UNIVERSIDADE ESTADUAL DE MARINGÁ

CENTRO DE CIÊNCIAS EXATAS

DEPARTAMENTO DE MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

(mestrado)

Azuaite Aramis Schneider

MÉTRICAS INTRÍNSECAS INVARIANTES À ESQUERDA EM GRUPOS DE LIE

Maringá - PR
2016

Universidade Estadual de Maringá
Departamento de Matemática

Azuaite Aramis Schneider

MÉTRICAS INTRÍNSECAS INVARIANTES À ESQUERDA EM GRUPOS DE LIE

Dissertação de mestrado apresentada ao Departamento de Matemática como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática.

Orientador: Ryuichi Fukuoka

Este exemplar corresponde à versão final da dissertação defendida pelo aluno Azuaite Aramis Schneider, e orientada pelo Prof. Dr. Ryuichi Fukuoka

Maringá
2016

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Central - UEM, Maringá, PR, Brasil)

S359m

Schneider, Azuaite Aramis
Métricas intrínsecas invariáveis à esquerda em Grupos de Lie / Azuaite Aramis Schneider -- Maringá, 2016.
58 f.

Orientador: Prof. Dr. Ryuichi Fukuoka.
Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciência Exatas, Departamento de Matemática, Programa de Pós-Graduação em Matemática, 2016.

1. Grupos de Lie. 2. Métricas intrínsecas. 3. Métricas de Carnot-Carathéodory. 4. Métricas de Finsler. 5. Geometria subriemanniana. I. Fukuoka, Ryuichi, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Departamento de Matemática. Programa de Pós-Graduação em Matemática. IV. Título.

CDD 21.ed. 512.55
AHS-002855

AZUAITE ARAMIS SCHNEIDER

MÉTRICAS INTRÍNSECAS INVARIANTES À ESQUERDA EM GRUPOS DE LIE

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Departamento de Matemática, Centro de Ciências Exatas da Universidade Estadual de Maringá, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática tendo a Comissão Julgadora composta pelos membros:

COMISSÃO JULGADORA:

Ryuichi Fukuoka

Prof. Dr. Ryuichi Fukuoka
DMA/Universidade Estadual de Maringá (Presidente)

Prof. Dr. Paulo Régis Caron Ruffino
Universidade Estadual de Campinas

Marcos A. Verdi

Prof. Dr. Marcos André Verdi
DMA/Universidade Estadual de Maringá

Aprovada em 08 de março de 2016.
Local de defesa: Auditório do DMA, Bloco F67, campus da Universidade Estadual de Maringá.

Dedico esse trabalho aos meus pais, aos meus amigos e a todos os que amam essa arte chamada Matemática.

Agradecimentos

Agradeço,

ao Prof. Ryuichi a sua disposição, dedicação e paciência durante esses dois anos de muita aprendizagem.

aos colegas de trabalho mais próximos: Hugo, Djeison e Anderson pelas trocas de experiências.

aos demais colegas do Departamento de Matemática a ótima convivência.

aos professores da UNIOESTE, campus de Foz do Iguaçu: Kelly Lübeck, Fernando Bando, Luciano Panek, Claiton Massarolo e Emerson Lazzarotto os ótimos cursos oferecidos e o incentivo para seguir na carreira acadêmica.

aos membros da banca examinadora os comentários, sugestões e contribuições, que ajudaram a melhorar a qualidade e a redação final do manuscrito.

ao PMA/UEM a ótima estrutura que oferece aos estudantes e pesquisadores.

à CAPES o suporte financeiro.

aos meus pais, Seni e Francisco, a confiança depositada em mim e as abdicações que fizeram durante toda a vida para me verem onde estou hoje.

aos amigos mais próximos: Wadley, Alessandro, Taís, Adrieli, Inara, Jean, Eduardo, Guilherme Futoshi, Guilherme Melluzzi, João, Everton, Francisco e Luiz Guilherme o carinho, o apoio incondicional nas horas difíceis e a presença em todos os momentos de alegria.

à Deise a incansável tarefa de querer me ver bem, ter me apresentado Maringá, me ajudado de todas as formas possíveis e estado do meu lado me fornecendo o apoio de que precisei.

a todos que de alguma forma contribuíram com o meu progresso como aluno e como ser.

Na maior parte das ciências uma geração põe abaixo o que outra construiu, e o que uma estabeleceu a outra desfaz. Somente na matemática é que cada geração constrói um novo andar sobre a antiga estrutura.

Hermann Hankel

Resumo

Neste trabalho tratamos de grupos de Lie com métricas intrínsecas invariantes à esquerda. Definimos métricas de Finsler invariantes à esquerda e métricas de Carnot-Carathéodory em distribuições completamente não-holonômicas, e chamamos a versão Finsler destas últimas de métricas de Carnot-Carathéodory-Finsler. O objetivo principal deste trabalho é provar que toda métrica intrínseca invariante à esquerda em um grupo de Lie é uma métrica de Carnot-Carathéodory-Finsler. Estudamos também em que condições as métricas intrínsecas invariantes à esquerda são de Finsler, mostrando que as métricas para as quais essa condição é satisfeita são caracterizadas pela retificabilidade dos subgrupos a 1-parâmetro do grupo de Lie.

Palavras-chave: Grupos de Lie, Métricas Intrínsecas, Métricas de Carnot-Carathéodory, Métricas de Finsler, Geometria subriemanniana.

Abstract

In this work we deal with Lie groups with left-invariant intrinsic metrics. We define left-invariant Finsler metrics and Carnot-Carathéodory metrics in completely nonholonomic distributions, and we call the Finsler version of the latter metrics by Carnot-Carathéodory-Finsler metrics. The main objective of this work is to prove that all left-invariant intrinsic metric in a Lie group is a Carnot-Carathéodory-Finsler metric. We also study conditions under which the left-invariant intrinsic metrics are Finsler, showing that the metrics that satisfy this condition are characterized by rectifiability of one-parameter subgroups of the Lie group.

Key-words: Lie groups, Intrinsec metrics, Carnot-Carathéodoy metrics, Finsler metrics, Subriemannian geometry.

Sumário

Introdução

No Capítulo 1 tratamos das métricas intrínsecas, que basicamente são métricas que provém de uma estrutura de comprimento, isto é, de uma estrutura que, respeitando uma série de propriedades, "mede" o comprimento de curvas num espaço topológico. Aproveitando o ensejo do caráter bastante geométrico da estrutura de comprimento apresentamos a noção de ângulo no sentido de Alexandrov num espaço métrico. A grosso modo, o problema da existência do ângulo no sentido de Alexandrov está intimamente relacionado ao problema de determinar se uma métrica de Finsler é Riemanniana.

Ainda são apresentados conceitos e resultados considerados básicos ao escopo do trabalho e que serão amplamente utilizados posteriormente. De forma sucinta apresentamos definições de variedade Riemanniana, campos de vetores e colchete de Lie, grupos e álgebras de Lie, distribuições e ações de grupos que são as ferramentas centrais de nosso estudo. Outro tópico importante que tratamos é a a série de Baker-Campbell-Hausdorff (BCH, chamada por alguns autores apenas como série de Campbell-Hausdorff).

No Capítulo 2 são feitos alguns cálculos com a série de Baker-Campbell-Hausdorff que são o prelúdio para uma sequência encadeada de lemas que são demonstrados, culminando no importante Lema 2.11. Este lema relaciona a estrutura métrica, mais precisamente as bolas abertas do grupo topológico localmente compacto e completo com a estrutura de grupo.

No capítulo 3 estudamos variedades munidas com distribuições completamente não-holonômicas, e com uma métrica sub-riemanniana. Apresentamos o Teorema de ball-box, que nos dá uma descrição qualitativa das bolas sub-riemannianas. Como consequência temos o Teorema de Chow, que estabelece que dados dois pontos em uma variedade sub-riemanniana conexa conforme descrito acima, existe um caminho horizontal C^1 por partes, que liga esses pontos. Na sequência do capítulo é definida métrica de Carnot-Carathéodory e a sua versão Finsler. É também provado o Teorema 3.5 que tem o seguinte enunciado:

Teorema. *Sejam G um grupo de Lie conexo, L_0 um subespaço vetorial da álgebra de Lie $\mathfrak{g}$, que gera $\mathfrak{g}$, F_0 uma norma arbitrária em L_0 e $\theta : TG \to \mathfrak{g}$ a aplicação definida por $\theta(p, v) = dL_{p^{-1}}(v)$. Então a fórmula*

$$d_c(g, h) = \inf_X \int_0^1 F_0(\theta(X'(t)))dt, \tag{1}$$

onde o ínfimo é tomado sobre todos os caminhos horizontais continuamente diferenciáveis por *partes* $X = X(t)$, $0 \leq t \leq 1$, *em* G *ligando* g *e* h, *define uma métrica (Carnot-Carathéodory-Finsler) intrínseca invariante à esquerda em* G *que é compatível com a topologia de* G.

No Capítulo 4 provamos que para um grupo de Lie conexo com métrica intrínseca invariante à esquerda existe um subespaço da álgebra de Lie deste grupo e uma norma nesse subespaço de forma que a métrica, definida pela fórmula (1), coincide com a métrica inicial. Além disso essa correspondência entre métricas intrínsecas invariantes à esquerda e os pares (L_0, F_0), onde L_0 é um subespaço que gera a álgebra $\mathfrak{g}$ e F_0 é uma norma em L_0, é uma bijeção. É provado também o Teorema 4.12 que dá uma condição suficiente para que toda métrica intrínseca invariante à esquerda num grupo de Lie conexo seja de Finsler.

Capítulo 1

Preliminares e Definições

1.1 Estrutura de comprimento

No que segue indicamos [2] como referência principal. Trataremos de métricas intrínsecas, mas antes disso é preciso entender a estrutura por trás disso.

Uma estrutura de comprimento em um espaço topológico X consiste em uma classe de caminhos admissíveis em X para os quais nós possamos medir seu comprimento, e o comprimento em si é uma correspondência atribuindo um número não negativo para cada caminho da classe. Tanto a classe quanto a correspondência têm que possuir várias propriedades naturais.

Lembrando que em nosso contexto um caminho γ em um espaço topológico X é uma aplicação contínua $\gamma : I \to X$ definida em um intervalo $I \subset \mathbb{R}$. Por um intervalo entendemos qualquer subconjunto conexo da reta real (ele pode ser aberto ou fechado, finito ou infinito, e um ponto é considerado um intervalo).

Definição 1.1 (Estrutura de comprimento)**.** *Uma estrutura de comprimento em um espaço topológico X é uma classe A de caminhos admissíveis, que é um subconjunto de todos os caminhos contínuos em X, junto com uma aplicação $\ell : A \to \mathbb{R}_+ \cup \{\infty\}$ que é chamada comprimento do caminho. A classe A deve satisfazer as seguintes condições:*

1. *A classe A é fechada para restrições: Se $\gamma : [a, b] \to X$ é um caminho admissível e $a \leq c \leq d \leq b$, então a restrição $\gamma\,|_{[c,d]}$ de γ ao intervalo $[c, d]$ é também admissível.*

2. *A é fechada para concatenações (produto) de caminhos. A saber, se um caminho γ : $[a, b] \to X$ é tal que as restrições $\gamma_1,\ \gamma_2$ para $[a, c]$ e $[c, b]$ são ambos caminhos admissíveis, então γ também o é. (Lembrando que γ é chamada o produto ou concatenação de γ_1 e γ_2, $\gamma = \gamma_1 \cdot \gamma_2$).*

3. *A é fechada para (pelo menos) reparametrizações lineares: para um caminho admissível $\gamma : [a, b] \to X$ e um homeomorfismo $\varphi : [c, d] \to [a, b]$ da forma $\varphi(t) = \alpha t+\beta$, a composição $\gamma \circ \varphi(t) = \gamma(\varphi(t))$ é também um caminho admissível.*

Exigimos que ℓ possua as seguintes propriedades:

1. *Comprimento de caminhos é aditivo:* $\ell(\gamma\ |_{[a,b]}) = \ell(\gamma\ |_{[a,c]}) + \ell(\gamma\ |_{[c,b]})$ *para qualquer* $c \in [a,b]$.

2. *O comprimento de um pedaço de um caminho depende continuamente do pedaço. Mais formalmente, para um caminho* $\gamma : [a,b] \to X$ *de comprimento finito, denote por* $\ell(\gamma, a, t)$ *o comprimento da restrição de* $\gamma : [a,b] \to X$ *para o segmento* $[a,t]$. *Exigimos que* $\ell(\gamma, a, \cdot)$ *seja uma função contínua. (Observe que a propriedade anterior implica que* $\ell(\gamma, a, a) = 0$.*)*

3. *O comprimento é invariante por reparametrizações:* $\ell(\gamma \circ \varphi) = \ell(\gamma)$ *para um homeomorfismo linear* φ.

4. *Exigimos que a estrutura de comprimento concorde com a topologia de* X *no seguinte sentido: para uma vizinhança* U_x *de um ponto* x, *o comprimento de caminhos conectando* x *com os pontos do complementar de* U_x *é limitado inferiormente por um número estritamente positivo:*
$$\inf\{\ell(\gamma)\ \mid\ \gamma(a) = x,\ \gamma(b) \in X \setminus U_x\} > 0.$$

Adotaremos a notação $\ell(\gamma, a, b)$ introduzida acima. Se $\gamma : I \to X$ é um caminho (admissível) e $[a,b] \subset I$, onde $a \leq b$, denotamos por $\ell(\gamma, a, b)$ o comprimento da restrição de γ à $[a,b]$, isto é, $\ell(\gamma, a, b) = \ell(\gamma\ |_{[a,b]})$. Adicionalmente definimos $\ell(\gamma, b, a) = \ell(\gamma, a, b)$. Esta convenção implica que $\ell(\gamma, a, b) = \ell(\gamma, a, c) + \ell(\gamma, c, b)$ para todos a, b, $c \in I$.

Tendo em mãos uma estrutura de comprimento é possível definir uma métrica no espaço topológico X associada a essa estrutura. Vamos assumir sempre que o espaço topológico X no qual estamos considerando a estrutura de comprimento seja um espaço de Hausdorff. Assim, dados dois pontos x, $y \in X$ definimos a distância $d(x,y)$ entre eles como sendo o ínfimo dos comprimentos de caminhos admissíveis conectando esses dois pontos:

$$d_\ell(x,y) = \inf\{\ell(\gamma)\ \mid\ \gamma : [a,b] \to X, \gamma \in A, \gamma(a) = x, \gamma(b) = y\}.$$

Se estiver claro no contexto qual a estrutura de comprimento ℓ que dá origem a d_ℓ, usualmente omitimos o ℓ na notação d_ℓ.

Proposição 1.2. (X, d_ℓ) *é um espaço métrico.*

Demonstração. 1. $d_\ell(x,x) = 0$ para todo $x \in X$;

De fato, para qualquer caminho $\gamma : [a,b] \to X$, $\gamma \in A$, temos

$$\ell(\gamma, a, b) = \ell(\gamma, a, c) + \ell(\gamma, c, b) \quad \forall c \in [a,b]$$

Tomando, em particular, $c = a$, temos

$$\ell(\gamma, a, b) = \ell(\gamma, a, a) + \ell(\gamma, a, b) \quad \Rightarrow \quad \ell(\gamma, a, a) = 0$$

Logo, $d_\ell(x,x) = 0$.

2. $d_\ell(x, y) > 0$ para quaisquer $x, y \in X$, com $x \neq y$;
   Se $x \neq y$, por X ser de Hausdorff, existe uma vizinhança U_x de x tal que $y \notin U_x$. Como ℓ concorda com a topologia de x,
   $$\inf\{\ell(\gamma) \mid \gamma \in A, \gamma(a) = x, \gamma(b) \in X \setminus U_x\} > 0,$$
   donde
   $$\inf\{\ell(\gamma) \mid \gamma \in A, \gamma(a) = x, \gamma(b) = y\} > 0 \quad \Rightarrow \quad d_\ell(x, y) > 0.$$

3. $d_\ell(x, y) = d_\ell(y, x)$ para quaisquer $x, y \in X$;
   De fato, considere o homeomorfismo linear $\varphi : [-b, -a] \to [a, b]$ onde $\varphi(t) = -t$.
   Se $\gamma : [a, b] \to X$ é um caminho que liga x a y, então $\gamma \circ \varphi$ é um caminho que liga y a x, e como ℓ é invariante por reparametrizações lineares, então $\ell(\gamma) = \ell(\gamma \circ \varphi)$. Logo é imediato que $d_\ell(x, y) = d_\ell(y, x)$.

4. $d_\ell(x, z) \leq d_\ell(x, y) + d_\ell(y, z)$, para quaisquer $x, y, z \in X$;
   Para simplificar a notação, indiquemos por $\gamma : x \leftrightarrow y$ uma curva γ que ligue x a y.
   Suponha, por absurdo, que $d_\ell(x, z) > d_\ell(x, y) + d_\ell(y, z)$ para alguns $x, y, z \in X$. Então
   $$\begin{aligned} \inf\{\ell(\gamma) \mid \gamma \in A, \gamma : x \leftrightarrow z\} &> \inf\{\ell(\gamma_1) \mid \gamma_1 \in A, \gamma_1 : x \leftrightarrow y\} \\ &\quad + \inf\{\ell(\gamma_2) \mid \gamma_2 \in A, \gamma_2 : y \leftrightarrow z\} \\ &= \inf\{\ell(\gamma_1) + \ell(\gamma_2) \mid \gamma_1 \in A, \gamma_1 : x \leftrightarrow y, \gamma_2 \in A, \gamma_2 : y \leftrightarrow z\} \\ &= \inf\{\ell(\gamma_1 \cdot \gamma_2) \mid \gamma_1 \cdot \gamma_2 \in A, \gamma_1 \cdot \gamma_2 : x \leftrightarrow y \leftrightarrow z\} \\ &= \inf\{\ell(\gamma) \mid \gamma \in A, \gamma : x \leftrightarrow y \leftrightarrow z\}, \end{aligned}$$
   o que é um absurdo, pois o conjunto das curvas admissíveis que ligam x a z contém o conjunto das curvas admissíveis que ligam x a z passando por y.

■

Note que d_ℓ não é necessariamente uma métrica finita. Por exemplo, se X é uma união desconexa de duas componentes, não há caminhos contínuos que podem ir de uma componente para a outra e portanto a distância entre pontos das diferentes componentes é infinita. Por outro lado, podem haver pontos tais que existam caminhos contínuos os conectando mas todos tenham comprimento infinito. Diz-se que dois pontos x, $y \in X$ pertencem a mesma componente de acessibilidade se podem ser ligados por um caminho de comprimento finito.

Proposição 1.3. *Caminhos admissíveis de comprimento finito são contínuos com respeito a* (X, d_ℓ).

Proposição 1.4. *A topologia determinada por* d_ℓ *é mais fina do que a de* X*: qualquer conjunto aberto em* X *é aberto em* (X, d_ℓ) *também.*

Definição 1.5. *Uma métrica que pode ser obtida como a função distância associada a uma estrutura de comprimento é chamada de métrica intrínseca. Um espaço métrico do qual a métrica é intrínseca é chamado de espaço métrico intrínseco.*

Nem toda métrica pode ser originada de uma métrica intrínseca. Mesmo se (X,d) é um espaço métrico intrínseco e $A \subset X$, a restrição de d ao conjunto A não é necessariamente intrínseca. Por exemplo, considere a esfera unitária S^1 em $\mathbb{R}^2$. Temos que a métrica usual em $\mathbb{R}^2$ restrita à S^1 difere da métrica intrínseca de S^1. Tomando pontos antipodais $x, y \in S^1$ a distância euclidiana entre eles é 2, enquanto a distância entre os mesmos na métrica intrínseca (associada a estrutura de comprimento usual) é π, já que o comprimento do menor caminho em S^1 que liga pontos antipodais tem comprimento π.

Definição 1.6. *Uma estrutura de comprimento é dita completa se para quaisquer dois pontos x, y existe um caminho admissível ligando-os cujo comprimento é igual a $d_\ell(x,y)$; em outras palavras, uma estrutura de comprimento é completa se existe um caminho mínimo ligando cada dois pontos.*

Métricas intrínsecas associadas com estruturas de comprimento completas são ditas estritamente intrínsecas.

Definição 1.7. *Um ponto $z \in X$ é chamado ponto médio entre os pontos x e y num espaço métrico (X,d) se $d(x,z) = d(y,z) = \frac{1}{2}d(x,y)$.*

Lema 1.8. *Se d é uma métrica estritamente intrínseca, então para quaisquer dois pontos x, y existe um ponto médio z.*

Demonstração. O comprimento de um caminho mais curto $\gamma : [a,b] \to X$ entre x e y é $\ell(\gamma) = d(x,y)$. Denotamos $\ell(t) = \ell(\gamma_{[a,t]})$. Como $\ell(t)$ é contínua em t e $\ell(0) = 0$, existe $c \in [a,b]$ tal que $\ell(c) = \frac{1}{2}\ell(b)$. Agora escolhendo $z = \gamma(c)$ e lembrando que o comprimento de um caminho nunca é menor que a distância entre seus pontos inicial e final vem logo que $d(x,z) = d(y,z) = \frac{1}{2}d(x,y)$. ∎

1.2 Estrutura de comprimento induzida

Podemos ainda induzir uma estrutura de comprimento caso já tenhamos uma métrica.

Definição 1.9. *Seja (X,d) um espaço métrico e γ um caminho em X, isto é, uma aplicação contínua $\gamma : [a,b] \to X$. Considere a partição Y de $[a,b]$, ou seja, uma coleção finita de pontos $Y = \{y_0, \ldots, y_N\}$ tais que $a = y_0 \le y_1 \le y_2 \le \cdots \le y_N = b$. O supremo das somas*

$$\Sigma(Y) = \sum_{i=1}^{N} d(\gamma(y_{i-1}), \gamma(y_i))$$

sobre todas as partições Y é chamado o comprimento de γ (com respeito a métrica d) e denotado por $\ell_d(\gamma)$. Uma curva é dita retificável se seu comprimento é finito.

A estrutura de comprimento induzida pela métrica d é definida como segue: Todos os caminhos contínuos (parametrizados por intervalos fechados) são admissíveis, e o comprimento é dado pela função ℓ_d.

1.3 Ângulo no sentido de Alexandrov

A definição de ângulo no sentido de Alexandrov é inspirada no caso euclidiano. Para tal, definamos primeiramente o chamado ângulo de comparação.

Definição 1.10. *Sejam x, y e z três pontos distintos em um espaço métrico (X, d). O ângulo de comparação xyz, denotado por $\tilde{\measuredangle}(x, y, z)$, é definido por*

$$\tilde{\measuredangle}(xyz) = \arccos\left(\frac{d^2(x,y) + d^2(y,z) - d^2(x,z)}{2d(x,y)d(y,z)}\right).$$

Definição 1.11. *Sejam $\alpha : [0, \varepsilon) \to X$ e $\beta : [0, \varepsilon) \to X$ dois caminhos em um espaço métrico intrínseco X, ambos saindo do mesmo ponto $p = \alpha(0) = \beta(0)$. Definimos o ângulo no sentido de Alexandrov $\measuredangle(\alpha, \beta)$ entre α e β como*

$$\measuredangle(\alpha, \beta) = \lim_{s,t \to 0} \tilde{\measuredangle}(\alpha(s), p, \beta(t)),$$

se o limite existir.

Assumindo que as curvas α e β estão fixas, definimos $\theta(s, t) = \tilde{\measuredangle}(\alpha(s), p, \beta(t))$. Nesta notação

$$\measuredangle(\alpha, \beta) = \lim_{s,t \to 0} \theta(s, t).$$

Em geral espaços com estrutura de comprimento podem ser bastante mal comportados, e quando nos restringimos à espaços de Alexandrov, ou seja, espaços de comprimento em que a curvatura é limitada, o ângulo entre dois caminhos de menor comprimento está sempre definido e por isso dizemos ângulo no sentido de Alexandrov. Não tratamos aqui de curvaturas pois não estamos interessados exatamente nos espaços de Alexandrov. Caso o leitor queira saber mais a respeito destes espaços confira [2].

