Curso de EDO

Augusto Armando de Castro Júnior (armando@impa.br)

06 de março de 2007

Conteúdo

0 Prolegômenos de Espaços Métricos e Análise **1**
0.1 Espaços métricos 1
0.2 O Teorema do Ponto Fixo para Contrações 4
0.3 Espaços de aplicações contínuas com domínio compacto 10
0.4 Integração de Caminhos em Espaços de Banach 19
0.5 Exercícios 22

1 O conceito de EDO **25**
1.1 O problema de Cauchy 28
1.2 Problemas de Contorno 29
1.3 Alguns métodos de Solução de Equações na Reta 30
1.4 Exercícios 34

2 Teoremas de existência e unicidade de soluções **36**
2.1 O Teorema de Picard 37
2.2 O Teorema de Peano 41
2.3 Intervalo maximal 43
2.4 Exercícios 48

3 Dependência das soluções em relação às condições iniciais e parâmetros. **52**
3.1 Dependência contínua 52
3.2 Dependência diferenciável 57
3.3 Exercícios 59

4 Campos de Vetores **61**
4.1 Equivalência e conjugação de campos vetoriais 67
4.2 O Teorema do Fluxo Tubular 70

4.3 Exercícios 77

5 Os conjuntos de $\omega-$limite e $\alpha-$limite 80
5.1 O teorema de Poincaré-Bendixson 83
5.2 Exercícios 91

6 Equações lineares 92
6.1 Caracterização das soluções 92
6.2 Campos lineares a coeficientes constantes 99
6.3 C^1-Conjugação de campos lineares a coeficientes constantes . . 102
6.4 Revisão de álgebra linear 104
6.5 Aplicações da Forma de Jordan 118
6.5.1 Classificação dos campos lineares hiperbólicos 121
6.5.2 Equações lineares de ordem superior na Reta 134
6.6 Exercícios 137

7 Noções Básicas de Teoria Espectral 138
7.1 Funções de um Operador 148
7.2 O Operador Adjunto e seu espectro 154
7.3 Operadores Compactos e Problemas de Contorno 165

8 O Teorema de Grobman-Hartman 166
8.1 O teorema de Grobman-Hartman para difeomorfismos 168
8.2 O teorema de Grobman-Hartman para campos 174
8.3 Apêndice: Classificação dos isomorfismos hiperbólicos 179

9 O Teorema da Variedade Estável 192
9.1 O Teorema da Variedade Estável para difeomorfismos 193
9.2 O Teorema da Variedade estável para campos 207

Capítulo 0

Prolegômenos de Espaços Métricos e Análise

Neste capítulo, relembramos alguns conceitos básicos em Topologia e demonstramos resultados que serão úteis desde o início do curso, mas que nem sempre recebem a ênfase devida nos cursos de Análise, dada a carga de assuntos que já mune tais cursos. Assim, enunciamos e provamos o Teorema do ponto fixo para Contrações, o Teorema de Aproximação de Weierstrass e o Teorema de Ascoli-Arzelá.

0.1 Espaços métricos

Definição 0.1.1. (Métrica e espaço métrico). Uma *métrica* em um conjunto Y é uma função $d : Y \times Y \to [0, +\infty)$ tal que, dados quaisquer $x, y, z \in Y$, valem:

d1) $d(x, y) = 0 \Leftrightarrow x = y$.

d2) $d(x, y) = d(y, x)$.

d3) $d(x, z) \leq d(x, y) + d(y, z)$ (desigualdade triangular).

O par ordenado (Y, d) é chamado de *espaço métrico*. Em geral, por um abuso de linguagem, diz-se que Y é um espaço métrico, subentendendo-se uma métrica d a ele associada.

Definição 0.1.2. (Bola aberta e conjuntos abertos de um espaço métrico). Seja (X, d) um espaço métrico. Dado $x \in X$ e $r \in \mathbb{R}^+$ qualquer definimos a *bola aberta* centrada em x e raio r como o conjunto

$$B(x, r) := \{y \in X; d(x, y) < r\}.$$

Dizemos que $A \subset X$ é um *conjunto aberto* de X se A pode ser escrito como união qualquer (inclusive não enumerável) de bolas abertas de X. Dizemos que um conjunto $F \subset X$ é *fechado* em X se $F^c := X \setminus F$ é aberto.

Observação 0.1.3. Lembramos que a coleção $\mathcal{T}$ acima definida dos abertos de um espaço métrico (X, d) possui as seguintes propriedades:

1. X e $\emptyset$ pertencem a $\mathcal{T}$.
2. $\mathcal{T}$ é fechada para uniões arbitrárias de seus elementos. Em outras palavras, dada uma família arbitrária (possivelmente não enumerável) $(A_\tau)_{\tau \in \Upsilon}$ de abertos $A_\tau \in \mathcal{T}$, a união $\cup_{\tau \in \Upsilon} A_\tau$ também é um conjunto aberto.
3. $\mathcal{T}$ é fechada para intersecções finitas de seu elementos. Isto é, dados abertos $A_1, \ldots, A_q \in \mathcal{T}$, $q \in \mathbb{N}$, a intersecção $\cap_{j=1}^{q} A_j$ também é um conjunto aberto.

As três propriedades acima fazem de $\mathcal{T}$ uma *topologia* , e do par $(X, \mathcal{T})$ um exemplo de *espaço topológico*. Embora não nos alonguemos sobre isso no presente texto, em algumas proposições lançaremos mão destas propriedades da coleção dos abertos de X.

Definição 0.1.4. (Norma). Seja $(E, +, ., \mathbb{R})$ um espaço vetorial real. Uma norma em E é uma aplicação $\| \cdot \| : E \to [0, +\infty)$ com as seguintes propriedades:

1. $\|v\| = 0 \Leftrightarrow v = 0$;
2. $\|\lambda v\| = |\lambda| \cdot \|v\|$; $\forall \lambda \in \mathbb{R}$, $\forall v \in E$.
3. $\|v + w\| \leq \|v\| + \|w\|; \forall v, w \in E$ (desigualdade triangular).

O exemplo mais comum de espaço métrico é dado pelos espaços vetoriais normados. Se E é um tal espaço, dotado de uma norma $\|\cdot\|$, então a aplicação $d : E \times E \to [0, +\infty)$ dada por

$$d(v, w) := \|v - w\|, \forall v \in E, w \in E;$$

define uma distância em E.

Outra classe importante de exemplos de espaços métricos é dada quando tomamos um subconjunto $Y \subset X$ de um espaço métrico (X, d). Nesse caso, a restrição $d|_{Y \times Y}$ define uma métrica em Y.

Definição 0.1.5. (Sequência e subsequência). Seja X um conjunto qualquer. Uma *sequência* em X é uma aplicação $x : \mathbb{N} \to X$. Denota-se $x_j := x(j)$ e $(x_j) := x$. Dada uma sequência $(x_j) : \mathbb{N} \to X$, uma *subsequência* (x_{j_k}) de (x_j) é qualquer restrição de (x_j) a um subconjunto infinito $\hat{\mathbb{N}} \subset \mathbb{N}$, $\hat{\mathbb{N}} = \{j_1, j_2, \ldots, \text{ com } j_1 < j_2 < \ldots\}$.

Definição 0.1.6. (Sequência convergente). Uma sequência (x_j) em um espaço métrico (Y, d) é dita *convergente* para $y \in Y$ se para toda bola aberta B tal que $y \in B$, tem-se um número finito de índices j tais que $x_j \notin B$. Em outras palavras, dado uma bola aberta $B \subset Y$ com $y \in B$, existe j_B tal que $x_j \in B, \forall n > n_B$. Escrevemos $x_j \to y$ quando $j \to +\infty$, ou simplesmente $x_j \to y$ para denotar que a sequência (x_j) converge a $y \in Y$. Dizemos que uma *subsequência* (x_{j_k}) é *convergente* se a sequência $(y_k) : \mathbb{N} \to Y$ definida por $y_k := x_{j_k}, \forall k \in \mathbb{N}$ for convergente.

Definição 0.1.7. (Sequência de Cauchy). Seja (Y, d) um espaço métrico. Uma sequência (y_n), com $y_n \in Y, \forall n \in \mathbb{N}$ é dita *sequência de Cauchy* se dado um real $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para todos $m, j \in \mathbb{N}$, com $m \geq n_0$ e $j \geq n_0$ temos $d(y_m, y_j) \leq \epsilon$.

Intuitivamente, dizer que (y_n) é uma sequência de Cauchy significa dizer que seus termos vão ficando mais e mais próximos para índices n suficientemente grandes.

Definição 0.1.8. (Aplicação contínua). Sejam $(X, d), (\tilde{X}, \tilde{d})$ espaços métricos. Uma aplicação $f : X \to \tilde{X}$ é dita *contínua* em um ponto $x \in X$ se dado $\epsilon > 0$ existe $\delta > 0$ tal que

$$y \in X, d(x, y) < \delta \Rightarrow \tilde{d}(f(x), f(y)) < \epsilon.$$

0.2 O Teorema do Ponto Fixo para Contrações

Definição 0.2.1. (Espaço métrico completo). Um espaço métrico (X, d) é dito *completo* se toda sequéncia de Cauchy (x_n), com $x_n \in X$, converge para um ponto $x \in X$.

Definição 0.2.2. (Espaço de Banach). Um espaço vetorial normado cuja métrica oriunda da norma é completa é chamado de *espaço de Banach*.

Exemplo 0.2.3. Seja $X = \mathbb{R}^k$, e $\|\cdot\| : \mathbb{R}^k \to [0, \infty)$ uma norma qualquer. Então é possível provar que X com a métrica dada por $d(v, w) := \|v - w\|, \forall v, w \in \mathbb{R}^k$ é um espaço métrico completo, e portanto, um espaço de Banach. Tal fato segue-se de que toda sequência limitada em $\mathbb{R}^k$ possui uma subsequência convergente (teorema de Bolzano-Weierstrass).

Exemplo 0.2.4. Seja X um conjunto qualquer, e seja $(E, \|\cdot\|)$ um espaço de Banach. Defina o conjunto

$$\mathcal{F} = \mathcal{F}(X) := \{f : X \to E, f \text{ é limitada}\}.$$

Então a aplicação $\|\cdot\|_\infty : \mathcal{F} \to [0, +\infty)$ dada por

$$\|f\|_\infty = \sup\{\|f(x)\|, x \in X\}$$

é uma norma de $\mathcal{F}$, chamada de norma da convergência uniforme, ou norma do sup. É possível mostrar (veja o exercício 3) que $\mathcal{F}$ é um espaço de Banach.

Definição 0.2.5. (Aplicação lipschitziana). Sejam (X, d) e $(\hat{X}, \hat{d})$ espaços métricos. Uma aplicação $F : X \to \hat{X}$ é dita ser *lipschitziana* ou simplesmente *Lipschitz* se existe $0 \leq \lambda$ tal que

$$\hat{d}(F(x), F(y)) \leq \lambda \cdot d(x, y), \forall x, y \in X.$$

Dizemos que λ é uma *constante de Lipschitz* de F. Denotamos o ínfimo das constantes de Lipschitz de F por $\mathrm{Lip}(F)$, o qual é, ele mesmo, uma constante de Lipschitz.

