UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO ACADÊMICO EM MATEMÁTICA

ANDRÉ GADELHA ROCHA

UMA VERSÃO LIPSCHITZ DO TEOREMA DE SARD

FORTALEZA
2021

ANDRÉ GADELHA ROCHA

UMA VERSÃO LIPSCHITZ DO TEOREMA DE SARD

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Teoria das Singularidades.

Orientador: Prof. Dr. José Edson Sampaio

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R571v Rocha, André Gadelha.
Uma versão Lipschitz do Teorema de Sard / André Gadelha Rocha. – 2022.
51 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Matemática, Fortaleza, 2022.
Orientação: Prof. Dr. José Edson Sampaio.

1. Conjuntos Subanalíticos. 2. Análise Convexa. 3. Teorema de Sard Lipschitz. 4. Derivada de Clarke. I. Título.

CDD 510

ANDRÉ GADELHA ROCHA

UMA VERSÃO LIPSCHITZ DO TEOREMA DE SARD

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Teoria das Singularidades.

Aprovada em: 09/ 04/ 2021

BANCA EXAMINADORA

Prof. Dr. José Edson Sampaio (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Alexandre César Fernandes Gurgel
Universidade Federal do Ceará (UFC)

Prof. Dr. Eurípedes Carvalho da Silva
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Prof. Dr. Lev Birbrair
Universidade Federal do Ceará (UFC)

Dedico este trabalho aos meus pais, esposa, filhos e demais familiares.

AGRADECIMENTOS

Agradeço primeiramente a Deus por seu imenso amor e por sua misericórdia e sua mãe santíssima.

Aos meus pais, Antônio Rocha e Claúdia Simone Gadelha Rocha por todo esforço que fizeram para que eu chegasse até aqui.

A minha esposa, Tamires Fontenele por desde a graduação me apoiar principalmente nos momentos difíceis, pelo carinho e amor dedicado a nossa família.

Aos meus filhos Alisson Kenedy e Isllohana Sousa.

Aos meus irmãos, Felipe Rocha, Matheus Rocha, Aline Rocha, Dayane Rocha e Antônio Rocha Filho.

Ao Prof. Dr. Edson Sampaio, pelo excelente apoio e grandes conselhos na produção desse trabalho e por ter sido desde a graduação uma inspiração para estudar muito.

Aos membros da banca, Edson Sampaio, Alexandre Fernandes e Eurípedes Carvalho, por se dispor a participar da banca e pelos excelentes ajustes ao trabalho.

Ao Prof. Dr. Tobias Rafael Fernandes Neto, coordenador do Laboratório de Sistemas Motrizes (LAMOTRIZ) onde este *template* foi desenvolvido.

Aos professores da UFC, pelos excelentes cursos ministrados e pela grande disponibilidade para exclarecer dúvidas fora do período de aula. Entre os quais destaco: Edson Sampaio, Alexandre Fernandes, Lev Birbrair, Vincent Granjean, Rodrigo Rodrigues, Ederson Braga e Antônio Caminha.

Aos meus colegas de curso, em especial a Felipe Fernandes que sempre, mesmo ocupado, esclarecia minhas dúvidas, ao Sérgio pelas companhias estudando até tarde da noite, ao Gauss e Josafá.

Ao Shalon Gonçalves pela correção ortográfica e pelos incentivos dados ao longo dessa jornada.

À Andrea Costa, pela eficiência e presteza na secretária do departamento.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

RESUMO

O objetivo desse trabalho é estudar uma versão Lipschitz do Teorema de Sard. Em relação ao Teorema de Sard em sua versão clássica, trocaremos a hipótese de ser C^1 por subanalítica e localmente Lipschitz. Como não estamos necessariamente trabalhando com funções diferenciáveis, os pontos críticos serão substituídos por pontos críticos de Clarke. Para finalizar, daremos exemplos que mostram que o teorema é falso, se omitimos a hipótese de ser localmente Lipschitz ou de ser subanalítica.

Palavras-chave: conjuntos subanalíticos; análise convexa; teorema de Sard Lipschitz; derivada de Clarke.

ABSTRACT

The aim of this work is to study a Lipschitz version of Sard's Theorem. In relation to the Sard Theorem in its classic version, we will change the hypothesis of being C^1 for subanalytic and locally Lipschitz. Since we are not necessarily working with differentiable functions, the critical points will be replaced by Clarke's critical points. Finally, we will give examples which show that the theorem is not true, if we omit the locally Lipschitz hypothesis or the subanalytic hypothesis.

Keywords: subanalytic sets; convex analysis; Sard Lipschitz's theorem; Clarke's derivative.

SUMÁRIO

1 INTRODUÇÃO

Em um curso de Análise na reta, um resultado clássico é que, se $f : I \subset \mathbb{R} \to \mathbb{R}$ é tal que $f'(x) = 0$, para todo $x \in I$, então temos que f é constante em I, onde I é um intervalo.

Na Matemática, uma pergunta natural é se o mesmo resultado vale para várias variavéis. Em outras palavras, considere $\Omega \subset \mathbb{R}^n$ um aberto e conexo (nesse capítulo Ω sempre será um aberto) e uma função $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ onde $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \ldots = \frac{\partial f}{\partial x_n} = 0$ para todo $(a_1, a_2, \cdots, a_n) \in \Omega$, então f é constante, resultado que também clássico. Esse resultado em particular implica que o conjunto dos valores críticos tem medida nula em $\mathbb{R}$. H. Whitney deu um exemplo, em $(WHITNEY, 1935)$, de uma função real de classe C^1, cujo o conjunto dos pontos críticos era apenas um subconjunto próprio do domínio e tinha como valores críticos $[0,1]$, ou seja, não tem medida nula em $\mathbb{R}$.

Anthony Morse demonstrou em $(MORSE, 1939)$, que se $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ é de classe $C^1(n \geq 1)$, então o conjunto dos valores críticos tem medida nula em $\mathbb{R}$.

Arthur Sard estendeu o resultado de Anthony Morse em $(SARD, 1942)$, provando que $f : \Omega \subset \mathbb{R}^m \to \mathbb{R}^n$ uma aplicação diferenciável de classe $C^k(k \geq 1)$, então o conjunto dos valores críticos tem medida nula, para $m > n$, desde que $k \geq m - n + 1$ e para qualquer k, se $m \leq n$.

Muitos resultados na Análise não diferenciável, tem início em tentativas de fazer resultados da Análise (com diferenciação), com suas devidas adaptações. Não seria diferente para o Teorema de Sard.

Itoh e Tanaka, provaram em $(ITOH; TANAKA, 2001)$ que o conjunto de todos os valores críticos da função distância de uma subvariedade de uma variedade Riemanniana completa com dimensão menor que 5, tem medida de Lebesgue nula.

Posteriormente, Rifford em $(RIFFORD, 2004)$ generalizou o resultado, mostrando que o conjuntos dos pontos críticos de Clarke da função distância de uma subvariedade fechada de uma variedade Riemanniana completa tem medida de Lebesgue nula.

Esse trabalho se baseia em $(BOLTE et\ al., 2005)$, que tem por objetivo mostrar, que se $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ é localmente Lipschitz e subanalítica, então o conjuntos do valores críticos de Clarke tem medida de Lebesgue nula.

A estrátegia para resolver esse problema, usa o fato de f ser constante em cada componente conexa dos pontos críticos de Clarke. Uma versão desse resultado já havia sido feito em $(BOLTE et\ al., 2006)$, mas supondo f apenas contínua e subanalítica e os pontos críticos não eram de Clarke, mas sim um subconjunto deste, a saber, os críticos limitantes.

Para finalizar, é natural questionar, se o resultado provado em ($BOLTE et\ al.$, 2005), vale para f contínua e subanalítica, com pontos críticos de Clarke, assim como vale em ($BOLTE et\ al.$, 2006) para críticos limitantes. Na seção 3.1, será estendida a ideia de ponto crítico de Clarke para funções contínuas e verificaremos através de um exemplo que a resposta para o questionamento é negativa.

2 PRELIMINARES

Teorema 2.0.1 (Teorema de Sard)**.** *Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma aplicação de classe C^∞. Os valores críticos de f constituem um conjunto de medida nula em $\mathbb{R}^n$.*

Demonstração. Veja $(JUDICE, 2011)$, Teorema de Sard.

□

2.1 Elementos de teoria da medida

Definição 2.1.1. *Chamamos $Q = [a,b] \times [a,b] \times ... \times [a,b] \subset \mathbb{R}^n$, de cubo n-dimensional fechado, ou simplesmente, de cubo fechado. Denotamos o volume de Q, por $|Q|$, que é dado por $|Q| = (b-a)^n$.*

Definição 2.1.2. *Seja $\mathscr{P}(\mathbb{R}^n)$ o conjunto dos subconjuntos do $\mathbb{R}^n$. Defina $m_* : \mathscr{P}(\mathbb{R}^n) \to [0, +\infty)$, dada por*

$$m_*(E) = \inf \sum_{j=1}^{\infty} |Q_j|,$$

onde o ínfimo é tomado sobre todas as coberturas enumeráveis $\cup_{j=1}^{\infty} Q_j \supset E$ de cubos fechados.

Definição 2.1.3. *Dizemos que $E \subset \mathbb{R}^n$ é mensuráveis a Lebesgue, se para todo $\varepsilon > 0$, existe um aberto $U_\varepsilon \supset E$, tal que $m_*(U_\varepsilon \setminus E) < \varepsilon$.*

Exemplo 1. *Qualquer aberto do $\mathbb{R}^n$ é mensurável a Lebesgue. Com efeito, basta tomar $U_\varepsilon = E$. Como casos particulares, $\mathbb{R}^n$ e $\emptyset$ são mensuráveis.*

Definição 2.1.4. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ é dita mensurável a Lebesgue, se para todo $a \in \mathbb{R}$, o conjunto $f^{-1}((a,\infty]) = \{x \in \mathbb{R}^n : f(x) > a\}$ é mensurável.*

Escreveremos apenas mensurável ao invés de mensurável a Lebesgue.

Proposição 2.1.1. *Se f é contínua, então f é mensurável.*

Demonstração. Com efeito, $f^{-1}((a,\infty]) = f^{-1}((a,\infty)) \cup \{f = \infty\}$ como $f^{-1}((a,\infty))$ é um aberto, portanto f é mensurável.

□

Proposição 2.1.2. *Suponha que $\{f_k\}_{k\in\mathbb{N}}$ é uma sequência de funções mensuráveis, então:*

i. f=$\sup f_k$ *e* g=$\inf f_k$ *são mensuráveis;*

ii. f=$\limsup f_k$ *e* g=$\liminf f_k$ *são mensuráveis.*

Demonstração. Veja $(STEIN; SHAKARCHI, 2009)$, Capítulo 1, Proposição 3. □

Definição 2.1.5. *Seja E um conjunto mensurável. Dizemos que uma propriedade P vale para quase todo ponto de E, se a medida dos pontos de E que não satisfazem P, é nula. Ao invés de escrever "quase todo ponto", escreveremos apenas q.t.p.*

Teorema 2.1.3 (Teorema da convergência limitada)**.** *Suponha que* $\{f_n\}$ *é uma sequência de funções mensuráveis que são todas limitadas por M e com suporte em um conjunto de medida finita E, e* $f_n(x) \to f(x)$ *q.t.p. quando* $n \to \infty$*. Então f é mensurável, limitada, com suporte em E q.t.p. e*

$$\int |f_n - f| \to 0 \text{ quando } n \to \infty$$

Consequentemente,

$$\int f_n \to \int f \text{ quando } n \to \infty$$

Demonstração. Veja $(STEIN; SHAKARCHI, 2009)$, Capítulo 2, Teorema 1.4. □

Definição 2.1.6. *Sejam X um conjunto mensurável e f uma função mensurável. Dizemos que* $f : X \to \mathbb{R}$ *é integrável (ou integrável à Lebesgue) se*

$$\int_X f^+ < \infty \text{ e } \int_X f^- < \infty.$$

Teorema 2.1.4 (Decaimento da integral)**.** *Suponha f integrável em* $\mathbb{R}^d$*. Então para todo* $\varepsilon > 0$*:*
(i) Há um conjunto de medida finita B tal que,

$$\int_{B^c} |f| < \varepsilon.$$

(ii) Há um $\delta > 0$ *tal que*

$$\int_E |f| < \varepsilon \text{ quando } m(E) < \delta.$$

Demonstração. Veja $(STEIN; SHAKARCHI, 2009)$, Capítulo 2, Proposição 1.12. □

Teorema 2.1.5 (Teorema da convergência dominada de Lebesgue)**.** *Seja* (ϕ_k) *uma sequência de funções mensuráveis e g integrável no* $\mathbb{R}^n$ *tal que* $\phi_k \to f$ *q.t.p e* $|\phi_k| \leq g$*. Então,* $\int_{\mathbb{R}^n} |\phi_k - f| dx \to 0$*, quando* $k \to \infty$*.*

Demonstração. Seja $E_k := B_k \cap \{x \in \mathbb{R}^n; g(x) \leq k\}$, onde B_k é a bola centrada na origem e de raio k. Note que $|f| \leq |f - \phi_k| + |\phi_k| \leq \varepsilon + g$ q.t.p., para qualquer que seja o ε, portanto $|f| \leq g$ q.t.p. Defina agora $\Psi_k = \phi_k \chi_{E_j}$ *com* $j \in \mathbb{N}$, onde χ_{E_j} é a função característica. Observe que $\Psi_k \to f\chi_{E_j}$ q.t.p., além disso $supp(\Psi_k) \subset B_j$, onde $supp(\Psi_k)$ é o suporte, pois $\chi_{E_j} \equiv 0$ fora de B_j, temos também que $|\Psi_k| \leq j$, pois $\{x \in \mathbb{R}^n; g(x) \leq j\}$. Logo estamos nas condições do Teorema da convergência limitada, que nos garante que $\int_{E_j} |\phi_k - f| \to 0$, quando $k \to \infty$, logo $\int_{\mathbb{R}^n} |\phi_k - f| = \int_{E_j} |\phi_k - f| + \int_{E_j^c} |\phi_k - f| \leq \varepsilon + \int_{E_j^c} 2g$, que pelo teorema do decaimento da integral $\exists j_0 \in \mathbb{N}$ tq $\int_{E_{j_0}^c} gdx < \varepsilon$ □

Teorema 2.1.6 (Teorema de Fubini)**.** *Seja $f(x,y)$ uma função integrável em $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Então,*

(i) $f^y(x) = f(x,y)$ é integrável em $\mathbb{R}^{n_1}$ q.t.p $y \in \mathbb{R}^{n_2}$;

(ii) Seja $F(y) = \int_{\mathbb{R}^{n_1}} f^y(x)dx$, F(y) é integrável em $\mathbb{R}^{n_2}$.

Além disso, $\int_{\mathbb{R}^{n_2}} \int_{\mathbb{R}^{n_1}} f^y(x)dxdy = \int_{\mathbb{R}^{n_2}} \int_{\mathbb{R}^{n_1}} f(x,y)dxdy = \int_{\mathbb{R}^n} f(z)dz$.

Demonstração. Veja $(STEIN; SHAKARCHI, 2009)$, Capítulo 2, Teorema 3.1. □

Definição 2.1.7. *Um função f definida em $[a,b]$ é dita absolutamente contínua, se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que*

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

desde que $\sum_{k=1}^{n}(b_k - a_k) < \delta$, e os intervalos $(a_k, b_k), k = 1, ..., n$, são disjuntos.