A seguir alguns resultados sobre ângulos de Alexandrov que serão utilizados posteriormente.

Proposição 1.12. i. *Todo caminho de menor comprimento forma ângulo nulo consigo mesmo.*

ii. *Se dois segmentos de menor comprimento $[a, b]$ e $[b, c]$ são tais que sua concatenação(“[abc]”) continua sendo um caminho de menor comprimento, então o ângulo entre $[b, a]$ e $[b, c]$ é π.*

Teorema 1.13. *Considere três curvas γ_1, γ_2 e γ_3 com mesmo ponto inicial p. Assuma que os ângulos $\alpha_1 = \measuredangle(\gamma_2, \gamma_3)$ e $\alpha_2 = \measuredangle(\gamma_1, \gamma_3)$ existam. Se o ângulo $\alpha_3 = \measuredangle(\gamma_1, \gamma_2)$ também existe, então ele satisfaz a seguinte desigualdade triangular:*

$$\alpha_3 \leq \alpha_1 + \alpha_2.$$

1.4 Variedades diferenciáveis, campos de vetores e colchete

A teoria sobre variedades diferenciáveis pode ser encontrada em [4] e em [14], dentre outros. Em nosso caso estamos considerando variedades diferenciáveis de Hausdorff e com base enumerável.

Definição 1.14. *Uma variedade diferenciável de dimensão n é um conjunto M e uma família de aplicações biunívocas $\boldsymbol{x}_\alpha : U_\alpha \subset \mathbb{R}^n \to M$ de abertos U_α de $\mathbb{R}^n$ em M tais que:*

1. $\bigcup_\alpha \boldsymbol{x}_\alpha(U_\alpha) = M$
2. *Para todo par α, β, com $\boldsymbol{x}_\alpha(U_\alpha) \cap \boldsymbol{x}_\beta(U_\beta) = W \neq \emptyset$, os conjuntos $\boldsymbol{x}_\alpha^{-1}(W)$ e $\boldsymbol{x}_\beta^{-1}(W)$ são abertos em $\mathbb{R}^n$ e as aplicações $\boldsymbol{x}_\beta^{-1} \circ \boldsymbol{x}_\alpha$ são diferenciáveis.*
3. *A família $\{(U_\alpha, \boldsymbol{x}_\alpha)\}$ é máxima relativamente às condições (1) e (2).*

$(U_\alpha, \boldsymbol{x}_\alpha)$ com $p \in \boldsymbol{x}_\alpha(U_\alpha)$ é chamada uma parametrização (ou sistema de coordenadas) de M em p. $\boldsymbol{x}_\alpha(U_\alpha)$ é chamada de vizinhança coordenada em p e a família $\{(U_\alpha, \boldsymbol{x}_\alpha)\}$ que satisfaz as condições (1) e (2) da definição anterior é chamada de estrutura diferenciável.

Uma estrutura diferenciável em um conjunto M induz uma topologia em M: $A \subset M$ é um aberto em M se $\boldsymbol{x}_\alpha^{-1}(A \cap \boldsymbol{x}_\alpha(U_\alpha))$ é um aberto de $\mathbb{R}^n$ para todo α.

Definição 1.15. *Sejam M_1^n e M_2^m variedades diferenciáveis. Uma aplicação $\varphi : M_1 \to M_2$ é diferenciável em $p \in M_1$ se, dada uma parametrização $\boldsymbol{y} : V \subset \mathbb{R}^m \to M_2$ em $\varphi(p)$, existe uma parametrização $\boldsymbol{x} : U \subset \mathbb{R}^n \to M_1$ em p tal que $\varphi(\boldsymbol{x}(U)) \subset \boldsymbol{y}(V)$ e a aplicação*

$$\boldsymbol{y}^{-1} \circ \varphi \circ \boldsymbol{x} : U \subset \mathbb{R}^n \to \mathbb{R}^m$$

é diferenciável em $\boldsymbol{x}^{-1}(p)$. Dizemos que φ é diferenciável em um aberto de M_1 se é diferenciável em todos os pontos deste aberto.

A condição (2) da definição de variedade diferenciável implica que a definição de aplicação diferenciável não depende da escolha da parametrização.

Definição 1.16. *Seja M uma variedade diferenciável. Uma aplicação diferenciável $\alpha : (-\varepsilon, \varepsilon) \to M$ é chamada de curva (diferenciável) em M. Suponha que $\alpha(0) = p \in M$, e seja $\mathcal{D}$ o conjunto de todas as funções de M que são diferenciáveis no ponto p. O vetor tangente à curva α em $t = 0$ é a função $\alpha'(0) : \mathcal{D} \to \mathbb{R}$ dada por*

$$\alpha'(0)f = \left.\frac{d(f \circ \alpha)}{dt}\right|_{t=0} \qquad f \in \mathcal{D}.$$

Um vetor tangente em p é o vetor tangente de uma curva $\alpha : (-\varepsilon, \varepsilon) \to M$ com $\alpha(0) = p$. O conjunto dos vetores tangentes a M em p indicaremos por T_pM e denominaremos de espaço tangente.

Quando nos interessa o ponto inicial e final, geralmente chamamos uma curva (diferenciável) $\alpha : [a,b] \to M$ de caminho (diferenciável) que liga $\alpha(a)$ a $\alpha(b)$ em M.

Pode-se mostrar que o vetor tangente a uma curva α num ponto $p \in M$ depende somente das derivadas de α em um sistema de coordenadas e o conjunto T_pM, com as operações usuais de soma e produto por escalar, é um espaço vetorial de mesma dimensão da variedade.

Se tomarmos M_1^n e M_2^m variedades diferenciáveis e $\varphi : M_1 \to M_2$ uma aplicação diferenciável e para cada par $p \in M_1$ e $v \in T_pM_1$, escolhermos uma curva diferenciável $\alpha : (-\varepsilon, \varepsilon) \to M_1$ com $\alpha(0) = p$, $\alpha'(0) = v$ é possível verificar que a aplicação $d\varphi_p : T_pM_1 \to T_{\varphi(p)}M_2$ dada por $d\varphi_p(v) = \beta'(0)$, onde $\beta = \varphi \circ \alpha$, é uma aplicação linear que não depende da escolha de α. Essa aplicação $d\varphi_p$ é chamada de diferencial de φ no ponto p.

Seja M^n uma variedade diferenciável e seja $TM = \{(p,v); p \in M, v \in T_pM\}$. Vamos munir o conjunto TM com uma estrutura diferenciável de dimensão $2n$, e com essa estrutura este conjunto será chamado de fibrado tangente de M.

Observamos que a forma de munir TM com uma estrutura diferenciável é a mais intuitiva possível: Tome $\{U_\alpha, \mathbf{x}_\alpha\}$ a estrutura diferenciável máxima de M. Indicaremos as coordenadas de U_α por $(x_1^\alpha, \ldots, x_n^\alpha)$ e as bases associadas nos espaços tangentes de $\mathbf{x}_\alpha(U_\alpha)$ por $\left\{\frac{\partial}{\partial x_1^\alpha}, \ldots, \frac{\partial}{\partial x_n^\alpha}\right\}$. Para cada α, definimos $\mathbf{y}_\alpha : U_\alpha \times \mathbb{R}^n \to TM$, por

$$\mathbf{y}_\alpha(x_1^\alpha, \ldots, x_n^\alpha, u_1, \ldots, u_n) = \left(\mathbf{x}_\alpha(x_1^\alpha, \ldots, x_n^\alpha), \sum_{i=1}^n u_i \frac{\partial}{\partial x_i^\alpha}\right), \quad (u_1, \ldots, u_n) \in \mathbb{R}^n.$$

Isso significa que as coordenadas de um par (p,v) são as coordenadas de p no sistema de coordenadas inicial juntamente com as coordenadas de v na base do espaço tangente de p. Assim $\{(U_\alpha \times \mathbb{R}^n, \mathbf{y}_\alpha)\}$ é uma estrutura diferenciável em TM.

Definição 1.17 (Campo de Vetores)**.** *Um campo de vetores X em uma variedade diferenciável M é uma correspondência que a cada ponto $p \in M$ associa um vetor $X(p) \in T_pM$. X é uma aplicação de M no fibrado tangente TM. O campo de vetores X é diferenciável se a aplicação $X : M \to TM$ é diferenciável.*

Definição 1.18. *Seja M uma variedade diferenciável de dimensão n e seja $p \in M$. Dada uma vizinhança U de p, uma família de n campos de vetores $E_1, \ldots, E_n \in \mathcal{X}(U)$ é chamada de referencial local em p se para cada $q \in U$, $\{E_1(q), \ldots, E_n(q)\}$ é base de T_qM.*

É possível escrever, dada uma parametrização $\mathbf{x} : U \subset \mathbb{R}^n \to M$,

$$X(p) = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x_i},$$

onde cada $a_i : U \to \mathbb{R}$ é uma função real em U e $\left\{\frac{\partial}{\partial x_i}\right\}$ é a base associada à parametrização, $i = 1, \ldots, n$. X é diferenciável se, e só se, as funções a_i são diferenciáveis para todo i e para qualquer parametrização $\mathbf{x}$.

Por vezes é útil pensar em um campo de vetores como uma aplicação $X : \mathcal{D} \to \mathcal{F}$, onde $\mathcal{D}$ é o conjunto das aplicações diferenciáveis de M e $\mathcal{F}$ é o conjunto de todas as funções de M.

Convém observar que X é diferenciável se, e somente se, $Xf \in \mathcal{D}$ para todo $f \in \mathcal{D}$. Observe que em relação a uma parametrização $\mathbf{x}$, temos que

$$(Xf)(p) = \sum_{i=1}^{n} a_i(p) \frac{\partial f}{\partial x_i}(p),$$

onde f se refere à expressão de f em $\mathbf{x}$.

Graças a essa interpretação dos campos diferenciáveis de vetores como operadores em $\mathcal{D}$ podemos considerar iterados de X. Por exemplo, se X e Y são campos diferenciáveis e $f \in \mathcal{D}$ podemos considerar as funções $X(Yf)$ e $Y(Xf)$.

Lema 1.19. *Sejam X e Y campos diferenciáveis de vetores em uma variedade diferenciável M. Então existe um único campo de vetores Z tal que $Zf = (XY - YX)f$, para toda $f \in \mathcal{D}$.*

O campo vetorial Z dado pelo Lema 1.19 é chamado o colchete de X e Y, e denotado por $[X, Y] = XY - YX$ e é também um campo diferenciável.

Proposição 1.20. *Se X, Y e Z são campos diferenciáveis de vetores em M, a, b são números reais, e f, g são funções diferenciáveis, então:*

1. *Bilinearidade sobre $\mathbb{R}$:*
$$[aX + bY, Z] = a[X, Z] + b[Y, Z];$$

2. *Anti-simetria: $[X, Y] = -[Y, X]$;*

3. *Identidade de Jacobi:*
$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0; \tag{1.1}$$

4. $[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X$.

Teorema 1.21. *Seja X um campo diferenciável de vetores em uma variedade diferenciável M, e seja $p \in M$. Então existem uma vizinhança $U \subset M$ de p, um intervalo $(-\delta, \delta)$, $\delta > 0$, e uma aplicação diferenciável $\varphi : (-\delta, \delta) \times U \to M$ tais que a curva $t \mapsto \varphi(t, q)$, $t \in (-\delta, \delta)$, $q \in U$, é a única curva que satisfaz $\dfrac{\partial \varphi}{\partial t} = X(\varphi(t, q))$ e $\varphi(0, q) = q$.*

Uma curva $\alpha : (-\delta, \delta) \to M$ que satisfaz às condições $\alpha'(t) = X(\alpha(t))$ e $\alpha(0) = q$ é chamada a trajetória do campo X que passa por q para $t = 0$. O teorema acima garante a existência e unicidade da trajetória satisfazendo a condição inicial e é comum utilizar a notação $\varphi_t(q) = \varphi(t, q)$ e chamar $\varphi_t : U \to M$ o fluxo local de X.

A proposição a seguir fornece uma interpretação analítica do colchete de Lie como primeiro termo na expansão de Taylor do comutador dos fluxos dos campos de vetores.

Proposição 1.22. *Sejam $X, Y : U \to \mathbb{R}^n$ campos de vetores diferenciáveis definidos no aberto $U \subset \mathbb{R}^n$. Fixe $x \in U$ e considere a curva*

$$\alpha(t) = X_t \circ Y_t \circ X_{-t} \circ Y_{-t}(x)$$

definida em um intervalo aberto de $0 \in \mathbb{R}$. Então, $\alpha'(0) = 0$ e $\alpha''(0) = 2[Y, X](x)$.

Um campo de vetores X ao longo de uma curva $\alpha : I \to M$ é uma aplicação que, a cada t no aberto $I \subset \mathbb{R}$, associa um vetor tangente $X(t) \in T_{\alpha(t)}M$. Diremos que X é diferenciável se para toda função diferenciável f em M, a função $t \mapsto X(t)f$ é uma função diferenciável no intervalo aberto $I \subset \mathbb{R}$.

O campo vetorial $d\alpha\left(\dfrac{d}{dt}\right)$, indicado por $\dfrac{d\alpha}{dt}$, é chamado campo tangente (ou velocidade) de α.

Na direção de definir a aplicação exponencial vamos definir alguns conceitos que nos serão úteis. Indicaremos por $\mathcal{X}(M)$ o conjunto dos campos diferenciáveis em M.

Definição 1.23. *Uma conexão afim ∇ em uma variedade diferenciável M é uma aplicação*

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M),$$

indicada por $\nabla(X,Y) = \nabla_X Y$, que, para quaisquer X, Y, $Z \in \mathcal{X}(M)$ e $f, g \in \mathcal{D}$, satisfaz as seguintes condições:

1. $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z$;
2. $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$;
3. $\nabla_X(fY) = f\nabla_X Y + X(f)Y$.

Proposição 1.24. *Seja M uma variedade diferenciável com uma conexão afim ∇. Então existe uma única correspondência que associa a um campo vetorial V ao longo de uma curva diferenciável $\alpha : I \subset \mathbb{R} \to M$ um outro campo vetorial $\frac{DV}{dt}$ ao longo da curva α, denominado derivada covariante de V ao longo de α, tal que:*

1. $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$;
2. $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$, *onde W é um campo de vetores ao longo de α e f é uma função diferenciável em I;*
3. *Se V é induzido por um campo de vetores $Y \in \mathcal{X}(M)$, isto é, $V(t) = Y(\alpha(t))$, então* $\frac{DV}{dt} = \nabla_{\frac{d\alpha}{dt}}Y$.

A seguir apresentaremos algumas definições de geometria Riemanniana, onde basicamente munimos uma variedade diferenciável com uma estrutura métrica.

Definição 1.25. *Uma métrica Riemanniana em uma variedade diferenciável M é uma correspondência que associa a cada ponto p de M um produto interno $\langle\cdot,\cdot\rangle_p$ no espaço tangente T_pM, que varia diferenciavelmente no seguinte sentido: Se $\boldsymbol{x} : U \subset \mathbb{R}^n \to M$ é um sistema de coordenadas locais em torno de p, com $\boldsymbol{x}(x_1,\ldots,x_n) = q \in \boldsymbol{x}(U)$ e $\frac{\partial}{\partial x_i}(q) = d\boldsymbol{x}(0,\ldots,1,\ldots,0)$, então*

$$\left\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) \right\rangle_q = g_{ij}(x_1,\ldots,x_n)$$

é uma função diferenciável em U para todos $i,j = 1,\ldots,n$.

Sempre que não houver possibilidade de confusão deixaremos de indicar o índice p de $\langle \cdot, \cdot \rangle_p$.

A métrica Riemanniana pode ser utilizada para medir o comprimento de curvas na variedade diferenciável.

A restrição de uma curva α a um intervalo fechado $[a,b] \subset I$ chama-se segmento. Se M é uma variedade Riemanniana, definimos o comprimento de um segmento por

$$\ell_a^b(\alpha) = \int_a^b \left\langle \frac{d\alpha}{dt}, \frac{d\alpha}{dt} \right\rangle^{\frac{1}{2}} dt.$$

Podemos introduzir uma distância numa variedade Riemanniana M conexa da seguinte maneira: A distância $d(p,q)$ é definida como $d(p,q) = \inf_{\gamma}\{\ell(\gamma)\}$, onde o ínfimo é tomado sobre todas as curvas diferenciáveis por partes que ligam p a q. Essas curvas existem graças à conexidade da variedade. Com essa função distância, (M,d) é um espaço métrico. Chamamos a métrica d de métrica associada ao tensor g.

Definição 1.26. *Sejam (M,d_1) e (M,d_2) espaços métricos. Se para qualquer $p \in M$ existem uma vizinhança $V(p)$ e constantes $c = c(p)$, $C = C(p) \in \mathbb{R}$, $c, C > 0$ tais que para todos $x, y \in V(p)$, $cd_2(x,y) \leq d_1(x,y) \leq Cd_2(x,y)$ dizemos que (M,d_1) e (M,d_2) são localmente Lipschitz equivalentes. Se existirem $c, C > 0$ tal que $cd_2(x,y) \leq d_1(x,y) \leq Cd_2(x,y)$ para todos $x, y \in M$ então (M,d_1) e (M,d_2) são Lipschitz equivalentes.*

Proposição 1.27. *Sejam (M,g_1) e (M,g_2) variedades Riemannianas de dimensão n. Sejam d_1 e d_2 as funções distância associadas aos tensores g_1 e g_2, respectivamente. Então (M,d_1) e (M,d_2) são localmente Lipschitz equivalentes. Se $K \subset M$ é compacto então (K,d_1) e (K,d_2) são Lipschitz equivalentes.*

Demonstração. Como estamos em um espaço de dimensão finita as normas induzidas por g_1 e g_2 são equivalentes para cada T_pM. Pela continuidade de g_1 e g_2, dado $p \in M$, existe uma vizinhança $V(p)$ de p e constantes positivas $c(p)$ e $C(p)$ tais que $c(p)g_2(v,v) \leq g_1(v,v) \leq C(p)g_2(v,v)$ para todo $v \in T_qM$, com $q \in V(p)$. Com isso, se γ é um caminho diferenciável em $V(p)$, temos que

$$c(p)\ell_2(\gamma) \leq \ell_1(\gamma) \leq C(p)\ell_2(\gamma).$$

Daí, se $x, y \in V(p)$, então $c(p)\inf_{\gamma}\{\ell_2(\gamma)\} \leq \ell_1(\gamma)$ para todo caminho diferenciável γ ligando x à y. Com isso $c(p)\inf_{\gamma}\{\ell_2(\gamma)\} \leq \inf_{\gamma}\{\ell_1(\gamma)\}$ e $c(p)d_2(x,y) \leq d_1(x,y)$. A desigualdade $d_1(x,y) \leq C(p)d_2(x,y)$ é análoga.

Suponha que $K \subset M$ seja compacto, então tome uma família $\{V(p)\}_{p\in K}$, onde para cada $p \in K$, $V(p)$ é um aberto, como na Definição 1.26, para o qual $(V(p),d_1)$ e $(V(p),d_2)$ são Lipschitz equivalentes. Tome uma subcobertura finta $\mathcal{A} = \{V(p_1), \ldots, V(p_m)\}$ da cobertura aberta $\{V(p)\}_{p\in K}$. Para cada p_i existem as constantes de Lipschitz $c(p_i)$ e $C(p_i)$ tais que se $x, y \in V(p_i)$ então $c(p_i)d_2(x,y) \leq d_1(x,y) \leq C(p_i)d_2(x,y)$, $i = 1, \ldots, m$.

Sejam ℓ_1 e ℓ_2 as aplicações de comprimento induzidas pelas métricas d_1 e d_2, respectivamente. Tomando $c = \min\{c(p_i) : 1 \leq i \leq m\}$ e $C = \max\{C(p_i) : 1 \leq i \leq m\}$, afirmamos que dados $x, y \in K$ temos

$$c\ell_2(\gamma) \leq \ell_1(\gamma) \leq C\ell_2(\gamma),$$

para qualquer caminho diferenciável $\gamma : [a,b] \to K$ ligando x a y. De fato, considere uma partição $P = \{s_1 = a < s_2 < \cdots < s_N = b\}$ do intervalo $[a,b]$ suficientemente refinada de forma que para cada $l = 1,\ldots,N$ os pontos $\gamma(s_{l-1})$ e $\gamma(s_l)$ pertençam a uma mesma vizinhança $V(p_k) \in \mathcal{A}$. Desta forma,

$$c\Sigma_2(P) = c\sum_{l=1}^{N} d_2(\gamma(s_{l-1}),\gamma(s_l)) = \sum_{l=1}^{N} cd_2(\gamma(s_{l-1}),\gamma(s_l)) \leq \sum_{l=1}^{N} d_1(\gamma(s_{l-1}),\gamma(s_l)) = \Sigma_1(P)$$

e

$$\Sigma_1(P) = \sum_{l=1}^{N} d_1(\gamma(s_{l-1}),\gamma(s_l)) \leq \sum_{l=1}^{N} Cd_2(\gamma(s_{l-1}),\gamma(s_l)) \leq C\sum_{l=1}^{N} d_2(\gamma(s_{l-1}),\gamma(s_l)) = C\Sigma_2(P).$$

Donde, aplicando o supremo sobre partições todas as partições Q de $[a,b]$ (mais refinadas que P), vem

$$c\ell_2(\gamma) = c\sup_Q\{\Sigma_2(Q)\} = \sup_Q\{c\Sigma_2(Q)\} \leq \sup_Q\{\Sigma_1(Q)\} = \ell_1(\gamma),$$

e

$$\ell_1(\gamma) = \sup_Q\{\Sigma_1(Q)\} \leq \sup_Q\{C\Sigma_2(Q)\} = C\sup_Q\{\Sigma_2(Q)\} = C\ell_2(\gamma).$$

Assim, devido a arbitrariedade da curva γ, decorre que

$$cd_2(x,y) \leq d_1(x,y) \leq Cd_2(x,y).$$

Portanto (K,d_1) e (K,d_2) são Lipschitz equivalentes.

■

A seguir um resultado sobre a existência de métricas Riemannianas, cuja demonstração pode ser encontrada em [4].

Proposição 1.28. *Uma variedade diferenciável M (de Hausdorff e com base enumerável) possui ao menos uma métrica Riemanniana.*

Definição 1.29. *Seja M uma variedade diferenciável com uma conexão afim ∇. Um campo vetorial X ao longo de uma curva $\alpha : I \to M$ é chamado paralelo quando $\dfrac{DX}{dt} = 0$, para todo $t \in I$.*

Definição 1.30. *Seja M uma variedade diferenciável com uma conexão afim ∇ e uma métrica Riemanniana $\langle\cdot,\cdot\rangle$. A conexão é dita compatível com a métrica $\langle\cdot,\cdot\rangle$, quando para toda curva diferenciável α e quaisquer pares de campos de vetores paralelos X e X' ao longo de α, tivermos $\langle X, X'\rangle$ constante.*

Definição 1.31. *Uma conexão afim ∇ em uma variedade diferenciável M é dita simétrica quando*

$$\nabla_X Y - \nabla_Y X = [X,Y],$$

para todo $X, Y \in \mathcal{X}(M)$.

Teorema 1.32 (Levi-Civita). *Dada uma variedade Riemanniana M, existe uma única conexão afim ∇ em M satisfazendo as condições:*

1. *∇ é simétrica;*

2. *∇ é compatível com a métrica Riemanniana.*

A conexão dada pelo teorema acima é chamada de conexão de Levi-Civita ou conexão Riemanniana de M. Essa conexão nos permite derivar campos de vetores em M. Em uma variedade Riemanniana, sempre usaremos a conexão Riemanniana.