Observação 0.2.6. Notamos que as aplicações lipschitzianas são contínuas: Se F é uma tal aplicação, supondo sem perda $\lambda > 0$, dados $x \in X$, $\epsilon > 0$, tomando $\delta = \epsilon/\lambda$, temos

$$d(x, y) < \delta \Rightarrow \hat{d}(F(x), F(y)) \leq \lambda \cdot d(x, y) < \lambda \cdot \epsilon/\lambda = \epsilon.$$

Observação 0.2.7. Se X, Y e Z são espaços métricos, com $f : X \to Y$ e $g : Y \to Z$ ambas lipschitzianas, então a composta $h = g \circ f : X \to Z$ também é Lipschitz com

$$\mathrm{Lip}(g \circ f) \leq \mathrm{Lip}(g) \cdot \mathrm{Lip}(f).$$

Uma subclasse relevante de aplicações Lipschitz é contituída pelas *contrações* de um espaço métrico nele mesmo:

Definição 0.2.8. (Contração). Seja (X, d) um espaço métrico. Uma aplicação $F : X \to X$ é dita ser uma *contração* se existe $0 \leq \lambda < 1$ tal que

$$d(F(x), F(y)) \leq \lambda \cdot d(x, y), \forall x, y \in X.$$

O próximo resultado corresponde à principal ferramenta para construir objetos em dimensão infinita, onde, ao contrário do que ocorre no $\mathbb{R}^n$, argumentos de compacidade são quase sempre inviáveis.

Teorema 0.2.9. *(Ponto fixo para contrações). Sejam (X, d) um espaço métrico completo e $F : X \to X$ uma contração. Então existe um único ponto fixo p por F, ou seja, existe um único ponto $p \in X$ tal que $F(p) = p$. Ademais, tal ponto fixo p é um atrator de F, isto é, fixado qualquer $x \in X$, $F^n(x) \to p$ quando $n \to +\infty$. ($F^n(x)$ é definido indutivamente por $F^n(x) := F(F^{n-1}(x))$.)*

Prova: Sejam $x \in X$ e $x_n = F^n(x), n \in \mathbb{N}$. Provaremos que x_n é uma sequência de Cauchy. Para tal, primeiro mostremos por indução que existe $0 \leq \lambda < 1$ tal que

$$d(x_{n+1}, x_n) \leq \lambda^n \cdot d(x_1, x_0), \forall n \in \mathbb{N}.$$

De fato, como F é contração, temos que existe $\lambda < 1$ tal que:

$$d(x_{n+1}, x_n) = d(F(x_n), F(x_{n-1})) \leq \lambda \cdot d(x_n, x_{n-1}),$$

o que já implica a fórmula de indução para $n = 1$ (o caso $n = 0$ é trivial. Supondo a fórmula válida para um certo $n \in \mathbb{N}$, para $n + 1$, da última desigualdade, temos:

$$d(x_{n+2}, x_{n+1}) \leq \lambda \cdot d(x_{n+1}, x_n) \underbrace{\leq}_{\text{hip. indução}} \lambda \cdot \lambda^n d(x_1, x_0) = \lambda^{n+1} \cdot d(x_1, x_0),$$

o que prova a indução desejada.

Dados $m \geq n$, temos portanto:

$$d(x_m, x_n) \leq (\lambda^n + \cdots + \lambda^m) \cdot d(x_1, x_0) \leq (\sum_{j=n}^{+\infty} \lambda_j) \cdot d(x_1, x_0) = \frac{\lambda^n}{1-\lambda} d(F(x), x),$$

o que prova que x_n é uma sequência de Cauchy, e como X é completo, tal sequência converge, digamos, para $p \in X$. Afirmamos que p é ponto fixo de F. Realmente,

$$F(p) = F(\lim_{n \to +\infty} x_n) = \lim_{n \to +\infty} F(x_n) = \lim_{n \to +\infty} x_{n+1} = p.$$

Notamos que a segunda igualdade acima se dá porque toda contração é contínua, e a última desigualdade se dá porque em uma sequência convergente toda subsequência converge para o mesmo limite.

É fácil ver que p é o único ponto fixo de F. De fato, se $p, q \in X$ são pontos fixos de F, temos:

$$d(p, q) = d(F(p), F(q)) \leq \lambda \cdot d(p, q) \Rightarrow$$

$$(1 - \lambda) \cdot d(p, q) \leq 0 \Rightarrow d(p, q) = 0 \Leftrightarrow p = q,$$

findando a prova do teorema.

□

Observação 0.2.10. Assinalamos que se p é o único ponto fixo de um iterado $F^m, m \geq 1$ de uma aplicação $F : X \to X$ qualquer, então p é o único ponto fixo de F. De fato:

$$F^m(p) = p \Rightarrow F^m(F(p)) = F(F^m(p)) = F(p),$$

ou seja, se p e $F(p)$ são pontos fixos de $F^m(p)$, logo $F(p) = p$. Isso é muito útil, pois nem sempre F é uma contração, mas muitas vezes um seu iterado é. Assim, a existência e unicidade preconizadas no teorema do ponto fixo para contrações continuam válidas para F se apenas um iterado positivo de F for contração.

Observação 0.2.11. Se $F : X \to X$, com X um espaço métrico completo, é tal que um iterado positivo $F^m, m > 1$ seu seja uma contração, então fixado $x \in X$, ainda vale que $F^n(x) \to p$ quando $n \to \infty$, onde p é o ponto fixo de

F. De fato, usando o Algoritmo da divisão de Euclides, dado $n \in \mathbb{N}$ podemos escrever $n = m \cdot j + r$, com $0 \leq r < m$ e $j \to +\infty$ quando $n \to \infty$. Daí,

$$d(F^n(x), p) = d(F^{m \cdot j + r}(x), p) = d((F^m)^j(F^r(x)), p).$$

Aplicando o teorema, já provado, a F^m, obtemos:

$$d(F^n(x), p) = d((F^m)^j(F^r(x)), p) \leq \lambda^j \frac{d(F^r(x), p)}{1 - \lambda} \leq$$

$$\frac{\lambda^{n/m-1}}{1-\lambda} \cdot \max_{s=0,\ldots,m-1}\{d(F^s(x), p)\} \leq \frac{(\sqrt[m]{\lambda})^n}{\lambda(1-\lambda)} \cdot \max_{s=0,\ldots,m-1}\{d(F^s(x), p)\},$$

ou seja, $F^n(x) \to p$ quando $n \to +\infty$, ainda que com uma taxa exponencial mais baixa que F^m.

Entre os inúmeros corolários do Teorema do Ponto Fixo encontram-se as versões não diferenciáveis do teorema da Função Inversa:

Teorema 0.2.12. *(Perturbação da Identidade). Sejam E um espaço vetorial normado completo (espaço de Banach), $I : E \to E$ a identidade em E e seja $\Phi : E \to E$ uma contração em E. Então $I + \Phi$ é um homeomorfismo sobre E.*

Prova: Sejam x, $y \in E$ e $h = I + \Phi$. Seja $0 < \lambda < 1$ a constante de Lipschitz de Φ. Então

$$\|I(x)+\Phi(x)-I(y)-\Phi(y)\| \geq \|x-y\|+\|\Phi(x)-\Phi(y)\| \geq \|x-y\|-\lambda\cdot\|x-y\| =$$

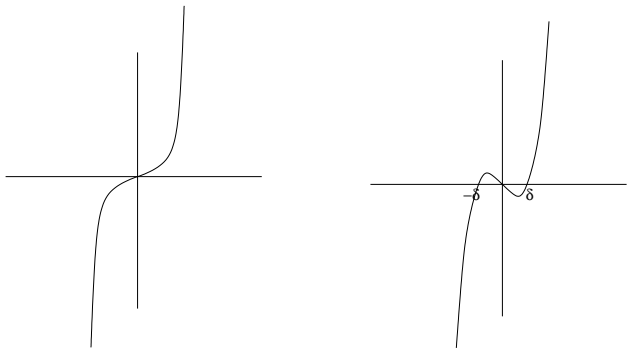
$$(1-\lambda)\cdot\|x-y\| \Rightarrow \|h(x)-h(y)\| \geq (1-\lambda)\cdot\|x-y\| \neq 0 \text{ se } x \neq y;$$

donde obtemos a injetividade de h, e também a continuidade de h^{-1}. Mostremos agora a sobrejetividade de h. Seja $z \in E$. Queremos ver que existe $p \in E$ tal que $h(p) = z \Leftrightarrow p + \Phi(p) = z \Leftrightarrow p = z - \Phi(p)$. Por conseguinte definamos $f_z : E \to E$ por $f_z(x) = z - \Phi(x)$. Basta então acharmos um ponto fixo p para f_z, que teremos $h(p) = z$. De fato, $f_z : E \to E$ é contração:

$$\|f_z(x) - f_z(y)\| = \|z - \Phi(x) - z + \Phi(y)\| = \|\Phi(y) - \Phi(x)\| \leq \lambda \cdot \|x - y\|.$$

Como E é espaço normado completo, segue-se do teorema do ponto fixo para contrações (teorema 0.2, página 5) que existe um único $p \in E$ tal que $h(p) = z$, como queríamos. Isso nos dá ao mesmo tempo a sobrejetividade e uma nova prova da injetividade.

$\square$



Os gráficos de $y = x^3$ e de $y = x^3 - \delta x^2$ nos mostram que somando−se uma contração a um homeomorfismo com inversa não lipschitziana, o resultado pode não ser um homeomorfismo. Mostra ademais que a soma de homeomorfismos pode não ser um homeomorfismo.

Observação 0.2.13. É possível melhorar ainda mais a última demonstração: se $a \in E$ e $b = h(a)$, mostremos que dado $\delta > 0$, então $h(B(a,\delta)) \supset B(b,(1-\lambda)\cdot\delta)$. De fato, vimos que dado $z \in B(b,(1-\lambda)\cdot\delta)$, existe um único $p \in E$ tal que $h(p) = z$. Daí,

$$\|p-a\| = \|f_z(p)-a\| = \|z-\Phi(p)-a\| = \|z-\Phi(p)\underbrace{-a-\Phi(a)}_{-b}+\Phi(a)\| \leq$$

$$\|z-b\| + \|\Phi(a)-\Phi(p)\| \leq \|z-b\| + \lambda\cdot\|p-a\| \Rightarrow$$
$$(1-\lambda)\cdot\|p-a\| \leq \|z-b\| \leq (1-\lambda)\cdot\delta \Rightarrow \|p-a\| \leq \delta.$$

Isso prova novamente a continuidade de h^{-1}. Além do mais, nos dá um controle sobre o comportamento local de h.

Lema 0.2.14. *Seja E um espaço de Banach, $L \in \mathcal{L}(E,E)$ satisfazendo $\|L\| \leq a < 1$ e $G \in \mathcal{L}(E,E)$ isomorfismo com $\|G^{-1}\| \leq a < 1$. Então:*

a) $(I+L)$ é isomorfismo e $\|(I+L)^{-1}\| \leq 1/(1-a)$;

b) $(I+G)$ é isomorfismo e $\|(I+G)^{-1}\| \leq a/(1-a)$.

