poderíamos ter escolhas inconsistentes em espaços tangentes próximos. Mas nem tudo está perdido! Note que a definição acima nos permite definir uma relação de equivalência em T_pM da seguinte forma:

$$v \sim w \Leftrightarrow g(v, w) > 0.$$

O quociente de T_pM por essa relação tem 2 elementos, que são os dois cones de tempo. Uma escolha de orientação seria então uma escolha de classe de equivalência que varie suavemente com o ponto p . Tornamos isso rigoroso da seguinte forma:

Definição 4.10. Uma variedade lorentziana (M, g) é dita *orientável no tempo* se existe um campo de vetores V tal que $V(p)$ é tipo tempo, para todo $p \in M$. Nesse caso, dizemos que um vetor $v \in T_pM$ aponta para o futuro se $g(v, V(p)) > 0$ e que uma curva diferenciável $\alpha : I \rightarrow M$ aponta para o futuro se $\alpha'(t)$ o faz, para todo $t \in I$. Dizemos também que v ou α estão *direcionados para o futuro*.

Note que nem toda variedade lorentziana é orientável no tempo. Veja a figura 4.3. Apesar disso, vale o seguinte resultado:

Proposição 4.11. Seja (M, g) uma variedade lorentziana não-orientável no tempo. Então existe uma variedade lorentziana (\tilde{M}, \tilde{g}) e uma isometria local $\pi : \tilde{M} \rightarrow M$ tal que \tilde{M} é orientável no tempo e cada $p \in M$ tem duas pré-imagens por π . O par (\tilde{M}, \tilde{g}) é dito o recobrimento duplo orientado no tempo de M .

Como não usaremos esse resultado aqui, não o provaremos. Sua prova pode ser encontrada em [Liv:Onl83].

Espaços-tempos não orientáveis no tempo são interpretados como aqueles onde não se consegue distinguir claramente as noções de “ir para o futuro” e “ir para o passado”, e queremos evitar essa patologia, do ponto de vista físico.

Daqui em diante sempre suporemos que M é orientável no tempo, a não ser que seja dito o contrário.

Mais adiante veremos que nossos três exemplos são todos orientáveis no tempo.

Sabemos que toda variedade diferenciável admite uma métrica riemanniana. Esse resultado é bem conhecido e uma demonstração pode ser facilmente encontrada em textos de Geometria Diferencial. Porém, nem toda variedade admite uma métrica lorentziana.

Teorema 4.12. São equivalentes:

1. M admite uma métrica lorentziana;
2. M admite uma métrica lorentziana orientada no tempo;

3. Existe um campo de vetores V que nunca se anula;
4. M é não compacta ou M é compacta e sua característica de Euler $\chi(M)$ é nula.

Demonstração. Façamos as seguintes equivalências:

- $3 \Leftrightarrow 4$: É um fato clássico da Topologia Geral. Sua prova pode ser encontrada em [Liv:Hat01].
- $2 \Rightarrow 1$: É trivial.
- $3 \Rightarrow 2$: Seja g uma métrica riemanniana em M e V um campo de vetores que nunca se anula. Podemos supor que $g(V, V) = 1$. Defina

$$h(X, Y) = -g(X, Y) + 2g(X, V)g(Y, V).$$

Provemos que h é uma métrica lorentziana orientável no tempo. Para isso, seja $\{V, E_2, \dots, E_n\}$ um referencial local ortonormal para g (ele sempre existe, pelo algoritmo de Gram-Schmidt). Temos então que:

$$\begin{aligned} h(E_i, E_j) &= -g(E_i, E_j) + 2g(E_i, V)g(E_j, V) = -\delta_{ij}; \\ h(V, E_i) &= -g(V, E_i) + 2g(V, V)g(E_i, V) = -g(V, E_i) = 0; \\ h(V, V) &= -g(V, V) + 2g(V, V)g(V, V) = 2 - 1 = 1. \end{aligned}$$

Portanto, h é uma métrica lorentziana. A última relação mostra também que V é um campo de vetores do tipo tempo, implicando em (M, h) ser orientável no tempo.

- $2 \Rightarrow 3$: Segue diretamente da definição de orientação no tempo.
- $1 \Rightarrow 4$: Suponha primeiro que M é orientável no tempo. Então vale o item 3 e, portanto, 4. Suponha agora que M não é orientável no tempo e seja (\tilde{M}, \tilde{g}) seu recobrimento duplo orientado no tempo. Então \tilde{M} é não-compacta ou é compacta e $\chi(\tilde{M}) = 0$. Nesse último caso, temos que $M = \pi(\tilde{M})$ é compacta e $\chi(M) = \chi(\tilde{M})/2 = 0$. Se \tilde{M} é não-compacta, devemos provar que M também não o é. Note que isso é equivalente ao fato de que se M é compacta então \tilde{M} também o é. Provemos esse último fato. De fato, seja (U_i) uma cobertura aberta de \tilde{M} e seja (V_j) um refinamento dessa cobertura tal que cada V_j é homeomorfo a sua imagem por π . Como $(\pi(V_j))$ forma uma cobertura aberta de M que estamos supondo compacta, podemos dizer que $M \subset \pi(V_{j_1}) \cup \dots \cup \pi(V_{j_n})$. Assim, temos que $\tilde{M} \subset \pi^{-1}(\pi(V_{j_1})) \cup \dots \cup \pi^{-1}(\pi(V_{j_n}))$, e cada uma dessas parcelas é composta por 2 abertos disjuntos, pois o recobrimento é duplo. Usamos essa cobertura para produzir uma sub-cobertura de (U_i) .

□

4.4.2 Condição cronológica

Lembremos que em M definimos uma relação de causalidade, denotada por \ll , da seguinte forma:

$$x_0 \ll x \text{ se e somente se } x - x_0 \text{ é tipo tempo e direcionado para o futuro.}$$

Provamos o Teorema 3.15, que nos dá uma equivalência para essa relação em termos de curvas do tipo tempo. Isso nos motiva a dizer que, dados dois pontos $p, q \in M$ então $p \ll q$ se existe uma curva

diferenciável $\alpha : [0, 1] \rightarrow M$ tipo tempo e direcionada para o futuro tal que $\alpha(0) = p$ e $\alpha(1) = q$. Ou seja, $p \ll q$ se e somente se é possível que um observador, não necessariamente inercial, esteja presente em ambos os eventos. Vamos, porém, fazer uma definição mais fácil de se trabalhar, mas que será equivalente a essa aqui proposta. Para isso, precisamos do conceito de uma *viagem* em M .

Definição 4.13. Uma *viagem* de p a q em M é uma curva contínua $\alpha : [0, 1] \rightarrow M$ direcionada para o futuro, tipo tempo e geodésica por partes satisfazendo $\alpha(0) = p$ e $\alpha(1) = q$.

Para ficar claro, o fato de α ser uma geodésica por partes quer dizer que existe uma partição $\{0 = t_0, t_1, \dots, t_n = 1\}$ do intervalo $[0, 1]$ tal que a restrição de α a qualquer intervalo $[t_k, t_{k+1}]$ acima é uma geodésica.

Pensamos em uma curva do tipo tempo direcionada para o futuro como o movimento de uma partícula não necessariamente livre (a não ser que a curva em questão seja uma geodésica). Podemos interpretar uma viagem como sendo o movimento de uma partícula que se desloca livremente durante um certo instante de tempo, muda de direção, volta a deslocar-se livremente durante mais um instante de tempo, e assim sucessivamente.

Agora podemos dar uma definição para \ll .

Definição 4.14. Dados $p, q \in M$, dizemos que p *antecede* q cronologicamente se existe uma viagem de p a q em M . Nesse caso, escrevemos que $p \ll q$.

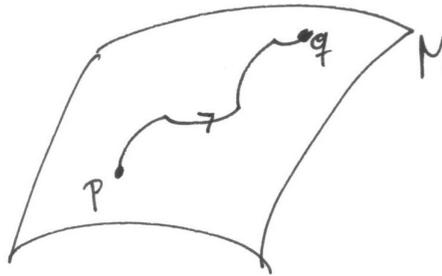


Figura 4.4: p antecede q cronologicamente.

Com essa definição é imediato verificar que \ll é uma relação transitiva. O conceito de viagem pode parecer um pouco diferente da primeira noção introduzida, mas na verdade eles são equivalentes:

Proposição 4.15. Sejam $p, q \in M$. Então $p \ll q$ se e somente se existe uma curva diferenciável $\alpha : [0, 1] \rightarrow M$ tipo tempo e direcionada para o futuro tal que $\alpha(0) = p$ e $\alpha(1) = q$.

Ainda não podemos provar esse resultado, pois ele usa alguns conceitos que serão introduzidos somente na Seção 5.8.

Vamos definir agora análogos dos cones de tempo em M .

Definição 4.16. Dado $p \in M$, definimos:

- $I^+(p) = \{q \in M \mid p \ll q\}$, dito o *futuro cronológico* de p ;
- $I^-(p) = \{q \in M \mid q \ll p\}$, dito o *passado cronológico* de p ;
- $S \subset M$ sub-conjunto arbitrário, então $I^\pm(S) = \cup_{p \in S} I^\pm(p)$ denota o *futuro/passado cronológico* de S , respectivamente.

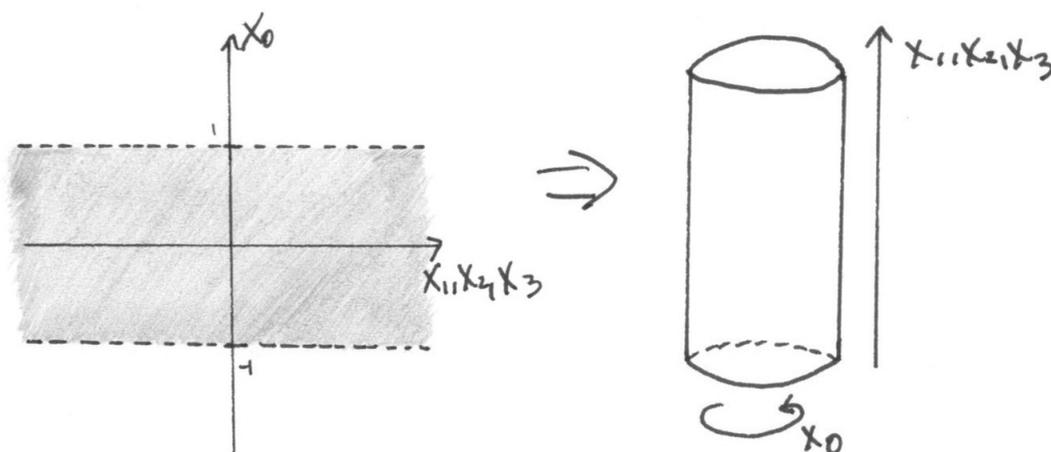


Figura 4.5: Espaço tempo no qual $p \in I^+(p)$.

Note que em M os conjuntos $I^\pm(p)$ e $C_T^\pm(p)$ coincidem, mas no caso geral tal conjunto pode ser bem complicado. Em particular, pode existir um espaço-tempo M tal que p pertença a $I^+(p)$! Isso quer dizer que você poderia tomar uma certa decisão, “voltar no tempo” e deixar de tomar essa decisão. A Figura 4.5 ilustra um desses exemplos.

Considere o subconjunto de M dado pelos pontos (x^i) tal que $1 \leq x^0 \leq 1$. Seja M o espaço tempo obtido identificando-se os pontos da forma $(-1, x^1, x^2, x^3)$ com os pontos da forma $(1, x^1, x^2, x^3)$. É fácil se convencer que nesse caso $I^+(p) = M$, para todo ponto $p \in M$.

Isso claramente é uma propriedade no mínimo estranha e devemos evitá-la, para que não aconteça situações como a ilustrada na Figura 4.6.

Definição 4.17. Dizemos que M satisfaz a *condição cronológica* se $p \notin I^+(p)$, para todo $p \in M$.

Provaremos que toda M compacta *não* satisfaz a condição cronológica. Para isso, precisamos do seguinte fato.

Proposição 4.18. $I^\pm(p)$ são abertos em M , para todo $p \in M$.

Demonstração. Provemos somente para $I^+(p)$. Para o outro caso basta reverter a orientação no tempo de M . A idéia da prova é a seguinte: Tome $q \in I^+(p)$ e seja α uma viagem de p a q . Devemos provar que existe uma vizinhança de q em M inteiramente contida em $I^+(p)$. Usando coordenadas normais, em torno de cada ponto, M é essencialmente M . Escolhendo então um ponto r suficientemente perto de q , a imagem de q através das coordenadas normais centradas em r estará no cone de tempo futuro em M , que sabemos que é aberto. Através desse artifício obteremos a vizinhança desejada, já que a aplicação exponencial é um difeomorfismo, em vizinhanças normais.

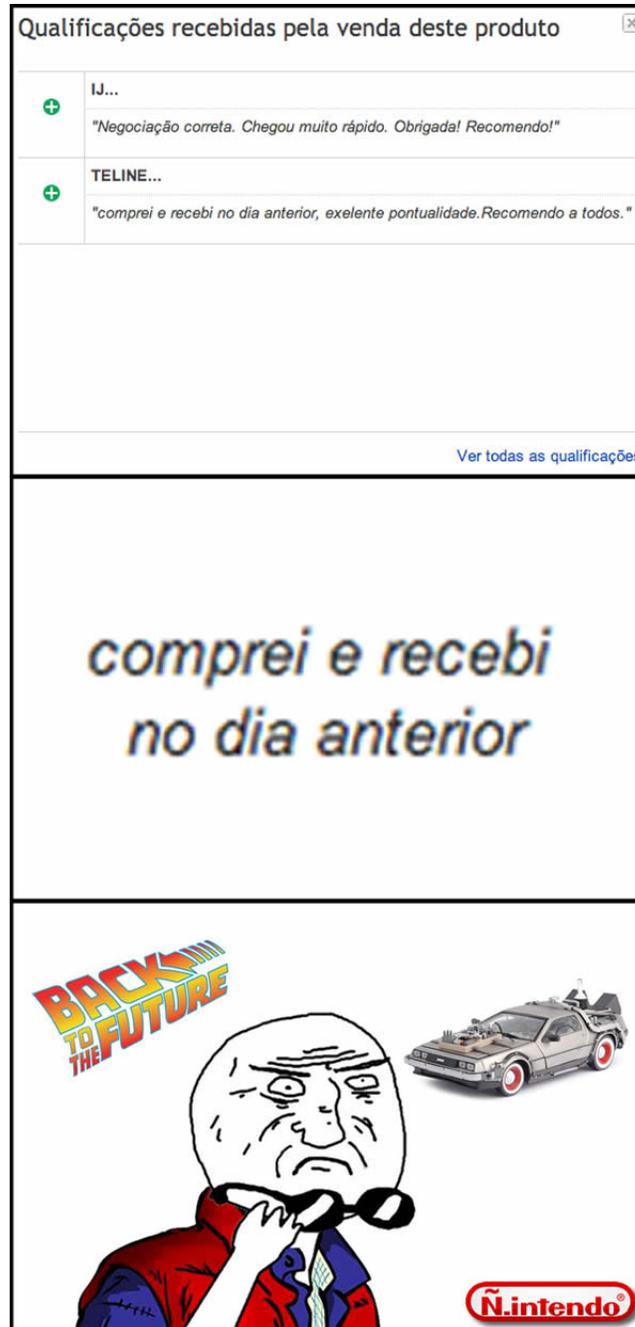


Figura 4.6: Falha na causalidade.

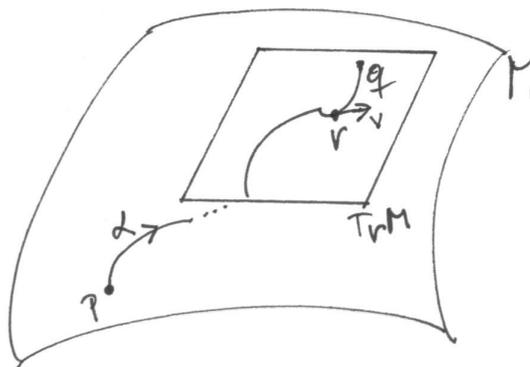


Figura 4.7: Prova da Proposição 4.18.

Passemos então à demonstração rigorosa. Tome $q \in I^+(p)$ e seja U uma vizinhança convexa de q em M . Tome $U \ni r \neq q$ na imagem do último segmento da viagem α ligando p a q . Adicionando outro segmento, se necessário, suponha que o último segmento de α começa em r e termina em q . Como U é uma vizinhança convexa, nas coordenadas de \exp_r o último segmento de α é um segmento de reta em $T_r M$. Como existe U_r vizinhança estrelada de $0 \in T_r M$ na qual $\exp_r : U_r \rightarrow U$ é um difeomorfismo, então o último segmento de α está inteiramente contido em U . Seja $v = \exp_r^{-1}(q) \in U_r$ e seja γ_v a geodésica com vetor tangente inicial v . Sabemos que $\gamma_v(1) = q$. Dessa forma, o último segmento de α e γ_v são geodésicas de r a q inteiramente contidas em uma vizinhança convexa, e portanto, são iguais, pelo Teorema 4.7. Portanto, o último segmento de α é uma reparametrização de γ_v com parâmetro da forma $t = ms + b$, para $m > 0$. Assim, os seus vetores tangentes são paralelos e têm a mesma orientação no tempo. O fato do último segmento de α ser do tipo tempo e direcionado para o futuro implica que v também o é, e portanto está contido no interior do cone de tempo futuro em $T_r M$. Denote por V a interseção do cone de tempo futuro em $T_r M$ com U_r . Como \exp_r é um difeomorfismo, temos que $\exp_r(V)$ é aberto em U (e portanto em M) e contém q . Mostremos que essa é a vizinhança desejada, ou seja, que $\exp_r(V) \subset I^+(p)$. De fato, tomando $x \in \exp_r(V)$, existe um único $y \in V$ tal que $\exp_r(y) = x$. Dessa forma, x é o ponto final de um segmento de geodésica do tipo tempo que começou em r , ou seja, $r \ll x$. Como sabemos que vale $p \ll r$ e a relação \ll é transitiva, concluímos que $p \ll x$, como desejado. \square

Corolário 4.19. *Toda M compacta falha na condição cronológica.*

Demonstração. Começemos observando que a $\{I^+(p), p \in M\}$ é uma cobertura de M . Para isso, devemos notar que todo ponto p de M está no futuro de algum outro ponto. Para isso tome $v \in T_p M$ direcionado para o futuro e tipo tempo e considere $\alpha : (-\varepsilon, \varepsilon) \rightarrow M$ geodésica com vetor tangente inicial v . Dessa forma, $p \in I^+(\alpha(t_0))$, para qualquer $t_0 \in (-\varepsilon, 0)$.

Como M é compacta temos que $M \subset I^+(p_1) \cup \dots \cup I^+(p_n)$, para alguns $p_1, \dots, p_n \in M$. Podemos supor que $I^+(p_1) \not\subset I^+(p_j)$, para todo $j \geq 2$ (caso contrário podemos excluir $I^+(p_1)$ da cobertura). Afirmamos que $p_1 \in I^+(p_1)$. De fato, se fosse falso, teríamos que $p_1 \in I^+(p_j)$, para algum $j \geq 2$. Isso implica que $I^+(p_1) \subset I^+(p_j)$, pois dado $q \in I^+(p_1)$ temos que $p_1 \ll q$. Como sabemos que $p_j \ll p_1$ e a relação \ll é transitiva, concluímos que $p_j \ll q$, e portanto, $q \in I^+(p_j)$, o que leva a uma contradição. \square

Daqui em diante sempre suporemos que M é orientável no tempo e satisfaz à condição cronológica, a não ser que seja dito o contrário.

Mais adiante veremos que nossos três exemplos também satisfazem à condição cronológica.

Isso nos impõe uma forte restrição topológica nos espaços-tempos razoáveis de considerarmos. Porém, não basta considerarmos somente espaços-tempos não compactos, pois existem espaços-tempo não

compactos que falham em satisfazer a condição cronológica, como o *universo de Gödel*, por exemplo. Veja [Art:Göd49] para mais detalhes. Devemos então considerar alguma outra restrição nos nossos espaços-tempos para que eles sejam fisicamente razoáveis.

4.4.3 Estabilidade causal

Note que a existência de geodésicas do tipo tempo direcionadas para o futuro e fechadas não é o único tipo de anomalia causal que pode acontecer. Curvas fechadas do tipo luz podem existir mesmo quando M satisfaz à condição cronológica (por exemplo, considere M como sendo o espaço-tempo obtido através do sub-conjunto de \mathcal{M} dado por $\{(x^0, x^1, x^2, x^3) \in \mathcal{M} \mid 0 \leq x^0 \leq 1\}$ identificando os pontos da forma $(0, x^1, x^2, x^3)$ e $(1, x^1 + 1, x^2, x^3)$).

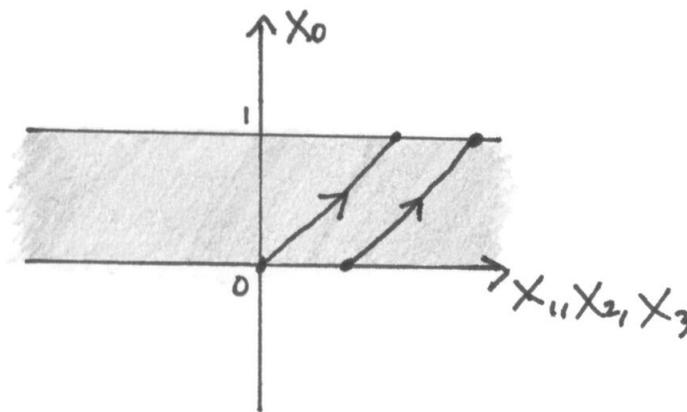


Figura 4.8: Espaço-tempo no qual existem curvas fechadas do tipo luz.

Além disso, podemos ter um espaço-tempo onde não ocorre nenhum dos dois casos acima, mas onde existem geodésicas do tipo tempo direcionadas para o futuro que entram em vizinhanças cada vez menores de um ponto dado da curva. Esses espaços-tempo estão “no limiar” de possuírem geodésicas do tipo tempo direcionadas para o futuro e fechadas. Dessa forma, alguma perturbação suficientemente pequena da métrica pode levar esses espaços-tempo a deixarem de ser estavelmente causais. Como a física é feita a partir de medições e nenhuma medição é totalmente precisa, queremos evitar anomalias desse tipo. Felizmente, Stephen Hawking⁶ nos deu a resposta para quais espaços-tempo são bem comportados nesse sentido [Liv:Haw75]. Daremos uma definição que a princípio nada tem a ver com a discussão acima e mais adiante a justificaremos.