Proposição 2.1.7. *Seja f uma função Lipschitz, então f é absolutamente contínua.*

Demonstração. Seja C a constante de Lipschitz de f, para todo $\varepsilon > 0$ dado tome $\delta = \varepsilon/C$, com efeito $\sum_{k=1}^{n} |f(b_k) - f(a_k)| \leq \sum_{k=1}^{n} C(b_k - a_k) = C\sum_{k=1}^{n}(b_k - a_k) < C\delta = \varepsilon$

□

Teorema 2.1.8. *Seja $f : [a,b] \to \mathbb{R}$ absolutamente contínua, então f' existe q.t.p e f' é integrável. Além disso, vale*

$$f(b) - f(a) = \int_a^b f'(x)dx$$

Demonstração. Veja $(STEIN; SHAKARCHI, 2009)$, Capítulo 3, Teorema 3.11. □

Definição 2.1.8. *Um funcional sublinear no espaço vetorial V é uma aplicação $p : V \to \mathbb{R}$ que satisfaz, para todos $\zeta, \eta \in V$,*

1. $p(\zeta+\eta) \leq p(\zeta)+p(\eta)$ (*subaditividade*);
2. $p(\alpha\zeta) = \alpha p(\zeta)$, $0 \leq \alpha$ (*positivamente homogênea*).

Teorema 2.1.9 (Hanh-Banach)**.** *Sejam V um espaço vetorial real e* $p : V \to \mathbb{R}$ *um funcional sublinear. Se* $f : W \to \mathbb{R}$ *é um funcional linear definido no subespaço* $W \subset V$ *tal que*

$$f(\zeta) \leq p(\zeta)\ \forall \zeta \in W,$$

então f possui uma extensão linear $F : V \to \mathbb{R}$ *que também satisfaz*

$$F(\zeta) \leq p(\zeta),\ \forall \zeta \in V.$$

Demonstração. A prova desse teorema pode ser vista em $(OLIVEIRA, 2001)$, Teorema 10.12. □

Definição 2.1.9. *Sejam B uma bola e*

$$\chi_B = \begin{cases} 1, & se\ x \in B \\ 0, & se\ x \notin B \end{cases}.$$

Dizemos que uma função mensurável f é localmente integrável em $\mathbb{R}^n$*, se* $f(x)\chi_B(x)$ *é integrável para toda bola B. Denotaremos por* $L^1_{loc}(\mathbb{R}^n)$ *o espaço de todas as funções localmente integráveis.*

Teorema 2.1.10. *Sejam* $C_0^\infty(\Omega)$ *o conjunto de todas as funções de classe* C^∞*, com suporte compacto e* $u \in L^1_{loc}(\Omega)$ *tal que*

$$\int_\Omega u(x)\phi(x)dx = 0,$$

para todo $\phi \in C_0^\infty(\Omega)$*. Então* $u = 0$ *q.t.p. em* Ω*.*

Demonstração. Veja $(CAVALCANTI; CAVALCANTI, 2009)$, Lema 1.48. □

2.1.1 Teorema de Radamacher

O teorema a seguir será de fundamental importância para o desenvolvimento desse trabalho, pois ele garantirá que uma função localmente Lipschitz possui uma sequência de pontos diferenciáveis convergindo para qualquer ponto desejado.

Definição 2.1.10. *Dado um conjunto* $E \subset \mathbb{R}^n$ *mensurável a Lebesgue, denotaremos por* $\mathscr{L}^n(E)$ *a medida de E em* $\mathbb{R}^n$*.*

Lema 2.1.11. *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ localmente Lipschitz, então existe a derivada direcional de f em quase todo ponto $x \in \mathbb{R}^n$.*

Demonstração. Como f é localmente Lipschitz, temos que é contínua, o que implica que f é mensurável a Lebesgue, portanto, $\frac{f(x+tv)-f(x)}{t}$ é mensurável. Como já mostrado acima, lim sup de funções mensuráveis é mensurável, então

$$\overline{D}_v f(x) = \limsup_{t\to 0} \frac{f(x+tv)-f(x)}{t}$$

é mensurável. Analogamente,

$$\underline{D}_v f(x) = \liminf_{t\to 0} \frac{f(x+tv)-f(x)}{t}$$

é mensurável. Fixado um $v \in \mathbb{R}^n$ arbitrário, seja $\mathscr{B}_v = \{x \in \mathbb{R}^n; D_v f(x) \text{ não } existe\}$, observe que lim sup e o lim inf, é o maior e o menor dos valores de aderência, quando existe o limite de uma função em um ponto, eles coincidem, logo podemos ver $\mathscr{B}_v = \{x \in \mathbb{R}^n; \overline{D}_v f(x) > \underline{D}_v f(x)\}$ que é mensurável a Lebesgue, pois é a imagem inversa de $(0,+\infty)$ pela função $h = \overline{D}_v f(x) - \underline{D}_v f(x)$ que é mensurável.

Defina $\phi : \mathbb{R} \to \mathbb{R}$, dada por $\phi(t) = f(x+tv)$, temos que $\lim_{h\to 0} \frac{\phi(t+h)-\phi(t)}{h} = \lim_{h\to 0} \frac{f(x+tv+hv)-f(x+tv)}{h}$, entretanto podemos observar que tomando $v \in \mathbb{R}^n$ tal que $|v| = 1$, temos $|\phi(t_1) - \phi(t_2)| = |f(x+t_1 v) - f(x+t_2 v)| \leq k|t_1 - t_2||v| = k|t_1 - t_2|$, sendo k a constante de Lipschitz de f em x. Logo ϕ é Lipschitz, portanto absolutamente contínua e consequentemente diferenciável q.t.p Observe que pedir $|v| = 1$, não tira a generalidade da questão, pois independente do valor do módulo do vetor, a função ainda seria Lipschitz. Sendo ϕ diferenciável q.t.p, implica a existência da derivada direcional de f q.t.p, então temos que f possui as derivadas direcionais q.t.p, portanto $\mathscr{L}^1(B_v \cap L) = 0$, para uma reta L, paralela a v. Assim, pelo Teorema de Fubini, item (ii), temos $\mathscr{L}^n(B_v) = 0$, o que demonstra o resultado desejado.

□

Lema 2.1.12. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ localmente Lipschitz. A derivada de f em x na direcão v é igual a $\langle v, \nabla f(x)\rangle$ para quase todo ponto $x \in \mathbb{R}^n$.*

Demonstração. Tome uma função $\eta \in C_0^\infty(\mathbb{R}^n)$, ou seja, infinitamente diferenciável e com suporte compacto em $\mathbb{R}^n$. Então,

$$\int_{\mathbb{R}^n} \frac{f(x+tv)-f(x)}{t}\eta(x)dx = \int_{\mathbb{R}^n} \frac{f(y)\eta(y-tv) - f(x)\eta(x)}{t}dx = -\int_{\mathbb{R}^n} \frac{\eta(x)-\eta(x-tv)}{t} f(x)dx$$

Na primeira igualdade acima usamos mudança de variável, fazendo $y = x + tv$, na segunda usamos o fato de f ser localmente Lipschitz e fizemos $y \to x$.

Para $t \neq 0$ pequeno o suficiente de modo que f seja Lipschitz nos pontos da forma $x + tv$, observe que

$$\left| f(x+tv) - f(x)t \right| \leq \frac{Ct|v|}{t}.$$

Podemos supor, sem perda de generalidade, que $|v| = 1$.

Seja $\{t_k\}_k$ uma sequência de números reais não nulos tal que $\lim t_k = 0$. Veja que η tem suporte compacto, digamos K, e isto junto com a continuidade de f e η, temos que

$$\int_{\mathbb{R}^n} \frac{f(x+tv) - f(x)}{t} \eta(x) dx = \int_K \frac{f(x+tv) - f(x)}{t} \eta(x) dx < \infty,$$

observe que estamos nas hipoteses do teorema da Convergência Dominada de Lebesgue, onde $g = C$, cada $\phi_k = \frac{f(x+t_k v) - f(x)}{t_k}$ e $D_v f$ faz o papel da função f em tal Teorema, sendo assim, obtemos

$$\lim_{t \to 0} \int_{\mathbb{R}^n} \frac{f(x+tv) - f(x)}{t} \eta(x) dx = \int_{\mathbb{R}^n} \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t} \eta(x) dx = \int_{\mathbb{R}^n} D_v f(x) \eta(x) dx,$$

pela primeira equação dessa demonstração podemos ver que

$$\int_{\mathbb{R}^n} D_v f(x) \eta(x) dx = -\int_{\mathbb{R}^n} f(x) D_v \eta(x) dx$$

mas como podemos escrever

$$-\int_{\mathbb{R}^n} f(x) D_v \eta(x) dx = -\sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} f(x) \frac{\partial \eta}{\partial x_i}(x) dx$$

pelo Teorema de Fubini, temos que

$$-\sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} f(x) \frac{\partial \eta}{\partial x_i}(x) dx = -\sum_{i=1}^{n} v_i \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} f(x) \frac{\partial \eta}{\partial x_i}(x) dx_i \right) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$$

como f é localmente Lipschitz, então f é absolutamente contínua, portanto f' é integrável, pelo Teorema 2.1.8. Seja $x = (x_1, x_2, ..., x_i, ..., x_n)$, fixando todas as coordenadas de x, deixando apenas x_i variando, podemos aplicar integração por partes em

$$\int_{\mathbb{R}} f(x) \frac{\partial \eta}{\partial x_i}(x) dx_i = f(x)\eta(x) \Big]_{-\infty}^{\infty} - \int_{\mathbb{R}} \frac{\partial f(x)}{\partial x_i} \eta(x) dx_i,$$

veja que devido ao η ter suporte compacto

$$f(x)\eta(x)\Big]_{-\infty}^{\infty} = 0$$

então segue que

$$-\sum_{i=1}^{n} v_i \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} f(x) \frac{\partial \eta}{\partial x_i}(x) dx_i \right) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n =$$

$$= -\sum_{i=1}^{n} v_i \int_{\mathbb{R}^{n-1}} \left(-\int_{\mathbb{R}} \frac{\partial f(x)}{\partial x_i} \eta(x) dx_i \right) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$$

novamente pelo Teorema de Fubini

$$-\sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} f(x) \frac{\partial \eta}{\partial x_i}(x) dx = \sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} \frac{\partial f(x)}{\partial x_i} \eta(x) dx$$

usando a linearidade da integral

$$\sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} \frac{\partial f(x)}{\partial x_i} \eta(x) dx = \int_{\mathbb{R}^n} \langle \nabla f(x), v \rangle \eta(x) dx.$$

Agora que temos as igualdades

$$\int_{\mathbb{R}^n} D_v f(x) \eta(x) dx = -\int_{\mathbb{R}^n} f(x) D_v \eta(x) dx$$

e

$$-\int_{\mathbb{R}^n} f(x) D_v \eta(x) dx = \int_{\mathbb{R}^n} \langle \nabla f(x), v \rangle \eta(x) dx,$$

podemos concluir que

$$\int_{\mathbb{R}^n} D_v f(x) \eta(x) dx = \int_{\mathbb{R}^n} \langle \nabla f(x), v \rangle \eta(x) dx,$$

como $\eta \in C_0^\infty(\mathbb{R}^n)$ foi tomada arbitrária, segue do Teorema 2.1.10 que $D_v f(x) = \langle \nabla f(x), v \rangle$, para quase todo ponto $x \in \mathbb{R}^n$.

$\square$

Lema 2.1.13. *Sejam $\{v_k\}_{k=1}^{+\infty} \subset \mathbb{S}^{n-1}$ um conjunto enumerável e denso,*

$$A_k := \{x \in \mathbb{R}^n | \exists D_{v_k} f(x), \exists \nabla f(x) \text{ e } D_{v_k} f(x) = \langle v_k, \nabla f(x) \rangle\}$$

e $A := \cap_{k=1}^{\infty} A_k$. A função localmente Lipschitz f é diferenciável em A.

Primeiramente veja que pelo Lema 2.1.11 $D_v f(x)$ e $\nabla f(x)$ existem q.t.p. e pelo Lema 2.1.12, temos que $D_v f(x) = \langle v, \nabla f(x)\rangle$ q.t.p., sendo assim $\{x \in \mathbb{R}^n; D_v f(x) \neq \langle v, \nabla f(x)\rangle\}$ tem medida nula, em particular, $\mathscr{L}^n(A_k^c) = 0$, logo como a medida de Lebesgue é subaditiva, temos

$$\mathscr{L}^n(A^c) \leq \sum_{k\in\mathbb{N}} \mathscr{L}^n(A_k^c) = 0,$$

sendo assim $\mathscr{L}^n(A^c) = 0$.

Defina agora $Q : A \times \mathbb{S}^{n-1} \times \mathbb{R} \setminus \{0\} \to \mathbb{R}$, dada por

$$Q(x,v,t) = \frac{f(x+tv) - f(x)}{t} - \langle v, \nabla f(x)\rangle$$

temos que $|Q(x,v,t) - Q(x,w,t)| \leq \left|\frac{f(x+tv)-f(x)}{t} - \langle v, \nabla f(x)\rangle - \frac{f(x+tw)-f(x)}{t} + \langle w, \nabla f(x)\rangle\right|$, por desigualdade triangular e linearidade do produto interno segue que

$$|Q(x,v,t) - Q(x,w,t)| \leq \left|\frac{f(x+tv) - f(x)}{t} - \frac{f(x+tw) - f(x)}{t}\right| + |\langle w - v, \nabla f(x)\rangle|,$$

pela desigualdade de Cauchy-Schwarz e por f ser localmente lipschtz (com constante de Lipschitz C) em x, temos $\left|\frac{f(x+tv)-f(x)}{t} - \frac{f(x+tw)-f(x)}{t}\right| \leq \frac{Ct|v-w|}{t} = C|v-w|$,para $t << 1$ e $|\langle w - v, \nabla f(x)\rangle| \leq |v-w||\nabla f(x)|$, logo

$$C|v-w| + |v-w||\nabla f(x)| = (C + |\nabla f(x)|)|v-w| = (C + \sqrt{(D_{e_1} f(x))^2 + \ldots + (D_{e_n} f(x))^2})|v-w|,$$

sendo $(D_{e_i})^2$ o máximo dos quadrados das derivadas parciais, temos $(C + |\nabla f(x)|)|v-w| = (C + \sqrt{(D_{e_1} f(x))^2 + \ldots + (D_{e_n} f(x))^2})|v-w| \leq (C + \sqrt{n}|D_{e_i} f(x)|)|v-w|$, mas como temos que $|D_{e_i} f(x)| = lim_{t\to 0} \frac{f(x+te_i)-f(x)}{t} \leq C$, resultando assim que

$$|Q(x,v,t) - Q(x,w,t)| \leq (1 + \sqrt{n})C|v-w|$$

sendo assim Q é Lipschitz em $\mathbb{S}^{n-1}$. Queremos mostrar que f é diferenciável em todos os pontos de A, como já verificamos que $\mathscr{L}^n(A^c) = 0$, então f será diferenciável q.t.p. Queremos mostrar que $\lim_{y\to x} \frac{f(y)-f(x)-\langle \nabla f(x),(y-x)\rangle}{|x-y|} = 0$, observe que é o mesmo que mostrar que $lim_{y\to x} Q\Big(x, \frac{y-x}{|y-x|}, |y-x|\Big) = 0$, chamando $t = y - x$ e $v = \frac{y-x}{|y-x|}$, então é equivalente a mostrar que $\lim_{t\to 0} Q(x,v,|t|) = 0$.

Veja que

$$\begin{aligned} |Q(x,v,t)| &\leq |Q(x,v_k,t)| + |Q(x,v,t) - Q(x,v_k,t)| \\ &\leq |Q(x,v_k,t)| + (1+\sqrt{n})C|v - v_k|, \end{aligned}$$

como $\{v_k\}_k \subset \mathbb{S}^{n-1}$ é denso, então podemos tomar $v_k \in \mathbb{S}^{n-1}$ tal que $|v - v_k| \leq \frac{\varepsilon}{2(1+\sqrt{n})C}$, para um $\varepsilon > 0$ fixo.

Para $x \in A$, temos por definição de A, que $lim_{t\to 0} Q(x, v_k, t) = 0$, ou seja, para o $\varepsilon > 0$ dado acima, existe um $\delta > 0$, tal que $|Q(x, v_k, t)| < \varepsilon/2$, desde que $0 < |t| < \delta$, sendo assim temos que

$$|Q(x, v, t)| < \varepsilon, \text{ se } 0 < |t| < \delta,$$

fazendo $t \to 0$, temos que $Q(x, v, t) \to 0$, o que prova que f é diferenciável em $x \in A$.