Prova:

a) Seja $y \in E$ qualquer fixado. Defina $u : E \to E$ por

$$u(x) := y - L(x).$$

Logo
$$|u(x_1) - u(x_2)| = |L(x_2 - x_1)| \leq a \cdot |x_2 - x_1|,$$
o que implica que $u : E \to E$ é uma contração. Pelo teorema do ponto fixo para contrações,

$$\exists! z \in E / \, u(z) = z \Leftrightarrow \exists! z \in E / \, z = y - L(z) \Leftrightarrow \exists! z \in E / \, y = z + L(z),$$

o que implica que $(I + L)$ é isomorfismo.

Seja $y \in E$ com $|y| = 1$ e seja $x \in E$ tal que $(L + I)^{-1}(y) = x$. Como $x + L(x) = y$, temos que $|x| - a \cdot |x| \leq 1 \Rightarrow |x| \leq 1/(1 - a)$, donde se conclui que $\|(I + L)^{-1}\| \leq 1/(1 - a)$.

b) $(I + G) = G \cdot (I + G^{-1})$. Como

$$\|G^{-1}\| \leq a < 1 \underbrace{\Rightarrow}_{\text{ítem a)}} (I + G^{-1}) \text{ é inversível.}$$

Daí, $(I + G)^{-1} = (I + G^{-1})^{-1} \cdot G^{-1}$, o que implica que

$$\|(I + G)^{-1}\| \leq \|(I + G^{-1})^{-1}\| \cdot \|G^{-1}\| \leq \frac{1}{1 - a} \cdot a = \frac{a}{1 - a}.$$

□

Corolário 0.2.15. *(Perturbação de uma aplicação bilipschitz). Sejam $E, \tilde{E}$ espaços de Banach e $\Psi : E \to \tilde{E}$ uma aplicação bilipschitz (sobrejetiva), isto é, f é invertível e lipschitziana com inversa também lipschitziana. Seja $\Phi : E \to \tilde{E}$ Lipschitz tal que sua constante de Lipschitz $\mathrm{Lip}(\Phi) < Lip(\Psi^{-1})^{-1}$. Então $\Psi + \Phi : E \to \tilde{E}$ é um homeomorfismo (sobrejetivo).*

Prova: Considere $h : \tilde{E} \to \tilde{E}$ dado por

$$h := (\Psi + \Phi)\Psi^{-1} = I + \Phi \circ \Psi^{-1}.$$

Dados $\tilde{x}, \tilde{y} \in \tilde{E}$,

$$\|\Phi(\Psi^{-1}(\tilde{x})) - \Phi(\Psi^{-1}(\tilde{y}))\| \leq \mathrm{Lip}(\Phi) \cdot \|\Psi^1(\tilde{x}) - \Psi^{-1}(\tilde{y})\| \leq$$

$$\mathrm{Lip}(\Phi) \cdot \mathrm{Lip}(\Psi^{-1})\|\tilde{x} - \tilde{y}\| \Rightarrow \|\Phi \circ \Psi^{-1}(\tilde{x}) - \Phi \circ \Psi^{-1}(\tilde{y})\| \leq \lambda\|\tilde{x} - \tilde{y}\|,$$

ou seja, $\Phi \circ \Psi^{-1}$ é uma $\lambda-$contração. Logo, pelo teorema da perturbação da identidade, $h = (\Psi + \Phi) \circ \Psi^{-1} = I + \Phi\Psi^{-1}$ é um homeomorfismo (injetivo e sobre $\tilde{E}$). Portanto a composição

$$(\Psi + \Phi)\Psi^{-1} \circ \Psi = \Psi + \Phi$$

é um homeomorfismo, como queríamos mostrar.

□

Corolário 0.2.16. *(Perturbação do Isomorfismo). Sejam $E, \tilde{E}$ espaços de Banach e $T : E \to \tilde{E}$ um isomorfismo linear (sobrejetivo). Seja $\Phi : E \to \tilde{E}$ Lipschitz tal que sua constante de Lipschitz* $\mathrm{Lip}(\Phi) < \|T^{-1}\|^{-1}$. *Então $T + \Phi : E \to \tilde{E}$ é um homeomorfismo (sobrejetivo).*

Prova: O corolário segue imediatamente do corolário anterior. Podemos ainda dar-lhe outra prova a partir do teorema da perturbação da identidade, adaptando a prova do Corolário anterior, conforme fazemos abaixo. A linearidade de T torna a adaptação mais simples:

Considere $h : E \to E$ dado por

$$h := T^{-1}(T + \Phi) = I + T^{-1}\Phi.$$

Dados $x, y \in E$,

$$\|T^{-1} \circ \Phi(x) - T^{-1} \circ \Phi(y)\| = \|T^{-1}(\Phi(x) - \Phi(y))\| \leq \|T^{-1}\| \cdot \|\Phi(x) - \Phi(y)\| \leq$$

$$\|T^{-1}\| \cdot \mathrm{Lip}(\Phi) \cdot \|x - y\| \Rightarrow \|T^{-1}\Phi(x) - T^{-1}\Phi(y)\| \leq \lambda\|x - y\|,$$

ou seja, $T^{-1} \circ \Phi$ é uma $\lambda-$contração. Logo, pelo teorema da perturbação da identidade, $h = T^{-1}(T + \Phi) = I + T^{-1}\Phi$ é um homeomorfismo (injetivo e sobre). Portanto a composição

$$T \circ (T^{-1}(T + \Phi)) = T + \Phi$$

é um homeomorfismo, como queríamos mostrar.

□

0.3 Espaços de aplicações contínuas com domínio compacto

Lembramos a definição de conjunto compacto:

Definição 0.3.1. (Conjunto compacto). Seja M um espaço métrico. Um conjunto $K \subset M$ é dito *compacto* se para toda união de bolas abertas $\cup_\lambda B_\lambda$ contendo K (também chamada cobertura aberta de K) podemos extrair uma subcoleção finita $B_1, \ldots, B_r$ tal que $\cup_{j=1}^r B_j \supset K$. Sucintamente, dizemos que um conjunto K é compacto se e só se toda cobertura por abertos de K admite uma subcobertura finita.

Uma caracterização muitas vezes útil dos conjuntos compactos, equivalente à definição acima, é dada com o auxílio do seguinte conceito:

Definição 0.3.2. (Propriedade da intersecção finita). Uma família (possivelmente não enumerável) $(F_\lambda), \lambda \in \Lambda$ de conjuntos fechados F_λ de um espaço topológico X é dita ter a *propriedade da intersecção finita* (abreviadamente, *p.i.f.*) se toda sua subcoleção finita $\{F_{\lambda_1}, \ldots, F_{\lambda_k}\} \subset \{F_\lambda, \lambda \in \Lambda\}$ possui intersecção

$$F_{\lambda_1} \cap \cdots \cap F_{\lambda_k} \neq \emptyset.$$

Proposição 0.3.3. *Seja X um espaço topológico. Então $K \subset X$ é compacto $\Leftrightarrow$ toda família $(F_\lambda), \lambda \in \Lambda$ de fechados $F_\lambda \subset K$ com a propriedade da intersecção finita possui intersecção*

$$\cap_{\lambda \in \Lambda} F_\lambda \neq \emptyset.$$

Prova: ($\Rightarrow$) Seja K compacto, e $(F_\lambda), \lambda \in \Lambda$ uma família de subconjuntos fechados de K com a propriedade da intersecção finita. Suponha por absurdo que $\cap_{\lambda \in \Lambda} F_\lambda = \emptyset$. Denotando por F_λ^c o complementar de F_λ em K, então

$$\cup_{\lambda \in \Lambda} {F_\lambda}^c = (\cap_{\lambda \in \Lambda} F_\lambda)^c = \emptyset^c = K,$$

o que significa que $\cup_{\lambda \in \Lambda} {F_\lambda}^c$ constitui uma cobertura por abertos de K. Entretanto, tal cobertura não admite subcobertura finita: dada qualquer coleção finita

$$\{F_{\lambda_1}^c, \ldots, F_{\lambda_q}^c\} \subset \{F_\lambda^c, \lambda \in \Lambda\},$$

temos

$$\cup_{j=1}^q F_{\lambda_j}^c = (\cap_{j=1}^q F_{\lambda_j})^c \neq (\emptyset)^c = K,$$

o que significa que a cobertura $\{F_\lambda^c, \lambda \in \Lambda\}$ de K não admite subcobertura finita, absurdo, pois K é compacto.

($\Leftarrow$) Agora seja $\cup_{\lambda \in \Lambda} A_\lambda = K$ uma cobertura por abertos (em K) e suponha por absurdo que $\{A_\lambda, \lambda \in \Lambda\}$ não admita subcobertura finita. Daí,

família de fechados $\{A_\lambda^c, \lambda \in \Lambda\}$ possui a propriedade da intersecção finita, pois dada qualquer subcoleção finita

$$\{A_{\lambda_1}^c, \ldots, A_{\lambda_q}^c\} \subset \{A_\lambda^c, \lambda \in \Lambda\},$$

temos

$$\cap_{j=1}^q A_{\lambda_j}^c = (\cup_{j=1}^q A_{\lambda_j})^c \neq (K)^c = \emptyset.$$

Todavia,

$$\cap_{\lambda \in \Lambda} {A_\lambda}^c = (\cup_{\lambda \in \Lambda} A_\lambda)^c = K^c = \emptyset,$$

o que é absurdo, pois por hipótese toda família de fechados de K com a propriedade da intersecção finita possui intersecção não vazia.

□

Embora, como dissemos, os argumentos de compacidade não sejam comuns em espaços de dimensão infinita, no contexto específico de espaço de funções contínuas com domínio compacto, algumas vezes tais argumentos são possíveis. Tal é o conteúdo do teorema de Ascoli-Arzelá, que classifica os conjuntos compactos desses espaços.

Definição 0.3.4. (Sequência equicontínua de funções). Sejam M, N espaços métricos. Uma sequência de funções $f_n : M \to N$, $n \in \mathbb{N}$ é dita *equicontínua* se dados $\epsilon > 0$ e $x \in M$, então existe $\delta > 0$ tal que

$$y \in M, d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \epsilon, \forall n \in \mathbb{N}.$$

Antes do teorema de Ascoli-Arzelá, vejamos alguns resultados intermediários, porém relevantes em si mesmos:

Teorema 0.3.5. *Seja M um espaço métrico e N um espaço métrico completo. Suponha que $f_n : M \to N$ seja uma sequência de funções equicontínuas que convirja pontualmente em todo x pertencente a um subconjunto $D \subset M$, denso em M. Então, f_n converge uniformemente em partes compactas de M a uma função contínua $f : M \to N$.*

Prova: Primeiramente vejamos que f_n converge pontualmente (em todo ponto de M) para uma função $f : M \to N$. Para tanto, basta mostrarmos que para cada $x \in M$, a sequência $(f_n(x))$ é de Cauchy em N. Seja $\epsilon > 0$ dado. Então, pela equicontinuidade das f_n, existe $\delta > 0$ tal que

$$d(x, y) < \delta \Rightarrow d(f_n(x), f_n(y)) < \epsilon/3, \forall n \in \mathbb{N}.$$

Seja então $y \in D \cap B(x, \delta)$. Daí, existe n_0 tal que $d(f_m(y), f_n(y)) < \epsilon/3$, $\forall m, n \geq n_0$. Temos portanto que, $\forall m, n \geq n_0$, vale:

$$d(f_m(x), f_n(x)) \leq d(f_m(x), f_m(y)) + d(f_m(y), f_n(y)) + d(f_n(y), f_n(x)) < \epsilon,$$

o que implica que $f_n(x)$ é de Cauchy e como N é completo, podemos definir $f(x) := \lim_{n \to +\infty} f_n(x)$.