Definição 1.33. *Uma curva parametrizada $\gamma : I \to M$ é uma geodésica em $t_0 \in I$ se $\frac{D}{dt}\left(\frac{d\gamma}{dt}\right) = 0$ no ponto $t = t_0$; Se γ é geodésica em t, para todo $t \in I$, diz-se que γ é uma geodésica.*

Proposição 1.34. *Dado $p \in M$, existem uma vizinhança V de p em M, um número $\varepsilon > 0$ e uma aplicação C^∞, $\gamma : (-2,2) \times \mathcal{U} \to M$, $\mathcal{U} = \{(q,w) \in TM; q \in V, w \in TqM, |w| < \varepsilon\}$, tal que $t \mapsto \gamma(t,q,w)$, $t \in (-2,2)$, é a única geodésica de M que no instante $t = 0$ passa por q com velocidade w, para cada $q \in V$ e cada $w \in T_qM$ com $|w| < \varepsilon$.*

Agora podemos apresentar a definição de exponencial. Seja $p \in M$ e $\mathcal{U} \subset TM$ um aberto dado pela Proposição 1.34. Então a aplicação $\exp : \mathcal{U} \to M$ dada por

$$\exp(q,v) = \gamma(1,q,v) = \gamma\left(|v|, q, \frac{v}{|v|}\right), \quad (q,v) \in \mathcal{U},$$

é chamada a aplicação exponencial em $\mathcal{U}$. Trata-se de uma aplicação diferenciável pela forma como foi definida.

É bastante usual restringirmos o domínio da exponencial para um aberto no espaço tangente T_qM, ou seja, considerar

$$\exp_q : B_\varepsilon(0) \subset T_qM \to M$$

com $\exp_q(v) = \exp(q,v)$, onde $B_\varepsilon(0)$ é a bola centrada em $0 \in T_qM$ e de raio ε.

Proposição 1.35. *Dado $q \in M$, existe um $\varepsilon > 0$ tal que $\exp_q : B_\varepsilon(0) \subset T_qM \to M$ é um difeomorfismo de $B_\varepsilon(0)$ sobre um aberto de M que contém q.*

1.5 Grupos e álgebras de Lie

Vamos apresentar a teoria de grupos de Lie e álgebras de Lie baseando-se principalmente em [12] onde podem ser encontradas as demonstrações omitidas aqui. A álgebra de Lie $\mathfrak{g}$ de um grupo de Lie G é basicamente o espaço dos campos de vetores invariantes à esquerda (ou à direita, conforme a escolha) com o colchete dado pelo colchete de Lie entre campos de vetores. Através do fluxo destes campos estabelecemos a exp, que é uma aplicação de $\mathfrak{g}$ em G e é a ferramenta mais utilizada para relacionar estes dois espaços.

Um grupo de Lie é um grupo G com uma estrutura diferenciável tal que a aplicação $G \times G \to G$ dada por $(x,y) \mapsto xy$, $x, y \in G$, é diferenciável. Se o produto é diferenciável a função inversa

$i : G \to G$, dada por $i(g) = g^{-1}$ também é diferenciável. As translações à esquerda L_x e à direita R_x dadas por: $L_x : G \to G$, $L_x(y) = xy$; $R_x : G \to G$, $R_x(y) = yx$ são difeomorfismos.

Dizemos que uma métrica Riemanniana em G é invariante à esquerda se

$$\langle u, v\rangle_y = \langle d(L_x)_y u, d(L_x)_y v\rangle_{L_x(y)}$$

para todo $x, y \in G$, $u, v \in T_y G$, isto é, L_x é uma isometria. Analogamente se define métrica Riemanniana invariante à direita.

Um campo diferenciável de vetores X em um grupo de Lie G é invariante à esquerda se $dL_x X = X$ para todo $x \in G$. Os campos invariantes à esquerda ficam completamente determinados pelos seus valores em algum ponto de G. Assim a cada vetor $X_e \in T_e G$, no espaço tangente do elemento neutro $e \in G$, associamos o campo invariante à esquerda X definido por $X_a = dL_a X_e$, $a \in G$. E o colchete de campos invariantes à esquerda é também invariante à esquerda.

Definição 1.36. *Se X_e e Y_e são vetores em $T_e G$, definimos $[X_e, Y_e] = [X, Y]_e$, onde X e Y são os campos diferenciáveis invariantes à esquerda associados a X_e e Y_e, respectivamente. $T_e G$ munido com essa operação é chamado de álgebra de Lie de G e será indicada por $\mathfrak{g}$. Eventualmente, $\mathfrak{g}$ denotará o espaço vetorial dos campos invariantes à esquerda munidos com o colchete.*

Definição 1.37. *Se $\mathfrak{g}$ é uma álgebra de Lie, então um subespaço $\mathfrak{a}$ de $\mathfrak{g}$ é dito uma subálgebra de Lie se é fechado para o colchete de Lie.*

Proposição 1.38. *Seja G um grupo de Lie e X um campo invariante à esquerda de G. Então o fluxo φ de X está definido em todo $\mathbb{R} \times TG$.*

Definição 1.39. *Seja $X \in \mathfrak{g}$ e $\varphi : \mathbb{R} \times TG \to G$ o fluxo associado a X. Defina* $\exp : \mathfrak{g} \to G$ *por* $\exp(X) = \varphi(1, e, X)$. *Como é usual* $\exp(X)$ *também se denota por e^X. Isso define a aplicação exponencial* $\exp : \mathfrak{g} \to G$, *onde $\mathfrak{g} = T_e G$ é a álgebra de Lie de G.*

Sempre que houver a possibilidade de confusão, especificaremos se exp é a exponencial Riemanniana ou se é a exponencial em grupos de Lie.

Definição 1.40. *A aplicação* $t \mapsto \exp(tX)$, *$X \in \mathfrak{g}$, é um homomorfismo, isto é,*

$$\exp(t+s)(X) = \exp(tX)\exp(sX) = \exp(sX)\exp(tX)$$

e sua imagem $\{\exp(tX) : t \in \mathbb{R}\}$ *é um subgrupo de G. Esse subgrupo é denominado de subgrupo a 1-parâmetro gerado por $X \in \mathfrak{g}$.*

Proposição 1.41.

A aplicação $\log = \exp^{-1} : V \to U$ é um difeomorfismo entre um aberto de G e um aberto de um espaço vetorial. Portanto, log pode ser considerado um sistema de coordenadas local de G. Esse sistema de coordenadas local é denominado de sistema de coordenadas de primeira espécie.

Um outro tipo de sistema de coordenadas nas vizinhanças do elemento neutro, obtida por exponenciais, é dada pela seguinte aplicação: tome uma base $\{X_1, X_2, \ldots, X_n\}$ de $\mathfrak{g}$ e considere a aplicação $\psi : \mathbb{R}^n \to G$ dada por $\psi(t_1, t_2, \ldots, t_n) = e^{t_1 X_1} \ldots e^{t_n X_n}$. ψ é chamada de sistema de coordenadas do segundo tipo (ou sistema de coordenadas de segunda espécie.)

Para introduzir uma métrica Riemanniana invariante à esquerda em G tome um produto interno qualquer $\langle \cdot, \cdot \rangle_e$ em $T_eG = \mathfrak{g}$ e defina

$$\langle u, v \rangle_x = \langle (dL_{x^{-1}})_x u, (dL_{x^{-1}})_x v \rangle_e,\ x \in G,\ u, v \in T_xG.$$

Note que se G estiver munido com uma métrica invariante à esquerda, a exponencial definida por campos invariantes à esquerda nem sempre coincide com a exponencial definida por geodésicas (Cf. Seção 1.4)

Uma representação de G em um espaço vetorial V é um homomorfismo $\rho : G \to Gl(V)$, onde $Gl(V)$ é o grupo das transformações lineares invertíveis de V. O espaço V é chamado de espaço da representação e $\dim V$ sua dimensão.

A seguir um importante teorema que nos facilita muito em termos de cálculo. Este resultado nos permite restringir as contas a álgebras de matrizes e a partir daí estender para qualquer grupo de Lie que tenha álgebra real de dimensão finita.

Teorema 1.42 (Teorema de Ado)**.** *Toda álgebra de Lie real de dimensão finita admite uma representação fiel (isto é, injetora), também de dimensão finita.*

1.6 Ações à esquerda de grupos

Trataremos brevemente sobre ações de grupos, mais especificamente ações à esquerda que é o que nos vai interessar mais tarde. Definiremos os conceitos básicos apenas.

Definição 1.43. *Diz-se que um grupo G age em uma variedade diferenciável M se existe uma aplicação $\varphi : G \times M \to M$ tal que:*

1. *Para cada $g \in G$, a aplicação $\varphi_g : M \to M$ dada por $\varphi_g(p) = \varphi(g, p)$, $p \in M$ é um difeomorfismo, e $\varphi_e = id_M$.*

2. *Se $g_1, g_2 \in G$, $\varphi_{g_1 g_2} = \varphi_{g_1} \circ \varphi_{g_2}$.*

φ é chamada de ação e quando estamos tratando apenas com uma única ação é usual indicar $\varphi(g, p) = gp$.

Definição 1.44. *Dado $p \in M$, sua órbita por G, denotada por $G \cdot p$ ou Gp, é definida como sendo o conjunto*

$$Gp = \{gp \in M : g \in G\}.$$

Mais geralmente, se $A \subset G$ então $Ap = \{gp : g \in A\}$, isto é, $Ap = \varphi(A, p)$.

Cada órbita é uma classe de equivalência da relação "$p \sim q$ se existe $g \in G$ tal que $q = gp$".

Definição 1.45. *O conjunto $G_p = \{g \in G : gp = p\}$ dos elementos de G que fixam p é denominado subgrupo de isotropia de p.*

Definição 1.46. *Uma ação de G em M é dita livre se os subgrupos de isotropia se reduzem ao elemento neutro de G, isto é, se $gp = p$ para algum $p \in M$, então $g = e$.*

Definição 1.47. *Dizemos que uma ação é propriamente descontínua se todo $p \in M$ possui uma vizinhança $U \subset M$ tal que $U \cap g(U) = \emptyset$ para todo $g \neq e$.*

1.7 Série de Baker-Campbell-Hausdorff (BCH)

Seja G um grupo de Lie com álgebra de Lie $\mathfrak{g}$. A aplicação exponencial $\exp : \mathfrak{g} \to G$ se restringe a um difeomorfismo $\exp : V \to U$ de uma vizinhança aberta da origem $0 \in V \subset \mathfrak{g}$ em uma vizinhança aberta da identidade $e \in U \subset G$.

Se $X, Y \in V$ são suficientemente pequenos então o produto $\exp(X)\exp(Y) = e^X e^Y$ ainda é um elemento de U, o que permite escrever

$$e^X e^Y = e^{c(X,Y)}.$$

A aplicação c é a expressão do produto em G em coordenadas locais do primeiro tipo.

A série de Baker-Campbell-Hausdorff fornece uma expressão para $c(X, Y)$ como a soma de uma série, cujos termos são colchetes sucessivos entre X e Y. Esta série é escrita como

$$c(X,Y) = X + Y + \sum_{n \geq 2} c_n(X,Y)$$

em que os termos $c_n(X, Y)$, $n \geq 2$, são homogêneos de ordem n, isto é, são somas de colchetes com n fatores (X ou Y). Por exemplo, os primeiros termos da fórmula com colchetes de campos invariantes à esquerda são

$$\begin{aligned} c_2(X,Y) &= \frac{1}{2}[X,Y] \\ c_3(X,Y) &= \frac{1}{12}[[X,Y],Y] - \frac{1}{12}[[X,Y],X]. \end{aligned}$$

A fórmula para a série BCH é universal, no sentido em que as expressões dadas para os termos homogêneos são as mesmas, independente do grupo de Lie.

Devido ao Teorema 1.42 a prova da convergência da série BCH é basicamente utilizando cálculo com exponenciais.

Faremos apenas considerações envolvendo matrizes. O primeiro membro da igualdade $e^X e^Y = e^{c(X,Y)}$ se escreve como a soma da série

$$e^X e^Y = \left(\sum_{n \geq 0} \frac{1}{n!} X^n \right) \left(\sum_{n \geq 0} \frac{1}{n!} Y^n \right) = \sum_{n \geq 0} \left(\sum_{j=0}^{n} \frac{1}{j!} \frac{1}{(n-j)!} X^j Y^{n-j} \right) = \sum_{n \geq 0} e_n(X,Y).$$

Essa última série converge normalmente para quaisquer X, Y, pois isso ocorre com a série de potências da exponencial (a convergência é com relação à uma norma pré-estabelecida).

Considere agora a série do logaritmo

$$\log(1+x) = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k} x^k,$$

que converge absolutamente se $|x| < 1$. Essa série inverte a exponencial, ou seja,

$$\log(\exp x) = \log(1 + (\exp x - 1)) = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k} (\exp x)^k = x.$$

Proposição 1.48 (Fórmula BCH)**.** *Existe $\rho > 0$ tal que se $|X|, |Y| < \rho$ então $c(X,Y)$ é dado pela série convergente $\sum_{n\geq 0} c_n(X,Y)$ em que o termo $c_n(X,Y)$ é um polinômio homogêneo em X, Y da forma*

$$c_n(X,Y) = \sum a_{I,J} X^{i_1} Y^{j_1} \ldots X^{i_s} Y^{j_s},$$

com $n = i_1 + j_1 + \cdots + i_s + j_s$ e $I = (i_1, \ldots, i_s)$, $J = (j_1, \ldots, j_s)$.

Explicitamente, dados $X, Y \in \mathfrak{g}$ temos que

$$\log(e^X e^Y) = X + Y + \frac{1}{2}, [X,Y] + \frac{1}{12}[X,[X,Y]] + \frac{1}{12}[Y,[Y,X]] + \ldots, \tag{1.2}$$

onde os termos são dados como diversos colchetes de X e Y.

1.8 Distribuições

Uma boa referência sobre a teoria de distribuições é [9].

Definição 1.49 (Distribuição)**.** *Uma distribuição $\mathcal{H}$ em uma variedade diferenciável M de dimensão n é uma aplicação que a cada $q \in M$ associa um subespaço $\mathcal{H}_q \subset T_qM$, do espaço tangente a q.*

Definição 1.50 (Distribuição diferenciável)**.** *Uma distribuição $\mathcal{H}$ é dita diferenciável em $q \in M$ se existem campos de vetores diferenciáveis $X_1, X_2, \ldots, X_k$ definidos em uma vizinhança de q que são:*

1. *tangentes à $\mathcal{H}$;*

2. *$\{X_1(q), X_2(q), \ldots, X_k(q)\}$ gera $\mathcal{H}_q$.*

Definição 1.51. *Uma geometria sub-riemanniana em uma variedade M consiste de uma distribuição $\mathcal{H} \subset TM$ no fibrado tangente de M junto com um produto interno $\langle\cdot,\cdot\rangle$ diferenciável nesta distribuição. Diferenciável no seguinte sentido: Se X e Y são campos diferenciáveis e horizontais à $\mathcal{H}$ então $\langle X, Y\rangle$ é diferenciável. A distribuição $\mathcal{H}$ é um subfibrado vetorial do fibrado tangente TM.*

Definição 1.52. *Uma curva $\gamma : [a,b] \subset \mathbb{R} \to M$ numa variedade diferenciável M é dita absolutamente contínua no intervalo $[a,b]$ se para qualquer número real $\varepsilon > 0$ existe um número $\delta > 0$ tal que*

$$\sum_{i=1}^{n} d(\gamma(x_i'), \gamma(x_i)) < \varepsilon,$$

sempre que

$$\sum_{i=1}^{n} |x_i' - x_i| < \delta,$$

onde $\{(x_i, x_i')\}$ é uma coleção de intervalos não sobrepostos contidos em $[a,b]$, g é uma métrica Riemanniana qualquer e d é a métrica associada a g.

A Proposição 1.27 nos garante que a definição anterior faz sentido, já que dadas duas métricas riemannianas quaisquer elas são Lipschitz equivalentes na imagem da curva, o que permite concluir que se a curva é absolutamente contínua com relação a uma das métricas Riemannianas também será com relação à outra.

Definição 1.53. *Dada uma distribuição $\mathcal{H} \subset TM$, uma curva absolutamente contínua $\gamma : I \subset \mathbb{R} \to M$ é chamada horizontal (à distribuição) se $\gamma'(t) \in \mathcal{H}_{\gamma(t)}$ sempre que $\gamma'(t)$ existe.*

Podemos através do produto interno definido na distribuição induzir uma função de comprimento para cada curva absolutamente contínua γ que seja horizontal à distribuição em quase toda parte, ou seja, $\gamma'(t) \in \mathcal{H}_{\gamma(t)}$ quase sempre. Para tais curvas o comprimento é dado por $\ell(\gamma) = \int \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt$, onde $\langle \cdot, \cdot \rangle$ é calculado em $\mathcal{H}_{\gamma(t)}$.

Usamos o comprimento para definir a distância sub-riemanniana, $d(p,q)$, entre os pontos $p, q \in M$

$$d(p,q) = \inf\{\ell(\gamma)\},$$

onde o ínfimo é tomado sob todos os caminhos absolutamente contínuos e horizontais que ligam p e q e caso não haja um caminho em tais condições dizemos que a distância é infinita.

Definição 1.54 (Distribuição completamente não-holonômica)**.** *Uma distribuição $\mathcal{H} \subset TM$ é chamada de completamente não-holonômica se para todo $p \in M$ e qualquer referencial X_i de $\mathcal{H}$ em uma vizinhança de p, temos que $X_i(p)$ e os colchetes de Lie iterados $[X_i, X_j](p)$, $[X_i, [X_j, X_k]](p)$, ..., geram T_pM.*

Estamos utilizando, até agora, a notação $\mathcal{H}$ para nos referir a uma distribuição em M. Vamos utilizar a mesma notação para o conjunto de todos os campos diferenciáveis de vetores tangentes a $\mathcal{H}$. Deste modo, temos que para cada aberto $U \subset M$ é atribuída uma coleção $\mathcal{H}_U$ de todos os campos de vetores diferenciáveis horizontais definidos em U. Note que já estamos trabalhando com propriedades locais visto que esses campos não precisam necessariamente estarem definidos para todo M, apenas em U.

Os colchetes de Lie de campos de vetores em $\mathcal{H}$ geram o que chamamos de *flag* de subfeixes

$$\mathcal{H} \subset \mathcal{H}^2 \subset \cdots \subset \mathcal{H}^r \subset \cdots \subset TM,$$

com $\mathcal{H}^2 = \mathcal{H} + [\mathcal{H}, \mathcal{H}]$, $\mathcal{H}^{r+1} = \mathcal{H}^r + [\mathcal{H}, \mathcal{H}^r]$, onde $[\mathcal{H}, \mathcal{H}^k] = \text{span}\{[X, Y]; X \in \mathcal{H}, Y \in \mathcal{H}^k\}$ e o gerado é tomado sobre $\mathcal{D}(M)$. Isto significa, que $\mathcal{H}^2$ é gerado pelos campos de vetores em $\mathcal{H}$ e seus colchetes de dois termos $[X, Y]$, $\mathcal{H}^3$ adiciona a isso os colchetes de três termos, e assim por diante.

No caso de M compacto, dizer que $\mathcal{H}$ é completamente não-holonômica nada mais é que dizer que existe um $r \in \mathbb{N}$ tal que $\mathcal{H}^r = TM$. Como o que faremos daqui em diante se encaixa nesse caso, suporemos a existência desse r supracitado.

Se tomarmos um ponto $p \in M$ o flag de subfeixes determina um flag de subespaços de TpM

$$\mathcal{H}_p \subset \mathcal{H}_p^2 \subset \cdots \subset \mathcal{H}_p^r = T_pM.$$

Definição 1.55. *O vetor $(n_1(p), n_2(p), \ldots, n_r(p))$, onde $n_i(p) = \dim \mathcal{H}_p^i$,$i = 1, \ldots, r$, é chamado vetor crescimento de $\mathcal{H}$ em p. Chamamos de grau de não-holonomia da distribuição em p o menor inteiro $r = r(p)$ tal que $\mathcal{H}_p^r = T_pM$.*

Note que as dimensões $n_i(p)$ podem variar conforme variamos o ponto p, o que nos leva à seguinte definição.

Definição 1.56. *Uma distribuição $\mathcal{H}$ em uma variedade diferenciável M^n é chamada regular em um ponto $p \in M$ se o vetor crescimento é constante em uma vizinhança de p.*

Note que na definição acima $n_1 = k$ é a dimensão da distribuição e $n_r = n$ é a dimensão da variedade.

Definição 1.57. *A função endpoint associada a uma distribuição $\mathcal{H}$ em M com base no ponto $p \in M$ é uma aplicação que leva cada curva horizontal começando em p para o seu ponto final.*

É possível, apesar de não ser feito aqui, provar que a função endpoint é diferenciável (Cf. [9], Apêndice E).

Definição 1.58. *Se $\mathcal{H}$ é uma distribuição (não necessariamente completamente não-holonômica) em M e $p \in M$, então o conjunto acessível associado a $\mathcal{H}$ em p, denotado por $Acc(p)$, é a imagem da função endpoint com base em p.*

Capítulo 2

Estimativas locais da aplicação exponencial

2.1 Fórmula de Baker-Campbell-Hausdorff com resto

Nesta seção utilizamos uma fórmula clássica, a Fórmula de Taylor com resto, aplicada na fórmula de BCH. Isso nos permite “truncar” a série de BCH fazendo com que possamos trabalhar apenas com seus termos inciais, ou seja, seus termos iniciais são uma boa aproximação.

Como já observado, a função exponencial $\exp : \mathfrak{g} \to G$ define um sistema canônico de coordenadas do primeiro tipo em uma vizinhança U da identidade $e \in G$. Denote o sistema de coordenadas por $\mathbf{x} = (x_1, \ldots, x_n)$. Introduzimos uma norma euclidiana arbitrária $|\cdot|$ no espaço vetorial $\mathfrak{g}$. Podemos assumir que $V = \exp^{-1}(U) = U(0, \delta)$ é uma bola aberta em $(\mathfrak{g}, |\cdot|)$ centrada em 0. Transferimos a estrutura local de grupo de $U \subset G$ para V através de $\exp^{-1}$.

Pela fórmula de Baker-Campbell-Hausdorff existe um número $\lambda > 0$, $\lambda < \delta$, de tal forma que se X e Y pertencem à bola $B(0, \lambda)$ de centro 0 e raio λ na métrica induzida pela norma euclidiana $|\cdot|$ temos

$$XY = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}\left([[X, Y], Y] - [[X, Y], X]\right) + \sum_{n \geq 4} c_n(X, Y), \tag{2.1}$$

onde os c_n são dados como na Proposição 1.48. Observe que todos os termos da série apresentam pelo menos um termo em X e um Y, já que o colchete de um campo com ele mesmo é o campo nulo.

Caso necessário, restrinja $\mathbf{x} : A \subset \mathbb{R}^n \to \mathfrak{g}$ de forma que $V \subset \mathbf{x}(A)$. Logo, $B(0, \lambda) \subset \mathbf{x}(A)$.

Para simplificar a notação neste trecho vamos indicar vetores de $\mathbb{R}^n$ por x'. Assim $\mathbf{x}(x') = \mathbf{x}(x_1, x_2, \ldots, x_n) = X \in \mathfrak{g}$.

Defina a aplicação produto por $p : B(0, \lambda) \times B(0, \lambda) \subset \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ onde $p(X, Y) = XY = \exp^{-1}(\exp(X)\exp(Y))$. A expressão de $p(X, Y)$ é dada pela fórmula de BCH e é um difeomorfismo de classe C^∞. Em particular, é de classe C^3 que é o que nos basta.