Teorema 2.1.14 (Teorema de Rademacher)**.** *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ localmente Lipschitz. Então f é diferenciável em quase todo ponto de $\mathbb{R}^n$.*

Demonstração. A demonstração segue direto do Lema 2.1.13.

□

2.2 Um pouco de análise convexa

2.2.1 Conjuntos convexos

Definição 2.2.1. *Um conjunto $X \subset \mathbb{R}^n$ é chamado convexo, quando dados $a, b \in X$ tem-se que $(1 - t)a + tb \in X$ para todo $t \in [0, 1]$.*

Em outras palavras, dizer que um conjunto X é convexo, significa dizer que se dois pontos pertencem a X, então o segmento de reta que os unem está contido em X.

Exemplo 2. *O conjunto $A = \{x \in \mathbb{R}^n; |x| < 1\}$ é convexo. Com efeito, tome $a, b \in A$, assim*

$$|(1 - t)a + tb| < (1 - t)|a| + t|b| < (1 - t) + t = 1$$

Exemplo 3. *Seja $B = \{(x, y, z) \in \mathbb{R}^3; a \leq x \leq b, c < y < d\}$ é convexo em $\mathbb{R}^3$. Com efeito, tome $w, v \in B$, logo podemos ver $v = (v_1, v_2, v_3)$ e $w = (w_1, w_2, w_3)$, com $a \leq v_1, w_1 \leq b$, sendo assim, $a = (1 - t)a + ta \leq (1 - t)v_1 + tw_1$, além disso $(1 - t)v_1 + tw_1 \leq (1 - t)b + tb = b$. Temos também que $c < v_2, w_2 < d$, de modo análogo temos que $c < (1 - t)v_2 + tw_2 < d$, mostrando assim que nosso conjunto é convexo.*

Definição 2.2.2. *Seja $C \subset \mathbb{R}^n$. Diremos que C é um cone, se para todo $v \in C$ e todo $t > 0$, tem-se $tv \in C$.*

Exemplo 4. *Um cone $C \subset \mathbb{R}^n$ é um conjunto convexo se, e somente se, $u, v \in C \Rightarrow u + v \in C$. Com efeito, suponha que o cone C é fechado para soma. Tome $u, v \in C \Rightarrow (1-t)u, tv \in C$, por definição de cone e pela hipotese de fechamento $(1-t)u + tv \in C$. Agora suponha C convexo, então $\frac{u+v}{2} = (1-\frac{1}{2})u + \frac{1}{2}v \in C \Rightarrow u + v = 2 \cdot \frac{u+v}{2} \in C$.*

Proposição 2.2.1. *Sejam $C_1, C_2, ..., C_k \subset \mathbb{R}^n$ conjuntos convexos, então $C_1 \cap C_2 \cap ... C_k$ é convexo.*

Demonstração. É suficiente provar para $C_1 \cap C_2$. Tome $a, b \in C_1 \cap C_2$, logo o segmento ligando a e b, $\overline{ab} \subset C_1$ e $\overline{ab} \subset C_2$, portanto $\overline{ab} \subset C_1 \cap C_2$. □

Exemplo 5. *Sejam $B_1 = B[(-2,0);1]$ e $B_2 = B[(2,0);1]$ as bolas euclidianas centradas em $P_1 = (-2,0)$ e $P_2 = (2,0)$ respectivamente, com raio 1. Veja que B_1 e B_2 são conjuntos convexos, mas $B_1 \cup B_2$ não é convexa, pois $\overline{P_1P_2} \not\subseteq B_1 \cup B_2$.*

Exemplo 6. *Seja $B = B[(0,0);1] \subset \mathbb{R}^2$ e B^c (o complementar de B) não é convexo, pois $\overline{(-1,0)(1,0)} \not\subseteq B^c$*

Definição 2.2.3. *Seja $X \subset \mathbb{R}^n$, chamamos de fecho convexo de X, ou envoltória convexa, o menor conjunto convexo C, tal que $X \subset C$.*

O termo menor não tem uma definição formal aqui, então com menor estamos dizendo exatamente que, C é o menor convexo que contém X, se dado um convexo D, tal que $X \subset D$, então $C \subset D$.

Exemplo 7. *Tome dois pontos distintos $a, b \in \mathbb{R}$, logo o fecho convexo do conjunto formado por a e b é o segmento de reta $[a, b]$.*

Exemplo 8. *O fecho convexo do conjunto $A = \{x \in \mathbb{R}^n; |x| < 1\}$ é o próprio conjunto A.*

Definição 2.2.4. *Seja $X \subset \mathbb{R}^n$, definimos o conjunto das combinações convexas de X, como $C(X) = \{\sum_{i=1}^{k} t_i v_i; t_i \geq 0, \sum_{i=1}^{k} t_i = 1\}$, com $v_1, v_2, ..., v_k \in X$.*

Proposição 2.2.2. *Seja $X \subset \mathbb{R}^n$, então $C(X)$ é convexo.*

Demonstração. Seja $C(X)$ o conjunto das combinações convexas, tome $u, v \in C(X)$, logo

$$u = a_1u_1 + a_2u_2 + ... + a_mu_m$$

e

$$v = b_1v_1 + b_2v_2 + ... + b_nv_n,$$

com $\sum_{i=1}^{m} a_i = 1$ e $\sum_{j=1}^{n} b_j = 1$. Seja $t \in (0,1)$, observe que

$$(1-t)u + tv = (1-t)\sum_{i=1}^{m} a_i u_i + t\sum_{j=1}^{n} b_j v_j = \sum_{i=1}^{m}(1-t)a_i u_i + \sum_{j=1}^{n} t b_j v_j,$$

veja que a somatória do lado direito é uma combinação convexa de elementos de X, pois

$$\sum_{i=1}^{m}(1-t)a_i + \sum_{j=1}^{n} t b_j = (1-t)\sum_{i=1}^{m} a_i + t\sum_{j=1}^{n} b_j = (1-t)\cdot 1 + t1 = 1,$$

portanto $C(X)$ é convexo.

□

Lema 2.2.3. *Se $C \subset \mathbb{R}^n$ é convexo e $X \subset C$, então $C(X) \subset C$.*

Demonstração. Faremos a prova por indução. Seja $k = 2$, logo dados $v_1, v_2 \in X \subset C$, temos pela convexidade de C que $(1-t)v_1 + tv_2 \in C$ e $(1-t)v_1 + tv_2$ é uma combinação convexa. Suponha verdade que dados $v_1, v_2, ..., v_n \in X$ e seja $t_1v_1 + t_2v_2 + ... + t_nv_n$ uma combinação convexa, então $t_1v_1 + t_2v_2 + ... + t_nv_n \in C$. Tome uma combinação convexa com $n+1$ elementos de X, ou seja,

$$\begin{aligned} t_1v_1 + t_2v_2 + ... + t_nv_n + t_{n+1}v_{n+1} &= (1-t_{n+1})\tfrac{t_1v_1 + t_2v_2 + ... + t_nv_n}{(1-t_{n+1})} + t_{n+1}v_{n+1} \\ &= (1-t_{n+1})\tfrac{t_1v_1 + t_2v_2 + ... + t_nv_n}{(t_1 + t_2 + ... + t_n)} + t_{n+1}v_{n+1} \end{aligned}$$

seja $w = \frac{t_1v_1 + t_2v_2 + ... + t_nv_n}{(t_1 + t_2 + ... + t_n)} \in C$, pois é uma combinação convexa de n elementos de X (hipotese de indução). Pela convexidade de C, temos $(1-t_{n+1})w + t_{n+1}v_{n+1} \in C$. Sendo assim

$$t_1v_1 + t_2v_2 + ... + t_nv_n + t_{n+1}v_{n+1} = (1-t_{n+1})w + t_{n+1}v_{n+1} \in C.$$

□

Teorema 2.2.4 (Caracterização de fecho convexo)**.** *As seguintes afirmações são equivalentes:*
(i) C é o fecho convexo de X.
(ii) Seja $\{C_\lambda; \lambda \in \Lambda\}$ a coleção de todos os conjuntos convexos que contém X, então $C = \bigcap_{\lambda \in \Lambda} C_\lambda$.
(iii) C=C(X).

Demonstração. $(i) \Leftrightarrow (ii)$
Se C é o fecho convexo, então é claro que $C \subset C_\lambda$ para todo $\lambda \in \Lambda$, consequentemente $C \subset \bigcap_{\lambda \in \Lambda} C_\lambda$. Claramente $\bigcap_{\lambda \in \Lambda} C_\lambda \subset C$. Recíprocamente tome um C' convexo, tal que $X \subset C'$, como $C = \bigcap_{\lambda \in \Lambda} C_\lambda \subset C'$ e $X \subset \bigcap_{\lambda \in \Lambda} C_\lambda = C$ concluíndo assim o resultado.
$(i) \Leftrightarrow (iii)$
Como C é o fecho convexo de X e $C(X)$ é um conjunto convexo, então $C \subset C(X)$. Além disso,

pelo Lema 2.2.3, temos que $C(X) \subset C$, portanto $C = C(X)$. Agora suponha que $C = C(X)$, novamente pelo Lema 2.2.3, $C \subset D$, para qualquer D convexo que contém X, então C é o fecho convexo. $\square$

Proposição 2.2.5. *Seja $F \subset \mathbb{R}^n$ compacto, então $C(F)$ é compacto.*

Demonstração. Veja Proposição 1.16 em ($MACEDO$, 2009). $\square$

2.2.2 Funções convexas

Definição 2.2.5. *Seja C um conjunto convexo. Uma função $f : C \to \mathbb{R}$ é chamada convexa, se para quaisquer $x, y \in C$ e $t \in [0,1]$, tem-se*

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

O que foi dito acima, é que uma função é convexa, se seu gráfico em $[x,y]$ perde para o segmento de reta que liga $f(x)$ a $f(y)$.

Definição 2.2.6. *Uma função $f : C \to \mathbb{R}$ diz-se côncava quando para quaisquer $x, y \in C$ e $t \in [0,1]$ tem-se $(1-t)f(x) + tf(y) \leq f((1-t)x + ty)$. Em outras palavras quando f ganha do segmento de reta.*

Equivalentemente podemos dizer que f é côncava quando $-f$ é convexa.

Os dois conceitos acima não são opostos, observe os exemplos a seguir.

Exemplo 9. *A função $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$, dada por $f(x) = cx + d$ é convexa e côncava.*

Exemplo 10. *A função convexa $f(x) = x^3$ em $[-1,1]$, não é nem concâva, nem convexa, pois em $x = \frac{1}{2}$, temos $f(\frac{1}{2}) = \frac{1}{8} < \frac{1}{2}$. Agora quando $x = -\frac{1}{2}$, temos $-\frac{1}{2} < -\frac{1}{8} = f(-\frac{1}{2})$*

Definição 2.2.7. *Seja $f : U \to \mathbb{R}$ com U aberto e convexo. Chamamos de epigráfico de f, o conjunto $E(f) = \{(x,y) \in U \times \mathbb{R}; f(x) \leq y\} \subset \mathbb{R}^{n+1}$*

Teorema 2.2.6. *Seja $f : U \to \mathbb{R}$ uma função. O E(f) é convexo se, e somente se, f é convexa.*

Demonstração. Suponha $E(f)$ convexo, logo dados $x, y \in U$, temos $(x, f(x)), (y, f(y)) \in E(f)$, pela convexidade do epigráfico, temos que $((1-t)x + ty, (1-t)f(x) + tf(y)) \in E(f)$ para todo $t \in [0,1]$, logo $f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$ o que mostra que f é convexa. Agora suponha f convexa, tome $(x,x'), (y,y') \in E(f)$, logo temos que $f(x) \leq x'$ e $f(y) \leq y'$, sendo assim dado um $t \in [0,1]$ temos $(1-t)f(x)+tf(y) \leq (1-t)x'+ty'$, mas como f é convexa, tem-se que $f((1-t)x+ty) \leq (1-t)f(x)+tf(y) \leq (1-t)x'+ty'$, sendo $((1-t)x+ty,(1-t)x'+ty') \in E(f)$, concluindo assim a demonstração. □

O príncipal resultado desse trabalho, trata-se de pontos críticos. Uma utilidade desses pontos é otimizar, portanto fazeremos uma pequena exposição de alguns resultado de otimização envolvendo funções convexas.

Teorema 2.2.7. *Seja $U \subset \mathbb{R}^n$ um aberto e convexo. Todo ponto crítico a de uma função convexa $f : U \to \mathbb{R}$ de classe C^1 é um ponto de mínimo global, isto é, $f(a) \leq f(x)$ para todo $x \in U$.*

Demonstração. Veja Capítulo 3, Corolário 7 de $(LIMA, 2010)$ □

Teorema 2.2.8. *Seja $U \subset \mathbb{R}^n$ um aberto e convexo. Toda função convexa $f : U \to \mathbb{R}$ é contínua.*

Demonstração. Veja Capítulo 3, Teorema 10 de $(LIMA, 2010)$ □

2.2.3 Funções quase convexas

Definição 2.2.8. *Seja $C \subset \mathbb{R}^n$ convexo. Dizemos que $f : C \to \mathbb{R}$ é quase-convexa quando, para todo $c \in \mathbb{R}$, o conjunto $C_c = \{x \in C; f(x) \leq c\}$ é convexo.*

Teorema 2.2.9 (Caracterização de funções quase-convexa)**.** *Seja $C \subset \mathbb{R}^n$ um conjunto convexo. $f : C \to \mathbb{R}$ é quase-convexa se, e somente se, $f((1-t)x+ty) \leq$ máx$\{f(x), f(y)\}$ para $x, y \in X$ e $t \in [0,1]$ quaisquer.*

Demonstração. Suponha f quase-convexa, logo C_c é um conjunto convexo, para todo $c \in \mathbb{R}$. Dados $x, y \in C_c$, com $c = \text{máx}\{f(x), f(y)\}$, pela convexidade de C_c, temos que $(1-t)x+ty \in C_c$, para todo $t \in [0,1]$. Pela definição de C_c, temos que $f((1-t)x+ty) \leq c = \text{máx}\{f(x), f(y)\}$.

Agora dados $x, y \in C_c$, então $f(x) \leq c$ e $f(y) \leq c$, consequentemente $\text{máx}\{f(x), f(y)\} \leq c$. Por hipotese $f((1-t)x+ty) \leq \text{máx}\{f(x), f(y)\} \leq c$, logo $(1-t)x+ty \in C_c$, portanto C_c é convexo e como c é arbitrário então o resultado está provado. □

Proposição 2.2.10. *Uma função convexa é quase-convexa.*

Demonstração. Se $f : C \to \mathbb{R}$ é convexa, dados quaisquer $x, y \in C$, seja $k = \text{máx}\{f(x), f(y)\}$. Pela convexidade de f temos que

$$\begin{aligned} f((1-t)x+ty) &\leq (1-t)f(x)+tf(y) \\ &\leq (1-t)k+tk \\ &\leq k \\ &\leq \text{máx}\{f(x), f(y)\} \end{aligned}$$

□

Exemplo 11. *Tome $f : \mathbb{R} \to \mathbb{R}$, dada por $f(x) = x^3$. Veja que f é uma função quase-convexa, mas como vimos acima, não é uma função convexa.*

2.3 Derivada de Clarke

Nessa seção falaremos de uma derivada que dá sentido a derivada de certas funções que não existem no sentido clássico, como por exemplo a função módulo. Tal sentido de derivada foi desenvolvida por Frank Clarke, a mesma leva seu nome. A derivada de Clarke será usada fortemente nesse trabalho para tentar fazer análogos ao cálculo para funções localmente Lipschitz. Infelizmente como veremos logo a diante, a derivada de Clarke não generaliza a derivada no sentido clássico, entretanto se a função for de classe C^1 temos essa generalização.

Nessa parte inicial, desenvolveremos a teoria para funções, ou seja, cujo contra domínio é a reta. Posteriormente para aplicações.