Mostremos que f, assim definida, é contínua: De fato, dados $\epsilon > 0$ e $x_0 \in M$, existe $\delta_0 > 0$ tal que

$$d(x, x_0) \leq \delta_0 \Rightarrow d(f_n(x), f_n(x_0)) < \epsilon, \forall n \in \mathbb{N} \Rightarrow d(f(x), f(x_0)) \leq \epsilon.$$

Afirmamos agora que se uma sequência equicontínua de aplicações $f_n : M \to N$ converge pontualmente em M. então sua convergência é uniforme em cada parte compacta $K \subset M$. Com efeito, para cada $x \in M$ seja $n_x \in \mathbb{N}$ tal que

$$d(f_n(x), f(x)) < \epsilon/3, \forall n \geq n_x.$$

Ademais, para cada $x \in M$, podemos tomar uma bola aberta $B_x = B(x, r_x)$ tal que $\forall n \in \mathbb{N}$,

$$y \in B_x \Rightarrow d(f_n(x), f_n(y)) < \epsilon/3, d(f(x), f(y)) < \epsilon/3.$$

Tomando $K \subset M$ um compacto, podemos extrair da cobertura $\cup_{x \in K} B_x \supset K$ uma subcobertura finita $K \subset B_{x_1} \cup \cdots \cup B_{x_p}$. Ponha $n_0 = \max\{n_{x_1}, \ldots, n_{x_p}\}$. Então para todo $x \in K$, temos $d(f_n(x), f(x)) < \epsilon$, $\forall n \geq n_0$. Realmente, dado $x \in K$, então existe $1 \geq j \geq p$ tal que $x \in B_{x_j}$. Portanto,

$$d(f_n(x), f(x)) \leq d(f_n(x), f_n(x_j)) + d(f_n(x_j), f(x_j)) + d(f(x_j), f(x)) < \epsilon, \forall n \geq n_0.$$

□

Observação 0.3.6. Vale uma espécie de recíproca do teorema 0.3.5 acima. Isto é, se M e N são espaços métricos e $f_n : M \to N$ é uma sequência de funções contínuas convergindo uniformemente em M para $f : M \to N$, então a sequência (f_n) é equicontínua. De fato, tomando $x_0 \in M$, dado $\epsilon > 0$ existe $\delta_0 = \delta(x_0, \epsilon) > 0$ tal que

$$d(x, x_0) < \delta_0 \Rightarrow d(f(x), f(x_0)) < \epsilon/3.$$

Usando da convergência uniforme, podemos tomar $n_0 \in \mathbb{N}$ tal que $d(f_n(x), f(x)) < \epsilon/2$, $\forall n > n_0$; daí obtemos

$$d(f_n(x), f_n(x_0)) \leq d(f_n(x), f(x)) + d(f(x), f(x_0)) + d(f(x_0), f_n(x_0)) < \epsilon,$$

$\forall x \in B(x_0, \delta_0), \forall n > n_0$. Usando da continuidade das f_n, para cada $n = 1 \dots n_0$, seja $\delta_n > 0$ tal que

$$d(x, x_0) < \delta_n \Rightarrow d(f(x), f(x_0)) < \epsilon.$$

Por conseguinte, tomando $\delta := \min\{\delta_0, \delta_1, \dots, \delta_{n_0}\}$, concluímos que

$$d(x, x_0) < \delta \Rightarrow d(f_n(x), f_n(x_0)) < \epsilon, \forall n \geq n_0.$$

Como $x_0 \in M$ é arbitrário, temos que f_n é equicontínua.

Uma vez que provamos um resultado em que a convergência pontual (ou simples) implica em convergência uniforme em partes compactas, cumpre agora obter condições para a convergência pontual de subsequências de funções contínuas.

Teorema 0.3.7. *(Cantor-Tychonov). Seja X um conjunto enumerável qualquer. Toda sequência pontualmente limitada de funções $f_n : X \to \mathbb{R}^m$ possui uma subsequência pontualmente convergente.*

Prova: Seja $X = \{x_1, x_2, \dots\}$. A sequência $(f_n(x_1))_{n\in\mathbb{N}}$, sendo uma sequência limitada em $\mathbb{R}^m$, possui uma subsequência convergente $(f_n(x_1))_{n\in\mathbb{N}_1}$, $\mathbb{N}_1 \subset \mathbb{N}$ um subconjunto enumerável. Igualmente é limitada a sequência $(f_n(x_2))_{n\in\mathbb{N}_1}$ é limitada, logo podemos achar um subconjunto infinito $\mathbb{N}_2 \subset \mathbb{N}_1$ tal que $(f_n(x_2))_{n\in\mathbb{N}_2}$ seja convergente. Prosseguindo analogamente, conseguimos, para cada $j \in \mathbb{N}$ um subconjunto infinito $\mathbb{N}_j \subset \mathbb{N}$, de tal modo que $\mathbb{N}_1 \supset \mathbb{N}_2 \supset \dots \supset \mathbb{N}_j \supset \dots$ e para cada $j \in \mathbb{N}$, a sequência $(f_n(x_j))_{n\in\mathbb{N}_j}$ é convergente a um certo $a_j \in \mathbb{R}^m$. Definimos então um subconjunto $\tilde{\mathbb{N}} \subset \mathbb{N}$ tomando como $j-$ésimo elemento de $\tilde{N}$ o $j-$ésimo elemento de $\mathbb{N}_j$. Desta maneira, fixado $x_j \in X$, a sequência $(f_n(x_j))_{n\in\tilde{\mathbb{N}}}$ é, a partir de seu $j-$ésimo elemento uma subsequência de $(f_n(x_j))_{n\in\mathbb{N}_j}$, logo converge. Portanto, f_n converge pontualmente em cada $x_j \in X$.

□

Observação 0.3.8. Lembramos que todo espaço métrico compacto possui uma base enumerável (exercício 1), o que para espaços métricos é equivalente a possuir um subconjunto denso enumerável (exercício 2).

Usamos a observação acima na prova do próximo teorema:

Teorema 0.3.9. *(Ascoli-Arzelá). Seja K um espaço métrico compacto, e $f_n : K \to \mathbb{R}^m$ uma sequência equicontínua de funções pontualmente limitada. Então (f_n) admite uma subsequência uniformemente convergente (a uma função contínua $f : K \to \mathbb{R}^m$).*

Prova: Como K é um espaço métrico compacto, segue-se que existe um subconjunto $D \subset K$ denso enumerável. Por Cantor-Tychonov (teorema 0.3.7), existe uma subsequência (f_{n_j}) de (f_n) convergindo pontualmente em cada ponto $x \in D$. Pelo teorema 0.3.5, isto implica que f_{n_j} converge uniformemente em K a uma função contínua $f : K \to \mathbb{R}^m$.

□

Teorema 0.3.10. *(Dini). Seja $f_n : K \to M$ uma sequência de funções contínuas definidas em um espaço métrico compacto K, tendo um espaço métrico M como contradomínio. Suponha que f_n converge pontualmente para uma função contínua $f : K \to M$ e, além disso, $\forall x \in K$ tem-se*

$$d(f(x), f_1(x)) \geq d(f(x), f_2(x)) \geq \cdots \geq d(f_n(x), f(x)) \geq \ldots$$

Então a convergência $f_n \to f$ é uniforme.

Prova: Dado $\epsilon > 0$, para cada $n \in \mathbb{N}$, definimos

$$F_n := \{x \in M; d(f_n(x), f(x)) \geq \epsilon\}.$$

Como d, f_n e f são aplicações contínuas, o mesmo vale para $d(f_n, f)$, o que implica que os conjuntos F_n acima definidos são fechados em K. Provar a convergência uniforme é o mesmo que mostrar que existe $n_0 \in \mathbb{N}$ tal que $F_n = \emptyset, \forall n \geq n_0$. Ora, como $\lim_{n\to\infty} f_n(x) = f(x), \forall x \in K$, segue-se que $\cap_{n=1}^{\infty} F_n = \emptyset$. Da *propriedade da intersecção finita* do compacto K, segue-se que existe n_0 tal que $F_{n_0} = \emptyset$, o que implica, do fato de que os F_n são encaixantes, que $F_n = \emptyset, \forall n \geq n_0$.

□

Uma proposição muito simples que nos será útil diversas vezes é a seguinte:

Proposição 0.3.11. *Sejam M, K e N espaços métricos, e $f_n : M \to K$, $g_n : K \to N$ sequências de aplicações contínuas convergindo uniformente para aplicações $f : M \to K$ e $g : K \to N$, respectivamente. Suponha que g seja uniformemente contínua (este é o caso, por exemplo de quando K é compacto). Então a sequência de aplicações $h_n : M \to N$ dada por $h_n(x) := g_n \circ f_n(x)$ converge uniformemente para $h = g \circ f$.*

Prova: Seja $\epsilon > 0$ dado.
Considere $\delta > 0$ da continuidade uniforme da g tal que

$$d(y,z) < \delta, y, z \in K \Rightarrow d(g(y), g(z)) < \epsilon/2.$$

Da convergência uniforme da sequência f_n temos que existe $n_1 \in \mathbb{N}$

$$d(f_n(x), f(x)) < \delta, \forall n \geq n_1, \forall x \in M$$

Da convergência uniforme da sequência g_n temos que existe $n_2 \in \mathbb{N}$ tal que

$$d(g_n(y), g(y)) < \epsilon/2, \forall n \geq n_2, \forall y \in K.$$

Tomando $n_0 \geq \max\{n_1, n_2\}$, obtemos $\forall n \geq n_0$ e $\forall x \in M$:

$$d(g_n(f_n(x)), g(f(x))) \leq d(g_n(f_n(x)), g(f_n(x))) + d(g(f_n(x)), g(f(x))) < \epsilon.$$

□

Lema 0.3.12. *Seja $0 < c < 1$ fixado e seja $f : [0,c] \to \mathbb{R}$ definida por $f(t) := \sqrt{t}$. Existe uma sequência de polinômios $p_n : [0,c] \to \mathbb{R}$ convergindo uniformemente a f em todo ponto $t \in [0,c]$.*

Prova: A prova se baseia em uma aplicação do teorema do ponto fixo, para mostrar a convergência pontual, seguida de uma aplicação do teorema de Dini para garantir a convergência uniforme.