Consideremos a expressão de p nas coordenadas de $\mathbf{x}$.

$$p' = (\mathbf{x})^{-1} \circ p \circ (\mathbf{x} \times \mathbf{x}) : W \times W \subset \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n,$$

onde $W = (\mathbf{x})^{-1}(B(0,\lambda))$ e

$$p'(x',y') = (x_1+y_1,\ldots,x_n+y_n) + \frac{1}{2}\left(\sum_{i,j=1}^{n} a_{ij}^1 x_i y_j, \ldots, \sum_{i,j=1}^{n} a_{ij}^n x_i y_j\right) + f(x',y'). \tag{2.2}$$

Lembrando que se $X = \sum_{i=1}^{n} x_i X_i$ e $Y = \sum_{j=1}^{n} y_j X_j$, onde $\{X_1,\ldots,X_n\}$ é a base associada à parametrização $\mathbf{x}$, então

$$\begin{aligned}[X,Y] &= \left[\sum_{i=1}^{n} x_i X_i, \sum_{j=1}^{n} y_j X_j\right] = \sum_{i,j=1}^{n} x_i y_j [X_i,X_j] \\ &= \sum_{i,j=1}^{n} x_i y_j \left(\sum_{k=1}^{n} a_{ij}^k X_k\right) = \sum_{k=1}^{n}\left(\sum_{i,j=1}^{n} a_{ij}^k x_i y_j\right) X_k\end{aligned}$$

onde os a_{ij}^k são as constantes de estrutura do colchete de Lie. Isto justifica o segundo termo de (2.2).

Em (2.2), $f(x',y')$ é a série $\sum_{n\geq 3} c_n(X(x_1,\ldots,\mathbf{x}_n), Y(y_1,\ldots,y_n))$.

Recordemos agora da Fórmula de Taylor com resto de 2ª ordem.

Teorema 2.1 (Fórmula de Taylor com resto de 2ª ordem)**.** *Seja $f : A \subset \mathbb{R}^n \to \mathbb{R}$ uma função de classe C^3, em que A é um aberto e $P_0 \in A$. Então,*

$$f(P_0+h) = f(P_0) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(P_0)h_j + \frac{1}{2}\sum_{j,i=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(P_0)h_j h_i + R_2(P_0,h),$$

com $h = (h_1,\ldots,h_n) \in \mathbb{R}^n$,

$$R_2(P_0,h) = \sum_{j,i,l=1}^{n} \int_0^1 \frac{\partial^3 f}{\partial x_j \partial x_i \partial x_l}(P_0+th)h_j h_i h_l \frac{(t-1)^2}{2} dt.$$

Note que

$$\lim_{h\to 0} \frac{|R_2(P_0,h)|}{||h||^2} = 0.$$

A função produto p' expressa nas coordenadas não satisfaz as hipóteses do Teorema 2.1, para tanto teremos que considerar as funções coordenadas, e nelas aplicar o teorema. Em termos mais precisos para cada $k = 1,\ldots,n$ e $\pi_k : \mathbb{R}^n \to \mathbb{R}$ a projeção canônica consideremos

$$p'_k = \pi_k \circ p' : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}.$$

Por (2.2), temos a expressão explícita de p'_k,

$$p'_k(x',y') = x_k + y_k + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}^k x_i y_j + \pi_k(f(x',y')). \tag{2.3}$$

Como p é de classe C^3, sua expressão em coordenadas p' é C^3 também, e por consequência suas funções coordenadas o são. Assim podemos aplicar a Fórmula de Taylor para cada p'_k, onde $P_0 = (0,0) \in \mathbb{R}^n \times \mathbb{R}^n$, $h = (x', y')$ e fazemos a identificação natural $\mathbb{R}^n \times \mathbb{R}^n \simeq \mathbb{R}^{2n}$. Para adequar a notação à identificação feita indiquemos

$$z_i = \begin{cases} x_i, & \text{se } 1 \leq i \leq n \\ y_i, & \text{se } n+1 \leq i \leq 2n \end{cases}$$

Então,

$$p'_k((0,0)+(x',y')) = p'_k(0,0) + \sum_{j=1}^{2n} \frac{\partial p'_k}{\partial z_j}(0,0) z_j + \frac{1}{2}\sum_{i,j=1}^{2n} \frac{\partial^2 p'_k}{\partial z_j \partial z_i}(0,0) z_j z_i + R_{2,k}((0,0),(x',y')).$$

Logo,

$$p'_k(x',y') = x_k + y_k + \frac{1}{2}\sum_{i,j=1}^{n} a^k_{ij} x_i y_j + R_{2,k}(x',y'),$$

onde

$$|R_{2,k}(x',y')| \leq \sum_{j,i,l=1}^{2n} \int_0^1 \left| \frac{\partial^3 p'_k(tx',ty')}{\partial z_j \partial z_i \partial z_l} \right| |z_j||z_i||z_l| \frac{(t-1)^2}{2} dt \tag{2.4}$$

Nos termos do somatório em que $i,j,l \leq n$ teríamos que trocar z_j, z_i e z_l por x_j, x_i e x_l, respectivamente, e isso faria os termos se anularem. Analogamente para o caso em que $i,j,l \geq n+1$, trocaríamos por y_l, y_i e y_j e anularíamos os termos. Portanto, todos os termos são nulos quando $i,j,l \leq n$ ou quando $i,j,l \geq n+1$. Isso quer dizer que em qualquer termo sempre teremos dois índices menores ou iguais a n e um índice maior ou igual a $n+1$, ou vice-e-versa.

Faremos uma mudança de índices introduzindo a notação z_s, onde s pode ser i, j ou l seguindo a seguinte escolha: Para fixar ideias, suponha que num dos termos tenhamos $i,l \leq n$ e $j \geq n+1$, escolhemos um índice em cada caso, por exemplo, i e l e majoramos $|z_i| \leq ||x'||_{Euc} = |X|$ e $|z_j| \leq ||y'||_{Euc} = |Y|$, onde $||\cdot||_{Euc}$ é a norma Euclidiana do $\mathbb{R}^n$ que coincide com a norma em $\mathfrak{g}$. Portanto nesse caso $z_s = z_l$. Assim podemos reescrever a desigualdade (2.4) como

$$|R_{2,k}(x',y')| \leq |X||Y| \sum_{j,i,l=1}^{2n} \int_0^1 \left| \frac{\partial^3 p'_k(tx',ty')}{\partial z_j \partial z_i \partial z_l} \right| |z_s| \frac{(t-1)^2}{2} dt.$$

Podemos majorar $|z_s|$ por $|z_s| \leq |z'|$, onde $|z'| = \max_{W'}\{||x'||_{Euc}, ||y'||_{Euc}\}$.

Considerando uma restrição de p' para um fechado $W' \times W' \subset W \times W$, temos que para cada $k = 1,\ldots,n$, p'_k é definida num fechado e limitado, logo compacto, de $\mathbb{R}^n \times \mathbb{R}^n$ e de classe C^3. Então $\frac{\partial^3 p'_k}{\partial z_j \partial z_i \partial z_l}$ é limitada, ou seja,

$$\left| \frac{\partial^3 p'_k(tx',ty')}{\partial z_j \partial z_i \partial z_l} \right| \leq C'; \quad C' > 0, \ \ (x',y') \in W' \times W',\ t \in [0,1] \ \text{ e } \ i,j,l \in \{1,\ldots,2n\}.$$

Então, temos que

$$|R_{2,k}(x',y')| \leq |X||Y| \sum_{j,i,l=1}^{2n} C'|z'| \int_0^1 \frac{(t-1)^2}{2} dt = |X||Y| 8n^3 C'|z'| \frac{1}{6}.$$

Decorre que $|R_{2,k}(x', y')| < \varepsilon |X||Y|$ se $\frac{4C'n^3}{3}|z'| < \varepsilon$, ou seja, $|z'| < \frac{3\varepsilon}{4C'n^3}$. Logo, dado ε, existe $\alpha = \frac{3\varepsilon}{4C'n^3}$ de forma que se $|X|, |Y| < \alpha$ então

$$|R_{2,k}(x', y')| < \varepsilon |X||Y|.$$

Agora, por (2.2), temos que $R_2(x', y') = (R_{2,1}(x', y'), R_{2,2}(x', y'), \ldots, R_{2,n}(x', y'))$ e portanto, corrigindo α se necessário, devido à equivalência das normas em $\mathbb{R}^n$, se $|X|, |Y| < \alpha$, então

$$|R_2(X, Y)| = ||R_2(x', y')||_\infty = \max_{k=1}^{n} \{R_{2,k}(x', y')\} < \varepsilon |X||Y|.$$

Portanto, para cada número $\varepsilon > 0$ e $\lambda \in (0, \delta)$ temos que se $X, Y \in \mathfrak{g}$, $|X| < \lambda$, $|Y| < \lambda$, então

$$XY = X + Y + \frac{1}{2}[X, Y] + R_2(X, Y), \quad |R_2(X, Y)| < \varepsilon |X||Y|.$$

Consequentemente, existe um número positivo C tal que para esses mesmos X e Y temos

$$XY = X + Y + z(X, Y), \quad |z(X, Y)| < C|X||Y|. \tag{2.5}$$

Mudando a norma proporcionalmente (através de homotetias), se for necessário, podemos assumir que $C = 1$, o que nos facilita daqui em diante.

2.2 Consequências de BCH

A partir de agora provaremos alguns lemas que nos levarão a provar resultados importantes mais tarde. Estamos considerando, a partir de agora sempre elementos na vizinhança $V = U(0, \lambda)$ tomada na Seção 2.1.

Lema 2.2. $|XY - (X + Y)| < |X|\,|Y|$.

Demonstração. Decorre diretamente de (2.5) depois de considerado $C = 1$, pois

$$|XY - (X + Y)| = |z(X, Y)| < |X|\,|Y|.$$

■

Imediatamente segue

Lema 2.3. $|XY| < |X| + |Y| + |X|\,|Y|$.

Demonstração. Utilizando a desigualdade triangular e o Lema 2.2 temos que

$$|XY| = |XY - (X + Y) + (X + Y)| \leq |XY - (X + Y)| + |X + Y| < |X|\,|Y| + |X| + |Y|.$$

■

Lema 2.4. *Para $|X|$ e $|Y|$ suficientemente pequenos ,*

$$X + Y = XY',$$

onde $|Y - Y'| < |X|\,|Y'|$, $|Y - Y'| < 2|X|\,|Y|$.

Demonstração. Vamos mostrar a existência de Y'. Note que a aplicação $m : G \times G \to G$ dada por $m(X,Y) = XY$ é tal que $d_e(m(e,\cdot))$ e $d_e(m(\cdot,e))$ são isomorfismos. Daí $d_X(m(X,\cdot))$ é um isomorfismo para $|X|$ suficientemente pequeno. Logo, pelo Teorema da Função Inversa, $m(X,\cdot)$ é um difeomorfismo local e isso mostra a existência de Y', tal que $XY' = X + Y$ para todo $|Y|$ suficientemente pequeno.

Devido ao Lema 2.2

$$|Y - Y'| = |(X + Y) - (X + Y')| = |XY' - (X + Y')| < |X|\,|Y'|.$$

Vamos estimar $|Y'|$. Temos que

$$|Y'| = |Y + Y' - Y| \leq |Y| + |Y' - Y| < |Y| + |X|\,|Y'|.$$

Consequentemente, $|Y'| - |X|\,|Y'| < |Y|$. Então $|Y'| < \dfrac{|Y|}{(1 - |X|)} < 2|Y|$ se $|X| < \dfrac{1}{2}$. Daí chegamos à desigualdade

$$|Y - Y'| < |X|\,|Y'| < 2|X|\,|Y|.$$

■

Lema 2.5. *Para $|X|$, $|Y|$ suficientemente pequenos $YX = XY'$, $|Y - Y'| < 4|X|\,|Y|$.*

Demonstração. Aplicando o Lema 2.2 para YX, temos

$$YX = Y + X + z(Y,X) = XY' \;\Rightarrow\; Y + z(Y,X) = XY' - (X + Y) + Y$$

$$\Rightarrow |Y + z(Y,X)| \leq |XY' - (X + Y)| + |Y| < |X|\,|Y| + |Y| = |Y|(1 + |X|).$$

Pelo Lema 2.4

$$|Y' - (Y + z(Y,X))| < 2|X|\,|Y + z(Y,X)| < 2|X|\,|Y|(1 + |X|).$$

Consequentemente,

$$\begin{aligned}
|Y - Y'| = |Y' - Y| &= |Y' - (Y + z(Y,X)) + z(Y,X)| \\
&\leq |Y' - (Y + z(Y,X))| + |z(Y,X)| \\
&< 2|X|(|Y| + |X|\,|Y|) + |X|\,|Y| \\
&= 2|X|\,|Y| + 2|X|^2|Y| + |X|\,|Y| \\
&< 4|X|\,|Y|, \text{ se } |X| < \frac{2}{3}.
\end{aligned}$$

■

Lema 2.6. *Para números suficientemente pequenos*

$$|X_1 X_2 \ldots X_j| < e^{\eta} - 1 \sim \eta$$

se $|X_i| < \frac{\eta}{j}$, $i = 1, \ldots, j$.

Demonstração. Introduzimos a notação $Y_m = X_1X_2 \dots X_m$.

Mostremos, por indução sobre m, que

$$|Y_m| < \left(1 + \frac{\eta}{j}\right)^m - 1.$$

Para $m = 1$ temos que $|Y_1| = |X_1| < \frac{\eta}{j} = \left(1 + \frac{\eta}{j}\right)^1 - 1$.

Suponha que $|Y_m| < \left(1 + \frac{\eta}{j}\right)^m - 1$. Então para $m+1$, devido ao Lema 2.3, temos que

$$|Y_{m+1}| = |Y_m X_{m+1}| < |Y_m| + |X_{m+1}| + |Y_m|\,|X_{m+1}|$$

$$< |Y_m| + \frac{\eta}{j} + |Y_m|\frac{\eta}{j} = |Y_m|\left(1 + \frac{\eta}{j}\right) + \frac{\eta}{j}.$$

Pela hipótese de indução,

$$|Y_{m+1}| < \left(\left(1 + \frac{\eta}{j}\right)^m - 1\right)\left(1 + \frac{\eta}{j}\right) + \frac{\eta}{j}$$

$$= \left(1 + \frac{\eta}{j}\right)^{m+1} - \left(1 + \frac{\eta}{j}\right) + \frac{\eta}{j}.$$

Assim,

$$|Y_{m+1}| < \left(1 + \frac{\eta}{j}\right)^{m+1} - 1. \tag{2.6}$$

Consequentemente

$$|X_1 \dots X_j| < \left(1 + \frac{\eta}{j}\right)^j - 1 < e^\eta - 1.$$

A última desigualdade decorre de que se $t \in \left[1, 1 + \frac{\eta}{j}\right]$, temos

$$\frac{1}{t} \leq 1 \;\Rightarrow\; \int_1^{1+\frac{\eta}{j}} \frac{1}{t}dt \leq \int_1^{1+\frac{\eta}{j}} dt$$

$$\Rightarrow\; \ln\left(1 + \frac{\eta}{j}\right) \leq \frac{\eta}{j} \;\Rightarrow 1 + \frac{\eta}{j} \leq e^{\frac{\eta}{j}}.$$

Portanto,

$$\left(1 + \frac{\eta}{j}\right)^j \leq e^\eta.$$

■

Lema 2.7. *Para números η, $\varepsilon > 0$ suficientemente pequenos,*

$$\prod_{i=1}^{j}(X_i + Y_i) = \left(\prod_{i=1}^{j} X_i\right) Y, \quad |Y| < 13\varepsilon\eta e^\eta, \tag{2.7}$$

se $|X_i| \leq \frac{\eta}{j}$ e $|Y_i| \leq \varepsilon\frac{\eta}{j}$.

Demonstração. Para $i = 1, \ldots, j$, pelo Lema 2.4

$$X_i + Y_i = X_i Y_i', \quad |Y_i - Y_i'| < 2|X_i|\,|Y_i|.$$

Então

$$\begin{aligned} |Y_i'| &= |Y_i' - Y_i + Y_i| \leq |Y_i' - Y_i| + |Y_i| < 2|X_i|\,|Y_i| + |Y_i| \\ &< 2\frac{\eta}{j}\varepsilon\frac{\eta}{j} + \varepsilon\frac{\eta}{j} = \left(\varepsilon\frac{\eta}{j}\right)\left(1 + 2\frac{\eta}{j}\right) < 3\varepsilon\frac{\eta}{j}. \end{aligned}$$

Desde que $\frac{\eta}{j} < 1$. Então,

$$\prod_{i=1}^{j}(X_i + Y_i) = \prod_{i=1}^{j}(X_i Y_i'), \quad |Y_i'| < 3\varepsilon\frac{\eta}{j}.$$

Pode-se escrever essa última expressão na forma

$$\prod_{i=1}^{j}(X_i Y_i') = \prod_{i=1}^{j} X_i \prod_{i=1}^{j} Y_i'' = \left(\prod_{i=1}^{j} X_i\right) Y,$$

onde $Y_j'' = Y_j'$ e Y_i'', $1 \leq i \leq j$, é definido por

$$Y'\left(\prod_{m=i+1}^{j} X_m\right) = \left(\prod_{m=i+1}^{j} X_m\right) Y_i''.$$

Para se convencer que essa forma vale, observe que

$$\begin{aligned} \prod_{i=1}^{j} X_i \prod_{i=1}^{j} Y_i'' &= X_1\left(\prod_{m=2}^{j} X_m\right) Y_1'' Y_2'' \ldots Y_j'' = X_1 Y_1' \left(\prod_{m=2}^{j} X_m\right) Y_2'' \ldots Y_j'' \\ &= X_1 Y_1' X_2 \left(\prod_{m=3}^{j} X_m\right) Y_2'' \ldots Y_j'' = X_1 Y_1' X_2 Y_2' \left(\prod_{m=3}^{j} X_m\right) Y_3'' \ldots Y_j'' \\ &= \cdots = X_1 Y_1' X_2 Y_2' \ldots X_j X_j'' = \prod_{i=1}^{j}(X_i Y_j'). \end{aligned}$$

Estimamos agora $|Y_i''|$. Pelo Lema 2.5 e pela desigualdade (2.6) na prova do Lema 2.6,

$$|Y_i' - Y_i''| < 4\left|\prod_{m=i+1}^{j} X_m\right| |Y_i'|, \tag{2.8}$$

$$\begin{aligned} |Y_i''| &= |Y_i'' - Y_i' + Y_i'| \leq |Y_i'' - Y_i'| + |Y_i'| < 4\left|\prod_{m=i+1}^{j} X_m\right| |Y_i'| + |Y_i'| \\ &< 3\varepsilon\frac{\eta}{j}\left(1 + 4\left((1 + \eta/j)^{j-i} - 1\right)\right) < 3\varepsilon\frac{\eta}{j}\left(1 + 4\left(e^{\eta} - 1\right)\right) \\ &= 12\varepsilon\frac{\eta}{j}e^{\eta} - 9\varepsilon\frac{\eta}{j} < 12\varepsilon\frac{\eta}{j}e^{\eta}. \end{aligned}$$

Usando o Lema 2.6 e a desigualdade (2.8), temos

$$|Y| = \left| \prod_{i=1}^{j} Y_i'' \right| < e^{12\varepsilon\eta e^{\eta}} - 1 \sim 12\varepsilon\eta e^{\eta}.$$

Consequentemente, para ε suficientemente pequeno

$$|Y| < 13\varepsilon\eta e^{\eta},$$

que é o que queríamos provar. ∎

Será fundamental daqui para a frente conseguirmos relacionar a operação do grupo com suas características métricas e topológicas. É o que faremos no Lema 2.11 onde provamos que o produto de bolas num grupo topológico localmente compacto com métrica intrínseca invariante à esquerda é ainda uma bola, e temos o controle sobre o raio desta bola resultante do produto. Isto nos será bastante útil como elo entre a operação do grupo e várias desigualdades desejadas.

Antes disso vejamos a definição de grupo topológico localmente compacto.

Definição 2.8. *Seja (G, ρ) um grupo topológico. G é dito localmente compacto se existe uma vizinhança U de $e \in G$ tal que $\overline{U}$ é compacto.*

Lema 2.9. *Todo grupo topológico localmente compacto (G, ρ) com métrica ρ invariante à esquerda é completo.*

Demonstração. Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de Cauchy em G e seja $U \in G$ uma vizinhança de $e \in G$ tal que $\overline{U}$ é compacto, que existe pela compacidade local de G.

Tome $\varepsilon > 0$ tal que $B_\varepsilon(e) \subset U$. Como (x_n) é de Cauchy então para o ε tomado existe um certo índice $N \in \mathbb{N}$ tal que se $n, m \geq N$ então $\rho(x_n, x_m) < \varepsilon$. Consequentemente, todos os termos da sequência (x_n), a partir do índice N, estão na bola $B_\varepsilon(x_N)$. Aplicando a translação à esquerda $L_{x_N^{-1}} : G \to G$, que é um difeomorfismo em um espaço homogêneo, temos que

$$L_{x_N^{-1}}(B_\varepsilon(x_N)) = B_\varepsilon(e),$$

pois a métrica é invariante à esquerda.

Logo, a sequência $\left(L_{x_N^{-1}}(x_n)\right)_{n\geq N}$ é de Cauchy no conjunto compacto (e completo como espaço métrico) $\overline{B_\varepsilon(e)}$, portanto possui limite, digamos $x \in \overline{B_\varepsilon(e)}$.

Assim, $x_n \to L_{x_N}(x)$, o que conclui a demonstração. ∎

Omitiremos a demonstração do próximo teorema, mas ela pode ser encontrada em [2].

Teorema 2.10. *Seja (G, ρ) um espaço métrico localmente compacto e completo, com métrica intrínseca. Então este espaço é estritamente intrínseco: para quaisquer $x, y \in G$ tais que $\rho(x, y) < \infty$ existe um caminho de menor comprimento γ ligando x a y, isto é, uma curva $\gamma : [a, b] \to G$ tal que $\gamma(a) = x$, $\gamma(b) = y$ e $\ell(\gamma) = \rho(x, y)$.*

Lema 2.11. *Seja (G, ρ) um grupo topológico localmente compacto com métrica intrínseca invariante à esquerda ρ. Então*

$$B_{r_1+r_2} = B_{r_1}B_{r_2},$$

onde B_r é uma bola fechada de raio r em (G, ρ) com centro na identidade $e \in G$, e do lado direito da igualdade temos o produto de conjuntos com respeito a operação do grupo G.

Demonstração. Pelo Lema 2.9, temos que (G, ρ) é completo e pelo Teorema 2.10 temos que qualquer dois pontos em (G, ρ) podem ser ligados por um caminho de menor comprimento.

Sejam $x \in B_{r_1}$ e $y \in B_{r_2}$. Então, devido à invariância à esquerda de ρ, temos

$$\rho(xy, e) \leq \rho(xy, x) + \rho(x, e) = \rho(x^{-1}xy, x^{-1}x) + \rho(x, e) = \rho(y, e) + \rho(x, e) \leq r_1 + r_2.$$

Portanto, $B_{r_1}B_{r_2} \subset B_{r_1+r_2}$

Provemos a inclusão contrária. Seja $z \in B_{r_1+r_2}$, então podemos ligar z à e por um caminho de comprimento mínimo γ. Denotemos por x o ponto de γ (aqui pensamos o caminho como a imagem da aplicação que o define) tal que $\rho(e, x) = r_1$ e por definição tomemos $y = x^{-1}z$. Então novamente pela invariância à esquerda

$$\rho(e, y) = \rho(e, x^{-1}z) = \rho(x, z) \leq (r_1 + r_2) - r_1 = r_2,$$

ou seja, $z = xy$, $x \in B_{r_1}$, $y \in B_{r_2}$, isto é, $z \in B_{r_1}B_{r_2}$ que é o que tínhamos que provar.