Definição 2.3.1. *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, localmente Lipschitz, definimos a derivada direcional de Clarke de f no ponto x na direção v como sendo*

$$f^o(x;v) = \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y+tv) - f(y)}{t}.$$

Observe que $|f^{\text{o}}(x;v)| = \left| \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y+tv)-f(y)}{t} \right| \leq \limsup_{\substack{y \to x \\ t \to 0^+}} \left| \frac{f(y+tv)-f(y)}{t} \right|$, como f é localmente Lipschitz para t suficientemente pequeno, temos que $\limsup_{\substack{y \to x \\ t \to 0^+}} \left| \frac{f(y+tv)-f(y)}{t} \right| \leq C \left| \frac{tv}{t} \right| = C|v|$, onde C é a constante de Lipschitz de f em y.

Definição 2.3.2. *Dizemos que uma função f é positivamente homogênea, se para todo $s > 0$, $f(sx) = sf(x)$*

Proposição 2.3.1. *A função $g_x(v) = f^o(x,v)$ é positivamente homôgenea.*

Demonstração. Sendo

$$g_x(sv) = f^{\circ}(x,sv) = \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+tsv)-f(y)}{t} = \limsup_{\substack{y\to x\\ t\to 0^+}} s\frac{f(y+tsv)-f(y)}{ts},$$

como s não depende de y nem de t, temos que

$$\limsup_{\substack{y\to x\\ t\to 0^+}} s\frac{f(y+tsv)-f(y)}{ts} = s\cdot \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+tsv)-f(y)}{ts} = sf^{\circ}(x,v).$$

□

Definição 2.3.3. *Chamamos uma função f de subaditiva, se* $f(x+y) \leq f(x)+f(y)$*.*

Proposição 2.3.2. *A função* $h_x(v) := f^o(x,v)$ *é subaditiva.*

Demonstração. Tome $v,w \in \mathbb{R}^n$, logo

$$h_x(v+w) = f^{\circ}(x,v+w) = \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+t(v+w))-f(y)}{t} =$$

$$= \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+tv+tw)-f(y+tv)+f(y+tv)-f(y)}{t} \leq \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+tv+tw)-f(y+tv)}{t} +$$

$$+ \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+tv)-f(y)}{t} = f^{\circ}(x,w)+f^{\circ}(x,v) = h_x(w)+h_x(v)$$

.

□

Definição 2.3.4. *O gradiente generalizado de f em x, denotado por* $\partial^o f(x)$*, é o fecho convexo do conjunto dos limites da forma* $\lim_{t\to\infty} \nabla f(x_t)$*, onde* $x_t \to x$ *e f é diferenciável em* x_t *para todo t.*

Exemplo 12. *Seja* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = |x|$*, vamos buscar a derivada de Clarke em* $x = 0$*, temos que buscar todos os* $\nabla f(0+h)$*, com* $h \to 0$*, observe que qualquer que seja o ponto h, teremos* $\nabla f(h) = 1$ *ou* -1*, portanto ao tomar o fecho convexo, concluísse que* $\partial^o f(0) = [-1,1]$*.*

Exemplo 13. *Seja* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = máx\{x,0\}$*, vamos determinar* $\partial^o f(0)$*. Observe que temos dois casos,* $lim_{t\to\infty}\nabla f(x_t) = 0$*, para* $x_t \to 0^-$ *e* $lim_{t\to\infty}\nabla f(x_t) = 1$*, para* $x_t \to 0^+$*, logo concluímos que* $\partial^o f(0) = [0,1]$*.*

Proposição 2.3.3. $f^o(x;v) = máx\{\langle \tau, v\rangle; \tau \in \partial^o f(x)\}$*.*

Demonstração. Veja ($CLARKE$, 1975), Proposição 1.4. □

Proposição 2.3.4. *A função $g_x(v) = f^o(x;v)$ é convexa.*

Demonstração. Veja ($CLARKE$, 1975), Corolário 1.9. □

Exemplo 14. *Tome $f(x) = x^2 sen(\frac{1}{x})$, se $x \neq 0$ e $f(0) = 0$ temos que f é diferenciavél em 0, com $\nabla f(0) = 0$ entretanto $\partial f(0)$, não coincide com o gradiente habitual. Com efeito, tome $h_t = \frac{1}{2\pi t}$ e $l_t = \frac{1}{\pi + 2\pi t}$, assim $\nabla f(0+h_t) = 2(h_t)sen(\frac{1}{(h_t)}) - cos(\frac{1}{(h_t)}) = -cos(2\pi t) = -1$ e observando a outra sequência que também converge para zero, temos $\nabla f(0+l_t) = 2(l_t)sen(\frac{1}{(l_t)}) - cos(\frac{1}{(l_t)}) = -cos(\pi + 2\pi t) = 1$, assim podemos ver que $0 = \nabla f(0) \neq \partial^o f(0) \supset [-1,1]$.*

O exemplo acima mostra que se uma função é diferenciável em um ponto, isso não implica que a derivada de Clarke, nesse ponto coincide com a derivada no sentido clássico.

Como foi dito anteriormente a derivada de Clarke não generaliza a derivada no sentido clássico, pois temos casos onde as funções são diferenciáveis porém a derivada de Clarke não se reduz a derivada no sentindo clássico, entretanto a proposição a seguir, diz que se f é de classe C^1 então tal resultado é verdadeiro.

Proposição 2.3.5. *Se $f : \mathbb{R}^n \to \mathbb{R}$, de classe C^1 em um $x \in \mathbb{R}^n$ então $\partial^o f(x) = \nabla f(x)$.*

Demonstração. Temos por definição que $\partial^o f(a) = C(\{T; lim_{j\to\infty} \nabla f(a_j), a_j \in \mathbb{R}^n, a_j \to a\})$, pela continuidade de ∇f, temos que $lim_{j\to\infty} \nabla f(a_j) = \nabla f(a)$, portanto $\partial^o f(a) = C\{\nabla f(a)\} = \{\nabla f(a)\}$. □

Um resultado elementar de cálculo diferencial é que a derivada da soma de duas funções, é a soma das derivadas. Entretanto o resultado não é mais verdadeiro quando falamos de derivada de Clarke, como veremos no exemplo a seguir.

Exemplo 15. *Tome as funções $f(x) = |x|$ e $g(x) = -|x|$, logo temos que $f + g = 0$, portanto $\partial^o(f+g) = 0$, entretanto $\partial^o f = [-1,1]$ e $\partial^o g = [-1,1]$, o que mostra que a derivada da soma, não é a soma das derivadas.*

Infelizmente, não podemos garantir igualdade, mas podemos garantir pelo menos uma inclusão, como mostra a proposição adiante.

Proposição 2.3.6. *Sejam $f, g : \mathbb{R}^n \to \mathbb{R}$ funções localmente lipschitz, então $\partial^o(f+g) \subset \partial^o f + \partial^o g$.*

Demonstração. Veja ($CLARKE$, 1975), Proposição 1.12. □

Proposição 2.3.7. *$\partial^o f(x)$ é não-vazio, convexo e compacto.*

Demonstração. Veja Teorema 17.2 de ($ROCKAFELLAR$, 1970). □

Definição 2.3.5. *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}$, com U aberto e f localmente Lipschitz. Dizemos que um ponto $x \in U$ é um ponto crítico de Clarke, se $0 \in \partial^o f(x)$.*

Proposição 2.3.8. *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}$, localmente Lipschitz, $x \in U$ é um máximo ou mínimo local, então x é um ponto crítico.*

Demonstração. Para o caso mínimo local, veja Teorema 10.1 de ($ROCKAFELLAR; WETS$, 2009). Feito o caso mínimo local, basta usar a propriedade $\partial^{\circ}(\lambda f)(x) \subset \lambda \partial^{\circ} f(x)$, para verificar o caso máximo local, com $\lambda = -1$. □

A partir de agora falaremos um pouco sobre a Teoria de derivada de Clarke, para aplicações.

Definição 2.3.6. *O Jacobiano generalizado de f em x_0, denotado por $J^o f(x_0)$, é o fecho convexo de todas as matrizes da forma M, onde $M = \lim_{t\to\infty} Jf(x_t)$, onde x_t converge para x_0 e f é diferenciável em x_t para cada t.*

Exemplo 16. *Seja $f : \mathbb{R}^2 \to \mathbb{R}^2$, definida por $f(x,y) = (|x| + y, 2x + |y|)$, temos que o jacobiano generalizado $J^o f(0,0)$ é o fecho convexo dos limites das matrizes da forma $\begin{bmatrix} \partial^o(|x|) & 1 \\ 2 & \partial^o(|y|) \end{bmatrix}$, como já visto no exemplo 12 esse conjunto é $\left\{ \begin{bmatrix} t & 1 \\ 2 & u \end{bmatrix}; -1 \leq t, u \leq 1 \right\}$.*

Proposição 2.3.9. *Se $f : \mathbb{R}^n \to \mathbb{R}^m$ é de classe C^1 então $J^o f(x_0) = Jf(x_0)$.*

Demonstração. A demonstração desse fato é feito análogo ao caso escalar. □

Teorema 2.3.10 (Regra da Cadeia)**.** *Sejam V um subconjunto aberto de $\mathbb{R}^m$ e $G : V \to U$ de classe C^1 e $f : U \to \mathbb{R}$ localmente Lipschitz. Além disso, seja g(x)=f(G(x)), então*

$$\partial^o g(x) \subset \nabla G(x)^T \partial^o f(G(x)),$$

onde $\nabla G(x)^T$ denota a transposta da matriz jacobiana de G em x.

Demonstração. Veja Teorema 10.6 de ($ROCKAFELLAR; WETS$, 2009). □

Definição 2.3.7. *$J^o f(x_0)$ é dita de rank maximal se toda matriz M em $J^o f(x_0)$ é de rank maximal.*

Teorema 2.3.11. *Seja $f : \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz e $J^o f(x_0)$ é de rank maximal, então existe uma vizinhanças U e V de x_0 e $f(x_0)$ respectivamente, e uma função Lipschitz $g : V \to \mathbb{R}^n$ tal que,*
(a)g(f(u))=u para todo $u \in U$;
(b)f(g(u))=u para todo $v \in V$.

Demonstração. Veja Teorema 1 de $(CLARKE, 1976)$. □

Corolário 2.3.11.1. *Seja $f : \mathbb{R}^n \to \mathbb{R}^n$,$f \in C^1$ e $Jf(x_0)$ é invertível, então existem vizinhas U e V de x_0 e $f(x_0)$, respectivamente, e uma função g de classe C^1 tal que, f e g são inversas em U.*

Demonstração. Como $f \in C^1$, então pela Proposição 2.3.9 temos que $J^o f(x_0) = Jf(x_0)$, portanto tem rank maximal já que $Jf(x_0)$ é invertível, novamente pelo fato de $f \in C^1$ temos que f é localmente Lipschipz, logo pelo Teorema 2.3.11, existem U,V e g nas condições desejadas, pela prova do Teorema 2.3.11 vemos que g é tomada de forma que sua regularidade depende somente de f, como $f \in C^1$, então $g \in C^1$. □

O teorema a seguir é um análogo ao Teorema da Função Implícita, mas na versão Lipschitz.

Teorema 2.3.12 (Teorema da Função Implícita Lipschitz)**.** *Seja $f : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$ uma função Lipschitz e $f(x_0, y_0) = z_0$. Suponha que para qualquer $l \in \partial^o f(x_0, y_0)$ tem-se $\ker l \cap (\{0\}^m \times \mathbb{R}^k) = \{0\}^{m+k}$. Então, há uma vizinhança $U \times V$ de (x_0, y_0) e uma única função Lipschitz $g : U \to V$ tal que $f^{-1}(z_0) \cap (U \times V) = \nabla g$.*

Demonstração. Defina $F : \mathbb{R}^{m+k} \to \mathbb{R}^{m+k}$, dada por $F(x,y) = (x, f(x,y))$. Temos que $|F(x_1,y_1) - F(x_0,y_0)| = |(x_1, f(x_1,y_1)) - (x_0, f(x_0,y_0))| = \sqrt{(x_1-x_0)^2 + (f(x_1,y_1) - f(x_0,y_0))^2} \leq \sqrt{((x_1-x_0)^2 + (y_1-y_0)^2 + k^2((x_1-x_0)^2 + (y_1-y_0)^2))} = \sqrt{k^2+1}|(x_1,y_1) - (x_0,y_0)|$, ou seja, F é Lipschitz. Além disso $L \in \partial^o F(x_0,y_0)$, é da forma $L(x,y) = (x, l(x,y))$, logo L tem rank maximal, pelo Teorema da Aplicação Inversa Lipschitz, temos uma função $G : \mathbb{R}^{m+k} \to \mathbb{R}^{m+k}$, definida em uma vizinhança $W \times V$ e $U \times U'$ e $G \circ F$ e $F \circ G$ são inversas, seja $G(x,y) := (x, h(x,y))$ veja que $|h(x_1,y_1) - h(x_0,y_0)| \leq |(x_1, h(x_1,y_1)) - (x_0, h(x_0,y_0))| = |G(x_1,y_1) - G(x_0,y_0)|$ como G é Lipschitz, h é Lipschitz, defina $g : U \to V$ $g(x) := h(x, z_0)$, claramente g é Lipschitz e $g(x_0) = h(x_0, z_0) = y_0$, pelo fato de F e G serem inversas. Observe ainda que $(x, f(x, g(x))) = F(x, g(x)) = F(x, h(x, z_0)) = F(G(x, z_0)) = (x, z_0) \Rightarrow f(x, g(x)) = z_0$. □

2.3.1 Subgradiente de Frechet e subgradiente limitante

Definição 2.3.8. *Sejam $U \subset \mathbb{R}^n$ aberto e uma função $f : U \to \mathbb{R}$ localmente Lipschitz. Definimos o subgradiente de Fréchet de f em x, e denotamos por*

$$\widehat{\partial} f(x) = \left\{ x^* \in \mathbb{R}^n ; \liminf_{y \to x, y \neq x} \frac{f(y) - f(x) - \langle x^*, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

Proposição 2.3.13. *Se f é diferenciável, então $\widehat{\partial} f(x) = \{\nabla f(x)\}$.*

Demonstração. Veja Observação 1.(c) de $(BOLTE et\ al., 2005)$. □

Definição 2.3.9. *Sejam $U \subset \mathbb{R}^n$ aberto e uma função $f : U \to \mathbb{R}$ localmente Lipschitz. Definimos o subgradiente limitante de f em x, e denotamos por $\partial f(x)$,*

$$x^* \in \partial f(x) \Leftrightarrow \exists x_n \in U, \exists x_n^* \in \widehat{\partial} f(x_n) : x_n \to x, x_n^* \to x^* \text{ quando } n \to \infty$$

Teorema 2.3.14 (Regra da Cadeia para subgradientes)**.** *Sejam V um subconjunto aberto de $\mathbb{R}^m$ e $G : V \to U$ de classe C^1 e $f : U \to \mathbb{R}$ localmente Lipschitz. Além disso, seja g(x)=f(G(x)), então*

$$\widehat{\partial} g(x) \supset \nabla G(x)^T \widehat{\partial} f(G(x)),$$

$$\partial g(x) \subset \nabla G(x)^T \partial f(G(x)),$$

onde $\nabla G(x)^T$ denota a transposta da matriz jacobiana de G em x.