Tomemos a sequência (p_n) de polinômios dada por $p_0 \equiv 0$ e

$$p_{n+1}(t) = p_n(t) + \frac{1}{2}(t - (p_n(t))^2), \forall t \in [0,c], n \in \mathbb{N}.$$

Seja $t \in [0,c]$ fixado, e considere a função

$$g(x) := x + (t - x^2)/2.$$

Como $g(0) = t/2$, $g'(x) = 1 - x \geq 0$ e $g(\sqrt{t}) = \sqrt{t}$, o que implica que g é uma bijeção crescente de $[0, \sqrt{t}]$ sobre $[t/2, \sqrt{t}]$ (note que aqui usamos que $t < 1$, o que nos dá $t/2 \leq \sqrt{t}$), que g é uma contração (devido ao teorema do valor médio) e que tem $\sqrt{t}$ como seu único ponto fixo no espaço métrico compacto $[0, \sqrt{t}]$. Como $p_{n+1}(t) = g(p_n(t))$ segue-se indutivamente que $p_n(t) = g^n(p_0(t)) = g^n(0)$, o que implica pelo teorema do ponto fixo que $p_n(t) \to \sqrt{t} = g(\sqrt{t})$ e que $p_n(t) \in [0, \sqrt{t}]$, t fixado qualquer em

$[0,1]$. Falta agora vermos que a convergência $p_n \to f$ é uniforme. Como já adiantamos, isso é feito usando o teorema de Dini, mostrando que cada sequência $(p_n(t)), t \in [0,c]$ é monótona. De fato, como vimos acima, temos $p_n(t)^2 \leq t, \forall t \in [0,c]$, donde tiramos que

$$p_{n+1}(t) = p_n(t) + \frac{1}{2}(t - p_n(t)^2) \geq p_n(t),$$

implicando a monotonicidade desejada.

□

Corolário 0.3.13. *Fixado um intervalo compacto $[a,b]$, a aplicação $|\cdot| : [a,b] \to \mathbb{R}$ é uniformemente aproximada por uma sequência de polinômios.*

Prova: Sem perda de generalidade, suponha que o intervalo dado é da forma $[-a,a]$, com $a > 0$, pois todo intervalo compacto está contido em um desse tipo. Tome $c = 1/2$ no último lema e seja $p_n : [0,c] \to \mathbb{R}$ uma sequência de polinômios convergindo uniformemente para $f = \sqrt{|}_{[0,c]}$. Então, a sequência $q_n : [-a,a] \to \mathbb{R}$ dada por

$$q_n(t) := 2 \cdot a \cdot p_n(\frac{t^2}{4a^2}) = 2 \cdot a \cdot p_n((\frac{t}{2a})^2)$$

define, pela proposição 0.3.11, uma sequência de polinômios que tende uniformemente para $\sqrt{t^2} = |t|$ quando $t \in [-a,a]$.

□

Lema 0.3.14. *Seja $K \subset \mathbb{R}^m$ um conjunto compacto e sejam $p : K \to \mathbb{R}$, $\hat{p} : K \to \mathbb{R}$ polinômios. Então, existe uma sequência de polinômios $p_n : K \to \mathbb{R}$ convergindo uniformemente para $h := \max\{p, \hat{p}\}$. (Analogamente, como $\min\{p(x), \hat{p}(x)\} = -\max\{-p(x), -\hat{p}(x)\}$, temos que também existe uma sequência de polinômios convergindo uniformemente para $\min\{p, \hat{p}\}$).*

Prova: Como p e $\hat{p}$ são, em particular, aplicações contínuas no compacto K, seus conjuntos-imagem estão contidos em algum intervalo aberto limitado do tipo $(-a,a)$. Notamos que o máximo de p e $\hat{p}$ é dado pela fórmula:

$$\max\{p(x), \hat{p}(x)\} = \frac{p(x) + \hat{p}(x) + |p(x) - \hat{p}|}{2}.$$

Daí se $q_n : [-a,a] \to \mathbb{R}$ é a sequência de polinômios obtida no corolário anterior que aproxima a função contínua módulo, temos que

$$p_n(x) := \frac{p(x) + \hat{p}(x) + q_n(p(x) - \hat{p}(x))}{2}$$

é uma sequência de polinômios convergindo uniformemente para $\max\{p, \hat{p}\}$.

□

Corolário 0.3.15. *Sejam $K \subset \mathbb{R}^m$ um conjunto compacto, $f : K \to \mathbb{R}$ e $\hat{f} : K \to \mathbb{R}$ funções que são limites uniformes de sequências de polinômios. Então $\max\{f, \hat{f}\}$ (respectivamente, $\min\{f, \hat{f}\}$) é limite uniforme de uma sequência de polinômios.*

Prova: De fato, suponha que $f_n \to f$ e $\hat{f}_n \to \hat{f}$ sejam sequências de polinômios convergindo uniformemente a f e $\hat{f}$, respectivamente. Então, para cada n, do lema anterior temos que para cada n existe p_n tal que

$$d(p_n(x), \max\{f_n(x), \hat{f}_n(x)\}) < 1/n, \forall n \in \mathbb{N}, \forall x \in K.$$

Pela proposição 0.3.11, dado $\epsilon > 0$, existe n_0 tal que

$$d(\max\{f_n(x), \hat{f}_n(x)\}, \max\{f(x), \hat{f}(x)\}) < \epsilon/2, \forall n \geq n_0, \forall x \in K.$$

Tomando $n_1 > n_0$ tal que $\epsilon/2 > 1/n_1$, $\forall n \geq n_1$, $\forall x \in K$, obtemos

$$d(p_n(x), \max\{f(x), \hat{f}(x)\}) \leq d(p_n(x), \max\{f_n(x), \hat{f}_n(x)\}) +$$

$$d(\max\{f_n(x), \hat{f}_n(x)\}, \max\{f(x), \hat{f}(x)\}) < \epsilon,$$

implicando que $p_n \to \max\{f, \hat{f}\}$ uniformemente.

□

Teorema 0.3.16. *(Aproximação de Weierstrass). Seja $K \subset \mathbb{R}^m$ um espaço métrico compacto. Então, dada uma aplicação contínua $f : K \to \mathbb{R}$, existe uma sequência de polinômios $p_n : K \to \mathbb{R}$ convergindo uniformemente para f.*

Prova: Para provar o teorema, basta mostrar que dado $\epsilon > 0$, existe um polinômio p tal que

$$d(f(x), p(x)) < \epsilon, \forall x \in K.$$

Dados $x, y \in K$, existe um polinômio p_{xy} tal que $p_{xy}(x) = f(x)$ e $p_{xy}(y) = f(y)$. Por exemplo, se $x \neq y$, existe pelo menos uma coordenada $x_j \neq y_j$. Basta tomar

$$p_{xy}(z_1, \ldots z_m) := \frac{f(x) \cdot (z_j - y_j)}{x_j - y_j} + \frac{f(y) \cdot (z_j - x_j)}{y_j - x_j}.$$

No caso $x = y$, basta tomar p_{xy} constante igual a $f(x)$. Por continuidade de $p_{xy} - f$, para cada $x \in K$, cada ponto $y \in K$ possui uma vizinhança V_{xy} tal que

$$z \in V_{xy} \Rightarrow p_{xy}(z) > f(z) - \epsilon/2.$$

Sendo K compacto, existem $y_1, \ldots, y_r$ tais que $K = V_{xy_1} \cup \cdots \cup V_{xy_r}$. Seja $g_x = \max\{p_{xy_1}, \ldots, p_{xy_r}\}$. Então, $g_x(x) = f(x)$ e $g_x(z) > f(z) - \epsilon/2, \forall z \in K$. Pelo último corolário, existe um polinômio p_x tal que

$$d(p_x(z), g_x(z)) < \epsilon/4, \forall z \in K.$$

Então $p_x(z) > f(z) - \epsilon/2$, e sem perda de generalidade, podemos supor $p_x(x) = f(x)$ (Caso não fosse, bastaria trocar p_x por $q_x := p_x + (f(x) - p_x(x))$). Por continuidade de $p_x - f$, cada ponto $x \in K$ possui uma vizinhança U_x tal que

$$z \in U_x \Rightarrow p_x(z) < f(z) + \epsilon/2.$$

Novamente, da compacidade de K, existe uma subcobertura finita $U_{x_1} \cup \cdots \cup U_{x_s} \supset K$. Como antes, definimos $g = \min\{p_{x_1}, \ldots, p_{x_s}\}$, donde obtemos que

$$f(z) - \epsilon/2 < p_{x_j}(z) \leq g(z) < f(z) + \epsilon/2, \forall z \in K, j = 1, \ldots, s.$$

Pelo último corolário, obtemos que existe um polinômio p tal que

$$d(p(z), g(z)) < \epsilon/2, \forall z \in K,$$

o que implica que

$$d(p(z), f(z)) \leq d(p(z), g(z)) + d(g(z), f(z)) < \epsilon, \forall z \in K,$$

como queríamos mostrar.

□

0.4 Integração de Caminhos em Espaços de Banach

Definição 0.4.1. (Partição de um intervalo). Uma *partição $\mathcal{P}$ de um intervalo* $[a, b] \subset \mathbb{R}$ é uma coleção finita $\mathcal{P} = \{I_1, \ldots, I_j\}$ de intervalos dois a dois disjuntos tais que $I_1 = [x_0, x_1), \ldots, I_{j-1} = [x_{j-2}, x_{j-1}), I_j = [x_{j-1}, x_j]$, com $x_0 = a, x_j = b$ e $x_0 \leq \cdots \leq x_j$. Note que uma partição $\mathcal{P}$ de

um intervalo $[a, b]$ fica inteiramente determinada pelo conjunto dos pontos $A_{\mathcal{P}} := \{a = x_0, \ldots, x_j = b\}$, o qual designaremos por conjunto dos pontos associados a $\mathcal{P}$.

Definição 0.4.2. (Diâmetro de uma partição de um intervalo). O *diâmetro de uma partição* P de um intervalo I é o máximo dos diâmetros (comprimentos) dos elementos de $\mathcal{P}$.

Definição 0.4.3. (Integral de Riemann). Seja $I = [a, b]$ e $f : I \to E$ um caminho limitado, tomando valores em um espaço de Banach E. A integral de Riemann $\int_I f(x)dx \in E$, se existir, é o limite

$$\int_I f(x)dx := \lim_{\text{diam}(\mathcal{P})\to 0} \sum_{j=1}^{\#\mathcal{P}} f(x_j) \cdot vol(I_j),$$

onde $x_j \in I_j$ e $\mathcal{P} = \{I_j, j = 1, \ldots, \#\mathcal{P}\}$, e vol é o volume (comprimento) do intervalo. Mais precisamente, dizer que existe o limite "$\lim_{\text{diam}(\mathcal{P})\to 0}$", quer dizer que, dado $\epsilon > 0$, $\exists \delta > 0$ tal que, para toda partição de I com $\text{diam}(\mathcal{P}) < \delta$, e para todo conjunto finito $\{x_1, \ldots, x_{\#\mathcal{I}}\}$, com $x_j \in I_j$, temos

$$\| \sum_{j=1}^{\#\mathcal{P}} f(x_j) \cdot \text{vol}(I_j) - \int_I f(x)dx \| < \epsilon.$$

Se existir a integral de Riemann de uma aplicação f, então dizemos que f é *integrável à Riemann*, ou simplesmente, *integrável*. Uma soma do tipo $\sum_{j=1}^{\#\mathcal{P}} f(x_j) \cdot \text{vol}(I_j)$, com $x_j \in P_j$ e $\mathcal{P} = \{C_1, \ldots, I_{\#\mathcal{P}}\}$ é chamada de *soma de Riemann de f em relação a $\mathcal{P}$*, e denotada por $s(f, \mathcal{P})$, ou apenas, por $s(\mathcal{P})$ nos contextos em que f puder ser subentendida sem ambiguidades.