■

Capítulo 3

Métrica de Carnot-Carathéodory-Finsler

3.1 Teorema de Chow e Teorema de ball-box

Nesta seção apresentaremos dois resultados intimamente ligados, Teorema de ball-box e Teorema de Chow. O primeiro é utilizado para a demonstração do segundo e será usado mais adiante neste capítulo. O Teorema de Chow tem muita importância no sentido que ele prova, sob certas condições, a existência de caminhos horizontais ligando quaisquer dois pontos de uma variedade sub-riemanniana munida com uma distribuição diferenciável, o que nos garante que métricas que dependam exclusivamente destes caminhos em sua construção sejam finitas. Isto ficará claro mais adiante ao tratarmos das métricas de Carnot-Carathéodory.

Teorema 3.1 (Teorema de Chow). *Se $\mathcal{H}$ é uma distribuição completamente não-holonômica em uma variedade conexa M então quaisquer dois pontos de M podem ser ligados por um caminho horizontal.*

Como já dito, esta demonstração exige o Teorema de ball-box e, portanto, será realizada no final desta seção. Nesta direção apresentamos a motivação do Teorema de ball-box e a seguir seu enunciado e sua demonstração.

Seja M uma variedade diferenciável munida com uma geometria sub-riemanniana, onde a distribuição correspondente é completamente não-holonômica. Sejam X_1 e X_2 campos de vetores diferenciáveis com fluxo (local) $\Phi_i(t) = \exp(tX_i)$. Usando o fato de que $\Phi_i^{-1}(t) = \Phi_i(-t)$ e pela Proposição 1.22, temos, para t suficientemente pequeno, que $\Phi_1(t) \circ \Phi_2(t) \circ \Phi_1^{-1}(t) \circ \Phi_2^{-1}(t)(q) = q + t^2[X_1, X_2](q) + o(t^2)$ em um sistema qualquer de coordenadas, onde $o(t^2)/t^2 \to 0$ quando $t \to 0$. Por brevidade escrevemos $[\Phi_1(t), \Phi_2(t)] = \Phi_1(t) \circ \Phi_2(t) \circ \Phi_1^{-1}(t) \circ \Phi_2^{-1}(t)$.

Fixe um ponto base $q_0 \in M$ e escolha um referencial local ortonormal X_i, $i = 1, 2, \ldots, k$, para nossa distribuição $\mathcal{H}$. Seja Φ_i o correspondente fluxo local. Podemos usá-los para nos mover facilmente em direções horizontais, para t pequeno,

$$\Phi_i(t)(q) = q + tX_i(q) + o(t^2).$$

Vamos chamar as curvas $t \mapsto \Phi_i(t)(q)$ de curvas horizontais simples.

Note que o comprimento de uma curva horizontal simples com $0 \leq t \leq \varepsilon$ é $\varepsilon + o(t^2)$.

Se considerarmos $\Phi_k(t_k) \circ \ldots \Phi_2(t_2) \circ \Phi_1(t_1)(q_0)$ e fazendo os t_i's variarem sobre o k-cubo $|t_i| \leq \varepsilon$ nós varremos um cubo k-dimensional curvilinear tangente à $\mathcal{H}_{q_0}$ em $t = 0$ com volume k-dimensional aproximadamente $(2\varepsilon)^k$. Cada ponto sobre este cubo é o ponto final de uma concatenação de k ou menos curvas simples: primeiro viaja ao longo de $\Phi_1(t)(q_0)$ de 0 até t_1, determina $q_1 = \Phi_1(t_1)$, então viaja ao longo de $\Phi_2(t)(q_1)$ de $t = 0$ até $t = t_2$, e assim por diante. Cada uma dessas curvas horizontais simples tem comprimento menor ou igual a ε. Então o cubo situa-se em uma bola sub-riemanniana de raio $k\varepsilon$.

Podemos mover em direções $\mathcal{H}^2/\mathcal{H}$ ao longo de caminhos horizontais pela aplicação do comutador $\Phi_{ij}(t) = [\Phi_i(t), \Phi_j(t)]$ para q_0.

Agora $\Phi_{ij}(t)(q_0) = q_0 + t^2[X_i, X_j](q_0) + o(t^2)$, então se restringirmos para $|t| \leq \varepsilon$ nos moveremos uma quantidade euclidiana ε^2 aproximadamente na direção de $\mathcal{H}^2/\mathcal{H}$.

Continuamos o processo de tomar o comutador e colchetes até termos esgotado o espaço tangente, visto que estamos trabalhando com uma distribuição completamente não-holonômica. A implementação desta ideia é principalmente uma questão de notação. Para os multi-índices $I = (i_1, i_2, \ldots, i_m)$, $1 \leq i_j \leq k$, defina os campos de vetores X_I indutivamente por $X_I = [X_{i_1}, X_J]$ onde $J = (i_2, \ldots, i_m)$. Podemos escrever $i_1 J = I$ e denotar o comprimento de um multi-índice I por $|I|$. Similarmente definimos os fluxos Φ_I por $\Phi_I = [\Phi_{i_1}(t), \Phi_J(t)]$.

Observe que $\Phi_I(t) = I + t^m X_I + \mathcal{O}(t^{m+1})$.

Pela suposição de $\mathcal{H}$ ser completamente não-holonômica podemos selecionar um referencial local para o fibrado tangente todo dentre os X_I: Escolhemos um tal referencial e o renomeamos de Y_i, $i = 1, \ldots, n$: $\{Y_1 = X_1, \ldots, Y_k = X_k\}$ gera $\mathcal{H}$ próximo de q_0; $\{Y_1, \ldots, Y_{n_2}\}$ gera $\mathcal{H}^2$ próximo de q_0; $\{Y_1, \ldots, Y_{n_2}, \ldots, Y_{n_3}\}$ gera $\mathcal{H}^3$ próximo de q_0; e assim por diante, onde $(k, n_2, n_3, \ldots, n_r)$ é o vetor crescimento da distribuição em q_0.

Para cada Y_i escolhido da forma X_I, seja w_i o comprimento $|I|$. Assim $w_i = m$ se, e somente se, $Y_i \in \mathcal{H}^m$ e $Y_i \notin \mathcal{H}^{m-1}$. A atribuição $i \mapsto w_i$ é chamada de ponderação associada ao vetor crescimento.

Similarmente, renomeamos os fluxos Φ_I como Φ_i, $i = 1, \ldots, n$, para aqueles multi-índices aparecendo na construção do nosso referencial Y_i. Agora para cada ponto $\Phi_i(t)(q_0) = \Phi_I(t)(q_0)$ é o ponto final de um caminho horizontal consistindo na concatenação de $C(w_i)$ caminhos horizontais simples, cada um com comprimento t, onde $C(w_i)$ conta os caminhos concatenados de cada iteração, ou seja, $C(1) = 1$ pois $\Phi_1(t) = \Phi_1(t)$, $C(2) = 4$ pois $\Phi_{12}(t) = \Phi_1(t) \circ \Phi_2(t) \circ \Phi_1^{-1}(t) \circ \Phi_2^{-1}(t)$, $C(3) = 10$ pois $\Phi_{123}(t) = \Phi_3(t) \circ \Phi_1(t) \circ \Phi_2(t) \circ \Phi_1^{-1}(t) \circ \Phi_2^{-1}(t) \circ \Phi_3^{-1}(t) \circ \Phi_1^{-1}(t) \circ \Phi_2^{-1}(t) \circ \Phi_1(t) \circ \Phi_2(t)$, $C(4) = 22$, e assim sucessivamente. Então se restringirmos t para $|t| \leq \varepsilon$ então $\Phi_i(t)(q_0)$ está na bola de raio $C(w_i)\varepsilon$ centrada em q_0.

Por outro lado, uma vez que $\Phi_i(t)(q_0) = q_0 + t^{w_i} Y_i(q_0) + \mathcal{O}(t^{w_i+1})$, esse ponto reside na caixa euclidiana cujas dimensões são de ordem ε^{w_i} nas direções $\mathcal{H}^{w_i}$. Isto sugere que a bola sub-riemanniana de raio ε contem uma caixa euclidiana cujos lados são de ordem ε^{w_i} na i-ésima direção coordenada. Este resultado é chamado de teorema de ball-box e será provado em parte depois da seguinte definição.

Definição 3.2 (Coordenadas linearmente adaptadas)**.** *Coordenadas $y_1, y_2, \ldots, y_n$ são ditas linearmente adaptadas à distribuição $\mathcal{H}$ em q_0 se $\mathcal{H}^i(q_0)$ é anulado pelas diferenciais $dy_{n_i+1}, \ldots, dy_n$ em q_0, onde $n_i = n_i(q_0)$ são as coordenadas do vetor crescimento em q_0.*

A caixa w-ponderada de tamanho ε é o conjunto de pontos

$$Box^w(\varepsilon) = \{y \in \mathbb{R}^n : \ |y_i| \leq \varepsilon^{w_i}, i = 1, \ldots, n\}.$$

Denotaremos os pontos da variedade diferenciável M pela expressão em coordenadas linearmente adaptadas y_i.

Por definição, as coordenadas y_i centradas em q_0 para as quais $dy_i(q_0)$ são as bases para $Y_i(q_0)$ são as mais simples coordenadas linearmente adaptadas e as utilizaremos no teorema.

Teorema 3.3 (Teorema de ball-box). *Para uma distribuição $\mathcal{H}$, existem coordenadas adaptadas $y_1, \ldots, y_n$ e constantes positivas $c < C$ e $\varepsilon_0 > 0$ tais que para todo $\varepsilon < \varepsilon_0$*

$$Box^w(c\varepsilon) \subset B(\varepsilon, q_0) \subset Box^w(C\varepsilon),$$

onde $B(\varepsilon, q_0)$ denota a bola sub-riemanniana de raio ε centrada em q_0.

Demonstração. Provaremos apenas que $Box(c\varepsilon) \subset B(\varepsilon, q_0)$, onde os y_i são coordenadas linearmente adaptadas. A prova de que $B(\varepsilon, q_0) \subset Box(C\varepsilon)$ pode ser encontrada em [9], na seção 2.6.

Precisamos contornar primeiramente o fato de que na aproximação $\Phi_{ij}(t) = I + t^2[X_i, X_j] + \ldots$ o termo importante, o colchete, tem coeficiente t^2 que é sempre positivo e não nos permite usá-lo para mover na direção negativa de $[X_i, X_j]$. Resolvemos isso pela troca da ordem dos comutadores. Defina

$$\Psi_{ij}(t) = \begin{cases} [\Phi_i(t), \Phi_j(t)], & t \geq 0 \\ [\Phi_j(t), \Phi_i(t)], & t < 0. \end{cases}$$

Então

$$\Psi_{ij}(t) = \begin{cases} I + t^2 X_{ij} + \mathcal{O}(t^3), & t \geq 0 \\ I - t^2 X_{ij} + \mathcal{O}(t^3), & t < 0. \end{cases}$$

O mesmo problema ocorre sempre que o número w_i do colchete de Lie é par. Usamos o mesmo truque para contornar isso. Assim, para $|I|$ par temos

$$\Psi_I(t) = \begin{cases} \Phi_I(t), & t \geq 0 \\ [\Phi_J(t), \Phi_{i_1}(t)], & t < 0, \end{cases}$$

onde $I = i_1 J$ como antes. Quando $|I|$ é ímpar mantemos $\Psi_I = \Phi_I$. Renomeamos então Ψ_I como Ψ_i, $i = 1, \ldots, n$, de acordo com a maneira que rotulamos o X_I como Y_i.

Considere as funções

$$\sigma_i(t) = \begin{cases} t^{w_i}, & w_i \text{ par}, t \geq 0; \\ -t^{w_i}, & w_i \text{ par}, t < 0; \\ t^{w_i}, & w_i \text{ ímpar}. \end{cases}$$

Então $\Psi_i(t) = I + \sigma_i(t) Y_i + \mathcal{O}(t^{w_i+1})$.

Defina a aplicação $F : \mathbb{R}^n \to M$ por $F(t_1, \ldots, t_n) = \Psi_n(t_n) \circ \cdots \circ \Psi_1(t_1) = (y_1, \ldots, y_n)$. Assim,

$$\begin{array}{rcl} Y_1 & = & (1 + \mathcal{O}(||y||), \ldots, \mathcal{O}(||y||), \ldots, \mathcal{O}(||y||)) \\ \vdots & & \qquad\qquad\qquad \vdots \\ Y_i & = & (\mathcal{O}(||y||), \ldots, 1 + \mathcal{O}(||y||), \ldots, \mathcal{O}(||y||)) \\ \vdots & & \qquad\qquad\qquad \vdots \\ Y_n & = & (\mathcal{O}(||y||), \ldots, \mathcal{O}(||y||), \ldots, 1 + \mathcal{O}(||y||)). \end{array}$$

Note que
$$\Psi_i(t_i) = I + \sigma_i(t_i)Y_i + \mathcal{O}(t_i^{w_i+1}).$$

Resta entender e limitar os termos do erro e isso é mais facilmente feito pela introdução de novas variáveis $s_i = \sigma_i(t_i)$. Note que

$$\Psi_i(s_i) = I + s_iY_i + o(s_i).$$

Nessas novas variáveis F é dada por $y_i = s_i + o(s_i)$. De fato,

$$\begin{aligned}
&\Psi_n(s_n) \circ \cdots \circ \Psi_1(s_1)(0, \ldots, 0)\\
&= \Psi_n(s_n) \circ \cdots \circ \Psi_2(s_2)\Big((0, \ldots, 0) + s_1(1 + \mathcal{O}(s), \mathcal{O}(s), \ldots, \mathcal{O}(s)) + (o(s), \ldots, o(s))\Big)\\
&= \Psi_n(s_n) \circ \cdots \circ \Psi_2(s_2)(s_1 + o(s), o(s), \ldots, o(s))\\
&= \Psi_n(s_n) \circ \cdots \circ \Psi_3(s_3)\Big((s_1 + o(s), \ldots, o(s)) + s_2(\mathcal{O}(s), 1 + \mathcal{O}(s), \ldots, \mathcal{O}(s)) + (o(s), \ldots, o(s))\Big)\\
&= \Psi_n(s_n) \circ \cdots \circ \Psi_3(s_3)\big(s_1 + o(s), s_2 + o(s), o(s), \ldots, o(s)\big)\\
&\vdots\\
&= (s_1 + o(s), s_2 + o(s), \ldots, s_n + o(s)).
\end{aligned}$$

A função $\sigma = (\sigma_1, \ldots, \sigma_n)$ é de classe C^1, mas não C^∞, e trata-se de um homeomorfismo entre vizinhanças de 0 em $\mathbb{R}^n$.

Escreva $S : \mathbb{R}^n \to \mathbb{R}^n$ para a inversa de σ de modo que $S = (s_1, \ldots, s_n)$ com $s_i = \pm|t|^{1/w_i}$. Então $y \circ F \circ S(s_1, \ldots, s_n) = (s_1, \ldots, s_n) + o(|S|)$. Daí segue que $F \circ S$ é diferenciável na origem, com derivada a identidade nas $s - y$ coordenadas. É também C^1 próximo de 0.

Pelo Teorema da função inversa, $F \circ S(Box^w(\varepsilon_0))$ é uma vizinhança de 0 com fecho compacto para todo ε_0. Logo existem $c(\varepsilon_0)$ e $C(\varepsilon_0)$ tais que

$$Box^w(c\varepsilon_0) \subset F \circ S(Box^w(\varepsilon_0)) \subset Box^w(C\varepsilon_0).$$

Mas $S(Box^w(\varepsilon_0)) \subset Box(\varepsilon_0)$. Logo

$$Box^w(c\varepsilon) \subset F(Box(\varepsilon)) \subset Box^w(C\varepsilon).$$

Cada curva em $F(Box(\varepsilon))$ é o ponto final de uma curva horizontal iniciada em q_0 da qual o comprimento é menor que $M\varepsilon$, onde $M = M(w)$ conta o número de concatenações envolvidas em F (por exemplo, Ψ_{123} envolve a concatenação de 10 curvas horizontais simples, e Ψ_{1234} envolve 22). Consequentemente $F(Box(\varepsilon)) \subset B(M\varepsilon, q_0)$, que produz a inclusão desejada $Box^w(c\varepsilon) \subset B(\varepsilon, q_0)$. ■

Podemos agora demonstrar o teorema de Chow:

Demonstração do Teorema de Chow. A consequência da metade do teorema de ball-box que provamos é que o conjunto acessível de p, que é o conjunto dos pontos ligados a p por caminhos horizontais, é uma vizinhança de p. Resta mostrar que este conjunto acessível é toda nossa variedade (conexa). Seja B qualquer outro ponto da nossa variedade M Ligue p à q por qualquer caminho diferenciável γ. A imagem de γ é compacta, então podemos cobri-la por uma quantidade finita sucessiva de vizinhanças "caixa" para as quais o teorema de ball-box vale. Chame essas vizinhanças $U_1, U_2, \ldots, U_m$, com os U_i centrados em pontos sucessivos p_i ao longo de γ, $p_1 = p$ e $p_m = q$. Podemos escolher estas vizinhanças tais que $U_i \cap U_{i+1} \neq \emptyset$ e tomar $q_i \in U_i \cap U_{i+1}$ em γ. Pelo teorema de ball-box nós temos caminhos "caixa" conectando p_i à q_i e q_i à p_{i+1}. Concatenando esses caminhos temos um caminho horizontal ligando p à q. ■

3.2 Métrica de Carnot-Carathéodory-Finsler

Seja M uma variedade diferenciável de dimensão n munida com uma geometria sub-riemanniana, onde a distribuição correspondente $\mathcal{H}$ é completamente não-holonômica e diferenciável, de k-planos (regular com dimensão $\dim \mathcal{H} = k$), ou seja, a distribuição $\mathcal{H}$ associa cada ponto $m \in M$ a um subespaço k-dimensional de $T_m M$.

Trabalharemos de agora em diante apenas com curvas absolutamente contínuas. Em [11] prova-se que curvas absolutamente contínuas tem derivada em quase toda parte. Uma curva absolutamente contínua α em M é dita horizontal se é tangente a distribuição $\mathcal{H}$ em quase toda parte.

Definição 3.4. *Por distância de Carnot-Carathéodory entre os pontos p, $q \in M$ se entende*

$$d_c(p,q) = \inf_{\omega \in C_{p,q}} \{\ell(\omega)\} \tag{3.1}$$

onde $C_{p,q}$ é o conjunto de curvas horizontais ligando p e q e $\ell(\omega)$ é o comprimento de ω.

Pelo Teorema 3.1 temos que qualquer dois pontos de uma variedade conexa M, munida com uma distribuição completamente não-holonômica, podem ser ligados por um caminho continuamente diferenciável por partes tangente a essa distribuição.

Como consequência disso, no nosso caso, a métrica d_c é finita. Pode-se provar ainda mais:

1. d_c é uma métrica intrínseca em M.

   De fato, para verificar que d_c é intrínseca basta verificar que ela foi obtida como a função distância associada a uma estrutura de comprimento. Mas basta tomar a classe dos caminhos admissíveis como todos os caminhos horizontais e a função comprimento ℓ como sendo função comprimento induzida pela métrica g. É fácil ver que a classe de caminhos admissíveis assim tomada satisfaz as propriedades:

   (a) ser fechada para restrições;

   (b) fechada para concatenações;

   (c) fechada para reparametrizações lineares.

Assim,

$$d_c(p,q) = \inf_{\omega \in C_{p,q}} \{\ell(\omega)\}.$$

2. A topologia métrica da métrica d_c é equivalente à topologia original de M.

   Lembremos que a métrica original da variedade Riemanniana é dada por

$$g(x,y) = \inf_{\gamma} \left\{ \int_a^b \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt : \gamma : [a,b] \to M, \gamma(a) = x, \gamma(b) = y \right\}.$$

   Então as métricas são quase a mesma exceto pela classe de caminhos admissíveis em que o ínfimo é tomado. Sejam $B_g(x,\varepsilon)$ e $B_{d_c}(x,\varepsilon)$ as bolas abertas de centro x e raio ε em (M,g) e (M,d_c), respectivamente.

   Note que $B_{d_c}(x,\varepsilon) \subset B_g(x,\varepsilon)$, pois $g \leq d_c$ já que o conjunto dos caminhos horizontais está contido no conjunto de todos os caminhos diferenciáveis.

   Pelo Teorema de ball-box (Teorema 3.3) temos que existe um número $k = k(\varepsilon) > 0$ tal que $B_g(x,k) \subset B_{d_c}(x,\varepsilon)$.

3. Se a variedade Riemanniana (M,g) é completa então (M,d_c) é também um espaço métrico completo.

   Esse fato decorre do fato das topologias coincidirem, logo, (M,g) é completa se, e somente se, (M,d_c) é completa.

Neste caso, devido ao Teorema de Chow (Teorema 3.1) quaisquer dois pontos p e q em (M,d_c) podem ser ligados por um caminho minimizante γ. Mais do que isso, se γ é parametrizado por comprimento de arco s, então $\gamma(s)$ será uma curva absolutamente contínua e horizontal.

Provaremos os resultados no caso especial de grupos de Lie com distribuição $\mathcal{H}$ invariante à esquerda.

É necessário notar que na definição de d_c, $C_{p,q}$ consiste de curvas absolutamente contínuas horizontais. Se obtém o mesmo resultado se (M,d_c) consiste da classe menor de curvas horizontais continuamente diferenciáveis por partes (ligando os pontos p e q).

Ao mesmo tempo não sabemos se uma curva minimizante em (M,d_c) parametrizada por comprimento de arco será uma curva continuamente diferenciável (ou pelo menos continuamente diferenciável por partes). d_c depende apenas dos valores do tensor métrico g na distribuição $\mathcal{H}$.

Precisamos da versão Finsler da métrica de Carnot-Carathéodory. Seja M uma variedade diferenciável de dimensão n munida com uma k-distribuição diferenciável completamente não holonômica $\mathcal{H}$ $(k \leq n)$.

Para cada $p \in M$, em $\mathcal{H}_p$ é dada uma norma $\mathcal{F}_p$ tal que para qualquer campo de vetores contínuo horizontal X em M a função $\mathcal{F}(X)(p) = \mathcal{F}_p(X(p))$ é contínua. Então a expressão (3.1) define, como diremos, uma métrica de Carnot-Carathéodory-Finsler.

Entende-se que o comprimento da curva ω é medido com respeito a função métrica $\mathcal{F}$, definida no subfibrado k-dimensional de TM.

Estamos interessados no caso especial dessa métrica da seguinte forma. Como $\mathcal{H}$ toma-se um distribuição invariante à esquerda no grupo de Lie conexo $G(= M)$ tal que o subespaço vetorial $L_0 = \mathcal{H}_e$ (e sendo a identidade do grupo G) da álgebra de Lie $\mathfrak{g}$ do grupo G gera a álgebra de Lie $\mathfrak{g}$.

Aqui $\mathcal{F}$ é uma norma invariante à esquerda na distribuição $\mathcal{H}$ definida por uma norma F_0 em L_0 no sentido que $\mathcal{F} = F_0 \circ \theta$, onde θ é diferencial da translação à esquerda, ou seja, é uma 1-forma em G, com valores em $\mathfrak{g}$, calculada pela forma $\theta(u) = dL_{g^{-1}}(u)$, $u \in T_gG$, onde $L_{g^{-1}}$ é a translação à esquerda em G pelo elemento g^{-1}.

Teorema 3.5. *Sejam G um grupo de Lie conexo, L_0 um subespaço vetorial da álgebra de Lie $\mathfrak{g}$, que gera $\mathfrak{g}$, F_0 uma norma arbitrária em L_0. Então a fórmula*

$$d_c(g,h) = \inf_X \int_0^1 F_0(\theta(X'(t)))dt, \tag{3.2}$$

onde o ínfimo é tomado sobre todos os caminhos horizontais continuamente diferenciáveis por partes $X = X(t)$, $0 \leq t \leq 1$, em G ligando g e h, define uma métrica (Carnot-Carathéodory-Finsler) intrínseca invariante à esquerda em G que é compatível com a topologia de G.

Demonstração. Provaremos a equivalência da topologia métrica τ_c de d_c e a topologia original τ em G.