Definição 4.20. Dizemos que M é *estavelmente causal* se existe uma função suave $T : M \rightarrow \mathbb{R}$ cujo gradiente⁷ $\text{grad } T$ é sempre do tipo tempo, ou seja, $g(\text{grad } T(p), \text{grad } T(p)) > 0$, para todo $p \in M$. Tal função é dita uma *função de tempo global*.

Observação. Note que a função T acima *NÃO* é uma função que representa uma espécie de tempo absoluto em M . Já vimos que a noção de tempo depende fortemente do observador. Daremos uma interpretação física razoável para T mais adiante.

⁶Stephen William Hawking, Oxford 08/01/1942.

⁷Lembremos que dada uma função suave $f : M \rightarrow \mathbb{R}$, onde (M, g) é uma variedade riemanniana ou lorentziana suave, seu gradiente $\text{grad } f$ é definido como o campo de vetores que satisfaz $g(\text{grad } f, V) = df(V) = V(f)$, para todo campo de vetores V . É imediato que o gradiente é único. Em coordenadas, ele é dado por $\text{grad } f = g^{ik} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^k}$, onde g^{ij} denota a inversa da matriz simétrica g_{ij} que representa a métrica em coordenadas locais.

Note que se um espaço-tempo é estavelmente causal então ele é orientável no tempo. Note também que estabilidade causal implica a condição cronológica, pois a função T goza de uma propriedade importante. De fato, pela definição do gradiente de uma função, temos que

$$g(\text{grad } T(p), V) = dT(V) = V(T).$$

Dessa forma, se $V = \alpha'(0)$, onde α é uma curva tipo tempo, então

$$\frac{d}{dt}(T \circ \alpha)(0) = \alpha'(0)(T) = g(\text{grad } T(p), \alpha'(0)).$$

Como ambos os vetores $\text{grad } T$ e $\alpha'(0)$ são do tipo tempo, não podemos ter $g(\text{grad } T(p), \alpha'(0)) = 0$, pelo Corolário 3.6. Isso prova que T é uma função estritamente monótona quando olhada sobre curvas do tipo tempo. Com isso podemos provar que a existência de tal T implica que o espaço-tempo em questão satisfaz a condição cronológica, pois a existência de uma curva fechada do tipo tempo contradiz a propriedade de T acima deduzida. De fato, se $\alpha : [0, 1] \rightarrow M$ é uma curva fechada do tipo tempo, então temos que $\alpha(0) = \alpha(1)$, mas por outro lado temos que $T(\alpha(0)) < T(\alpha(1))$ ou $T(\alpha(0)) > T(\alpha(1))$. Tomando então $\text{grad } T$ como campo de vetores que dá uma orientação no tempo em M , podemos assumir que T cresce sobre curvas do tipo tempo direcionadas para o futuro (caso contrário, troque T por $-T$).

Agora podemos provar que nossos três exemplos são orientáveis no tempo e satisfazem à condição cronológica. Para isso, basta exibirmos uma função de tempo global para eles. No caso de \mathcal{M} e \mathcal{E} , essa função é bem clara. De fato, defina $T_1 : \mathcal{M} \rightarrow \mathbb{R}$ como $T_1(u^0, u^1, u^2, u^3) = u^0$ e $T_2 : \mathcal{E} \rightarrow \mathbb{R}$ também como $T_2(u^0, u^1, u^2, u^3) = u^0$. Uma verificação fácil nos diz que $\text{grad } T_1 = (1, 0, 0, 0)$ e $\text{grad } T_2 = (1, 0, 0, 0)$, ambos do tipo tempo nos seus respectivos espaços tangentes. No caso de \mathcal{D} , defina primeiro $T : \mathbb{R}^5 \rightarrow \mathbb{R}$ como $T(u^0, u^1, u^2, u^3, u^4) = u^0$ e considere $T_3 : \mathcal{D} \rightarrow \mathbb{R}$ como a restrição de T a \mathcal{D} . Temos então que

$$\text{grad } T_3 = g^{ik} \frac{\partial(T_3 \circ \varphi)}{\partial u^i} \frac{\partial}{\partial u^k} = g^{ik} \frac{\partial \sinh(u^0)}{\partial u^i} \frac{\partial}{\partial u^k} = g^{0k} \cosh(u^0) \frac{\partial}{\partial u^k} = \cosh(u^0) \frac{\partial}{\partial u^0},$$

onde φ é a parametrização de \mathcal{D} que usamos no exemplo. Note que só provamos esse fato para essa parametrização, porém, como a função T_3 só leva em consideração a primeira coordenada e ela é sempre a mesma no atlas que construímos para \mathcal{D} , mostramos que T_3 é de fato uma função de tempo global em \mathcal{D} .

Lembremos que nossa motivação para definir estabilidade causal foi o fato de existirem exemplos de espaços-tempo no qual existem geodésicas do tipo tempo “quase”fechadas tais que pequenas perturbações da métrica poderiam tornar essas curvas fechadas e que queríamos evitar esse tipo de anomalia, não só por não serem razoáveis do ponto de vista físico, mas também porque a métrica é um objeto físico construído a partir de medições no espaço e no tempo, e medições são naturalmente imprecisas⁸. Dessa forma, não podemos considerar a métrica como uma quantidade absoluta, mas sim passível de erros em relação à realidade. Portanto qualquer hipótese fisicamente razoável sobre o espaço-tempo não pode ser sensível a pequenas perturbações da métrica. Isso nos leva a acreditar que a condição cronológica é uma condição um pouco fraca e que devemos impor uma condição mais forte em M . Mas o que a função T tem a ver com essa história? Nossa idéia para corrigir essa falha seria, a princípio, definir uma topologia adequada no conjunto das métricas lorentzianas que M admite e dizer que aceitamos g como uma métrica razoável no sentido discutido acima se existe uma vizinhança de g na qual toda métrica satisfaz a condição cronológica. De fato, essa é a idéia introduzida por Hawking em [Liv:Haw75]. Pode-se provar que essa construção e a que fizemos aqui são equivalentes. Isso é bastante difícil e foge do escopo desse texto. Uma prova pode ser encontrada em [Liv:Wal84] e [Liv:Haw75].

Vamos estudar um pouco mais a fundo as propriedades de uma função de tempo global T em M . Como $\text{grad } T$ é sempre do tipo tempo, ele nunca se anula. Dessa forma, todo ponto r na imagem de T é um valor

⁸De fato, o Princípio da Incerteza de Heisenberg nos dá uma cota inferior para o “tamanho” dessa imprecisão, mas não entraremos nesses detalhes sobre Mecânica Quântica aqui. Além disso Relatividade Geral e Mecânica Quântica são difíceis de se juntar em uma Teoria consistente.

regular de T , e portanto, o conjunto $S = T^{-1}(r)$ é uma subvariedade de codimensão 1 em M . Pela continuidade de T , é uma subvariedade fechada. Mais ainda: dados $p, q \in S$ temos que nem $p \ll q$ nem $q \ll p$ acontece. De fato, suponha, por exemplo, que $p \ll q$. Então existe uma curva diferenciável $\alpha : [0, 1] \rightarrow M$ do tipo tempo e que aponta para o futuro tal que $\alpha(0) = p$ e $\alpha(1) = q$, pela Proposição 4.15. Sabemos que T é escritamente crescente ao longo de curvas de tal tipo. Então teríamos que $T(p) < T(q)$, uma contradição.

Um conjunto que satisfaz a propriedade de que nenhum ponto seu antecede cronologicamente outro é dito *acronal*. Provamos acima que S é um conjunto acronal.

Tome agora $p \in S$ e $V \in T_p S$. Lembremos que podemos supor que $T_p S$ é um sub-espaço de $T_p M$. Dessa forma, podemos calcular

$$g(\text{grad } T(p), V) = V(T) = \frac{d}{dt}(T \circ \alpha)(0),$$

onde α é uma curva diferenciável tal que $\alpha(0) = p$ e $\alpha'(0) = V$. Porém, como podemos tomar a curva α tal que sua imagem está contida em S e T é constante em S , temos que $g(\text{grad } T(p), V) = 0$. Como $p \in S$ e $V \in T_p S$ são arbitrários, isso mostra que $\text{grad } T$ é perpendicular à S . Como $\text{grad } T$ é do tipo tempo, o Corolário 3.6 nos diz que todos os vetores de $T_p S$ são do tipo espaço. Uma subvariedade de M satisfazendo tal propriedade é dita uma *subvariedade do tipo espaço*.

Resumindo, concluímos que toda superfície de nível de T que não seja vazia é uma subvariedade fechada de codimensão 1, acronal e tipo espaço. Apesar das dificuldades em dar uma interpretação física razoável para a função T , encontramos que suas superfícies de nível têm propriedades muito próximas do que gostaríamos de chamar de um “espaço instantâneo”, ou seja, todo o espaço-tempo em um certo instante de tempo fixado. Por isso chamaremos qualquer curva de nível não vazia de uma função tempo global de uma *fatia do tipo espaço de M* . Os exemplos mais imediatos que temos em mãos são as superfícies com $u^0 = \text{constante}$ em \mathcal{M} , \mathcal{D} e \mathcal{E} . Essas fatias são difeomorfas a \mathbb{R}^3 no caso de \mathcal{M} e \mathcal{E} e a \mathbb{S}^3 no caso de \mathcal{D} .

Observação. ATENÇÃO! Vale ressaltar mais uma vez que NÃO estamos dizendo que a função T representa uma noção de tempo universal, independente do observador. Não estamos dizendo também todos os observadores concordam que, fixado um instante de tempo, o conjunto dos eventos simultâneos é dado por S . Repare que essa afirmação sequer faz sentido, pois no contexto da Relatividade Geral um observador só introduz coordenadas (ou seja, sabe medir tempo e distância) localmente. Como podemos notar pelos exemplos acima, no caso particular de \mathcal{M} de fato as superfícies S coincidem com as superfícies de simultaneidade para algum observador (aquele cuja trajetória está contida no eixo e_0). Mas qualquer outro observador discordará desse primeiro em relação às superfícies de simultaneidade. Dessa forma, usaremos a função de tempo global pensando somente na sua influência na geometria do espaço-tempo em questão, e não tentando elocubrar sobre o fato de sua existência implicar em consequências físicas exóticas.

4.4.4 Hiperbolicidade global

Lembremos que o conjunto $I^+(p)$ representa, de certa forma, os eventos de M que podem ser influenciados por p . Como uma superfície S suave fechada, do tipo espaço, de codimensão 1 e acronal é interpretada como sendo todo o espaço-tempo em um certo instante de tempo fixado, queremos saber quanto que da informação contida em M é influenciada pela informação em S . Para isso vamos precisar de uma técnica, o conceito de uma *curva sem pontos finais*.

Definição 4.21. Seja $\alpha : I \rightarrow M$ uma curva contínua, onde $I \subset \mathbb{R}$ é um intervalo, $a = \inf I$ e $b = \sup I$ (possivelmente teremos $a = -\infty$ e $b = \infty$). Um ponto $p \in M$ é dito um *ponto final* de α se temos que $\lim_{t \rightarrow a^+} \alpha(t) = p$ ou $\lim_{t \rightarrow b^-} \alpha(t) = p$. Uma curva é dita *sem pontos finais* se nenhum ponto de M é ponto final

da curva em questão. Se α é uma curva do tipo tempo direcionada para o futuro e $p = \lim_{t \rightarrow a^+} \alpha(t)$, então dizemos que p é dito um *ponto final passado* de α . Definimos um *ponto final futuro* de α de maneira análoga. Finalmente, dizemos que α não tem pontos finais passados (respectivamente futuros) se nenhum $p \in M$ é ponto final passado (respectivamente futuro) de α .

Note que uma curva ter ou não pontos finais depende do espaço-tempo no qual estamos olhando a curva. Por exemplo, a curva $(0, \infty) \ni t \mapsto (t, 0, 0, 0) \in M$ tem $(0, 0, 0, 0)$ como seu ponto final passado, mas quando a consideramos em $M \setminus \{(0, 0, 0, 0)\}$ ela deixa de ter um ponto final passado.

Definição 4.22. Seja $S \subset M$ uma superfície fechada, do tipo espaço, de codimensão 1 e acronal. Definimos o *domínio de dependência futuro* de S , denotado por $D^+(S)$, como o conjunto de todos os pontos $p \in M$ tais que toda viagem que passa por p e não tem pontos finais passados contém pelo menos um ponto de S . O *domínio de dependência passado* de S é definido de maneira análoga, trocando a palavra “passado” por “futuro”. O denotamos por $D^-(S)$. Finalmente, o *domínio de dependência* de S é definido como $D(S) = D^+(S) \cup D^-(S)$.

A interpretação dessa definição é algo um pouco sutil. Analisemos primeiro e em mais detalhes o caso de $D^+(S)$, pois o outro caso é análogo. Primeiramente, note que seria mais razoável usarmos curvas do tipo tempo direcionadas para o futuro em vez de usarmos viagens nas definições acima. Porém, a Proposição 4.15 nos garante que teríamos definições equivalentes, no caso de S acronal⁹. No caso de um conjunto S arbitrário a definição usando viagens e curvas suaves não é equivalente. Para mais detalhes veja [Liv:Pen87].

Apesar disso, o uso de viagens é mais conveniente para se trabalhar, como já tivemos a oportunidade de notar na prova de alguns teoremas. O fato de *toda* viagem (e portanto, toda curva direcionada para o futuro do tipo tempo) que passa por p e não tem pontos finais passados conter pelo menos um ponto de S quer dizer que tudo que aconteça que possa influenciar algo no ponto p estará “registrado” em S . Note também que quando escolhemos viagens ou curvas do tipo tempo direcionadas para o futuro e interpretamos $D^+(S)$ como acima estamos automaticamente assumindo que não existem informações que se propaguem mais rápido do que a luz, ou seja, informações que influenciam p não se propagam em curvas com a seguinte cara:

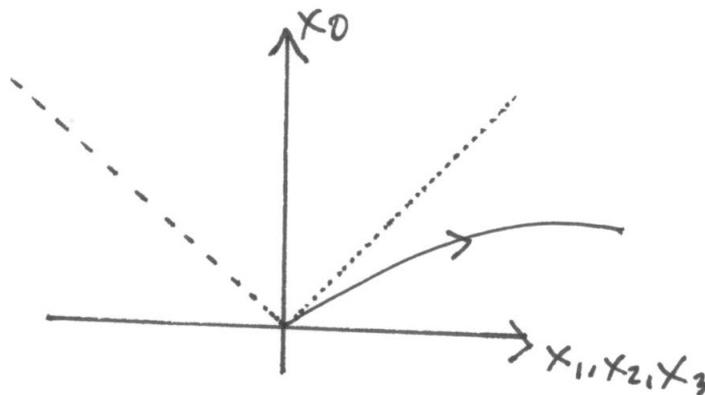


Figura 4.9: Informações não se propagam sobre curvas assim.

A interpretação de $D^-(S)$ é análoga. Dessa forma, deveríamos imaginar que se tivermos informação apropriada em S , então deveríamos conseguir prever o que acontece em $p \in D^+(S)$. Analogamente, o conhecimento de condições sobre S nos deveria permitir “retroceder” e deduzir condições em pontos

⁹Repare que poderíamos ter definido $D^\pm(S)$ para um conjunto S arbitrário, mas isso não tem interpretação física razoável.

$q \in D^-(S)$. Por outro lado, se tivermos $p \in I^+(S)$ e $p \notin D^+(S)$, então seria possível enviar um sinal que influencia p de modo que ele não esteja registrado em S , e dessa forma o conhecimento de informações somente em S não seria suficiente para determinar o que ocorre em p . Essa intuição está correta, e o interesse no estudo do conjunto $D^+(S)$ reside no fato de que através dele é possível definirmos um problema de valor inicial, também chamado de *Problema de Cauchy*, bem-posto¹⁰ com as Equações de Einstein.

Tendo em mente o raciocínio do parágrafo acima, um espaço-tempo M que contém uma superfície S que é fechada, acronal do tipo espaço e de codimensão 1 tal que $D(S) = M$ é o que temos de mais determinístico possível, pois o conhecimento de condições sobre S nos permite inferir informações sobre todos os pontos de M .

Definição 4.23. Uma superfície S fechada, acronal, do tipo espaço e de codimensão 1 tal que $D(S) = M$ é dita uma *superfície de Cauchy*. M é dito *globalmente hiperbólico* se contém uma superfície de Cauchy.

Um exemplo de superfície de Cauchy são os subconjuntos de M dados por $S_c = \{(x^0, x^1, x^2, x^3) \in M \mid x^0 = c\}$, onde c é um número real qualquer. Uma propriedade extremamente importante de espaços-tempo globalmente hiperbólicos é a seguinte:

Teorema 4.24. *Seja M um espaço-tempo globalmente hiperbólico. Então M é estavelmente causal e existe uma função de tempo global T tal que todas suas superfícies de nível $T^{-1}(r)$ são superfícies de Cauchy difeomorfas.*

Esse é dito o *Teorema de Geroch*¹¹. Sua prova no caso contínuo pode ser encontrada em [Art:Ger70]. A prova rigorosa no caso diferenciável é recente e pode ser encontrada em [Art:Ber03] e [Art:Ber05].

4.5 A Equação de Einstein

No começo desse capítulo, quando deduzimos que a trajetória de partículas livres sob a influência de campos gravitacionais devem ser geodésicas, somos automaticamente levados a pensar que existe alguma relação entre campos gravitacionais e a geometria do espaço-tempo, ou seja, sua curvatura. Para introduzi-la, precisamos relembrar brevemente alguns conceitos de Geometria Riemanniana que valem também no contexto Lorentziano, pois eles dependem só da existência de uma conexão na variedade.

4.5.1 Alguns conceitos de Geometria (semi)-Riemanniana

Nada aqui será provado nem motivado, já que isso fugiria do escopo desse texto. Motivações e demonstrações podem ser encontradas em [Liv:Lee97], [Liv:Spi99], [Liv:Onl83] e [Liv:Man08].

Nessa seção, M é uma variedade diferenciável munida de uma conexão linear ∇ .

Definição 4.25. O *tensor de curvatura* é definido como

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z.$$

Proposição 4.26. *Em coordenadas, o tensor de curvatura é dado por*

$$R^l_{ijk} = \frac{\partial \Gamma^l_{jk}}{\partial x^i} - \frac{\partial \Gamma^l_{ik}}{\partial x^j} + \Gamma^m_{jk} \Gamma^l_{im} - \Gamma^m_{ik} \Gamma^l_{jm},$$

onde $\Gamma^k_{ij} \partial_k = \nabla_{\partial_i} \partial_j$ são os símbolos de Christoffel associados à conexão ∇ .

¹⁰O que é um *problema bem-posto*? Não existe uma definição rigorosa e absoluta que permeia todos os problemas da Física e da Matemática. Cada problema tem seu próprio conceito de bem-posto.

¹¹Robert Geroch, Akron 01/06/1942.

O tensor de curvatura nos dá informação completa sobre a geometria da variedade, porém precisamos de muitas funções para determiná-lo (a princípio n^4 funções, mas esse número pode ser reduzido usando as simetrias do tensor de curvatura). Podemos obter informação suficiente sobre a geometria através de uma variação do tensor de curvatura, o *tensor de Ricci*. Para isso, vamos relembrar alguns conceitos sobre contração tensorial. Seja A um $(1, 1)$ -tensor. Em coordenadas, A se escreve como

$$A = A^i_j \partial_i \otimes dx^j.$$

Defina $C(A) \in C^\infty(M)$ como

$$C(A) = A^i_i = A(dx^i, \partial_i).$$

Tal definição é independente de coordenadas, pois

$$A\left(dy^m, \frac{\partial}{\partial y^m}\right) = A\left(\frac{\partial y^m}{\partial x^i} dx^i, \frac{\partial x^j}{\partial y^m} \frac{\partial}{\partial x^j}\right) = \frac{\partial y^m}{\partial x^i} \frac{\partial x^j}{\partial y^m} A\left(dx^i, \frac{\partial}{\partial x^j}\right) = \delta^j_i A\left(dx^i, \frac{\partial}{\partial x^j}\right) = A\left(dx^i, \frac{\partial}{\partial x^i}\right).$$

Dessa forma, $C(A)$ está bem definida em toda M .

Seja agora A um (r, s) -tensor. $C^i_j(A)$, a *contração de A nos índices i e j* , é definida da seguinte forma: fixados $\omega^1, \dots, \omega^{r-1}$ 1-formas e X_1, \dots, X_{s-1} campos de vetores, a aplicação

$$(\omega, X) \mapsto A(\omega^1, \dots, \omega, \dots, \omega^{r-1}, X_1, \dots, X, \dots, X_{s-1})$$

é um $(1, 1)$ -tensor (na expressão acima, ω ocupa a i -ésima posição nas 1-formas e X ocupa a j -ésima posição nos campos de vetores). Contraindo esse $(1, 1)$ -tensor, temos uma função real

$$(C^i_j(A))(\omega^1, \dots, \omega^{r-1}, X_1, \dots, X_{s-1}).$$

Para vermos como é a expressão de uma contração em coordenadas, vejamos um exemplo. Seja A um $(2, 3)$ -tensor. $C^1_3 A$ é o $(1, 2)$ -tensor dado por

$$(C^1_3 A)(\omega, X, Y) = C(A(\cdot, \omega, X, Y, \cdot)).$$

Em coordenadas:

$$(C^1_3 A)^k_{ij} = (C^1_3 A)(dx^k, \partial_i, \partial_j) = C(A(\cdot, dx^k, \partial_i, \partial_j, \cdot)) = A(dx^m, dx^k, \partial_i, \partial_j, \partial_m) = A^m_{ijm}.$$

Definição 4.27. Na notação acima, o *tensor de Ricci* é definido como $\text{Ric} = C^1_1 R$. Em coordenadas, $\text{Ric}_{ij} = R^k_{kij}$. Para uma interpretação do que o tensor de Ricci mede, veja [Liv:Man08].

Definição 4.28. A *curvatura escalar* é definida como o traço de Ric com respeito a g , ou seja

$$S = \text{Tr}_g \text{Ric} = R^i_i = g^{ij} R_{ij}.$$

Para uma interpretação desse conceito, veja [Liv:Man08]

4.5.2 O Tensor de Energia-Momento e as Equações de Einstein

No Capítulo 3 definimos o que é um tensor de energia-momento em M e demos uma interpretação física para tal objeto matemático. Motivados por essa discussão, a definição no caso geral é exatamente a mesma:

Definição 4.29. Um *tensor de energia-momento* em um espaço-tempo M é um $(0, 2)$ -tensor simétrico, denotado por T , que satisfaz $T(V, V) \geq 0$ para todo V do tipo tempo.