Demonstração. Veja Teorema 10.6 de $(ROCKAFELLAR; WETS, 2009)$. □

Definição 2.3.10. *Sejam X e Y dois conjuntos não vazios. Uma multivalorada $T : X \rightrightarrows Y$, é uma regra que permite associar, de modo bem determinado, a cada elemento $x \in X$, um $Z \subset Y$.*

Exemplo 17. *Tome $T : \mathbb{Z} \rightrightarrows \mathbb{Z}$ dada por, $T(k) = \{p \in \mathbb{Z}; p = k - 5q\}$, com $q \in \mathbb{Z}$. T é uma multivalorada.*

Exemplo 18. *Seja $P : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, dada por $P(x) = \widehat{\partial} f(x)$. P é uma multivalorada.*

Exemplo 19. *Seja $Q : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, dada por $Q(x) = \partial f(x)$. Q é uma multivalorada.*

Exemplo 20. *Seja $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, dada por $S(x) = \partial^o f(x)$. S é uma multivalorada.*

Definição 2.3.11. *Seja $T : U \rightrightarrows \mathbb{R}^n$ uma multivalorada, definimos*

$$dom(T) := \{x \in U : T(x) \neq \varnothing\}$$

Definição 2.3.12. *Seja $T : U \rightrightarrows \mathbb{R}^n$ uma multivalorada, definimos*

$$Graf(T) := \{(x,y) \in U \times \mathbb{R}^n : y \in T(x)\}.$$

Teorema 2.3.15. *Seja $f : U \to \mathbb{R}$ localmente Lipschitz. A multivalorada que associa cada $x \in U$ a $\partial^o f(x)$ é limitada em qualquer subconjunto compacto de U.*

Demonstração. Veja Observação 1.(b) de $(BOLTEet\ al., 2005)$. □

2.4 Conjuntos e funções subanalíticas

Definição 2.4.1. *Um conjunto $A \subset \mathbb{R}^n$ é chamado semialgébrico básico, se existem funções polinomiais f, g_i, com $1 \leq i \leq n$, tais que:*

$$A = \{x \in \mathbb{R}^n ; f(x) = 0\} \cap (\cap_{i=1}^n \{x \in \mathbb{R}^n ; g_i(x) > 0\}).$$

Definição 2.4.2. *Um conjunto semialgébrico é a reunião finita de conjuntos semialgébricos básicos.*

Exemplo 21. *A bola euclidiana $B = \{x \in \mathbb{R}^n ; \|x\| < 1\}$, é semialgébrica. Basta tomar $f(x) \equiv 0$ e $g(x_1, x_2, ..., x_n) = 1 - x_1^2 - x_2^2 - ... - x_n^2$, onde $x = (x_1, x_2, ..., x_n)$.*

Definição 2.4.3. *Um conjunto $A \subset \mathbb{R}^n$ é chamado semianalítico básico se para todo $x \in \mathbb{R}^n$ existem uma vizinhança U_x de x e funções analítica f, g_i, com $1 \leq i \leq n$ analíticas nessa vizinhança tal que:*

$$A \cap U_x = \{x \in \mathbb{R}^n ; f(x) = 0\} \cap (\cap_{i=1}^n \{x \in \mathbb{R}^n ; g_i(x) > 0\}).$$

Definição 2.4.4. *Um conjunto semianalítico é a reunião finita de conjuntos semianalíticos básicos.*

Exemplo 22. *Todo conjunto analítico X, é semianalítico. Com efeito, basta tomar a f da definição como a soma dos quadrados das funções analíticas que definem X e $g \equiv 1$. Em particular, como todo algébrico é analítico, então todo algébrico é semianalítico.*

Exemplo 23. *Todo conjunto semialgébrico é semianalítico, pois todo polinômio é uma função analítica.*

Exemplo 24. *Seja $g : \mathbb{R}^2 \to \mathbb{R}$ dada por $g(x,y) = e^x - y$, que é uma função analítica, logo $A = \{(x,y) \in \mathbb{R}^2 ; g(x,y) > 0\}$ é um conjunto semianalítico.*

Definição 2.4.5. *Seja $A \subset \mathbb{R}^n$ um conjunto semianalítico. Dizemos que $F : A \subset \mathbb{R}^n \to \mathbb{R}$ é uma aplicação semianalítica se $Graf(F) \subset \mathbb{R}^{n+1}$ é semianalítico.*

Exemplo 25. *A função $f : \mathbb{R}_+ \to \mathbb{R}$, dada por $f(x) = \sqrt{x}$ é semianalítica. Tome $F(x,y) = x - y^2$ e $g(x,y) = y$, logo temos que $Graf(f) = \{(x,y) \in \mathbb{R}^2; F(x,y) = 0 \text{ e } g(x,y) > 0\} \cup \{(0,0)\}$, logo f é semianalítica.*

Exemplo 26. *A função $f : \mathbb{R} \to \mathbb{R}$, dada por $f(x) = |x|$ não é analítica, no entanto, tome $F(x,y) = x^2 - y^2$ e $g(x,y) = y$, logo temos que $Graf(f) = \{(x,y) \in \mathbb{R}^2; F(x,y) = 0 \text{ e } g(x,y) > 0\} \cup \{(0,0)\}$, logo f é semianalítica.*

Definição 2.4.6. *Um subconjunto $A \subset \mathbb{R}^n$ é dito subanalítico, se para todo $x \in \mathbb{R}^n$ admite uma vizinhança U_x de x, tal que $A \cap U_x$ é a projeção linear e própria de um subconjunto semianalítico. Em outras palavras, existe $B \subset \mathbb{R}^m$ semianalítico, $n \leq m$, tal que a projeção $\pi : \mathbb{R}^m \to \mathbb{R}^n$ restrita a B é uma projeção própria, linear e $A \cap U_x = \pi(B)$.*

Exemplo 27. *Todo semianalítico A é subanalítico, basta tomar a projeção, sendo a aplicação identidade. Claramente a aplicação identidade é própria e linear.*

Proposição 2.4.1. *Sejam A e B subconjuntos do $\mathbb{R}^n$ subanalíticos, então $A \cup B$ e $A \cap B$ são subanalíticos.*

Demonstração. Veja Teorema 1.2.2-(i) de $(TAMM, 1981)$.

□

Proposição 2.4.2. *Sejam $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}^m$ subanalíticos, então $A \times B \subset \mathbb{R}^n \times \mathbb{R}^m$ é subanalítico.*

Demonstração. Veja Teorema 1.2.2-(ii) de $(TAMM, 1981)$.

□

Teorema 2.4.3 (Teorema de Gabrielov)**.** *O complemento de um subanalítico em $\mathbb{R}^n$ é um subanalítico.*

Demonstração. Veja Teorema 3.10 de $(BIERSTONE; MILMAN, 1988)$.

□

Proposição 2.4.4. *Seja $X \subset \mathbb{R}^n$. Suponha que $X \cap B$ é subanalítico para todo semialgébrico limitado B. Então X é subanalítico.*

Demonstração. Suponha que X não é subanalítico. Então, existe $x_0 \in \mathbb{R}^n$ tal que para todo aberto U contendo x_0, não existe semianalítico S e uma projeção própria $\pi : S \to U \cap X$ sobrejetiva.

Tome $B = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^n; |x_i| < 2|x_0|\}$, por hipotese temos que $X \cap B$ é subanalítico. Portanto existe $V \subset \mathbb{R}^n$ contendo x_0 e um semianalítico $S \subset \mathbb{R}^m$ tal que $\pi : S \to (X \cap B) \cap V$ é sobrejetiva e própria. Mas como $W = B \cap V$ é um aberto, temos que existe uma projeção própria e sobrejetiva $\pi : S \to X \cap W$, absurdo. □

Exemplo 28. *O gráfico da função $f : \mathbb{R}^*_+ \to \mathbb{R}$, dada por $f(x) = sen(\frac{1}{x})$ em $\mathbb{R} \setminus \{0\}$ é um conjunto analítico em $\mathbb{R}^*_+ \times \mathbb{R}$, pois é os zeros da função $F(x, y) = y - f(x)$, que é analítica, pois f é analítica.*

Exemplo 29. *Sejam $p(x, y) = y$ e $q(x, y) = x$. O conjunto $A = \{(x, y) \in \mathbb{R}^2; x > 0 \text{ e } y = 0\}$ é subanalítico, pois $A = B \cap C$, onde $B = \{(x, y) \in \mathbb{R}^2; p(x, y) = 0\}$ que é analítico, em particular subanalítico, e C=$\{(x, y) \in \mathbb{R}^2; q(x, y) > 0\}$ que é semianalítico, em particular subanalítico.*

Os dois exemplos acima, foram abordados para da suporte ao próximo exemplo, que terá importância teórica. Em Geometria Algébrica Real, os conjuntos semialgébricos da reta, podem ser descritos como união finita de pontos e intervalos, no entanto isso não é verdade para conjuntos subanalíticos, como veremos no exemplo abaixo. Mas se adicionarmos a hipótese de limitação teremos tal resultado.

Exemplo 30. *Sejam V e W os conjuntos dos exemplos* 28 *e* 29 *respectivamente. Logo temos que $Y = V \cap W$ é subanalítico, mas Y é dado por uma união infinita de pontos.*

Proposição 2.4.5. *A família de componentes conexas de um conjunto subanalítico é localmente finita em cada ponto do $\mathbb{R}^n$ e cada componente conexa é subanalítica.*

Demonstração. Veja Propriedade $I.2.1.3$ de $(SHIOTA, 2012)$ □

Proposição 2.4.6. *O número de componentes conexas de um conjunto subanalítico $X \subset \mathbb{R}^n$ é enumerável.*

Demonstração. É possível tomar um conjunto enumerável e denso em X, assim pela proposição anterior, temos um número finito de componentes conexa intersectando uma vizinhança desse ponto, que por sua vez é enumerável. □

Proposição 2.4.7. *Se $X \subset \mathbb{R}^n$ é um conjunto subanalítico e limitado, então X tem um número finito de componentes conexas.*

Demonstração. Suponha que X tenha uma quantidade infinita de componentes conexas. Tome um ponto em cada componente conexa, formando uma sequência (x_n). Como X é limitado, existe $a \in X$ que é um ponto de acumulação de (x_n), logo dado qualquer vizinhança de a, essa vizinhança tem uma infinidade de pontos de (x_n), ou seja, intersecta uma quantidade infinita de componentes conexa, contradizendo Proposição 2.4.5. Absurdo, que veio de supor que X tem uma quantidade infinita de componentes conexas. $\square$

Proposição 2.4.8. *Um subconjunto subanalítico e limitado da reta é a união finita de intervalos abertos, fechados, semiabertos e pontos.*

Demonstração. Como os conjuntos conexos da reta, são os intervalos, o resultado segue de Proposição 2.4.7. $\square$

Existe a noção de função subanalítica, sem que seu domínio seja subanalítico, porém aqui sempre assumiremos que o domínio de uma função subanalítica é subanalítico. Isso nos leva a seguinte definição.

Definição 2.4.7. *Seja $A \subset \mathbb{R}^n$ subanalítico. Dizemos que a função $F : A \longrightarrow \mathbb{R}$ é subanalítica se seu gráfico $Graf(F) = \{(x, F(x)); x \in A\} \subset \mathbb{R}^{n+1}$ é um conjunto subanalítico.*

Definição 2.4.8. *Sejam $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}^m$. Uma aplicação $F : A \to B$ é subanalítica se suas funções componentes são subanalíticas.*

Proposição 2.4.9. *Sejam $f : A \to \mathbb{R}^n$ tal que $f(A) \subset B$ e $g : B \to \mathbb{R}^m$ funções subanalíticas. Se $f(A \cap D)$ é limitada para qualquer conjunto limitado $D \subset \mathbb{R}^n$, então $g \circ f$ é uma função subanalítica.*

Demonstração. Veja Propriedade I.2.1.10 de ($SHIOTA$, 2012). $\square$

Proposição 2.4.10. *Se $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ é de classe C^1 e subanalítica, tal que para qualquer conjunto limitado $B \subset \mathbb{R}^n$, $f(B \cap U)$ é limitado. Então Df é uma aplicação subanalítica.*

Demonstração. Veja Propriedade I.2.1.13 de ($SHIOTA$, 2012). $\square$

Proposição 2.4.11. *Se f é subanalítica, então $\partial^o f(x)$, $\widehat{\partial} f(x)$ e $\partial f(x)$ tem o gráfico subanalítico.*

Demonstração. Veja Observação 1.(d) de ($BOLTE$ *et al.*, 2005). $\square$

Proposição 2.4.12. *Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}$ subanalítica e $S = \{x \in U; 0 \in \partial^o f(x)\}$. Então, S é um conjunto subanalítico.*

Demonstração. Temos que U é subanalítico por definição, $\{0\}_n$ também é subanalítico, pois é um ponto e $Graf(\partial^o f)$ é subanalítico pela Proposição 2.4.11. Como temos que produto cartesiano e intesecção de conjuntos subanalíticos é subanalítico, segue que S é subanalítico, pois $S = (U \times \{0\}_n) \cap Graf(\partial^o f)$. □

Proposição 2.4.13. *Cada componente conexa de um conjunto subanalítico é um conjunto subanalítico e além disso é subanaliticamente conexa por caminhos, istó é, existe um caminho subanalítico contínuo ligando quaisquer dois pontos da componente conexa.*

Demonstração. Veja Proposição 3 de ($BOLTEet\ al.$, 2006). □

Agora introduziremos outra classe de conjuntos importante, que é a classe de globalmente subanalíticos.

Definição 2.4.9. *Seja* $\Phi : \mathbb{R}^k \to \mathbb{R}^k$ *a aplicação dada por*

$$\Phi(x_1, x_2, \ldots, x_k) = \left(\frac{x_1}{\sqrt{1+x_1^2}}, \ldots, \frac{x_k}{\sqrt{1+x_k^2}} \right)$$

Um conjunto $X \subset \mathbb{R}^n$ *é dito globalmente subanalítico, se* $\Phi(X)$ *é um conjunto subanalítico.*

Proposição 2.4.14. *Todo conjunto globalmente subanalítico é subanalítico.*

Demonstração. Veja ($BOLTEet\ al.$, 2007). □

Teorema 2.4.15. *Um conjunto* $X \subset \mathbb{R}^n$ *subanalítico e limitado é globalmente subanalítico.*

Demonstração. Veja ($BOLTEet\ al.$, 2007). □

Definição 2.4.10. *Seja X um conjunto globalmente subanalítico. Uma aplicação* $f : X \subset \mathbb{R}^m \to \mathbb{R}^n$ *é dita globalmente subanalítico, se* $Graf(f) \subset \mathbb{R}^{m+n}$ *é um conjunto globalmente subanalítico.*

Teorema 2.4.16 (Lema da Monotonicidade)**.** *Seja* $\alpha < \beta \in \mathbb{R}$. *Se* $\phi : (\alpha, \beta) \to \mathbb{R}$ *é uma função globalmente subanalítica, então existe uma partição* $t_0 := \alpha < t_1 < \ldots < t_{l+1} := \beta$ *de* (α, β), *tal que* $\phi|_{(t_i, t_{i+1})}$ *é* C^∞ *e ou é constante ou é estritamente monótona, para* $i \in \{0, \ldots, l\}$.

Demonstração. Veja ($BIERSTONE; MILMAN$, 1988). □

3 VERSÃO LIPSCHITZ DO TEOREMA DE SARD

Nesse capítulo, mostraremos o principal resultado do trabalho, a saber, o Teorema de Sard em sua versão Lipschitz, ou seja, que a imagem do pontos críticos de Clarke por uma função subanalítica e localmente Lipschitz tem medida de Lebesgue nula. E finalizaremos com um exemplo que mostra que o resultado não é verdadeiro se a função não for localmente Lipschitz.

Lema 3.0.1 (Lema da perturbação para caminhos). *Seja $F \subset \mathbb{R}^n$ um conjunto não-vazio, globalmente subanalítico, $\gamma : [0,1] \to \overline{F}$ um caminho subanalítico contínuo injetivo e $\eta > 0$. Então há um caminho subanalítico contínuo $z : [0,1] \to \overline{F}$ tal que:*
(i) $|z'(t) - \gamma'(t)| < \eta$ para quase todo $t \in (0,1)$;
(ii) O conjunto $\Delta := \{t \in [0,1] : z(t) \in \overline{F} \setminus F\}$ tem medida de Lebesgue menor que η;
(iii) $z(t) = \gamma(t)$ para todo $t \in \Delta \cup \{0,1\}$.

