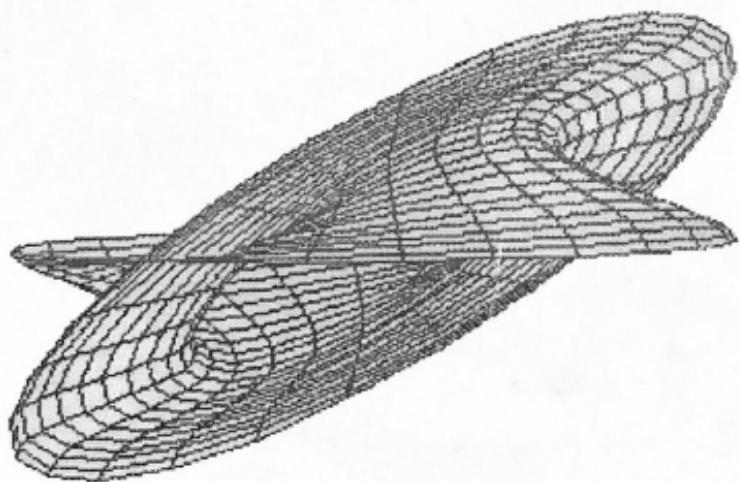


**BOLETIM DE
INICIAÇÃO CIENTÍFICA
EM MATEMÁTICA
- BICMat -**



**VOLUME I
SETEMBRO DE 1998
DEPARTAMENTO DE MATEMÁTICA
IGCE - UNESP - RIO CLARO**

BOLETIM DE INICIAÇÃO CIENTÍFICA EM MATEMÁTICA

BICMat

Recentemente, o número de alunos matriculados em atividades extracurriculares da UNESP cresceu e dobrou. No Rio Claro, o número de alunos inscritos no BICMat é de 800, que é o maior número de inscritos em todos os projetos da UNESP. O BICMat é um projeto de extensão da UNESP que visa estimular o interesse dos alunos pelas matemáticas e suas aplicações. O projeto é coordenado por professores da Faculdade de Ciências Exatas da UNESP, que trabalham com diferentes temas de matemática, como álgebra, geometria, cálculo, estatística, probabilidade e teoria dos jogos. O projeto também oferece palestras, workshops e outras atividades para os alunos.

A participação dos alunos em atividades de extensão é muito significativa, englobando trabalhos de pesquisa, elaboração de artigos científicos, participação em competições e outras atividades.

**DEPARTAMENTO DE MATEMÁTICA
IGCE-UNESP-RIO CLARO**

ICCC

**BOLETIM DE
INICIAÇÃO CIENTÍFICA
EM MATEMÁTICA
BICMat**

Objetivo

Divulgar trabalhos de iniciação científica em Matemática, incentivando o envolvimento do aluno com a pesquisa e propiciando um canal de informações para a comunidade acadêmica.

Orientação aos autores

Ao redigir o material a ser divulgado o autor deve observar que o alvo principal é o aluno devendo a redação ser clara e objetiva incentivando-o à leitura.

O trabalho deve ser enviado para a Comissão Editorial em disquete, utilizando o programa LATEX , em estilo "article " ou S.W.P., em estilo "modified articles style for short texts ".

A responsabilidade de cada artigo é exclusiva do autor e respectivo orientador.

Comissão Editorial:

Prof. Sergio Roberto Nobre

Prof. Marcelo José Saia

Profa. Nativi Viana Pereira Bertolo

Apoio:

Projeto PADCT-SPEC-01/95-01-CAPES

Departamento de Matemática

IGCE

Colaboração:

Paulo Henrique Bertolo - Aluno do curso de computação
- UNESP - Bauru

EDITORIAL

Em toda a história do Curso de Matemática do IGCE-UNESP-Rio Claro, desde a época da antiga Faculdade de Filosofia Ciências e Letras de Rio Claro, a iniciação científica tem sido uma atividade constante e muito importante na formação dos alunos. É uma atividade importante não só por envolver e iniciar o aluno na pesquisa matemática, como também por propiciar um elo importante de ligação entre alunos e professores da graduação.

Atualmente, o número de alunos envolvidos em atividades de iniciação científica, como estágios de iniciação científica, desenvolvimento de projetos através de bolsas de iniciação científica, desenvolvimento de projetos que integram o PET (Projeto Especial de Treinamento) etc, é bem grande, onde a maior parte tem recebido suporte financeiro da FAPESP, CNPq, CAPES. Nos últimos cinco anos este número aumentou de forma bastante expressiva, principalmente quando se leva em consideração o nível do aluno ingressante e a evasão nos cursos de Matemática.

A participação dos alunos em congressos de iniciação científica tem sido significativa, originando trabalhos interessantes para uma maior divulgação.

Fundamenta-se assim, a criação da série "Boletim de Iniciação Científica Em Matemática "- BICMat, um veículo pelo qual o aluno pode divulgar seu trabalho e ao mesmo tempo desenvolver a linguagem escrita e a redação científica.

Neste Boletim também poderão ser divulgados trabalhos e palestras interessantes, concernentes com a graduação.

Este número da revista do projeto visa dar uma formação ao aluno em sua carreira, fazendo-o ter mais tempo em que se dê a uma utilização das competências adquiridas.

"Boletim CIEP-PUCRS"

ÍNDICE

Equações de Lagrange para Queda-Livre Sob o ponto de Vista da Topologia Diferencial Mickel Abreu de Ponte.....	1
Sobre Algumas Desigualdades Importantes da Análise Monica Regina Gaiotto.....	49
Teorema do Ponto Fixo de Brower no Plano Andréia Cristina Aravéquia.....	74
Uma Máquina de Catástrofe Gravitacional Luciana de Fátima Martins.....	78
COLABORADORES	
Paulo Henrique Ribeiro	
UNESP - Bauru	

Colaboradores

Paulo Henrique Ribeiro
UNESP - Bauru

Equações de Lagrange para Queda Livre Sob o Ponto de Vista da Topologia Diferencial

Mickel Abreu de Ponte*

Orientadora: Profa. Dra. Alice Kimie Miwa Libardi

1 Introdução

“Os métodos de Poincaré se caracterizam, antes de mais nada, pelo seu ponto de vista geométrico. Ele visualizou um sistema dinâmico como um campo de vetores em um espaço de fase, onde a solução é uma curva suave. A teoria qualitativa é baseada nas propriedades geométricas do retrato de fase: a família de curvas que preenchem todo o espaço de fase. Para questões como estabilidade é necessário estudar o retrato de fase inteiro, incluindo o comportamento das soluções em todos os valores do parâmetro t . Assim é melhor considerar o espaço de fase como um objeto geométrico. Fazendo isto, Poincaré descobriu que o modelo matemático usado, tendo como espaço subjacente o espaço Euclidiano, era inadequado. A visão global na teoria qualitativa conduziram Poincaré a noção de uma variedade diferencial como o espaço de fase em Mecânica”.([4])

Esta monografia nasceu da preocupação em se dar uma formalização rigorosa aos conceitos físicos ao mesmo tempo em que se dava uma aplicação aos conceitos matemáticos.

*Bolsista CNPq-PIBIC

Iremos apresentar uma formalização das equações de Lagrange para queda livre utilizando conceitos matemáticos particularmente da Topologia Diferencial tais como: variedades, subvariedades, espaços e fibrados tangentes. Colocaremos a resolução de vários exercícios propostos no livro texto [1] que nos ajudaram a compreender a teoria estudada.

2 Preliminares

2.1 Noções Básicas

Apresentamos abaixo algumas noções básicas utilizadas para o desenvolvimento da teoria nos capítulos seguintes.

Definição 2.1.1: Sejam V e W dois espaços vetoriais reais. Uma aplicação linear é uma função de V em W , $F : V \rightarrow W$, que satisfaz as seguintes condições:

i) quaisquer que sejam $u, v \in V$

$$F(u + v) = F(u) + F(v)$$

ii) quaisquer que sejam $k \in \mathbb{R}$ e $v \in V$

$$F(kv) = kF(v).$$

Definição 2.1.2: Seja $T : V \rightarrow W$ uma aplicação linear. O conjunto de todos os vetores $v \in V$ tais que $T(v) = 0$ é chamado núcleo de T , sendo denotado por $\text{Ker}(T)$. Isto é

$$\text{Ker}(T) = \{v \in V; T(v) = 0\}.$$

$\text{Ker}(T) = \{0\}$ se, e somente se, T é injetora.

Definição 2.1.3: Seja $T : V \rightarrow W$ uma aplicação linear. A imagem de T é o conjunto dos vetores $w \in W$ tal que existe um vetor $v \in V$, que satisfaz $T(v) = w$. Ou seja

$$Im(T) = \{w \in W; T(v) = w \text{ para algum } v \in V\}.$$

Se $Im(T) = W$, diremos que T é sobrejetora.

OBS: Também denotamos $Im(T)$ por $T(V)$.

Teorema 2.1.4: Seja $T : V \rightarrow W$ uma aplicação linear entre espaços vetoriais de dimensão finita.

$$\text{Então } \dim Ker(T) + \dim T(V) = \dim V.$$

Definição 2.1.5: (Relação de equivalência) Seja A um conjunto, $A \neq \emptyset$. Dada uma relação R em A , ela é de *equivalência* se:

- i) $\forall a \in A, aRa$ (propriedade reflexiva)
- ii) $\forall a, b \in A, aRb \Rightarrow bRa$ (propriedade simétrica)
- iii) $\forall a, b, c \in A, aRb \text{ e } bRc \Rightarrow aRc$ (propriedade transitiva).

Exemplo 2.1.6: Seja $A = \mathbb{Z}$. Definimos a relação $aRb \Leftrightarrow a - b = 2k$, para $k \in \mathbb{Z}$ e $\forall a, b \in \mathbb{Z}$.

Temos:

- i) $\forall a \in \mathbb{Z}; a - a = 0$ que é par, logo aRa
- ii) $\forall a, b \in \mathbb{Z} | aRb \Rightarrow a - b = 2k, k \in \mathbb{Z}$ e então
 $b - a = -2k = 2(-k), -k \in \mathbb{Z}$, que é par
 logo $aRb \Rightarrow bRa$
- iii) $\forall a, b, c \in \mathbb{Z} | aRb \Rightarrow a - b = 2k_1, k_1 \in \mathbb{Z}$ e
 $bRc \Rightarrow b - c = 2k_2, k_2 \in \mathbb{Z}$ então
 $a - c = 2(k_1 + k_2), (k_1 + k_2) \in \mathbb{Z}$, que é par
 logo aRb e $bRc \Rightarrow aRc$.

Portanto, esta relação é de equivalência.

Definição 2.1.7: (Classe de equivalência). Se R é uma relação de equivalência em A , chamamos *classe de equivalência* do elemento $a \in A$ ao conjunto

$$\bar{a} = \{b \in A, bRa\}$$

e $\frac{A}{R} = \{\bar{a}, a \in A\}$ é dito *conjunto quociente*.

Exemplo 2.1.8: Pela relação dada no *Exemplo 2.1.6* temos

$$\begin{aligned}\bar{0} &= \{b \in Z, bR0 \Leftrightarrow b - 0 = 2k, k \in Z\}, \text{ então} \\ \bar{0} &= \{0, \pm 2, \pm 4, \dots\}.\end{aligned}$$

Da mesma maneira temos para

$$\begin{aligned}\bar{1} &= \{\pm 1, \pm 3, \pm 5, \dots\} \text{ e então o conjunto quociente é} \\ Z_2 &= \frac{Z}{R} = \{\bar{0}, \bar{1}\}.\end{aligned}$$

Definição 2.1.9: (Espaço Topológico). Sejam X um conjunto e $\mathcal{U} \subset P(X)$ tal que

$$\mathcal{U} = \{A_i, i \in I \mid A_i \in P(X)\}.$$

Se

- i) $\emptyset, X \in \mathcal{U}$
- ii) $A_1, \dots, A_n \in \mathcal{U} \Rightarrow A_1 \cap \dots \cap A_n \in \mathcal{U}$
- iii) $A_i \in \mathcal{U} \Rightarrow \cup A_i \in \mathcal{U}$

então \mathcal{U} é uma topologia para X e (X, \mathcal{U}) é um espaço topológico.

Definição 2.1.10: Um espaço topológico (X, U) é de *Hausdorff* se dados x, y em X , $x \neq y$, $\exists Ux$ e Uy elementos de U contendo x e y respectivamente e tal que $Ux \cap Uy = \emptyset$.

2.2 Derivadas de Frechét

Começaremos descrevendo como a idéia de uma derivada pode ser formulada para funções entre espaços vetoriais normados.

A idéia básica é a seguinte: uma função é *diferenciável* em um ponto se existir uma *aplicação afim* que dá uma boa aproximação da função perto do ponto. A *derivada* da função neste ponto é a *parte linear* desta aplicação afim.

Exemplo 2.2.1: Seja $f : R^2 \rightarrow R^2$ dada por

$$f(x, y) = (x^2 + y^2 + 1, xy + 1).$$

Pela Série de Taylor obtemos

$$f(1 + h, 2 + k) = f(1, 2) + (2h + 4k, 2h + k) + (h^2 + k^2, hk).$$

Quando (h, k) aproxima-se de $(0, 0)$ o termo quadrático torna-se desprezível comparado com h e k e assim f é aproximada pela *aplicação afim* $A : R^2 \rightarrow R^2$ dada por

$$A(1 + h, 2 + k) = f(1, 2) + (2h + 4k, 2h + k).$$

A *parte linear* de A é a *aplicação linear* $L : R^2 \rightarrow R^2$ dada por

$$L(h, k) = (2h + 4k, 2h + k).$$

O significado preciso de “uma boa aproximação de f ” é esclarecido na definição a seguir pelo uso de limites.

Definição 2.2.2: Seja $f : U \rightarrow F$ onde E e F são espaços vetoriais normados e U um subconjunto aberto de E . Dizemos que f é *diferenciável* em $a \in U$ se existir uma *aplicação afim contínua* $A : E \rightarrow F$ tal que

$$f(a) = A(a) \quad \lim_{x \rightarrow a} \frac{\|f(x) - A(x)\|}{\|x - a\|} = 0.$$

Definição 2.2.3: Para uma aplicação f como dada na *Definição 2.2.2* definimos sua *derivada de Fréchet* em a como sendo a única *aplicação linear contínua* $Df(a) : E \rightarrow F$ tal que

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - Df(a)(h)\|}{\|h\|} = 0.$$

Agora vamos retornar ao *Exemplo 2.2.1* e verificar em que caso

$$Df(1, 2)(h, k) = (2h + 4k, 2h + k).$$

Note que a norma de um vetor $(x, y) \in \mathbb{R}^2$ é dada por $\| (x, y) \| = \sqrt{x^2 + y^2}$ e então neste caso temos

$$\begin{aligned} & \lim_{(h,k) \rightarrow (0,0)} \frac{\| f(1+h, 2+k) - f(1, 2) - (2h+4k, 2h+k) \|}{\|(h, k)\|} \\ &= \lim_{(h,k) \rightarrow (0,0)} \frac{\| (h^2+k^2, hk) \|}{\|(h, k)\|} = 0. \end{aligned}$$

OBS: Note que a matriz Jacobiana de f no ponto $(1, 2)$ é justamente a matriz da aplicação linear $Df(1, 2)$, como veremos mais adiante.

Teorema 2.2.4: (Regra da Cadeia I). Sejam E, F e G espaços vetoriais normados e U e V subconjuntos abertos de E e F , respectivamente. Se $f : U \rightarrow F$ é diferenciável em $a \in U$ com $f(U) \subset V$ e $g : V \rightarrow G$ é diferenciável em $f(a) \in V$ então a *aplicação composta* $g \circ f : U \rightarrow G$ é diferenciável em a e

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Definição 2.2.5: A *aplicação identidade* de um espaço vetorial normado E nele mesmo é denotada por $id_E(x) = x, \forall x \in E$.

Definição 2.2.6: Sejam E e F espaços vetoriais normados e U um subconjunto aberto de E . Uma aplicação $f : U \subset E \rightarrow F$ é um difeomorfismo se f é diferenciável, bijetora e tem inversa diferenciável.

2.3 O Funtor Tangente

As idéias desta seção serão posteriormente generalizadas para conjuntos que são um pouco mais abrangentes do que espaços vetoriais normados.

Definição 2.3.1: Para um conjunto aberto U no espaço vetorial normado E definimos o *espaço tangente* $T_a E$ em $a \in U$ como o conjunto

$$T_a E = \{(a, h); h \in E\}.$$

Definição 2.3.2: A adição vetorial e multiplicação por um escalar em $T_a E$ são definidas por:

$$\begin{aligned}(a, h) + (a, k) &= (a, h + k) \\ \lambda(a, h) &= (a, \lambda h)\end{aligned}$$

para todos vetores (a, h) e $(a, k) \in T_a E$ e escalares $\lambda \in R$.

OBS: Estas operações definem uma estrutura de espaço vetorial para $T_a E$.

Definição 2.3.3: O fibrado tangente de U , sendo U um aberto, em a é definido por

$$\bigcup_{a \in U} T_a E = U \times E.$$

OBS: Veremos no estudo de fibrados tangentes que nem sempre o fibrado tangente pode ser escrito na forma de um produto cartesiano.

Duas partes importantes da idéia de aproximação afim são a *aplicação linear* $Df(a)$ e o *ponto* $f(a)$. Na formulação usual de diferenciabilidade o ponto $f(a)$ é perdido restando apenas $Df(a)$.

A idéia é que enquanto a derivada $Df(a)$ leva E em F , a *aplicação tangente* $T_a f$ leva $T_a E$ em $T_{f(a)} F$.

Note que, veja *Fig 2.3.5*, podemos considerar os vetores h e $Df(a)(h)$ localizados em a e $f(a)$, respectivamente.

Definição 2.3.4: Seja $f : U \rightarrow F$ diferenciável em $a \in U$ onde $U \subseteq E$ é aberto (E e F são espaços vetoriais normados). A *aplicação tangente* em a de f é a aplicação

$$T_a f : T_a E \rightarrow T_{f(a)} F$$

definida por:

$$T_a f(a, h) = (f(a), Df(a)(h))$$

para cada $(a, h) \in T_a E$.

Fica claro que a *aplicação tangente é linear*. Agora permitindo o ponto variar em U nós obtemos uma aplicação definida no fibrado tangente.

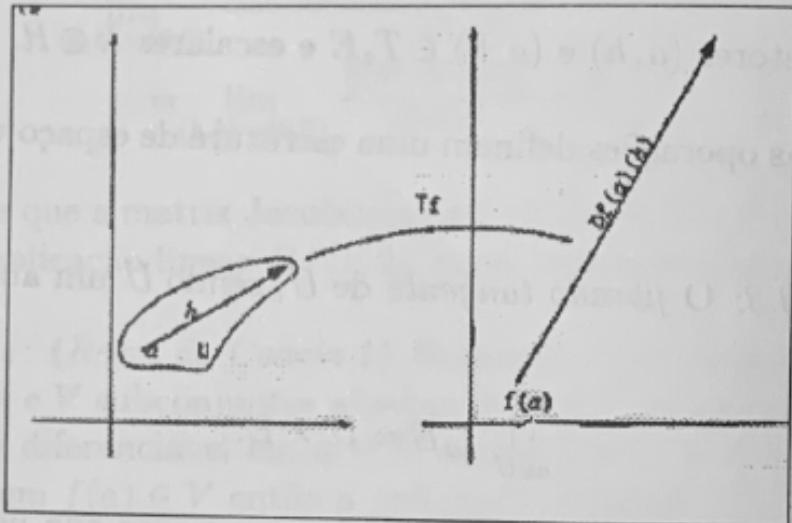


Fig 2.3.5: A aplicação tangente em a .

Definição 2.3.6: Sejam f, U, E, F e a como na definição acima. O funtor tangente

$$Tf : U \times E \rightarrow F \times F$$

é definido por:

$$Tf(a, h) = (f(a), Df(a)(h)).$$

Teorema 2.3.7: (Regra da Cadeia II). Suponha $g : U \subset E \rightarrow V \subset F$ e $f : V \rightarrow G$ onde E, F e G são espaços vetoriais normados, U e V são abertos de E e F respectivamente e f, g são diferenciáveis. Então

$$T(f \circ g) = Tf \circ Tg.$$

Prova: Sejam $a \in U$ e $h \in E$, então

$$\begin{aligned}
T(f \circ g)(a, h) &= (f \circ g(a), D(f \circ g)(a)(h)) \\
&= (f(g(a)), Df(g(a)) \circ Dg(a)(h)) \\
&= Tf(g(a), Dg(a)(h)) \\
&= Tf(Tg(a, h)).
\end{aligned}$$

Assim $T(f \circ g) = Tf \circ Tg$, como queríamos. ■

2.4 Diferenciação Parcial

Aqui a idéia da derivada de Fréchet é generalizada para incluir derivadas parciais.

Definição 2.4.1: Sejam E_1, E_2 e F espaços vetoriais normados, U_1 e U_2 subconjuntos abertos de E_1 e E_2 respectivamente e suponha que $f : U_1 \times U_2 \rightarrow F$ seja uma função diferenciável e que $(a, b) \in U_1 \times U_2$. A derivada parcial de f com respeito a sua primeira variável, $\partial_1 f(a, b)$, é a única *aplicação linear contínua* de E_1 para F tal que

$$\lim_{h \rightarrow 0} \frac{\|f(a+h, b) - f(a, b) - \partial_1 f(a, b)(h)\|}{\|h\|} = 0.$$

OBS: A derivada parcial $\partial_2 f(a, b)$ pode ser definida da mesma maneira.

Teorema 2.4.2: (Regra da cadeia para diferenciação parcial I). Sejam E_1, E_2 , E e F espaços vetoriais normados, U_1, U_2 e U subconjuntos abertos de E_1, E_2 e E respectivamente e seja

$$f : U_1 \times U_2 \rightarrow U \subseteq E$$

diferenciável em $(a, b) \in U_1 \times U_2$ e

$$g : U \rightarrow F$$

diferenciável em $f(a, b) \in U$. Então para $i = 1, 2$

$$\partial_i(g \circ f)(a, b) = Dg(f(a, b)) \circ \partial_i f(a, b).$$

É fácil de se verificar que a *Definição 2.4.1* e o *Teorema 2.4.2* podem ser generalizados para funções de n-variáveis.

Lema 2.4.3: Seja $f \subset U : R^n \rightarrow R$ diferenciável em $a \in U$, onde U é aberto do R^n . Então

$$Df(a)(\sum_{j=1}^n x_j e_j) = \sum_{j=1}^n \partial_j f(a)(x_j)$$

onde $\{e_j\}$ denota a base canônica do R^n .

Prova: Seja $h = \sum_{j=1}^n x_j e_j \in R^n$. Para provar este lema, lembremos que $Df(a)$ pode ser escrito em sua forma matricial como

$$\left(\frac{\partial f(a)}{\partial x_1} \dots \frac{\partial f(a)}{\partial x_n} \right).$$

Assim temos que

$$\begin{aligned} Df(a)(h) &= Df(a)\left(\sum_{j=1}^n x_j e_j\right) \\ &= \left(\frac{\partial f(a)}{\partial x_1} \dots \frac{\partial f(a)}{\partial x_n} \right) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \partial_1 f(a)x_1 + \dots + \partial_n f(a)x_n \\ &= Df(a)\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n \partial_j f(a)(x_j) \end{aligned}$$

como queríamos. ■

2.5 Cálculo de Componentes

Muitas vezes o estudo das derivadas recai no estudo das derivadas de suas componentes. Assim é interessante estudar a regra da cadeia em suas diferentes formas.

Estudaremos agora as derivadas de aplicações entre R^n e R^m .

Começaremos pela introdução de uma notação para derivadas parciais de funções a valores reais. Nesta seção, as funções estão definidas em um aberto do seu domínio.

Definição 2.5.1: Seja $f : R^n \rightarrow R$ diferenciável em $a \in R^n$. A “inclinação” de f com respeito a sua i -ésima variável ($i = 1, 2, \dots, n$) é denotada por f'^i e definida por

$$f'^i(a) = \partial_i f(a)(1).$$

OBS: Nos textos clássicos, escreve-se $\frac{\partial f}{\partial x_i} \Big|_{(x=a)}$ em vez de $f'^i(a)$ e $f'^i(a)$ é um número real enquanto que $\partial_i f(a)$ é uma aplicação linear.

Definição 2.5.2: Seja $f : R^n \rightarrow R^m$ diferenciável. A matriz da derivada de f em $a \in R^n$, que é a matriz de $Df(a)$, é chamada *matriz Jacobiana* de f em a e é denotada por $f'(a)$.

Teorema 2.5.3: Seja $f : R^n \rightarrow R^m$ diferenciável. Então a ij -ésima componente da matriz Jacobiana de f em a é dada por $f_i'^j(a)$.