Definição 0.4.4. (Refinamento de uma partição). Seja $\mathcal{P}$ uma partição de um intervalo $I \subset \mathbb{R}^n$. Uma partição $\hat{\mathcal{P}}$ de I é dita ser um *refinamento de $\mathcal{I}$* se todo elemento de $\hat{\mathcal{P}}$ estiver contido em algum elemento de $\mathcal{P}$. Também escrevemos que $\hat{\mathcal{P}}$ *refina* $\mathcal{P}$ significando o mesmo que $\hat{\mathcal{P}}$ é *refinamento de $\mathcal{P}$*.

Observação 0.4.5. Dizer que $\hat{\mathcal{P}}$ é refinamento de $\mathcal{P}$ é o mesmo que dizer que todo elemento de $\mathcal{P}$ se escreve como união (disjunta) de elementos de $\hat{\mathcal{P}}$. De fato, dado $I_j \in \mathcal{P}$, temos:

$$I = \cup_{\hat{I}_l \in \hat{\mathcal{P}}} \hat{I}_1 \supset I_j \Rightarrow \cup_{\hat{I}_l \in \hat{\mathcal{P}}, \hat{I}_l \cap I_j \neq \emptyset} I_l \supset I_j.$$

Mas como $\hat{\mathcal{P}}$ refina $\mathcal{P}$, temos em particular que fixado $\hat{I}_l \in \hat{\mathcal{P}}$ tal que $\hat{I}_l \cap I_j \neq \emptyset$, existe $i \in \mathbb{N}$ tal que $\hat{I}_l \subset I_i \in \mathcal{P}$, implicando que $I_i \cap I_j \neq \emptyset$ e portanto que $I_j = I_i \supset \hat{I}_l$. Por conseguinte,

$$\cup_{\hat{I}_l \in \hat{\mathcal{P}}, \hat{I}_l \cap I_j \neq \emptyset} I_l \subset I_j,$$

concluindo a prova da observação.

Observação 0.4.6. Dadas duas partições $\mathcal{P}$ e $\hat{\mathcal{P}}$ de um intervalo $[a, b]$ É fácil ver que $\tilde{\mathcal{P}}$ refina $\mathcal{P}$ e $\hat{\mathcal{P}}$. Basta tomarmos o conjunto dos pontos extremos de todos os intervalos que so elementos de $\mathcal{P}$ e $\hat{\mathcal{P}}$, ordená-los obtendo um conjunto finito de reais. Daí, basta considerar os intervalos cujos extremos são pontos consecutivos desse conjunto, tomando-se apenas o cuidado para que sejam disjuntos e sua união seja I.

Proposição 0.4.7. *Sejam I um intervalo compacto, E um espaço de Banach e $f : I \to E$ uma aplicação contínua. Então $\exists \int_I f(x)dx \in E$.*

Prova: Como f é contínua em I compacto, é uniformemente contínua. Seja $\epsilon > 0$ e tome $\delta > 0$ tal que

$$\|f(x) - f(y)\| < \epsilon/(2\,\mathrm{vol}(I)), \forall x, y \in I, d(x, y) < \delta.$$

Sejam $\mathcal{P}$ e $\hat{\mathcal{P}}$ partições quaisquer, com $\mathrm{diam}(\mathcal{P}) < \delta$ e $\mathrm{diam}(\hat{\mathcal{P}}) < \delta$. Seja $\tilde{\mathcal{P}}$ uma partição que refina tanto $\mathcal{P}$ como $\hat{\mathcal{P}}$ (tal partição existe, pela observação 0.4.6). Daí, comparando somas de Riemann em $\mathcal{P}$ e $\tilde{\mathcal{P}}$, obtemos:

$$\|s(\mathcal{P}) - s(\tilde{\mathcal{P}})\| = \|\sum_j^{\#\mathcal{P}} f(x_j) \cdot \mathrm{vol}(I_j) - \sum_j^{\#\tilde{\mathcal{P}}} f(\tilde{x}_j) \cdot \mathrm{vol}(\tilde{I}_j)\|.$$

Para cada $I_j \in \mathcal{P}$, tomemos $\tilde{I}_{j,1}, \ldots, \tilde{I}_{j,r(j)} \in \tilde{\mathcal{P}}$ tais que $I_j = \dot{\cup}_{i=1}^{r(j)} \tilde{I}_{j,i}$. Por conseguinte, reenumerando a soma de Riemann em $\tilde{\mathcal{P}}$, chegamos a

$$\|s(\mathcal{P}) - s(\tilde{\mathcal{P}})\| = \|\sum_j^{\#\mathcal{P}} f(x_j) \cdot \mathrm{vol}(I_j) - \sum_j^{\#\mathcal{P}} (\sum_{i=1}^{r(j)} f(\tilde{x}_{j,i}) \cdot \mathrm{vol}(\tilde{I}_{j,i}))\| \leq$$

$$\sum_j^{\#\mathcal{P}} \|f(x_j) \cdot \mathrm{vol}(I_j) - \sum_{i=1}^{r(j)} f(\tilde{x}_{j,i}) \cdot \mathrm{vol}(\tilde{I}_{j,i}))\| = \sum_j^{\#\mathcal{P}} \|\sum_{i=1}^{r(j)} (f(x_j) - f(\tilde{x}_{j,i})) \cdot \mathrm{vol}(\tilde{I}_{j,i}))\| \leq$$

$$\sum_{j}^{\#\mathcal{P}}\sum_{i=1}^{r(j)}\|f(x_j)-f(\tilde{x}_{j,i})\|\cdot\mathrm{vol}(\tilde{I}_{j,i})\leq\frac{\epsilon}{2\,\mathrm{vol}(I)}\cdot\sum_{j}^{\#\mathcal{P}}\sum_{i=1}^{r(j)}\mathrm{vol}(\tilde{I}_{j,i})=\epsilon/2.$$

Trocando $\mathcal{P}$ por $\hat{\mathcal{P}}$ nas contas acima, temos que $\|s(\hat{\mathcal{P}})-s(\tilde{\mathcal{P}})\|<\epsilon/2$, logo

$$\|s(\mathcal{P})-s(\hat{\mathcal{P}})\|\leq\|s(\mathcal{P})-s(\tilde{\mathcal{P}})\|+\|s(\tilde{\mathcal{P}})-s(\hat{\mathcal{P}})\|<\epsilon,$$

implicando que f é integrável.

□

0.5 Exercícios

1. Seja τ uma topologia de um espaço topológico X. Dizemos que uma coleção de abertos $\mathcal{B}=\{B_\lambda,\lambda\in\Lambda\}$ é uma *base da topologia* de X se todo aberto em τ pode ser expresso como união de abertos em $\mathcal{B}$. Por exemplo, se X é um espaço métrico, por definição o conjunto das bolas abertas constitui uma base de τ. Mostre que todo espaço métrico compacto possui uma base enumerável.

2. Seja M um espaço métrico. Mostre que são equivalentes:

   (i) M contém um subconjunto enumerável denso;

   (ii) M possui uma base enumerável de abertos;

   (iii) (Propriedade de Lindelöf). Toda cobertura aberta de M admite uma subcobertura enumerável.

3. Seja X um conjunto qualquer, e $(E,\|\cdot\|)$ um espaço de Banach. Mostre que o conjunto

$$\mathcal{F}=\mathcal{F}(X):=\{f:X\to\mathbb{R}^k, f \text{ é limitada}\}.$$

   dotado da aplicação $\|\cdot\|_\infty:\mathcal{F}\to[0,+\infty)$ dada por

$$\|f\|_\infty=\sup\{\|f(x)\|,x\in X\}$$

   é um espaço de Banach.

4. Seja X um espaço métrico qualquer e $(E, \|\cdot\|)$ um espaço de Banach. Mostre que o conjunto

$$C_b^0 = C_b^0(X) := \{f : X \to E, f \text{ é contínua e limitada}\}.$$

dotado da aplicação $\|\cdot\|_\infty : C_b^0 \to [0, +\infty)$ dada por

$$\|f\|_\infty = \sup\{\|f(x)\|, x \in X\}$$

é um espaço de Banach.

5. Seja X um espaço métrico compacto e $(E, \|\cdot\|)$ um espaço de Banach. Mostre que o conjunto

$$C^0 = C^0(X) := \{f : X \to \mathbb{R}^k, f \text{ é contínua}\}.$$

dotado da aplicação $\|\cdot\|_\infty : C^0 \to [0, +\infty)$ dada por

$$\|f\|_\infty = \sup\{\|f(x)\|, x \in X\}$$

é um espaço de Banach.

6. Seja E um espaço vetorial. Dizemos que um conjunto C é convexo se dados $x, y \in C$, então $t \cdot x + (1-t) \cdot y \in C, \forall t \in [0, 1]$. Em outras palavras, C é convexo se dados dois quaisquer de seus pontos, o segmento de reta que os une está contido em C. Mostre que em um espaço vetorial normado, toda bola é convexa.

7. Seja X um espaço métrico, Y um espaço métrico completo, e seja $f : X \to Y$ uma aplicação uniformemente contínua. Mostre que f admite uma única extensão contínua $\overline{f}$ a $\overline{X}$; ademais, $\overline{f}$ é uniformemente contínua.

8. Sejam X e Y espaços métricos e seja $f : X \to Y$ uma aplicação localmente lipschitziana. Mostre que para todo conjunto compacto $K \subset X$, $f|_K$ é lipschitziana, isto é, existe $\tilde{M} > 0$ tal que

$$\|f(x) - f(y)\| \leq \tilde{M} \cdot \|x - y\|, \forall x, y \in K.$$

9. Sejam X, Y espaços métricos. Uma aplicação $f : X \to Y$ é dita Hölder contínua se existem constantes $c > 0$ e $0 < \gamma < 1$ tais que

$$\|f(x) - f(y)\| \leq c \cdot \|x - y\|^{\gamma}, \forall x, y \in X.$$

Similarmente, f é dita localmente Hölder contínua se dado $x \in X$, existe uma vizinhança $V_x \ni x$, $V_x \subset X$ tal que $f|_{V_x}$ é Hölder contínua. Seja f uma aplicação localmente Hölder contínua. Mostre que para todo conjunto compacto $K \subset X$, $f|_K$ é Hölder contínua.

10. O que aconteceria se na definição de aplicação Hölder contínua, a constante γ fosse tomada maior que 1?

11. Seja $r > 1$ fixado e considere a aplicação $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) := x^r$. Mostre que f é localmente lipschitziana, mas não lipschitziana. f é Hölder contínua? Se $0 < r < 1$, mostre que f não é lipschitziana. Nesse caso, f é Hölder contínua?

12. Dê exemplos de aplicações Hölder contínuas. Dê exemplos de aplicações lipschitzianas. Dê exemplos de aplicações Hölder contínuas que não são lipschitzianas.