4.5. A EQUAÇÃO DE EINSTEIN

Lembremos do Capítulo 3 que a quantidade $T(V, V)$ representa a densidade de energia em relação a um observador que tenha velocidade V . Assim, é razoável supormos que essa quantidade é positiva. É possível impor uma condição mais forte no tensor de energia-momento, mas para nossos propósitos essa será suficiente.

As equações mais importantes para a teoria da Relatividade Geral são sem dúvida as *equações de Einstein*. Como já observamos, são elas que dizem a maneira como matéria e geometria integrem entre si em um determinado espaço-tempo. Na notação da seção acima, as equações se escrevem como

$$\text{Ric} - \frac{1}{2}Sg = 8\pi T.$$

Apesar de sua forma elegante, esse conjunto de equações é extremamente complicado de se trabalhar. Usando simetrias do tensor de curvatura, temos um sistema não-linear de 10 equações diferenciais parciais para a métrica g , já que todas as quantidades envolvidas do lado esquerdo podem ser escritas em função da métrica. É um bom exercício escrever as equações de Einstein em um sistema de coordenadas¹².

Logo após a publicação dessas equações em 1915 no artigo [Art:Ein15], várias soluções surgiram, algumas delas modelando espaços-tempo não estacionários¹³. Einstein acreditava fortemente que isso não era fisicamente razoável, e introduziu um novo termo nas equações, chamado de *constante cosmológica*. A nova forma das equações era então

$$\text{Ric} - \frac{1}{2}Sg - \Lambda g = 8\pi T.$$

Mal podia imaginar Einstein que em 1929 Hubble¹⁴ comprovaria experimentalmente que nosso universo está em expansão (veja [Art:Hub29]), fazendo Einstein voltar atrás e dizer que a constante cosmológica foi o maior erro científico de toda a sua vida. Apesar disso, algumas observações recentes indicam que a constante cosmológica é positiva porém muito pequena, da ordem de 10^{-122} em unidades adimensionais. Para mais detalhes, veja [Art:Car01].

Note que uma solução para as equações de Einstein não é somente uma métrica, mas também um espaço-tempo no qual essa métrica vive. Mas qual seria esse espaço-tempo, se não conhecemos a solução para as equações? Para quebrar esse raciocínio circular e resolver as equações de Einstein, deve-se ter um “chute”, baseado na intuição física, de qual deve ser mais ou menos a cara do nosso espaço-tempo, ou pelo menos como deve ser uma carta local para tal espaço-tempo, já que as quantidades envolvidas nas equações podem ser todas expressas em coordenadas. Para isso, se usa alguma hipótese que nunca é encontrada na natureza, como simetria esférica, homogeneidade ou isotropia, mas que nos permite encontrar uma solução com um mínimo de significado físico. Isso é uma tarefa extremamente difícil. Veja [Liv:Ste09] para uma ótima referência em soluções exatas das equações de Einstein. Não exploraremos mais a fundo aqui as equações de Einstein, pois o resultado principal desse trabalho não envolve toda sua força. De fato, os trabalhos de Stephen Hawking e Roger Penrose¹⁵ em resultados globais em Relatividade Geral foi motivado pelo fato de querer saber se algumas propriedades de soluções exatas das equações de Einstein são recorrentes ou somente consequência de hipóteses fisicamente não-realísticas.

Por fim, é claro que as equações de Einstein não surgiram da noite pro dia, por mais genial que Einstein fosse. Ela foi consequência de muito trabalho duro, e pode ser interpretada como a equação de Euler-Lagrange para o *funcional de Hilbert-Einstein*. Para mais detalhes, veja [Liv:Haw75].

¹²Eu fiz isso em um dado momento. A expressão é muito feia e grande para ser copiada aqui.

¹³Pode-se dar uma definição exata para espaços-tempo estacionários usando derivadas de Lie, mas para nossos propósitos aqui basta imaginarmos que um espaço-tempo é estacionário se ele “não muda” com o tempo.

¹⁴Edwin Powell Hubble, Marshfield 20/11/1889 - San Marino 28/09/1953.

¹⁵Sir Roger Penrose, Colchester 08/08/1931.

Capítulo 5

Teoremas de Singularidade

5.1 Motivação: Porque singularidades são importantes em Relatividade Geral?

As Equações de Einstein foram publicadas pela primeira vez em 1915 no artigo [Art:Ein15]. Desde então se procuram soluções exatas para tais equações, pois elas representam potenciais modelos para nosso universo. Algumas das primeiras soluções exatas a serem encontradas foram as *métrica de Schwarzschild*¹ [Art:Sch16] e a *métrica de Reissner²-Nordström*³. [Art:Rei16] [Art:Nor18]. A primeira modela um buraco negro proveniente do colapso gravitacional de um corpo sem rotação e sem carga elétrica, enquanto que a segunda permite que o corpo colapsado tenha carga elétrica mas não esteja em rotação⁴.

Note que todas essas soluções apresentam singularidades⁵. Uma pergunta natural é se qualquer solução das Equações de Einstein admite singularidades. A resposta é obviamente não, pois M modela um universo sem presença de matéria e não admite nenhuma singularidade. Além disso, o espaço tempo de Gödel também não apresenta singularidades, mas tem o leve inconveniente de permitir “viagens no tempo”, ou seja, existem curvas do tipo tempo fechadas. Então será que, em algum sentido, a “maioria” das soluções “razoáveis” para as Equações de Einstein são singulares? Que propriedades devemos esperar do espaço-tempo para que ele apresente ou não alguma singularidade? Surpreendentemente o primeiro a pensar em algo nessa direção foi Laplace⁶. Lógico que por mais brilhante que Laplace tenha sido, ele supostamente não conhecia a Teoria da Relatividade e não estava pensando em termos relativísticos. Mas ele provou, usando a teoria da gravitação newtoniana e fazendo a hipótese que a luz tem massa, um resultado bem interessante: se existir uma estrela que satisfaça algumas condições de compatibilidade entre sua massa e raio, então existe uma distância dessa estrela da qual nenhum raio luminoso partindo dela poderá passar. Provaremos esse resultado nesse Capítulo. O primeiro resultado desse tipo em Relatividade Geral é devido a Raychaudhuri⁷ [Art:Ray55], [Liv:Haw75], que provou que uma métrica proveniente de um tensor de energia-momento que modele uma nuvem de partículas em um universo irrotacional e sem cargas elétricas apresenta uma singularidade. Claro que esse resultado

¹Karl Schwarzschild, Frankfurt am Main, 09/10/1873 - Postdam 11/05/1916.

²Hans Jacob Reissner, Berlin 18/01/1874 - Colton 02/10/1967.

³Gunnar Nordström, Helsinki 12/03/1881 - Helsinki 24/12/1923

⁴Aqui poderia começar uma discussão infinita sobre existência de buracos negros, evidências experimentais, problemas psicológicos causados por eles, real possibilidade do colapso gravitacional de um corpo, etc... mas não entrarei nesses detalhes. Existe também uma solução que admite que o corpo colapsado esteja em rotação mas não tenha carga elétrica, chamada de *métrica de Kerr* (Roy Patrick Kerr, Kurow 16/05/1934) [Art:Ker63] e uma mais geral, que permite corpos tanto em rotação quanto com carga elétrica, a *métrica de Kerr-Newman* (Ezra “Ted” Newman, New York 17/10/1929) [Art:New65].

⁵Repare que ainda não sabemos o que é de fato uma singularidade, mas essas soluções modelam espaços-tempo no qual existem buracos negros. Isso é suficientemente estranho para chamarmos seu comportamento de singular.

⁶Pierre-Simon marquis de Laplace, Beaumont-en-Auge 23/03/1749 - Paris 05/03/1827.

⁷Amal Kumar Raychaudhuri, Barisal 14/09/1923 - ?? 18/06/2005.

é um caso muito particular, mas sua importância é muito grande por ser o primeiro resultado nessa direção. Os grandes teoremas de singularidade foram provados nas décadas de 60 e 70 por Stephen Hawking e Roger Penrose. Eles dão conta de espaços-tempo bem mais gerais, assumindo condições bem mais plausíveis do ponto de vista físico, por exemplo, a expansão do universo (que é experimentalmente verificada [Art:Hub29]) e condições sobre a cronologia do espaço-tempo como as vistas no Capítulo 4. A importância de tais teoremas não é só física, mas também matemática, pois são resultados globais em Geometria Lorentziana.

5.2 Mas o que é uma singularidade?

A princípio parece fazer sentido definir uma singularidade em um espaço-tempo como um ponto da variedade correspondente onde a métrica de Lorentz ou não está definida ou não é diferenciável. Essa definição tem alguns inconvenientes: o primeiro é que na nossa definição uma métrica de Lorentz está definida em todo espaço-tempo. Além disso, a definição também nos diz que ela é diferenciável em todo ponto da variedade em questão. Para que nossa definição contenha o caso especial de espaços-tempo singulares deveríamos tratar métricas de Lorentz que não sejam diferenciáveis, mas sim que satisfaçam condições mais fracas, talvez diferenciáveis a menos de um conjunto de medida nula, o que tornaria as coisas tecnicamente mais difíceis. Além disso as equações físicas deixam de ter sentido nas singularidades, pois nelas temos impossibilidades de efetuar medidas sobre a métrica. Mas o maior problema é que essa definição não reproduz a nossa imagem mental de uma singularidade. Se, na definição anterior, retirarmos os pontos singulares, teríamos um novo espaço-tempo, agora sem singularidades. Isso quer dizer que toda singularidade tem um caráter removível, implicando que tais objetos constituem um problema que é facilmente contornado. Dessa forma, identificar uma singularidade em um espaço-tempo se resumiria a identificar se ele veio de um espaço-tempo “maior”, do qual retiramos alguns pontos e obtemos o espaço-tempo em questão. Queremos uma definição de singularidade que traduza matematicamente o pensamento “aqui tudo dá errado e coisas muito estranhas acontecem”. Então supondo removidos os pontos aonde a métrica não é diferenciável ou não está definida, como fazemos para detectar “buracos”, ou pontos do espaço-tempo aonde as coisas dão errado? É muito simples! Basta verificarmos se coisas “caem” através deles. Como assim? Uma indicação bem forte de que “falta” alguma coisa em nosso espaço-tempo é que existe uma geodésica cujo domínio de definição não pode ser estendido para todo \mathbb{R} . Essa definição é totalmente razoável, pois ela não faz referência a conceitos exóticos como os primeiros sugeridos, e é coerente com nosso propósito: se uma geodésica simplesmente deixa de existir a partir de um certo tempo quer dizer que algo de errado aconteceu em nosso espaço-tempo. Mas note que temos três tipos de geodésicas: tipo tempo, espaço e luz. Vamos discutir mais a frente sobre isso. Agora vamos fazer uma breve pausa e discutir o conceito de completude geodésica em variedades Riemannianas e Lorentzianas para compará-las. Aqui vemos a primeira grande diferença entre as duas geometrias.

5.3 Completude Geodésica

No caso Riemanniano, temos o seguinte resultado:

Teorema 5.1. (Hopf⁸-Rinow⁹) *Seja (M, g) uma variedade Riemanniana conexa. São equivalentes:*

- *Os conjuntos fechados e limitados de M são compactos;*
- *M é um espaço métrico completo, com a distância induzida por g ;*
- *M é geodesicamente completa, ou seja, para todo $p \in M$, a aplicação exponencial \exp_p está definida em todo T_pM .*

⁸Heinz Hopf, Gräbschen 19/11/1894 - Zollikon 03/06/1971.

⁹Willi Rinow - Berlin 28/02/1907 - Greifswald 29/03/1979.

Mais ainda, cada uma dessas afirmações implica que dados dois pontos $p, q \in M$ existe uma geodésica minimizante ligando p a q .

Podemos então dizer simplesmente que uma certa variedade Riemanniana conexa é *completa* sem problemas de ambiguidade em relação às completudes geodésicas e como espaço métrico, já que elas são equivalentes. Porém, esse Teorema não vale no caso Lorentziano, pois uma métrica Lorentziana não induz uma função distância. Um Corolário imediato de tal Teorema é que toda variedade Riemanniana compacta é completa, já que todo espaço métrico compacto é completo no sentido das seqüências de Cauchy. Será que podemos provar que variedades Lorentzianas compactas são geodesicamente completas? Note que só podemos falar de completude geodésica, pois uma variedade Lorentziana não tem uma função distância natural, como ocorre nas variedades Riemannianas. Para isso, vamos tentar caracterizar a completude geodésica no caso particular de uma variedade Riemanniana compacta de outra forma. Considere em TM o *fluxo geodésico*¹⁰: se $U \subset M$ é o domínio de uma carta de M e $\pi : TM \rightarrow M$ é a projeção canônica, seja (u^i, v^i) as coordenadas padrão em $\pi^{-1}(U) \subset TM$. Defina o seguinte campo de vetores em $\pi^{-1}(U)$:

$$G_{(x,v)} = v^k \frac{\partial}{\partial x^k} - v^i v^j \Gamma_{ij}^k(x) \frac{\partial}{\partial v^k}.$$

As curvas integrais de G satisfazem o seguinte sistema de EDO's:

$$\begin{aligned} \dot{x}^k(t) &= v^k(t) \\ \dot{v}^k(t) &= -v^i(t)v^j(t)\Gamma_{ij}^k(x(t)). \end{aligned}$$

Note que a equação que uma geodésica satisfaz pode ser transformada no sistema acima através da substituição $v^k = \dot{x}^k$. A importância fundamental desse campo de vetores é que ele pode ser estendido para um campo de vetores definido em todo TM . Essa extensão é chamada de *campo de vetores geodésicos*, e seu fluxo é chamado de *fluxo geodésico*. Tal nome vem do fato que as curvas integrais se projetam através de π em geodésicas de M . E, reciprocamente, toda geodésica $\gamma(t) = (x^1(t), \dots, x^n(t))$ pode ser levantada para uma curva integral de G colocando-se $v^i(t) = \dot{x}^i(t)$. Porém, o fato mais importante que usaremos sobre G é que se f é uma função em $C^\infty(TM)$, então

$$Gf(p, V) = \left. \frac{d}{dt} \right|_{t=0} f(\gamma_V(t), \dot{\gamma}_V(t)),$$

onde γ_V é a única geodésica satisfazendo $\gamma_V(0) = p$ e $\dot{\gamma}_V(0) = V$.

Levando em consideração que $g(\dot{\gamma}(t), \dot{\gamma}(t))$ é constante para toda geodésica γ , pois $\frac{d}{dt}g(\dot{\gamma}(t), \dot{\gamma}(t)) = 2g(D_t\dot{\gamma}(t), \dot{\gamma}(t)) = 0$, podemos considerar, sem perda de generalidade, a restrição de G ao fibrado tangente unitário de M , a saber,

$$UM = \{(p, V) \in TM \mid g(V, V) = 1\}.$$

Como M é compacta e as fibras também o são, temos que UM é uma variedade compacta. Dessa forma, todas as curvas integrais de G restrito a UM estão definidas para todo tempo¹¹. Dessa forma, as geodésicas de M também estão definidas para todo tempo, o que implica que M é geodesicamente completa¹². Será que podemos considerar essa abordagem em uma variedade lorentziana? Infelizmente não, pois usamos fortemente o fato de que UM é compacta. Em uma variedade lorentziana (ou semi-riemanniana em geral) podemos definir os conjuntos

$$U^+M = \{(p, V) \in TM \mid g(V, V) = 1\} \text{ e}$$

¹⁰Tudo que será dito aqui está provado em [Liv:Lee97].

¹¹É verdade que toda curva integral de um campo de vetores sobre uma variedade compacta está definida pra todo tempo. Para uma prova disso, veja [Liv:Spi99].

¹²Note que, como não estamos usando o Teorema de Hopf-Rinow, não podemos simplesmente concluir que M é completa sem problemas de ambiguidade. Além disso, estamos interessados principalmente na completude geodésica de M , e não em sua completude como espaço métrico.

$$U^-M = \{(p, V) \in TM \mid g(V, V) = -1\}.$$

Mas como as fibras desses fibrados não são compactas, a compacidade de M não implica na compacidade de $U^\pm M$, e portanto, o argumento acima não é válido no caso lorentziano. De fato, existem exemplos de variedades lorentzianas compactas que não são geodesicamente completas. Nos perguntamos então quando uma variedade lorentziana compacta é geodesicamente completa. A nível de curiosidade, vamos enunciar aqui alguns resultados nessa direção:

Teorema 5.2 ([Art:Mar73]). *Toda variedade lorentziana (semi-riemanniana, em geral) homogênea¹³ é completa.*

A idéia da prova consiste em mostrar que TM é a união disjunta de compactos que são invariantes pelo fluxo geodésico. Dessa forma, toda curva integral de G está contida em um compacto de TM e, portanto, está definida para todo tempo. Resultados um pouco diferentes são apresentados abaixo:

Teorema 5.3 ([Art:Rom95]). *Uma variedade lorentziana compacta que admite um campo de vetores X que é conforme¹⁴ e do tipo tempo é completa.*

Teorema 5.4 ([Art:Car89], [Art:Kli96]). *Toda variedade lorentziana compacta de curvatura seccional constante é geodesicamente completa.*

Temos também o seguinte resultado, que é deveras surpreendente:

Teorema 5.5 ([Art:Che63], [Art:Cal62]). *Não existe variedade lorentziana compacta de curvatura seccional constante positiva.*

Para terminar de ilustrar quão estranha pode ser a questão da completude em variedades lorentzianas, temos o seguinte resultado:

Teorema 5.6 ([Art:Rom94]). *Existem duas métricas lorentzianas no toro com a mesma curvatura de Gauß em todo ponto tais que uma é completa e a outra não.*

A estratégia a ser adotada então é contrária. Vamos tentar caracterizar quando que uma variedade lorentziana é *incompleta*. Para aquecer, vamos começar com o Teorema de Laplace que citamos na introdução.

5.4 O Teorema de Laplace

Lembremos qual é o espírito do resultado provado por Laplace. Se uma estrela satisfaz certas condições sobre sua massa e raio, então existe uma distância dessa estrela da qual qualquer raio luminoso não pode passar. Vamos tentar deduzir quais são essas condições. Como esse resultado é baseado na mecânica newtoniana, temos que supor que um fóton tem massa $\mu > 0$, o que sabemos ser falso na Teoria da Relatividade. Vamos considerar também que a velocidade do fóton ao sair da estrela independe da estrela, o que sabemos ser verdadeiro. Pela conservação de energia, temos que

$$E_{p_1} + E_{c_1} = E_{p_2} + E_{c_2},$$

onde E_{p_i} denota a energia potencial no ponto P_i e E_{c_i} representa a energia cinética no ponto P_i . Lembremos que o potencial gravitacional é dado por $-GmM/d$, onde G é a constante gravitacional, m e M são as massas dos corpos em questão e d é sua distância. Seja m a massa da estrela e R, r como na figura acima. Temos então que

¹³Uma variedade semi-riemanniana é dita *homogênea* se admite uma ação suave, transitiva e por isometrias de um grupo de Lie. Intuitivamente, ela é "igual" em todas as direções.

¹⁴Um campo de vetores é dito *conforme* se $\mathcal{L}_X g = \sigma g$, onde \mathcal{L} denota a derivada de Lie e $\sigma \in C^\infty(M)$.

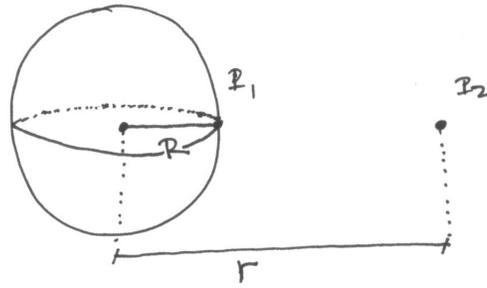


Figura 5.1: Teorema de Laplace.

$$\frac{-Gm\mu}{R} + \frac{1}{2}\mu|v_0|^2 = \frac{-Gm\mu}{r} + \frac{1}{2}\mu|v|^2,$$

onde v é a velocidade do fóton no ponto P_2 . Queremos ter que $v = 0$, ou seja,

$$\frac{1}{2}\mu|v|^2 = \frac{Gm\mu}{r} - \frac{Gm\mu}{R} + \frac{1}{2}\mu|v_0|^2 = 0.$$

Isso implica que

$$\begin{aligned} \frac{2Gm}{r} &= \frac{2Gm}{R} - |v_0|^2 \Leftrightarrow \\ r &= \frac{2Gm}{\frac{2Gm}{R} - |v_0|^2}. \end{aligned}$$

Devemos ter que $r > 0$. para isso, devemos ter que

$$|v_0|^2 < \frac{2Gm}{R}.$$

Como v_0 não depende da estrela, temos que ter que m e R satisfaçam a condição acima para que r seja um número admissível. Além disso, devemos ter que $r > R$, o que é facilmente verificado ser verdadeiro. Dessa forma, provamos o seguinte resultado:

Teorema 5.7 (Laplace). *Considere uma estrela de massa m e raio R satisfazendo $|v_0|^2 < 2Gm/R$, onde v_0 é a velocidade da luz. Então em $r = \frac{2Gm}{\frac{2Gm}{R} - |v_0|^2} > R$ temos que a velocidade de um fóton que sai da estrela é nula, ou seja, o fóton não passará dessa distância, ficando confinado no conjunto $\{R < x < r\}$.*

Esse Teorema é muito simpático por motivos históricos, mas se baseia em conceitos que sabemos estarem errados. Para estrelas muito densas a Teoria da gravitação newtoniana não se aplica mais, sendo necessária a Relatividade Geral. Para ilustrar, vamos usar esse resultado com o Sol para ver o que obtemos. Primeiro, note que:

- $v_0 =$ velocidade da luz no vácuo $\approx 3 \times 10^8 m/s$;
- $G =$ constante gravitacional $\approx 6.67 \times 10^{-11} N(m/kg)^2$;
- $m =$ massa do Sol $\approx 1.98 \times 10^{30} kg$;

- $R =$ raio do Sol $\approx 6.95 \times 10^8 m$.

Suponha agora que a gente queira verificar a validade desse Teorema em nosso Sistema Solar. A quantidade $2Gm/R$ vale aproximadamente 3.80×10^{11} . Deveríamos ter que a velocidade da luz no vácuo fosse menor que 6.16×10^5 para que o resultado fosse verdade, ou seja, três ordens de grandeza menor do que realmente é.