Prova: Denotando f pelas suas componentes (funções coordenadas), isto é

$$f = (f_1, f_2, \dots, f_m)$$

e sabendo que cada função coordenada f_i

$$f_i : R^n \rightarrow R$$

então aplicando o *Lema 2.4.3* temos

$$Df_i(a)\left(\sum_{j=1}^n e_j\right) = \sum_{j=1}^n \partial_j f_i(a)(1) = \sum_{j=1}^n f_i'^j(a).$$

Portanto a ij -ésima componente da matriz Jacobiana de f em a é dada por $f_i'^j$, como queríamos. ■

Corolário 2.5.4: Seja $f : R \rightarrow R^m$ diferenciável. Então para cada $x \in R$

$$Df(x)(1) = f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h}.$$

Prova: Seja $x \in R$ e denotemos f por suas coordenadas, $f = (f_1, f_2, \dots, f_m)$. Pela *Definição 2.5.2* temos que $Df(x)(1) = f'(x)$ e usando o *Teorema 2.5.3* para cada componente f_i temos

$$Df_i(x)(1) = f'_i(x).$$

Agora lembrando que $f_i : R \rightarrow R$ então

$$f'_i(a) = \lim_{h \rightarrow 0} \frac{f_i(x+h)-f_i(x)}{h} \quad \text{para cada componente } f_i, \text{ e portanto}$$

$$Df(x)(1) = f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h},$$

como queríamos. ■

Teorema 2.5.5: (Regra da Cadeia para Derivada Parcial II). Sejam $h : R^n \rightarrow R^P$ e $g : R^m \rightarrow R^n$ diferenciáveis. Então para cada $i = 1, 2, \dots, m$

$$(h \circ g)^{/i} = \sum_{j=1}^n h^{/j} \circ g \cdot g_j^{/i}.$$

Prova: Seja $a \in R^m$, Então para cada $i = 1, 2, \dots, m$ temos

$$\begin{aligned} (h \circ g)_j^{/i}(a) &= \partial_i(h \circ g)_j(a)(1) \\ (h \circ g)^{/i}(a) &= \partial_i(h \circ g)(a)(1) \\ &= Dh(g(a)) \circ \partial_i g(a)(1). \end{aligned}$$

Agora usando novamente a *Definição 2.5.1* temos

$$Dh(g(a)) \circ \partial_i g(a)(1) = Dh(g(a))(g^{/i}(a)).$$

Observe que $g^{/i}(a) = (g_1^{/i}(a), g_2^{/i}(a), \dots, g_n^{/i}(a))$ e aplicando o *Teorema 2.5.3* temos

$$\begin{aligned}
& \begin{pmatrix} h_1(g(a))^{/1}(a) & \cdots & h_1(g(a))^{/n}(a) \\ \vdots & \ddots & \vdots \\ h_p(g(a))^{/1}(a) & \cdots & h_p(g(a))^{/n}(a) \end{pmatrix} \cdot \begin{pmatrix} g_1^{/i}(a) \\ \vdots \\ g_n^{/i}(a) \end{pmatrix} \\
& = \left(\sum_{j=1}^n h_1(g(a))^{/j}(a) \cdot g_j^{/i}, \dots, \sum_{j=1}^n h_p(g(a))^{/j}(a) \cdot g_j^{/i} \right), \text{ logo temos que} \\
& \quad (h \circ g)^{/i} = \sum_{j=1}^n h^{/j} \circ g \cdot g_j^{/i}
\end{aligned}$$

como queríamos. ■

3 Variedades Diferenciáveis

3.1 Cartas e Atlas

Embora a superfície da terra seja aproximadamente uma esfera, regiões bastante pequenas irão parecer como um plano. Um atlas geográfico é de fato uma coleção de mapas ou figuras, cada qual estando em um plano.

Cada figura determina uma função ϕ de alguma região U da superfície da terra no plano R^2 .

Esta idéia conduz a nossa primeira definição, em que substituimos a superfície da terra por um conjunto arbitrário M .

Definição 3.1.1: Seja M um conjunto. Uma *carta* de M consiste de um subconjunto U de M junto com uma função injetora $\phi : U \rightarrow R^n$ tal que $\phi(U)$ é aberto em R^n .

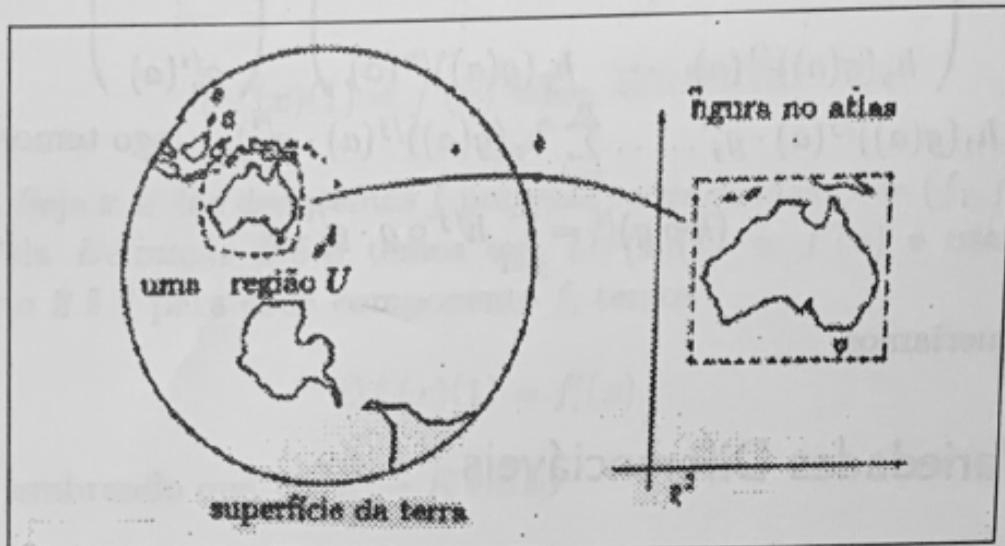


Fig 3.1.2: Uma carta (U, ϕ) .

Antes de reunirmos todas as cartas individuais para construir um atlas para todo o conjunto M , introduziremos uma certa condição de compatibilidade para garantir que as várias cartas juntam-se satisfatoriamente.

Por fim, consideremos um par de cartas para M , digamos (U, ϕ) e (V, ψ) . No caso em que estas cartas se sobrepõem, isto é, se $U \cap V \neq \emptyset$, podemos formar a aplicação composta $\phi \circ \psi^{-1}$. Esta aplica o conjunto $\psi(U \cap V) \subseteq R^n$ no conjunto $\phi(U \cap V) \subseteq R^n$ como mostra a figura abaixo. A aplicação $\phi \circ \psi^{-1}$ é chamada uma *aplicação de transição*.

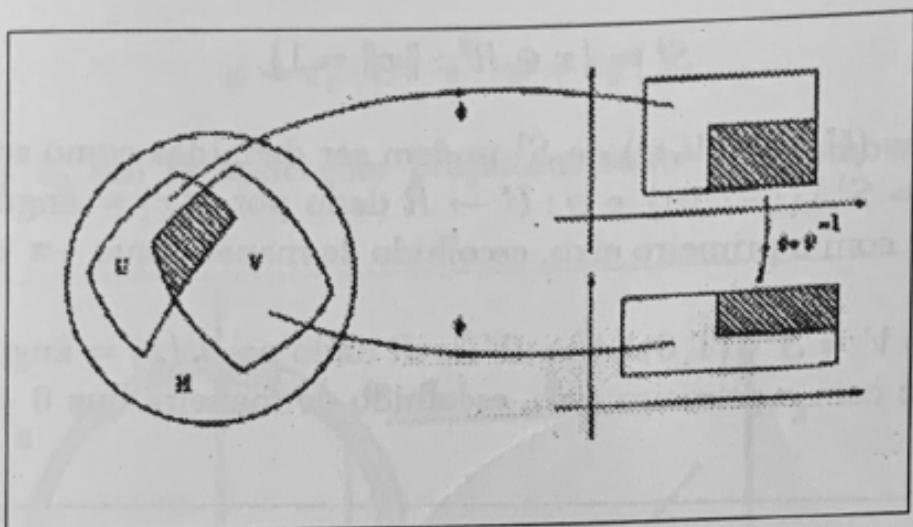


Fig 3.1.3: Aplicação de transição.

Assim, compatibilidade como definiremos abaixo nos diz que podemos transformar cartas suavemente.

Definição 3.1.4: Dizemos que as cartas (U, ϕ) e (V, ψ) são *compatíveis* se para $U \cap V \neq \emptyset$ tivermos:

- i) os conjuntos $\phi(U \cap V)$ e $\psi(U \cap V)$ são subconjuntos abertos do R^n .
- ii) a aplicação $\phi \circ \psi^{-1}$ é um difeomorfismo entre esses dois conjuntos.

As cartas serão chamadas compatíveis de classe C^r se a aplicação é um difeomorfismo de classe C^r ($1 \leq r \leq \infty$).

Definição 3.1.5: Um *atlas* para o conjunto M é uma coleção de cartas mutuamente compatíveis

$$\{(U_i, \phi_i) \mid i \in I\}$$

tal que

$$\bigcup_{i \in I} U_i = M.$$

Exemplo 3.1.6: Seja S^1 o círculo unitário do R^2 dado por

$$S^1 = \{x \in R^2 : \|x\| = 1\}.$$

Duas cartas (U, ϕ) e (V, ψ) de S^1 podem ser definidas como segue:
 Pondo $U = S^1 \setminus \{(-1, 0)\}$ e $\phi : U \rightarrow R$ dado por $\phi(x) = \text{ângulo que o vetor } x \text{ faz com o primeiro eixo, escolhido de maneira que } -\pi < \phi(x) < \pi.$
 Colocando $V = S^1 \setminus \{(1, 0)\}$ e $\psi : V \rightarrow R$ dado por $\psi(x) = \text{ângulo que o vetor } x \text{ faz com o primeiro eixo, escolhido de maneira que } 0 < \psi(x) < 2\pi.$

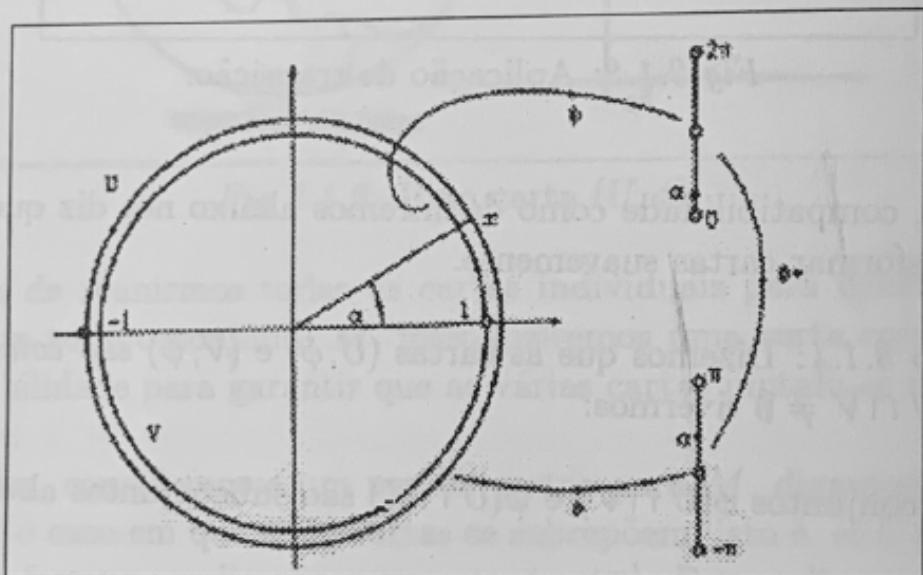


Fig 3.1.7: Um atlas para S^1 .

Com a ajuda da figura acima pode-se ver que

$$\phi \circ \psi^{-1} = \begin{cases} id & \text{sobre } (0, \pi) \\ id - 2\pi & \text{sobre } (\pi, 2\pi). \end{cases}$$

Esta aplicação é assim um difeomorfismo entre os conjuntos $\psi(U \cap V)$ e $\phi(U \cap V)$. Assim as cartas são compatíveis, onde S^1 é coberto por U e V . Portanto estas duas cartas formam um atlas para S^1 .

Exercício 3.1.8: Seja U consistido de todos os pontos (x_1, x_2) de um círculo unitário com $x_2 > 0$ e seja V consistido de todos os pontos com $x_1 > 0$. Escolha aplicações

$$\phi = \pi_1 | U \quad \text{e} \quad \psi = \pi_2 | V$$

onde π_1 e π_2 são as aplicações projetoras sobre R^2 , como na Figura 3.1.9.

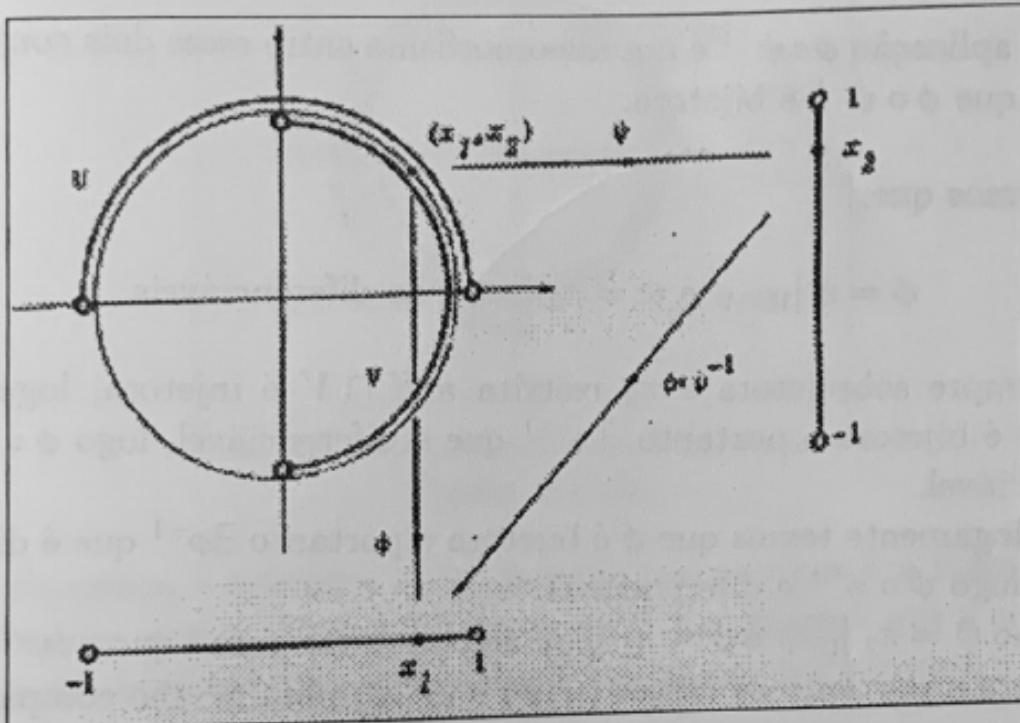


Figura 3.1.9.

a) Diga o domínio de $\phi \circ \psi^{-1}$ e encontre uma fórmula dando seus valores.

Pela figura acima vemos que o domínio de $\phi \circ \psi^{-1}$ é $(-1, 1)$, uma fórmula para esta função é encontrada pela seguinte maneira:

Temos pela figura acima que $\tan \theta = \frac{x_2}{x_1}$ e que $\sin \theta = \frac{x_2}{1} \Leftrightarrow \theta = \arcsen(x_2)$, assim substituindo θ na primeira equação encontramos uma função que relaciona x_1 com x_2 , portanto, sendo x um ponto qualquer de $(-1, 1)$ obtemos que:

$$\phi \circ \psi^{-1}(x) = \frac{x}{\tan(\arcsen(x))}.$$

b) Deduza que as cartas (U, ϕ) e (V, ψ) para S^1 são compatíveis.

Para duas cartas serem compatíveis mostraremos que:

i) os conjuntos $\phi(U \cap V)$ e $\psi(U \cap V)$ são abertos em R .

Sabendo que $U \cap V$ é aberto e que ϕ e ψ são homeomorfismos, concluimos então que $\phi(U \cap V)$ e $\psi(U \cap V)$ são abertos em R .

ii) a aplicação $\phi \circ \psi^{-1}$ é um difeomorfismo entre esses dois conjuntos e que $\phi \circ \psi^{-1}$ é bijetora.

Sabemos que:

$$\phi = \pi|_{U \cap V} \text{ e } \psi = \pi_2|_{U \cap V} \text{ são diferenciáveis}$$

π_2 é sempre sobrejetora e π_2 restrita a $U \cap V$ é injetora, logo $\psi = \pi_2|_{U \cap V}$ é bijetora e portanto $\exists \psi^{-1}$ que é diferenciável, logo $\phi \circ \psi^{-1}$ é diferenciável.

Analogamente temos que ϕ é bijetora e portanto $\exists \phi^{-1}$ que é diferenciável, logo $\psi \circ \phi^{-1}$ é diferenciável.

Como $\psi = \pi_2|_{U \cap V}$ e $\phi = \pi_1|_{U \cap V}$ são bijetoras temos que $\phi \circ \psi^{-1}|_{U \cap V}$ é bijetora e portanto as cartas (U, ϕ) e (V, ψ) para S^1 são compatíveis.

c) Como você poderá obter um atlas contendo essas duas cartas?

Para formar um atlas precisamos que:

i) as cartas sejam compatíveis dois a dois

ii) $\bigcup_{i \in I} U_i = S^1$

Apresentamos abaixo um atlas contendo as duas cartas.

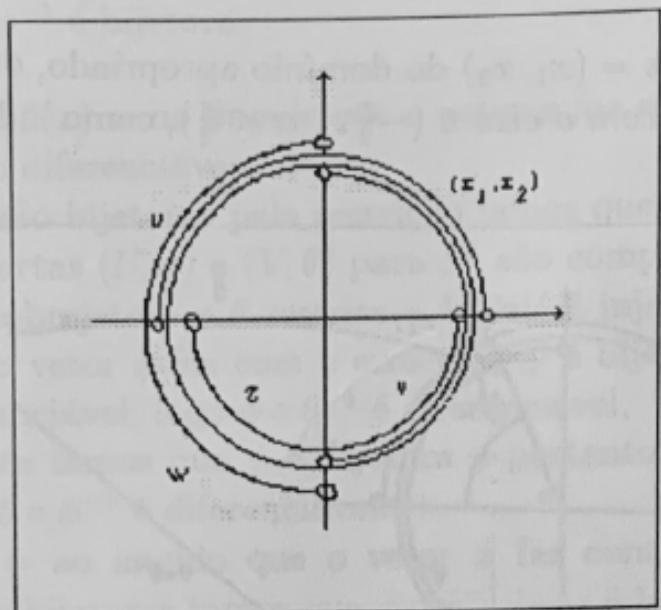


Figura 3.1.10.

Verifiquemos o primeiro item; mostramos no item b a compatibilidade de duas cartas (U, ϕ) e (V, ψ) para S^1 , verificamos que a compatibilidade das outras cartas que formam o atlas são por simetria análogas as anteriores.

Para mostrarmos o segundo item denotemos as cartas da Figura 3.1.10 na forma complexa:

$$U = e^{i\theta}, \quad 0 < \theta < \pi$$

$$W = e^{i\theta}, \quad \pi/2 < \theta < 3\pi/2$$

$$Z = e^{i\theta}, \quad \pi < \theta < 0$$

$$V = e^{i\theta}, \quad 3\pi/2 < \theta < \pi/2$$

agora fazendo a união dessas cartas temos:

$$U \cup W \cup Z \cup V = e^{i\theta}, \quad 0 \leq \theta \leq 2\pi = S^1,$$

portanto encontramos um atlas contendo as duas cartas mencionadas acima.

Exercício 3.1.11: Sejam os conjuntos U, V e ϕ como no exercício acima e sejam

$\theta : V \rightarrow R$ e $\phi : U \rightarrow R$ definido por;

para cada ponto $x = (x_1, x_2)$ do domínio apropriado, $\theta(x) = \text{ao ângulo que o vetor } x \text{ faz com o eixo } u (-\frac{\pi}{2} < \alpha < \frac{\pi}{2})$, como indicado na Figura 3.1.10.

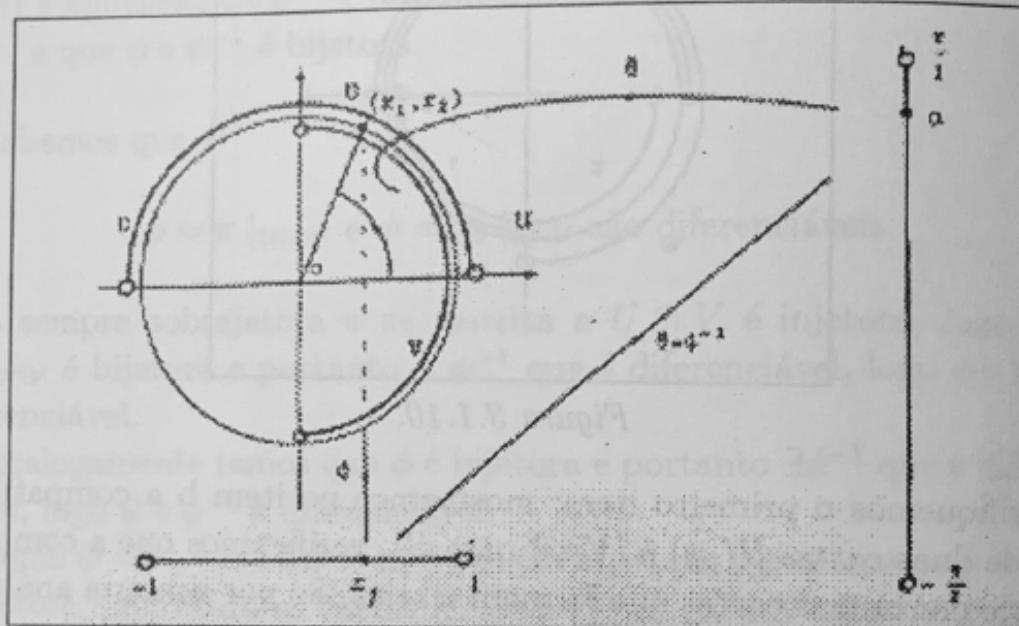


Figura 3.1.12.

a) Diga o domínio de $\theta \circ \phi^{-1}$ e encontre uma fórmula para esta aplicação.

Pela figura acima vemos que o domínio de $\theta \circ \phi^{-1}$ é $(-1, 1)$, para encontrar uma fórmula basta vermos que $x_1 = \cos \alpha$ e portanto pegando um ponto qualquer x do domínio temos que $\theta \circ \phi^{-1}(x) = \arccos(x)$.

b) Portanto mostre que as cartas (V, θ) e (U, ϕ) para S^1 são compatíveis.

Para duas cartas serem compatíveis mostraremos que:

i) os conjuntos $\theta(U \cap V)$ e $\phi(U \cap V)$ são abertos em R .

Sabendo que $U \cap V$ é aberto e que θ e ϕ são homeomorfismos, concluimos então que $\phi(U \cap V)$ e $\psi(U \cap V)$ são abertos em R .

ii) a aplicação $\theta \circ \phi^{-1}$ é um difeomorfismo entre esses dois conjuntos e que $\theta \circ \phi^{-1}$ é bijetora.

Sabemos que $\theta(x) = \text{ao ângulo que o vetor } x \text{ faz com o eixo } u|_{U \cap V}$ e $\phi = \pi_2|_{U \cap V}$ são diferenciáveis.

Como θ e ϕ são bijetoras pela restrição temos que $\theta \circ \psi^{-1}$ é bijetora e portanto as cartas (U, ϕ) e (V, θ) para S^1 são compatíveis.

θ é sempre sobrejetora e θ restrita a $U \cap V$ é injetora, logo $\theta(x) = \text{ao ângulo que o vetor } x \text{ faz com o eixo } u|_{U \cap V}$ é bijetora e portanto $\exists \theta^{-1}$ que é diferenciável, logo $\phi \circ \theta^{-1}$ é diferenciável,

Analogamente temos que ϕ é bijetora e portanto $\exists \phi^{-1}$ que é diferenciável, logo $\theta \circ \phi^{-1}$ é diferenciável.

Como $\theta(x) = \text{ao ângulo que o vetor } x \text{ faz com o eixo } u|_{U \cap V}$ e $\phi = \pi_2|_{U \cap V}$ são bijetoras temos que $\phi \circ \psi^{-1}|_{U \cap V}$ é bijetora e portanto as cartas (U, ϕ) e (V, ψ) para S^1 são compatíveis.

3.2 Definição de Uma Variedade Diferenciável

Dado um conjunto M podemos ter muitos atlas diferentes. Dizemos que um atlas A para o conjunto M é equivalente a outro atlas B para o mesmo conjunto se cada carta em A é compatível com toda carta em B .

É fácil verificar que isto define uma *relação de equivalência* sobre o conjunto de todos os atlas para o conjunto M . Cada *classe de equivalência* S do atlas é uma *estrutura diferenciável* para o conjunto M .

Uma *variedade diferenciável* é um par ordenado (M, S) onde M é um conjunto e S é uma *estrutura diferenciável* para M .

O *atlas maximal* de uma *variedade diferenciável* (M, S) é a coleção de todas as cartas que pertencem a pelo menos um atlas na *estrutura diferenciável* S .

Uma carta do atlas maximal é chamada uma *carta admissível* para a variedade diferenciável.

OBS: A definição acima é dada pelo livro texto [1]¹. A equivalente definição abaixo é mais usual.

Definição 3.2.1: Uma variedade diferenciável de dimensão n é um par ordenado (M, U) onde M é um espaço topológico de Hausdorff, com base enumerável e U um atlas maximal de dimensão n .

Ao falarmos de variedade diferenciável omitiremos a palavra diferenciável e chamaremos simplesmente de variedade.

Rigorosamente falando, uma variedade consiste de um conjunto M junto com uma estrutura diferenciável S . Onde não existir risco de ambigüidade, omitiremos a referência S e chamaremos M próprio de variedade.

Um atlas constituído de uma única carta com a função identidade $id : R^n \rightarrow R^n$ determina uma estrutura diferenciável para R^n , em que as cartas admissíveis são difeomorfismos entre subconjuntos abertos do R^n .

Quando falarmos de R^n como uma variedade, queremos dizer que R^n tem uma estrutura diferenciável.

Os conceitos mais gerais de classe C^r -equivalente, C^r -estrutura diferenciável e C^r -variedade diferenciável são definidas de modo análogo, substituindo a palavra compatível por C^r -compatível no primeiro parágrafo desta seção. Ser diferenciável para nós é ser C^∞ .

4 Subvariedades

Uma subvariedade do R^n é essencialmente um subconjunto do R^n que pode, ao menos localmente “achatar-se” em um subespaço do R^n . Um critério útil para um conjunto ser uma subvariedade do R^n é dado pelo Teorema da função implícita

¹Jones, A. and Gray, A. Manifolds and Mechanics. Cambridge University Press

4.1 O que é uma subvariedade

Queremos ser capazes de chamar uma esfera S^2 de uma subvariedade do R^3 . Na Figura 4.1.1 o difeomorfismo ϕ envia o conjunto aberto $U \subseteq R^3$ no conjunto aberto $\phi(U) \subseteq R^3$. Ao mesmo tempo o pedaço da esfera $S^2 \cap U$ “achata-se” em um subconjunto do plano $R^2 \times \{0\}$.