13. Sejam X, Y espaços métricos e sejam $c > 0$, $0 < \gamma < 1$ constantes. Mostre que os conjuntos

$$\{f : X \to Y; f \text{ é Lipschitz com constante de Lipschitz } c\}$$

e

$$\{f : X \to Y; f \text{ é Hölder contínua com constantes de Hölder } c \text{ e } \gamma\}$$

são exemplos de conjuntos equicontínuos de aplicações (uniformemente contínuas).

Capítulo 1

O conceito de EDO

Neste capítulo, ao contrário do restante do nosso texto, faremos uma discussão mais ou menos informal sobre o conceito de Equação Diferencial Ordinária e os dois tipos de problemas que aparecem precocemente na teoria clássica, quais sejam, problemas de Valor Inicial (ou de Cauchy) e problemas de Contorno. A despeito da informalidade com que esses conceitos serão inicialmente introduzidos, procuraremos imediatamente em seguida prover definições rigorosas dos mesmos como é a tônica deste livro.

Considere E um espaço vetorial normado completo (espaço de Banach), por exemplo $E = \mathbb{R}^n$. Seja Ω um aberto conexo em $\mathbb{R} \times E^{k+1}$; um ponto em Ω será denotado por (t, X), com t real e $X = (X_0, X_1, \cdots X_k)$ em E^{k+1}.

Definição 1.0.1. (Solução de EDO). Seja $F : \Omega \to E$ uma aplicação contínua e seja I um intervalo não degenerado da reta. Uma curva k vezes diferenciável $\varphi : I \to E$ chama-se solução da equação diferencial ordinária de ordem k dada formalmente por $F(t, x, \frac{dx}{dt}, \cdots, \frac{d^k x}{dt^k}) = 0$ se e só se:

i) O "gráfico estendido" de φ, isto é, o conjunto $\{(t, \varphi(t), \frac{d\varphi(t)}{dt}, \cdots, \frac{d^k\varphi(t)}{dt^k}),$ com $t \in I\}$ está contido em Ω.

ii) $F(t, \varphi(t), \frac{d\varphi(t)}{dt}, \cdots, \frac{d^k\varphi(t)}{dt^k}) = 0$, $\forall t \in I$. (No caso de t ser um ponto extremo do intervalo I, a derivada acima é a derivada lateral respectiva.)

Observação 1.0.2. Seja $F : \Omega \to E$ uma aplicação contínua e seja I um intervalo não degenerado da reta. Uma função diferenciável $\varphi_0 : I \to E$ é solução da equação diferencial ordinária de ordem k dada formalmente por $F(t, x, \frac{dx}{dt}, \cdots, \frac{d^k x}{dt^k}) = 0$ se e só se:

i) Existem funções $\varphi_i : I \to E^n, 1 \leq i \leq k$, tais que o conjunto

$$\{t, \varphi_0(t), \cdots, \varphi_k(t), \text{ com } t \in I\}$$

está contido em Ω.

ii) $F(t, \varphi_0(t), \cdots, \varphi_k(t)) = 0$, $\forall t \in I$.

iii) $\frac{d\varphi_i(t)}{dt} = \varphi_{i+1}(t)$, com $0 \leq i < k$. (No caso de t ser um ponto extremo do intervalo I, a derivada acima é a derivada lateral respectiva.)

Isso significa que toda EDO de ordem k é equivalente a uma EDO de ordem 1. Para a equação expressa acima, por exemplo, sua correspondente de ordem 1 é dada pela aplicação $G : \Omega \to E^{k+1}$ definida como $G(t, X) = (F(t, X), \frac{dX_0}{dt} - X_1, \cdots, \frac{dX_{k-1}}{dt} - X_k)$.

A definição dada de EDO e suas soluções é contudo excessivamente geral. No sentido de provar os teoremas fundamentais (por exemplo, de existência e unicidade) de nossa teoria, precisamos nos restringir a classe de equações nas quais a derivada de mais alta ordem aparece isolada em um dos lados da equação:

$$\frac{d^k x}{dt^k} = G(t, x, \frac{dx}{dt}, \cdots, \frac{d^{k-1}x}{dt^{k-1}}),$$

onde $G : \Omega' \to E$ é uma aplicação (pelo menos) contínua e $\Omega' \subset \mathbb{R} \times E^k$ é um aberto conexo.

Esse tipo de equação de ordem k, como observamos, é facilmente reduzido a um sistema de equações de ordem 1:

$$\begin{cases} \frac{dX_{k-1}}{dt} = G(t, X_0, \cdots, X_{k-1}) \\ \frac{dX_i}{dt} = X_{i+1}, \text{ para } i = 0 \cdots k-2 \end{cases}$$

Fazendo $X = (X_0, \cdots, X_{k-1})$, escrevemos sinteticamente:

$$\frac{dX}{dt} = \begin{pmatrix} X_1 \\ \vdots \\ X_{k-1} \\ G(t, X) \end{pmatrix} = H(t, X),$$

isto é, qualquer EDO (de qualquer ordem) da classe que estudaremos pode ser reduzida a uma equação de ordem 1 da forma $\frac{dx}{dt} = f(t, x)$, com $f : U \to \tilde{E}$ contínua no aberto $U \subset \mathbb{R} \times \tilde{E}$, onde $\tilde{E}$ é um espaço de Banach.

Observação 1.0.3. (Campos de Vetores). Seja E um espaço de Banach e $V \subset E$ um aberto. Chamamos de *Campo de Vetores em* E qualquer aplicação contínua $g : V \to E$. É óbvio que se temos um campo de vetores g, podemos definir uma aplicação $f : \mathbb{R} \times V \to E$ dada por $f(t, x) := g(x)$, à qual temos associada a Equação Diferencial Ordinária:

$$\frac{dx}{dt} = f(t, x) = g(x).$$

Pela definição de solução dada anteriormente, temos que $\varphi I \to V$ é uma solução da equação acima se, e só se, $\varphi(I) \subset V$ e $\frac{d\varphi}{dt}(t) = g(\varphi(t))$, $\forall t \in I$, onde I é o intervalo (não degenerado) domínio de φ. Qualquer equação diferencial dada por um campo (cuja expressão, portanto, não dependa da variável $t \in \mathbb{R}$) é chamada de *autônoma*. . Do que acabamos de ver, as equações autônomas são um tipo de Equação Diferencial Ordinária. Notamos que toda Equação Diferencial Ordinária do tipo $\frac{dx}{dt} = f(t, x)$ pode também ser reduzida a uma Equação Autônoma. Para tanto, basta isar da substituição $y = (t, x)$, que associa á equação original uma equação autônoma particular (note que com dimensão mai alta): $\frac{dy}{dt} = (1, g(y))$. Desse modo, nosso modelo geral de Equação Ordinária poderia ser afinal as Equações autônomas, já que qualquer outra pode ser reduzida a uma deste tipo. A razão de não o adotarmos como modelo geral logo de início é porque a importante classe das *Equações Diferenciais Ordinárias Lineares* tem sua forma mais natural como equações dependentes do tempo (vide exemplo 1.3.2 na página 31).

Exemplo 1.0.4. Seja $U = I \times \mathbb{R}$, onde I é um intervalo, e $f : U \to \mathbb{R}$ dada por $f(t, x) = g(t)$, onde g é uma função contínua em I. Então, φ é uma solução de $\frac{dx}{dt} = g(t)$ em I se e só se $\varphi(t) = c + \int_{t_0}^{t} g(s)ds$, onde $t_0 \in I$ e c é uma constante.

Exemplo 1.0.5. Seja $U = \mathbb{R}^2$, $f(t, x) = x^2 + 1$. Para todo $t \in \mathbb{R}$, a função $\varphi_c : (-\pi/2 - c, \pi/2 - c) \to \mathbb{R}$ dada por $\varphi_c(t) = \tan(t + c)$ é uma solução da equação $\frac{dx}{dt} = x^2 + 1$. Note que como $\tan(t + c) \to \pm\infty$ quando $t \to \pm\pi/2 - c$, respectivamente, o domínio de φ_c não pode ser prolongado, embora o domínio da equação seja todo o $\mathbb{R}^2$, sem qualquer restrição. Em outras palavras, embora a equação esteja globalmente definida, é possível provar que suas soluções não podem ser estendidas a uma função globalmente definida na reta.

Exemplo 1.0.6. Seja $U = \mathbb{R}^2$, $f(t,x) = 3x^{2/3}$. Para todo $c \in \mathbb{R}$ a função $\varphi_c : \mathbb{R} \to \mathbb{R}$ dada por

$$\varphi_c(t) := \begin{cases} (t-c)^3, & \text{se} \quad t \geq c \\ 0, & \text{se} \quad t < c \end{cases}$$

é uma solução da equação $\frac{dx}{dt} = 3 \cdot x^{2/3}$. Note que, nesse caso, dado $t_0 \in \mathbb{R}$ e para $x_0 = 0$ existe mais de uma curva solução $x(t)$ tal que $x(t_0) = x_0 = 0$. Observe ainda que f é contínua, mas a derivada parcial $\partial_x f(t, x_0)$ não está definida em nenhum ponto tal que $x_0 = 0$.

1.1 O problema de Cauchy

Conforme vimos no exemplo 1.0.4 do final da última seção, mesmo quando a aplicação f que fornece a expressão da EDO estudada depende apenas de t, necessitamos acrescentar mais dados ao problema para termos unicidade das soluções. Naquele exemplo, era necessário que arbitrássemos o valor da solução para algum $t = t_0$ fixado, para termos a unicidade. Isso enseja a seguinte definição:

Definição 1.1.1. (Problema de Valor Inicial, ou problema de Cauchy). Seja U um aberto contido em $I \times E$, onde I é um intervalo não degenerado da reta e E é um espaço vetorial normado completo (espaço de Banach). Seja $f : U \to E$ uma aplicação pelo menos contínua. Fixado um par $(t_0, x_0) \in U$, chamado de *valor inicial* para a equação diferencial ordinária dada por f, chamamos de *problema de Cauchy* associado a f com valor inicial (t_0, x_0) ao problema definido formalmente por:

$$\begin{cases} \frac{dx}{dt} = f(t,x) \\ x(t_0) = x_0, \end{cases}$$

Em outras palavras, dizemos que uma aplicação $\varphi : I \to E$ (onde I é intervalo não degenerado da reta contendo t_0) é uma solução do problema de Cauchy dado por f, com valor inicial (t_0, x_0) se φ é solução da EDO dada por f e ademais, temos $\varphi(t_0) = x_0$.

Como mostra o exemplo 1.0.6, mesmo o problema de Cauchy pode não ter solução única. No próximo capítulo veremos entretanto que hipóteses relativamente gerais sobre a aplicação f darão condições suficientes para

termos existência e unicidade de soluções dos problemas de Cauchy, pelo menos localmente. Por exemplo, se f é Lipschitz, veremos que tais soluções existem e são únicas localmente.