Tudo bem, não podemos mudar a velocidade da luz. Mas quais deveriam ser a massa e o raio do Sol para satisfazer a estimativa do Teorema? Caso tivéssemos a liberdade de alterar a massa do Sol, deveríamos tê-la maior que $4.69 \times 10^{35} kg$, 5 ordens de grandeza maior do que sua massa real. Para efeito de comparação, a estrela mais massiva que se conhece é a R136a1 [Int:Wik], com massa aproximada de 5.247×10^{32} , 256 vezes a massa solar. Ainda assim estamos muito longe do previsto pelo Teorema.

Agora, caso pudéssemos alterar o raio do Sol, teríamos que comprimi-lo muito, pois ele deveria ser menor que $2936 km$. O raio da Terra no Equador é de aproximadamente $6353 km$.

5.5 Um Teorema de S. Hawking

O Teorema que provaremos agora, devido a S. Hawking em [Art:Haw66A], [Art:Haw66B] e [Art:Haw67], mostra a existência de uma singularidade no passado, provando que certas geodésicas do tipo tempo, inextensíveis e que apontam para o futuro são incompletas pela esquerda. É claro nas demonstrações do Teorema que essa incompletude não depende da orientação temporal do espaço-tempo, mas somente da conexão de Levi-Civita associada à métrica, mas para a interpretação física desses resultados a orientação temporal é crucial.

Como a demonstração do Teorema é longa e tem vários pré-requisitos, vamos fazer aqui um pequeno roteiro:

- Na Seção 5.5.1 vamos fazer algumas definições, para fixar a notação e relembrar alguns conceitos. Além disso, provaremos alguns resultados auxiliares sobre o funcional comprimento de arco.
- A Seção 5.5.2 contém a demonstração do Teorema de Hawking, porém lança mão do Lema 5.17, cuja demonstração é bastante técnica.
- Finalmente, a Seção 5.5.3 traz a prova do Lema 5.17. Essa é a Seção mais dura do texto. Primeiramente começamos provando alguns resultados auxiliares. O primeiro ponto chave na prova de tal resultado é o Lema 5.29, que traz uma equivalência parcial para o conceito de espaços-tempo globalmente hiperbólicos. Não menos importante são os Lemas 5.30 e 5.35, que falam sobre a compacidade de um certo conjunto e a semi-continuidade superior de um funcional definido sobre tal conjunto, respectivamente. Por fim, o Lema 5.17 segue desses resultados acima e do Teorema de Avez-Seifert (Teorema 5.39), que aqui não provamos.

5.5.1 Preliminares matemáticas

Definição 5.8. Dada uma curva $\gamma : [a, b] \rightarrow M$, uma *variação* de γ é uma aplicação suave $\Gamma : [a, b] \times (-\delta, \delta) \rightarrow M$ tal que $\Gamma(u, 0) = \gamma(u)$.

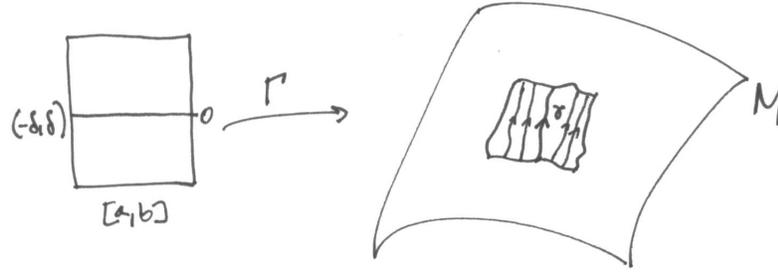


Figura 5.2: Variação de uma curva γ .

Pensamos em Γ como sendo uma família suave de curvas que estão “próximas” de γ .

Toda variação produz duas famílias de curvas: as *curvas principais* $u \mapsto \Gamma(u, v)$ definidas em $[a, b]$ com v constante e denotadas por $\Gamma_v(u)$ e as *curvas transversais* $v \mapsto \Gamma(u, v)$ definidas em $(-\delta, \delta)$ com u constante e denotadas por $\Gamma^{(u)}(v)$. Portanto, temos dois campos de vetores ao longo de Γ , que são os vetores tangentes a tais curvas. Os denotamos por

$$\partial_u \Gamma(u, v) := \frac{d}{du} \Gamma_v(u); \quad \partial_v \Gamma(u, v) := \frac{d}{dv} \Gamma^{(u)}(v).$$

Se V é um campo de vetores ao longo de Γ , podemos calcular as derivadas covariantes de V tanto sobre as curvas principais quanto sobre as curvas transversais. Denotamos essas derivadas da seguinte forma:

$$D_u V := \nabla_{\partial_u \Gamma} V; \quad D_v V := \nabla_{\partial_v \Gamma} V.$$

Lema 5.9 (Lema de simetria). *Seja $\Gamma : [a, b] \times (-\delta, \delta) \rightarrow M$ uma variação. Então $D_v \partial_u \Gamma = D_u \partial_v \Gamma$.*

A prova desse lema é muito simples e pode ser encontrada em [Liv:Lee97]. A idéia da demonstração é escrever $D_v \partial_u \Gamma$ e $D_u \partial_v \Gamma$ em coordenadas e usar a simetria da conexão, ou seja, $\Gamma_{ij}^k = \Gamma_{ji}^k$.

Observação 5.10. Não confunda os símbolos de Christoffel com uma variação! Apesar de eles serem denotados pela mesma letra Γ , os símbolos de Christoffel estão sempre acompanhados dos seus índices, permitindo uma fácil distinção entre eles.

O seguinte Lema nos diz como que se relacionam os operadores $D_v D_u$ e $D_u D_v$:

Lema 5.11. *Se Γ é uma variação e V é um campo de vetores sobre Γ , então $D_v D_u V - D_u D_v V = R(\partial_v \Gamma, \partial_u \Gamma)V$.*

Novamente a prova se resume a escrever o lado esquerdo da igualdade em coordenadas, fazer algumas contas e obtermos o lado direito. Para mais detalhes, veja [Liv:Lee97].

Dada uma variação Γ , denote por $V(u)$ o campo de vetores dado por $\partial_v \Gamma(u, 0)$. Esse campo é chamado de *campo de variação* de Γ . Note que V é um campo de vetores ao longo de $\Gamma_0(u)$. Vale uma recíproca desse resultado:

Lema 5.12. *Se γ é uma curva suave em M e V é um campo de vetores sobre γ , então V é o campo de variação de alguma variação de γ .*

A idéia da prova é construir uma variação da seguinte forma: fixado u , ande um pouco sobre a geodésica que tem como vetor velocidade inicial $V(u)$, ou seja, defina $\Gamma(u, v)$ como $\exp(vV(u))$. Para mais detalhes, veja [Liv:Lee97].

No nosso caso particular aonde M é um espaço-tempo, temos o seguinte resultado:

Proposição 5.13. *Seja $\gamma : [a, b] \rightarrow M$ uma curva tipo tempo e $\Gamma : [a, b] \times (-\delta, \delta) \rightarrow M$ uma variação. Então podemos tomar δ suficientemente pequeno de modo que as curvas principais $\Gamma_v(u)$ são sempre do tipo tempo.*

Demonstração. Suponha falso o resultado. Então para todo $n \in \mathbb{N}$ existe $0 < v_n < 1/n$ tal que $\Gamma_{v_n}(u)$ não é tipo tempo. Isto implica na existência de um $u_n \in [a, b]$ tal que $\partial_u \Gamma(u_n, v_n)$ não é tipo tempo. Como a sequência (u_n, v_n) está contida em $[a, b] \times [-1, 1]$, que é compacto, existe uma sub-sequência (u_{n_k}, v_{n_k}) convergente para, digamos, (u_0, v_0) . Como $0 < v_n < 1/n$, temos que $v_0 = 0$. Dessa forma, temos que $\partial_u \Gamma(u_0, 0) = \gamma'(u_0)$ não é tipo tempo, uma contradição com o fato de γ ser do tipo tempo. \square

Dada uma variação $\Gamma : [a, b] \times (-\delta, \delta) \rightarrow M$, temos uma função associada, chamada de *comprimento de arco associado a Γ* . Ela é dada por

$$L(v) = \int_a^b |g(\partial_u \Gamma(u, v), \partial_u \Gamma(u, v))|^{1/2} du$$

e mede o comprimento das curvas principais. Sendo uma função de uma variável, podemos calcular suas derivadas. Nos será importante a segunda derivada de L em um caso particular da variação Γ . Para resultados mais gerais sobre a primeira a segunda variação do comprimento de arco, veja [Liv:Lee97].

Considere então $\gamma : [a, b] \rightarrow M$ uma geodésica do tipo tempo e direcionada para o futuro e $\Gamma : [a, b] \times (-\delta, \delta) \rightarrow M$ uma variação de γ tal que todas as curvas principais são direcionadas para o futuro. Denote por V o campo de variação de Γ e suponha que ele é ortogonal a γ , ou seja, $g(\gamma'(u), V(u)) = 0$. Para simplificar a notação, denote o integrando de L por $H(u, v)$. Como as curvas principais são do tipo tempo, podemos tirar o módulo na definição de H e temos que ela é suave, já que nunca se anula. Temos então que

$$L''(0) = \frac{d^2}{dv^2} \Big|_{v=0} \int_a^b H(u, v) du = \int_a^b \frac{\partial^2}{\partial v^2} \Big|_{v=0} H(u, v) du.$$

Calculemos então quem é $\partial^2 / \partial v^2 H(u, v)$ quando $v = 0$.

$$\frac{\partial H}{\partial v} = \frac{\partial}{\partial v} g(\partial_u \Gamma, \partial_u \Gamma)^{1/2} = \frac{2g(D_v \partial_u \Gamma, \partial_u \Gamma)}{2g(\partial_u \Gamma, \partial_u \Gamma)^{1/2}} = \frac{g(D_v \partial_u \Gamma, \partial_u \Gamma)}{H}.$$

$$\begin{aligned} \frac{\partial^2 H}{\partial v^2} &= \frac{\partial}{\partial v} \frac{g(D_v \partial_u \Gamma, \partial_u \Gamma)}{H} \\ &= \frac{H \frac{\partial}{\partial v} g(D_v \partial_u \Gamma, \partial_u \Gamma) - g(D_v \partial_u \Gamma, \partial_u \Gamma) \frac{\partial}{\partial v} H}{H^2} \\ &= \frac{1}{H^2} \left[H(g(D_v D_v \partial_u \Gamma, \partial_u \Gamma) + g(D_v \partial_u \Gamma, D_v \partial_u \Gamma)) - \frac{g(D_v \partial_u \Gamma, \partial_u \Gamma)g(D_v \partial_u \Gamma, \partial_u \Gamma)}{H} \right] \\ &= \frac{1}{H} \left[g(D_v D_v \partial_u \Gamma, \partial_u \Gamma) + g(D_v \partial_u \Gamma, D_v \partial_u \Gamma) - \frac{g(D_v \partial_u \Gamma, \partial_u \Gamma)^2}{H^2} \right]. \end{aligned}$$

Note que todos os termos dessa expressão são bem comportados, exceto $D_v D_v \partial_u \Gamma$, que a princípio não fazemos a menor idéia do que seja. Mas nem tudo está perdido! Pelo lema de simetria, temos que $D_v D_v \partial_u \Gamma = D_v D_u \partial_v \Gamma$. Pelo Lema 5.11, temos que

$$D_v D_u \partial_v \Gamma - D_u D_v \partial_v \Gamma = R(\partial_v \Gamma, \partial_u \Gamma) \partial_v \Gamma.$$

Dessa forma, concluímos que

$$D_v D_v \partial_u \Gamma = D_v D_u \partial_v \Gamma = D_u D_v \partial_v \Gamma + R(\partial_v \Gamma, \partial_u \Gamma) \partial_v \Gamma.$$

Usando o Lema de simetria em alguns termos, temos que:

$$\frac{\partial^2 H}{\partial v^2} = \frac{1}{H} \left[g(D_u D_v \partial_v \Gamma, \partial_u \Gamma) + g(R(\partial_v \Gamma, \partial_u \Gamma) \partial_v \Gamma, \partial_u \Gamma) + g(D_u \partial_v \Gamma, D_u \partial_v \Gamma) - \frac{g(D_u \partial_v \Gamma, \partial_u \Gamma)^2}{H^2} \right].$$

Antes de avaliarmos essa segunda derivada quando $v = 0$, note que:

- $H(u, 0) = g(\partial_u \Gamma(u, 0), \partial_u \Gamma(u, 0))^{1/2} = g(\gamma'(u), \gamma'(u)) = 1$;
- $\partial_u \Gamma(u, 0) = \gamma'(u)$;
- $\partial_v \Gamma(u, 0) = V(u)$;
- $D_u \partial_v \Gamma(u, 0) = D_u V(u) = D_\gamma V(u) =$ derivada covariante de V sobre γ ;
- $D_v \partial_v \Gamma(u, 0) = D_v V(u) := A(u) =$ derivada covariante do campo de variação sobre as curvas transversais. Podemos pensar nessa quantidade como uma “aceleração” do campo de variação. Daí a letra A para denotá-la;
- $D_u D_v \partial_v \Gamma(u, 0) = D_u A(u) = D_\gamma A(u) := A'(u) =$ derivada covariante de A sobre γ .

Temos então que:

$$\begin{aligned} \left. \frac{\partial^2 H}{\partial v^2} \right|_{v=0} &= g(A'(u), \gamma'(u)) + g(R(V(u), \gamma'(u))V(u), \gamma'(u)) + g(D_\gamma V(u), D_\gamma V(u)) - g(D_\gamma V(u), \gamma'(u))^2 \\ &= g(A'(u), \gamma'(u)) - g(R(V(u), \gamma'(u))\gamma'(u), V(u)) + g(D_\gamma V(u), D_\gamma V(u)) - g(D_\gamma V(u), \gamma'(u))^2. \end{aligned}$$

Note que, como V é ortogonal a γ , temos que $g(V(u), \gamma'(u)) = 0$, e portanto,

$$0 = \gamma'(u)g(V(u), \gamma'(u)) = g(D_\gamma V(u), \gamma'(u)) + g(V(u), \underline{D}_\gamma \gamma'(u) \Big|_0).$$

Logo, $g(D_\gamma V(u), \gamma'(u)) = 0$. Note também que

$$\frac{d}{du} g(\gamma'(u), A(u)) = g(\underline{D}_\gamma \gamma'(u) \Big|_0, A(u)) + g(\gamma'(u), D_\gamma A(u)) = g(\gamma'(u), D_u A(u)) = g(\gamma'(u), A'(u)).$$

Juntando esses dois novos resultados na nossa longa conta, concluímos que

$$\left. \frac{\partial^2 H}{\partial v^2} \right|_{v=0} = \frac{d}{du} g(\gamma'(u), A(u)) - g(R(V(u), \gamma'(u))\gamma'(u), V(u)) + g(D_\gamma V(u), D_\gamma V(u)),$$

e portanto,

$$L''(0) = \int_a^b -g(R(V(u), \gamma'(u))\gamma'(u), V(u)) + g(D_\gamma V(u), D_\gamma V(u)) du + g(\gamma'(u), A(u)) \Big|_{u=a}^{u=b}.$$

Lembremos que nossas superfícies de Cauchy são imagens inversas de valores regulares de uma função real suave T , que tinha como propriedade o fato de que seu gradiente é um campo de vetores do tipo tempo. Podemos supor que além disso $\text{grad } T$ aponta para o futuro (caso contrário, troque T por $-T$). Dessa forma, o campo N definido por $g(\text{grad } T, \text{grad } T)^{-1/2} \text{grad } T$ também é do tipo tempo, aponta para o futuro, é normal às superfícies de Cauchy e além disso é unitário. Se queremos alguma quantidade que meça a “taxa de expansão” do universo em um dado momento, é razoável considerarmos uma superfície de Cauchy fixa S e olharmos para a quantidade

$$H_S := (\text{div} N)|_S.$$

Sabemos que o divergente de um campo de vetores mede de fato o quanto ele está “divergindo”¹⁵. Como o campo N está intimamente relacionado com S , é razoável interpretarmos a quantidade H_S como uma taxa de expansão do universo em S . Sendo então uma quantidade importante, ela merece um nome. Chamamos H_S de *curvatura média* de S .

Então estamos interessados em algumas informações sobre a geometria de S , uma subvariedade de M . É razoável que tal informação tenha relação com a geometria de M , e isso se expressa pelas chamadas *Fórmula de Gauß* e *Equação de Weingarten*. Vejamos o que elas dizem.

Primeiramente, vamos lembrar um resultado importante sobre conexões.

Proposição 5.14. *Seja ∇ uma conexão em uma variedade M e X, Y campos de vetores sobre M . Então o valor de $\nabla_X Y$ em um ponto $p \in M$ depende somente de X_p e dos valores de Y sobre uma curva γ qualquer satisfazendo $\gamma(0) = p$ e $\gamma'(0) = X_p$.*

Uma prova pode ser encontrada em [Liv:Lee97].

Tome então X, Y campos de vetores sobre S . Dado $p \in S$, o vetor $X_p \in T_p S$ está bem definido e existe uma curva $\gamma : (-\varepsilon, \varepsilon) \rightarrow S$ tal que $\gamma(0) = p$ e $\gamma'(0) = X_p$. Dessa forma, pela Proposição acima, a quantidade $(\nabla_X Y)(p)$ está bem definida. Note que isso nos dá um vetor em $T_p M$. Mas como S é uma subvariedade do tipo espaço, então para todo $p \in S$ vale que $T_p M = T_p S \oplus (T_p S)^\perp$. Podemos decompor então $\nabla_X Y$ na sua parte tangencial e normal, ou seja,

$$\nabla_X Y = (\nabla_X Y)^\top + (\nabla_X Y)^\perp,$$

onde $(\nabla_X Y)^\top$ é a projeção ortogonal de $\nabla_X Y$ em $T_p S$ e $(\nabla_X Y)^\perp$ é a projeção ortogonal de $\nabla_X Y$ em $(T_p S)^\perp$.

Dessa forma, estamos motivados a definirmos duas aplicações: a primeira leva dois campos de vetores sobre S em um campo de vetores normal a S , é chamada de *segunda forma fundamental* de S e denotada por

$$II(X, Y) := (\nabla_X Y)^\perp.$$

A segunda aplicação em questão é relativa a outra parte da nossa decomposição. Vamos denotar $(\nabla_X Y)^\top$ por $\nabla_X^S Y$. Essa notação nos sugere que $\nabla_X^S Y$ é a conexão de Levi-Civita relativa à métrica induzida por g sobre S . De fato isso é verdade, e é o conteúdo da prova da Fórmula de Gauß.

Teorema 5.15 (Fórmula de Gauß). *∇^S como definida acima é a conexão de Levi-Civita da métrica induzida por g em S e se X, Y são campos de vetores sobre S vale que*

$$\nabla_X Y = \nabla_X^S Y + II(X, Y).$$

Uma prova pode ser encontrada em [Liv:Lee97]. A última relação segue imediatamente das definições acima. A prova de que ∇^S é a conexão de Levi-Civita é feita mostrando-se que ∇^S é simétrica e compatível com a métrica induzida. A unicidade da conexão de Levi-Civita nos garante a conclusão.

5.5.2 O Teorema

Agora estamos prontos para enunciarmos e provarmos o Teorema de Hawking. Vamos ao enunciado:

Teorema 5.16. *Seja M um espaço-tempo globalmente hiperbólico satisfazendo:*

- *$Ric(V, V) \geq 0$, para todo V do tipo tempo;*
- *Existe uma superfície de Cauchy S tal que $H_S(p) \geq k > 0$, para todo $p \in S$, onde k é uma constante.*

¹⁵Pode se provar que se ω é uma forma de volume em uma variedade (semi-)riemanniana M (tal forma sempre existe localmente), então $\mathcal{L}_X \omega = (\operatorname{div} X)\omega$, onde $\mathcal{L}_X \omega$ representa a derivada de Lie da forma ω ao longo do campo X . Veja [Liv:Onl83].

Então M é geodesicamente incompleta. Mais precisamente, se $\mu : (-u_0, 0] \rightarrow M$ é uma geodésica direcionada para o futuro tal que $g(\mu', \mu') = 1$, $\mu(0) \in S$ e $\mu'(0) \perp S$, então $-u_0 \geq -3/k$.

A conclusão desse Teorema é que cada observador que intersecte S de maneira ortogonal não pode ter um “tempo de vida” ilimitado inferiormente, e existe um limite inferior para esse “tempo de vida” que é independente do observador e depende só de S . Note que as hipóteses desse Teorema são razoáveis: as hipóteses na causalidade fazem sentido, como já vimos no Capítulo 4; a hipótese $\text{Ric}(V, V) \geq 0$, para todo V do tipo tempo nos diz que a gravidade tem um efeito, em média, atrativo¹⁶, o que é razoável do ponto de vista experimental; finalmente, a última hipótese nos diz que o universo está em expansão, o que é comprovado experimentalmente. Então não estamos em um contexto fisicamente improvável. Vamos então a sua prova.

Demonstração. A idéia da prova é fixar um $-v_0 \in (-u_0, 0]$ arbitrário, mostrarmos que $-v_0 \geq -3/k$ e concluirmos que $-u_0 \geq -3/k$. Mas como fazemos isso? O ponto chave é usar o seguinte Lema, que provaremos depois, pois sua demonstração é bastante técnica e complicada.

Lema 5.17. *Seja M um espaço-tempo globalmente hiperbólico, S uma superfície de Cauchy e $p \in M \setminus S$. Então existe pelo menos uma curva suave, tipo tempo $\lambda : [0, 1] \rightarrow M$ tal que $\lambda(0) = p$, $\lambda(1) \in S$ e $\int_0^1 g(\lambda'(t), \lambda'(t))^{1/2} dt$ = tempo próprio de λ é maior ou igual que o tempo próprio de qualquer outra curva suave do tipo tempo que começa em p e termina em S . Mais ainda, λ é uma geodésica tal que $\lambda'(1) \perp T_{\lambda(1)}S$.*

Escolhendo p como um ponto em $(M \setminus S) \cap D^-(S)$ e normalizando a velocidade da geodésica, temos que:

Corolário 5.18. *Nas condições do Lema, existe uma geodésica direcionada para o futuro e do tipo tempo $\gamma : [-a, 0] \rightarrow M$ tal que $\gamma(-a) = p$, $\gamma(0) \in S$, $\gamma'(0) \perp T_{\gamma(0)}S$ e $g(\gamma', \gamma') = 1$ tal que o tempo próprio de γ , dado por a , é maior ou igual que o tempo próprio de qualquer outra curva suave do tipo tempo em M que parte de p e vai até um ponto de S .*

Denote por p o ponto $\mu(-v_0)$. Como a é o tempo próprio de γ de p até S e v_0 é o tempo próprio de μ de p até S , pela maximalidade de γ basta mostrarmos que $-a \geq -3/k$, pois teremos então que $-u_0 > -v_0 \geq -a \geq -3/k$. Podemos então esquecer de μ e nos concentrarmos em γ . Vamos considerar três variações espertas de γ . Cada uma dela nos dará uma função comprimento de arco, denotada por L_i . Pela maximalidade, teremos que $L_i''(0) \leq 0$. Brincando um pouco com as fórmulas para essas segundas derivadas, obteremos que

$$0 \leq \frac{3}{a} - \int_{-a}^0 \left(\frac{a+u}{a} \right)^2 \text{Ric}(\gamma'(u), \gamma'(u)) du - H_S(\gamma(0)).$$

Usando as hipóteses do Teorema, concluímos que $-a \geq -3/k$, que é o que queríamos provar.