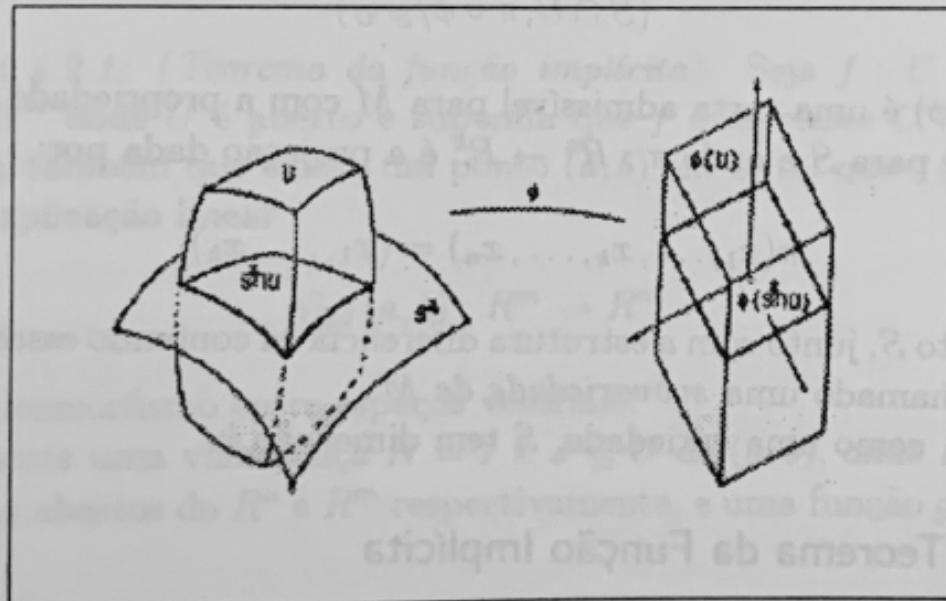


Fig 4.1.1: Uma carta de uma subvariedade para S^2 .

Definição 4.1.2: Seja (U, ϕ) uma carta para uma variedade M onde $\phi : U \rightarrow R^n$. Dizemos que esta carta tem a *propriedade de subvariedade* para um subconjunto S de M se para algum inteiro $k \leq n$,

$$(\forall x \in S \cap U) \phi(x) = (\phi_1(x), \dots, \phi_k(x), 0, \dots, 0)$$

e além disso,

$$\phi(S \cap U) = \phi(U) \cap (R^k \times \{0\}).$$

Gostaríamos de afirmar que a aplicação restrita $\phi|_{S \cap U}$ é uma carta para o conjunto S . Porém não existe esperança para isto se $k < n$, pois $\phi|_{S \cap U}$ não seria sobrejetora em um conjunto aberto do R^n . Podemos, porém, obter esta sobrejeção jogando fora os zeros do fim da n -upla. Neste caso, $\phi|_{S \cap U}$ pode ser visto como uma carta para o conjunto S .

Pode ocorrer que existam bastantes cartas para S da forma acima completando um atlas para S .

Definição 4.1.3: Suponha que S é um subconjunto de uma variedade M que tem um atlas consistindo de cartas da forma

$$(S \cap U, \pi \circ \phi|_{S \cap U})$$

onde (U, ϕ) é uma carta admissível para M com a propriedade de subvariedade para S e onde $\pi : R^n \rightarrow R^k$ é a projeção dada por:

$$\pi(x_1, \dots, x_k, \dots, x_n) = (x_1, \dots, x_k).$$

O conjunto S , junto com a estrutura diferenciável contendo essas cartas, é então chamado uma *subvariedade de M* .

Note que, como uma variedade, S tem dimensão k .

4.2 O Teorema da Função Implícita

Este *Teorema* surge quando consideramos a seguinte questão: dada uma função $f : R^2 \rightarrow R$, o conjunto de pontos

$$\{(x, y) \mid f(x, y) = 0\}$$

é o gráfico de alguma função?

Infelizmente a resposta é não. Considere a função $f : R^2 \rightarrow R$ definida por

$$f(x, y) = x^2 + y^2 - 1.$$

O conjunto de pontos $f^{-1}(0)$ é então o círculo unitário do R^2 e este não pode ser um gráfico de alguma função.

Portanto substituiremos a *questão global* por uma *questão local*: dado um ponto $(a, b) \in f^{-1}(0)$, existe alguma vizinhança $N = I \times J$ de (a, b) tal que o conjunto

$$N \cap f^{-1}(0)$$

forma o gráfico de alguma função?

No caso do círculo a resposta é sim desde que o ponto (a, b) não seja $(1, 0)$ e nem $(-1, 0)$. Note que nestes dois pontos excluídos, a derivada parcial $\partial_2 f$ é zero.

O Teorema da função implícita responde nossa questão local para uma função arbitrária $f : R^{n+m} \rightarrow R^m$.

Teorema 4.2.1: (Teorema da função implícita). Seja $f : U \subseteq R^n \times R^m \rightarrow R^m$ onde U é aberto e suponha que f é de classe C^r ($r \geq 1$). Suponha também que exista um ponto (a, b) em U tal que $f(a, b) = 0$ e que a aplicação linear

$$\partial_2 f(a, b) : R^m \rightarrow R^m$$

seja um isomorfismo entre espaços vetoriais.

Então existe uma vizinhança $N = I \times J \subseteq U$ de (a, b) , onde I e J são conjuntos abertos do R^n e R^m respectivamente, e uma função $g : I \rightarrow J$ tal que

$$\forall (x, y) \in I \times J; f(x, y) = 0 \Leftrightarrow y = g(x).$$

A função g é de classe C^r .

O seguinte *Corolário* é um importante passo para o nosso objetivo de determinar um critério para que $f^{-1}(0)$ seja uma subvariedade do R^{n+m} .

Corolário 4.2.2: Seja a aplicação f e a vizinhança N de (a, b) como no teorema acima. Então existe uma aplicação $\phi : N \rightarrow R^{n+m}$ tal que (N, ϕ) é uma carta para R^{n+m} com a propriedade de subvariedade para $f^{-1}(0)$.

4.3 Um Teste de Subvariedade

O seguinte teorema assegura um critério útil para mostrarmos que certos conjuntos são subvariedades do R^{n+m} .

Teorema 4.3.1: Seja $f : U \subset R^n \times R^m \rightarrow R^m$ onde U é aberto e suponha que f é de classe C^r ($r \geq 1$). Se em todo ponto (a, b) de $f^{-1}(0)$ a aplicação linear

$$Df(a, b) : R^{n+m} \rightarrow R^m$$

tem dimensão m , então o conjunto $f^{-1}(0)$ é uma subvariedade de R^{n+m} cuja dimensão é n .

5 Diferenciabilidade

Em cálculo avançado, diferenciabilidade foi estudada para aplicações entre conjuntos abertos no espaço vetorial normado. Neste capítulo, o conceito de diferenciabilidade será estendido para aplicações entre variedades diferenciáveis.

5.1 Representante Local

Uma variedade é um conjunto que pode ser localmente identificado com um subconjunto aberto de um espaço Euclidiano.

Esta propriedade permite-nos definir propriedades de funções entre variedades arbitrárias.

Definição 5.1.1: Sejam M e N duas variedades diferenciáveis e seja $f : M \rightarrow N$. Suponha além disso que (U, ϕ) e (V, ψ) são cartas admissíveis para M e N respectivamente, com $f(U) \subseteq V$. A aplicação

$$\psi \circ f \circ \phi^{-1} : \phi(U) \rightarrow \psi(V)$$

é chamada de *representante local* de f relativamente as duas cartas dadas e será denotada por $f_{\phi\psi}$.

Int 13 é o mao 14 que (6.13) impõe que este seja V ⊃ (a)

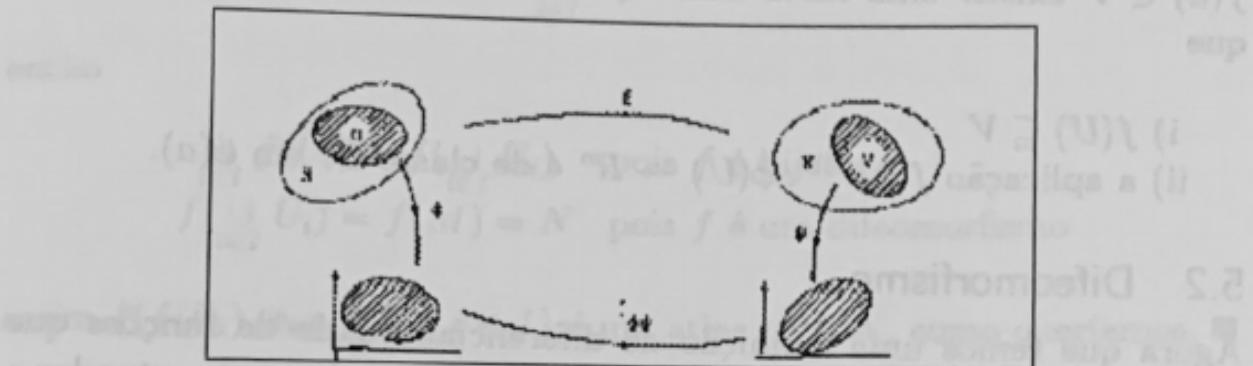


Fig 5.1.2: Representante local.

Observe que o representante local $f_{\phi\psi}$ aplica um conjunto aberto de algum espaço Euclidiano em outro. Sendo assim, faz sentido falarmos sobre diferenciabilidade entre variedades levando-se em conta o fato da diferenciabilidade de $f_{\phi\psi}$ que é estudado em cálculo avançado.

Definição 5.1.3: Seja $f : M \rightarrow N$ e seja $a \in M$. Dizemos que f é *diferenciável* em a se para cada carta admissível (V, ψ) de N com $f(a) \in V$ existir uma carta admissível (U, ϕ) para M com $a \in U$ tal que

$$f(U) \subseteq V$$

e o representante local $f_{\phi\psi}$ seja diferenciável em $\phi(a)$.

Dizemos que f é *diferenciável sobre um subconjunto* S de M se f for diferenciável em todo ponto de S . Dizemos que f é *diferenciável* se este for diferenciável sobre M .

Podemos considerar o R^n como uma variedade diferenciável com sua usual estrutura diferenciável contendo a aplicação identidade como uma carta. Assim se M for qualquer variedade faz sentido falarmos sobre diferenciabilidade de uma aplicação $f : M \rightarrow R^n$ ou uma aplicação $g : R^n \rightarrow M$.

Definição 5.1.4: Seja $f : M \rightarrow R^n$ e seja $a \in M$. A aplicação f é de classe C^r em a se e somente se para cada conjunto aberto V do R^n com

$f(a) \in V$ existir uma carta admissível (U, ϕ) para M com $a \in U$ tal que

- i) $f(U) \subseteq V$
- ii) a aplicação $f \circ \phi^{-1} : \phi(U) \rightarrow \mathbb{R}^n$ é de classe C^r em $\phi(a)$.

5.2 Difeomorfismo

Agora que temos uma definição de diferenciabilidade de funções que aplicam variedades em variedades, estamos preparados para estender a idéia de um difeomorfismo para compreender tais funções.

Definição 5.2.1: Uma aplicação bijetora $f : M \rightarrow N$, onde M e N são variedades é um *difeomorfismo* se for diferenciável e sua inversa também. Se f e sua inversa são de classe C^r então dizemos que f é um *difeomorfismo de classe C^r* .

Definição 5.2.2: Duas variedades M e N são *difeomorfas* se existir um difeomorfismo de M para N .

Podemos considerar duas variedades como sendo essencialmente a mesma se elas forem difeomorfas entre si. Esta idéia é resumida pelo lema seguinte.

Lema 5.2.3: Suponha que a variedade M é difeomorfa a variedade N e que $\{(U_i, \phi_i) / i \in I\}$ seja um atlas para M . Seja $f : M \rightarrow N$ um difeomorfismo.

Então $\{(f(U_i), \phi_i \circ f^{-1}) / i \in I\}$ é um atlas para N .

Prova: Como $\{(U_i, \phi_i) / i \in I\}$ é um atlas, temos que ϕ_i é um difeomorfismo e portanto $\phi_i \circ f^{-1}$ é também um difeomorfismo, então $(f(U_i), \phi_i \circ f^{-1})$ é uma carta para N .

Como as cartas (U_i, ϕ_i) e (U_j, ϕ_j) para $i, j \in I$ são compatíveis para M e f é um difeomorfismo, temos que $(f(U_i), \phi_i \circ f^{-1})$ e $(f(U_j), \phi_j \circ f^{-1})$ são duas cartas compatíveis para N .

Agora lembrando que

então

$$\begin{aligned}\bigcup_{i \in I} f(U_i) &= f\left(\bigcup_{i \in I} U_i\right) \quad \text{pois } f \text{ é bijetora} \\ f\left(\bigcup_{i \in I} U_i\right) &= f(M) = N \quad \text{pois } f \text{ é um difeomorfismo}\end{aligned}$$

assim $\{(f(U_i), \phi_i \circ f^{-1}) \mid i \in I\}$ é um atlas para N , como queríamos. ■

Lema 5.2.4: Seja $L : R^n \rightarrow R^m$ uma aplicação linear injetora (onde $m \geq n$). Suponha que M seja uma subvariedade do R^n . Então $L(M)$ é uma subvariedade do R^m que é difeomorfa a M .

6 Espaços e Aplicações Tangentes

No capítulo anterior a idéia de diferenciabilidade foi generalizada para aplicações entre variedades. Aqui a derivação será generalizada para variedades.

Para conservarmos a idéia de diferenciação como um processo de linearização, será conveniente introduzir em cada ponto da variedade um espaço vetorial conhecido como espaço tangente em cada ponto.

Uma formulação para a regra da cadeia será possível através da generalização do funtor tangente.

Embora todos os conceitos desta seção possam ser formulados para variedades arbitrárias, nossa discussão será restrita à subvariedades do R^n . Isto proporcionará um conhecimento adequado para as seções posteriores sobre Mecânica, evitando abstrações desnecessárias.

6.1 Espaço Tangente

Em geometria elementar estudamos tangencias à círculos e planos tangentes à esferas. A definição de espaço tangente generalizará essas idéias para curvas e superfícies arbitrárias.

Seja M uma subvariedade do R^n e seja $a \in M$. Intuitivamente, o espaço tangente em a consiste de todos os vetores localizados em

a , cujas direções são “tangenciais” à M em a , como na Figura 6.1.1. Os vetores podem ser considerados como elementos do espaço vetorial $T_a R^n$, introduzido na Seção 2.3. Para compreendermos a idéia de “tangência” usaremos curvas parametrizadas na variedade.

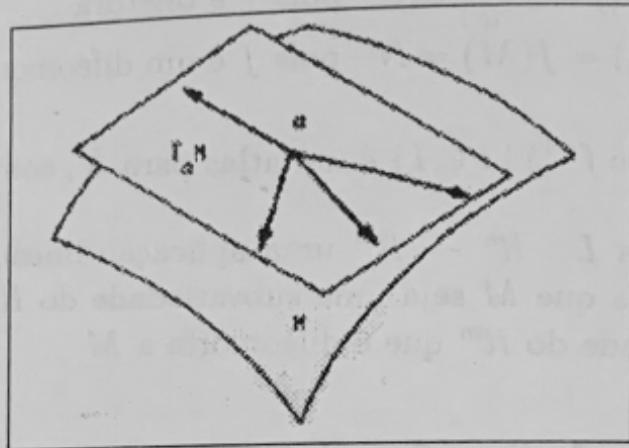


Fig 6.1.1: Um espaço tangente em a .

Definição 6.1.2: Uma curva parametrizada de M em a é uma aplicação contínua $\gamma : I \rightarrow M$ com $\gamma(0) = a$, onde I é um intervalo aberto contendo 0. A imagem $\gamma(I)$ da aplicação γ será chamada uma curva de M em a e γ será dita parametrização desta curva. A curva parametrizada será chamada diferenciável (ou de classe C^r) se a aplicação γ é diferenciável (ou de classe C^r).

Visto que $M \subseteq R^n$, a notação do Corolário 2.5.4 é aplicável à γ e pode ser interpretado como segue: se $\gamma(t)$ é a posição da partícula movendo-se ao longo da variedade no tempo t então $\gamma'(t)$ é a velocidade da partícula. A intuição requer que esta direção seja “tangencial” à variedade, levando-nos a seguinte definição, ilustrada na Figura 6.1.4.

Definição 6.1.3: Um vetor tangente à M em a é um elemento de $T_a R^n$ da forma $(a, \gamma'(0))$ onde γ é uma curva parametrizada suave de M com base em a . Denotaremos, para simplificar, este vetor tangente por $[\gamma]_a$.

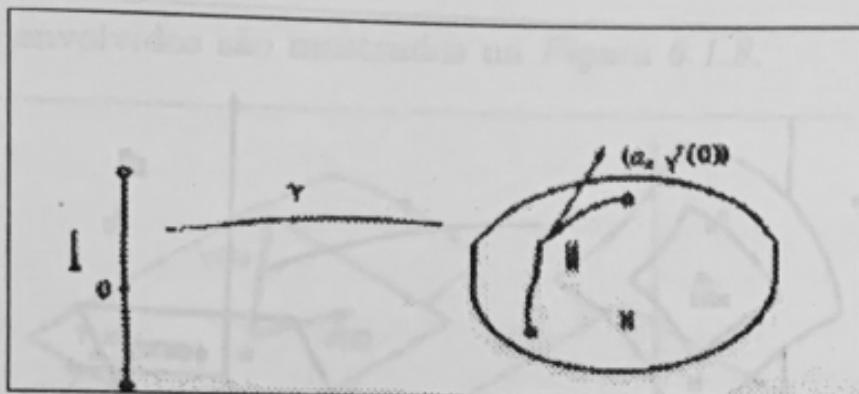


Fig 6.1.4: Um vetor tangente.

Definição 6.1.5: O espaço tangente à M em a é o conjunto de todos vetores tangentes à M em a e será denotado por $T_a M$. O fibrado tangente de M é a união de todos estes espaços tangentes e será denotado por TM tal que

$$TM = \bigcup_{a \in M} T_a M.$$

A definição de $T_a M$ implica que este seja um subconjunto do espaço vetorial $T_a R^n$ da Seção 2.3. O seguinte Teorema confirma a expectativa que o espaço tangente deverá ser “plano”, ao invés de “curvo” semelhante a uma típica variedade.

Teorema 6.1.6: O espaço tangente $T_a M$ é um subespaço vetorial de $T_a R^n$ de mesma dimensão que a variedade M .

Prova: Para provarmos, usaremos uma carta (U, ϕ) para o R^n em a que tem a propriedade de subvariedade para M . Portanto, na notação da Seção 4.1,

$$\phi(M \cap U) = R^k \times \{0\} \cap \phi(U)$$

como ilustrado na Figura 6.1.7. Nós podemos assumir que $\phi(a) = 0$.

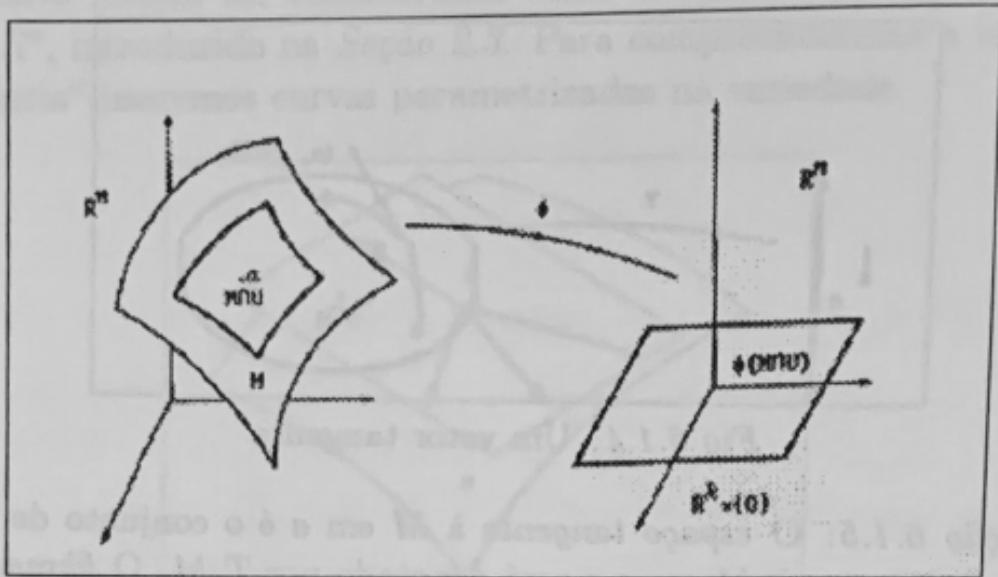


Fig 6.1.7: Uma carta de subvariedade.

Verificaremos que em $T_a M$ a adição vetorial e a multiplicação escalar são fechados. Assim sejam $(a, \gamma'(0))$ e $(a, \delta'(0))$ um par de vetores arbitrários em $T_a M$, sendo γ e δ curvas parametrizadas em M com base em a . Como ϕ é uma carta de subvariedade segue que ambas

$$\phi \circ \gamma \text{ e } \phi \circ \delta \text{ aplicam em } R^k \times \{0\}$$

e portanto também

$$(\phi \circ \gamma)' \text{ e } (\phi \circ \delta)' \text{ aplicam em } R^k \times \{0\}.$$

Pelo *Corolário 2.5.4* isto implica que

$$D(\phi \circ \gamma)(0)(1) \text{ e } D(\phi \circ \delta)(0)(1) \in R^k \times \{0\}$$

e portanto pela regra da cadeia, *Teorema 2.2.4*,

$$D\phi(a)(\gamma'(0)) \text{ e } D\phi(a)(\delta'(0)) \in R^k \times \{0\}.$$

Portanto pela linearidade de $D\phi(a)$

$$D\phi(a)(\gamma'(0) + \delta'(0)) \in R^k \times \{0\}.$$

Os vetores envolvidos são mostrados na *Figura 6.1.8*.

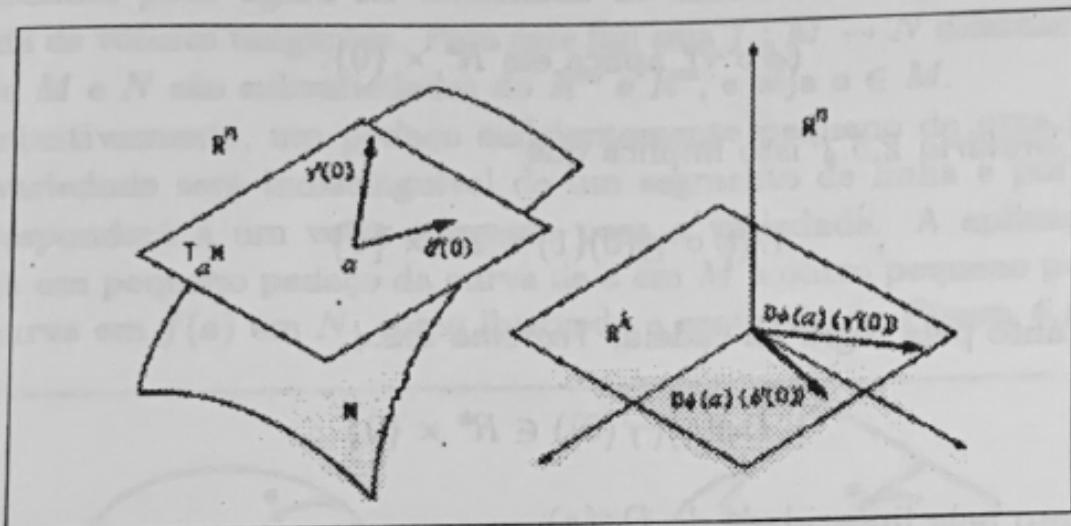


Fig 6.1.8.

Podemos deste modo definir uma aplicação ε em um intervalo de M pondo

$$\varepsilon(t) = \phi^{-1}(tD\phi(a)(\gamma'(0) + \delta'(0))).$$

Visto que ε é uma curva de M em a , o vetor $(a, \varepsilon'(0))$ é um elemento de $T_a M$. Pela regra da cadeia, porém,

$$\begin{aligned}\varepsilon'(0) &= D\phi^{-1}(0) \circ D\phi(a)(\gamma'(0) + \delta'(0)) \\ &= \gamma'(0) + \delta'(0)\end{aligned}$$

e assim, pela definição de adição vetorial em $T_a R^n$,

$$(a, (\gamma'(0))) + (a, (\delta'(0))) = (a, \varepsilon'(0)) \in T_a M.$$

Seguindo o mesmo raciocínio temos para a multiplicação escalar:

Sejam $(a, \gamma'(0))$ um vetor arbitrário em $T_a M$ e $\alpha \in R$ um escalar qualquer, sendo γ uma curva parametrizada em M com base em a . Como ϕ é uma carta de subvariedade segue que

$\phi \circ \gamma$ aplica em $R^k \times \{0\}$

e portanto

$(\phi \circ \gamma)'$ aplica em $R^k \times \{0\}$.

Pelo *Corolário 2.5.4* isto implica que

$$D(\phi \circ \gamma)(0)(1) \in R^k \times \{0\}$$

e portanto pela regra da cadeia, *Teorema 2.2.4*,

$$D\phi(a)(\gamma'(0)) \in R^k \times \{0\}.$$

Portanto pela linearidade de $D\phi(a)$

$$\alpha D\phi(a)(\gamma'(0)) = D\phi(a)(\alpha\gamma'(0)) \in R^k \times \{0\}.$$

Definindo uma aplicação ε em um intervalo de M pondo

$$\varepsilon(t) = \phi^{-1}(tD\phi(a)(\alpha\gamma'(0))).$$

Visto que ε é uma curva de M em a , o vetor $(a, \varepsilon'(0))$ é um elemento de $T_a M$. Pela regra da cadeia, porém,

$$\begin{aligned}\varepsilon'(0) &= D\phi^{-1}(0) \circ D\phi(a)(\alpha\gamma'(0)) \\ &= \alpha\gamma'(0)\end{aligned}$$

e assim, pela definição de multiplicação por escalar em $T_a R^n$,

$$(a, \alpha(\gamma'(0))) = (a, \varepsilon'(0)) \in T_a M.$$

Observe que uma carta para $T_a M$ tem a mesma dimensão que M . Deste modo $T_a M$ é um subespaço vetorial de $T_a R^n$. ■

6.2 Aplicações Tangentes

A idéia de uma “aproximação linear local” para uma aplicação entre variedades pode agora ser formulada de uma maneira precisa com a ajuda de vetores tangentes. Para este fim seja $f : M \rightarrow N$ diferenciável, onde M e N são subvariedades do \mathbb{R}^m e \mathbb{R}^n , e seja $a \in M$.