Demonstração. Veja Lema 12 de (*BOLTE et al.*, 2006). □

Observação. *Observe que no Lema 3.0.1 podemos usar qualquer intervalo $[a,b]$, com $a < b$, no lugar de $[0,1]$.*

Teorema 3.0.2 (Teorema de Sard para pontos críticos limitantes). *Seja $f : U \to \mathbb{R}$ uma função subanalítica contínua. Então, f é constante em cada componente conexa do conjunto dos pontos críticos limitantes, a saber,*

$$(\partial f)^{-1}(0) = \{x \in U; 0 \in \partial f(x)\}.$$

Demonstração. Veja Teorema 13 de (*BOLTE et al.*, 2006). □

Proposição 3.0.3. *Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}$ subanalítica e localmente Lipschitz e $e := (1,0,\ldots,0)$. Defina $\Gamma_\delta = \{x \in [0,1]e; \forall x^* \in \partial f(x), |\langle x^*, e\rangle| > \delta\}$ é subanalítico.*

Demonstração. Inicialmente temos que $Graf(\partial f)$ é subanalítico, pela Proposição 2.4.11. Além disso, os conjuntos $A = \{y \in \mathbb{R}^n; |\langle y, e\rangle| > \delta\}$ e $B = \{te; 0 \le t \le 1\}$ são semialgébricos. Como $C = (Graf(\partial f)) \cap (\mathbb{R}^n \times A)$ é subanalítico, pois pelas Proposições 2.4.1 e 2.4.2 a intersecção de subanalíticos é subanalítico e o produto cartesiano de conjuntos subanalíticos é subanalítico.

Seja $\pi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, dada por $\pi(x,y) = x$. Vamos mostrar que π é própria. Como π é contínua, resta mostrar que a imagem inversa de limitado é limitado. Tome W_1 limitado, como f é localmente Lipschitz, segue do Teorema 2.3.15, que $|x^*| \le M$, para todo $x \in W_1$. Assim π é própria, logo $\pi(C)$ é subanalítica. Para finalizar observe que $\pi(C) \cap [0,1]e = \Gamma_\delta$. □

Lema 3.0.4. *Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}$ uma função subanalítica, localmente Lipschitz e $e := (1,0,...,0) \in \mathbb{R}^n$ e assuma que $[0,1]e \subset U$, com $0 \in \partial^o f(te)$ para todo $t \in [0,1]$. Então Γ_δ é um conjunto finito.*

Demonstração. A demonstração será feita por contradição. Suponha que Γ_δ seja infinito. Como é subanalítico da reta e limitado, então é a união finita de intervalos e pontos, então existe um intervalo $(a,b) \subset [0,1]$ tal que $(a,b)e \subset \Gamma_\delta$. Tome V um subconjunto aberto limitado de U tal que $[0,1]e \subset V \subset \overline{V} \subset U$ e defina

$$\widehat{\Gamma_\delta}^+ = \{x \in \overline{V}; \exists x^* \in \widehat{\partial} f(x), \langle x^*, e\rangle > \delta\}$$

e

$$\widehat{\Gamma_\delta}^- = \{x \in \overline{V}; \exists x^* \in \widehat{\partial} f(x), \langle x^*, e\rangle < -\delta\}$$

onde $\widehat{\partial} f(x)$ denota a derivada de Fréchet de f. Como para cada $x \in \Gamma_\delta$, tem-se $0 \in \partial^o f(x) = \overline{co}\partial f(x)$ e $0 \notin \partial f(x)$, então obtemos o seguinte

$$max\{\langle x^*, e\rangle; x^* \in \partial f(x)\} > \delta \text{ e } min\{\langle x^*, e\rangle; x^* \in \partial f(x)\} < -\delta.$$

Além do mais, veja que $(a,b)e \subset \overline{\widehat{\Gamma_\delta}^+}$, pois dado $t \in (a,b)e \subset \Gamma_\delta$ existe $t^* \in \partial f(t)$ tal que $\langle t^*, e\rangle > \delta$, pelo máximo estabelecido acima, além disso, $t^* \in \partial f(t)$ significa que existem $(t_n) \subset V$ e $t_n^* \in \widehat{\partial} f(t_n)$ onde $t_n \to t$ e $t_n^* \to t^*$, portanto

$$\delta < \langle t^*, e\rangle = \langle \lim_{n\to\infty} t_n^*, e\rangle = \lim_{n\to\infty} \langle t_n^*, e\rangle,$$

logo existe n_0 tal que $\langle t_n^*, e\rangle > \delta$ para todo $n \geq n_0$, o que implica em $t_n \in \widehat{\Gamma_\delta}^+$ para todo $n \geq n_0$ e, assim, $t \in \overline{\widehat{\Gamma_\delta}^+}$, mostrando que $(a,b)e \subset \overline{\widehat{\Gamma_\delta}^+}$. De modo análogo, temos que $(a,b)e \subset \overline{\widehat{\Gamma_\delta}^-}$.

Como $\widehat{\Gamma_\delta}^+$ é um subconjunto de $\mathbb{R}^n$ não vazio, subanalítico e limitado (portanto globalmente subanalítico) e $\gamma(t) = te$, $t \in (a,b)$, é um caminho subanalítico injetivo e contínuo, tomando $\eta > 0$, como estamos nas hipoteses do Lema de pertubação para caminhos, temos que existe um caminho $z : [a,b] \to \overline{\widehat{\Gamma_\delta}^+}$ contínuo e subanalítico, tal que

- $\|z'(t) - e\| < \eta$ para quase todo $t \in (a,b)$,
- O conjunto subanalítico $\Delta := \{t \in [a,b] | z(t) \in \overline{\widehat{\Gamma_\delta}^+} \setminus \widehat{\Gamma_\delta^+}\}$ tem medida de Lebesgue menor que η.
- $z(t) = \gamma(t)$ para todo $t \in \Delta \cup \{a,b\}$.

Pela Proposição 2.4.9, a função $f(z(t))$ é subanalítica, além de ser contínua em $[a,b]$, portanto tem o gráfico subanalítico limitado, o que implica pelo Teorema 2.4.15 ser globalmente subanalítico. Seja $g(t) = f(z(t))$, pelo fato de ser globalmente subanalítica, temos que $g \in C^\infty$ em $(a,b) \setminus \Delta$ a menos de um número finito de pontos, relembre o Teorema 2.4.16. Logo nesses pontos, pela Proposição 2.3.13, temos que $\{g'(t)\} = \widehat{\partial} g(t)$. Pela regra da cadeia, Proposição 2.3.14, temos $\{g'(t)\} = \widehat{\partial} g(t) \supset \langle z'(t), \widehat{\partial} f(z(t))\rangle \supset \{\langle z'(t), z_+^*(t)\rangle\}$, onde $\langle z'(t), \widehat{\partial} f(z(t))\rangle = \{\langle z'(t), w\rangle; w \in \widehat{\partial} f(z(t))\}$ e $z_+^*(t) \in \widehat{\partial} f(z(t))$. Logo, para todo $[a,b] \setminus \Delta$ a menos de uma quantidade finita, temos

$$g'(t) = \langle z'(t), z_+^*(t)\rangle = \langle e + z'(t) - e, z_+^*(t)\rangle = \langle z'(t) - e, z_+^*(t)\rangle + \langle e, z_+^*(t)\rangle.$$

Podemos fazer $\langle e, z_+^*(t)\rangle > \delta$, basta tomar $z(t) \in \widehat{\Gamma_\delta^+}$, logo

$$g'(t) = \langle z'(t) - e, z_+^*(t)\rangle + \langle e, z_+^*(t)\rangle > \delta + \langle z'(t) - e, z_+^*(t)\rangle,$$

além do mais, podemos estimar

$$\langle z'(t) - e, z_+^*(t)\rangle = -\langle -z'(t) + e, z_+^*(t)\rangle \geq -\|\langle -z'(t) + e, z_+^*(t)\rangle\| \geq -\| - z'(t) + e\|\|z_+^*(t)\|,$$

pelo Teorema 2.3.15, temos que $sup\{\|x^*\| : x^* \in \partial^o f(x), x \in \overline{V}\}$ é finito, chamaremos de M tal supremo, mais pelo Lema da pertubação para caminhos segue que

$$\langle z'(t) - e, z_+^*(t)\rangle \geq -\| - z'(t) + e\|\|z_+^*(t)\| \geq -\eta M,$$

sendo assim podemos concluir que

$$g'(t) = \langle z'(t) - e, z_+^*(t)\rangle + \langle e, z_+^*(t)\rangle \geq \delta - \eta M,$$

Calculando,

$$f(be) - f(ae) = \int_a^b \frac{d}{dt} f(z(t))dt = \int_{[a,b]\setminus\Delta} \frac{d}{dt} f(z(t))dt + \int_\Delta \frac{d}{dt} f(z(t))dt,$$

como

$$\int_\Delta \frac{d}{dt} f(z(t))dt = -\int_\Delta -\frac{d}{dt} f(z(t))dt \geq -\int_\Delta \left|\frac{d}{dt} f(z(t))\right| dt$$

e pela regra da cadeia, temos

$$\int_u^v \left|\frac{d}{dt} f(z(t))\right| dt \leq (v-u) sup\{|\langle e, x^*\rangle| : t \in [u,v], x^* \in \partial f(te)\} \leq (v-u)M$$

lembre-se de que $g(t) = f(z(t))$ e então somando os fatos da medida de Δ ser menor que η, $g'(t) \geq \delta - \eta M$ e $\left|\frac{d}{dt}f(z(t))\right| \leq M$, segue que

$$f(be) - f(ae) \geq \int_{[a,b]\setminus\Delta} \frac{d}{dt}f(z(t))dt - \int_{\Delta}\left|\frac{d}{dt}f(z(t))\right|dt \geq (l-\eta)(\delta - \eta M) - \eta M,$$

onde $l = b - a$. Veja que se tomarmos η suficientemente pequeno, temos que $(l-\eta)(\delta-\eta M) - \eta M > 0$, pois $l\delta > 0$ e portanto teremos $f(be) > f(ae)$. Repetindo o mesmo argumento para $\widehat{\Gamma_\delta}^{-}$ temos

$$f(be) - f(ae) \leq (l-\eta)(\eta M - \delta) + \eta M,$$

novamente tomando η suficientemente pequeno, segue que $(l-\eta)(\eta M - \delta) + \eta M < 0$ e consecutivamente $f(ae) > f(be)$, absurdo, que veio do fato de supor Γ_δ infinito, o que finaliza nossa prova. $\square$

Teorema 3.0.5. *Sejam $f : U \subset \mathbb{R}^n \to \mathbb{R}$ uma função subanalítica, localmente Lipschitz. Seja $e := (1,0,...,0) \in \mathbb{R}^n$ e assuma que $[0,1]e \subset U$, com $0 \in \partial^o f(te)$ para todo $t \in [0,1]$. Então f é constante em $[0,1]e$.*

Demonstração. Seja $S_L = \{x \in [0,1]e; 0 \in \partial f(x)\}$, onde ∂f denota o limitante subgradiente de f. Observe que ele é limitado e pela Proposição 2.4.11, o conjunto S_L é subanalítico, logo pela Proposição 2.4.8 é a união finita de pontos e intervalos. Pelo Teorema 3.0.2, nós temos que f é constante em cada componente conexa de S_L. Como f é contínua é suficiente mostrar que f é constante em cada segmento não-degenerado de $[0,1]e \setminus S_L$. Suponha sem perda de generalidade que S_L é vazio, ou seja, $0 \notin \partial f(te)$, $\forall t \in [0,1]$. Tome $\delta > 0$ e seja

$$\Gamma_\delta = \{x \in [0,1]e; \forall x^* \in \partial f(x), |\langle x^*, e\rangle| > \delta\},$$

pelo Lema 3.0.4 vimos que ele é finito. E considere também o conjunto

$$\Gamma_0 = \{x \in [0,1]e; \exists x^* \in \partial f(x), \langle x^*, e\rangle = 0\},$$

como o subgradiente limitante é um conjunto construído com seus valores de aderência, ele é um conjunto fechado, ou seja, $\partial f(te)$ é fechado $\forall t \in [0,1]$. Veja que $[0,1]e = \Gamma_0 \cup \left(\bigcup_{1\leq i} \Gamma_{\frac{1}{i}}\right)$ e novamente pelo Lema 3.0.4, cada um dos $\Gamma_{\frac{1}{i}}$ é finito. Então, $\bigcup_{1\leq i}\Gamma_{\frac{1}{i}}$ é enumerável e coincide com o conjunto $[0,1]e \setminus \Gamma_0$ que pela Proposição 2.4.3 é subanalítico, além disso é limitado. Portanto pela Proposição 2.4.8 somado ao lema anterior $\bigcup_{1\leq i}\Gamma_{\frac{1}{i}}$ é um conjunto finito. Assim

$\{t \in [0,1] : te \in \Gamma_0\}$ é uma união finita de intervalos, cujo complementar em [0,1] é finito, vamos mostrar que f é constante em cada intervalo de Γ_0. Dado um $\varepsilon > 0$ arbitrário, defina

$$\widehat{\Gamma}_0^{\varepsilon} := \{x \in \overline{V} : \exists x^* \in \widehat{\partial} f(x), |\langle x^*, e\rangle| < \varepsilon\},$$

onde $\overline{V}$ é o mesmo do lema anterior.

Tomemos um intervalo $(a,b)e \subset \Gamma_0$. Dado $x \in (a,b)e \subset \Gamma_0$ temos que existe $x^* \in \partial f(x)$ tal que $\langle x^*, e\rangle = 0$ pela definição de subgradiente limitante existem $x_n \in V$ e $x_n^* \in \widehat{\partial} f(x_n)$ tal que $x_n \to x$ e $x_n^* \to x^*$, logo temos

$$|\langle x_n^*, e\rangle| \leq |\langle x_n^* - x^*, e\rangle| + |\langle x^*, e\rangle| \leq |x_n^* - x^*||e| + |\langle x^*, e\rangle| < \varepsilon,$$

para n suficientemente grande, temos que $x_n \in \widehat{\Gamma}_0^{\varepsilon}$, o que implica $(a,b)e \subset \overline{\widehat{\Gamma}_0^{\varepsilon}}$.
Aplicando o Lema da pertubação para caminhos, para o conjunto $\widehat{\Gamma}_0^{\varepsilon}$, para $\eta < \varepsilon$ e para o caminho $\gamma(t) = te$, então existe $z : [a,b] \to \widehat{\Gamma}_0^{\varepsilon}$ e um conjunto $\Delta \subset [a,b]$.

Seja $h(t) = f(z(t))$, pelos mesmos argumentos feitos no lema anterior, temos que $\{h'(t)\} = \{\langle z'(t), z_\varepsilon^*(t)\rangle\}$ em $[a,b] \setminus \Delta$ a menos de um número finito, onde $z_\varepsilon^*(t) \in \widehat{\partial} f(z(t))$ pode ser tomado em $\widehat{\Gamma}_0^{\varepsilon}$ para que $|\langle z_\varepsilon^*(t), e\rangle| < \varepsilon$, sendo assim

$$|h'(t)| = |\langle z'(t), z_\varepsilon^*(t)\rangle| = |\langle e + z'(t) - e, z_\varepsilon^*(t)\rangle| = |\langle e, z_\varepsilon^*(t)\rangle + \langle z'(t) - e, z_\varepsilon^*(t)\rangle|$$

pela desigualdade triangular e consecutivamente pela desigualdade de Cauchy-Schwarz, temos

$$|h'(t)| \leq |\langle e, z_\varepsilon^*(t)\rangle| + |\langle z'(t) - e, z_\varepsilon^*(t)\rangle| \leq \varepsilon + \eta M,$$

onde M é uma constante que limita $|\langle e, x^*\rangle|$ e a última desigualdade vem do fato de $z_\varepsilon^*(t) \in \widehat{\Gamma}_0^{\varepsilon}$, somado ao Lema da pertubação para caminhos para $\gamma(t) = te$, recorde que segue do Lema que existe um caminho $z : [a,b] \to \overline{\widehat{\Gamma_0}}^{\varepsilon}$ contínuo e subanalítico, tal que

- $\|z'(t) - e\| < \eta$ para quase todo $t \in (a,b)$,
- O conjunto subanalítico $\Delta := \{t \in [a,b] | z(t) \in \overline{\widehat{\Gamma_0}}^{\varepsilon} \setminus \widehat{\Gamma}_0^{\varepsilon}\}$ tem medida de Lebesgue menor que η.
- $z(t) = \gamma(t)$ para todo $t \in \Delta \cup \{a,b\}$.