Para construir essas variações, note que $\gamma'(-a)$ é do tipo tempo. Podemos encontrar então $W_i(-a) \in T_pM$, $i = 1, 2, 3$ do tipo espaço tal que $\{\gamma'(-a), W_i(-a)\}$ forma uma base ortonormal de T_pM . Seja $\{\gamma'(u), W_i(u)\}$ o transporte paralelo de $\{\gamma'(-a), W_i(-a)\}$ ao longo de γ . Note que como γ é geodésica então o transporte paralelo de $\gamma'(-a)$ através de γ nos dá $\gamma'(u)$. Como o transporte paralelo é uma isometria, temos que $\{\gamma'(u), W_i(u)\}$ é uma base ortonormal para $T_{\gamma(u)}M$, para todo $u \in [-a, 0]$. Defina $V_i(u) = \frac{a+u}{a} W_i(u)$, $i = 1, 2, 3$, $u \in [-a, 0]$. Note que $V_i(-a) = 0$ e $V_i(0) = W_i(0) \in T_{\gamma(0)}S$, $i = 1, 2, 3$. Isso nos permite definir variações Γ_i de γ tais que suas curvas principais começam em p e terminam em algum ponto de S da seguinte forma: estenda os campos V_i a uma vizinhança aberta U da imagem de γ , de modo que em $U \cap S$ as extensões de V_i sejam tangentes a S . Seja φ_i os fluxos associados a esses campos¹⁷. Agora, vamos deixar a curva γ “andar” um pouco sobre esses fluxos. Essas serão as nossas variações. Para isso, defina $\Gamma_i(u, v) = \varphi_i(v, \gamma(u))$. Note que $\Gamma_i(u, 0) = \varphi_i(0, \gamma(u)) = \gamma(u)$, e além disso $\partial_v \Gamma_i(u, 0) = \partial_v \varphi_i(0, \gamma(u)) = V_i(u)$.

¹⁶Para ver mais sobre isso, veja [Liv:Haw75] e [Liv:Wal84].

¹⁷Lembremos que $\varphi(t, p)$ representa deixar o ponto p “andar” sob a influência do campo durante um intervalo de tempo t .

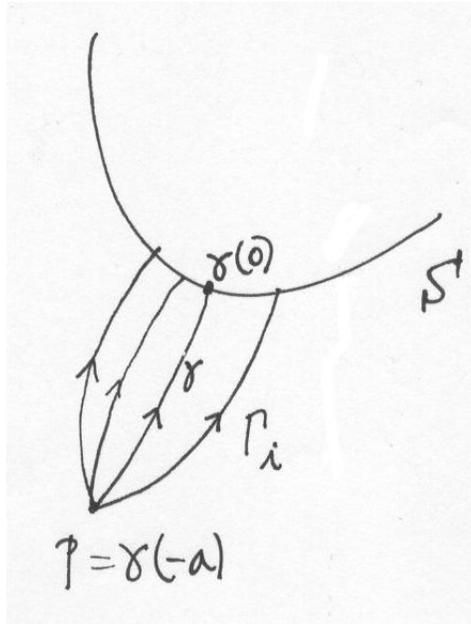


Figura 5.3: Variação Γ_i da geodésica γ .

Finalmente, note que as curva transversais de cada variação em $u = -a$, dadas por $\Gamma_i(-a, v)$, são constantes em p , implicando que $A_i(-a) = 0$, e além disso as curvas transversais $\Gamma_i(0, v)$ são curvas em S , pois as extensões dos campos V_i foram escolhidas de modo a serem tangentes a S . Com isso, temos que as curvas principais $u \mapsto \Gamma(u, v)$ começam em p e terminam em S , que é o que queríamos.

Cada variação Γ_i nos dá origem a uma função comprimento de arco associada L_i . Pela maximalidade de γ , temos que $L_i''(0) \leq 0, i = 1, 2, 3$, e portanto, $\sum_i L_i''(0) \leq 0$, ou seja,

$$\begin{aligned} 0 &\geq \sum_i \int_{-a}^0 -g(R(V_i(u), \gamma'(u))\gamma'(u), V_i(u)) + g(D_\gamma V_i(u), D_\gamma V_i(u))du + g(\gamma'(u), A_i(u)) \Big|_{u=-a}^{u=0} \\ &= \sum_i \int_{-a}^0 -g(R(V_i(u), \gamma'(u))\gamma'(u), V_i(u)) + g(D_\gamma V_i(u), D_\gamma V_i(u))du + g(\gamma'(0), A_i(0)). \end{aligned}$$

Vamos ver quem é cada termo dessa soma separadamente.

- $\sum_i \int_{-a}^0 g(D_\gamma V_i(u), D_\gamma V_i(u))du$: Por definição de V , temos que:

$$D_\gamma V_i(u) = D_\gamma \left(\frac{a+u}{a} W_i \right) (u) = \left(\frac{a+u}{a} \right)' W_i(u) + \frac{a+u}{a} \cancel{D_\gamma W_i(u)}^0 = \frac{1}{a} W_i(u).$$

Portanto:

$$g(D_\gamma V_i(u), D_\gamma V_i(u)) = g\left(\frac{1}{a} W_i(u), \frac{1}{a} W_i(u)\right) = \frac{1}{a^2} g(W_i(u), W_i(u)) = -\frac{1}{a^2}.$$

Concluimos então que

$$\sum_i \int_{-a}^0 g(D_\gamma V_i(u), D_\gamma V_i(u)) du = \sum_i \int_{-a}^0 -\frac{1}{a^2} du = \sum_i -\frac{1}{a} = -\frac{3}{a}.$$

- $\sum_i \int_{-a}^0 -g(R(V_i(u), \gamma'(u))\gamma'(u), V_i(u)) du$: Novamente pela definição de V temos que:

$$\begin{aligned} \sum_i g(R(V_i(u), \gamma'(u))\gamma'(u), V_i(u)) &= \sum_i g\left(R\left(\left(\frac{a+u}{a}\right)W_i(u), \gamma'(u)\right)\gamma'(u), \left(\frac{a+u}{a}\right)W_i(u)\right) = \\ &= \sum_i \left(\frac{a+u}{a}\right)^2 g(R(W_i(u), \gamma'(u))\gamma'(u), W_i(u)) = \\ &= -\left(\frac{a+u}{a}\right)^2 \text{Ric}(\gamma'(u), \gamma'(u)). \end{aligned}$$

Dessa forma, concluímos que

$$\sum_i \int_{-a}^0 -g(R(V_i(u), \gamma'(u))\gamma'(u), V_i(u)) du = \int_{-a}^0 \left(\frac{a+u}{a}\right)^2 \text{Ric}(\gamma'(u), \gamma'(u)) du.$$

- $\sum_i g(A_i(0), \gamma'(0))$: Esse é o termo mais trabalhoso da nossa análise. Primeiramente, note que como $\gamma'(0)$ aponta para o futuro, é tipo tempo e unitário então ele coincide com $N(\gamma(0))$. Dessa forma, como $\{W_i(0), N(\gamma(0))\}$ é uma base ortonormal em $T_{\gamma(0)}M$, com $\{W_i(0)\}$ base ortonormal de $T_{\gamma(0)}S$, podemos estender tal conjunto a campos $\{\bar{W}_i\}$ em uma vizinhança U de $\gamma(0)$ em M tal que para cada $P \in U \cap S$, $\{\bar{W}_i(p)\}$ é base ortonormal de T_pS . Temos então que:

$$\begin{aligned} \sum_i g(\nabla_{\bar{W}_i} \bar{W}_i, N)(\gamma(0)) &= \sum_i g(\nabla_{\bar{W}_i(\gamma(0))} \bar{W}_i, N(\gamma(0))) \stackrel{\nabla_X Y \text{ só depende de } X_p}{=} \\ &= \sum_i g(\nabla_{V_i(0)} \bar{W}_i, N(\gamma(0))) \stackrel{\text{Fórmula de Gauß}}{=} \\ &= \sum_i g(\nabla_{V_i(0)}^S \bar{W}_i + II(V_i(0), \bar{W}_i(\gamma(0)), N(\gamma(0)))) \stackrel{\nabla_{V_i(0)}^S \bar{W}_i \text{ é normal a } N(\gamma(0))}{=} \\ &= \sum_i g(II(V_i(0), \bar{W}_i(\gamma(0)), N(\gamma(0)))) \stackrel{II \text{ só é calculada em vetores nos espaços tangentes a } S}{=} \\ &= \sum_i g(II(V_i(0), V_i(0)), N(\gamma(0))) \stackrel{\text{Definição de } II}{=} \\ &= \sum_i g((\nabla_{V_i(0)} V)^{\perp}, N(\gamma(0))) \stackrel{\text{Estamos projetando } (\nabla_{V_i(0)} V)^{\perp} \text{ sobre um vetor colinear}}{=} \\ &= \sum_i g(\nabla_{V_i(0)} V, \gamma'(0)) \stackrel{\text{Definição de } A_i}{=} \\ &= \sum_i g(A_i(0), \gamma'(0)). \end{aligned}$$

Dessa forma, temos que

$$\begin{aligned} \sum_i g(A_i(0), \gamma'(0)) &= \sum_i g(\nabla_{\bar{W}_i} \bar{W}_i, N)(\gamma(0)) = \\ &= \sum_i g(\nabla_{\bar{W}_i(\gamma(0))} \bar{W}_i, N(\gamma(0))) = \\ &= \sum_i -g(\bar{W}_i(\gamma(0)), \nabla_{\bar{W}_i(\gamma(0))} N). \end{aligned}$$

Lembremos que $\text{div}(X) = \sum_i \varepsilon_i g(\nabla_{E_i} X, E_i)$, onde $\{E_i\}$ é um referencial ortonormal e $\varepsilon_i = g(E_i, E_i)$. Dessa forma, temos que

$$\text{div}(N(\gamma(0))) = \sum_i -g(\nabla_{\bar{W}_i(\gamma(0))} N, \bar{W}_i(\gamma(0))) + g(\nabla_{N(\gamma(0))} N, N(\gamma(0))).$$

Note que $g(\nabla_{N(\gamma(0))} N, N(\gamma(0))) = 0$, pois como $g(N, N)$ é constante igual a 1, temos que

$$Ng(N, N) = 2g(\nabla_N N, N) = 0.$$

Dessa forma, concluímos que

$$H_S(\gamma(0)) = \text{div}(N(\gamma(0))) = - \sum_i g(\nabla_{\bar{W}_i(\gamma(0))} N, \bar{W}_i(\gamma(0))) = \sum_i g(A_i(0), \gamma'(0)).$$

Juntando tudo feito acima, temos que:

$$\begin{aligned} 0 &\geq \sum_i \int_{-a}^0 -g(R(V_i(u), \gamma'(u))\gamma'(u), V_i(u)) + g(D_\gamma V_i(u), D_\gamma V_i(u)) du + g(\gamma'(0), A_i(0)) \\ &= \int_{-a}^0 \left(\frac{a+u}{a}\right)^2 \text{Ric}(\gamma'(u), \gamma'(u)) du - \frac{3}{a} + H_S(\gamma(0)) \geq \\ &\geq -\frac{3}{a} + H_S(\gamma(0)) \geq \\ &\geq -\frac{3}{a} + k. \end{aligned}$$

Portanto, concluímos, depois de algum esforço, que $a \leq 3/k$, que é o desejado. \square

Terminamos então a prova do Teorema de Hawking? Não! Ainda estamos longe disso, pois ainda temos que provar o fatídico Lema 5.17. A prova dele, objetivo da próxima seção, é longa e técnica.

5.5.3 Prova do Lema 5.17

Lembremos a definição feita no Capítulo 4, de vizinhança convexa.

Definição 5.19. Uma vizinhança N de um ponto $p \in M$ é dita *convexa* se é vizinhança normal de todos seus pontos, ou seja, para todo $q \in N$, existe $U_q \in T_q M$ vizinhança convexa e estrelada de zero tal que $\exp_q : U_q \rightarrow N$ é um difeomorfismo.

A partir disso, definimos um conceito um pouco mais forte, a ser usado mais adiante.

Definição 5.20. Uma vizinhança convexa de p é dita uma *região simples* se o seu fecho em M é compacto e está contido em outra vizinhança convexa de p .

Denotamos por U_p o aberto maximal de T_pM tal que $\exp_p : U_p \rightarrow N$ é um difeomorfismo.

Defina uma função $\Omega : N \times N \rightarrow \mathbb{R}$ como $\Omega(p, q) = g(\exp_p^{-1}(q), \exp_p^{-1}(q))$. Tal função é chamada de *função energia*. Note que $\Omega(p, q)$ mede o quadrado do comprimento da única geodésica $\mu : [0, 1] \rightarrow N$ tal que $\mu'(0) = \exp_p^{-1}(q)$. De fato:

$$\int_0^1 g(\mu'(t), \mu'(t)) dt \stackrel{\mu \text{ é geodésica}}{=} \int_0^1 g(\mu'(0), \mu'(0)) dt \stackrel{\text{Definição de } \mu'(0)}{=} g(\exp_p^{-1}(q), \exp_p^{-1}(q)) = \Omega(p, q).$$

Note que $\Omega(p, q) = \Omega(q, p)$, pois a única geodésica de q a p em N é uma reparametrização de μ , e sabemos que comprimento de arcos são invariantes por reparametrizações. Além disso, $\Omega(p, q)$ é estritamente positivo, estritamente negativo ou zero, se μ for uma geodésica do tipo tempo, espaço ou luz, respectivamente. Finalmente, Ω é uma função suave, pois tanto g quanto \exp o são.

Fixado $p \in N$, definimos $\Omega_p : N \rightarrow \mathbb{R}$ como $\Omega_p(q) = \Omega(p, q)$. Para $k \in \mathbb{R}$ seja

$$H_{p,k} = \{q \in N \mid q \neq p \text{ e } \Omega_p(q) = k\}.$$



Figura 5.4: $H_{p,k}$.

Note que $H_{p,k}$ é uma subvariedade de dimensão 3 de N (e portanto de M), pois é a imagem por \exp do conjunto

$$\{\tilde{q} \in T_pM \mid g_p(\tilde{q}, \tilde{q}) = k\},$$

que é uma subvariedade de T_pM .

Denote por $C_T^\pm(p) \subset T_pM$ os cones de tempo futuro (+) e passado (-) em T_pM . Finalmente, seja

$$N_T^\pm(p) = \exp_p(C_T^\pm(p) \cap U_p) \subset N.$$

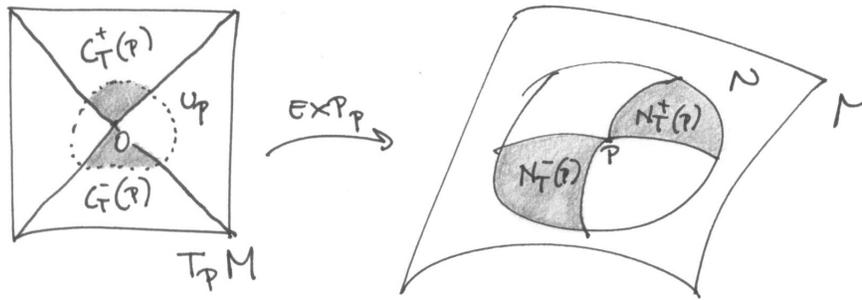


Figura 5.5: $N_T^\pm(p)$.

Note que $N_T^+(p)$ e $N_T^-(p)$ são abertos disjuntos (pois $C_T^\pm(p) \cap U_p$ o são e \exp_p é difeomorfismo). Além disso, $q \in N_T^+(p) \cup N_T^-(p)$ se e somente se μ é uma geodésica do tipo tempo. De fato:

$$\begin{aligned} q \in N_T^+(p) \cup N_T^-(p) &\Leftrightarrow \exists \tilde{q} \in (C_T^+(p) \cup C_T^-(p)) \cap U_p \mid \exp_p(\tilde{q}) = q \\ &\Leftrightarrow \Omega_p(q) = g(\exp_p^{-1}(q), \exp_p^{-1}(q)) = g(\exp_p^{-1}(\exp_p(\tilde{q})), \exp_p^{-1}(\exp_p(\tilde{q}))) = g(\tilde{q}, \tilde{q}) > 0. \end{aligned}$$

Lema 5.21 (de Gauß). *Sejam $p \in M$ e $x \in T_p M$. Se $v_x, w_x \in T_x(T_p M) \cong T_p M$ com v_x co-linear com x , então*

$$g(D \exp_p(x).v_x, D \exp_p(x).w_x) = g(v_x, w_x).$$

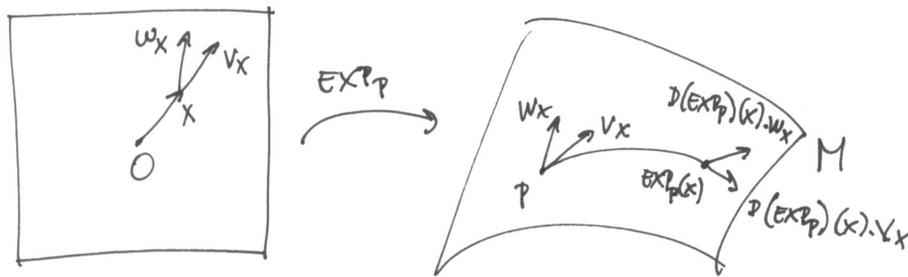


Figura 5.6: Lema de Gauß.

Esse lema nos diz que ao longo de geodésicas radiais partindo de p , $D \exp_p$ preserva a componente tangencial de vetores.

Corolário 5.22. *Se $q \in H_{p,k}$, então a única geodésica em N de p até q é ortogonal a $H_{p,k}$ em q .*

Para uma prova desses resultados, veja [Liv:Onl83].

Corolário 5.23. *Sejam N uma região simples em M , $p \in N$, $q \in N \setminus \{p\}$ e $\mu : [0, 1] \rightarrow N$ a única geodésica em N de p até q . Então $\text{grad } \Omega_p(q) = 2\mu'(1)$.*

Demonstração. Primeiramente, note que $\text{grad } \Omega_p(q) = 0$ se e somente se $p = q$, de modo que $\text{grad } \Omega_p$ não se anula em $N \setminus \{p\}$. Fazendo $k = \Omega_p(q)$, temos que $q \in H_{p,k}$. Pelo Lema de Gauß, temos que $\mu'(1)$ é um vetor não-nulo que é ortogonal a $H_{p,k}$. Mas $\text{grad } \Omega_p(q)$ também é um vetor não-nulo ortogonal a

$H_{p,k}$. Como $T_q(H_{p,k})$ tem dimensão 3 temos que $\text{grad } \Omega_p(q)$ e $\mu'(1)$ são linearmente dependentes, ou seja, $\text{grad } \Omega_p(q) = a\mu'(1)$, onde a é um parâmetro que, a princípio, depende continuamente de q . Mostremos que a é constante igual a 2. Para isso, calculemos, para $t \in [0, 1]$,

$$\begin{aligned} \Omega_p(\mu(t)) &= g(\exp_p^{-1}(\mu(t)), \exp_p^{-1}(\mu(t))) = \\ &= g(t \exp_p^{-1}(q), t \exp_p^{-1}(q)) = \\ &= t^2 k. \end{aligned}$$

Dessa forma,

$$\frac{d}{dt} \Omega_p(\mu(t)) = g(\text{grad } \Omega_p(\mu(t)), \mu'(t)) = 2kt, \quad t \in [0, 1].$$

Fazendo $t = 1$ temos que

$$2k = g(\text{grad } \Omega_p(q), \mu'(q)) = ak.$$

Se $k \neq 0$, ou seja, μ não é uma geodésica do tipo luz, temos que $a = 2$. Por outro lado, se μ é do tipo luz também temos que $a = 2$, por continuidade. \square

Finalmente, podemos provar a generalização há muito prometida do Teorema 3.15:

Teorema 5.24. *Seja N uma região simples com $p, q \in N$. Suponha que existe $\alpha : [0, 1] \rightarrow N$ suave, direcionada para o futuro e tipo tempo tal que $\alpha(0) = p$ e $\alpha(1) = q$. Então a única geodésica em N de p a q também é tipo tempo e direcionada para o futuro.*

Demonstração. Primeiramente, seja $\beta : [0, 1] \rightarrow T_p M$ o “levantamento” de α , ou seja, $\beta = \exp_p^{-1} \circ \alpha$. Temos então que

$$\beta'(0) = D \exp_p^{-1}(\alpha(0)) \cdot \alpha'(0) = D \exp_p(\exp_p^{-1}(p)) \cdot \alpha'(0) = D \exp_p(0) \cdot \alpha'(0) = \alpha'(0) \in C_T^+(p),$$

pois α é tipo tempo e direcionada para o futuro. Como $C_T^+(p)$ é aberto, existe $\varepsilon > 0$ tal que $\beta([0, \varepsilon]) \subset C_T^+(p)$. Dessa forma, $\alpha([0, \varepsilon]) \subset N_T^+(p)$, pela definição de $N_T^+(p)$. Mostremos que uma vez em $N_T^+(p)$, α não pode mais escapar, ou seja, $\alpha(1) \in N_T^+(p)$. Note que isso é suficiente para provar o Teorema, pois vimos acima que $q \in N_T^+(p)$ ou $q \in N_T^-(p)$ se e somente se a única geodésica de p até q em N é tipo tempo e direcionada para o futuro ou passado, respectivamente. Provemos então tal afirmação.