Intuitivamente, um pedaço suficientemente pequeno de uma curva na variedade será indistinguível de um segmento de linha e portanto corresponderá a um vetor tangente para a variedade. A aplicação f envia um pequeno pedaço da curva de a em M à outro pequeno pedaço da curva em $f(a)$ em N , como ilustrado à esquerda da Figura 6.2.1.

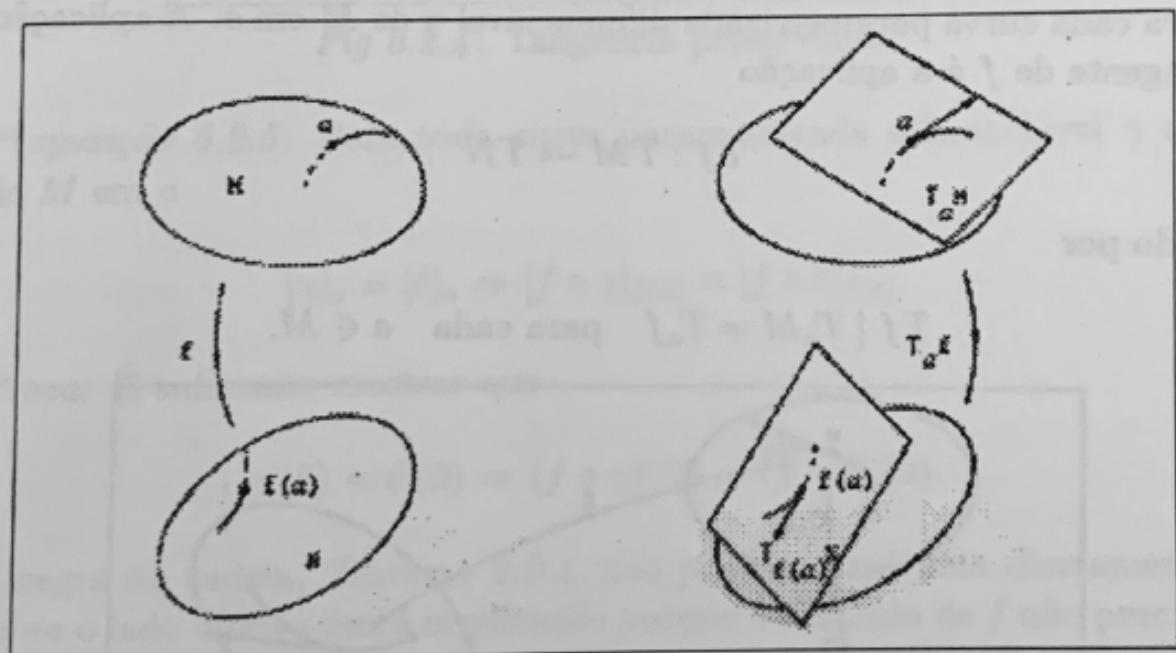


Fig 6.2.1: A aplicação tangente de f em a .

Agora imaginemos um pedaço da curva tão pequeno que podemos considerá-lo como vetores tangentes, obteremos a aplicação, mostrada na figura como $T_a f$, que leva vetores tangentes a $T_a M$ em vetores tangentes a $T_{f(a)} N$. Esta aplicação é uma aproximação de f perto de a e, felizmente, esta é linear.

Introduzindo parametrizações para estas curvas como na Figura 6.2.3 e usando suas derivadas para definir os vetores tangentes, estamos preparados para apresentar a seguinte *Definição*.

Definição 6.2.2: A aplicação tangente de f em a é a aplicação

$$T_a f : T_a M \rightarrow T_{f(a)} N$$

com

$$T_a f(a, \gamma'(0)) = (f(a), (f \circ \gamma)'(0))$$

ou, equivalentemente,

$$Tf([\gamma]_a) = [f \circ \gamma]_{f(a)}$$

para cada curva parametrizada diferenciável γ de M em a . A aplicação tangente de f é a aplicação

$$Tf : TM \rightarrow TN$$

dado por

$$Tf |_{T_a M} = T_a f \quad \text{para cada } a \in M.$$

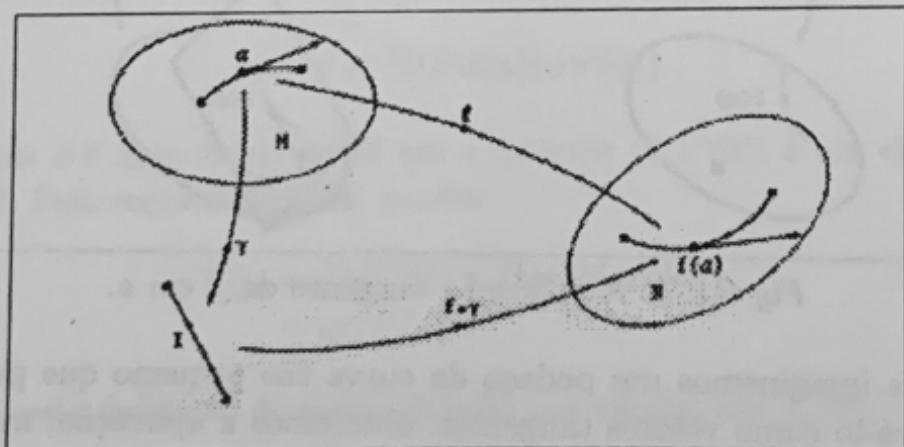


Fig 6.2.3.

Isto é necessário para checar que a definição proposta evidencia que duas curvas distintas γ e δ possam ter a mesma derivada em 0. Queremos estar certos que quando isto ocorrer as derivadas das curvas $f \circ \gamma$ e $f \circ \delta$ serão também iguais a 0, como ilustrado na Figura 6.2.4.

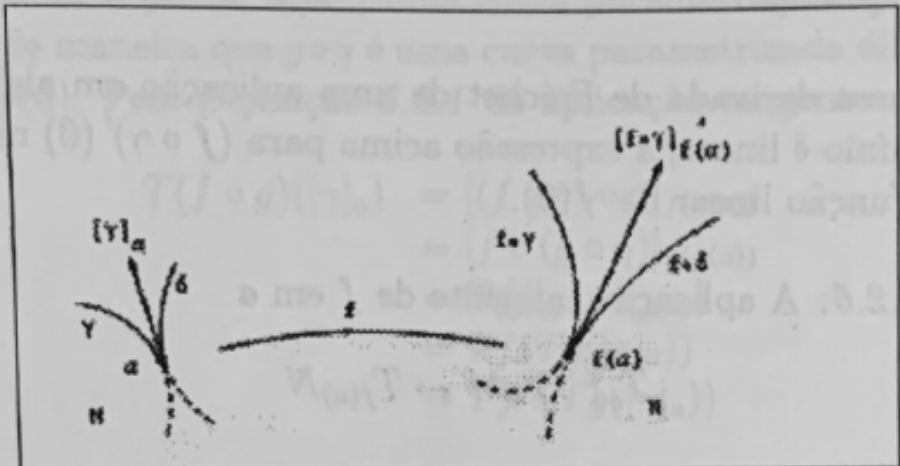


Fig 6.2.4: Tangência preservada.

Proposição 6.2.5: Para toda curva parametrizada diferenciável γ e δ de M em a

$$[\gamma]_a = [\delta]_a \Rightarrow [f \circ \gamma]_{f(a)} = [f \circ \delta]_{f(a)}.$$

Prova: É suficiente mostrar que

$$\gamma'(0) = \delta'(0) \Rightarrow (f \circ \gamma)'(0) = (f \circ \delta)'(0).$$

A regra da cadeia, *Teorema 2.2.4*, não pode ser aplicada diretamente sobre o lado direito desta implicação porque o domínio de f não precisa ser um conjunto aberto em um espaço vetorial normado. Para vencermos esta dificuldade, introduzimos uma carta (U, ϕ) do R^n em a com a propriedade de subvariedade para M e então escrevemos, ao menos sobre uma vizinhança de $0 \in R$,

$$f \circ \gamma = (f \circ \phi^{-1}) \circ \phi \circ \gamma.$$

Agora pelo *Corolário 2.5.4* e a regra da cadeia

$$\begin{aligned} (f \circ \gamma)'(0) &= D(f \circ \gamma)(0)(1) \\ &= D(f \circ \phi^{-1})(\phi(a)) \circ D\phi(\gamma(0)) \circ D\gamma(0)(1) \\ &= D(f \circ \phi^{-1})(\phi(a)) \circ D\phi(a)(\gamma'(0)). \end{aligned}$$

Isto, junto com uma expressão similar para $(f \circ \delta)'(0)$, estabelece a desejada implicação. ■

Visto que a derivada de Fréchet de uma aplicação em algum ponto de seu domínio é linear, a expressão acima para $(f \circ \gamma)'(0)$ mostra que ela é uma função linear de $\gamma'(0)$.

Teorema 6.2.6: A aplicação tangente de f em a

$$T_a f : T_a M \rightarrow T_{f(a)} N$$

é linear.

Com a ajuda da aplicação tangente, a regra da cadeia para aplicações entre variedades pode agora ser expressa da seguinte forma.

Teorema 6.2.7: Sejam $g : L \rightarrow M$ e $f : M \rightarrow N$ diferenciáveis, onde L, M e N são subvariedades do *Espaço Euclidiano*. A composta $f \circ g : L \rightarrow N$ é então diferenciável e

$$T(f \circ g) = T_f \circ T_g.$$

Se $a \in M$ então podemos escrever

$$T_a(f \circ g) = T_{g(a)} f \circ T_a g.$$

Prova: Primeiro provaremos a diferenciabilidade da composta.

Como f é diferenciável então para cada carta admissível (V, ϕ) de N com $f(a) \in V$ existe uma carta admissível (U, ψ) de M com $a \in U$ tal que $f(U) \subseteq V$ e o representante local $f_{\psi\phi}$ é diferenciável em $\psi(a)$. Analogamente como g é diferenciável temos que para cada carta admissível (U, ψ) de M com $g(a) \in U$ existe uma carta admissível (Q, δ) de L com $a \in Q$ tal que $g(Q) \subseteq U$ e o representante local $g_{\delta\psi}$ é diferenciável em $\delta(a)$. Assim para cada carta admissível (V, ϕ) de N com $f(a) \in V$ existe uma carta admissível (Q, δ) de L com $a \in Q$ tal que $f \circ g(Q) = f(g(Q)) \subseteq V$ e o representante local $(f \circ g)_{\delta\phi} = f_{\psi\phi} \circ g_{\delta\psi}$ é diferenciável (como já sabemos do cálculo avançado).

Assim fica provada a diferenciabilidade da composta.

Para completar a prova, seja γ uma curva parametrizada diferenciável de L em a de maneira que $g \circ \gamma$ é uma curva parametrizada diferenciável de M em $g(a)$. Pela Definição 6.2.1 da aplicação tangente

$$\begin{aligned} T(f \circ g)([\gamma]_a) &= [(f \circ g) \circ \gamma]_{f(g(a))} \\ &= [f \circ (g \circ \gamma)]_{f(g(a))} \\ &= Tf([g \circ \gamma]_{g(a)}) \\ &= Tf(Tg([\gamma]_a)) \\ &= Tf \circ Tg([\gamma]_a) \end{aligned}$$

como queríamos. ■

7 Fibrados Tangentes

Neste capítulo mostraremos que o fibrado tangente TM de uma variedade M é uma variedade em si mesmo. Além disso ele tem uma estrutura diferenciável que é induzida naturalmente pela estrutura diferenciável de M . Estas idéias podem ser estendidas para a forma $T(TM)$ que é naturalmente a situação para equações diferenciais de segunda ordem que são fundamentais na Mecânica e em outras aplicações.

7.1 Cartas para TM

Lembremos que o fibrado tangente de uma subvariedade M do R^n é definido por

$$TM = \bigcup_{a \in M} T_a M.$$

Este é um subconjunto do TR^n e portanto pode ser considerado como um subconjunto do $R^n \times R^n$.

Por exemplo, o fibrado tangente TS^1 é a coleção de todos espaços tangentes ao círculo unitário. Embora este seja rigorosamente um subconjunto do R^4 , ele é útil para visualização no plano pela projeção de linhas tangentes anexadas a S^1 , como na Figura 7.1.1.

Inte, juntó com *análise* que é o estudo das propriedades de um sistema de equações diferenciais. A teoria das equações diferenciais é dividida em duas partes principais: a teoria das equações ordinárias e a teoria das equações parciais.

Vista da perspectiva da matemática, a teoria das equações diferenciais é uma forma de resolver sistemas de equações. A teoria das equações diferenciais é uma forma de resolver sistemas de equações.

Teorema 6.1.6: Seja $\phi : U \rightarrow M$ uma carta de M .

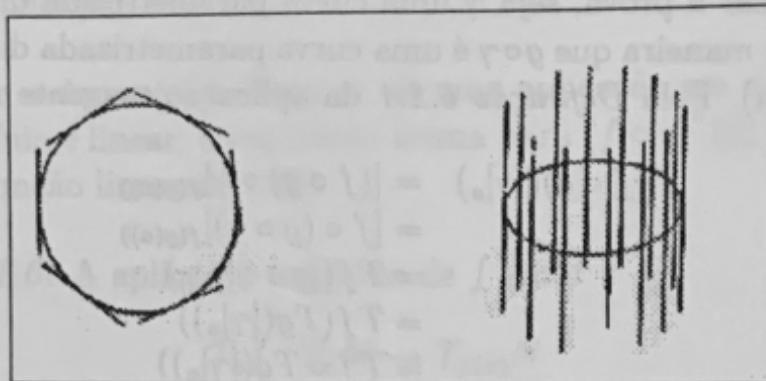


Fig 7.1.1: O fibrado tangente TS^1 .

Ele é facilmente visualizado se dermos a cada linha tangente uma rotação perpendicular ao plano do círculo para formar um cilindro, veja *Figura 7.1.1*. Entretanto o seguinte resultado mostra como cartas de M conduz, pela diferenciação, a cartas de uma subvariedade para TM .

Lema 7.1.2: Seja M uma subvariedade do R^n de classe C^r ($r \geq 2$) de dimensão k . Se (U, ϕ) é uma carta do R^n com a propriedade de subvariedade para M , então $(TU, T\phi)$ é uma carta para R^{2n} tal que

$$T\phi(TU \cap TM) = T(\phi(U)) \cap (R^k \times \{0\} \times R^k \times \{0\}).$$

Prova: Seja (U, ϕ) como na hipótese do *Lema*. A aplicação $T\phi : TU \rightarrow T\phi(TU)$ é bijetora e o conjunto

$$T\phi(TU) = T(\phi(U)) = \phi(U) \times R^n,$$

é aberto em R^{2n} . Portanto $(TU, T\phi)$ é uma carta para R^{2n} . Finalmente,

$$\begin{aligned} T\phi(TU \cap TM) &= T\phi(T(U \cap M)) \\ &= T(\phi(U \cap M)) \\ &= T(\phi(U) \cap (R^k \times \{0\})) \\ &= (\phi(U) \cap (R^k \times \{0\})) \cap (R^k \times \{0\} \times R^k \times \{0\}) \\ &= T\phi(U) \cap (R^k \times \{0\} \times R^k \times \{0\}). \end{aligned}$$

Assim tudo que resta para obter uma carta para TM é mudar os zeros para a direita. Isto pode ser obtido aplicando, por exemplo, a permutação $P : R^{2n} \rightarrow R^{2n}$ com

$$P(x, w, y, z) = (x, y, w, z)$$

para $x, y \in R^k$ e $w, z \in R^{n-k}$. ■

Corolário 7.1.3: Se (U, ϕ) é uma carta para R^n com a propriedade de subvariedade para M então $(TU, P \circ (T\phi))$ é uma carta para R^{2n} com a propriedade de subvariedade para TM .

Corolário 7.1.4: Se M é uma subvariedade do R^n de classe C^r ($r \geq 2$) então TM é uma subvariedade do R^{2n} de classe C^{r-1} . A dimensão de TM é duas vezes a dimensão de M .

Teorema 7.1.5: Se (U, ϕ) é uma carta de classe C^r ($r \geq 2$) para uma subvariedade M do R^n então $(TU, T\phi)$ é uma carta de classe C^{r-1} para TM .

7.2 Paralelizabilidade

Na seção anterior mostramos que $TS^1 = S^1 \times R$ e fizemos a observação que $TR^n = R^n \times R^n$. Por esta razão parece-nos natural presumir que em geral $TM = M \times R^K$. Esta presunção é falsa.

Definição 7.2.1: Seja M uma subvariedade K -dimensional de classe C^r ($r \geq 2$) do R^n . Se existir um difeomorfismo

$$\phi : TM \rightarrow M \times R^k$$

que leva cada espaço tangente $T_a M$ por um isomorfismo em $\{a\} \times R^k$ então M é chamado paralelizável.

Um importante exemplo de uma variedade que não é paralelizável, o qual surge no estudo do pêndulo esférico, é o S^2 .

Lema 7.2.2: Seja M uma subvariedade K -dimensional de classe C^r ($r \geq 2$) e seja (U, ϕ) uma carta admissível para M . Então $TU = U \times R^k$.

7.3 Aplicações Tangentes e Suavidade

Definição 7.3.1: Seja $\{(U_i, \phi_i) : i \in I\}$ um atlas para uma subvariedade M de classe C^r ($r \geq 2$) do R^n . Então a coleção de cartas para TM dado por $\{(TU_i, T\phi_i) : i \in I\}$ é chamado atlas natural para TM .

É agora significativo discutir diferenciabilidade de aplicações que tem fibrados tangentes como domínio. Um importante exemplo de tal aplicação é dada pela seguinte definição.

Definição 7.3.2: Seja M uma subvariedade de classe C^r ($r \geq 2$) do R^n . A projeção natural sobre o fibrado tangente de M é a aplicação $\tau_M : TM \rightarrow M$ com

$$\tau_M(a, h) = a$$

para cada $(a, h) \in T_a M$.

Teorema 7.3.3: A projeção natural τ_M é de classe C^∞ .

Teorema 7.3.4: Sejam M e N subvariedades de classe C^r ($r \geq 2$) do R^m e R^n respectivamente. Se $f : M \rightarrow N$ é de classe C^r então $Tf : TM \rightarrow TN$ é de classe C^{r-1} .

8 Equações de Lagrange Para Queda Livre

Primeiramente deduziremos as *equações de Lagrange* para um sistema mecânico muito simples consistido de uma partícula caindo livremente sob ação da gravidade.

A lei de Newton que rege o movimento de uma partícula de massa m é $\vec{F} = m\vec{a}$ (desde que as medidas sejam feitas em um *sistema inercial de referência*), onde a é a aceleração da partícula.

Para o estudo do movimento perto da superfície da terra é muitas vezes útil supor que eixos fixos na terra forneçam tal sistema de referência, porém para o estudo de movimentos planetários é mais conveniente fixar os eixos relativo as estrelas distantes. Consideremos então $\{\vec{i}, \vec{j}, \vec{k}\}$ base ortonormal positiva do R^3 fixo relativo a terra com os eixos x e y na horizontal e o eixo z na vertical. Consideremos também uma partícula de massa m movendo-se perto da superfície da terra sob ação de uma força constante de magnitude mg verticalmente descendente, veja *Figura 8.1.1*.

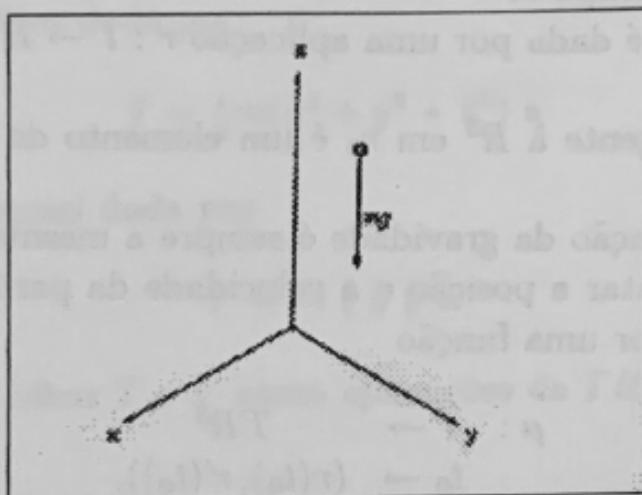


Fig 8.1.1: Movimento de uma partícula sob a ação de uma força gravitacional.

Seja $\vec{F} = m\vec{g}$ esta força atuando sobre uma partícula de massa m , onde \vec{g} é a aceleração da gravidade da terra.

Seja $\vec{r} = (r_x \vec{i} + r_y \vec{j} + r_z \vec{k})$ a posição da partícula no instante t , isto é $\vec{r} = r(t)$, onde $r : I \rightarrow R^3$ é a equação do movimento da partícula.

Temos que:

$$\vec{F} = m\vec{a} \text{ é equivalente a } \vec{F} = m \frac{d^2 \vec{r}}{dt^2}.$$

Como na queda livre o vetor \vec{g} tem apenas uma componente não nula, aquela dada pela direção de \vec{k} , a força \vec{F} também terá apenas a componente F_z não nula.

A equação (1) se reduz a

$$F_z = m \frac{d^2 r_z}{dt^2} = -mg_z.$$

Portanto as equações de Newton para queda livre são:

$$\begin{cases} F_x = m \frac{d^2 r_x}{dt^2} = 0 \\ F_y = m \frac{d^2 r_y}{dt^2} = 0 \\ F_z = m \frac{d^2 r_z}{dt^2} = -mg_z. \end{cases}$$

Consideremos agora uma partícula descrevendo uma curva no R^3 em um intervalo de tempo I .

Então a curva é dada por uma aplicação $r : I \rightarrow R^3$ tal que $r(t_0) = \vec{r}_0$.

Um vetor tangente à R^3 em r_0 é um elemento de $T_{r_0}R^3$ da forma $(r_0, r'(t_0))$.

Como a aceleração da gravidade é sempre a mesma em cada ponto podemos representar a posição e a velocidade da partícula em relação ao tempo $t \in I$ por uma função

$$\begin{aligned} \mu : & I \rightarrow TR^3 \\ & t_0 \rightarrow (r(t_0), r'(t_0)). \end{aligned}$$

Note que

$$\begin{array}{ccc} I \xrightarrow{\mu} & TR^3 \xrightarrow{\tau_{R^3}} & R^3 \xrightarrow{X} R^3 \\ t \rightarrow & (r(t), r'(t)) \rightarrow & r(t) \rightarrow r(t) \end{array}$$

onde $\tau_{R^3} : TR^3 \rightarrow R^3$ é a projeção natural sobre o fibrado tangente e X é uma carta para o R^3 .

Portanto a coordenada cartesiana da partícula no tempo t será $(X \circ \tau_{R^3} \circ \mu)(t)$ e assim as leis de Newton para o movimento implicam que

$$\begin{cases} m(x \circ \tau_{R^3} \circ \mu)(t) = 0 \\ m(y \circ \tau_{R^3} \circ \mu)(t) = 0 \\ m(z \circ \tau_{R^3} \circ \mu)(t) = -mg \end{cases} \quad (1)$$

OBS: No curso de Física as equações são apresentadas na forma

$$\begin{cases} mx''(t) = 0 \\ my''(t) = 0 \\ mz''(t) = -mg. \end{cases}$$

Isto porque na Física, o espaço tangente é considerado como sendo o R^3 .

Existem quantidades dinâmicas associadas com o movimento da partícula, são elas:

-Energia cinética dada por

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

-Energia potencial dada por

$$V = m \parallel \vec{g} \parallel z.$$

Nós podemos olhar T e V como aplicações de TR^3 e R^3 respectivamente em R .

$$\begin{array}{ccc} T : TR^3 & \rightarrow & R \\ (a, (\dot{x}(t_0), \dot{y}(t_0), \dot{z}(t_0))) & \rightarrow & \frac{1}{2}m(\dot{x}^2(t_0) + \dot{y}^2(t_0) + \dot{z}^2(t_0)) \\ (r(t_0), r'(t_0)) & \rightarrow & \frac{1}{2}m(r'^2(t_0)) \\ V : R^3 & \rightarrow & R \\ (x, y, z) & \rightarrow & m \parallel \vec{g} \parallel z. \end{array}$$

Agora definimos a função lagrangiana $L : TR^3 \rightarrow R$ como

$$L = T - V \circ \tau_{R^3}.$$

A equação (1) pode ser escrita em termos de L como

$$\begin{cases} \left(\frac{\partial L}{\partial \dot{x}} \circ \mu\right)' - \frac{\partial L}{\partial x} \circ \mu = 0 \\ \left(\frac{\partial L}{\partial \dot{y}} \circ \mu\right)' - \frac{\partial L}{\partial y} \circ \mu = 0 \\ \left(\frac{\partial L}{\partial \dot{z}} \circ \mu\right)' - \frac{\partial L}{\partial z} \circ \mu = 0. \end{cases} \quad (2)$$

As equações (2) são as *equações de Lagrange* para o movimento de uma partícula em queda livre. Elas tem uma vantagem sobre as *equações de Newton* (1) para o movimento devido ao fato que sua forma é preservada sob transformações de coordenadas. A força gravitacional no exemplo acima é um exemplo simples de um campo de força conservativo, como definimos abaixo.

Definição 8.1.2: Se a força que age sobre uma partícula tem a forma

$$F(a) = \left(a, -\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z} \right)(a) \right) \in T_a R^3$$

em cada $a \in U$, U um subconjunto aberto de R^3 , para alguma função $V : U \rightarrow R$, nós dizemos que a aplicação $F : U \rightarrow TR^3$ é um *campo de força conservativo*.

Esta definição tem uma evidente generalização para o R^n .

9 Aplicações

Apresentamos a solução de alguns exercícios que consideramos interessantes.

- 1) Mostre que o campo de força gravitacional descrito no texto é conservativo.

O campo de força gravitacional é da forma

$$F(a, b, c) = ((a, b, c), -(0, 0, mg)).$$

Seja U um aberto do R^3 contendo (a, b, c) e seja $V : U \rightarrow R^3$ dado por:

$$V(x, y, z) = mgz.$$

Temos que $\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z} \right)(a, b, c) = (0, 0, mg)$ e portanto podemos escrever a força $F(a, b, c)$ como

$$\begin{aligned} F(a, b, c) &= ((a, b, c), -(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z})(a, b, c)) \\ &= ((a, b, c), -(0, 0, mg)) \in T_{(a,b,c)}R^3. \end{aligned}$$

Segue pela *Definição 8.1.2* que este campo de força é conservativo.