Lembrando que é para um $\eta < \varepsilon$. Além do mais, temos ainda que

$$|f(be) - f(ae)| \leq \int_a^b \left|\frac{d}{dt} f(z(t))\right| dt = \int_{[a,b]\setminus\Delta} \left|\frac{d}{dt} f(z(t))\right| dt + \int_\Delta \left|\frac{d}{dt} f(z(t))\right| dt,$$

como $h(t) = f(z(t))$ e $\int_\Delta \left|\frac{d}{dt} f(z(t))\right| dt \leq \eta M$, segue que

$$|f(be) - f(ae)| \leq (l - \eta)(\varepsilon + \eta M) + \eta M,$$

tomando ε suficientemente pequeno, o que implica η pequeno, pois $\eta < \varepsilon$, temos que $f(be) = f(ae)$, como $a \leq b$ foram tomados quaisquer em Γ_0, temos que f é constante em cada componente conexa de Γ_0 e pela continuidade de f, temos que f é constante em todo $[0,1]e$, o que conclui nossa demonstração. □

Teorema 3.0.6. *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e não vazio e $f : U \longrightarrow \mathbb{R}$ uma aplicação localmente lipschitz subanalítica. Seja S o conjunto dos pontos críticos de Clarke de f, ou seja,*

$$S := \{x \in U | 0 \in \partial^o f(x)\}.$$

Então, f é constante em cada componente conexa de S.

Demonstração. Tomemos dois pontos x e y em uma mesma componente conexa de S. Pela Proposição 2.4.12 temos que S é subanalítico, então pela Proposição 2.4.13 cada componente conexa de S é subanaliticamente conexa por caminhos, sendo assim existe um caminho contínuo e subanalítico $\gamma : [0,1] \to S$ ligando x e y. É suficiente provarmos que f é constante em $\gamma((0,1))$, pois pela continuidade de f, teremos f constante em $\gamma([0,1])$ o que implica $f(x) = f(y)$.

Podemos também assumir que $\gamma((0,1))$ é uma subvariedade subanalítica, pois pelo Lema da monotonicidade $\gamma((0,1))$ é a união finita de subvariedades subanalítica, nós podemos verificar f constante em cada subvariedade, pela continuidade de f, será constante em todo $\gamma((0,1))$. Além disso existe um difeomorfismo subanalítico G de uma vizinhança V de $\gamma((0,1))$ sobre um subconjunto aberto $W \subset \mathbb{R}^n$ tal que $G(\gamma((0,1))) = (0,1)e$.

Como queremos mostrar que f é constante em $\gamma((0,1))$, é suficiente mostrar que $f \circ G^{-1}$ é constante em $(0,1)e$, o que podemos concluir via o Teorema 3.0.5, se mostrarmos que $(0,1)e \subset \{x \in V; 0 \in \partial^o(f \circ G^{-1})(G(x))\}$. Contudo, como

$$0 \in \partial^o f(x) \Leftrightarrow 0 \in \partial^o[f \circ G^{-1}](G(x)),$$

é suficiente mostrar que $\gamma((0,1)) \subset \{x \in U; 0 \in \partial^o f(x)\}$, o que é verdade, pois γ é uma curva cujo contradomínio é S.

□

Teorema 3.0.7 (Teorema de Sard para pontos críticos de Clarke)**.** *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e não vazio e $f : U \longrightarrow \mathbb{R}$ uma aplicação localmente lipschitz subanalítica. Seja S o conjunto dos pontos críticos de Clarke de f, ou seja,*

$$S := \{x \in U | 0 \in \partial^o f(x)\},$$

o conjunto $f(S)$ dos valores críticos de Clarke de f é enumerável e consequentemente tem medida zero. Além disso, se U é limitado, então $f(S)$ é um conjunto finito.

Demonstração. Lembre-se de que S é subanalítico (para isso veja Proposição 2.4.12). Pela Proposição 2.4.6, S possui uma quantidade enumerável de componentes conexas e pelo Teorema anterior, f é constante em cada componente conexas, tornando assim $f(S)$ em apenas uma quantidade enumerável de pontos. Em particular, $f(S)$ possui medida nula. No caso U limitado, pela Proposição 2.4.7, temos que U tem uma quantidade finita de componentes conexas, como f é constante em cada componente de $S \subset U$, segue que $f(S)$ é finito. $\square$

3.1 Um exemplo de uma função subanalítica contínua, que não é constante em uma componente conexa dos seus pontos críticos

A evolução no estudo da Matemática, por diversas vezes é dada omitindo uma hipótese e se questionando sobre o que acontece. Seguindo esse pensamento, é natural nos perguntarmos se retirarmos a hipótese de f ser localmente Lipschitz, então ainda continuaremos a ter f constante em cada componente conexas dos "pontos críticos"? E o conjunto dos "valores críticos", ainda tem medida nula? A resposta para essa pergunta é não, o exemplo a seguir justificará.

Mas antes de construir o exemplo, vamos extender a ideia de ponto crítico de Clarke para funções contínuas.

Definição 3.1.1. *Seja $\{t_n\}$ uma sequência monótona, que converge para $0 \in \mathbb{R}$ com valores positivos para todo $n \in \mathbb{N}$. Denotaremos a convergência de tal sequência por*

$$\{t_n\} \searrow 0_+$$

Definição 3.1.2. *Chamaremos de subgradiente limitante assintótico de $f : \mathbb{R}^n \to \mathbb{R}$ em $x \in \mathbb{R}^n$ e denotaremos por $\partial^\infty f(x)$, o conjunto de todos os $y^* \in \mathbb{R}^n$ tal que existe $\{t_n\}_n \subset \mathbb{R}_+$ com $\{t_n\} \searrow 0_+$, $\{y_n\}_n \subset \mathbb{R}^n$, $y_n^* \in \widehat{\partial} f(y_n)$ tal que $y_n \to x$ e $t_n y_n^* \to y^*$.*

Proposição 3.1.1. *Se $f : \mathbb{R}^n \to \mathbb{R}$ é localmente Lipschitz, então $\partial^\infty f(x) = 0$.*

Demonstração. Veja que como $y_n \to x$, ou seja, é uma sequência convergente, logo segue que $\{y_n\} \cup \{x\}$ é compacto. Pelo Teorema 2.3.15 segue que $\widehat{\partial} f(y_n)$ é limitada, logo como $\{t_n\} \searrow 0_+$, temos que $t_n y_n^* \to 0$, o que conclui a demonstração. $\square$

Definição 3.1.3. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ contínua. Definimos o subdiferencial de Clarke de f em $x \in \mathbb{R}^n$, denotado por $\partial^o f(x)$, como sendo,*

$$\partial^o f(x) = \overline{co}\{\partial f(x) + \partial^\infty f(x)\}.$$

Observação. *Veja que quando f é localmente Lipschitz, pela Proposição 3.1.1 temos $\partial^\infty f(x) = 0$, o que implica que a definição acima coincide com a definição de subdiferencial de Clarke inicial.*

Agora já conhecendo a ideia de subdiferencial de Clarke para funções contínuas, vamos estender a ideia de ponto crítico.

Definição 3.1.4. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ contínua. Dizemos que um ponto $x_0 \in \mathbb{R}^n$ é um ponto largamente crítico, se $0 \in T_f(x_0)$, onde*

$$T_f(x_0) = \bigcap_{\varepsilon>0} \overline{co}\Big\{ \bigcup_{x \in B(x_0,\varepsilon)} \widehat{\partial} f(x) \Big\}.$$

Finalmente vamos a construção da função desejada. Daremos início construindo funções auxiliares, que ajudará a definir a função do exemplo desejado.

Observação. *Considere $\theta_0 : [0,\pi) \to [0,\pi/2]$, dada por*

$$\theta_0(z) := \begin{cases} z, & se\ 0 \leq z \leq \pi/2 \\ \pi - z, & se\ \pi/2 < z < \pi. \end{cases}$$

Agora vamos extender o domínio de θ_0 para todos os reais:

$$z \mapsto \tilde{\theta}_0(z) := \theta_0(z(mod\pi))$$

e defina $\sigma : [0,\pi/2] \times \mathbb{R} \to \{-1,1\}$, dada por

$$\sigma(\theta,z) := \begin{cases} 1, & se\ \theta \geq \tilde{\theta}_0(z), \\ -1, & se\ \theta < \tilde{\theta}_0(z). \end{cases}$$

Agora defina $\Phi_1 : \mathbb{R}_+^* \times [0,\pi/2] \times \mathbb{R} \to \mathbb{R}$, como sendo

$$\Phi_1(\rho,\theta,z) := \begin{cases} \frac{2}{\pi}\tilde{\theta}_0(z) + \sigma(\theta,z)\rho, & se\ \rho \leq \frac{2}{\pi}|\theta - \tilde{\theta}_0(z)|, \\ \frac{2}{\pi}\theta, & se\ \rho > \frac{2}{\pi}|\theta - \tilde{\theta}_0(z)|. \end{cases}$$

Agora para $(\rho,\theta,z) \in \mathbb{R}_+^* \times [0,\pi) \times \mathbb{R}$, defina

$$\Phi_2(\rho,\theta,z) := \begin{cases} \Phi_1(\rho,\theta,z), & se\ 0 \leq \theta \leq \pi/2, \\ \Phi_1(\rho,\pi-\theta,z), & se\ \pi/2 < \theta \leq \pi. \end{cases}$$

Por último, defina $\Phi : \mathbb{R}_+^* \times [0,2\pi) \times \mathbb{R} \to [0,1]$ *dada por,*

$$\Phi(\rho,\theta,z) := \begin{cases} \Phi_2(\rho,\theta,z), & se\ 0 \leq \theta \leq \pi, \\ \Phi_2(\rho,\theta-\pi,z), & se\ \pi < \theta < 2\pi. \end{cases}$$

E finalmente,

Exemplo 31. *Defina* $f : \mathbb{R}^3 \to [0,1]$, *como sendo a função cujo o gráfico em coordenadas cartesianas é o mesmo de* Φ *em coordenadas cilindrícas. Ou seja,*

$$f(x,y,z) := \begin{cases} \Phi(\sqrt{x^2+y^2}, arctg\left(\frac{y}{x}\right), z), & se\ x > 0, y \geq 0 \\ \lim_{a\to y^+} \Phi(a, \frac{\pi}{2}, z), & se\ x = 0, y \geq 0 \\ \Phi(\sqrt{x^2+y^2}, arctg\left(\frac{y}{x}\right) + 2\pi, z), & se\ x > 0, y < 0 \\ \Phi(\sqrt{x^2+y^2}, arctg\left(\frac{y}{x}\right) + \pi, z), & x < 0 \\ \lim_{a\to y^-} \Phi(-a, \frac{3\pi}{2}, z), & se\ x = 0, y \leq 0 \end{cases}$$

Afirmação 1. A função f é contínua e subanalítica.

Inicialmente, mostraremos que f é contínua. Para isso não é difícil ver que Φ_1 é contínua. Com a continuidade de Φ_1, pela definição de f, resta mostrar a continuidade nas fronteiras dos octantes. Então começaremos verificando a continuidade de f no eixo z. Veja que calcular o limite de f no eixo z, é fazer $(x,y) \to (0,0) \Rightarrow \rho \to 0$, em ambas as situações, teremos $\Phi_1(\rho,\theta,z) = \frac{2}{\pi}\tilde{\theta}_0(z) + \sigma(\theta,z)\rho \to \frac{2}{\pi}\tilde{\theta}_0(z)$, pois caso $\rho > \frac{2}{\pi}|\theta - \tilde{\theta}_0(z)| \Rightarrow \theta \to \tilde{\theta}_0(z)$

Como $f(0,0,z) = \lim_{a\to 0^+} \Phi(a, \frac{\pi}{2}, z) = \frac{2}{\pi}\tilde{\theta}_0(z)$, resta calcular os limites aproximando-se pelos quadrantes e pelos eixos x e y.

$$\begin{cases} \lim_{(x,y)\to(0,0)} \Phi(\sqrt{x^2+y^2}, arctg\left(\frac{y}{x}\right), z) = \frac{2}{\pi}\tilde{\theta}_0(z), & \text{se } x > 0, y \geq 0 \\ \lim_{(x,y)\to(0,0)} (\lim_{a\to y^+} \Phi(a, \frac{\pi}{2}, z)) = \frac{2}{\pi}\tilde{\theta}_0(z), & \text{se } x = 0, y > 0 \\ \lim_{(x,y)\to(0,0)} \Phi(\sqrt{x^2+y^2}, arctg\left(\frac{y}{x}\right) + 2\pi, z) = \frac{2}{\pi}\tilde{\theta}_0(z), & \text{se } x > 0, y < 0 \\ \lim_{(x,y)\to(0,0)} \Phi(\sqrt{x^2+y^2}, arctg\left(\frac{y}{x}\right) + \pi, z) = \frac{2}{\pi}\tilde{\theta}_0(z), & x < 0 \\ \lim_{(x,y)\to(0,0)} (\lim_{a\to y^-} \Phi(-a, \frac{3\pi}{2}, z)) = \frac{2}{\pi}\tilde{\theta}_0(z), & \text{se } x = 0, y \leq 0 \end{cases}$$

Veja que f é contínua no primeiro e quinto octante. Além disso, também é contínua no semieixo positivo x. Vamos mostrar que é contínua no semieixo positivo y. Observe que é suficiente fazer para o caso $y_0 > 0$, logo

$$\lim_{(x,y)\to(0,y_0)} \Phi(\sqrt{x^2+y^2}, arctg(\frac{y}{x}), z) = \Phi(y_0, \frac{\pi}{2}, z) = \lim_{a\to y_0^+} \Phi(a, \frac{\pi}{2}, z) = f(0,y_0,z)$$

Agora verificaremos a continuidade nos octantes 2,3,6 e 7. Observe que é suficiente estudar os limites de f, em $\{(0,y,z); z \in \mathbb{R} \ e \ y \in \mathbb{R}^*\}$. Veja que

$$\lim_{x\to 0^-} arctg(\frac{y}{x}) = \begin{cases} -\frac{\pi}{2}, & \text{se } y > 0 \\ \frac{\pi}{2}, & \text{se } y < 0 \end{cases}$$

e

$$\lim_{x\to 0} \Phi(\rho, \theta + \pi, z) = \lim_{x\to 0} f(x,y,z)$$

Assim segue que

$$\lim_{x\to 0} \Phi(\rho, \theta + \pi, z) = \begin{cases} \lim_{x\to 0} \Phi(\rho, \frac{\pi}{2}, z), & \text{se } y > 0 \\ \lim_{x\to 0} \Phi(\rho, \frac{3\pi}{2}, z), & \text{se } y < 0 \end{cases} = \begin{cases} \Phi(y, \frac{\pi}{2}, z), & \text{se } y > 0 \\ \Phi(-y, \frac{3\pi}{2}, z), & \text{se } y < 0 \end{cases}$$

Como

$$f(0,y,z) = \begin{cases} \lim_{a\to y^+} \Phi(a, \frac{\pi}{2}, z), & \text{se } y > 0 \\ \lim_{a\to y^-} \Phi(-a, \frac{3\pi}{2}, z), & \text{se } y < 0 \end{cases} = \begin{cases} \Phi(y, \frac{\pi}{2}, z), & \text{se } y > 0 \\ \Phi(-y, \frac{3\pi}{2}, z), & \text{se } y < 0 \end{cases}$$

Finalizaremos, verificando a continuidade no quarto e oitavo octante. Observe que precisamos mostrar a continuidade na fronteira desses octantes com o primeiro e o quinto, posteriormente na fronteira com o terceiro e sétimo octante.