Suponha que $\alpha(1) \notin N_T^+(p)$. Pela definição de β , temos que $\beta(1) \notin C_T^+(p)$. Mas para sair de $C_T^+(p)$, β deve passar ou por $0 \in T_p M$ ou pelo cone de luz futuro em $T_p M$ (pois a união do cone de luz futuro com $\{0\}$ divide $T_p M$ em duas componentes conexas, e uma delas é $C_T^+(p)$). A afirmação acima segue do famoso Teorema da Alfândega, da Topologia Geral). Dessa forma, existe $t_0 \in (\varepsilon, 1]$ tal que $\Omega_p(\alpha(t_0)) = 0$. Note que não podemos ter $\beta(t_0) = 0$. Caso contrário, teríamos que

$$\alpha(t_0) = \exp_p(\beta(t_0)) = \exp_p(0) = \alpha(0),$$

contradição, pois M é suposto como sendo estavalmente causal. Dessa forma, $\beta(t_0)$ está no cone de luz futuro em $T_p M$, de modo que se $q = \alpha(t_0)$ e $\mu : [0, 1] \rightarrow N$ é a única geodésica em N de p a q , então $\mu'(1)$ é tipo luz e direcionada para o futuro. Como $\text{grad } \Omega_p(q) = 2\mu'(1)$, temos que $\text{grad } \Omega_p(q)$ também é tipo luz e direcionado para o futuro. Lembremos que $\alpha'(t_0)$ é tipo tempo e aponta para o futuro. Dessa forma, $g(\text{grad } \Omega_p(q), \alpha'(t_0)) > 0$, de modo que

$$\frac{d}{dt} \Big|_{t=t_0} \Omega_p(\alpha(t)) = D\Omega_p(\alpha(t_0)) \cdot \alpha'(t_0) = g(\text{grad } \Omega_p(\alpha(t_0)), \alpha'(t_0)) > 0.$$

Dessa forma, $\Omega_p(\alpha(t))$ é crescente em torno de t_0 . Como $\Omega_p(\alpha(t_0)) = 0$ temos que $\Omega_p(\alpha(t))$ deve ser negativo “um pouco antes”¹⁸ de t_0 . Assim, quando α atinge o bordo de $N_T^+(p)$ (ou seja, quando β atinge

¹⁸O termo “um pouco antes” não é nada rigoroso, mas está bem entendido o que ele quer dizer. A prova desse Teorema é suficientemente enrolada pra nos prendermos com detalhes do tipo “fixado ε existe δ tal que...”

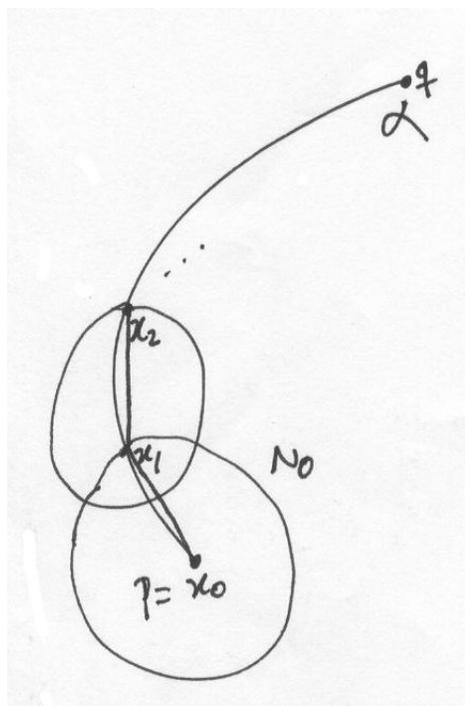


Figura 5.7: Ilustração da prova da Proposição 4.15.

o cone de luz superior em $T_p M$, $\Omega_p(\alpha(t))$ deve ser negativo um pouco antes de t_0 .

Considere o conjunto $\{t \in [0, 1] \mid \alpha(t) \in N_T^+(p)\}$. Note que por hipótese ele é não-vazio e limitado superiormente. Seja então t_1 o seu supremo. Note que $\alpha(t_1)$ deve estar no bordo superior de $N_T^+(p)$, de modo que $\Omega_p(\alpha(t))$ deve ser negativo um pouco antes de t_1 . Então temos que

$$\Omega_p(\alpha(t)) = \Omega(p, \alpha(t)) = g(\exp_p^{-1}(\alpha(t)), \exp_p^{-1}(\alpha(t))) = g(\beta(t), \beta(t)) < 0$$

um pouco antes de t_1 . Dessa forma, $\beta(t) \notin C_T^+(p)$ um pouco antes de t_1 , o que implica que $\alpha(t) \notin N_T^+(p)$ um pouco antes de t_1 . contradição com a definição de supremo. \square

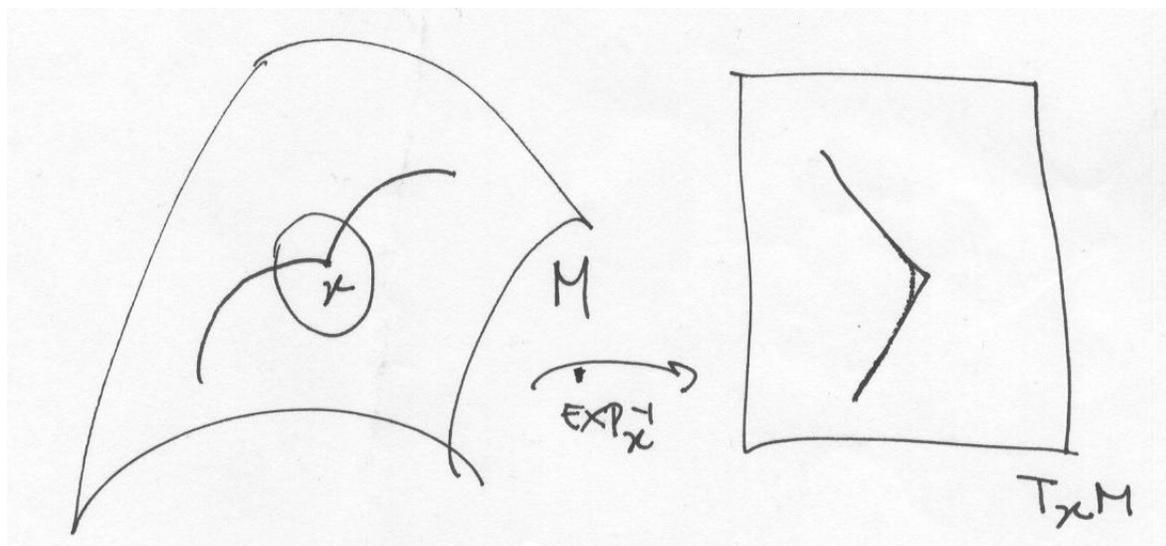
Podemos então provar o Teorema 3.15:

Corolário 5.25. *Sejam $p, q \in M$. Então $p \ll q$ se e somente se existe uma curva suave $\alpha : [a, b] \rightarrow M$ direcionada para o futuro e tipo tempo tal que $\alpha(a) = p, \alpha(b) = q$.*

Demonstração. Se $p \ll q$ então a existência de tal curva é imediata. Basta tomar a semi-reta que vai de p até q . Agora, se existe a curva α , como M é uma região simples para M em torno de p , segue do Teorema 5.24 que a única geodésica em M de p até q é tipo tempo e direcionada para o futuro. Mas tal geodésica, como já vimos no Teorema 3.17, é dada exatamente pela semi-reta que vai de p até q . Segue então que $p \ll q$. \square

Podemos também provar a Proposição 4.15, que fala sobre a equivalência entre usar viagens e curvas direcionadas para o futuro do tipo tempo na definição da relação \ll em um espaço-tempo arbitrário:

Demonstração. (da Proposição 4.15) Suponha que exista tal curva α . Cubra sua imagem com finitas regiões simples N_i .


 Figura 5.8: Suavização do levantamento da viagem em $T_x M$.

Seja $x_0 = p \in N_0$ e x_1 o ponto final futuro da componente conexa de $\text{Imagem}(\alpha) \cap \overline{N_0}$ que contém x_0 . Então, pelo Teorema 5.24 temos que $x_0 x_1$ ¹⁹ é direcionada para o futuro e do tipo tempo. Caso $x_1 = q$, o resultado está provado. Caso contrário, $x_1 \notin \overline{N_0}$ e $x_1 \in N_1$. Defina x_2 de maneira análoga ao x_1 e teremos que $x_1 x_2$ é direcionada para o futuro e do tipo tempo. Repita esse processo n vezes até que $x_n = q$. Tal número de passos é de fato finito, pois $\text{Imagem}(\alpha)$ é um conjunto compacto.

Suponha agora que $p \ll q$. Devemos provar que os vértices da viagem que liga p até q podem ser “suavizados” de maneira a obtermos uma curva do tipo tempo e direcionada para o futuro. É fácil de nos convencermos que isso é verdade com um desenho, porém construir tal curva é bastante enfadonho. Seja x um vértice da viagem e considere coordenadas normais \exp_x^{-1} em uma região simples em torno de x . Suavize o levantamento da viagem em $T_x M$ de modo a obter uma curva direcionada para o futuro e tipo tempo, como na Figura 5.8. Trazendo a curva suavizada de volta a M obtermos a suavização desejada.

□

Muitas vezes precisamos lidar com curvas que não são nem do tipo tempo nem do tipo luz, mas que podem assumir os dois aspectos ao longo de sua trajetória. Fazemos então a seguinte definição:

Definição 5.26. Uma curva $\alpha : I \rightarrow M$ é dita *causal* se $\alpha'(t)$ é do tipo tempo ou luz, para $t \in I$.

A partir daí é imediato adivinharmos qual é a definição de uma geodésica causal, viagem causal, etc... É claro que uma geodésica causal ou é do tipo tempo, ou é do tipo luz. A mesma demonstração do Teorema 5.24 prova o seguinte resultado:

Teorema 5.27. Seja N uma região simples com $p, q \in N$. Suponha que existe $\alpha : [0, 1] \rightarrow N$, curva ou viagem causal direcionada para o futuro que não seja uma geodésica do tipo luz e que satisfaz $\alpha(0) = p$ e $\alpha(1) = q$. Então a única geodésica em N de p a q é tipo tempo e direcionada para o futuro.

Mas para que estamos fazendo tudo isso? Vamos usar esses resultados preliminares para exibir uma característica vital de espaços-tempo globalmente hiperbólicos, que muitas vezes é tomado como sendo

¹⁹Essa notação representa a única geodésica em N_0 partindo de x_0 que vai até x_1 .

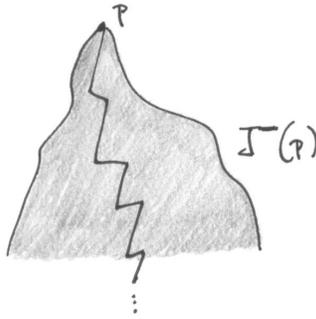


Figura 5.9: $J^-(p)$.

sua definição (veja [Liv:Haw75]).

Dizemos que p precede casualmente q (e denotamos isso por $p \leq q$) se e somente se existe uma viagem causal (ou de maneira equivalente, uma curva causal) de p até q . O futuro causal de qualquer ponto p de M é denotado por $J^+(p)$ e definido como

$$J^+(p) = \{q \in M \mid p \leq q\}.$$

O passado causal de p , denotado por $J^-(p)$, é definido de maneira análoga. Dado qualquer $S \subset M$, definimos $J^\pm(S) = \cup_{p \in S} J^\pm(p)$.

Proposição 5.28. *Valem as seguintes propriedades para os conjuntos J e I :*

- Para todo $p \in M$ vale que $\overline{I^+(p)} = \{q \in M \mid I^+(q) \subset I^+(p)\}$;
- $p \leq q$ implica que $I^+(q) \subset I^+(p)$;
- $\overline{J^+(p)} = \overline{I^+(p)}$;
- $J^+(p)$ não é necessariamente fechado.

Como essas propriedades foram colocadas aqui somente como uma curiosidade, não as provaremos. Para uma prova veja [Liv:Wal84].

Um passo importante na prova do Lema 5.17 é o seguinte resultado:

Lema 5.29. *Seja S uma superfície de Cauchy em M . Então para todo p no interior de $D^+(S)$, o conjunto $J^-(p) \cap J^+(S)$ é compacto. Analogamente, para todo p no interior de $D^-(S)$, o conjunto $J^+(p) \cap J^-(S)$ é compacto*

Mas o que esse resultado quer dizer? Por que essa compactidade é tão importante para a prova do Lema 5.17? Primeiramente vamos interpretar quem são os conjuntos J :

$$\begin{aligned} J^-(p) &= \{q \in M \mid q \leq p\} = \text{pontos de } M \text{ que precedem causalmente } p \\ &= \text{pontos de } M \text{ para os quais existe uma curva ou viagem causal} \\ &\quad \text{partindo do ponto em questão e chegando até } p. \end{aligned}$$

$$\begin{aligned} J^+(S) &= \cup_{p \in S} J^+(p) = \text{pontos de } M \text{ que são precedidos causalmente por algum ponto de } S \\ &= \text{pontos de } M \text{ para os quais existe uma curva ou viagem causal} \\ &\quad \text{partindo de algum ponto de } S \text{ e chegando no ponto em questão.} \end{aligned}$$

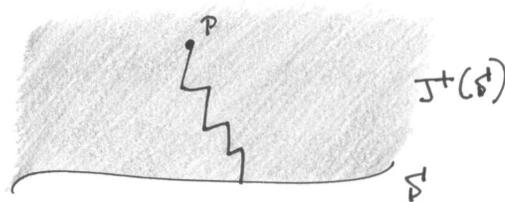


Figura 5.10: $J^+(S)$.

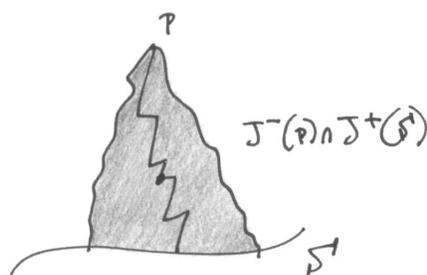


Figura 5.11: $J^-(p) \cap J^+(S)$.

E portanto:

$J^-(p) \cap J^+(S)$ = pontos de M que precedem causalmente p e são precedidos causalmente por algum ponto de S
 = pontos de M para os quais existe uma curva ou viagem causal partindo do ponto em questão e chegando até p e existe uma curva ou viagem causal partindo de algum ponto de S e chegando no ponto em questão.

O Lema 5.17 tem um espírito análogo ao famoso Teorema do cálculo, que afirma que em um intervalo da forma $[a, b]$ toda função contínua atinge um máximo e um mínimo. Note que o importante nessa afirmação é o fato do intervalo $[a, b]$ ser compacto. De fato, esse resultado se generaliza para qualquer espaço topológico compacto. Isso nos sugere abordar o Lema 5.17 da seguinte forma: considere a coleção C de todas as curvas causais que ligam qualquer ponto de S com um ponto fixado p em, digamos $D^-(S)$. Defina uma métrica em C de modo que esse espaço métrico seja compacto e a função real que a cada curva associa seu comprimento seja contínua. Então essa função deve atingir um máximo em alguma curva de C , que tentaríamos provar que é uma geodésica. Mas infelizmente as coisas não são tão fáceis quanto parece. Não há uma maneira razoável de definirmos uma estrutura de espaço métrico compacto sobre o conjunto de todas as curvas causais suaves de p até S , pois o limite de curvas (ou mais geralmente, funções) suaves não é necessariamente suave. Precisamos estender, então, a noção de "causal" para curvas que sejam contínuas mas não necessariamente suaves. Nesse conjunto maior de curvas encontraremos uma métrica natural que torna esse espaço métrico compacto, e essa compacidade depende de maneira crucial da compacidade de $J^-(p) \cap J^+(S)$. É claro que devemos também estender a noção de comprimento de curvas para esse conjunto mais geral, já que tais curvas não necessariamente possuem vetor tangente em todos os pontos. Ainda temos que ter esperança de que depois de tudo isso a função comprimento de arco seja contínua na nossa métrica. Infelizmente isso não é verdade, mas ainda assim teremos que ela é semicontínua superiormente, de modo que tal função em geral não possui

um mínimo sobre um espaço compacto, mas sempre possui um máximo, e isso resolve nosso problema. Vamos aos detalhes. Começemos provando o Lema 5.29.

Demonstração. (do Lema 5.29) Como essa prova é longa e fácil de se perder nela, colocarei as justificativas para algumas afirmações mais simples como notas de rodapé, para facilitar a leitura.

Primeiramente, note que se S é uma superfície de Cauchy em M , então $\partial D^+(S) = \partial D^-(S) = S$, e dessa forma, se $p \notin S$ então $p \in D^\pm(S)$.

A idéia é provar o resultado por contradição. Então assumiremos que p está no interior de $D^+(S)$ mas que $K = J^-(p) \cap J^+(S)$ não seja compacto. A partir disso vamos construir uma viagem intextendível para o passado que passe por um ponto de $D^+(S)$ mas que não encontre S , uma contradição com a definição de $D^+(S)$.

Portanto, apesar de estarmos supondo que K não é compacto, ele é paracompacto (pois M é metrizável, e K sendo um subconjunto de M é também metrizável e logo paracompacto²⁰) e Lindelöf (pois como M é segundo-contável e K sendo um subconjunto de M é também segundo-contável. Prova-se (e não é tão difícil. Veja [Liv:Mun00]) que todo espaço segundo-contável é Lindelöf.

Dessa forma, dada $\{U_\alpha\}$ cobertura aberta de K , existe uma outra cobertura aberta enumerável $\{V_n\}$ de K tal que todo V_n está contido em algum U_α e $\{V_n\}$ é localmente finita. Como as regiões simples formam uma base para a topologia de M , temos que existe uma cobertura enumerável $\{N_i\}$ de K por regiões simples de M que é localmente finita mas não tem subcobertura finita.

Para cada $i \in \mathbb{N}$, tome $a_i \in N_i \cap K$, com $a_i \neq a_j$, se $i \neq j$. Note que como $\{N_i\}$ é localmente finita, então a sequência (a_i) não tem ponto de acumulação²¹.

Agora vamos começar a construção da viagem de maneira indutiva. Seja $x_0 = p$. Como $\{N_i\}$ é uma cobertura de K , temos que existe i_0 de modo que $x_0 \in N_{i_0}$. Tome $y_0 \in N_{i_0} \cap I^+(x_0) \cap \text{interior}(D^+(S))$ ²². Repare que existem viagens causais em M de cada a_i até x_0 , pois $a_i \in K$. Como $\overline{N_{i_0}}$ é compacto, temos que infinitos a_i estão fora de N_{i_0} ²³. Portanto, infinitas dessas viagens causais de a_i até x_0 devem passar por ∂N_{i_0} .

Como ∂N_{i_0} é compacto (pois $\overline{N_{i_0}}$ o é), esses pontos de interseção têm um ponto de acumulação $z_0 \in \partial N_{i_0}$. Como cada um desses pontos é ligado a x_0 por uma viagem causal para o futuro temos que $\Omega(\text{esses pontos}, x_0) \geq 0$ ²⁴. Pela continuidade de Ω , temos que $\Omega(z_0, x_0) \geq 0$, e portanto, z_0 está em $J^-(x_0)$, ou seja $z_0 \geq x_0$. Como $x_0 \ll y_0$ (por definição de y_0), temos que $z_0 \ll y_0$. Logo, existe curva tipo tempo e direcionada para o futuro de z_0 até y_0 . Como z_0 e y_0 estão em alguma região simples, o Teorema 5.24 nos diz que a única geodésica em N_{i_0} de z_0 até y_0 (denotada aqui por $z_0 y_0$) é direcionada para o futuro e tipo tempo.

Note que $z_0 \notin N_{i_0}$, mas $z_0 \in J^-(x_0)$. Temos então dois casos: $z_0 \notin J^+(S)$ ou $z_0 \in J^+(S)$. Se $z_0 \notin J^+(S)$, então nenhuma viagem causal inextendível para o passado passando por z_0 encontra S . Uma tal viagem

²⁰É um Teorema (difícil!) de Topologia Geral que todo espaço metrizável é paracompacto. Para mais informação, veja [Liv:Mun00]. No caso de espaços não separáveis (ou seja, aqueles que não contém um enumerável denso) é necessário o uso do Axioma da Escolha. Não que isso seja algo muito grave ou me incomode muito, mas como variedades são sempre separáveis não precisamos recorrer a essa entidade mística aqui.

²¹De fato, se a fosse ponto de acumulação de (a_i) teríamos que se V é uma vizinhança de a então V contém infinitos a_i , ou seja, $V \cap N_i \neq \emptyset$, para infinitos índices i , uma contradição com o fato da cobertura $\{N_i\}$ ser localmente finita.

²²Note que esse conjunto é não-vazio, pois N_{i_0} é uma vizinhança de x_0 de $I^+(x_0)$ é um aberto do qual x_0 pertence ao seu bordo. Dessa forma, $N_{i_0} \cap I^+(x_0)$ é aberto, diferente de vazios e x_0 pertence ao seu fecho. Como o interior de $D^+(S)$ é uma vizinhança de x_0 , concluímos que $N_{i_0} \cap I^+(x_0) \cap \text{interior}(D^+(S))$ é não vazio.

²³Caso contrário, (a_i) teria um ponto de acumulação, o que vimos que não é possível.

²⁴Repare que faz sentido falar em Ω para elementos do bordo de N_{i_0} , pois podemos tomar uma região simples um pouco maior que N_{i_0} contendo-a, e nessa nova região simples a função Ω está bem definida.

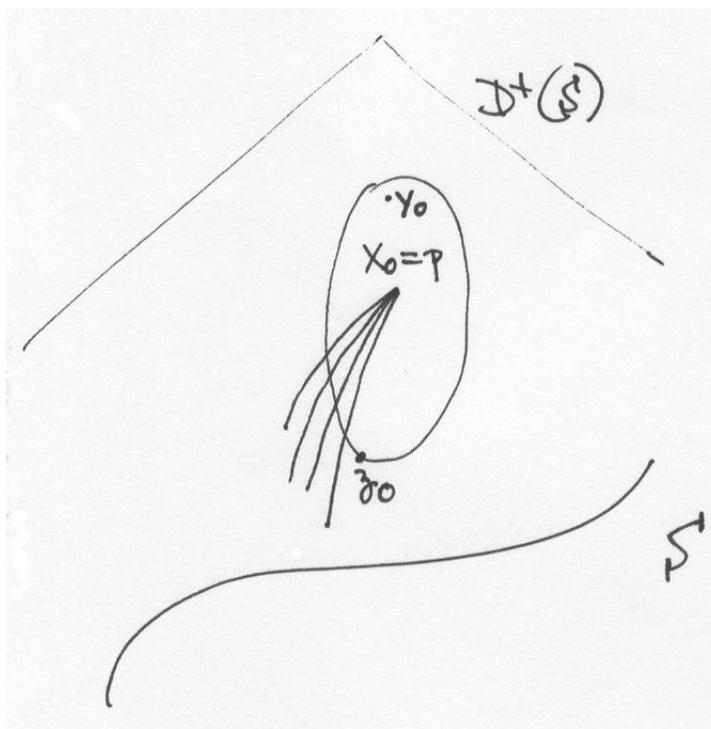


Figura 5.12: Prova do Lema 5.29.