- 2) Mostre que o campo de força definido sobre R^1 por $F(a) = (a, -ka)$ é conservativo.

Seja U um aberto de R contendo a e seja $V : U \rightarrow R$ dado por:

$$V(x) = \frac{1}{2}kx^2.$$

Temos que $\frac{dV}{dx}(a) = ka$ e portanto podemos escrever a força $F(a)$ como

$$F(a) = (a, -\frac{dV}{dx}(a)) = (a, -ka) \in T_a R.$$

Segue pela *Definição 8.1.2* que este campo de força é conservativo.

- 3) Mostre que o campo de força definido sobre $R^3 \setminus \{0\}$ por

$$F(a, b, c) = ((a, b, c), -\frac{1}{(\sqrt{a^2+b^2+c^2})^3}(a, b, c))$$

é conservativo. (Este campo de força corresponde a atração gravitacional produzida por uma partícula de massa unitária na origem).

Seja U um aberto do $R^3 \setminus \{0\}$ contendo (a, b, c) e seja $V : U \rightarrow R^3$ dado por:

$$V(x, y, z) = -\frac{1}{\sqrt{x^2+y^2+z^2}}.$$

Temos que $(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z})(a, b, c) = (\frac{1}{(\sqrt{a^2+b^2+c^2})^3}(a, b, c))$ e portanto podemos escrever a força $F(a, b, c)$ como

$$\begin{aligned} F(a, b, c) &= ((a, b, c), -(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z})(a, b, c)) \\ &= ((a, b, c), -\frac{1}{(\sqrt{a^2+b^2+c^2})^3}(a, b, c)) \in T_{(a,b,c)}R^3 \end{aligned}$$

Pela Definição 8.1.2 este campo de força é conservativo.

Bibliografia

- [1] Jones, A. and Gray, A. Manifolds and Mechanics. Cambridge University Press.
- [2] Hirsch, M. W. Differential Topology. Springer Verlag. New York.
- [3] Wells, D. A. Theory and Problems of Lagrangian Dynamics. McGraw Hill Book Company, New York.
- [4] Abraham, Ralph and Marsden, J. E. Foundations of Mechanics. The Benjamin Cummings Publ. Co.

3 Médias Geométrica e Aritmética

Sobre Algumas Desigualdades Importantes na Análise

Monica Regina Gaiotto*

Orientadoras: Profa. Dra. Nativi Viana Pereira Bertolo
Profa. Dra. Rosa Lúcia Swerzut Baroni

1 Introdução

Um dos objetivos destas notas é exibir uma série de desigualdades, muitas delas com interpretação geométrica. Vale a pena colocar algumas questões neste momento: por que estudar desigualdades? que papel elas desempenham na Matemática? Não é à-toa que desde que o estudante inicia seus estudos de Matemática elas estão presentes no seu dia a dia.

Vejamos algumas situações bem simples onde elas aparecem. Digamos que desejamos traçar a região \mathbb{R} do plano limitada por duas circunferências concêntricas de raio $r_2 > r_1$, respectivamente e centro $C = (a, b)$. Claro que essa região pode ser descrita da seguinte maneira:

$$\mathcal{R} = \{(x, y) \in \mathbb{R}^2 / r_1 < |(x, y) - (a, b)| < r_2\}$$

Outro exemplo seria determinar o domínio da função

$$f(x, y) = \sqrt{x - y}$$

que é dado por

*Bolsista PET-CAPES

$$\mathcal{D} = \{(x, y) \in \mathbb{R}^2 / x - y \geq 0\}.$$

De um modo geral as desigualdades estão presentes no estudo do comportamento de funções, no estudo de espaços normados, no estudo de crescimento populacional onde estão presentes as funções exponencial e logarítmica, no conceito de limite, convergência, etc.

Motivados por tudo isso selecionamos algumas desigualdades clássicas que serão objeto de estudo neste trabalho. Para isso, inicialmente faremos uma parte preliminar sobre as noções básicas de ordem em \mathbb{R} .

2 Preliminares

Nestas notas \mathbb{R} denotará o corpo dos números reais em relação à adição e multiplicação, $+$ e \cdot , usuais dos números reais. Vamos admitir que o leitor já tenha conhecimento dos axiomas da adição e da multiplicação. Além disso, assumiremos que o leitor saiba que \mathbb{R} é um corpo ordenado, pois aí temos uma relação de ordem, a relação de ordem usual, $<$, que satisfaz os axiomas:

(i) Dado $x \in \mathbb{R}$, apenas uma das três alternativas ocorre:

$$x > 0, x = 0, x < 0.$$

(ii) $\forall x, y \in \mathbb{R}$, com $x > 0, y > 0$, então

$$x + y > 0 \text{ e } x \cdot y > 0.$$

(iii) $\forall x, y \in \mathbb{R}$, então $x < y \Leftrightarrow y - x > 0$.

Esses axiomas implicam numa série de propriedades relativas às operações $+$ e \cdot e à relação de ordem, $<$ que serão utilizadas na sequência de nosso trabalho. Lembramos que a notação $a \leq b$ significa que $a = b$ ou que $a < b$.

3 Médias Geométrica e Aritmética

Suponhamos que um aluno obteve as notas P_1, P_2, \dots, P_n nas n provas de seu curso. O professor lhe deu o direito de optar por uma das alternativas para o cálculo da média final: a primeira é simplesmente a média aritmética das notas

$$M_A = \frac{P_1 + P_2 + \dots + P_n}{n}$$

e a segunda, é a média geométrica das notas

$$M_G = \sqrt[n]{P_1 \cdot P_2 \cdot \dots \cdot P_n}$$

Colocamos: qual é a média mais vantajosa? A resposta será dada a seguir, em várias etapas, com demonstração de cada afirmação feita.

3.1 Relação entre as médias geométrica e aritmética para dois números não negativos

Sejam a e b números reais não-negativos, então:

$$(1) \quad \sqrt{a.b} \leq \frac{a+b}{2},$$

e a igualdade vale se, e somente se, $a = b$.

Demonstração: desde que $a \geq 0$ e $b \geq 0$, então

$$\begin{aligned} \sqrt{a.b} &\leq \frac{a+b}{2} \Leftrightarrow a.b \leq \frac{(a+b)^2}{4} \Leftrightarrow 4.a.b \leq a^2 + b^2 + 2.a.b \Leftrightarrow \\ &\Leftrightarrow a^2 + b^2 - 2.a.b \geq 0 \Leftrightarrow (a-b)^2 \geq 0, \end{aligned}$$

e a última desigualdade é sempre válida, pois o quadrado de qualquer número real é sempre um número não-negativo. Claro que a igualdade vale, se, e somente se, $a = b$, daí a segunda parte da afirmação está provada.

Dessa forma, concluímos que a vantagem do aluno está em calcular a média final do curso através da média aritmética, isso se o número de provas for 2, por enquanto.

Antes de passarmos para outra etapa, vejamos uma interpretação geométrica da desigualdade (1) acima. Ela diz que a altura de um triângulo retângulo cuja base é a hipotenusa não ultrapassa a metade da hipotenusa. Ver figura 1

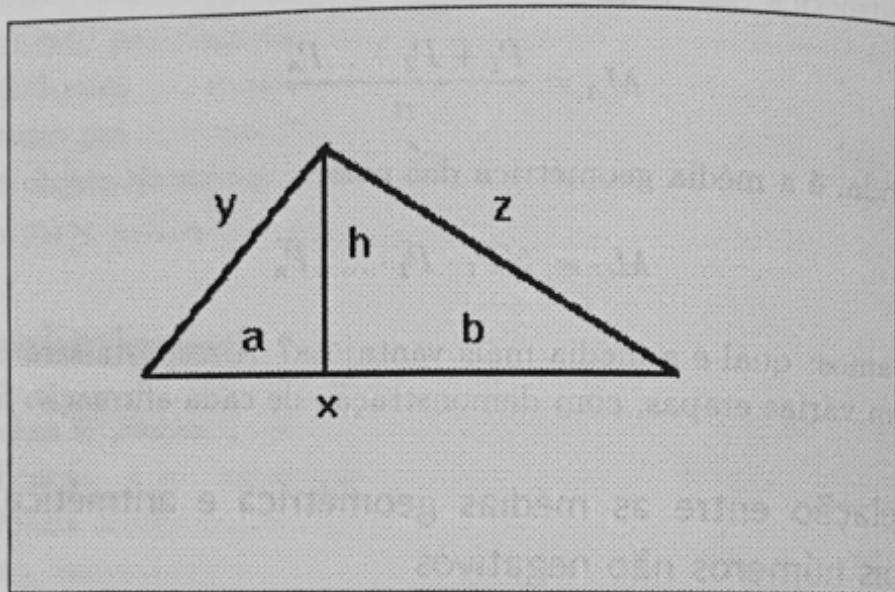


Figura 1

Com efeito, considerando a figura acima, pelo Teorema de Pitágoras, temos que

$$h^2 = y^2 - a^2 \text{ e } h^2 = z^2 - b^2$$

onde a e b são as respectivas projeções dos lados y e z sobre o lado x . Logo, é fácil concluirmos que:

$$2 \cdot h^2 = y^2 + z^2 - (a^2 + b^2)$$

e pelo Teorema de Pitágoras, novamente, temos que

$$2 \cdot h^2 = 2 \cdot a \cdot b$$

onde

$$h = \sqrt{a \cdot b} \leq \frac{a+b}{2}$$

Logo,

$$h \leq \frac{x}{2}$$

que é exatamente o que foi anteriormente afirmado.

3.2 Relação entre as médias geométrica e aritmética para n números não negativos

Sejam a_1, a_2, \dots, a_n números reais não-negativos. Então

$$(2) \quad \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n} \leq \frac{a_1 + a_2 + \dots + a_n}{n}, \forall n \in \mathbb{N},$$

e a igualdade vale se, e somente se, $a_1 = a_2 = \dots = a_n$.

Demonstração. Esta demonstração será feita em dois estágios que explicaremos a seguir:

(i) No primeiro estágio mostraremos que (2) vale para todo n da forma 2^k , isto é, $n = 2^k$, para algum $k \geq 1$, k inteiro.

(i i) No segundo estágio provaremos que se $n > 1$ e $P(n)$ for verdadeira, então $P(n-1)$ será verdadeira, onde $P(n)$ é a desigualdade (2) para n .

Claro que de (i) e (i i) concluimos que (2) é verdadeira para todo número natural n . Com efeito. Consideremos $n \in \mathbb{N}$. Se existir k tal que $n = 2^k$, então estamos no caso (i), e não há nada a provar. Se $n \neq 2^k$, para todo $k \in \mathbb{Z}$, $k \geq 1$, então desde que $\mathbb{N}^* \subset \bigcup_{k=1}^{\infty} [2^{k-1}, 2^k)$, então, $\exists k \in \mathbb{N}^* / 2^{k-1} < n < 2^k$, onde \mathbb{N}^* é o conjunto dos naturais positivos.

Seja $m = 2^k - n$. Se $m = 1$ isso nos dá que $n = 2^k - 1$ e então, desde que $P(2^k)$ é verdadeira por (i) então, por (i i) $P(2^k - 1)$ também é verdadeiro, logo $P(n)$ é também verdadeira. Agora faremos indução sobre m , para isso, suponhamos que (2) seja verdadeira para $n = 2^k - m$ e provemos que também é verdadeira para $n - 1 = 2^k - (m + 1)$, ou

seja, suponhamos que (2) seja verdadeira para m tal que $m = 2^k - n$ e provemos que também é verdadeira para $m + 1 = 2^k - n + 1$. Por (i i) temos que $P(n - 1) = P(2^k - m - 1)$ é verdadeira desde que $P(n) = P(2^k - m)$ é também verdadeira. Assim temos que $P(m)$ é verdadeira $\forall m$, desde que $P(n) = P(2^k - m)$.

Logo, para concluirmos a demonstração de (2) para todo $n \in \mathbb{N}^*$, basta que provemos que (i) e (i i) são verdadeiras. Demonstremos inicialmente (i). Para isso, faremos a demonstração por indução sobre k onde $n = 2^k$. Se $k = 1$, desde que $a_1 \geq 0$ e $a_2 \geq 0$ então (2) é exatamente a desigualdade (1) de 3.1. Agora, seja $r \in \mathbb{N}^*$. Suponhamos que $P(2^r)$ seja verdadeira e provemos que $P(2^{r+1})$ é também verdadeiro. Consideremos $n = 2^r$, então $2^{r+1} = 2n$.

Mostremos que se a_1, a_2, \dots, a_{2n} são números positivos então

$$\frac{a_1 + a_2 + \dots + a_{2n}}{2n} \geq \sqrt[2n]{a_1 \cdot a_2 \cdot \dots \cdot a_{2n}},$$

com igualdade se, e somente se, $a_1 = a_2 = \dots = a_{2n}$. Para isso, tomemos

$$b_1 = \frac{a_1 + a_2}{2}, b_2 = \frac{a_3 + a_4}{2}, \dots, b_n = \frac{a_{2n-1} + a_{2n}}{2}$$

Então

$$b_i = \frac{a_{2i-1} + a_{2i}}{2} \geq \sqrt{a_{2i-1} \cdot a_{2i}}$$

com igualdade se, e somente se, $a_{2i-1} = a_{2i}$, conforme foi demonstrado em 3.1.

Utilizando a hipótese de indução sobre r , temos:

$$\frac{a_1 + a_2 + \dots + a_{2n}}{2n} = \frac{\frac{a_1 + a_2 + \dots + a_{2n}}{2}}{\sqrt[n]{b_1 \cdot b_2 \cdot \dots \cdot b_n}} = \frac{b_1 + b_2 + \dots + b_n}{n} \geq$$

onde a igualdade vale se, e somente se, os b_i s forem iguais. Mas, por 3.1, novamente, temos

$$\sqrt[n]{b_1 \cdot b_2 \cdot \dots \cdot b_n} = \sqrt[n]{\left(\frac{a_1 + a_2}{2}\right) \cdot \left(\frac{a_3 + a_4}{2}\right) \cdot \dots \cdot \left(\frac{a_{2n-1} + a_{2n}}{2}\right)} =$$

$$\begin{aligned}
&= \sqrt[n]{\frac{a_1 + a_2}{2}} \cdot \sqrt[n]{\frac{a_3 + a_4}{2}} \cdot \dots \cdot \sqrt[n]{\frac{a_{2n-1} + a_{2n}}{2}} \geq \\
&\geq \sqrt[n]{\sqrt{a_1 \cdot a_2}} \cdot \sqrt[n]{\sqrt{a_3 \cdot a_4}} \cdot \dots \cdot \sqrt[n]{\sqrt{a_{2n-1} \cdot a_{2n}}} = \\
&= \sqrt[2^n]{a_1 \cdot a_2} \cdot \sqrt[2^n]{a_3 \cdot a_4} \cdot \dots \cdot \sqrt[2^n]{a_{2n-1} \cdot a_{2n}} = \sqrt[2^n]{a_1 \cdot a_2 \cdot \dots \cdot a_{2n}},
\end{aligned}$$

com igualdade se, e somente se, os a_i s forem iguais, o que é facilmente verificado. Assim, $P(2^k)$ é verdadeiro para todo k inteiro e positivo.

Agora, provemos (i i), isto é, se $n > 1$ e $P(n)$ for verdadeira, então $P(n - 1)$ será verdadeira. Para isso, tomemos

$$b_1 = a_1, b_2 = a_2, \dots, b_{n-1} = a_{n-1}, b_n = \frac{a_1 + a_2 + \dots + a_{n-1}}{n-1}$$

Então, por hipótese indutiva

$$\frac{b_1 + b_2 + \dots + b_n}{n} \geq \sqrt[n]{b_1 \cdot b_2 \cdot \dots \cdot b_n}$$

e por substituição, temos

$$\begin{aligned}
&\frac{1}{n} \left\{ a_1 + a_2 + \dots + a_{n-1} + \frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \right\} \geq \\
&\geq \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_{n-1} \cdot \left(\frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \right)}
\end{aligned}$$

Como todos os a_i s são não-negativos, a última desigualdade é equivalente a elevar à potência n -ésima, que resulta em:

$$\begin{aligned}
&\frac{1}{n^n} \left\{ a_1 + a_2 + \dots + a_{n-1} + \frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \right\}^n \geq \\
&\geq \left\{ a_1 \cdot a_2 \cdot \dots \cdot a_{n-1} \cdot \left(\frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \right) \right\} \Leftrightarrow \\
&\Leftrightarrow \frac{(a_1 + a_2 + \dots + a_{n-1})^n}{n^n} \left\{ 1 + \frac{1}{n-1} \right\}^n \geq \\
&\geq \left\{ a_1 \cdot a_2 \cdot \dots \cdot a_{n-1} \cdot \left(\frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \right) \right\} \Leftrightarrow \\
&\Leftrightarrow \frac{(a_1 + a_2 + \dots + a_{n-1})^{n-1}}{n^n} \left\{ \frac{n}{n-1} \right\}^n \geq \frac{a_1 \cdot a_2 \cdot \dots \cdot a_{n-1}}{n-1} \Leftrightarrow \\
&\Leftrightarrow \left(\frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \right)^{n-1} \geq a_1 \cdot a_2 \cdot \dots \cdot a_{n-1} \Leftrightarrow \\
&\Leftrightarrow \frac{a_1 + a_2 + \dots + a_{n-1}}{n-1} \geq \sqrt[n-1]{a_1 \cdot a_2 \cdot \dots \cdot a_{n-1}},
\end{aligned}$$

provando que $P(n - 1)$ é também verdadeira. ■

Qual é a média final mais vantajosa para o aluno?

3.3 Uma consequência de 3.2

Seja $r = \frac{m}{n}$ um número racional, com $m, n > 0$ e $m < n$ e sejam $x, y \in \mathbb{R}^+$. Então

$$(1) \quad r \cdot x + (1 - r) \cdot y \geq x^r \cdot y^{1-r}.$$

Demonstração. Basta fazermos em (2) de 3.2

$$a_1 = a_2 = \dots = a_m = x \text{ e } a_{m+1} = \dots = a_n = y.$$

Então, por (2) de 3.2 temos:

$$\frac{a_1 + a_2 + \dots + a_n}{n} \geq \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n}$$

ou seja,

$$\frac{m \cdot x + (n - m) \cdot y}{n} \geq (x^m \cdot y^{n-m})^{\frac{1}{n}}$$

o que implica:

$$r \cdot x + (1 - r) \cdot y \geq x^r \cdot y^{1-r} \blacksquare$$

4 Desigualdade de Cauchy-Schwarz

4.1 Identidade de Lagrange

Inicialmente vejamos a Identidade de Lagrange, que será usada na Desigualdade de Cauchy-Schwarz

Sejam a_1, a_2, \dots, a_n e b_1, b_2, \dots, b_n dois conjuntos de números reais. A identidade de Lagrange é expressa por:

$$(1) \quad \left(\sum_{j=1}^n a_j \cdot b_j \right)^2 = \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right) - \frac{1}{2} \left[\sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 \right]$$

Demonstração. Faremos a demonstração por indução sobre n . Para $n = 1$ a igualdade é trivial, pois:

$$(a_1 \cdot b_1)^2 = (a_1^2) \cdot (b_1^2) - \frac{1}{2} \left[(a_1 \cdot b_1 - a_1 \cdot b_1)^2 \right].$$

Claro que essa última igualdade é verdadeira. Suponhamos, agora que $P(n)$ seja verdadeira e provemos que $P(n + 1)$ é também verdadeira, onde $P(n)$ é a igualdade (1) para $n \geq 1$. Então, por hipótese indutiva, temos:

$$\left(\sum_{j=1}^n a_j \cdot b_j \right)^2 = \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right) - \frac{1}{2} \left[\sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 \right]$$

Provemos que:

$$(i) \quad \left(\sum_{j=1}^{n+1} a_j \cdot b_j \right)^2 = \left(\sum_{j=1}^{n+1} a_j^2 \right) \cdot \left(\sum_{k=1}^{n+1} b_k^2 \right) - \frac{1}{2} \left[\sum_{j,k=1}^{n+1} (a_j \cdot b_k - a_k \cdot b_j)^2 \right]$$

que é equivalente a:

$$\begin{aligned} \left(\sum_{j=1}^{n+1} a_j \cdot b_j \right)^2 &= \left(\sum_{j=1}^n a_j^2 + a_{n+1}^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 + b_{n+1}^2 \right) - \\ &- \frac{1}{2} \left[\sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 + \sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 \right] - \\ &- \frac{1}{2} \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 - \frac{1}{2} (a_{n+1} \cdot b_{n+1} - a_{n+1} \cdot b_{n+1})^2 = \\ &\sum_{j=1}^n a_j^2 \cdot \sum_{k=1}^n b_k^2 + \\ &+ b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 + a_{n+1}^2 \cdot \sum_{k=1}^n b_k^2 + a_{n+1}^2 \cdot b_{n+1}^2 - \frac{1}{2} \left[\sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 \right] - \\ &- \frac{1}{2} \left[\sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 + \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 \right] - \\ &- \frac{1}{2} (a_{n+1} \cdot b_{n+1} - a_{n+1} \cdot b_{n+1})^2 \end{aligned}$$

Iniciaremos agora a demonstração de (i). Para isso, observemos inicialmente o segundo membro da igualdade acima, que mostra-nos o que devemos obter desenvolvendo o primeiro membro de (i).

Dessa forma, desenvolvendo o primeiro membro de (i), temos, pela hipótese indutiva:

$$\begin{aligned} \left(\sum_{j=1}^{n+1} a_j \cdot b_j \right)^2 &= \left[\left(\sum_{j=1}^n a_j \cdot b_j \right) + a_{n+1} \cdot b_{n+1} \right]^2 = \\ &= \left(\sum_{j=1}^n a_j \cdot b_j \right)^2 + (a_{n+1} \cdot b_{n+1})^2 + 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1}) = \end{aligned}$$

$$\begin{aligned}
& \left(\sum_{j=1}^{n+1} a_j \cdot b_j \right)^2 = \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right) - \frac{1}{2} \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 + \\
& + (a_{n+1} \cdot b_{n+1})^2 + 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1}) = \\
& = \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right) - \frac{1}{2} \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 + \\
& + (a_{n+1} \cdot b_{n+1})^2 + 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1}) + b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 + a_{n+1}^2 \cdot \\
& \sum_{k=1}^n b_k^2 - \\
& - b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 - a_{n+1}^2 \cdot \sum_{k=1}^n b_k^2 - \frac{1}{2} \sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 - \\
& - \frac{1}{2} \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 + \\
& + \frac{1}{2} \sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 + \frac{1}{2} \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 = \\
& = \left(\sum_{j=1}^n a_j^2 + a_{n+1}^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 + b_{n+1}^2 \right) - b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 - a_{n+1}^2 \cdot \sum_{k=1}^n b_k^2 - \\
& - \frac{1}{2} \left[\sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 + \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 \right] - \\
& - \frac{1}{2} \sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 + \\
& + \frac{1}{2} \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 + \frac{1}{2} \sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 + \\
& + 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1}) = \\
& = \left(\sum_{j=1}^{n+1} a_j^2 \right) \cdot \left(\sum_{k=1}^{n+1} b_k^2 \right) - b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 - a_{n+1}^2 \cdot \sum_{k=1}^n b_k^2 - \\
& - \frac{1}{2} \sum_{j,k=1}^{n+1} (a_j \cdot b_k - a_k \cdot b_j)^2 + \\
& + \frac{1}{2} \sum_{k=1}^n (a_{n+1} \cdot b_k - a_k \cdot b_{n+1})^2 + \frac{1}{2} \sum_{j=1}^n (a_j \cdot b_{n+1} - a_{n+1} \cdot b_j)^2 + \\
& + 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1})
\end{aligned}$$

$$\begin{aligned}
&= \left(\sum_{j=1}^{n+1} a_j^2 \right) \cdot \left(\sum_{k=1}^{n+1} b_k^2 \right) - b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 - a_{n+1}^2 \cdot \sum_{k=1}^n b_k^2 - \\
&- \frac{1}{2} \sum_{j,k=1}^{n+1} (a_j \cdot b_k - a_k \cdot b_j)^2 + \\
&+ \frac{1}{2} \left[\sum_{k=1}^n (a_{n+1} \cdot b_k)^2 + (a_k \cdot b_{n+1})^2 + 2 \cdot (a_{n+1} \cdot b_k) \cdot (a_k \cdot b_{n+1}) \right] + \\
&+ \frac{1}{2} \left[\sum_{j=1}^n (a_j \cdot b_{n+1})^2 + (a_{n+1} \cdot b_j)^2 + 2 \cdot (a_j \cdot b_{n+1}) \cdot (a_{n+1} \cdot b_j) \right] + \\
&+ 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1}) = \\
&= \left(\sum_{j=1}^{n+1} a_j^2 \right) \cdot \left(\sum_{k=1}^{n+1} b_k^2 \right) - \frac{1}{2} \sum_{j,k=1}^{n+1} (a_j \cdot b_k - a_k \cdot b_j)^2 - b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 - a_{n+1}^2 \cdot \\
&\sum_{k=1}^n b_k^2 + \\
&+ \frac{1}{2} \cdot b_{n+1}^2 \cdot \sum_{j=1}^n a_j^2 + \frac{1}{2} \cdot a_{n+1}^2 \cdot \sum_{k=1}^n b_k^2 + \frac{1}{2} \cdot a_{n+1}^2 \cdot \sum_{j=1}^n b_j^2 - a_{n+1} \cdot b_{n+1} \cdot \sum_{j=1}^n a_j \cdot b_j + \\
&+ \frac{1}{2} \cdot b_{n+1}^2 \cdot \sum_{k=1}^n a_k^2 + (a_{n+1} \cdot b_{n+1}) \cdot \sum_{k=1}^n a_k \cdot b_k + 2 \left(\sum_{j=1}^n a_j \cdot b_j \right) (a_{n+1} \cdot b_{n+1}) = \\
&= \left(\sum_{j=1}^{n+1} a_j^2 \right) \cdot \left(\sum_{k=1}^{n+1} b_k^2 \right) - \frac{1}{2} \sum_{j,k=1}^{n+1} (a_j \cdot b_k - a_k \cdot b_j)^2, \blacksquare
\end{aligned}$$

A seguir veremos a desigualdade de Cauchy-Swartz de grande importância nas aplicações.