Como L_0 pode ser visto como um espaço vetorial podemos definir, graças à equivalência das normas em espaços vetoriais de dimensão finita, dois produtos internos $\langle \cdot, \cdot \rangle_1$ e $\langle \cdot, \cdot \rangle_2$ em L_0 tais que para todos os vetores $u \in L_0$ tenhamos $\langle u, u \rangle_2 \leq F_0(u) \leq \langle u, u \rangle_1$. Sejam g_1 e g_2 os tensores métricos invariantes à esquerda em G correspondendo à $\langle \cdot, \cdot \rangle_1$ e $\langle \cdot, \cdot \rangle_2$, respectivamente. Logo, para qualquer caminho horizontal continuamente diferenciável por partes $X = X(t)$, $0 \leq t \leq 1$, em G ligando g e h temos que

$$\sqrt{\langle \theta(X'(t)), \theta(X'(t)) \rangle_2} \leq F_0(\theta(X'(t))) \leq \sqrt{\langle \theta(X'(t)), \theta(X'(t)) \rangle_1}.$$

Portanto

$$d_{g_2}(g,h) \leq d_c(g,h) \leq d_{g_1}(g,h),$$

onde d_{g_1} e d_{g_2} são as métricas sub-riemannianas associadas aos tensores g_1 e g_2, respectivamente.

Assim, para quaisquer $\delta > 0$ e $g_0 \in G$,

$$B_{d_{g_1}}(\delta, g_0) \subset B_{d_c}(\delta, g_0) \subset B_{d_{g_2}}(\delta, g_0), \tag{3.3}$$

onde $B_{d_{g_i}}(\delta, g_0)$ denota a bola sub-riemanniana com relação à métrica d_{g_i}, $i = 1, 2$, centrada em g_0 e raio δ e $B_{d_c}(\delta, g_0)$ denota a bola sub-riemanniana com relação à métrica d_c centrada em g_0 e raio δ.

Aplicando o Teorema de ball-box (Teorema 3.3) para as variedades sub-riemannianas (G, Δ, g_1) e (G, Δ, g_2), existem coordenadas $y_1, y_2, \ldots, y_n$ linearmente adaptadas à distribuição Δ e constantes positivas $c_1 < C_1$, $c_2 < C_2$ e $\varepsilon_1, \varepsilon_2 > 0$ tais que para todo $\varepsilon < \min\{\varepsilon_1, \varepsilon_2\}$

$$\mathrm{Box}(c_1\varepsilon) \subset B_{d_{g_1}}(\varepsilon, g_0) \subset \mathrm{Box}(C_1\varepsilon) \tag{3.4}$$

$$\mathrm{Box}(c_2\varepsilon) \subset B_{d_{g_2}}(\varepsilon, g_0) \subset \mathrm{Box}(C_2\varepsilon). \tag{3.5}$$

Assim, de (3.3), (3.4) e (3.5) temos que

$$\mathrm{Box}(c_1\varepsilon) \subset B_{d_{g_1}}(\varepsilon, g_0) \subset B_{d_c}(\varepsilon, g_0) \subset B_{d_{g_2}}(\varepsilon, g_0) \subset \mathrm{Box}(C_2\varepsilon). \tag{3.6}$$

Como as caixas $\mathrm{Box}(c_1\varepsilon)$ e $\mathrm{Box}(C_2\varepsilon)$ são identificadas por caixas n-dimensionais em $\mathbb{R}^n$, podemos tomá-las como abertos básicos de τ. Portanto, devido à (3.6) decorre que $\tau = \tau_{d_c}$.

$\blacksquare$

Capítulo 4

Métricas Intrínsecas Invariantes à Esquerda em Grupos de Lie

Neste capítulo nos encaminhamos para a caracterização das métricas intrínsecas invariantes à esquerda em grupos de Lie.

Teorema 4.1. **(i)** *Seja (G, ρ) um grupo de Lie conexo com métrica intrínseca invariante à esquerda ρ. Então existe um subespaço vetorial L_0 da álgebra de Lie $\mathfrak{g}$ que gera $\mathfrak{g}$, e uma norma F_0 em L_0 tal que ρ coincide com a métrica de Carnot-Carathéodory-Finsler definida por*

$$d_c(g,h) = \inf_X \int_0^1 F_0(\theta(X'(t)))dt.$$

O ínfimo é tomado sobre todos os caminhos horizontais continuamente diferenciáveis por partes $X = X(t)$, $0 \leq t \leq 1$, em G ligando os pontos g e h.

(ii) *A correspondência $\rho \to (L_0, F_0)$ é biunívoca.*

Antes de demonstrarmos o Teorema 4.1, demonstraremos alguns resultados preliminares.

Lema 4.2. *Considere $(U, |\cdot|)$ conforme definido nas Seções 2.1 e 2.2. Se $v \in \mathfrak{g}$ denote por X_v o campo invariante à esquerda correspondente a v. Então dado $\varepsilon > 0$, existe $r > 0$ tal que $|X_v(p) - X_v(e)| < \varepsilon$ para todo $p \in B_{|\cdot|}(e, r)$ e todo v com $|v| = 1$.*

Demonstração. Primeiramente note que $X_v(p) = \frac{d}{dt}(p\exp(tv))$. Escrevendo $p = \exp(w)$, temos que

$$p\exp(tv) = \exp(w)\exp(tv) = \exp(w + tv + z(w, tv)),$$

onde $|z(w, tv)| \leq |w||tv|$. Então

$$X_v(p) = \left.\frac{d}{dt}\right|_{t=0} v + \frac{z(w, tv)}{t}.$$

Logo, se $|w| = |p| < \varepsilon$, então $|X_v(p) - X_v(e)| < \varepsilon$. ∎

Lema 4.3. *Seja G um grupo de Lie e $(U, |\cdot|)$ conforme definido anteriormente. Seja $\gamma : [a,b] \to G$ um caminho de classe C^1 parametrizado por comprimento de arco em relação a $|\cdot|$, ou seja, tal que $F_0(\gamma'(t)) = 1$ para todo t. Então dado $\varepsilon > 0$, existe $\delta > 0$ tal que se $t \in [t_1, t_1+\delta)$, então $|\gamma(t_1)^{-1}\gamma(t) - \exp(\gamma'(t_1))(t-t_1)| < \varepsilon|t-t_1|$.*

Demonstração. Transladaremos a curva para a origem em $t = t_1$, ou seja, consideraremos a curva $\gamma(t_1)^{-1}\gamma(t)$. Considere $\bar{\delta} > 0$ tal que $|\theta(\gamma'(t_2)) - \theta(\gamma'(t_1))| < \varepsilon/2$ se $|t_2 - t_1| < \delta$. Seja $\delta = \min\{\bar{\delta}, r\}$, onde r é tomado conforme o lema 4.2, só que com ε substituído por $\varepsilon/2$. Fixe $t_1 \in [a,b]$. Considere a curva $\bar{\gamma}(t) := \int_{t_1}^t \theta(\gamma'(t))dt$ em U. Note que ele em geral não é a curva $\eta(t) := \gamma(t_1)^{-1}\gamma(t)$. A curva η em U é a curva integral de

$$\frac{d}{dt}\eta(t) = X_{\theta(\eta'(t))}(\eta(t)),$$

ou seja

$$\eta(t) = \int_{t_1}^t X_{\theta(\eta'(t))}(\eta(t))dt.$$

Pela escolha de δ, temos que $|X_{\theta(\eta'(t))}(\eta(t)) - X_{\theta(\eta'(t))}(e)| < \varepsilon/2$. Daí temos que

$$\begin{aligned} |\gamma(t_1)^{-1}\gamma(t) - \exp(\gamma'(t_1))(t-t_1)| &= |\eta(t) - \theta(\gamma'(t_1))(t-t_1)| \\ &\leq |\eta(t) - \bar{\gamma}(t)| + |\bar{\gamma}(t) - \theta(\gamma'(t_1))(t-t_1)| \\ \leq \left| \int_{t_1}^t X_{\theta(\eta'(t))}(\eta(t))dt - \int_{t_1}^t \theta(\gamma'(t))dt \right| &+ \left| \int_{t_1}^t \theta(\gamma'(t))dt - \int_{t_1}^t \theta(\gamma'(t_1))dt \right| \\ &< \varepsilon(t-t_1) \end{aligned}$$

sempre que $t \in [t_1, t_1+\delta)$. ∎

Corolário 4.4. *Nas condições do lema 4.3 temos que dado $\varepsilon > 0$, existe $\delta > 0$ tal que se $t \in [t_1, t_1+\delta)$, então $d(\gamma(t), \gamma(t_1)\exp(\gamma'(t_1))(t-t_1)) < \varepsilon|t-t_1|$, onde d é a métrica correspondente a uma métrica de Finsler invariante.*

Demonstração. É só notar que $|\cdot|$ e d são Lipschitz equivalentes. ∎

Proposição 4.5. *Para qualquer número positivo ε existe um caminho horizontal diferenciável por partes $X_\varepsilon = X_\varepsilon(t)$, $0 \leq t \leq 1$, com as seguintes propriedades:*

(a) *$\theta(X'_\varepsilon(t))$ é constante por partes;*

(b) *nos pontos comuns de existência dos tangentes X' e X'_ε temos*

$$|F_0(\theta(X'(t))) - F_0(\theta(X'_\varepsilon(t)))| < \varepsilon;$$

(c) *$X_\varepsilon(0) = g$ e $\rho(X_\varepsilon(1), h) < \varepsilon$.*

Demonstração. Fixe $\varepsilon > 0$.

Para fixar ideias, suponha que $X : [0,1] \to G$ seja de classe C^1 e parametrizado por comprimento de arco em relação a $(U, |\cdot|)$. A ideia da demonstração é particionar $[0,1]$ em k partes iguais, e para essa partição definir a curva $\gamma_k : [0,1] \to G$ dada por

$$\gamma_k(t) :=$$

$$\begin{cases} X(0)\exp(w_1 t) & \text{se} \quad t \in [0, \frac{1}{k}]; \\ X(0)\exp(\frac{w_1}{k})\exp(w_2(t-\frac{1}{k})) & \text{se} \quad t \in [\frac{1}{k}, \frac{2}{k}]; \\ \vdots & \vdots \qquad \vdots \\ X(0)\exp(\frac{w_1}{k})\exp(\frac{w_2}{k})\ldots\exp(\frac{w_{k-1}}{k})\exp(w_k(t-\frac{k-1}{k})) & \text{se} \quad t \in [\frac{k-1}{k}, 1], \end{cases}$$

onde $w_i = \theta(\gamma'(\frac{i-1}{k}))$. É claro que para k suficientemente grande, as condições a) e b) são satisfeitas. Resta mostrar a condição c). Para isso, sejam $v_1, v_2, \ldots, v_k \in \mathfrak{g}$ tais que $\exp(v_i) = X(\frac{i-1}{k})^{-1}X(\frac{i}{k})$. Então

$$X(1) = X(0)\exp(v_1)\exp(v_2)\ldots\exp(v_k).$$

Por sua vez,

$$\gamma_k(1) = X(0)\exp(w_1)\exp(w_2)\ldots\exp(w_k).$$

Devemos controlar $d(\gamma_k(1), X(1))$, ou seja, $d(e, X(1)^{-1}\gamma_k(1))$. Faremos as contas "a partir dos termos centrais".

$$X(1)^{-1}\gamma_2(1) = \ldots\exp(-v_2)\exp(-v_1)\exp(w_1)\exp(w_2)\ldots$$

$$= \ldots\exp(-v_2)\exp(w_1 - v_1 + z(w_1, -v_1))\exp(w_2)\ldots;$$

Pelo Lema 4.3, existe $\delta > 0$ suficientemente pequeno (ou k suficientemente grande) tal que $|v_1 - w_1| < \varepsilon/k$. Prosseguindo os cálculos:

$$= \ldots\exp(-v_2 + w_1 - v_1 + 2z(-v_1, -v_2, w_1))\exp(w_2)\ldots$$

onde

$$\begin{aligned} |z(-v_1, -v_2, w_2)| &\leq |v_2|\varepsilon/k + |v_1||v_2||w_1| + |w_1||v_1| \\ &\leq |v_2|\varepsilon/k + 2\max_i\{|v_i|\}.\max_i\{|w_i|\}. \end{aligned}$$

Prosseguindo, temos

$$\leq \exp(w_k - v_k + \ldots + w_1 - v_1 + z)$$

onde

$$|z| \leq \max_{i,j}\{|v_i|, |w_j|\}(k-1)\varepsilon + 2k\max_i\{|v_i|\}.\max_i\{|w_i|\}.$$

Logo, considerando que $|w_i| = 1/k$ e que $|v_i| \approx 1/k$, temos que

$$|w_k - v_k + \ldots + w_1 - v_1 + z| \leq 2\varepsilon + \frac{4}{k}$$

que converge a zero quando k tende a infinito. $\blacksquare$

Demonstração do Teorema 4.1. Demonstraremos *a priori* apenas a parte (i). A parte (ii) seguirá do Lema 4.6 e sua demonstração será feita, portanto, após o lema. **(i)** A aplicação exponencial $\exp : \mathfrak{g} \to G$ define um difeomorfismo de uma vizinhança V de 0 em $\mathfrak{g}$ em uma vizinhança U de e em G. Como no capítulo anterior, pela $\exp^{-1}$ transferimos a estrutura de grupo local e métrica ρ de U para V e vamos assumir que $V = U_{|\cdot|}(0,\lambda)$, $0 < \lambda < 1$ e o Lema 2.2 vale em V.

Para $r_0 > 0$ suficientemente pequeno temos que $B_{r_0}(e) \subset U$.

Para $r < r_0$ introduzimos a notação

$$B_r = \exp^{-1}(B_r(e)) \subset U_{|\,|}(0,\lambda). \tag{4.1}$$

Pelo Lema 2.11, para qualquer número natural k tem-se

$$B_{\frac{r}{2^k}} \subset \frac{1}{2^k} B_r, \tag{4.2}$$

pois

$$2^k B_{\frac{r}{2^k}} \subset B_{\frac{r}{2^k}} \dots B_{\frac{r}{2^k}} = B_r \;\Rightarrow\; B_{\frac{r}{2^k}} \subset \frac{1}{2^k} B_r.$$

Consequentemente, a sequência de conjuntos compactos $\beta_k = 2^k B_{\frac{r}{2^k}}$, $k = 1, 2, \dots$, não aumenta, isto é,

$$\beta_{k+1} = 2^{k+1} B_{\frac{r}{2^{k+1}}} = 2^{k+1} B_{\frac{r}{2.2^k}} \subset \frac{2^{k+1}}{2} B_{\frac{r}{2^k}} = \beta_k. \tag{4.3}$$

Definimos então

$$\beta = \bigcap_{k=1}^{\infty} \beta_k. \tag{4.4}$$

Provemos as seguintes afirmações:

1. β é um conjunto compacto centralmente simétrico de V com centro em $0 \in V$;

2. β é um conjunto convexo;

3. $L_0 = \bigcup\limits_{\lambda > 0} \lambda\beta$ é um subespaço vetorial da álgebra de Lie $\mathfrak{g}$ que a gera;

4. β é um compacto centralmente simétrico e convexo em β com centro em 0.

1. A compacidade de β segue do fato que β é a interseção de uma sequência não-crescente de conjuntos compactos β_k, $k = 1, 2, \dots$ (Que são compactos pois estamos considerando $B_{\frac{r}{2^k}}(e)$ como bolas fechadas). Cada bola $B_{\frac{r}{2^k}}$ junto com o ponto x também contém o ponto x^{-1} (isto é, $-x$). Consequentemente vale o mesmo para β_k e, por conseguinte, também β é centralmente simétrico, com centro em 0.

2. Devido à compacidade de β é suficiente provar que β contem juntamente com os pontos x e y seu ponto médio $\frac{x+y}{2}$. De fato, suponha que para qualquer par de pontos x, $y \in \beta$ o ponto $\frac{x+y}{2} \in \beta$ e suponha, por absurdo, que β não seja convexo. Então existem $x, y \in \beta$ e $t_0 \in (0,1)$ tais que $\alpha = (1-t_0)x + t_0 y \notin \beta$. Construímos uma sequência $(\alpha_n)_{n\in\mathbb{N}}$, pondo $\alpha_1 = \frac{x+y}{2}$, $\alpha_2 = \frac{x+\alpha_1}{2}$ se α está entre x e α_1 ou $\alpha_2 = \frac{\alpha_1+y}{2}$ se α_1 está entre α_1 e y. Para

fixar ideias, suponha que $\alpha_2 \in [x, \alpha_1]$. Se $\alpha \in [x, \alpha_2]$ faça $\alpha_3 = \frac{x+\alpha_2}{2}$; caso contrário faça $\alpha_3 = \frac{\alpha_1+\alpha_2}{2}$. Seguindo o raciocínio, colocando os pontos da sequência como o ponto médio do intervalo que contiver α teremos então a sequência $(\alpha_n)_{n\in\mathbb{N}}$ em β que, por construção, converge para $\alpha \notin \beta$, o que é um absurdo pelo fato de β ser compacto. Portanto para mostrar a convexidade é suficiente mostrar que para quaisquer dois pontos do conjunto o ponto médio também pertence ao conjunto.

Por definição de β para cada $k > 1$, $x, y \in \beta_k = 2^k B_{\frac{r}{2^k}}$, isto é, $\frac{x}{2^k}, \frac{y}{2^k} \in B_{\frac{r}{2^k}}$. Devido ao Lema 2.11, $\left(\frac{x}{2^k}\right)\left(\frac{y}{2^k}\right) \in B_{\frac{r}{2^k}} B_{\frac{r}{2^k}} = B_{\frac{r}{2^{k-1}}}$, isto é,

$$\left(\frac{x}{2^k}\right) + \left(\frac{y}{2^k}\right) + z\left(\frac{x}{2^k}, \frac{y}{2^k}\right) \in B_{\frac{r}{2^{k-1}}}, \tag{4.5}$$

onde pelo Lema 2.2 e por (4.1)

$$\left|z\left(\frac{x}{2^k}, \frac{y}{2^k}\right)\right| < \frac{\lambda^2}{2^{2k}} < \frac{1}{2^{2k}}. \tag{4.6}$$

Multiplicamos (4.5) por 2^{k-1}

$$z_k = \frac{x}{2} + \frac{y}{2} + 2^{k-1} z\left(\frac{x}{2^k}, \frac{y}{2^k}\right) \in \beta_{k-1}.$$

Por (4.6),

$$\left|z_k - \left(\frac{x}{2} + \frac{y}{2}\right)\right| < \frac{2^{k-1}}{2^{2k}} = \frac{1}{2^{k+1}}.$$

Como β_k forma uma sequência não-crescente de conjuntos compactos, segue das duas últimas relações que para qualquer número natural m

$$\frac{x}{2} + \frac{y}{2} = \lim_{k\to\infty} z_k \ \in 2^m B_{\frac{r}{2^m}} = \beta_m.$$

Isto é,

$$\frac{x}{2} + \frac{y}{2} \in \bigcap_{m=1}^{\infty} \beta_m = \beta.$$

3. Para um número ε, $0 < \varepsilon < 1$, introduzimos a notação $\beta_\varepsilon = \beta + \varepsilon \tilde{U}_\lambda$, onde $\tilde{U}_\lambda = \{x \in \mathfrak{g}; |x| < \lambda\}$. Obviamente β_ε é uma vizinhança aberta do conjunto β.

   Afirmamos que, devido a compacidade de β_k e (4.3) e (4.4), existe k_0 tal que para todo $k \geq k_0$ temos a inclusão

$$\beta_k \subset \beta + \varepsilon \tilde{U}_\lambda. \tag{4.7}$$

   Suponha, por absurdo, que exista ε_0, $0 < \varepsilon_0 < 1$, de forma que para todo $k \in \mathbb{N}$ temos $\beta_k \nsubseteq \beta_{\varepsilon_0} = \beta + \varepsilon_0 \tilde{U}_\lambda$. Então para: $k = 1$, existe $x_1 \in \beta_1$ tal que $x_1 \notin \beta + \varepsilon_0 \tilde{U}_\lambda$; $k = 2$, existe $x_2 \in \beta_2$ tal que $x_2 \notin \beta + \varepsilon_0 \tilde{U}_\lambda$; assim por diante.

   Desta forma, construímos uma sequência $(x_k)_{k\in\mathbb{N}}$ tal que $x_k \in \beta_k$ e $x_k \notin \beta + \varepsilon_0 \tilde{U}_\lambda$ para todo $k \in \mathbb{N}$. Em particular $x_k \notin \beta$.

Devido à compacidade de β_k, quando $k \to \infty$, existe uma subsequência $x_k \to a$, tal que $a \in \beta_k$ para todo k. Como $\beta = \bigcap_{k=1}^{\infty} \beta_k$, temos que $a \in \beta$. Mas como $x_k \notin \beta_{\varepsilon_0}$, para todo k, então

$$|x_k - a| \geq \varepsilon_0 \lambda,$$

o que é um absurdo e prova o afirmado.

Assim, para todo $k \geq k_0$, temos pela equação (4.7) que

$$B_{\frac{r}{2^k}} \subset \left(\frac{1}{2^k}\right)\left(\beta + \varepsilon \tilde{U}_\lambda\right).$$

Ao mesmo tempo, pelo Lema 2.11, $B_r = B_{\frac{r}{2^k}} B_{\frac{r}{2^k}} \dots B_{\frac{r}{2^k}}$, onde o lado direito é o produto de 2^k conjuntos idênticos.

Devido às últimas duas relações, cada elemento x de B_r pode ser representado na forma $x = (x_1 + y_1)(x_2 + y_2) \dots (x_{2^k} + y_{2^k})$, onde $x_i \in \left(\frac{1}{2^k}\right)\beta$, $y_i \in \left(\frac{\varepsilon}{2^k}\right)\tilde{U}_\lambda$, $i = 1, 2, \dots, 2^k$. Pelo Lema 2.7, para ε suficientemente pequeno temos a representação

$$x = \left(\prod_{i=1}^{2^k} x_i\right) y, \tag{4.8}$$

onde

$$x_i \in \left(\frac{1}{2^k}\right)\beta, \quad |y| < 13\varepsilon\lambda e^{\lambda}. \tag{4.9}$$

Em particular, o conjunto de produtos da forma

$$\prod_{i=1}^{2^k} x_i, \quad x_i \in \left(\frac{1}{2^k}\right)\beta, \quad k = 1, 2, \dots$$

é denso em B_r. Note que, pelo Lema 2.11, esses produtos residem em B_r. De fato, para todo $k = 1, 2, \dots$, $\beta \subset \beta_k = 2^k B_{\frac{r}{2^k}}$. Assim

$$\prod_{i=1}^{2^k} x_i \in \left(\frac{1}{2^k}\right)\beta \dots \left(\frac{1}{2^k}\right)\beta \subset \frac{1}{2^k}\left(B_{\frac{r}{2^k}}\right) \dots \frac{1}{2^k}\left(B_{\frac{r}{2^k}}\right) = B_{\frac{r}{2^k}} \dots B_{\frac{r}{2^k}} = B_{2^k \frac{r}{2^k}} = B_r.$$

Em virtude da convexidade e da simetria central de β o conjunto $L_0 = \bigcup_{\lambda > 0} \lambda\beta$ é um subespaço vetorial da álgebra de Lie $\mathfrak{g}$. Seja $\tilde{L}_0$ a álgebra de Lie gerada pelo espaço vetorial L_0. Então o conjunto $\tilde{L}_0 \cap B_r \subset V$ é fechado e todos os produtos $x_1 x_2 \dots x_{2^k}$ da forma indicada estão no conjunto $\tilde{L}_0 \cap B_r$. Consequentemente, devido a (4.8) e (4.9), $x \in B_r \cap \tilde{L}_0$, $B_r \subset B_r \cap \tilde{L}_0$. Isto só pode ser se $\tilde{L}_0 = \mathfrak{g}$.