“colada” com $z_0 y_0$ é uma viagem tipo tempo inextendível para o passado através de um ponto de $D^+(S)$, que é y_0 , mas que não encontra S , uma contradição²⁵. Assuma então que $z_0 \in J^+(S)$. Como $z_0 \geq x_0$ (ou seja, $z_0 \in J^-(x_0 = p)$), temos que $z_0 \in K$. Portanto, existe N_{i_1} (diferente de N_{i_0} , já que $z_0 \notin N_{i_0}$) tal que $z_0 \in N_{i_1}$. Tome $x_1, y_1 \in z_0 y_0 \cap N_{i_1}$, de modo que $z_0 \ll x_1 \ll y_1 \ll y_0$.

Agora, repare que $I^-(x_1)$ é uma vizinhança aberta de z_0 . Portanto, infinitas das viagens causais de a_i até x_0 entram em $I^-(x_1)$ ²⁶. Tome um ponto em casa viagem dessas que esteja em $I^-(x_1)$. Então temos que $a_i \ll$ esse ponto $\ll x_1$ ²⁷, e portanto, $a_i \in I^-(x_1)$.

Se todos, exceto finitos, desses a_i estão em uma única viagem até x_0 então eles se acumulariam, o que sabemos que não ocorre. Dessa forma, infinitos desses a_i em infinitas viagens distintas, estão em $I^-(x_1)$ e infinitos desses a_i estão fora de N_{i_1} (caso contrário eles também se acumulariam, como já vimos no caso de a_0). Então novamente existe um $z_1 \in \partial N_{i_1}$ que é ponto de acumulação das interseções das viagens tipo tempo de a_i até x_0 com ∂N_{i_1} . Encontramos novamente então que $z_1 \leq x_1$ (mesmo argumento anterior). Como $x_1 \ll y_1$, temos que $z_1 \ll y_1$, e novamente pelo Teorema 5.24 a geodésica $z_1 y_1$ é tipo tempo e direcionada para o futuro. Como $z_1 \notin N_{i_1}$, temos que existe N_{i_2} tal que $z_1 \in N_{i_2}$. Como antes escolhemos $x_2, y_2 \in z_1 y_1 \cap N_{i_2}$ tal que $z_1 \ll x_2 \ll y_2 \ll y_1$.

Procedendo dessa forma, obtemos uma sequência (y_i) com $\dots \ll y_2 \ll y_1 \ll y_0 \in D^+(S)$. Considere a viagem cujos segmentos são $\dots, y_2 y_1, y_1 y_0$. Essa viagem tem ponto final futuro $y_0 \in D^+(S)$. Note que

²⁵Repare aqui a importância do ponto y_0 . Sem ele teríamos relações somente do tipo \geq , e a definição de $D^+(S)$ envolve a relação \ll , que só foi obtida graças ao tal ponto y_0 .

²⁶Pois infinitos dos pontos de interseção de tais viagens com ∂N_{i_0} estão em $I^-(x_1)$, já que eles convergem para um elemento de tal conjunto.

²⁷A primeira “desigualdade” segue do fato de a_i e tal ponto estarem na mesma viagem e a segunda segue do fato desse ponto estar em $I^-(x_1)$.

ela é também inextendível para o passado, pois se q fosse um ponto final passado, seria um ponto de acumulação dos y_i . Como cada $y_k \in N_{i_k}$, isso é uma contradição com o fato de $\{N_{i_k}\}$ ser localmente finita. Finalmente, mostremos que essa viagem não passa por S . Suponha então que ela passe por S . Então existe $s \in S$ que está na imagem dessa viagem. Então algum y_i está em $I^-(S)$, ou seja, $y_i \ll s$. Mas sabemos que existe algum a_j tal que $a_j \ll x_i \ll y_i$. Dessa forma, $a_j \ll s$. Mas como $a_j \in J^+(S)$ (pois todo a_j está em K), temos que existe algum $s_0 \in S$ tal que $s_0 \leq a_j$. Portanto, $s_0 \leq a_j \leq s$, e logo, $s_0 \leq s \in S$, uma contradição, pois S é acronal. \square

Provado esse Lema, vamos generalizar alguns conceitos para curvas que não sejam suaves. Primeiramente, dizemos que uma curva contínua $\alpha : (a, b) \rightarrow M$ é *direcionada pra o futuro e causal* se para cada t_0 em (a, b) existe um $\varepsilon > 0$ e uma vizinhança convexa U de $\alpha(t_0)$ tal que $\alpha(t_0 - \varepsilon, t_0 + \varepsilon) \subset U$ e tal que para todos t_1 e t_2 satisfazendo $t_0 - \varepsilon < t_1 < t_2 < t_0 + \varepsilon$ existe uma curva diferenciável, direcionada para o futuro e causal, contida em U , que liga $\alpha(t_1)$ a $\alpha(t_2)$. Lógico que se a curva for definida em um intervalo da forma $[a, b)$ não usamos $t_0 - \varepsilon$ no ponto a , mas somente t_0 . O caso do intervalo fechado no outro extremo é análogo. Dessa forma, de agora em diante o termo “curva direcionada para o futuro e causal” será usado nesse contexto mais geral, e explicitaremos que a curva é suave quando necessário. Claramente no caso de α ser uma curva diferenciável a nova definição implica na antiga.

Considere uma superfície de Cauchy $S \subset M$ e um ponto p no interior de $D^-(S)$. Para simplificar a notação, denote o conjunto $J^+(p) \cap J^-(S)$ por K . Lembre que provamos que K é compacto. Denote por $C_K(p, S)$ o conjunto de todas as curvas em K de p até S que são direcionadas para o futuro e causais. Defina uma base para a topologia de $C_K(p, S)$ como $C_R(p, Q)$, onde $R \subset K$ e $Q \subset S$ são abertos relativos. Repare que tal coleção é de fato uma base, pois se $\alpha \in C_R(p, Q)$ e $\alpha \in C_{R'}(p, Q')$, então $\alpha \in C_{R''}(p, Q'')$, onde $R'' = R \cap R'$ e $Q'' = Q \cap Q'$. A idéia dessa topologia é generalizar a topologia C^0 para o nosso caso particular. Munido dessa topologia, o conjunto $C_K(p, S)$ é compacto:

Lema 5.30. $C_K(p, S)$ é compacto.

Demonstração. Seja $\{\alpha_i\}$ uma seqüência em $C_K(p, S)$. Devemos mostrar que existe uma curva $\alpha \in C_K(p, S)$ que é ponto de acumulação de $\{\alpha_i\}$. Seja $\{q_i\}$ em $K \cap S$ a seqüência que representa os pontos finais futuros das curvas α_i . Como $K \cap S$ é compacto, seja q um ponto de acumulação de $\{q_i\}$. Escolha $\{\alpha_i\}$ uma sub-sequência²⁸ de modo que seus pontos finais futuros convirjam para q .

Cubra K com finitas vizinhanças locais de causalidade. Seja N_0 uma dessas vizinhanças tal que $p = p_0 \in N_0$. Como as curvas α_i têm interseção com ∂N_0 , que é um conjunto compacto, escolha outra sub-sequência $\{\alpha_i\}$ cujos pontos de interseção com N_0 convirjam para, digamos $p_1 \in N_1 \cap \partial N_0$. Repetindo o argumento para N_2, N_3, \dots temos uma coleção finita de pontos $p = p_0, p_1, \dots, p_k = q$ tal que $p_{i-1} \leq p_i$, com cada par consecutivo $p_{i-1}, p_i \in N_{i-1}$. Queremos construir curvas causais de p_{i-1} até p_i de modo que a curva resultante seja ponto de acumulação de $\{\alpha_i\}$.

Para isso, suponha que N_0 foi escolhida pequena o suficiente de modo que as superfícies $x^0 = \text{constante}$, em um sistema de coordenadas normais com origem em p_0 sejam do tipo espaço. Seja ε o valor da coordenada x^0 de p_1 . As interseções das curvas α_i com a superfície $x^0 = \varepsilon/2$ tem ponto de acumulação, denotado por r_0 , e uma sub-sequência $\{\alpha_i\}$ cujas interseções com $x^0 = \varepsilon/2$ convergem para r_0 . Trocando p_0 por p_1 obtemos $r_{1,1}$ entre p_1 e p_2 e uma outra sub-sequência de $\{\alpha_i\}$. Repita esse procedimento para $p_3, \dots, p_k = q$ e obtenha $r_{i,1}, i = 2, \dots, k-1$ entre p_i e p_{i+1} .

Voltando a N_0 e repetindo a construção com $x^0 = \varepsilon/4$ e com $x^0 = 3/4\varepsilon$, obtemos $r_{0,01}$ e $r_{0,11}$ respectivamente, e assim sucessivamente. Tal construção nos dá um ponto r_{β} , para cada número real entre 0 e k cuja expansão binária β é finita. Por construção temos que $r_{\beta_1} \leq r_{\beta_2}$, se $\beta_1 \leq \beta_2$ e cada r_{β} é ponto de acumulação de alguma sub-sequência de $\{\alpha_i\}$. Dessa forma, a curva $\overline{\cup_{\beta} \{r_{\beta}\}}$ é ponto de acumulação de

²⁸Ao longo dessa prova vamos fazer esse abuso de notação frequentemente, pois teremos que tomar diversas sub-sub-...-sub-sequências.

$\{\alpha_i\}$. Vejamos que ela está em $C_K(p, S)$. Claramente ela começa em p e vai até S (mais precisamente, $K \cap S$). Além disso, ela está contida em K , por construção. Basta provarmos que é uma curva causal. Tome β_0 e seja N_i tal que $r_{\beta_0} \in N_i$. Como vale que $r_{\beta_1} \leq r_{\beta_2}$ se $\beta_1 \leq \beta_2$, temos que se β_1 e β_2 são tais que $r_{\beta_1}, r_{\beta_2} \in N_i$ então $\Omega(r_{\beta_1}, r_{\beta_2}) \geq 0$. Por continuidade, $\Omega(x, y) \geq 0$, se $x, y \in (\cup_{\beta} \{r_{\beta}\}) \cap N_i$. \square

Vamos então definir o funcional de comprimento de arco em $C_K(p, S)$. O problema agora é que dada uma curva em tal espaço, não necessariamente o seu vetor tangente está bem definido. Para contornar esse problema, vamos aproximar uma curva em $C_K(p, S)$ por viagens do tipo tempo, calcular o comprimento dessas curvas e tomar o limite, da mesma forma que fazemos para calcular o comprimento de curvas em \mathbb{R}^n . Note que essa aproximação por viagens será análoga à aproximação de curvas por poligonais. Vamos então ao trabalho.

Dada $\alpha \in C_K(p, S)$ (considere-a parametrizada sobre o intervalo $[0, 1]$, por simplicidade), temos que $p = \alpha(0)$ e seja $q = \alpha(1)$. Temos que $p \leq q$. Seja $\xi = \{x_i\}$ uma sequência finita de pontos em α com $x_0 = p$ e $x_k = q$ = último ponto da sequência = q de modo que cada par x_i, x_{i+1} esteja contido em uma região simples N_i que também contenha a porção de α que vai de x_i até x_{i+1} ²⁹. Temos então que $x_i \leq x_{i+1}$ e o único segmento de geodésica em N_i ligando x_i até x_{i+1} , aqui denotado por $x_i x_{i+1}$ é direcionado para o futuro e causal (ambas as afirmações seguem da definição de curva direcionada para o futuro e causal). Denote por α_{ξ} a viagem direcionada para o futuro e causal $x_0 x_1 \cup x_1 x_2 \cup \dots \cup x_{k-1} x_k$. Seja Ξ o conjunto de todas tais sequências ξ . As notações $\xi \subset \xi'$ e $\xi \cup \xi' = \xi''$ têm o significado óbvio. Vale então o seguinte:

- $\xi \subset \xi' \Rightarrow L(\xi') \leq L(\xi)$ ³⁰: Primeiramente, sejam $\xi = \{p = x_0, x_1, \dots, x_k = q\}$ e $\xi' = \{p = x_0, \tilde{x}, x_1, \dots, x_k = q\}$. Claramente temos que $L(\xi \setminus \{x_0\}) = L(\xi' \setminus \{x_0, \tilde{x}\})$. Provemos então que $L(\{x_0, x_1\}) \geq L(\{x_0, \tilde{x}, x_1\})$. Por definição, $x_0 x_1$ é a única geodésica de x_0 até x_1 em N , e $x_0 \tilde{x} \cup \tilde{x} x_1$ é uma viagem direcionada para o futuro e causal de x_0 até x_1 em N . Vamos parar e pensar um pouco: Em \mathbb{R}^n a menor distância entre dois pontos é uma geodésica, e qualquer outra curva que ligue esses dois pontos tem comprimento estritamente maior. Mas espere! Provamos um resultado semelhante a esse para \mathcal{M} , o Teorema 3.17. Como essas questões são locais e temos em nossas mãos as coordenadas normais, imaginamos que em uma região simples vale um resultado análogo. De fato:

Lema 5.31. *Sejam N uma região simples em M , $p, q \in N$ e $\mu : [0, 1] \rightarrow N$ a única geodésica de p a q em N . Suponha que μ é direcionada para o futuro e causal. Então se α é outra viagem causal em N de p a q , então $L(\mu) > L(\alpha)$.*

Em outras palavras, esse Lema nos diz que, pelo menos localmente, geodésicas causais maximizam a distância entre dois pontos. Vamos adiar um pouco a sua demonstração e continuar aonde tínhamos parado. Pela maneira como ξ foi construída, podemos usar o Lema acima e concluímos que $L(\{x_0, x_1\}) \geq L(\{x_0, \tilde{x}, x_1\})$ ³¹. Procedendo por indução, temos que $L(\xi') \leq L(\xi)$.

- $L(\xi'') \leq \min\{L(\xi), L(\xi')\}$, onde $\xi'' = \xi \cup \xi'$: Como $\xi \subset \xi''$, temos que $L(\xi'') \leq L(\xi)$. Analogamente, $L(\xi'') \leq L(\xi')$. Daí segue o resultado.

Defina então

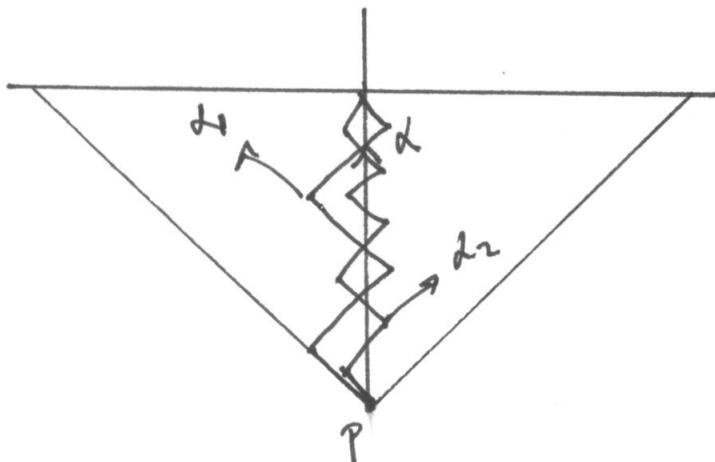
$$L(\alpha) = \inf_{\xi \in \Xi} L(\alpha_{\xi}).$$

Note que essa quantidade está bem definida, pois $L(\alpha_{\xi}) \geq 0$. Infelizmente essa função não é contínua sequer no espaço-tempo mais simples que temos, que é \mathcal{M} :

²⁹Para se convencer de que tal região simples existe basta lembrar da definição que demos para uma curva direcionada para o futuro e causal.

³⁰Essa notação significa o comprimento de arco da geodésica por partes cujos segmentos têm como extremos os pontos de ξ .

³¹Note que pode haver igualdade, se o ponto \tilde{x} estiver sobre a geodésica $x_0 x_1$.

Figura 5.13: L não é contínua.

Repare que as curvas α_n convergem para α em $C_K(p, S)$, mas cada $L(\alpha_n) = 0$ enquanto que $L(\alpha) > 0$. Note que isso não é inconsistente com nossa definição de L : nela aproximamos uma curva α por viagens cujos vértices estão sobre a curva, e entre dois vértices adjacentes a viagem é uma geodésica, exatamente da mesma forma que aproximamos uma curva por poligonais nela inscrita para calcularmos seu comprimento de arco em \mathbb{R}^n . A figura acima não aproxima a curva α dessa forma. Porém, temos a sorte de L ser semi-contínua superiormente. Com essa definição, uma curva em $C_K(p, S)$ é automaticamente retificável, de modo que satisfaz uma condição de Lipschitz-continuidade e, portanto, é diferenciável em quase todo ponto, pelo Teorema de Rademacher³². Dessa forma, podemos calcular também $L(\alpha)$ pela sua fórmula usual, apenas observando que existem pontos onde α falha em ser diferenciável. Antes de provar que L é semi-contínuo superiormente, vamos introduzir alguns conceitos.

Definição 5.32. Um subconjunto Q de M é dito *causalmente convexo* se Q não intersecta nenhuma viagem em um conjunto desconexo.

Definição 5.33. Uma *vizinhança local de causalidade* é um aberto causalmente convexo cujo fecho está contido em uma região simples.

Teorema 5.34. Se M é estavelmente causal, então em todo ponto existe uma base local de vizinhanças locais de causalidade.

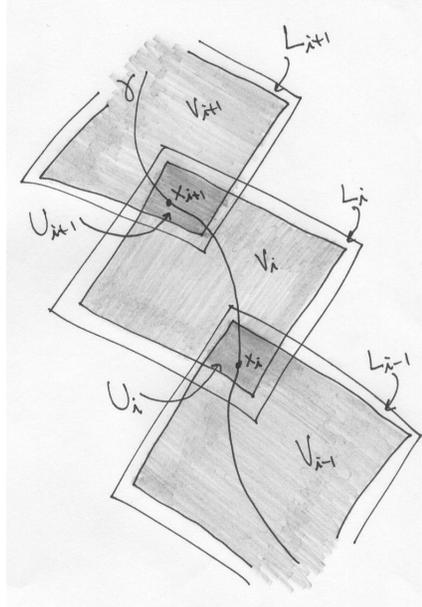
A prova desse Teorema pode ser encontrada em [Liv:Haw75].

Lema 5.35. L é semi-contínuo superiormente em $C_K(p, S)$.

Demonstração. Durante essa demonstração vamos construir uns conjuntos que podem ser meio confusos. É bom se guiar pela Figura 5.14.

Devemos mostrar que $L^{-1}(-\infty, a)$ é aberto em $C_K(p, S)$, para qualquer $a \in \mathbb{R}$. Para isso, tome uma curva causal $\alpha \in C_K(p, S)$ satisfazendo $L(\alpha) < a$. Sendo mais preciso, devemos mostrar que existe U , vizinhança de α em $C_K(p, S)$ tal que toda curva $\beta \in U$ também satisfaz $L(\beta) < a$. Seja $b = L(\alpha) < a$. Escolha $\xi \in \Xi$ tal que $L(\alpha_\xi) < b + (a - b)/2$. Tome os x_i perto o suficiente de modo que cada par x_i, x_{i+1} esteja contido em uma vizinhança local de causalidade L_i , com $\bar{L}_i \subset N_i$ região simples e além disso $L_i \cap L_j \neq \emptyset$ se e somente se $j = i \pm 1$.

³²Hans Adolph Rademacher, Wandsbeck 03/04/1892 - Havford 07/02/1969.


 Figura 5.14: Conjuntos U_i , V_i e L_i .

Como o comprimento de geodésicas em N_i é uma função contínua dos seus pontos finais (pois Ω é suave), temos que podemos escolher uma vizinhança local de causalidade U_i de cada x_i (com $U_i \subset L_i$, $U_{i+1} \subset L_i$, $i = 0, 1, \dots, k-1$), pequena o suficiente para o comprimento de qualquer geodésica de um ponto de U_i até um ponto de U_{i+1} seja diferente de $L(x_i x_{i+1})$ por no máximo $(b-a)/2k$. Seja

$$V_i = \bigcup_{y \in U_i, z \in U_{i+1}} I^+(y) \cap I^-(z).$$

Note que V_i é aberto. Além disso, $V_i \subset L_i$. De fato, todo ponto de V_i está na trajetória de uma viagem de $y \in U_i$ até $z \in U_{i+1}$. Como $U_i, U_{i+1} \subset L_i$ e L_i é vizinhança local de causalidade, tal viagem deve estar contida em L_i e, portanto, $V_i \subset L_i$.

Defina $Q = V_{k-1} \cap S$ e $R = \cup_i V_i$. Mostremos que $C_R(p, Q)$ é a vizinhança desejada de α . Para isso, tome $\beta \in C_R(p, Q)$. Então β passa por todos os V_i consecutivamente. Mais ainda, β passa por cada U_i . Isso segue da convexidade causal de U_i .

Dessa forma, β contém pontos x'_0, x'_1, \dots, x'_k , com $x'_i \in U_i$, de modo que a viagem causal $\gamma = x'_0 x'_1 \cup \dots \cup x'_{k-1} x'_k$ satisfaz

$$\begin{aligned} L(\gamma) - \left(b + \frac{1}{2}(a-b)\right) &< k \frac{a-b}{2k} \Rightarrow \\ L(\gamma) &< \frac{a-b}{2} + b + \frac{a-b}{2} = a. \end{aligned}$$

□

Vamos então provar o Lema 5.31.

Demonstração. (do Lema 5.31) Se μ é uma geodésica do tipo luz, então não há nada a ser provado³³. Suponha então que μ é do tipo tempo. Tome $r \in N$ um ponto na extensão passada de μ em p . Sejam (x^0, x^1, x^2, x^3) coordenadas normais em N e em

$$N_T^+(r) = \{(x^0, x^1, x^2, x^3 \in N \mid x^0 > ((x^1)^2 + (x^2)^2 + (x^3)^2)^{1/2}\}$$

³³Isso será justificado mais abaixo.