4.2 Desigualdade de Cauchy-Schwarz

Consideremos a_1, a_2, \dots, a_n e b_1, b_2, \dots, b_n dois conjuntos de números reais. Vale a seguinte desigualdade, chamada Desigualdade de Cauchy-Schwarz

$$(1) \quad \left(\sum_{j=1}^n a_j \cdot b_j \right)^2 \leq \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right)$$

Demonstração. Pela Identidade de Lagrange, temos:

$$\begin{aligned}
&\left(\sum_{j=1}^n a_j \cdot b_j \right)^2 = \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right) - \frac{1}{2} \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 \\
&\text{como} \\
&-\frac{1}{2} \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 \leq 0 \\
&\text{então,}
\end{aligned}$$

$$\left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right) - \frac{1}{2} \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 \leq \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right),$$

ou seja,

$$\left(\sum_{j=1}^n a_j \cdot b_j \right)^2 \leq \left(\sum_{j=1}^n a_j^2 \right) \cdot \left(\sum_{k=1}^n b_k^2 \right)$$

Observando novamente a Identidade de Lagrange, vemos que uma condição necessária e suficiente a fim de que a Desigualdade de Cauchy-Schwarz se reduza a uma igualdade é:

$$\begin{aligned} -\frac{1}{2} \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 = 0 &\Leftrightarrow \sum_{j,k=1}^n (a_j \cdot b_k - a_k \cdot b_j)^2 = 0 \Leftrightarrow \\ &\Leftrightarrow a_j \cdot b_k - a_k \cdot b_j = 0, \forall j, k = 1, 2, \dots, n \Leftrightarrow a_j \cdot b_k = a_k \cdot b_j, \\ &\quad \forall j, k = 1, 2, \dots, n \Leftrightarrow \\ &\Leftrightarrow \frac{a_j}{b_j} = \frac{a_k}{b_k} = \lambda, \lambda \in \mathbb{R}, \forall j, k = 1, 2, \dots, n, \text{ com } b_j \neq 0 \text{ e } b_k \neq 0, \end{aligned}$$

ou seja, que o quociente $\frac{a_j}{b_j} = \lambda$ seja constante $\forall j = 1, 2, \dots, n$, com $b_j \neq 0$,

onde, segue-se que

$$a_1 = \lambda b_1, a_2 = \lambda b_2, \dots, a_n = \lambda b_n$$

Portanto, temos que uma condição necessária e suficiente para que a desigualdade de Cauchy-Swartz se reduza a uma igualdade é que

$$(a_1, a_2, \dots, a_n) = \lambda (b_1, b_2, \dots, b_n),$$

ou seja, que as duas sequências de números reais sejam proporcionais ■.

5 Desigualdade de Young

A seguir veremos a desigualdade de Young sob vários aspectos, o geométrico e o analítico: sejam p e q elementos de $(1, \infty)$ tais que

$$\frac{1}{p} + \frac{1}{q} = 1,$$

e sejam $a, b \in \mathbb{R}^+$, então vale a desigualdade :

$$(1) \quad a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q}$$

denominada Desigualdade de Young.

Demonstração.

(i) Inicialmente vejamos sob o aspecto geométrico. Para isso, consideremos a curva no plano dada por

$$y = f(x) = x^{p-1}, x > 0$$

Isso nos dá que $x = y^{\frac{1}{p-1}} = y^{q-1}$. Claro que essa curva é estritamente crescente, convexa para cima se $p \geq 2$ e côncava para baixo se $1 < p < 2$. Na fig.(2) acima, representamos o caso em que $y = x^{p-1}, p > 2$. Tomando como exemplo $p = 3$, então temos $y = x^2$, tal que $x \in \mathbb{R}^+$.

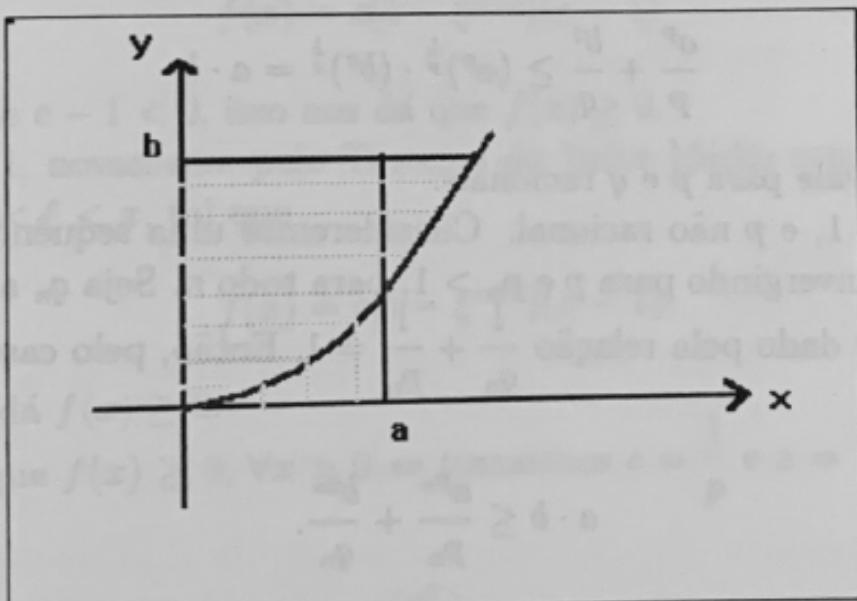


Figura 2

Assim, a área do retângulo de lados cujos comprimento são a e b é menor ou igual à soma das áreas das figuras hachuradas na vertical e na horizontal, isto é,

$$a \cdot b \leq \int_0^a x^{p-1} dx + \int_0^b y^{q-1} dy$$

ou seja,

$$a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q},$$

que é a desigualdade (1).

(i i) Vejamos uma demonstração da desigualdade de Young usando a desigualdade dada em 3.2. que é uma consequência da desigualdade entre as médias geométrica e aritmética dada.

Demonstraremos inicialmente essa desigualdade, para o caso em que p e q sejam racionais. Por (1) de (2.3) fazendo

$$x = a^p, y = b^q \text{ e } r = \frac{1}{p},$$

temos que

$$\frac{a^p}{p} + \frac{b^q}{q} \geq (a^p)^{\frac{1}{p}} \cdot (b^q)^{\frac{1}{q}} = a \cdot b$$

Assim, (1) vale para p e q racionais.

Seja $p > 1$, e p não racional. Consideremos uma sequência (p_n) de racionais convergindo para p e $p_n > 1$, para todo n . Seja q_n a sequência de racionais dado pela relação $\frac{1}{q_n} + \frac{1}{p_n} = 1$. Então, pelo caso anterior, temos que:

$$a \cdot b \leq \frac{a^{p_n}}{p_n} + \frac{b^{q_n}}{q_n}.$$

Desde que $\lim_{n \rightarrow \infty} q_n = q$ e usando a continuidade da função exponencial, temos que,

$$a \cdot b \leq \lim_{n \rightarrow \infty} \left(\frac{a^{p_n}}{p_n} + \frac{b^{q_n}}{q_n} \right) = \frac{a^p}{p} + \frac{b^q}{q}$$

o que demonstra (1).

(i i i) Agora, vejamos uma demonstração analítica. Consideremos a função

$$f(x) = (1 - c) + c \cdot x - x^c, \text{ com } x > 0 \text{ e } 0 < c < 1.$$

Temos então que $f'(x) = c - c \cdot x^{c-1}$, onde por f' estamos representando a derivada de f . Provemos inicialmente que $f(x) \geq 0$. Se $0 < x < 1$, então pelo Teorema do Valor Médio, segue-se que existe ξ em $(0, 1)$ tal que

$$f(1) - f(x) = f'(\xi)(1 - x).$$

Como $f(1) = 0$ então

$$f(x) = f'(\xi)(x - 1) = c(1 - \xi^{c-1})(x - 1)$$

Daí, concluímos que

$$f(x) = c(1 - \xi^{c-1})(x - 1)$$

e desde que $c - 1 < 0$, isso nos dá que $f(x) \geq 0$.

Se $x \geq 1$, novamente pelo Teorema do Valor Médio concluímos que existe $\xi, 1 < \xi < x$, tal que

$$f(x) = c(1 - \xi^{c-1})(x - 1)$$

o que nos dá $f(x) \geq 0$.

Desde que $f(x) \geq 0, \forall x > 0$, se tomarmos $c = \frac{1}{p}$ e $x = \frac{a^p}{b^q}$, teremos que

$$f\left(\frac{a^p}{b^q}\right) \geq 0$$

o que equivale a

$$\left(1 - \frac{1}{p}\right) + \frac{1}{p} \left(\frac{a^p}{b^q}\right) - \left(\frac{a^p}{b^q}\right)^{\frac{1}{p}} \geq 0$$

e isso nos dá que

$$\frac{1}{q} + \frac{1}{p} \left(\frac{a^p}{b^q} \right) - \frac{a}{b^{\frac{q}{p}}} \geq 0$$

Se multiplicarmos por b^q ambos os lados da desigualdade anterior teremos:

$$\frac{a^p}{p} + \frac{b^q}{q} \geq a \cdot b, \blacksquare$$

É interessante observar que a desigualdade de Young vale num contexto mais geral, mas não trataremos disso no presente trabalho e para os interessados nessa parte indicamos [4], pg. 96, ex.1.1.

6 Desigualdades de Hölder e de Minkowski

A próxima desigualdade, chamada Desigualdade de Hölder aparece no estudo dos espaços de funções.

6.1 Desigualdade de Hölder.

Se $p > 1$, q tal que $\frac{1}{p} + \frac{1}{q} = 1$ e se $a_1, a_2, \dots, a_n; b_1, b_2, \dots, b_n$ são números reais não-negativos, então

$$(1) \quad (a_1^p + a_2^p + \dots + a_n^p)^{\frac{1}{p}} \cdot (b_1^q + b_2^q + \dots + b_n^q)^{\frac{1}{q}} \geq$$

$$a_1 \cdot b_1 + a_2 \cdot b_2 + \dots + a_n \cdot b_n$$

que é uma generalização da Desigualdade de Cauchy-Schwarz.

Demonstração. Na desigualdade de Young, tomemos

$$a = \frac{a_i}{(a_1^p + \dots + a_n^p)^{\frac{1}{p}}}, \quad b = \frac{b_i}{(b_1^q + \dots + b_n^q)^{\frac{1}{q}}}$$

com $i = 1, \dots, n$ sucessivamente.

Assim, temos que

$$\frac{1}{p} \left(\frac{a_i^p}{(a_1^p + \dots + a_n^p)} \right) + \frac{1}{q} \left(\frac{b_i^q}{(b_1^q + \dots + b_n^q)} \right) \geq$$

$$\left(\frac{a_i}{(a_1^p + \dots + a_n^p)^{\frac{1}{p}}} \right) \left(\frac{b_i}{(b_1^q + \dots + b_n^q)^{\frac{1}{q}}} \right)$$

Adicionando as desigualdades resultantes, segue-se que,

$$\frac{1}{p} \left(\frac{a_1^p + \dots + a_n^p}{a_1^p + \dots + a_n^p} \right) + \frac{1}{q} \left(\frac{b_1^q + \dots + b_n^q}{b_1^q + \dots + b_n^q} \right) \geq$$

$$\frac{a_1 \cdot b_1 + \dots + a_n \cdot b_n}{(a_1^p + \dots + a_n^p)^{\frac{1}{p}} \cdot (b_1^q + \dots + b_n^q)^{\frac{1}{q}}}.$$

A partir da desigualdade acima e como $\frac{1}{p} + \frac{1}{q} = 1$, chegamos à desigualdade desejada. ■

6.2 Desigualdade de Minkowski

Se $p > 1$ e $a_1, a_2, \dots, a_n; b_1, b_2, \dots, b_n$ são números reais não-negativos, então

$$(1) \quad (a_1^p + a_2^p + \dots + a_n^p)^{\frac{1}{p}} + (b_1^p + b_2^p + \dots + b_n^p)^{\frac{1}{p}} \geq [(a_1 + b_1)^p + (a_2 + b_2)^p + \dots + (a_n + b_n)^p]^{\frac{1}{p}}$$

Demonstração. Consideremos $S_n = \sum_{r=1}^n (a_r + b_r)^p$, então

$$S_n = \sum_{r=1}^n (a_r + b_r) \cdot (a_r + b_r)^{p-1} = \sum_{r=1}^n a_r \cdot (a_r + b_r)^{p-1} + \sum_{r=1}^n b_r \cdot (a_r + b_r)^{p-1}$$

Aplicando-se a Desigualdade de Hölder a cada um dos termos da direita, temos

$$S_n \leq \sum_{r=1}^n (a_r^p)^{\frac{1}{p}} \cdot \left[\sum_{r=1}^n (a_r + b_r)^{(p-1) \cdot q} \right]^{\frac{1}{q}} + \sum_{r=1}^n (b_r^p)^{\frac{1}{p}} \cdot \left[\sum_{r=1}^n (a_r + b_r)^{(p-1) \cdot q} \right]^{\frac{1}{q}}$$

e, como $(p-1) \cdot q = p$, temos:

$$S_n \leq \left[\sum_{r=1}^n (a_r^p)^{\frac{1}{p}} + \sum_{r=1}^n (b_r^p)^{\frac{1}{p}} \right] \cdot \left[\sum_{r=1}^n (a_r + b_r)^p \right]^{\frac{1}{q}}$$

ou seja,

$$S_n \leq \left[\sum_{r=1}^n (a_r^p)^{\frac{1}{p}} + \sum_{r=1}^n (b_r^p)^{\frac{1}{p}} \right] \cdot [S_n]^{\frac{1}{q}} \blacksquare$$

Essa desigualdade é a desigualdade triangular que aparece nos Espaços de Funções, por exemplo em $C[a, b]$ com a norma $\| \cdot \|_p$, isto é, Espaço das Funções Contínuas em $[a, b]$, normado por

$$\| f \|_p = \left(\int_a^b |f(x)|^p dx \right)^{\frac{1}{p}}.$$

7 Desigualdade de Bernoulli

A desigualdade que trataremos agora é a Desigualdade de Bernoulli. Esta desigualdade é de grande importância na matemática devido ao grande número de aplicações que ela possui. Uma de suas aplicações está na Análise Combinatória, como segue no exemplo: calcular aproximadamente $(1,003)^{20}$. Temos que:

$$(1,003)^{20} = (1 + 0,003)^{20} = \sum_{k=0}^{20} \binom{20}{k} (0,003)^k \simeq 1 + 20 \cdot 0,003 = 1,06$$

onde

$$\binom{n}{p} = \frac{n!}{(n-p)!p!}$$

pois $(0,003)^2, (0,003)^3$, etc, são números tão pequenos comparados com $\binom{20}{2}, \binom{20}{3}$, que todo o resto da soma pode ser desprezado. Em geral, se nx é um número muito próximo de 0, então

$$(1 + x)^n \simeq 1 + n \cdot x.$$

7.1 Desigualdade de Bernoulli

Para todo $x \in \mathbb{R}$ com $x \geq -1$ e $n \in \mathbb{N}$ com $n \geq 1$ temos que:

$$(1) \quad (1+x)^n \geq 1 + nx$$

Faremos a demonstração dessa desigualdade inicialmente por indução sobre n , para $x \geq -1$ e posteriormente através de resultados do Cálculo Diferencial. Neste último método é interessante observar que o domínio no qual a desigualdade é válida é bem mais amplo.

(i) Demonstração por indução: se $n = 1$, claro que a desigualdade (6) é verdadeira. Então, seja $n \geq 1$ e suponhamos que $P(n)$ é verdadeira e provemos que $P(n+1)$ é verdadeira, onde $P(n)$ é a desigualdade (1).

Por hipótese indutiva, segue-se que:

$$(1+x)^n \geq 1 + n \cdot x.$$

Mas, como $x \geq -1$, então $1+x \geq 0$. Dessa forma, multiplicando ambos os membros da desigualdade acima por $1+x$, e desde que $x \geq -1$ temos que

$$(1+x)^n(1+x) \geq (1+n \cdot x)(1+x).$$

Assim,

$$(1+x)^{n+1} \geq 1 + (n+1)x + n \cdot x^2.$$

Como $nx^2 \geq 0$, temos que

$$1 + (n+1)x + n \cdot x^2 \geq 1 + (n+1)x.$$

Logo

$$(1+x)^{n+1} \geq 1 + (n+1)x.$$

Portanto, $\forall x \geq -1$ e $n \geq 1$, temos que:

$$(1+x)^n \geq 1 + nx.$$

(i i) Para desenvolvermos a demonstração através de fatos do Cálculo Diferencial, faremos um estudo geométrico da função definida por $f(x) = g(x) - h(x)$, onde $g(x) = (1+x)^n$ e $h(x) = 1 + nx$.

Verificaremos para que valores de x , $(1+x)^n \geq 1 + n \cdot x$, isto é, $f(x) \geq 0$, para comparar com a desigualdade de Bernoulli, no que se refere ao domínio de variação de x .

Inicialmente observaremos o gráfico da função $f(x)$ para $n = 1, 2, 3, 4, 5$, e posteriormente estudaremos o gráfico da função $f(x)$ para n em geral.

- Se $n = 1$, $f(x) = (1+x) - (1+x) = 0$. Portanto $f(x) = 0, \forall x$.

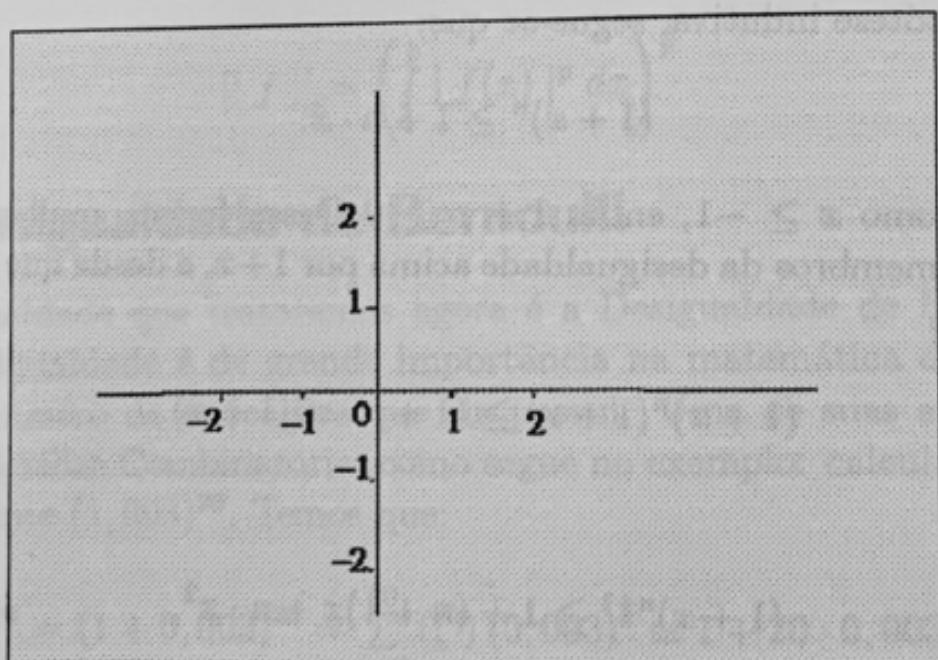


Figura 3

- Se $n = 2$, $f(x) = (1+x)^2 - (1+2x)$. Portanto $f(x) = x^2 \geq 0, \forall x$.

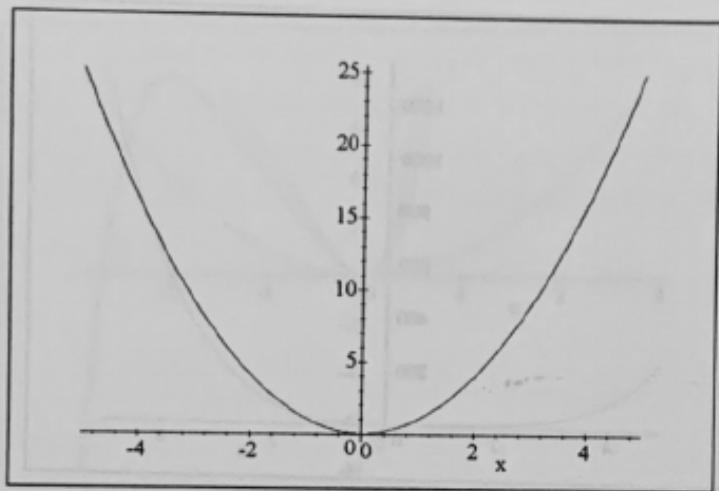


Figura 4

- Se $n = 3$, $f(x) = (1+x)^3 - (1+3x) = 1+x^3 + 3x^2 + 3x - 1 - 3x = x^3 + 3x^2$. Portanto $f(x) \geq 0$, para $x \geq x_0$, onde $x_0 < -2$.

Mas,

$$\bullet f''(x) = 0 \Leftrightarrow (1+x)^{n-1} - (1+nx)^{n-1} = 0 \Leftrightarrow (1+x)^{n-1} = (1+nx)^{n-1}$$

Se n é par, então $n-1$ é ímpar.

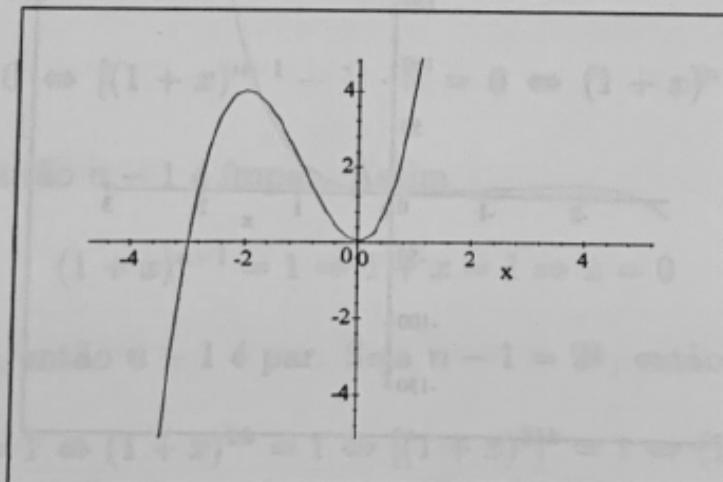


Figura 5

Analisando o crescimento da função, temos que

E assim por diante,

- Se $n = 4$, $f(x) = (1+x)^4 - (1+4x) = (x)$

Se n é ímpar, então $n-1$ é par, logo (x) é par ((x) é par).

para $x < -2 \Rightarrow f''(x) > 0$

para $-2 < x < 0 \Rightarrow f''(x) < 0$

para $x > 0 \Rightarrow f''(x) > 0$

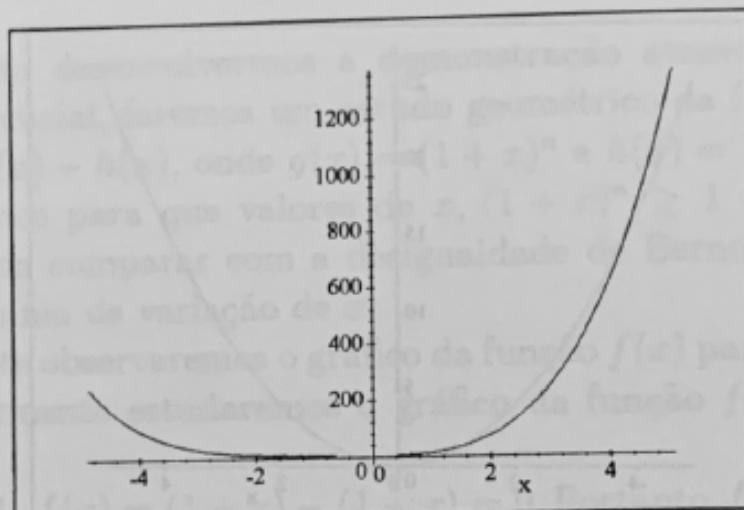


Figura 6

- Se $n = 5$, $f(x) = (1+x)^5 - (1+5x)$

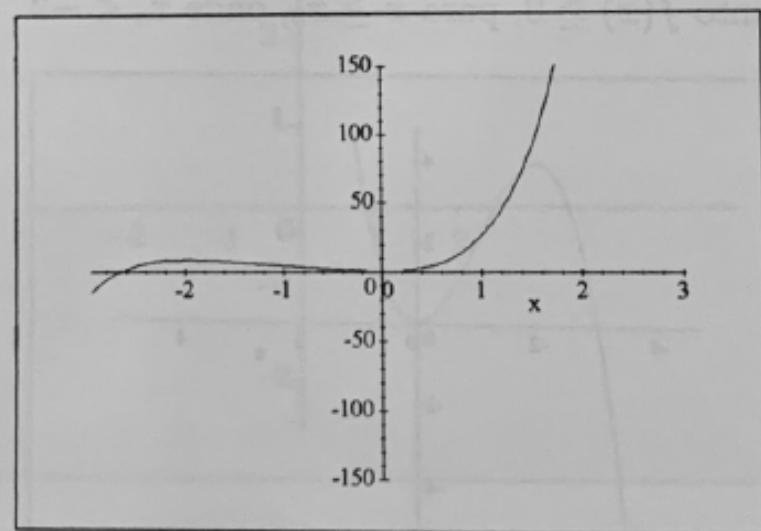


Figura 7

No caso geral, temos que:

$$f(x) = (1+x)^n - (1+nx)$$

$$f(x) = (1+x)^n - (1+nx), \quad 1 \leq n \leq 6.$$

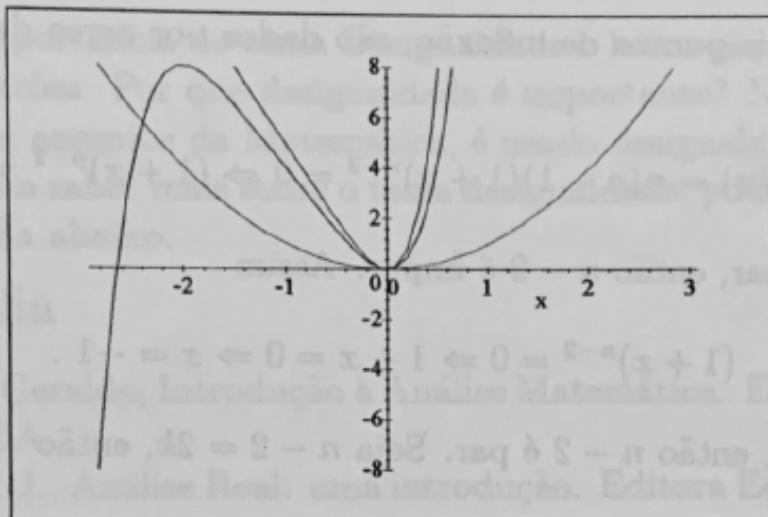


Figura 8

Os possíveis pontos críticos de $f(x)$ são valores de x que satisfazem a equação $f'(x) = 0$.