4. Segue imediatamente de (1), (2) e (3).

A condição (4) permite definir a norma F_0 em L_0 pela fórmula usual de Minkowski:

$$F_0(x) = r \inf_{\alpha>0} \left\{ \frac{1}{\alpha};\ \alpha x \in \beta \right\}, \quad x \in L_0. \tag{4.10}$$

Pelo Teorema 3.5 podemos definir uma métrica intrínseca invariante à esquerda de Carnot-Carathéodory-Finsler d_c por (3.2). Resta provar que $\rho = d_c$.

Primeiro vamos provar que $d_c \leq \rho$. Seja $\rho(e, x) = r$. Para qualquer $\varepsilon > 0$, pelo que foi provado, existe um número natural k e uma representação $x = \left(\prod_{i=1}^{2^k} x_i\right) y$ satisfazendo (4.9).

Uma vez que $2^k x_i \in \beta$, segue da definição de F_0 que $F_0(x_i) \leq \frac{r}{2^k}$, $x_i \in L_0$.

Definimos o caminho $X = X(t) = \exp(tx_i)$, $0 \leq t \leq 1$. Então

$$\int_0^1 F_0(\theta(X'(t)))dt = \int_0^1 F_0(x_i)dt = F_0(x_i) \leq \frac{r}{2^k}.$$

Consequentemente, $d_c(e, \exp(x_i)) \leq \frac{r}{2^k}$. Devido à invariância à esquerda de d_c, temos

$$d_c(x,e) = d_c\left(\left(\prod_{i=1}^{2^k} x_i\right) y, e\right) \leq d_c\left(\left(\prod_{i=1}^{2^k} x_i\right) y, \prod_{i=1}^{2^k} x_i\right) + d_c\left(\prod_{i=1}^{2^k} x_i, e\right)$$

$$= d_c(y,e) + d_c\left(\prod_{i=1}^{2^k} x_i, e\right) \leq d_c(y,e) + r.$$

De (4.9) e da continuidade de d_c temos $d_c(x, e) \leq r$, isto é,

$$d_c(x,e) \leq \rho(x,e), \text{ se } \rho(x,e) = r. \tag{4.11}$$

Analogamente se prova (4.11) se $\rho(x, e) = \xi r$, com o número ξ; $0 \leq \xi \leq 1$.

Daí (4.11) vale para todo x de B_r. Devido à invariância à esquerda das métricas ρ e d_c nós temos a desigualdade desejada, $d_c \leq \rho$.

Provemos a desigualdade oposta. Sejam g, $h \in G$ e $X = X(t)$, $0 \leq t \leq 1$, um caminho horizontal diferenciável por partes arbitrário em G ligando g e h. É suficiente provar que

$$\rho(g,h) \leq \int_0^1 F_0(\theta(X'(t)))dt. \tag{4.12}$$

Consideremos a curva X_ε definida pela Proposição 4.5. Por (a) e pelo Lema 4.6 dado abaixo, o caminho X_ε é retificável na métrica ρ e seu comprimento é igual a

$$\ell_\rho(X_\varepsilon) = \int_0^1 F_0(\theta(X'_\varepsilon(t)))dt.$$

Agora, segue das propriedades (b) e (c) que

$$\rho(g,h) = \rho(X_\varepsilon(0), h) \leq \rho(X_\varepsilon(0), X_\varepsilon(1)) + \rho(X_\varepsilon(1), h) \leq \ell_\rho(X_\varepsilon) + \varepsilon$$

e
$$\int_0^1 F_0(\theta(X'_\varepsilon(t)))dt + \varepsilon \leq \int_0^1 F_0(\theta(X'(t)))dt + 2\varepsilon.$$

Devido à escolha arbitrária de ε temos (4.12) e isso completa a prova do ponto (i) do teorema.

■

Vamos agora caracterizar L_0 através dos subgrupos a 1-parâmetro.

Lema 4.6. *Seja (G,ρ) um grupo de Lie com métrica intrínseca invariante à esquerda; $x \in \mathfrak{g}$. Então, $x \in L_0$ se, e somente se, o subgrupo à 1-parâmetro $\gamma_x(t) = \exp(tx)$, $0 \leq t \leq 1$, é retificável na métrica ρ. Neste caso temos*

$$F_0(x) = \lim_{t\to 0} \frac{\rho(e, \exp(tx))}{t}.$$

Demonstração. Seja a curva γ_x retificável na métrica ρ. Então para algum número α e qualquer número natural k
$$2^k \rho\left(e, \gamma_x(\frac{\alpha}{2^k})\right) \leq r,$$
isto é, $\exp\left(\dfrac{\alpha}{2^k}x\right) \in B_{\frac{r}{2^k}}$, $\alpha x \in 2^k B_{\frac{r}{2^k}}$ e $\alpha x \in \beta \subset L_0$.

Seja $x \in L_0$, $x \neq 0$. Por definição de F_0 e por β ser fechado temos que para qualquer número k
$$\frac{rx}{F_0(x)} \in \beta \subset 2^k B_{\frac{r}{2^k}}, \quad \frac{rx}{2^k F_0(x)} \in B_{\frac{r}{2^k}}, \quad \rho\left(\gamma_x\left(\frac{r}{2^k F_0(x)}\right), e\right) \leq \frac{r}{2^k}. \tag{4.13}$$

Note que
$$\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^k}\right), e\right) \leq \rho\left(\gamma_x\left(\frac{r}{F_0(x)2^k}\right), \gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right)\right) + \rho\left(\gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right), e\right)$$
$$= \rho\left(\gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right), e\right) + \rho\left(\gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right), e\right).$$

Então,
$$\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^k}\right), e\right) \leq 2\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right), e\right) = \frac{\frac{r}{2^k}\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right), e\right)}{\frac{r}{2^{k+1}}}.$$

De onde temos imediatamente que
$$\frac{\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^k}\right), e\right)}{\frac{r}{2^k}} \leq \frac{\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^{k+1}}\right), e\right)}{\frac{r}{2^{k+1}}}.$$

Segue que a sequência $\left(\dfrac{\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^k}\right), e\right)}{\frac{r}{F_0(x)2^k}}\right)_{k\in\mathbb{N}}$ é uma sequência não-decrescente de números reais. Pela equação (4.13) temos que a sequência é limitada superiormente por $\frac{1}{F_0(x)}$, portanto converge.

Defina

$$a := \lim_{k\to\infty} \left(\frac{\rho\left(\gamma_x\left(\frac{r}{F_0(x)2^k}\right), e\right)}{\frac{r}{F_0(x)2^k}} \right).$$

Afirmamos que

$$\frac{\rho(\gamma_x(t), e)}{t} \leq a. \tag{4.14}$$

De fato, todo $t \in \mathbb{R}$ pode ser escrito como $\sum_{i=1}^{\infty} \dfrac{r}{F_0(x)2^{ki}}$. Logo

$$\rho(\gamma_x(t), e) \leq \sum_{i=1}^{\infty} \rho\left(\gamma\left(\frac{r}{2^k F_0(x)}\right), e\right) \leq \frac{ar}{F_0(x)} \sum_{i=1}^{\infty} \frac{1}{2^{ki}} = at$$

o que mostra a afirmação.

Mostremos que $a = F(x)$. Para isto fixe $\varepsilon > 0$. Tome $N \in \mathbb{N}$ tal que se $k \geq N$, então

$$\frac{\rho\left(\gamma_x\left(\frac{r}{2^k F_0(x)}\right), e\right)}{\frac{r}{2^k F_0(x)}} > a - \frac{\varepsilon}{2}.$$

Afirmamos que se $t \in \left(0, \frac{r}{2^N F_0(x)}\right)$, então $\dfrac{\rho(\gamma_x(t), e)}{t} \geq a - \varepsilon$. Com efeito, para tal t, tome o menor $\bar{t}$ tal que $t + \bar{t}$ é da forma $\frac{r}{2^m F_0(x)}$ Note que $\bar{t} \leq t$ e $m \geq N$. Então

$$\begin{aligned}\rho(\gamma_x(t), e) &\geq \rho(\gamma_x(t+\bar{t}), e) - \rho(\gamma_x(\bar{t}), e) \geq \rho(\gamma_x(t+\bar{t}), e) - a\bar{t} \\ &\geq \left(a - \frac{\varepsilon}{2}\right)(t + \bar{t}) - a\bar{t} = at - \frac{\varepsilon}{2}(t+\bar{t}) \geq (a - \varepsilon)t.\end{aligned}$$

Portanto, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que se $t \in (0, \delta)$, então

$$\left| \frac{\rho(\gamma_x(t), e)}{t} - a \right| < \varepsilon.$$

Em vista da aleatoriedade de t segue daí que a curva γ_x é retificável e

$$\lim_{t\to 0} \frac{\rho(\gamma_x(t), e)}{t} = a \leq F_0(x). \tag{4.15}$$

Devido à equação (4.14) temos

$$\rho\left(\gamma_x\left(\frac{r}{a2^k}\right), e\right) \leq \frac{r}{2^k},$$

isto é, $\frac{rx}{a} \in 2^k B_{\frac{r}{2^k}}$, $\frac{rx}{a} \in \beta$, e por definição de F_0

$$a = r\frac{a}{r} \geq F_0(x).$$

Juntamente com (4.15) isto dá a equação desejada. O lema está provado. ∎

Agora, utilizando o Lema 4.6 vamos provar o ponto (ii) do Teorema 4.1.

Demonstração do Teorema 4.1. **(ii)** Suponha que existam dois subespaços vetoriais, L_0 e L_1, da álgebra de Lie $\mathfrak{g}$ e duas normas F_0 e F_1 em L_0 e L_1, respectivamente, que satisfaçam as condições do item (i) do Teorema 4.1, mas que $L_0 \neq L_1$.

Suponha, sem perda de generalidade, que $x \in L_0$ e que $x \notin L_1$. Pelo Lema 4.6, por x não pertencer à L_1, o subgrupo a 1-parâmetro $\gamma_x(t) = \exp(tx)$, $0 \leq t \leq 1$, não é retificável na métrica ρ, mas isto é um absurdo, pois $x \in L_0$ e então, pelo mesmo lema, γ_x é retificável em ρ. Logo $L_0 = L_1$.

Ainda pelo Lema 4.6, temos que

$$F_0(x) = \lim_{t \to 0} \frac{\rho(e, \exp(tx))}{t} = F_1(x).$$

Assim, $F_0 = F_1$.

Por outro lado, dado o par (L_0, F_0), a métrica de Carnot-Carathéodory-Finsler é unicamente definida.

■

Lema 4.7. *Seja $\mathbb{R}^n$ munido com uma norma F e com uma métrica intrínseca d, onde $d(x,y) = \inf\limits_{\gamma}\{\ell(\gamma)\}$, $\ell(\gamma) = \int_a^b F(\gamma'(t))dt$ e o ínfimo é tomado sobre todos os caminhos continuamente diferenciáveis por partes ligando x e y definidos no intervalo $[a,b]$. Então, para todos x, $y \in \mathbb{R}^n$, $d(x,y) = F(x-y)$.*

Demonstração. Considere a curva α como sendo o segmento que liga x e y, onde $\alpha : [0,1] \to \mathbb{R}^n$ é dada por $\alpha(t) = (1-t)x + ty$. Assim,

$$d(x,y) = \inf_{\gamma}\{\ell(\gamma)\} \leq \inf_{\gamma}\{\ell(\alpha)\} = \ell(\alpha) = \int_0^1 F(\alpha'(t))dt = \int_0^1 F(x-y)dt = F(x-y).$$

Portanto, $d(x,y) \leq F(y-x)$.

Agora para provar a desigualdade contrária considere um caminho diferenciável por partes $\gamma : [a,b] \to \mathbb{R}^n$ qualquer ligando x e y. Considere agora a aplicação $t \mapsto F(\gamma'(t))$. Note que para todo $\varepsilon > 0$ existe um $\delta > 0$ tal que se tomarmos uma partição $P = \{a = t_0 < t_1 < \cdots < t_m = b\}$ do intervalo $[a,b]$ com $|P| < \delta$[1] e $\bar{t}_i \in [t_{i-1}, t_i]$, $i = 1, \ldots, m$, então

$$\left| \sum_{i=1}^{m} F(\gamma'(\bar{t}_i))(t_i - t_{i-1}) - \int_a^b F(\gamma'(t))dt \right| < \frac{\varepsilon}{2}.$$

Além disso, para todo $t \in [a,b]$ existe $\beta_t < \delta$ de forma que

$$\left| F\left(\frac{\gamma(t+\Delta t) - \gamma(t)}{\Delta t}\right) - F(\gamma'(t)) \right| < \frac{\varepsilon}{2(b-a)}$$

se $|\Delta t| < \beta_t$.

Cubra $[a,b]$ por intervalos do tipo $(t - \beta_t, t + \beta_t)$. Tome uma subcobertura finita e considere a partição $P = \{a = \tilde{t}_0 < \tilde{t}_1 < \cdots < \tilde{t}_r = b\}$ tal que $\tilde{t}_i$ ou $\tilde{t}_{i-1}$ é o centro do intervalo e o outro é um ponto nesse mesmo intervalo. Então.

[1] Isso significa que o comprimento euclidiano de cada subintervalo da partição é estritamente menor do que δ.

$$
\begin{aligned}
|F(\gamma(b)) - F(\gamma(a))| &= \sum_{i=1}^{r} \left|F(\gamma(\tilde{t}_i)) - \gamma(\tilde{t}_{i-1})\right| \\
&\leq \sum_{i=1}^{r} \frac{F(\gamma(t_i) - \gamma(t_{i-1}))}{|t_i - t_{i-1}|}(t_i - t_{i-1}) \\
&\leq \sum_{i=1}^{r} \left(F\frac{(\gamma(t_i) - \gamma(t_{i-1}))}{t_i - t_{i-1}}\right)(t_i - t_{i-1}) \\
&\leq \sum_{i=1}^{r} \left(F(\gamma'(\bar{t}_i)) + \frac{\varepsilon}{2(b-a)}\right)(t_i - t_{i-1}) \\
&= \left(\sum_{i=1}^{r}(F(\gamma'(\bar{t}_i)))(t_i - t_{i-1})\right) + \frac{\varepsilon}{2} \\
&\leq \int_a^b F(\gamma'(t))dt + \varepsilon.
\end{aligned}
$$

Fazendo ε tender a zero temos que $F(y-x) \leq \ell(\gamma)$ para qualquer curva diferenciável por partes γ que ligue x e y, logo temos $F(y-x) \leq d(x,y)$. ■

Lema 4.8. *Seja $\mathbb{R}^n$ com uma norma $||\cdot||$ que induz uma métrica não euclidiana. Então existem vetores $v_1,\ v_2 \in \mathbb{R}^n$ tais que não existe ângulo no sentido de Alexandrov entre seus subgrupos a 1-parâmetro.*

Demonstração. Já que a métrica não é euclidiana então não vale a identidade do paralelogramo para todos os vetores de $\mathbb{R}^n$. Então tome $v_1,\ v_2 \in \mathbb{R}^n$, ambos não nulos, tais que para eles não vale a identidade do paralelogramo, ou seja, existe $r(v_1, v_2) = r \neq 0$ tal que

$$||v_1 + v_2||^2 + ||v_1 - v_2||^2 = 2||v_1||^2 + 2||v_2||^2 + r.$$

Assim,

$$||v_1||^2 + ||v_2||^2 - ||v_1 - v_2||^2 = -(||v_1||^2 + ||v_2||^2 - ||v_1 + v_2||^2 + r). \tag{4.16}$$

Considere os ângulos de comparação entre v_1 e v_2 e entre v_1 e $-v_2$, respectivamente dados por

$$
\begin{aligned}
\tilde{\measuredangle}(v_1, v_2) &= \arccos\left(\frac{||v_1||^2 + ||v_2||^2 - ||v_1 - v_2||^2}{2||v_1||\,||v_2||}\right) \\
\tilde{\measuredangle}(v_1, -v_2) &= \arccos\left(\frac{||v_1||^2 + ||v_2||^2 - ||v_1 + v_2||^2}{2||v_1||\,||v_2||}\right).
\end{aligned} \tag{4.17}
$$

Da equação (4.16) temos que

$$\tilde{\measuredangle}(v_1, v_2) = \arccos\left(-\left(\frac{||v_1||^2 + ||v_2||^2 - ||v_1 + v_2||^2 + r}{2||v_1||\,||v_2||}\right)\right). \tag{4.18}$$

Suponha, por absurdo, que $\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) = \pi$. Assim, $\tilde{\measuredangle}(v_1, v_2) = \pi - \tilde{\measuredangle}(v_1, -v_2)$ e aplicando o cosseno vem que

$$\cos\left(\tilde{\measuredangle}(v_1, v_2)\right) = \cos\left(\pi - \tilde{\measuredangle}(v_1, -v_2)\right) = -\cos\left(\tilde{\measuredangle}(v_1, -v_2)\right).$$

Das equações (4.17) e (4.18)

$$-\left(\frac{||v_1||^2 + ||v_2||^2 - ||v_1 + v_2||^2 + r}{2||v_1||\,||v_2||}\right) = -\left(\frac{||v_1||^2 + ||v_2||^2 - ||v_1 + v_2||^2}{2||v_1||\,||v_2||}\right)$$

O que implica em $r = 0$, o que seria uma contradição. Portanto, $\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) \neq \pi$.

Suponha, por absurdo, que $\measuredangle(\gamma_{v_1}, \gamma_{v_2})$ e $\measuredangle(\gamma_{v_1}, \gamma_{-v_2})$ existam. Então $\measuredangle(\gamma_{v_1}, \gamma_{v_2}) = \measuredangle(v_1, v_2)$ e $\measuredangle(\gamma_{v_1}, \gamma_{-v_2}) = \measuredangle(v_1, -v_2)$. Pelo Teorema 1.13 temos que

$$\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) > \pi. \tag{4.19}$$

Denotando por

$$\theta^+ = \frac{||v_1||^2 + ||v_2||^2 - ||v_1 - v_2||^2}{2||v_1||\,||v_2||}$$

e

$$\theta^- = \frac{||v_1||^2 + ||v_2||^2 - ||v_1 + v_2||^2}{2||v_1||\,||v_2||},$$

temos que $\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) = \pi$ se, e somente se, $\theta^+ + \theta^- = 0$. De fato,

$$\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) = \pi \Leftrightarrow \tilde{\measuredangle}(v_1, v_2) = \pi - \tilde{\measuredangle}(v_1, -v_2) \Leftrightarrow \cos(\tilde{\measuredangle}(v_1, v_2)) = \cos(\pi - \tilde{\measuredangle}(v_1, -v_2))$$

$$\Leftrightarrow \cos(\tilde{\measuredangle}(v_1, v_2)) = -\cos(\tilde{\measuredangle}(v_1, -v_2)) \Leftrightarrow \theta^+ = -\theta^- \Leftrightarrow \theta^+ + \theta^- = 0.$$

Afirmamos ainda que

$$\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) > \pi \Leftrightarrow \theta^+ + \theta^- < 0. \tag{4.20}$$

De fato, como $\tilde{\measuredangle}(v_1, v_2), \tilde{\measuredangle}(v_1, -v_2) \in [0, \pi]$ e a função cosseno é decrescente neste intervalo temos que

$$\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) > \pi \Leftrightarrow \tilde{\measuredangle}(v_1, v_2) > \pi - \tilde{\measuredangle}(v_1, -v_2) \Leftrightarrow \cos(\tilde{\measuredangle}(v_1, v_2)) < \cos(\pi - \tilde{\measuredangle}(v_1, -v_2))$$

$$\Leftrightarrow \cos(\tilde{\measuredangle}(v_1, v_2)) < -\cos(\tilde{\measuredangle}(v_1, -v_2)) \Leftrightarrow \theta^+ < -\theta^- \Leftrightarrow \theta^+ + \theta^- < 0.$$

Analogamente temos que

$$\tilde{\measuredangle}(v_1, v_2) + \tilde{\measuredangle}(v_1, -v_2) < \pi \Leftrightarrow \theta^+ + \theta^- > 0. \tag{4.21}$$

Note que

$$\theta^+ + \theta^- = \frac{2||v_1||^2 + 2||v_2||^2 - ||v_1 - v_2||^2 - ||v_1 + v_2||^2}{2||v_1||\,||v_2||}. \tag{4.22}$$

Considere agora os seguintes ângulos de comparação

$$\tilde{\measuredangle}\left(\frac{v_2 - v_1}{2}, \frac{-v_1 - v_2}{2}\right) = \arccos\left(\frac{\left|\left|\frac{v_1 - v_2}{2}\right|\right|^2 + \left|\left|\frac{v_1 + v_2}{2}\right|\right|^2 - ||v_2||^2}{2\left|\left|\frac{v_1 - v_2}{2}\right|\right| \left|\left|\frac{v_1 + v_2}{2}\right|\right|}\right) \tag{4.23}$$

e

$$\tilde{\measuredangle}\left(\frac{v_1 - v_2}{2}, \frac{-v_1 - v_2}{2}\right) = \arccos\left(\frac{\left|\left|\frac{v_1 - v_2}{2}\right|\right|^2 + \left|\left|\frac{v_1 + v_2}{2}\right|\right|^2 - ||v_1||^2}{2\left|\left|\frac{v_1 - v_2}{2}\right|\right| \left|\left|\frac{v_1 + v_2}{2}\right|\right|}\right). \tag{4.24}$$

E chame de φ^+ e φ^- os argumentos do arccos das equações (4.23) e (4.24), respectivamente.

Assim, temos

$$\begin{aligned}\varphi^+ + \varphi^- &= \frac{2\left|\left|\frac{v_1 - v_2}{2}\right|\right|^2 + 2\left|\left|\frac{v_1 + v_2}{2}\right|\right|^2 - ||v_1||^2 - ||v_2||^2}{2\left|\left|\frac{v_1 - v_2}{2}\right|\right| \left|\left|\frac{v_1 + v_2}{2}\right|\right|} \\ &= \frac{||v_1 - v_2||^2 + ||v_1 + v_2||^2 - 2||v_1||^2 - 2||v_2||^2}{2\,||v_1 - v_2||\,||v_1 + v_2||}. \end{aligned} \tag{4.25}$$

Comparando as equações (4.22) e (4.25) vemos que o sinal de $\theta^+ + \theta^-$ é sempre contrário ao sinal de $\varphi^+ + \varphi^-$.

Pela desigualdade (4.19) e pela equivalência em (4.20) temos que $\theta^+ + \theta^- < 0$ e pelo que observamos acima teremos $\varphi^+ + \varphi^- > 0$. Pela equivalência em (4.21) temos que

$$\tilde{\measuredangle}\left(\frac{v_2 - v_1}{2}, \frac{-v_1 - v_2}{2}\right) + \tilde{\measuredangle}\left(\frac{v_1 - v_2}{2}, \frac{-v_1 - v_2}{2}\right) < \pi,$$

o que é um absurdo, por contradizer a desigualdade triangular do Teorema 1.13.

Conclui-se que se existisse ângulo no sentido de Alexandrov entre v_1 e v_2 teríamos um absurdo com os ângulos de comparação. Então não existe o referido ângulo. ■

Vamos apresentar brevemente um resultado de topologia conhecido como Lema do Tubo, cuja demonstração pode ser encontrada em [10], e nos será útil na demonstração do próximo teorema.

Sejam X e Y espaços topológicos, onde Y é compacto. Seja $x_0 \in X$ e W uma vizinhança de x_0 em X. Então temos que $W \times Y$ é um tubo que contem a "fatia" $\{x_0\} \times Y$ de $X \times Y$.