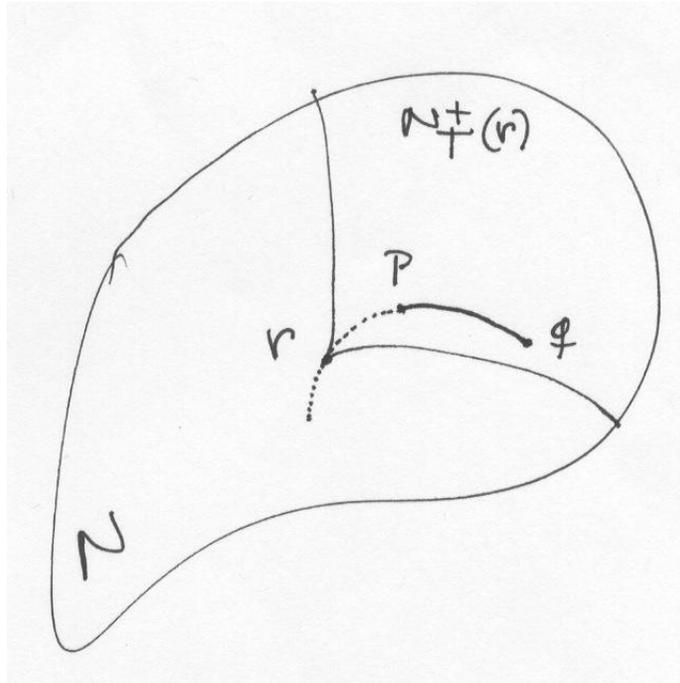


Figura 5.15: Prova do Lema 5.31.

introduza novas coordenadas dadas por

$$X^0 = ((x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2)^{1/2}$$

$$X^i = x^i/x^0, i = 1, 2, 3.$$

Repare que $X^i = \text{constante}$ são geodésicas radiais partindo de r que são ortogonais às superfícies $X^0 = \text{constante}$, pelo Lema de Gauß. Dessa forma, as superfícies $X^0 = \text{constante}$ são do tipo espaço e, nas novas coordenadas a métrica é dada por

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & G_{11} & G_{12} & G_{13} \\ 0 & G_{21} & G_{22} & G_{23} \\ 0 & G_{31} & G_{32} & G_{33} \end{pmatrix},$$

onde G_{ij} é uma matriz negativa-definida. Portanto, dada uma curva causal α direcionada pra o futuro e tipo tempo que passa por p e q temos que, usando X^0 como parâmetro:

$$\begin{aligned} L(\alpha) &= \int_{X_p^0}^{X_q^0} G_{ij} \left(\frac{dX^i}{dX^0} \frac{dX^j}{dX^0} \right)^{1/2} dX^0 = \\ &= \int_{X_p^0}^{X_q^0} G_{ij} \left(1 + \sum_{i,j=1}^3 \frac{dX^i}{dX^0} \frac{dX^j}{dX^0} \right)^{1/2} dX^0 \leq \\ &\leq \int_{X_p^0}^{X_q^0} dX^0 = \\ &= L(\mu), \end{aligned}$$

pois o somatório dentro da segunda integral é máximo se e somente se os X^i são zero, que é o caso da geodésica ligando p até q , nessas coordenadas. \square

Justifiquemos agora a primeira afirmativa nessa prova. Para isso, precisaremos de alguns resultados.

Proposição 5.36. *Seja N uma região simples. Suponha que $p, q, r \in \bar{N}$ são tais que pq e qr são ambas causais e direcionadas para o futuro, de modo que tenham um vetor tangente distintos em q se ambas são tipo luz ou existe α uma curva ou viagem tipo tempo em \bar{N} de p até r . Então pr é direcionada para o futuro e tipo tempo.*

Demonstração. Note que se existe tal curva γ , então o resultado está provado, como consequência do Teorema 5.27. Suponha então a outra hipótese. Note que o fato de pq e qr terem vetores tangentes distintos em q é crucial para nosso resultado, pois caso isso não acontecesse teríamos que $pr = pq \cup qr$ é do tipo luz, o que não é o que queremos provar. Suponha então que pr não seja do tipo tempo. Então é do tipo luz. Portanto, pela unicidade de geodésicas em N temos que $pr = pq \cup qr$, uma contradição com o fato dos vetores tangentes serem distintos. \square

Proposição 5.37. *Se vale qualquer uma das duas afirmativas abaixo*

$$\begin{aligned} p \ll q \text{ e } q \leq r \\ p \leq q \text{ e } q \ll r \end{aligned}$$

então $p \ll r$.

Demonstração. Como a imagem de uma curva é um conjunto compacto, basta cobri-la com finitas regiões simples e usar a Proposição anterior repetidas vezes. \square

Proposição 5.38. *Seja α uma geodésica do tipo luz de p até q e β uma geodésica de q até r . Então ou $p \ll r$ ou $\alpha \cup \beta$ é uma geodésica tipo luz de p até r .*

Demonstração. Se $\alpha \cup \beta$ não é uma geodésica, então os vetores tangentes a α e β em q são diferentes. Pelas proposições anteriores, se x em α e y em β são suficiente próximos porém distintos de q , então existe uma geodésica do tipo tempo de x até y . Dessa forma, temos que $p \leq x \ll y \leq q$, o que implica em $p \ll q$. \square

Agora podemos justificar o fato do Lema 5.31 ser trivial se μ é uma geodésica do tipo luz. Seja α uma viagem causal de p até q em N . Se α é uma geodésica, então ela deve coincidir com μ . Suponha então que α não é geodésica. Dessa forma, existe um ponto x em α na qual a curva falha em ser geodésica e px e xq são geodésicas do tipo luz. Pela Proposição acima, ou $p \ll q$, o que sabemos não ser verdade, ou $px \cup xq$ é uma única geodésica do tipo luz de p até q , contradizendo o fato da existência de tal ponto x . Assim, a única curva causal de p até q é a geodésica μ e o resultado é trivial por vacuidade.

Nosso trabalho está quase terminado. Resta juntarmos as peças e concluirmos o Lema 5.17.

Demonstração. (do Lema 5.17) Como o funcional L é semi-contínuo superiormente em $C_K(p, S)$ e tal conjunto é compacto, temos que existe uma curva λ tal que $L(\lambda) \geq L(\alpha)$, para toda curva $\alpha \in C_K(p, S)$. Provemos que λ é uma geodésica do tipo tempo que corta S ortogonalmente. Isso segue do Teorema de Avez-Seifert:

Teorema 5.39 (Avez-Seifert). *Seja (M, g) um espaço-tempo globalmente hiperbólico.*

- Se $p, q \in M$ são tais que $p \leq q$, então existe uma geodésica causal ligando p a q cujo comprimento é dado por $d(p, q)^{34}$.
- A função d definida como acima é contínua e finita em todo M .

³⁴ $d(p, q)$ é definido como $\sup_{\alpha} L(\alpha)$, onde o supermo é tomado sobre todas as curvas causais direcionadas para o futuro que ligam p a q . Caso não exista nenhuma tal curva, definimos $d(p, q)$ como sendo zero.

A prova desse Teorema pode ser encontrada em [Liv:Bee96] e [Liv:Vic10]. Aqui vamos usá-lo livremente para nosso propósito.

A primeira afirmativa do Teorema nos diz que λ é uma geodésica do tipo tempo. Para provar que ela corta S ortogonalmente, podemos proceder por contradição e construirmos uma variação de α que tem comprimento maior, contradizendo sua maximalidade. Isso segue diretamente da fórmula da primeira variação do funcional L . \square

5.6 Considerações Finais

Problemas matemáticos em Relatividade Geral podem ser divididos grosseiramente em 4 categorias:

- Construção de soluções especiais para as Equações de Einstein
- Propriedades matemáticas das equações de restrições
- Causalidade e problemas globais
- Problema de Cauchy em Relatividade Geral

Nessa dissertação tratamos um pouco sobre causalidade e problemas globais. No Capítulo 1 demos uma perspectiva história da noção de espaço e tempo. O Capítulo 2 trouxe uma introdução à Relatividade Restrita. O Capítulo 3 traz uma visão do Capítulo anterior de um ponto de vista mais rigoroso matematicamente, e prova resultados sobre o Espaço de Minkowski, espaço-tempo da Relatividade Restrita. No Capítulo 4 vimos uma breve introdução à Relatividade Geral, dando enfoque nos aspectos causais dos espaços-tempo. Finalmente, o Capítulo 5 trouxe a prova de um Teorema de Stephen Hawking sobre a incompletude geodésica de espaços-tempo globalmente hiperbólicos. Tais espaços-tempo são determinísticos no sentido que a Equação de Einstein é um problema bem posto.

Existe uma outra versão do Teorema de Hawking que supõe que a superfície de Cauchy S é compacta e não impõe condições fortes na cronologia do espaço-tempo. Sua prova é bem mais simples do que a do Teorema aqui apresentado. Nessa mesma direção existem os Teoremas de Hawking-Penrose, que tratam de singularidades cuja natureza é diferente da aqui apresentada.

Para mais referências de resultados clássicos sobre a existência de singularidades, veja [Liv:Pen05], [Liv:Wal84] e [Liv:Haw98]. Para uma referência de problemas em abertos interessantes nessa área veja [Art:Sen06].

A *hipótese do censor cósmico* é uma conjectura em Relatividade Geral que afirma que toda singularidade está escondida em um horizonte de eventos, ou seja, um “contorno” no espaço-tempo tal que eventos “dentro” desse contorno não podem afetar eventos “fora” do contorno. Uma singularidade que não satisfaz tal condição é dita *nua*. Pode-se construir exemplos de espaço-tempo nos quais isso falha, mas são exemplos artificiais e sem significado físico aparente. Isso está intimamente relacionado com a noção de hiperbolicidade global. De fato, pode-se provar que a definição aqui dada é equivalente ao fato de $J^+(p) \cap J^-(q)$ ser compacto, para todos $p, q \in M$. Caso isso não aconteça, pode-se provar também que o espaço-tempo em questão apresenta uma singularidade nua. Veja [Liv:Vic10] para mais detalhes.

Existem diversos resultados também sobre o Problema de Cauchy em Relatividade Geral, existência de soluções para condições iniciais suficientemente regulares, estabilidade do espaço de Minkowski, entre outros. Veja [Art:Rod06]. Pode-se também relacionar tais resultados com existências de singularidades.

Outro resultado interessante é o Teorema da Massa Positiva, provado por Schoen³⁵ e Yau³⁶. Uma curiosidade é que tal resultado foi usado na resolução do Problema de Yamabe³⁷ [Art:Lee87], e na demonstração da Desigualdade de Penrose Riemanniana [Art:Bra01], ilustrando que a Física também pode ajudar a Matemática a resolver alguns problemas.

³⁵Richard Schoen, Fort Recovery 32/10/1950

³⁶Shing-Tung Yau, Shantou 04/04/1949

³⁷Hidehiko Yamabe, Ashiya 22/08/1923 - Evanston 20/11/1960

Referências Bibliográficas

[] Artigos:

- [Art:Bae99] Baeßler, S, et. al. - *Improved Test of the Equivalence Principle for Gravitational Self-Energy*, Physical Review Letters, **86** (18) (1999), 3585–3588.
- [Art:Ber03] Bernal, A. N., Sánchez, M. - *On Smooth Cauchy Hypersurfaces and Geroch's Splitting Theorem*, Communications in Mathematical Physics, **243** (2003), 461–470.
- [Art:Ber05] Bernal, A. N., Sánchez, M. - *Smoothness of Time Functions and the Metric Splitting of Globally Hyperbolic Spacetimes*, Communications in Mathematical Physics, **257** (2005), 43–50.
- [Art:Bra01] Bray, H. - *Proof of the Riemannian Penrose Inequality Using the Positive Mass Theorem*, Journal of Differential Geometry, **59** (2001) 177–267.
- [Art:Bre77] Brecher, K. - *Is the Speed of Light Independent of the Velocity of the Source?*, Physical Review Letters, **39** (1977), 1051–1054.
- [Art:Car89] Carrière, Y. - *Autour de la conjecture de L. Markus sur les variétés affines*, Inventiones Mathematicae **95** (1989), 615–628.
- [Art:Car01] Carroll, Sean M. - *The Cosmological Constant*, Living Reviews in Relativity **4** (2001).
- [Art:Cal62] Calabi, E., Markus, L. - *Relativistic space forms*, Annals of Mathematics, **75** (1962), 63–76.
- [Art:CERN11] CERN - *Measurement of the neutrino velocity with the ICARUS detector at the CNGS beam*, <http://www.http://arxiv.org/abs/1203.3433>.
- [Art:Che63] Chern, S. S. - *Pseudo-Riemannian geometry and the Gauss Bonnet formula*, Anais da Academia Brasileira de Ciência, **35** (1963), 17–26.
- [Art:Eas91] Easwar, N., MacIntire, D. A. - *Study of the effect of relativistic time dilation on cosmic ray muon flux - An undergraduate modern physics experiment*, American Journal of Physics **59** (1991), 589–592.
- [Art:Ein05] Einstein, Albert - *Zur Elektrodynamik bewegter Körper*, Annalen der Physik **17** (1905), 891–921.
- [Art:Ein15] Einstein, Albert - *Die Feldgleichungen der Gravitation*, Königlich Preussische Akademie der Wissenschaften (1915), 844–847.
- [Art:Ein16] Einstein, Albert - *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen der Physik **7** (1916), 769–822.
- [Art:Fal07] Falciano, F. T. - *Cinemática Relativística: paradoxo dos gêmeos*, Revista Brasileira de Ensino de Física **29** (2007), 19–24.
- [Art:Ger70] Geroch, R. - *Domain of Dependence*, Journal of Mathematical Physics, **11** (1970), 437–450.

- [Art:Göd49] Gödel, Kurt - *An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation*, *Reviews of Modern Physics* **21** (1949), 447–450.
- [Art:Haw66A] Hawking, Stephen - *The Occurrence of Singularities in Cosmology*, *Proceedings of the Royal Society of London A* **294** (1966), 511–521.
- [Art:Haw66B] Hawking, Stephen - *The Occurrence of Singularities in Cosmology II*, *Proceedings of the Royal Society of London A* **295** (1966), 490–493.
- [Art:Haw67] Hawking, Stephen - *The Occurrence of Singularities in Cosmology III - Causality and Singularities*, *Proceedings of the Royal Society of London A* **300** (1967), 187–201.
- [Art:Hub29] Hubble, Edwin - *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*, *Proceedings of the National Academy of Science of The United States of America*, **15**(3) (1929), 168–173.
- [Art:Ker63] Kerr, R. P. - *Gravitational field of a spinning mass as an example of algebraically special metrics*. *Physical Review Letters* **11** (1963) 237.
- [Art:Kli96] Klingler, B. - *Complétude des variétés lorentziennes à courbure constante*, *Mathematische Annalen*, **306** (1996), 353–370.
- [Art:Lee87] Lee, John M., Parker, Thomas H. - *The Yamabe Problem*, *Bulletin of the American Mathematical Society*, **17** (1987), 37–91.
- [Art:Mar73] Marsden, J. E. - *On completeness of pseudo-Riemannian manifolds*, *Indiana University Mathematics Journal*, **22** (1973), 1065–1066.
- [Art:New65] Newman, E., et al. - *Metric of a Rotating, Charged Mass*, *Journal of Mathematical Physics*, **6** (1965), 918–919.
- [Art:Nor18] Nordström, G. - *On the Energy of the Gravitational Field in Einstein's Theory*, *Verhandl. Koninkl. Ned. Akad. Wetenschap., Afdel. Natuurk., Amsterdam* **26** (1918), 1201–1208.
- [Art:Ray55] Raychaudhuri, Amal K. - *Relativistic Cosmology I*, *Physical Review*, **98** (1955), 1123–1126.
- [Art:Rei16] Reissner, H. - *Über die Eigengravitation des elektrischen Feldes nach der Einsteinschen Theorie*, *Annalen der Physik*, **50** (1916), 106–120.
- [Art:Rie1854] Riemann, Bernhard - *Über die Hypothesen welche der Geometrie zu Grunde liegen*, (1854).
- [Art:Rod06] Rodnianski, I. - *The Cauchy Problem in General Relativity*, *Proceedings of the International Congress of Mathematics* (2006).
- [Art:Rom94] Romero, A., Sánchez, M. - *New properties and examples of incomplete Lorentzian tori*, *Journal of Mathematical Physics*, **35** (1994), 1992–1997.
- [Art:Rom95] Romero, A., Sánchez, M. - *Completeness of compact Lorentz manifolds admitting a timelike conformal-Killing vector field*, *Proceedings of the American Mathematical Society*, **123** (1995), 2831–2833.
- [Art:Sch16] Schwarzschild, K. - *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie*, *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, **1** (1916), 189–196.
- [Art:Sen06] Senovilla, J. M. M. - *Singularity Theorems in General Relativity: Achievements and Open Questions*, <http://arXiv.org/abs/physics/0605007v1>.
- [Art:Tau87] Taubes, Clifford H. - *Gauge theory on asymptotically periodic 4-manifolds*, *Journal of Differential Geometry*, **25**(3) (1987), 363–430.

[Art:Whi36] Whitney, H. - *Differentiable Manifolds*, The Annals of Mathematics, **37**(3) (1936), 645–680.

[Art:Zee64] Zeeman, Eric Christopher - *Causality implies the Lorentz group*, Journal of Mathematical Physics **5** (1964), 490–493.

Livros:

[Liv:Ast08] “Astronomica”, H. F. Ullmann, 2008.

[Liv:Bee96] Beem, John K., Ehrlich, Paul E., Easley, Kevin L. - “Global Lorentzian Geometry”, CRC Press, 1996.

[Liv:Cod84] Coddington, Earl A., Levinson, Norman - “Theory of Ordinary Differential Equations”, Krieger, 1984.

[Liv:Euc09] Euclides - “Os Elementos”, Editora Unesp, 2009.

[Liv:Fey08] Feynman, Richard P., Leighton, Robert B., Sands, Matthew - “Lições de Física de Feynman - Edição Definitiva”, 2008.

[Liv:Hat01] Hatcher, Allen - “Algebraic Topology”, Cambridge University Press, 2001.

[Liv:Haw75] Hawking, Stephen W., Ellis, G. F. R. - “The Large Scale Structure of Space-Time”, Cambridge University Press, 1975.

[Liv:Haw98] Hawking, Stephen W. - “A Brief History of Time”, Bantam, 1998.

[Liv:Jam10a] Jammer, Max - “Conceitos de Espaço”, Contraponto, 2010.

[Liv:Jam10b] Jammer, Max - “Concepts of Mass in Classical and Modern Physics”, Dover Publications, 2010.

[Liv:Mal07] Asselmeyer-Maluga, T., Brans, Carl H. - “Exotic Smoothness and Physics: Differential Topology and Space-time Models”, World Scientific Publishing, 2007.

[Liv:Man08] do Carmo, Manfredo Perdigão - “Geometria Riemanniana”, IMPA, 2008.

[Liv:Mar09] Levi, Mark - “The Mathematical Mechanic: Using Physical Reasoning to Solve Problems”, Princeton University Press, 2009.

[Liv:Mun00] Munkres, James - “Topology”, Prentice-Hall, 2000.

[Liv:Kel79] Lord Kelvin (Sir William Thomson), Tait, P. G. - *Treatise on Natural Philosophy*, Cambridge University Press, 1879.

[Liv:Lee97] Lee, John M. - “Riemannian Manifolds: An Introduction to Curvature”, Springer-Verlag, 1997.

[Liv:Lor63] Lorentz, H. A, et. al. - *The Principle of Relativity*, Dover Publications, 1963.

[Liv:MTW73] Misner, C., Thorne, K. S., Wheeler, J. A. - *Gravitation*, W. H. Freeman, 1973.

[Liv:Nab88] Naber, Gregory L. - “Spacetime and singularities - An Introduction”, Cambridge University Press, 1988.

[Liv:Nab92] Naber, Gregory L. - “The Geometry of Minkowski Spacetime - An Introduction to the Mathematics of the Special Theory of Relativity”, Springer-Verlag, 1992.

[Liv:Nai64] Naimak, M. A. - “Linear Representation of the Lorentz Group”, Pergamon, 1964.

- [Liv:Onl83] O'Neill, Barret - "Semi-Riemannian Geometry With Applications to Relativity", Academic Press, 1983.
- [Liv:Pen05] Penrose, Roger - "The Road to Reality", Vintage, 2005.
- [Liv:Pen87] Penrose, Roger - "Techniques of Differential Topology in Relativity", Society for Industrial and Applied Mathematics, 1987.
- [Liv:Sch09] Schutz, Bernard - "A First Course in General Relativity", Cambridge University Press, 2009.
- [Liv:Spi99] Spivak, Michael - "A Comprehensive Introduction to Differential Geometry, Vols. I-V", Publish or Perish, 1999.
- [Liv:Spi10] Spivak, Michael - "Physics for Mathematicians: Mechanics I", Publish or Perish, 2010.
- [Liv:Ste09] Stephani, H., et. al. - "Exact Solutions of Einstein's Field Equations", Cambridge University Press, 2009.
- [Liv:Syn72] Synge, John L. - "Relativity: The Special Theory", North-Holland, 1972.
- [Liv:Syn60] Synge, John L. - "Relativity: The General Theory", North-Holland, 1960.
- [Liv:Tun07] Tung, K. K. - "Topics in Mathematical Modelling", Princeton University Press, 2007.
- [Liv:Vic10] Victoria, Miguel A. J., Caja, Miguel S. - "An Introduction to Lorentzian Geometry and its Applications", XVI Escola de Geometria Diferencial, 2010.
- [Liv:Wal84] Wald, Robert M. - "General Relativity", The University of Chicago Press, 1984.
- [Liv:Wal92] Wald, Robert M. - "Space, Time and Gravity: The Theory of the Big Bang and Black Holes", The University of Chicago Press, 1992.
- [Liv:Wei09] Weyl, H. - "The Concept of a Riemann Surface", Dover Publications, 2009.
- [Liv:Whe92] Wheeler, John A., Taylor, Edwin F. - "Spacetime Physics", W. H. Freeman - 1992.
- [Liv:Woo07] Woodhouse, N. M. J. - "Special Relativity", Springer-Verlag, 2007.
- [Liv:Woo06] Woodhouse, N. M. J. - "General Relativity", Springer-Verlag, 2006.
- [Liv:Zee10] Zee, A. - "Quantum Field Theory in a Nutshell", Princeton University Press, 2010.

Obras clássicas:

- [Cls:Ari] Aristóteles - "Physica".
- [Cls:Gal1632] Galilei, Galileu - "Dialogo sopra i due massimi sistemi del mondo", 1632.
- [Cls:Kep1609] Kepler, Johannes - "Astronomia Nova", 1609.
- [Cls:Kep1619] Kepler, Johannes - "Harmonices Mundi", 1619.
- [Cls:Max1862] Maxwell, James Clerk - "On Physical Lines of Force", 1861-1862.
- [Cls:New1687] Newton, Isaac - "Philosophiae Naturalis Principia Mathematica", 1687.

Internet:

- [Int:Wik] Wikipedia, <http://www.wikipedia.org>
- [Int:Bae] John Baez' Usenet Physics FAQ, <http://math.ucr.edu/home/baez/physics/>