Mas,

$$f'(x) = 0 \Leftrightarrow [(1+x)^{n-1} - 1] \cdot n = 0 \Leftrightarrow (1+x)^{n-1} - 1 = 0 \Leftrightarrow (1+x)^{n-1} = 1$$

Se n é par, então $n-1$ é ímpar. Assim

$$(1+x)^{n-1} = 1 \Rightarrow 1+x = 1 \Rightarrow x = 0.$$

Se n é ímpar, então $n-1$ é par. Seja $n-1 = 2k$, então

$$(1+x)^{n-1} = 1 \Leftrightarrow (1+x)^{2k} = 1 \Leftrightarrow [(1+x)^2]^k = 1 \Leftrightarrow (1+x)^2 = 1 \Leftrightarrow |1+x| = 1 \Leftrightarrow \pm(1+x) = 1 \Leftrightarrow x = 0 \text{ ou } x = -2.$$

Analizando o crescimento da função, temos:

Se n é par, então $n-1$ é ímpar, portanto segue-se que:

para $x < 0 \Rightarrow f'(x) < 0$

para $x > 0 \Rightarrow f'(x) > 0$

Se n é ímpar, então $n-1$ é par, portanto segue-se que:

para $x < -2 \Rightarrow f'(x) > 0$

para $-2 < x < 0 \Rightarrow f'(x) < 0$

para $x > 0 \Rightarrow f'(x) > 0$

Os possíveis pontos de inflexão, são dados por zeros de $f''(x)$, isto é,

$$f''(x) = n(n-1)(1+x)^{n-2} = 0 \Rightarrow (1+x)^{n-2} = 0$$

Se n é ímpar, então $n-2$ é ímpar. Assim

$$(1+x)^{n-2} = 0 \Rightarrow 1+x = 0 \Rightarrow x = -1.$$

Se n é par, então $n-2$ é par. Seja $n-2 = 2k$, então

$$\begin{aligned} (1+x)^{n-2} = 0 &\Leftrightarrow (1+x)^{2k} = 0 \Leftrightarrow [(1+x)^2]^k = 0 \Leftrightarrow (1+x)^2 = 0 \Leftrightarrow \\ &\Leftrightarrow |1+x| = 0 \Leftrightarrow \pm(1+x) = 0 \Leftrightarrow x = -1. \end{aligned}$$

Calculando os valores que a função assume em $x=0$ e em $x=-2$, temos que:

$$f(0) = 1^n - 1 = 0, \text{ se } n \text{ é par e ímpar};$$

$$f(-2) = (-1)^n - (1-2n) = (-1)^n - 1 + 2n, \text{ se } n \text{ é ímpar } f(-2) = 2n-2 \text{ e se } n \text{ é par } f(-2) = 2n.$$

O comportamento da função $f(x)$ é descrito também por:

$$\begin{aligned} \lim_{\substack{x \rightarrow -\infty \\ n \text{ par}}} (1+x)^n - (1+nx) &= \infty \\ \lim_{\substack{x \rightarrow -\infty \\ n \text{ ímpar}}} (1+x)^n - (1+nx) &= -\infty \end{aligned}$$

Dessa forma, concluímos que:

- Se n é ímpar e $n \geq 3$, $f(-2) > 0$ e como

$$\lim_{\substack{x \rightarrow -\infty \\ n \text{ ímpar}}} (1+x)^n - (1+nx) = -\infty,$$

segue que a função $f(x)$ deve se anular em algum $x_o < -2$. Temos ainda que $x=0$ é ponto de mínimo, logo a função $f(x)$ é negativa para $x < x_o$ e positiva para $x > x_o$.

- Se n é par, temos que $f(x)$ é côncava para cima, $\forall x$. Segue-se, portanto que $f(x) \geq 0, \forall x$.

Visto a importância do tema Desigualdade o leitor já poderá refletir sobre as questões. Por que desigualdade é importante? No desenvolvimento de que assuntos da Matemática, é usado desigualdade? O leitor interessado em saber mais sobre o tema desigualdade, poderá utilizar-se da bibliografia abaixo.

Bibliografia

- [1] Ávila, Geraldo, Introdução à Análise Matemática. Editora Edgard Blücher LTDA.
- [2] White, A. J., Análise Real: uma introdução. Editora Edgard Blücher LTDA.
- [3] Figueiredo, Djairo Guedes, Análise I. Livros técnicos e científicos Editora S.A.,IMPA.
- [4] Hönig, Chaim Samuel, Análise Funcional e Aplicações. Publicações do Instituto de Matemática e Estatística da USP, 1970.
- [5] Bartle, Robert G., The Elements of Real Analysis. Copyright, 1964.

Teorema do Ponto Fixo de Brouwer no plano

Andréia Cristina Aravéquia*

Orientador : Prof. Dr. João Peres Vieira

1 Introdução

Este trabalho tem por objetivo demonstrar o Teorema do Ponto Fixo de Brouwer no Plano.

2 Preliminares

Para uma melhor compreensão da demonstração deste Teorema, o leitor deverá tomar conhecimento de alguns conceitos e resultados que relacionaremos abaixo.

Neste texto, denotaremos por D^2 o disco unitário no plano. Definimos a aplicação $e : \mathbb{R} \rightarrow S^1$, onde $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, por

$$e(t) = \exp(2\pi i t) = e^{2\pi i t}$$

Pode-se verificar facilmente que as propriedades usuais de função exponencial mostram que e é contínua e sobrejetiva e que é um homomorfismo do grupo aditivo \mathbb{R} para o grupo multiplicativo S^1 , ou seja, $e(t_1 + t_2) = e(t_1).e(t_2)$.

Seja X um espaço topológico e $f : X \rightarrow S^1$ uma aplicação contínua. Consideremos o diagrama:

*Bolsista Fapesp- PROC. 96/06156-5

Se existir $\tilde{f} : X \rightarrow \mathbb{R}$ tal que $f = e \circ \tilde{f}$, dizemos que \tilde{f} é um levantamento de f .

Um resultado importante no estudo mais detalhado deste assunto é que qualquer aplicação contínua $f : I \rightarrow S^1$ admite um levantamento $\tilde{f} : I \rightarrow \mathbb{R}$, que é único a menos de uma translação por um inteiro.

Consideremos agora $f : S^1 \rightarrow S^1$ e o diagrama abaixo

$$I \xrightarrow{g} \mathbb{R}$$

$$e|I \downarrow \quad \quad \quad \downarrow e$$

$$S^1 \xrightarrow{f} S^1$$

tal que $f \circ e|I = e \circ g$ (I)

Desse modo, g é um levantamento de $f \circ e|I$.

Temos $e(0) = e(1) = 1$. Substituindo em (I):

$$e \circ g(1) = f \circ e(1) = f \circ e(0) = e \circ g(0).$$

Logo,

$e(g(1)) = e(g(0))$ e portanto $1 = e(g(1)).(e(g(0)))^{-1} =$

$e(g(1)) - e(g(0))$ e daí $g(1) - g(0) \in \text{Ker}(e) = \mathbf{Z}$.

Assim, $g(1) - g(0)$ é um inteiro, chamado o **grau de f** e denotado por $\deg(f)$.

Dizemos que duas aplicações $f, g : Y \rightarrow X$ são aplicações homotópicas se existe uma aplicação contínua $F : Y \times I \rightarrow X$ tal que $\forall y \in Y$ temos

$$F(y, 0) = f(y)$$

$$F(y, 1) = g(y)$$

e F é chamada uma homotopia entre f e g .

Homotopia é uma relação de equivalência entre aplicações $Y \rightarrow X$. As classes de equivalência para a relação de homotopia são chamadas classes de homotopia.

Denotaremos por $[Y, X]$ o conjunto das classes de homotopia de aplicações $f : Y \rightarrow X$ e por $[f]$ a classe de equivalência da aplicação f .

Dizemos que uma aplicação $f : Y \rightarrow X$ é uma equivalência de homotopia, com homotopia inversa g , se

$$f \circ g \tilde{=} 1_X \quad e \quad g \circ f \tilde{=} 1_Y$$

Um exemplo importante de equivalência de homotopia ocorre quando X tem apenas um ponto, ou seja, quando $f(Y) = \{x\}$. Neste caso,

$$f \circ g \tilde{=} 1_X = x \quad e \quad g \circ f \tilde{=} 1_Y$$

Se Y é homotopicamente equivalente a um ponto dizemos que Y é **contrátil**.

Finalmente só nos resta explicitar que **aplicações homotópicas têm o mesmo grau**.

3 Teorema do Ponto Fixo de Brouwer no Plano

Agora estamos prontos para demonstrar o

Teorema do Ponto Fixo de Brouwer no Plano. Toda aplicação contínua $f : D^2 \rightarrow D^2$ tem ponto fixo.

Prova. Suponhamos que f não tenha ponto fixo.

Então $\forall x \in D^2$ temos $x \neq f(x)$.

Defina $\phi : D^2 \rightarrow S^1$ onde $\phi(x)$ é o ponto no qual o segmento que une x a $f(x)$ encontra S^1 .

Obviamente, ϕ depende continuamente de x e $\phi|S^1 = 1_{S^1}$.

Consideremos o diagrama

$$S^1 \xrightarrow{i} D^2 \xrightarrow{\phi} S^1$$

Como D^2 é contrátil temos que $\phi \equiv cte$ e então $\phi|S^1 = \phi \circ i \equiv cte \circ i = cte$, o que é um absurdo, pois $\deg(\phi|S^1) = 1$ e $\deg(cte) = 0$ e aplicações homotópicas têm o mesmo grau.
Então f tem ponto fixo.

Bibliografia

Wall,C.T.C.; **A Geometric Introduction To Topology** , Addison-Wesley, reading, Mass.; 1972.

Uma máquina de catástrofe Gravitacional

Luciana de Fátima Martins *

Orientadores : Profa. Dra. Solange Mancini

Prof. Dr. Marcelo José Saia

1 Introdução

Neste trabalho apresentamos uma discussão de um exemplo físico que nos fornece uma motivação intuitiva sobre o que vem a ser a Teoria das Catástrofes, focalizando o instrumental matemático nela utilizado e o contorno de outros aspectos sem, entretanto, entrar no terreno formal de detalhe técnico ou filosófico. Após a leitura do mesmo, ficará fácil de se convencer da necessidade de uma teoria matemática que explique este exemplo e de apreciar o valor da Teoria das Singularidades na explicação do mesmo.

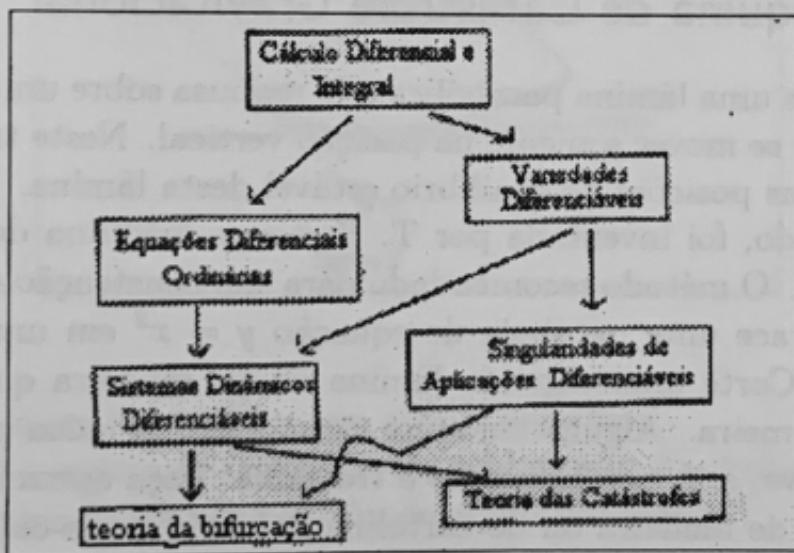
2 Sobre a Teoria das Catástrofes

A teoria das Catástrofes, idealizada pelo matemático francês René Thom, apresenta diversos aspectos compatíveis com as diferentes acepções que o termo teoria admite. Primeiramente, constitui teoria no sentido de um "grupo coerente de proposições gerais usadas como princípios para a explicação de uma determinada classe de fenômenos". Neste contexto, assim como a Teoria da Gravitação de Newton formula-se

*Bolsista PET-CAPES

para explicar os fenômenos de atração entre os corpos celestes e sua evolução, a Teoria das Catástrofes aborda a descrição, explicação e classificação dos fenômenos descontínuos ou catástrofes, na terminologia de Thom. Estas catástrofes não têm, necessariamente, a conotação de desastres, incluindo em seu estudo fenômenos tais como a ebullição da água, o rolar de uma pedra, a mudança de atitude de ataque para fuga nos animais sob a ação da raiva e do medo e muitos outros. No terreno puramente matemático, onde o termo teoria designa "um corpo de princípios e teoremas sobre um determinado objeto matemático", a teoria das Catástrofes tem como objeto preferencial de estudo as famílias de funções potenciais.

O quadro abaixo situa a Teoria das Catástrofes no conjunto das outras disciplinas matemáticas e focaliza suas diversas interrelações:

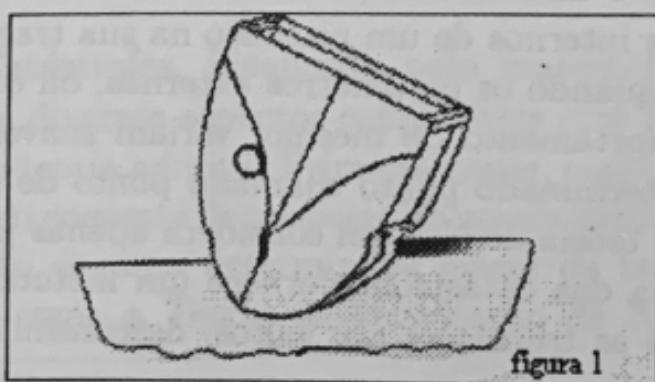


Uma catástrofe é uma mudança descontínua ou súbita experimentada pelos estados internos de um processo na sua transição entre dois regimes estáveis, quando os parâmetros externos, ou controles, que influenciam o comportamento do mesmo, variam suavemente na vizinhança de uma determinado ponto chamado ponto de catástrofe. Portanto, enquanto a teoria de Newton considera apenas processos suaves, contínuos, a teoria das catástrofes fornece um método universal para o estudo de todas as transições por saltos, descontinuidades e súbitas mudanças qualitativas.

Para entender o que vem a ser a teoria das catástrofes, é preciso em primeiro lugar se familiarizar com elementos da teoria das singularidades. Depois do trabalho fundamental de Whitney, a teoria das singularidades desenvolveu-se rapidamente e é agora uma das áreas centrais da matemática, onde se unem os ramos mais abstratos - geometria e topologia diferencial e algébrica, teoria das reflexões, grupos, álgebra comutativa, teoria do espaço complexo etc. - e os mais característicos de matemática aplicada - estabilidade de sistemas dinâmicos, bifurcação de estados de equilíbrio, óptica ondulatória e geométrica etc. René Thom e E.C.Zeeman sugeriram que a teoria das singularidades e suas aplicações se denominasse teoria das catástrofes.

3 A Máquina de Catástrofe Gravitacional

Consideremos uma lâmina parabólica que repousa sobre um plano horizontal e pode se mover somente na posição vertical. Neste trabalho iremos estudar as posições de equilíbrio estável desta lâmina. Na prática, para tal estudo, foi inventada por T. Poston a máquina de catástrofe gravitacional. O método recomendado para sua construção é o seguinte (figura 1): trace uma parábola de equação $y = x^2$ em uma cartolina e recorte-a. Corte uma segunda lâmina de tal maneira que seja congruente à primeira. Alguns cuidados devem ser tomados para fazer o contorno suave, sem achatamentos e tremidas. Faça aproximadamente seis suportes de madeira ou de cartolina dobrada e use-os para firmar as peças com suas fronteiras alinhadas. Coloque um pequeno ímã atrás de uma das lâminas e um pedaço de metal na frente.



O estudo é feito sob uma das lâminas da máquina. Consideremos então uma de suas lâminas e a coloquemos em repouso sobre um plano horizontal, deixando-a mover-se somente na posição vertical. Suponhamos seu centro de gravidade estar no ponto (a, b) (na máquina, o centro de gravidade será determinado pela posição do ímã). Ao soltarmos a lâmina, esta se movimenta até atingir uma posição de equilíbrio determinada pelo ponto (t, t^2) (figura 2).

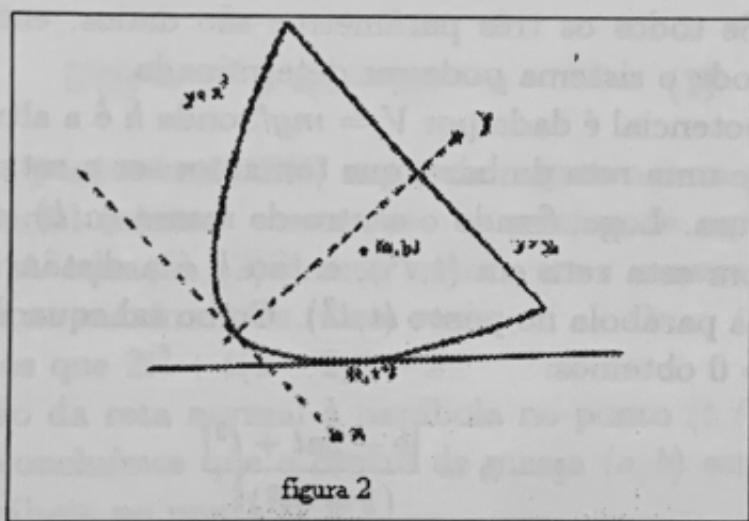


figura 2

Algumas questões que surgem são as seguintes:

- (a) Fixando o centro de massa (a, b) , qual o valor (ou valores) de t para o qual a lâmina ficará em posição de equilíbrio estável?
- (b) Como varia esta posição ao variarmos a posição do centro de massa da lâmina?

4 À procura do equilíbrio estável

Se uma partícula é deslocada um pouco de uma ponto de equilíbrio P , e tende a voltar a este ponto, então P é chamado de ponto de estabilidade ou ponto estável e o equilíbrio é dito ser estável. Uma condição necessária e suficiente para que um ponto seja de equilíbrio estável é que a energia potencial V seja mínima neste ponto. (Ver [6])

para maiores detalhes). Assim, para responder à questão (a) do item anterior, i.e., dado (a, b) encontrar o valor (ou valores) de t para o qual a lâmina ficará em posição de equilíbrio estável, devemos encontrar o valor, ou valores, de t para os quais $V(t)$ é mínimo.

O estudo da máquina de catástrofe é descrito por três parâmetros: a posição do centro de gravidade que é dada por duas coordenadas (chamadas de parâmetros externos ou de controle) e a posição da lâmina que é determinada por outro parâmetro, denominado parâmetro interno do sistema. Se todos os três parâmetros são dados, então a energia potencial de todo o sistema pode ser determinada.

A energia potencial é dada por $V = mgh$ onde h é a altura do centro de massa sobre uma reta da base, que tomamos ser a reta sobre a qual a lâmina repousa. Logo, fixado o centro de massa (a, b) , se a parábola faz contato com esta reta em (t, t^2) , então h é a distância de (a, b) à reta tangente à parábola no ponto (t, t^2) . Como tal equação é dada por $y - 2tx + t^2 = 0$ obtemos:

$$1 - h = \frac{|b - 2at + t^2|}{(1 + 4t^2)^{\frac{1}{2}}}$$

Naturalmente, para a lâmina física, o centro de massa estará em algum lugar da própria lâmina, i.e., (a, b) satisfará $b \geq a^2$. Assim,

$$b - 2at + t^2 \geq a^2 - 2at + t^2 = (a - t)^2 \geq 0$$

e, consequentemente,

$$V_{ab}(t) = mg \frac{b - 2at + t^2}{(1 + 4t^2)^{\frac{1}{2}}} \quad (1).$$

Como (a, b) está fixado, daqui por diante escreveremos $V(t)$ ao invés de $V_{ab}(t)$.

Como t_0 é ponto de mínimo se, e somente se, $V'(t_0) = 0$ e $V''(t)$ passa de negativo para positivo, para t em uma vizinhança de t_0 , devemos primeiramente derivar (1) e depois, estudar o seu sinal.

Seja

$$\frac{V'(t)}{mg} = \frac{(-2at + 2t)(1 + 4t^2) - 4bt + 8at^2 - 4t^3}{(1 + 4t^2)^{\frac{3}{2}}}.$$

Observando que a expressão à esquerda da igualdade acima é sempre positiva então, para estudarmos o sinal de $V'(t)$ e encontrar os valores de t para os quais $V'(t)$ passa de negativo para positivo, basta estudar a expressão

$$U(t) = 2t^3 + t(1 - 2b) - a \quad (2)$$

Assim, nos pontos onde $U(t)$ muda de negativo para positivo temos energia potencial mínima e estes são os pontos de equilíbrio estável procurados. A equação (2) é uma cúbica em t e assim, deve ter ao menos uma e no máximo três raízes reais.

Observemos que $2t^3 + t(1 - 2b) - x$

é a equação da reta normal à parábola no ponto (t, t^2) . Disto e da equação (3) concluímos que o centro de massa (a, b) está sobre a reta normal à parábola no ponto (t, t^2) .

4.1 Geometria da máquina gravitacional

Observemos que $U(t) = 0$ depende de a e b dados. Quando a e b também variam nós obtemos em \mathbb{R}^3 com coordenadas (t, a, b) uma superfície, que denotaremos por M (chamada a superfície catástrofe de V). O que faremos agora é esboçar esta superfície para, fixado (a, b) , encontrar os pontos (t, a, b) da superfície (i.e., dados (a, b) , encontrar as imagens inversas do 0) e analisar quais são pontos de mínimo.

Seja $U : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}$ onde $U(t, a, b) = 2t^3 + t(1 - 2b) - a$ e seja $M = U^{-1}(0)$. Assim M é uma superfície no espaço tridimensional com coordenadas (t, a, b) e 0 é um valor regular de U desde que sua matriz Jacobiana tem posto máximo, quaisquer que sejam t, a e b .

O plano horizontal $t = t_0$ encontra M na reta $t = t_0 = 2t^3 + t(1 - 2b) - a$, que é a reta normal à parábola $b = a^2$ no ponto $t = t_0$, $a = t_0$ e $b = t_0^2$

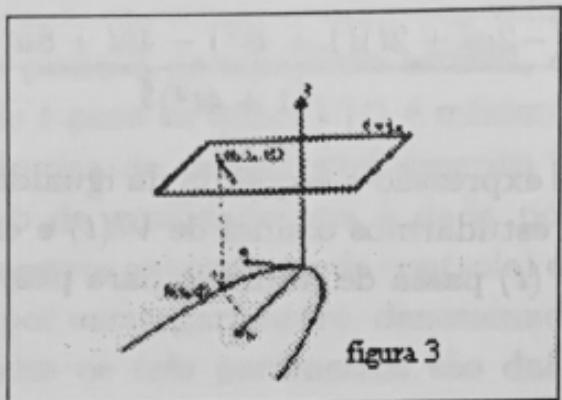


figura 3

Geometricamente, isto significa que M é obtida tomando as retas normais à parábola $b = a^2$ no plano (a, b) e movendo estas normais verticalmente até a altura t_0 . A figura 4(a) mostra as normais à parábola. Imagine estas normais se desdobrando até a altura t_0 , isto é mostrado na figura 4(b).

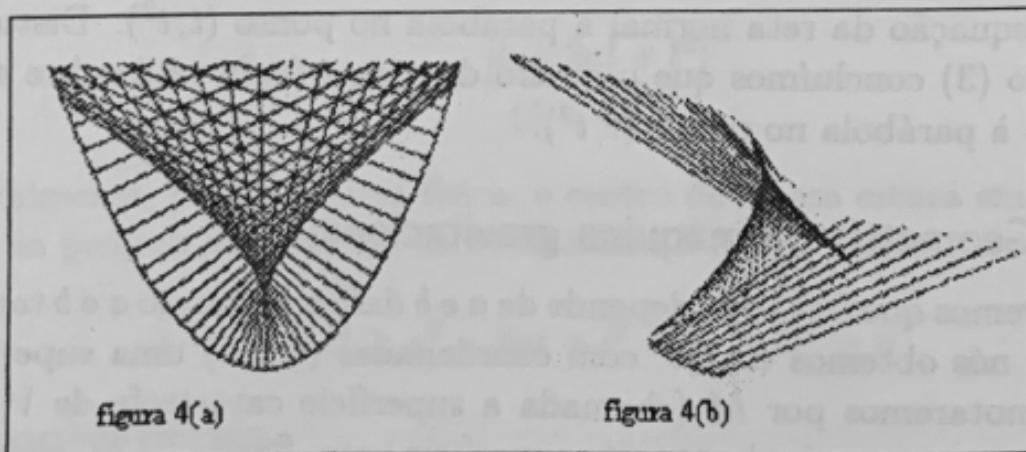


figura 4(a)

figura 4(b)

Continuando com o estudo da superfície M , tomamos o plano vertical $b = b_0$. O ponto (t, a, b_0) está em M se e somente se, $2t^3 + t(1 - 2b_0) - a = 0$, ou seja, $a = 2t^3 - t(1 - 2b_0)$.