Lema 4.9 (Lema do Tubo)**.** *Considere o espaço produto $X \times Y$, onde Y é compacto. Se N é um conjunto aberto de $X \times Y$ contendo a "fatia" $\{x_0\} \times Y$ de $X \times Y$, então N contem algum tubo $W \times Y$ ao redor de $\{x_0\} \times Y$, onde W é uma vizinhança de x_0 em X.*

Teorema 4.10. *Uma métrica intrínseca invariante à esquerda em um grupo de Lie conexo G será de Finsler se, e somente se, qualquer subgrupo a 1-parâmetro de (G, ρ) é retificável (em um segmento). Uma métrica de Finsler invariante à esquerda ρ será uma métrica Riemanniana invariante à esquerda se, e somente se, entre quaisquer dois subgrupos a 1-parâmetro de (G, ρ) existir um ângulo no sentido de A. D. Alexandrov na identidade $e \in G$.*

Demonstração. Se a métrica ρ é de Finsler então $L_0 = \mathfrak{g}$ e pelo Lema 4.6 segue que qualquer subgrupo a 1-parâmetro é retificável.

Reciprocamente, se qualquer subgrupo a 1-parâmetro é retificável, pelo Lema 4.6, $\mathfrak{g} = L_0$ e portanto ρ é uma métrica de Finsler.

Seja F a norma em $\mathfrak{g}$ que dá origem à métrica ρ, e considere a aplicação $\exp : \mathfrak{g} \to G$ e uma parametrização de G, $\mathbf{x} : A \subset \mathbb{R}^n \to G$, onde A é um aberto. Considere a identificação $\mathbb{R}^n \approx \mathfrak{g}$. Então $\exp : A \subset \mathbb{R}^n \to G$ é um sistema de coordenadas de primeira espécie em uma vizinhança da origem. Identificaremos A e $\exp(A)$. Suponha que A seja uma vizinhança convexa.

Considere também a família de métricas Finsler contínuas $\{d_t\}_{t>0}$ induzida pela família de normas $\{F_t\}_{t>0}$, em A, onde $F_t(x,v) = \frac{F(tx,tv)}{t} = F(tx,v)$.

Chamamos de F_0 a norma dada por $F_0(x,v) = \lim_{t\to 0^+} F_t(x,v)$. F_0 é uma norma Finsler contínua em $\mathbb{R}^n$. F_0 induz uma métrica d_0 em $\mathbb{R}^n$ e pelo Lema 4.7, $d_0(x,y) = F_0(x-y)$, para quaisquer $x,\ y \in \mathbb{R}^n$.

Sejam $v_1,\ v_2 \in \mathbb{R}^n$ e considere seus subgrupos a 1-parâmetro, γ_{v_1} e γ_{v_2}, respectivamente.

Provemos agora a seguinte afirmação: Para todo $\varepsilon > 0$ existe $t \in (0,1)$ tal que

$$(1-\varepsilon)F_0(tx,v) \leq F_t(x,v) \leq (1+\varepsilon)F_0(tx,v)$$

para todos $x \in V$, $t \in [0, t_\varepsilon)$.

Defina a aplicação $\tilde{F} : [0,1] \times S\overline{V} \to \mathbb{R}$ dada por

$$\tilde{F}(t,x,v) = F_t(x,v) - F_0(x,v) = F(tx,v) - F_0(x,v),$$

onde $S\overline{V} \subset T\overline{V}$ e S é a esfera unitária centrada em x.

Temos que $\tilde{F}(0,x,v) \equiv 0$ e $\{0\} \times S\overline{V} \subset \tilde{F}^{-1}\big((-\varepsilon,\varepsilon)\big)$.

Usando o Lema 4.9 temos que existe t_ε tal que

$$[0,t_\varepsilon) \times S\overline{V} \subset \tilde{F}^{-1}\big((-\varepsilon,\varepsilon)\big).$$

Assim, para todo $\overline{\varepsilon} > 0$ existe t_ε tal que se $(t,x,v) \in [0,t_\varepsilon) \times S\overline{V}$ então

$$|F_t(x,v) - F_0(x,v)| < \overline{\varepsilon}.$$

Logo,

$$\begin{gathered} F_0(0,v) - \overline{\varepsilon} \leq F_t(x,v) \leq F_0(0,v) + \overline{\varepsilon} \\ \Rightarrow F_0(0,v)\left(1 - \frac{\overline{\varepsilon}}{F_0(0,v)}\right) \leq F_t(x,v) \leq F_0(0,v)\left(1 + \frac{\overline{\varepsilon}}{F_0(0,v)}\right). \end{gathered}$$

Tomando $C = \min_{||v||=1}\{F_0(0,v)\}$ temos que para todos $(t,x,v) \in [0,t_\varepsilon) \times S\overline{V}$

$$F_0(0,v)\left(1 - \frac{\overline{\varepsilon}}{C}\right) \leq F_t(x,v) \leq F_0(0,v)\left(1 + \frac{\overline{\varepsilon}}{C}\right).$$

Se vale para todos $(t,x,v) \in [0,t_\varepsilon) \times S\overline{V}$ também vale para todos $(t,x,v) \in [0,t_\varepsilon) \times T\overline{V}$.

Chamando ℓ_0 a função de comprimento relativa à norma F_0, temos que para uma curva γ, $\ell_0(\gamma) = \int_0^1 F_0(\gamma(s),\gamma'(s))ds$. O mesmo para ℓ_t e F_t, $\ell_t(\gamma) = \int_0^1 F_t(\gamma(s),\gamma'(s))ds$.

Daí decorre que, para todo $t < t_\varepsilon$,

$$(1-\varepsilon)\ell_0(\gamma) \leq \ell_t(\gamma) \leq (1+\varepsilon)\ell_0(\gamma),$$

pois

$$\ell_t(\gamma) = \int_0^1 F_t(\gamma(s), \gamma'(s))ds \geq \int (1-\varepsilon)F_0(\gamma(s), \gamma'(s))ds = (1-\varepsilon)\ell_0(\gamma).$$

Analogamente $\ell_t(\gamma) \leq (1+\varepsilon)\ell_0(\gamma)$.

Assim, segue que

$$(1-\varepsilon)d_0(x,y) = (1-\varepsilon)\inf_\gamma\{\ell_0(\gamma)\} = \inf_\gamma\{(1-\varepsilon)\ell_0(\gamma)\} \leq \inf_\gamma\{\ell_t(\gamma)\} = d_t(x,y)$$

$$(1+\varepsilon)d_0(x,y) = (1+\varepsilon)\inf_\gamma\{\ell_0(\gamma)\} = \inf_\gamma\{(1+\varepsilon)\ell_0(\gamma)\} \geq \inf_\gamma\{\ell_t(\gamma)\} = d_t(x,y).$$

Portanto, $(1-\varepsilon)d_0 \leq d_t \leq (1+\varepsilon)d_0$, para todo $t \in [0, t_\varepsilon)$.

Da relação acima temos que

$$\frac{(1-\varepsilon)^2\left(d_0^2(x,0)+d_0^2(y,0)\right)}{(1+\varepsilon)^2 2d_0(x,0)d_0(y,0)} \leq \frac{d_t^2(x,0)+d_t^2(y,0)}{2d_t(x,0)d_t(y,0)} \leq \frac{(1+\varepsilon)^2\left(d_0^2(x,0)+d_0^2(y,0)\right)}{(1-\varepsilon)^2 2d_0(x,0)d_0(y,0)}. \tag{4.26}$$

Além disso,

$$\frac{-(1+\varepsilon)^2 d_0^2(x,y)}{(1-\varepsilon)^2 2d_0(x,0)d_0(y,0)} \leq \frac{-d_t^2(x,y)}{2d_t(x,0)d_t(y,0)} \leq \frac{-(1-\varepsilon)^2 d_0^2(x,y)}{(1+\varepsilon)^2 2d_0(x,0)d_0(y,0)}. \tag{4.27}$$

Somando termo a termo as desigualdades (4.26) e (4.27) temos

$$\frac{(1-\varepsilon)^4\left(d_0^2(x,0)+d_0^2(y,0)\right)-(1+\varepsilon)^4 d_0^2(x,y)}{(1+\varepsilon)^2(1-\varepsilon)^2 2d_0(x,0)d_0(y,0)} \leq \frac{d_t^2(x,0)+d_t^2(y,0)-d_t^2(x,y)}{2d_t(x,0)d_t(y,0)} \tag{4.28}$$

e

$$\frac{d_t^2(x,0)+d_t^2(y,0)-d_t^2(x,y)}{2d_t(x,0)d_t(y,0)} \leq \frac{(1+\varepsilon)^4\left(d_0^2(x,0)+d_0^2(y,0)\right)-(1-\varepsilon)^4 d_0^2(x,y)}{(1+\varepsilon)^2(1-\varepsilon)^2 2d_0(x,0)d_0(y,0)}. \tag{4.29}$$

Lembrando da definição de F_t, temos que dada uma curva γ e sendo ℓ a estrutura de comprimento associada à norma F, então

$$\ell_t(\gamma) = \int_0^1 F_t(\gamma(s), \gamma'(s))ds = \int_0^1 F(t\gamma(s), t\gamma'(s))\frac{ds}{t} = \frac{\ell(t\gamma)}{t}.$$

Assim, $\ell(t\gamma) = t\ell_t(\gamma)$ e, portanto

$$d_t(p,q) = \inf_\gamma\{\ell_t(\gamma)\} = \inf_\gamma\left\{\frac{\ell(t\gamma)}{t}\right\} = \frac{d(tp,tq)}{t}.$$

Assim, para $t < t_\varepsilon$, não nulo, temos que

$$\frac{d_t^2(x,0)+d_t^2(y,0)-d_t^2(x,y)}{2d_t(x,0)d_t(y,0)} = \frac{t^2d_t^2(x,0)+t^2d_t^2(y,0)-t^2d_t^2(x,y)}{2td_t(x,0)td_t(y,0)}$$

$$= \frac{d^2(tx,0)+d^2(ty,0)-d^2(tx,ty)}{2d(tx,0)d(ty,0)}.$$

Substituindo a igualdade acima em (4.28) e (4.29) e fazendo ε tender a zero, temos que t_ε também tende a zero, logo

$$\frac{d_0^2(x,0)+d_0^2(y,0)-d_0^2(x,y)}{2d_0(x,0)d_0(y,0)} \leq \lim_{t\to 0^+}\left(\frac{d^2(tx,0)+d^2(ty,0)-d^2(tx,ty)}{2d(tx,0)d(ty,0)}\right)$$

$$\leq \frac{d_0^2(x,0)+d_0^2(y,0)-d_0^2(x,y)}{2d_0(x,0)d_0(y,0)}.$$

Portanto vale a igualdade e assim

$$\measuredangle_{d_0}(\gamma_{V_1},\gamma_{V_2}) = \lim_{\bar{s},\tilde{s}\to 0}\arccos\left(\frac{d_0^2(\gamma_{V_1}(\bar{s}),0)+d_0^2(\gamma_{V_2}(\tilde{s}),0)-d_0^2(\gamma_{V_1}(\bar{s}),\gamma_{V_2}(\tilde{s}))}{2d_0(\gamma_{V_1}(\bar{s}),0)d_0(\gamma_{V_2}(\tilde{s}),0)}\right)$$

$$= \lim_{\bar{s},\tilde{s}\to 0}\arccos\left(\lim_{t\to 0}\frac{d^2(t\gamma_{V_1}(\bar{s}),0)+d^2(t\gamma_{V_2}(\tilde{s}),0)-d^2(t\gamma_{V_1}(\bar{s}),t\gamma_{V_2}(\tilde{s}))}{2d(t\gamma_{V_1}(\bar{s}),0)d(t\gamma_{V_2}(\tilde{s}),0)}\right)$$

$$= \lim_{\bar{s},\tilde{s}\to 0}\lim_{t\to 0}\arccos\left(\frac{d^2(\gamma_{V_1}(t\bar{s}),0)+d^2(\gamma_{V_2}(t\tilde{s}),0)-d^2(\gamma_{V_1}(t\bar{s}),\gamma_{V_2}(t\tilde{s}))}{2d(\gamma_{V_1}(t\bar{s}),0)d(\gamma_{V_2}(t\tilde{s}),0)}\right)$$

$$= \lim_{\bar{t},\tilde{t}\to 0}\arccos\left(\frac{d^2(\gamma_{V_1}(\bar{t}),0)+d^2(\gamma_{V_2}(\tilde{t}),0)-d^2(\gamma_{V_1}(\bar{t}),\gamma_{V_2}(\tilde{t}))}{2d(\gamma_{V_1}(\bar{t}),0)d(\gamma_{V_2}(\tilde{t}),0)}\right) = \measuredangle_{d}(\gamma_{V_1},\gamma_{V_2}).$$

Ou seja, a existência de um ângulo entre os subgrupos a 1-parâmetro γ_{V_1} e γ_{V_2} em G está condicionada à existência do ângulo, em $\mathbb{R}^n$, entre eles. ■

Uma questão de interesse é de descobrir em que grupos de Lie G qualquer métrica intrínseca invariante à esquerda será Finsler. A partir dos Teoremas 4.1 e 4.10 e do Lema 4.6 vê-se que isto é verdade se, e somente se, qualquer subespaço vetorial L_0 da álgebra de Lie $\mathfrak{g}$ do grupo G que gera $\mathfrak{g}$ coincide com $\mathfrak{g}$. Isso é equivalente a dizer que qualquer subespaço vetorial L_0 de $\mathfrak{g}$ é uma subálgebra.

De fato, suponha que qualquer subespaço da álgebra de Lie $\mathfrak{g}$ que a gere coincida com $\mathfrak{g}$. Suponha L_0 um subespaço vetorial que não é uma subálgebra. Então existem X, $Y \in L_0$ tais que $Z = [X,Y] \notin L_0$. Tome o subespaço E de codimensão 1 que contenha L_0 mas tal que $Z \notin E$. Assim, E é um subespaço que gera $\mathfrak{g}$ mas que não coincide com $\mathfrak{g}$, o que seria um absurdo. Logo L_0 é uma subálgebra.

Reciprocamente se qualquer subespaço vetorial é subálgebra, então para qualquer subespaço L_0 que gera $\mathfrak{g}$ temos que $L_0 = [L_0] = \mathfrak{g}$.

Como também é fácil ver, a caracterização acima é equivalente à afirmação de que qualquer subespaço $L_0 \subset \mathfrak{g}$ de dimensão 2 é uma subálgebra de Lie. De fato, se todo subespaço é subálgebra, em particular, os subespaços de dimensão 2 também o são. Por outro lado, se qualquer subespaço de dimensão 2 é subálgebra então dado um subespaço vetorial L_0 de $\mathfrak{g}$, para quaisquer $X, Y \in L_0$, $[X,Y] \in L_0$, pois o subespaço de dimensão 2 gerado por X e Y é uma subálgebra e está contido em L_0. Consequentemente, $\mathfrak{g}$ é caracterizada por esta propriedade:

Para todos X, $Y \in \mathfrak{g}$
$$[X,Y] = \alpha(X,Y)X + \beta(X,Y)Y \tag{4.30}$$
onde $\alpha(X,,Y)$ e $\beta(X,Y)$ são números reais.

Todas essas álgebras estão classificadas em [7]. Em particular, todas as álgebras de Lie comutativas têm essa propriedade.

Para cada dimensão $n > 1$ existe exatamente uma álgebra de Lie não-comutativa L_n com a propriedade 4.30, L_n tem uma descrição simples: em L_n existe um ideal comutativo $(n-1)$-dimensional I_{n-1} e um elemento $Z \in L_n \setminus I_{n-1}$ tal que $ad(Z) \mid_{I_{n-1}} = id_{I_{n-1}}$.

É provado em [7] que qualquer métrica Riemanniana invariante à esquerda g_n em um grupo de Lie conexo G_n com álgebra de Lie L_n tem curvatura seccional constante negativa.

Pode-se dar outra descrição de G_{n+1}, $n \geq 1$: Ele é o grupo afim A^n, isto é, o grupo gerado por translações paralelas e homotetias de A^n.

Exemplo 4.11. *Vamos considerar o espaço hiperbólico $\mathbb{H}^n$ visto da seguinte maneira: $\mathbb{H}^n = \{(x_1,\ldots,x_n); x_n > 0\}$. Considere o grupo afim $G = \text{span}\{(t,d) \in \mathbb{R}^{n-1} \times \mathbb{R}^*_+\}$ onde t é translação paralela ao hiperplano $x_n = 0$ e d é dilatação. Assim, $(t,d)(x) = dx + t$, ou seja, $(t_1,t_2,\ldots,t_{n-1},d)(x_1,\ldots,x_n) = (dx_1+t_1, dx_2+t_2,\ldots,dx_{n-1}+t_{n-1},dx_n)$.*

Dados os pares (t_1,d_1) e (t_2,d_2) temos que
$$(t_1,d_1)(t_2,d_2)(x) = (t_1,d_1)(d_2x+t_2) = d_1(d_2x+t_2)+t_1$$
$$= d_1d_2x + d_1t_2 + t_1 = (d_1d_2, d_1t_2+t_1)(x).$$
Portanto, essa é a operação do grupo em $\mathbb{H}^n$.

Note que $e = (0,0,\ldots,0,0,1)$ é o elemento neutro do grupo afim G.

Vamos calcular os campos invariantes à esquerda em $\mathbb{H}^n$. Considere $v = (v_1,\ldots,v_n) \in \mathfrak{g}_n$, onde $\mathfrak{g}_n$ é a álgebra de Lie do grupo G, e pode ser identificada com $\mathbb{R}^n$. Note que $\gamma : I \subset \mathbb{R} \to \mathbb{H}^n$ dada por $\gamma(s) = vs + e$ é uma curva tal que $\gamma'(0) = v$.

O campo invariante à esquerda X_v correspondente a v em um ponto $p = (t,d) \in \mathbb{H}^n$ é dado por
$$X_v(p) = \frac{d}{ds}\left(p.\gamma(s)\right)\bigg|_{s=0} = \frac{d}{ds}\left((t,d).(vs+e)\right)\bigg|_{s=0}$$
$$= \frac{d}{ds}\left((t,d)(v_1s, v_2s,\ldots,v_{n-1}s, v_ns+1)\right)\bigg|_{s=0}$$
$$= \frac{d}{ds}\left((dv_1s+t_1, dv_2s+t_2,\ldots,dv_{n-1}s+t_{n-1}, dv_ns+d)\right)\bigg|_{s=0}$$
$$= d(v_1,v_2,\ldots,v_{n-1},v_n) = dv.$$

Geometricamente isso significa que um campo é invariante por qualquer translação paralela ao hiperplano $x_n = 0$ e que o comprimento euclidiano dos vetores do campo X_v varia proporcionalmente ao parâmetro d sem mudar de sentido e de direção.

Considere a parametrização $\boldsymbol{y} : U \subset \mathbb{R}^n \to \mathbb{H}^n$, onde $U = \{(x_1,\ldots,x_n); x_n > 0\}$ e $\boldsymbol{y}$ é a aplicação identidade restrita a U. Podemos escrever $v = (v_1,\ldots,v_n) = \sum_{i=1}^{n} v_i \frac{\partial}{\partial y_i}(e)$, logo

$$X_v(y_1,\ldots,y_n)=\sum_{i=1}^{n}v_iy_n\frac{\partial}{\partial y_i}(y_1,\ldots,y_n). \tag{4.31}$$

Sejam dois vetores $v,w\in\mathfrak{g}_n$ e os campos de vetores invariantes à esquerda X_v e X_w correspondentes a $v=(v_1,\ldots,v_n)$ e $w=(w_1,\ldots,w_n)$, respectivamente. Pela expressão em (4.31) o colchete destes campos é dado por

$$[X_v,X_w]=\left[\sum_{i=1}^{n}v_iy_n\frac{\partial}{\partial y_i},\sum_{j=1}^{n}w_jy_n\frac{\partial}{\partial y_j}\right]$$

$$=\sum_{i=1}^{n}v_iy_n\frac{\partial}{\partial y_i}\left(\sum_{j=1}^{n}w_jy_n\frac{\partial}{\partial y_j}\right)-\sum_{j=1}^{n}w_jy_n\frac{\partial}{\partial y_j}\left(\sum_{i=1}^{n}v_iy_n\frac{\partial}{\partial y_i}\right)$$

$$=v_n\sum_{j=1}^{n}w_jy_n\frac{\partial}{\partial y_j}-w_n\sum_{i=1}^{n}v_iy_n\frac{\partial}{\partial y_i}=v_nX_w-w_nX_v.$$

Isso significa que o colchete de campos invariantes à esquerda neste exemplo satisfaz a propriedade da equação (4.30), logo podemos determinar para a álgebra $\mathfrak{g}_n$ um ideal comutativo $(n-1)$-dimensional I_{n-1} e um elemento $Z\in\mathfrak{g}_n\setminus I_{n-1}$ tal que $ad(Z)\mid_{I_{n-1}}=id_{I_{n-1}}$. Tome o ideal $I_{n-1}=\operatorname{span}\left\{y_n\frac{\partial}{\partial y_1},y_n\frac{\partial}{\partial y_2},\ldots,y_n\frac{\partial}{\partial y_{n-1}}\right\}$ e o elemento $Z=y_n\frac{\partial}{\partial y_n}$. Dado um campo $\sum_{i=1}^{n-1}a_iy_n\frac{\partial}{\partial y_i}\in I_{n-1}$ qualquer temos que

$$\operatorname{ad}(Z)\left(\sum_{i=1}^{n-1}a_iy_n\frac{\partial}{\partial y_i}\right)=\left[y_n\frac{\partial}{\partial y_n},\sum_{i=1}^{n-1}a_iy_n\frac{\partial}{\partial y_i}\right]$$

$$=y_n\frac{\partial}{\partial y_n}\left(\sum_{i=1}^{n-1}a_iy_n\frac{\partial}{\partial y_i}\right)-\sum_{i=1}^{n-1}a_iy_n\frac{\partial}{\partial y_i}\left(y_n\frac{\partial}{\partial y_n}\right)$$

$$=\sum_{i=1}^{n-1}a_iy_n\frac{\partial}{\partial y_i}.$$

Portanto, $\operatorname{ad}(Z)|_{I_{n-1}}=id|_{I_{n-1}}$.

Consequentemente, qualquer grupo de Lie conexo G_n com álgebra de Lie L_n, não comutativa e com a propriedade (4.30), será isométrico ao espaço hiperbólico e portanto será simplesmente conexo.

Em vista do exemplo e da descrição de G_{n+1}, fica provado:

Teorema 4.12. *Se o grupo de Lie conexo G é comutativo ou é o grupo afim n-dimensional $(n\geq 1)$ A^n, então qualquer métrica intrínseca invariante à esquerda em G será Finsler. Caso contrário existe uma métrica intrínseca invariante à esquerda que não é Finsler (Carnot-Carathéodory-Finsler) em G.*

Referências Bibliográficas

[1] V. Berestovskii. Homogeneous manifolds with intrinsic metric. i. *Siberian Mathematical Journal*, 29(6):887–897, 1988.

[2] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33. American Mathematical Society Providence, 2001.

[3] S. Cohn-Vossen. Existence of minimizing paths. *Questions of Differential Geometry in the Large*, 1959.

[4] M. P. do Carmo. *Geometria Riemanniana*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 2008.

[5] R. L. Faber. The Lie bracket and the curvature tensor. *L'Enseignement Mathématique*, 22:29–35, 1976.

[6] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80. Academic press, 1979.

[7] J. Milnor. Curvatures of left invariant metrics on Lie groups. *Advances in mathematics*, 21(3):293–329, 1976.

[8] J. Mitchell et al. On Carnot-Carathéodory metrics. *Journal of Differential Geometry*, 21(1):35–45, 1985.

[9] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications.* Number 91. American Mathematical Soc., 2006.

[10] J. R. Munkres. Topology: a first course, 1975.

[11] H. L. Royden and P. Fitzpatrick. *Real analysis*, volume 198. Macmillan New York, 1988.

[12] L. A. San Martin. *Grupos de Lie*. 2014 (a ser publicado).

[13] J.-P. Serre. *Lie algebras and Lie groups: 1964 lectures given at Harvard University*. Springer, 2009.

[14] M. Spivak. *Calculus on manifolds*, volume 1. WA Benjamin New York, 1965.

[15] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94. Springer Science & Business Media, 2013.

[16] J. A. Wolf. *Spaces of constant curvature*, volume 372. American Mathematical Soc., 2011.