Suponha $x > 0$, temos que

$$\lim_{y\to 0^-} f(x,y,z) = \lim_{y\to 0^-} \Phi(\rho, \theta + 2\pi, z) = \lim_{y\to 0^-} \Phi(\rho, \theta, z) =$$

$$= \Phi(x, \theta, z) = f(x,0,z)$$

e

$$\lim_{x\to 0^+} f(x,y,z) = \lim_{x\to 0^+} \Phi(\rho, \theta + 2\pi, z) = \Phi(-y, -\frac{\pi}{2} + 2\pi, z) =$$

$$= \Phi(-y, \frac{3\pi}{2}, z)$$

como

$$f(0,y,z) = \lim_{a\to y^-} \Phi(-a, \frac{3\pi}{2}, z) = \Phi(-y, \frac{3\pi}{2}, z),$$

segue pelo Lema da colagem que f é contínua.

Agora vamos mostrar que, f é subanalítica. Como as funções $(x,y,z) \mapsto \sqrt{x^2+y^2}$ e $(x,y,z) \mapsto z$ são semialgébricas, temos que elas são subanalíticas. Além disso, sendo $h\colon (\mathbb{R}\setminus\{0\}) \times \mathbb{R}^2 \to \mathbb{R}$ dada por $h(x,y,z) = arctg(\frac{y}{x})$, temos que seu gráfico

$$\begin{aligned} Graf(h) &= \{(x,y,z,w) \in (\mathbb{R}\setminus\{0\}) \times \mathbb{R}^3; w = arctg(\frac{y}{x})\} \\ &= \{(x,y,z,w) \in \mathbb{R}^4; x \cdot tg(w) - y = 0, -\pi/2 < w < \pi/2 \text{ e } x \neq 0\} \end{aligned}$$

é semianalítico. Logo, a aplicação $(x,y,z) \mapsto (\sqrt{x^2+y^2}, arctg(\frac{y}{x}), z)$ é subanalítica em $x \neq 0$.

Pela Proposição 2.4.9, resta provar que Φ é subanalítica. Perceba que o gráfico de $\tilde{\theta}_0(z)$, quando intersectados com conjuntos semialgébricos limitados é um conjunto semialgébrico. Já σ e Φ_1 é dado por desigualdades e igualdades de funções semialgébricas, logo são semialgébricas. Veja que para Φ ser subanalítica, é suficiente que Φ_2 seja subanalítica, que depende apenas de Φ_1 para ser subanalítica. Quando intersectamos com semialgébricos limitados o gráfico de Φ_2 e Φ, temos que eles são subanalíticos, veja 2.4.4.

Afirmação 2. A restrição de f ao conjunto $Z = \{(0,0,z); z \in \mathbb{R}\}$ não é constante.

Vamos calcular $f(0,0,0)$ e $f(0,0,2)$, verificando que são valores diferentes.

Veja que $f(0,0,0) = \Phi(0,0,0) = \Phi_2(0,0,0) = \Phi_1(0,0,0) = \frac{2}{\pi}\tilde{\theta}_0(0) + \sigma(0,0) \cdot 0$. Essa última igualdade acontece pois, não podemos ter $0 > \frac{2}{\pi}|\theta - \tilde{\theta}_0(z)| \geq 0$. Logo $\Phi_1(0,0,0) = \frac{2}{\pi}\tilde{\theta}_0(0)$ e como $\tilde{\theta}_0(0) = \theta_0(0) = 0$, logo $f(0,0,0) = 0$.

Agora veja que $f(0,0,2) = \Phi(0,0,2) = \Phi_2(0,0,2) = \Phi_1(0,0,2) = \frac{2}{\pi}\tilde{\theta}_0(2) + \sigma(0,2) \cdot 0$. Essa última igualdade acontece pois, não podemos ter $0 > \frac{2}{\pi}|\theta - \tilde{\theta}_0(z)| \geq 0$. Logo $\Phi_1(0,0,2) = \frac{2}{\pi}\tilde{\theta}_0(2)$ e como $\tilde{\theta}_0(2) = \theta_0(2) = \pi - 2 \neq 0$, logo $f(0,0,0) \neq f(0,0,2)$.

Afirmação 3. Todo ponto de Z é largamente crítico, ou seja, $Z \subset \{u \in \mathbb{R}^3; 0 \in T_f(u)\}$

Queremos mostrar que, todo ponto da forma $u_0 = (0,0,z_0)$ é um ponto largamente crítico. Tome z_0 de modo que $0 < z_0 < \pi/2$ e seja $\theta_0 = \tilde{\theta}_0(z_0)$, veja que pela definição de $\tilde{\theta}_0(z_0)$, temos $0 < \theta_0 < \pi/2$.

Defina uma sequência $\theta_n = \theta_0 + \frac{\pi}{2^{n+2}}$ e uma outra $x_n = \frac{1}{2^n\sqrt{1+a_n^2}}$, onde $a_n = tg(\theta_n)$. Por último, defina $y_n = a_n x_n$ e portanto segue que $\rho_n = \sqrt{x_n^2 + y_n^2} = \sqrt{x_n^2 + a_n x_n^2} = x_n\sqrt{1+a_n^2} = \frac{1}{2^n\sqrt{1+a_n^2}}\sqrt{1+a_n^2} = \frac{1}{2^n}$.

Agora defina

$$u_n := (x_n, y_n, z_0) \text{ e } \bar{u}_n := (-x_n, -y_n, z_0).$$

Veja que as sequências (u_n) e $(\bar{u}_n)$ convergem para u_0. Além disso, temos que $f(u_n) = f(x_n, y_n, z_0)$, como temos $x_n, y_n > 0$, segue que $f(u_n) = \Phi(\sqrt{x_n^2 + y_n^2}, arctg(\frac{y_n}{x_n}), z_0)$. Veja que, como $y_n = a_n x_n$, temos que $arctg(\frac{y_n}{x_n}) = arctg(a_n) = arctg(tg(\theta_n)) = \theta_n$, como $\theta_n \to \theta_0$, podemos conseguir um $n_0 \in \mathbb{N}$ de modo que, $\theta_n < \frac{\pi}{2}$ para todo $n \geq n_0$. Portanto, para $n \geq n_0$, temos

$$\begin{aligned} f(u_n) &= \Phi(\sqrt{x_n^2 + y_n^2}, arctg(\tfrac{y_n}{x_n}), z_0) \\ &= \Phi_2(\sqrt{x_n^2 + y_n^2}, arctg(\tfrac{y_n}{x_n}), z_0) \\ &= \Phi_1(\sqrt{x_n^2 + y_n^2}, arctg(\tfrac{y_n}{x_n}), z_0) \\ &= \tfrac{2}{\pi}\theta_n, \end{aligned}$$

a última igualdade acontece, pois $\rho_n = \sqrt{x_n^2 + y_n^2} = \frac{1}{2^n} > \frac{1}{2^{n+1}} = \frac{2}{\pi}\frac{\pi}{2^{n+2}} = \frac{2}{\pi}|\theta_n - \theta_0|$.

Calculemos as derivadas parciais de f, temos que $\frac{\partial \Phi}{\partial \rho}(u_n) = 0 = \frac{\partial \Phi}{\partial z}(u_n)$ e $\frac{\partial \Phi}{\partial \theta}(u_n) = \frac{2}{\pi}$.

Calculando a matriz jacobiana da aplicação de mudança de variavéis $(x, y, z) \longmapsto (\sqrt{x^2 + y^2}, arctg(\frac{y}{x}), z)$, obtemos na segunda linha da matriz $\frac{\partial arctg(\frac{y}{x})}{\partial z} = 0, \frac{\partial arctg(\frac{y_n}{x_n})}{\partial x} = \frac{-y_n}{y_n^2 + x_n^2}$ e $\frac{\partial arctg(\frac{y_n}{x_n})}{\partial y} = \frac{x_n}{y_n^2 + x_n^2}$, veja que as demais linhas não importa pois $\frac{\partial \Phi}{\partial \rho}(u_n) = 0 = \frac{\partial \Phi}{\partial z}(u_n)$, logo obtemos

$$\nabla f(u_n) = \frac{2}{\pi}\left(\frac{-y_n}{y_n^2 + x_n^2}, \frac{x_n}{y_n^2 + x_n^2}, 0\right)$$

e também conseguimos obter,

$$\nabla f(\bar{u}_n) = \frac{2}{\pi}\left(\frac{y_n}{y_n^2 + x_n^2}, \frac{-x_n}{y_n^2 + x_n^2}, 0\right)$$

observe que f é diferenciável em u_n e $\bar{u}_n$, pois suas derivadas parciais existem e são contínuas no primeiro octante do $\mathbb{R}^3$, essa continuidade só depende de $\Phi_1 = \frac{2}{\pi}\theta$ no aberto, $\{(\rho, \theta, z) \in \mathbb{R}_+^* \times [0, \frac{\pi}{2}] \times \mathbb{R}; \rho > \frac{2}{\pi}|\theta - \tilde{\theta}_0(z)|\}$.

Segue agora que, se $u_n \in B(u_0, \varepsilon)$, então $\nabla f(u_n), \nabla f(\bar{u}_n) = -\nabla f(u_n) \in \{\nabla f(u); u \in B(u_0, \varepsilon) \cap D_f\}$, ao tomar o fecho convexo, temos que o segmento de reta que liga $\nabla f(u_n), \nabla f(\bar{u}_n) = -\nabla f(u_n)$, passa pela origem e pertence ao fecho convexo. Consequentemente, o fecho contém a origem e

$$0 \in \bigcap_{\varepsilon > 0} \overline{co}\{\nabla f(u); u \in B(u_0, \varepsilon) \cap D_f\},$$

onde D_f denota o conjunto de todos os pontos onde f é diferenciável. Agora resta observar que

$$\bigcap_{\varepsilon>0} \overline{co}\{\nabla f(u); u \in B(u_0,\varepsilon)\cap D_f\} \subset \bigcap_{\varepsilon>0} \overline{co}\Big\{\bigcup_{u\in B(u_0,\varepsilon)} \widehat{\partial} f(u)\Big\},$$

concluímos assim u_0 é um ponto largamente crítico de f e como u_0 foi tomado arbitrário em Z, a afirmação está provada.

3.2 Um exemplo de função Lipschitz com domínio limitado, com o conjunto dos valores críticos não finito

No Teorema 3.0.7, vimos que se U é limitado, então o conjunto dos valores críticos é finito. Nessa seção, veremos um exemplo que se f for apenas localmente Lipschitz e não tiver a hipótese de ser subanalítica, então não é verdade que o conjunto dos valores crítico é finito, mesmo que U seja limitado.

Definiremos inicialmente funções auxiliares. Seja k um inteiro não negativo. Defina $f_k\colon (\frac{1}{2^k}, \frac{1}{2^{k-1}}) \to \mathbb{R}$, dada por $f_k(x) = -|x-P_k| + (P_k - \frac{1}{2^k})$, onde P_k é o ponto médio do segmento $[\frac{1}{2^k}, \frac{1}{2^{k+1}}]$.

Exemplo 32. *Defina $f\colon (-2,2) \to \mathbb{R}$ de modo que $f|_{(0,2)}$ é dada pela justaposição das funções f_k nos seus respectivos domínios, $f(0) = 0$ e em $(-2,0)$ satisfaz $f(x) = f(-x)$.*

Veja que a função é claramente Lipschitz, mas não é subanalítica em 0. Pois se seu gráfico fosse subanalítico, então a intersecção com $(-r,r)\times\{0\}$, para $r > 0$ qualquer, seria um conjunto subanalítico, pela Proposição 2.4.1. Mas pela Proposição 2.4.8, todo subanalítico e limitado da reta, deve ser composto apenas por uma quantidade finita de intervalo (degenerado ou não degenerado), logo não é subanalítica.

Veja que a sequência (a_k), dada por, $a_k = P_k - \frac{1}{2^k}$ é uma sequência formada pelos valores críticos de f, mostrando assim que o Teorema 3.0.7, não é mais veridico se f não for subanalítica.

4 CONCLUSÃO

Foram estudados inicialmente alguns tópicos, como Derivada de Clarke e Geometria Subanalítica. Posteriormente mostramos que se o vetor e_1, está contido dentro do conjunto dos pontos críticos de Clarke de uma função f subanalítica e localmente Lipschitz, então f é constante ao longo desse segmento. Depois estendemos o resultado para cada componente conexa do conjunto dos pontos críticos de Clarke. Como uma implicação desse resultado, mostramos que o conjunto dos valores críticos de Clarke é enumerável e consequentemente, possui medida de Lebesgue nula. Por fim, foi dado um exemplo que mostra a não veracidade desse resultado, caso f seja apenas subanalítica e contínua.

REFERÊNCIAS

BIERSTONE, E.; MILMAN, P. D. Semianalytic and subanalytic sets. **Publications Mathématiques de l'IHÉS**, v. 67, p. 5–42, 1988.

BOLTE, J.; DANIILIDIS, A.; LEWIS, A. A nonsmooth morse–sard theorem for subanalytic functions. **Journal of mathematical analysis and applications**, Elsevier, v. 321, n. 2, p. 729–740, 2006.

BOLTE, J.; DANIILIDIS, A.; LEWIS, A. The łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. **SIAM Journal on Optimization**, SIAM, v. 17, n. 4, p. 1205–1223, 2007.

BOLTE, J.; DANIILIDIS, A.; LEWIS, A. S.; SHIOTA, M. Clarke critical values of subanalytic lipschitz continuous functions. Centre de Recerca Matemàtica, 2005.

CAVALCANTI, M. M.; CAVALCANTI, V. N. D. Introdução a teoria das distribuições e aos espaços de sobolev. **Maringá: UEM**, 2009.

CLARKE, F. On the inverse function theorem. **Pacific Journal of Mathematics**, Mathematical Sciences Publishers, v. 64, n. 1, p. 97–102, 1976.

CLARKE, F. H. Generalized gradients and applications. **Transactions of the American Mathematical Society**, v. 205, p. 247–262, 1975.

ITOH, J.-i.; TANAKA, M. A sard theorem for the distance function. **Mathematische Annalen**, Springer, v. 320, n. 1, p. 1–10, 2001.

JUDICE, E. D. O Teorema de Sard e suas Aplicações. **Rio de Janeiro: Instituto de Matemática Pura e Aplicada**, v. 18, 2011.

LIMA, E. L. Análise real-vol. 2. **Rio de Janeiro: IMPA**, 2010.

MACEDO, A. C. Método de ponto proximal para otimização. **Dissetação de Mestrado em Matemática, UFG, Goiânia**, 2009.

MORSE, A. P. The behavior of a function on its critical set. **Annals of Mathematics**, JSTOR, p. 62–70, 1939.

OLIVEIRA, C. R. D. Introdução à análise funcional. **Impa**, 2001.

RIFFORD, L. A morse-sard theorem for the distance function on riemannian manifolds. **manuscripta mathematica**, Springer, v. 113, n. 2, p. 251–265, 2004.

ROCKAFELLAR, R. T. **Convex analysis**. [S.l.]: Princeton university press, 1970. v. 36.

ROCKAFELLAR, R. T.; WETS, R. J.-B. **Variational analysis**. [S.l.]: Springer Science & Business Media, 2009. v. 317.

SARD, A. The measure of the critical values of differentiable maps. **Bulletin of the american Mathematical Society**, v. 48, n. 12, p. 883–890, 1942.

SHIOTA, M. **Geometry of subanalytic and semialgebraic sets**. [S.l.]: Springer Science & Business Media, 2012. v. 150.

STEIN, E. M.; SHAKARCHI, R. **Real analysis: measure theory, integration, and Hilbert spaces**. [S.l.]: Princeton University Press, 2009.

TAMM, M. Subanalytic sets in the calculus of variation. **Acta mathematica**, Institut Mittag-Leffler, v. 146, p. 167–199, 1981.

WHITNEY, H. A function not constant on a connected set of critical points. **Duke Mathematical Journal**, Duke University Press, v. 1, n. 4, p. 514–517, 1935.