Pensando em (t, a) como um sistema de coordenadas no plano $b = b_0$ verificamos que este plano encontra M na curva cúbica $2t^3 + t(1 - 2b_0) - a = 0$. A forma da cúbica depende do sinal de b_0 e a figura 5 mostra o esboço desta curva para alguns valores de b_0 . Note que quando $b_0 > \frac{1}{2}$ a cúbica adquire um máximo e um mínimo.

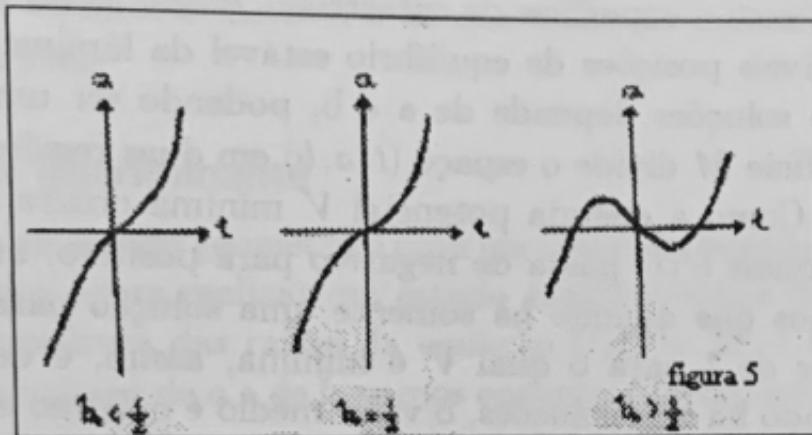


figura 5

Empilhando estas curvas para vários valores de b , (figura 6(a)) , fica claro que M tem a forma da superfície dobra (figura 6(b)).

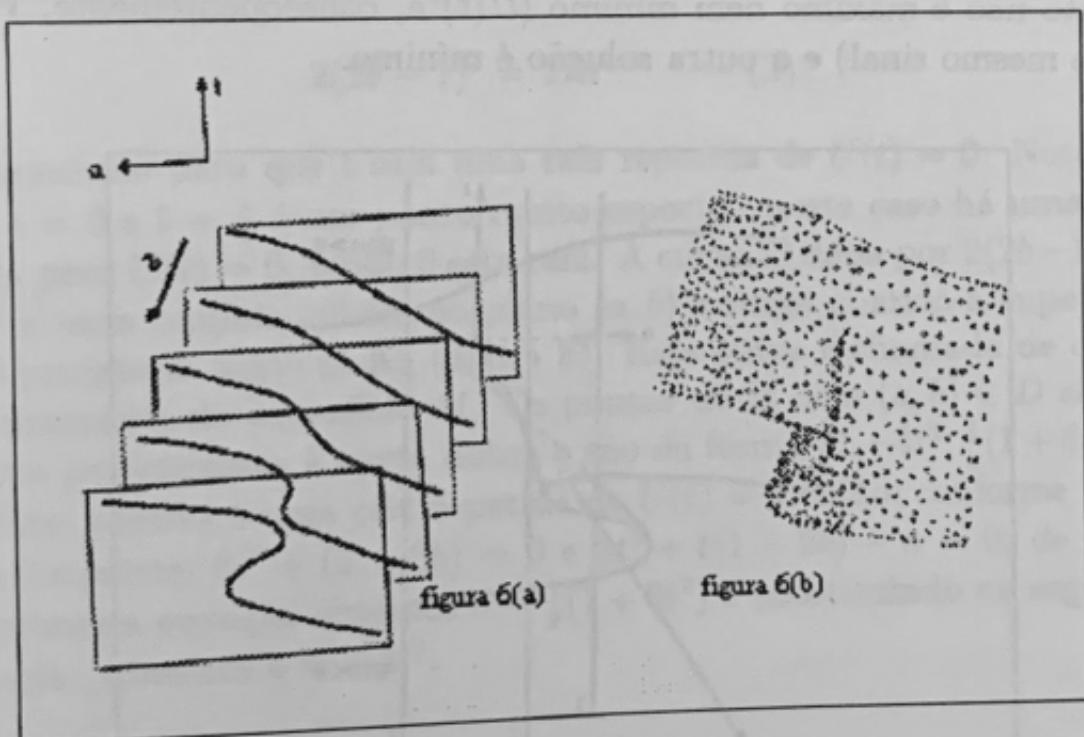
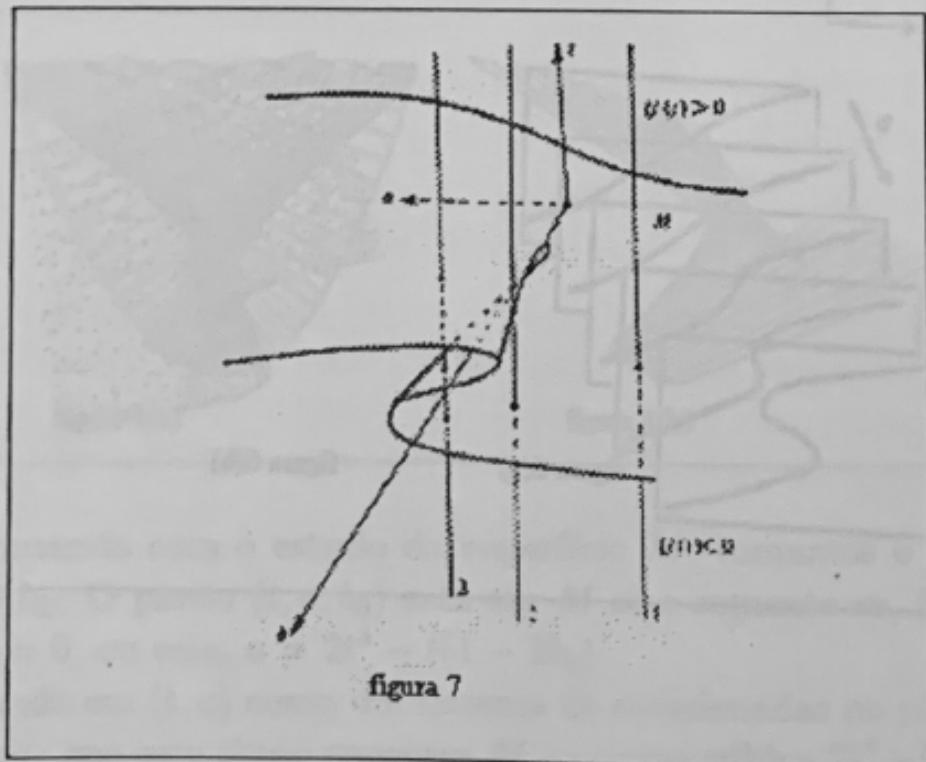


figura 6(a)

figura 6(b)

Estamos agora em condições de dar uma interpretação geométrica para a natureza do equilíbrio do sistema. Fixando um par de parâmetros (a, b) , os pontos de equilíbrio são obtidos pela resolução da equação $U(t) = 2t^3 + t(1 - 2b) - a = 0$. Eles podem, portanto, ser descritos como

os correspondentes t dos pontos de interseção da reta vertical sobre o ponto $(0, a, b)$ com a superfície de catástrofe. Assim, estes valores para t dão as possíveis posições de equilíbrio estável da lâmina parabólica. O número de soluções depende de a e b , podendo ser uma, duas ou três. A superfície M divide o espaço (t, a, b) em duas regiões: $U(t) > 0$ e $U(t) < 0$. Como a energia potencial V mínima ocorre nos valores de t para os quais $U(t)$ passa de negativo para positivo, analisando a figura 7, vemos que quando há somente uma solução para $U(t) = 0$ esta é o valor de t para o qual V é mínima, assim, é de equilíbrio estável. Quando há três soluções, o valor médio é máximo e, portanto, de equilíbrio instável e os outros dois são mínimos, logo, de equilíbrio estável. Quando há duas soluções, a reta vertical em um de seus pontos apenas toca M (M tem plano tangente vertical), sendo que o ponto de contato não é máximo nem mínimo ($U(t)$ e, consequentemente, $V'(t)$ tem o mesmo sinal) e a outra solução é mínimo.



A descrição do equilíbrio através da figura 7 é geometricamente

óbvia. Observemos que os pontos de M que estão dentro da dobra são todos de equilíbrio instável e os que estão do lado de fora são de equilíbrio estável.

4.2 Curva discriminante

Tendo feito um estudo geométrico para encontrar os pontos de equilíbrio estável, iremos agora realizar um estudo analítico sobre tais pontos.

Como a natureza das raízes da equação $U(t) = 2t^3 + t(1 - 2b) - a$ depende dos valores de a e de b , vamos encontrar o conjunto D no plano (a, b) , que consiste dos pontos onde há duas soluções para $U(t) = 0$, i.e., se $(a, b) \in D$, a equação $U(t) = 0$ tem uma raiz repetida.

A condição para t ser uma raiz repetida é: $U(t) = U'(t) = 0$. Um simples cálculo nos fornece que

$$2(2b - 1)^3 = 27a^2 \quad (3)$$

é a condição para que t seja uma raiz repetida de $U(t) = 0$. Notemos que $a = 0$ e $b = \frac{1}{2}$ é um ponto muito especial: neste caso há uma raiz tripla para $U(t) = 0$, sendo 0 esta raiz. A curva D dada por $2(2b - 1)^3 = 27a^2$ é uma cúspide cúbica no plano (a, b) e surge quando a superfície M é projetada neste plano (figura 8). Esta curva é chamada de curva discriminante da superfície M . Os pontos de M com $(a, b) \in D$ são os pontos pertencentes à curva dobra e são da forma $(t, -4t^3, \frac{1}{2}(1 + 6t^2))$. De fato, como t é uma raiz repetida de $U(t) = 0$ então, conforme visto anteriormente, $6t^2 + (1 - 2b) = 0$ e $2t^3 + t(1 - 2b) - a = 0$, de onde da primeira equação obtemos $b = \frac{1}{2}(1 + 6t^2)$ e substituindo na segunda equação, obtemos $a = -4t^3$.

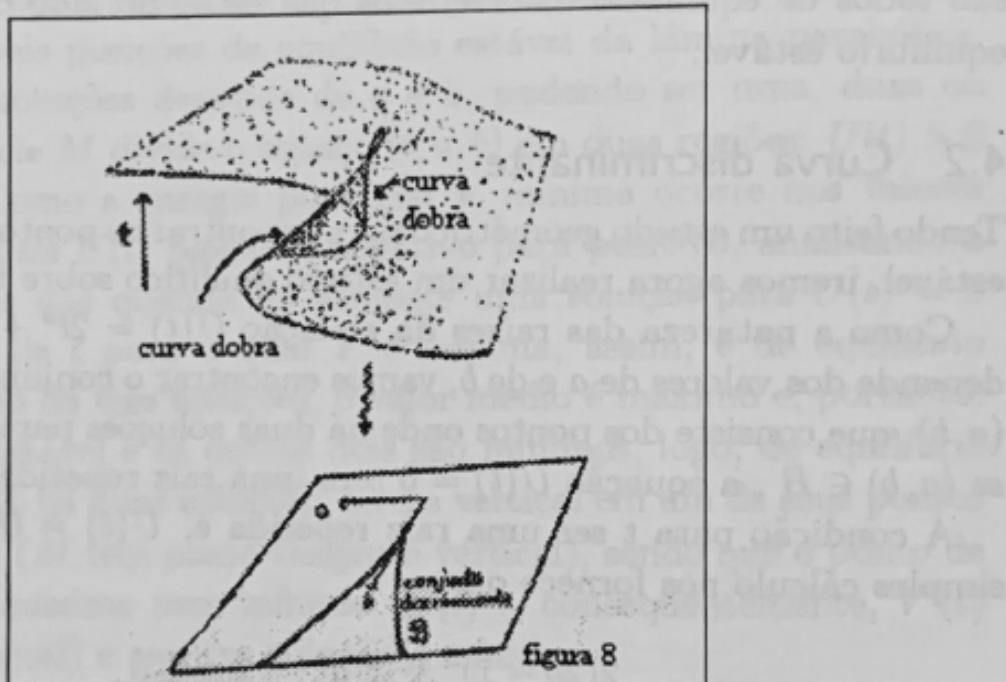


figura 8

Esta curva divide o plano (a, b) em duas regiões. Para os pontos da região menor, a equação (3) têm três raízes reais distintas, i.e., três pontos distintos da superfície M projetam-se em um único ponto do plano. Apelando novamente à parte geométrica, ou seja, esboçando V para (a, b) nesta região, podemos perceber que das três raízes, a raiz do meio é o máximo e as outras duas são os mínimos. Enquanto isso, para os pontos da região maior (fora da curva cuspide) a equação (2) tem somente uma raiz real, sendo esta o mínimo (a equação (2) tem uma raiz real e um par conjugado de raízes complexas). Para pontos sobre a curva há duas raízes (a equação (2) tem três raízes reais, porém duas delas coincidem), assim, uma é o mínimo e a outra é um ponto de inflexão. Quando nos movimentamos da região de dentro para fora da curva, duas das três raízes fundem-se ao atingirmos a curva e desaparecem em seguida quando penetrarmos na região de fora, onde para cada ponto (a, b) a equação (2) só tem uma raiz real. Se atingirmos a curva exatamente no ponto de cuspide, todas as três raízes coalescem de uma só vez, sendo este um ponto de mínimo. A figura 9(a) mostra o acima exposto e a figura 9(b) esboça V para várias posições de (a, b) .

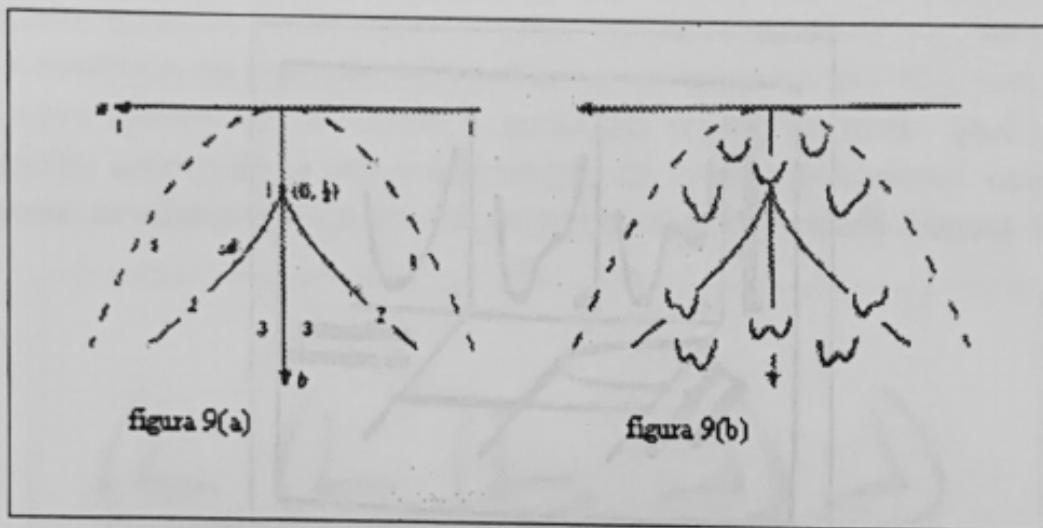


figura 9(a)

figura 9(b)

Observemos que o conjunto discriminante D foi obtido eliminando t em $U(t) = U'(t) = 0$, sendo $U(t) = 0$ a equação da normal à parábola $b = a^2$ em (t, t^2) . D é também chamado de envoltória das normais à parábola $b = a^2$. A envoltória das normais a uma curva é também chamada de evoluta da curva, assim D é a evoluta da parábola $b = a^2$.

5 Dinâmica da máquina gravitacional

Em parte já foi analisada nesta última seção a questão (b) feita na seção 3, ou seja, como varia a posição de equilíbrio estável quando variamos a posição do centro de massa da lâmina. Vimos o que acontece com as raízes da função $U(t)$ quando variamos (a, b) . É um interesse particular verificar o que acontece ao equilíbrio estável se movermos o centro de gravidade (a, b) ao longo do conjunto D .

Em síntese, sabemos que para um centro de massa fixado, a energia potencial é uma função da posição da lâmina. Essa função pode ter um, dois ou três mínimos, dependendo dos parâmetros de controle. Se ao variarmos os parâmetros de controle, a posição do mínimo variar suavemente, então não há salto. Este ocorrerá para valores dos parâmetros de controle nos quais um mínimo local desaparece por combinação com um máximo local; após o salto a lâmina atinge uma posição determinada por outro mínimo local.

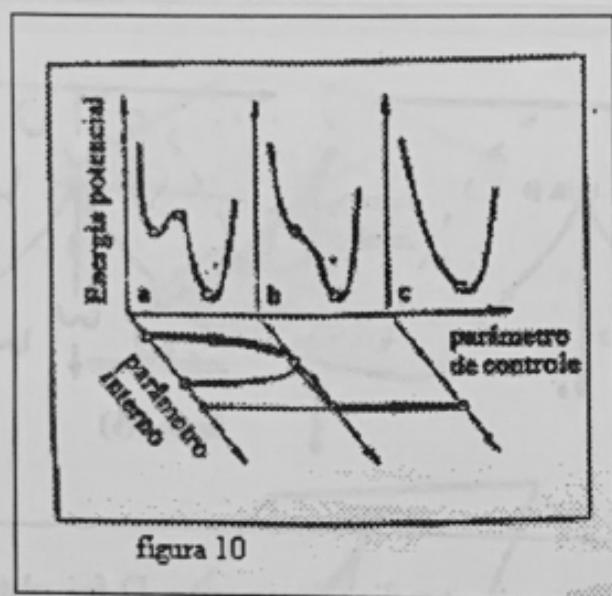


figura 10

~~é obviamente que~~

Suponhamos, por exemplo, que o ponto (a, b) move-se continuamente através do conjunto D ao longo da reta 12345, conforme figura 11.

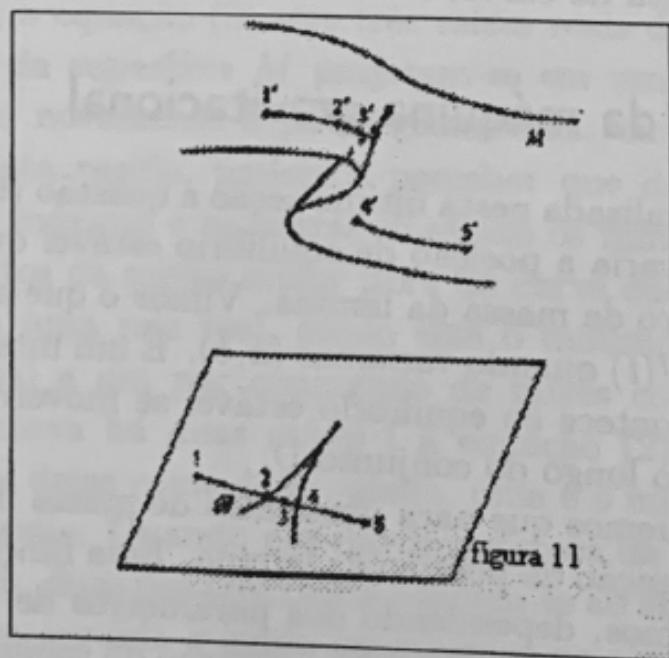


figura 11

Correspondendo a 1 e 2 há somente um ponto de equilíbrio estável, dado pelo ponto 1' e 2' de M , respectivamente. Em 3 há duas posições de equilíbrio estável, mas na prática, a parábola inclina-se suavemente

(vagarosamente), assim, o ponto $3'$ de M é o correspondente de 3. Contudo, quando (a, b) move-se para 4, não é possível que haja uma troca contínua na posição de equilíbrio: o caminho em M sobre $3'4'5'$ tem uma quebra em $4'$, onde a parábola repentinamente gira para a esquerda; uma grande troca na posição de equilíbrio ocorreu com uma pequena mudança em (a, b) , i.e., ocorreu uma catástrofe (figura 12).

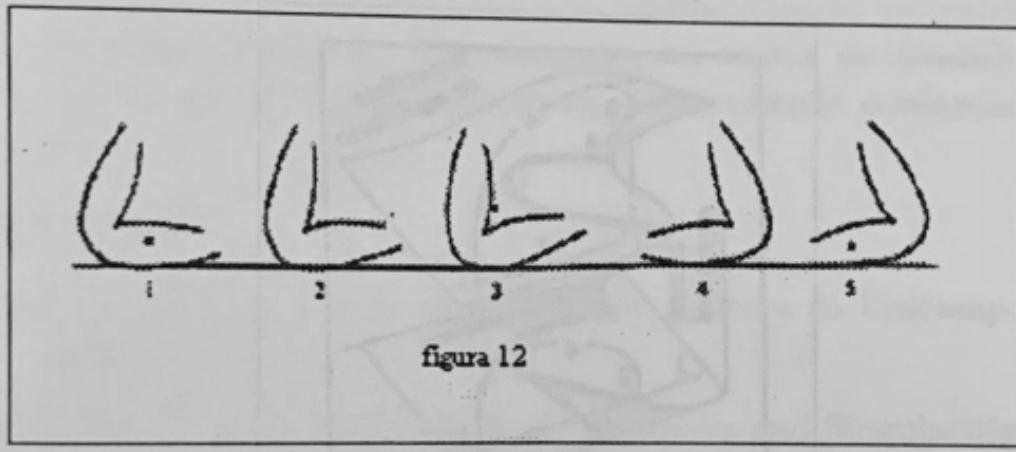


figura 12

Pode ser feita uma analogia do que está descrito acima da seguinte maneira: esboçando a função $V(t)$ para vários valores de (a, b) , imaginemos a curva ser de material flexível e o sistema sendo representado por uma bola; uma possível seqüência dos eventos descritos acima é ilustrada na figura 13.

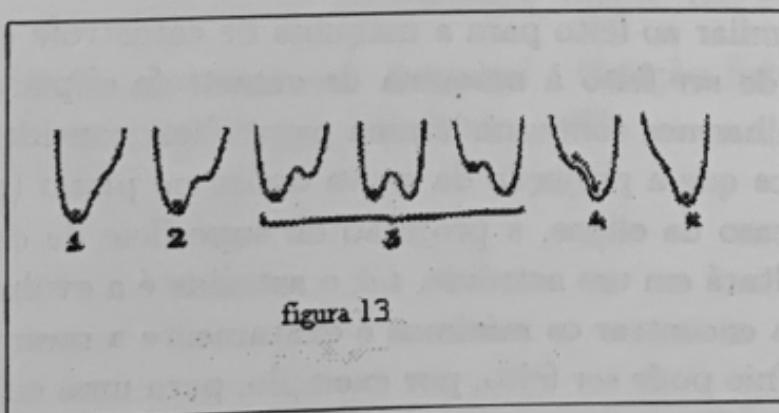


figura 13

A figura 14 mostra claramente por que uma transição dos parâmetros de controle de uma lado para o outro da curva cúspide algumas vezes provoca uma salto e outras não (isso dependerá de que ponto da superfície está em correspondência com o centro de massa). Com o uso do diagrama é possível deslocar-se de um ponto da superfície para outro sem que ocorram saltos.

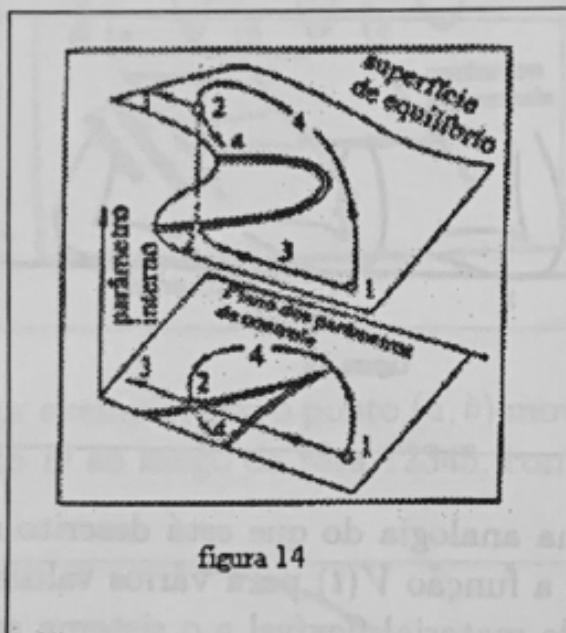


figura 14

6 Conclusão

Um cálculo similar ao feito para a máquina de catástrofe gravitacional parabólica pode ser feito à máquina de catástrofe elíptica, i.e., se ao invés de trabalharmos com uma lámina parabólica, considerarmos uma elíptica. Vimos que a projeção da curva dobra no plano (a, b) dá uma cúspide. No caso da elipse, a projeção da superfície de catástrofe em tal plano resultará em um astróide, i.e., o astróide é a evoluta da elipse. A análise para encontrar os mínimos é exatamente a mesma. Assim, o mesmo raciocínio pode ser feito, por exemplo, para uma circunferência, onde o conjunto discriminante (a evoluta) se reduz ao centro da circunferência.

A situação, na maioria das aplicações da teoria das catástrofes, é similar à do exemplo dado. A teoria das singularidades prediz a geometria das catástrofes, isto é, transições de um estado de equilíbrio para outro com a mudança dos parâmetros de controle. Na maioria das aplicações, a singularidade é uma cúspide de Whitney e os resultados já eram conhecidos antes do advento da teoria das catástrofes.

Aplicações do tipo descrito têm uma fundamentação tão sólida quanto suas premissas originais. Por exemplo, na teoria de estabilidade de embarcações, as previsões teóricas são inteiramente confirmadas pela experiência.

Bibliografia

- [1] ARNOLD, V. I. *Teoria da Catástrofe*. Editora da Unicamp. Campinas, 1989.
- [2] BRUCE, J. W. & GIBLIN, P. J., *it Curves and Singularities*. Cambridge University Press. Cambridge, 1984.
- [3] POSTON, T. & STEWART I., *Catastrophe Theory and its Applications*. Pitman. Boston, 1976.
- [4] SAUNDERS, P. T., *An Introduction to Catastrophe Theory*. Cambridge University Press. Cambridge, 1980.
- [5] SOTOMAYOR, J., *Singularidades de Aplicações Diferenciáveis*. III Escola Latino Americana de Matemática. IMPA. Rio Janeiro, 1976.
- [6] SPIEGEL, M.R. *Mecânica Racional*. Coleção Schaum. Editora McGraw-Hill do Brasil Ltda. São Paulo, 1973.