1.2 Problemas de Contorno

Vimos na seção anterior que o problema de Cauchy diz respeito à existência de uma solução de uma EDO $\frac{dx}{dt} = f(t, x)$ que em um momento t_0 dado tenha x_0 (também dado) como valor. Como já antecipamos, sob hipóteses de regularidade da f bastante gerais, existe e é única a solução de um problema de Cauchy. Assim, o problema de Cauchy é um problema (matematicamente) bem posto. Por outro lado, um problema de Contorno para a mesma EDO $\frac{dx}{dt} = f(t, x)$ é dado quando perguntamos por alguma solução cujos valores v e w em dois instantes dados $t_0 \neq t_1$ satisfaçam a uma dada relação $g(v, w) = 0$. Por ter condições tão amplas, um problema de contorno pode ter qualquer número de soluções ou mesmo nenhuma, sendo muitas vezes, (matematicamente) mal posto. Relações que aparecem com mais frequência em problemas de Contorno estão nos exemplos abaixo:

- Para $f : U \subset \mathbb{R} \times E \to E$,

$$\begin{cases} \frac{dx}{dt} = f(t, x) \\ x(t_0) = x(t_1) \end{cases}$$

- Supondo que E se escreva como a soma direta de dois espaços de Banach, $E = \hat{E} \oplus \tilde{E}$, com projeções canônicas $\hat{P} : E \to \hat{E}$ e $\tilde{P} : E \to \tilde{E}$, e $f : U \subset \hat{E} \oplus E \to E$. Outro exemplo típico de problema de Contorno é:

$$\begin{cases} \frac{dx}{dt} = f(t, x) \\ \hat{P}(x(t_0)) = p_0 \\ \hat{P}(x(t_1)) = p_1, \end{cases}$$

  onde p_0 e p_1 são pontos de $\hat{E}$ dados. Outra variação desse exemplo podem ser obtida trocando $\hat{P}$ por $\tilde{P}$ (obviamente, p_0 e p_1 deverão ser dados em $\tilde{P}$, neste caso).

Note que no último exemplo, buscamos curvas soluções que tenham algumas de suas entradas em dois tempos diferentes fixadas, ficando as outras, livres. Mesmo no exemplo mais simples de EDO que vimos, quando f só

depende de t, fica claro que esse tipo de exigência muito provavelmente pode não ser satisfeita ou ser satisfeita mesmo por uma infinidade de soluções. Para alguém que começa a estudar Equações Diferenciais sob um viés matemático, tal pode parecer estranho. Afinal, em problemas de Contorno, alimentamos a equação diferencial com informações provavelmente insuficientes para obtermos uma solução única ou tão demasiadas que talvez não haja nenhuma solução que realize as condições requeridas. Contudo, problemas de Contorno afloram naturalmente na Física, onde as equações são, em sua maioria, de segunda ordem. Se lembrarmos que a aceleração de uma partícula, por exemplo, nada mais é que a derivada segunda de sua posição e uma vez que força é definida como massa vezes aceleração, fica claro que a condição inicial para a equação de movimento de uma partícula na mecânica tem que conter a informação de sua posição e velocidade em um dado instante inicial. Todavia, em algumas situações práticas, como por exemplo, o lançamento de uma sonda espacial a outro planeta (modelado pela Lei da Gravitação Universal de Newton) temos o instante do lançamento e aquele em que desejamos ver a sonda chegar ao outro planeta, bem como as posições de partida (Terra) e de chegada (o outro corpo). Mas o que não sabemos é exatamente qual o vetor velocidade (em módulo e direção) em que devemos lançar a sonda inicialmente, para que uma tal missão seja bem sucedida. Tal contexto é o de um problema de Contorno.

Ao contrário do que ocorre com o problema de Cauchy, não existe uma teoria geral sobre problemas de Contorno. Ou seja, cada equação e problema de Contorno precisa ser estudado particularmente. Apenas no caso de Equações Diferenciais Lineares, a teoria está mais desenvolvida.

1.3 Alguns métodos de Solução de Equações na Reta

Nesta seção, mostraremos alguns exemplos de equações na reta, bem como alguns dos poucos métodos existentes para explicitarmos suas soluções. Como ainda não demonstramos os teoremas de Existência e Unicidade (assunto do próximo capítulo), daremos ênfase à procura do candidato à solução, ainda que seja necessária a adição de hipóteses fortes para encontrá-lo.

Exemplo 1.3.1. (Variáveis separáveis). Seja $f : [a, b] \times [c, d] \to \mathbb{R}$ dada por $f(t, x) = g(t) \cdot h(x)$, com g e h contínuas e $h(x) \neq 0$, $\forall x \in [c, d]$. A equação

diferencial ordinária dada por uma tal f é dita uma *equação a variáveis separáveis*. Para encontrarmos sua solução geral (que ainda não sabemos sequer se existe) procedamos heuristicamente. Suponha que $\varphi : I \to \mathbb{R}$ seja uma tal solução. Daí, temos:

$$\frac{d\varphi(t)}{dt} = g(t) \cdot h(\varphi(t)); \varphi(t_0) = x_0 \Leftrightarrow \frac{1}{h(\varphi(t))} \cdot \frac{d\varphi(t)}{dt} = g(t).$$

Note que definindo $H(x) := \int_{x_0}^{x} \frac{1}{h(s)} ds$, pela regra da cadeia, a última equação acima é o mesmo que:

$$(H \circ \varphi)'(t) = g(t)$$

Aplicando o Teorema Fundamental do Cálculo, integrando ambos os lados da igualdade acima de t_0 a t, obtemos:

$$H(\varphi(t)) - H(\varphi(t_0)) = \int_{t_0}^{t} g(s)ds \Leftrightarrow H(\varphi(t)) = H(x_0) + \int_{t_0}^{t} g(s)ds = \int_{t_0}^{t} g(s)ds.$$

Agora, note que como $1/h$ é a derivada de H, e não muda de sinal, segue-se que H é estritamente monótono (estritamente crescente ou decrescente, conforme $h(x)$ seja sempre positivo ou negativo, respectivamente). Portanto, $H([c, d])$ é um intervalo contendo x_0 e existe a inversa $H^{-1} : H([c, d]) \to [c, d]$. Compondo H^{-1} com cada lado da última igualdade, obtemos:

$$\varphi(t) = H^{-1}(H(x_0) + \int_{t_0}^{t} g(s)ds) = H^{-1}(\int_{t_0}^{t} g(s)ds). \tag{1.1}$$

Portanto, se existe uma solução da equação a variáveis separáveis, ela possui a fórmula acima. Por outro lado, definindo uma aplicação φ pela expressão 1.1, retrocedendo na cadeia de implicações acima, vemos que esta é solução da correspondente equação a variáveis separáveis.

Exemplo 1.3.2. (Equações lineares na reta). Seja $J \subset \mathbb{R}$ um intervalo e $a : J \to \mathbb{R}$, $b : J \to \mathbb{R}$ duas funções contínuas. A *equação linear não homogênea* dada por a e b é a equação:

$$\frac{dx}{dt} = a(t) \cdot x + b(t).$$

Vamos supor que $\varphi : I \to \mathbb{R}$, $I \subset J$, seja uma solução (a determinar) da equação acima, com $\varphi(t_0) = x_0$ para um certo $t_0 \in I$ dado. Daí, agrupando os

termos envolvendo φ de um lado da equação, e os demais termos (conhecidos) do outro, obtemos

$$\frac{d\varphi(t)}{dt} - a(t) \cdot \varphi(t) = b(t).$$

Para determinarmos φ, precisamos fazer aparecer no primeiro lado da equação a derivada de uma expressão conhecida, para aplicarmos a integração em ambos os lados e eliminarmos a derivada com o uso do teorema Fundamental do Cálculo. Neste exemplo, a expressão em questão será a da derivada de um produto de funções, uma das quais, φ. Contudo, para que uma expressão de derivada de produto de funções apareça no primeiro lado da equação necessitamos multiplicar ambos os lados por uma função que não se anule h, a determinar. Dada uma tal função, a equação anterior é equivalente a:

$$h(t)\frac{d\varphi(t)}{dt} - a(t)h(t) \cdot \varphi(t) = b(t)h(t).$$

Para que o primeiro lado da equação seja a derivada $(h(t) \cdot \varphi(t))'$, devemos ter:

$$h'(t) = -a(t)h(t), h(t_0) \neq 0.$$

Como $h(t_0) \neq 0$, podemos supor, pelo menos próximo a t_0, que $h(t)$ não se anula. Por conseguinte, ao menos localmente, a última equação equivale a

$$\frac{1}{h(t)} \cdot h'(t) = -a(t) \Leftrightarrow (\log(h(t)))' = -a(t) \Rightarrow$$

(integrando de t_0 a t, e aplicando o Teorema Fundamental do Cálculo)

$$\log(h(t)) - \log(h(t_0)) = \int_{t_0}^{t} -a(s)ds \Leftrightarrow h(t) = h(t_0) \cdot e^{-\int_{t_0}^{t} a(s)ds}.$$

Note que a expressão obtida para $h(t)$ é válida para todo $t \in I$ e possui o mesmo sinal que $h(t_0)$. Chamemos de $h_0 = h(t_0) \neq 0$. Nossa equação linear original é portanto equivalente a:

$$(h(t) \cdot \varphi(t))' = b(t) \cdot h(t).$$

Integrando de t_0 a t ambos os lados da equação e aplicando mais uma vez o Teorema Fundamental do Cálculo, obtemos:

$$h(t)\cdot\varphi(t) - h_0 \cdot x_0 = \int_{t_0}^{t} b(s)\cdot h(s)ds \Leftrightarrow \varphi(t) = (h(t))^{-1}\cdot(h_0 \cdot x_0 + \int_{t_0}^{t} b(s)\cdot h(s)ds) =$$

(substituindo a expressão de h)

$$\varphi(t) = h_0^{-1} \cdot e^{\int_{t_0}^{t} a(s)ds} \cdot (h_0 \cdot x_0 + \int_{t_0}^{t} b(s) \cdot (h_0 \cdot \int_{t_0}^{s} -a(u)du)ds) =$$

$$e^{\int_{t_0}^{t} a(s)ds} \cdot (x_0 + \int_{t_0}^{t} b(s) \cdot (\int_{t_0}^{s} -a(u)du)ds).$$

Conforme veremos com mais profundidade no próximo capítulo, embora com uma hipótese mais forte que os demais, o seguinte é de longe o mais importante método de solução de Equações Diferenciais Ordinárias na reta:

Exemplo 1.3.3. (Solução por Séries de Taylor). Seja $U = I \times J$, onde I, J são intervalos não degenerados e seja $f : U \to \mathbb{R}$ uma função tal que para toda função analítica $\psi : I \to J$, a função $t \mapsto f(t, \psi(t))$ é analítica (ou seja, é dada localmente em torno de cada ponto $t_0 \in I$ por sua série de Taylor em t_0). Neste caso, se $\varphi : \hat{I} \to J$ é uma solução analítica de

$$\cdot x = f(t, x); x(t_0) = x_0,$$

podemos determiná-la facilmente usando da seguinte técnica:

1. Escrevemos formalmente a série de Taylor de φ em torno de t_0 por

$$\varphi(t) = \sum_{n=0}^{\infty} b_n (t - t_0)^n \Rightarrow \frac{d\varphi}{dt}(t) = \sum_{n=0}^{\infty} (n+1) b_{n+1} (t - t_0)^n,$$

   onde os coeficientes de Taylor, exceto $b_0 = x_0$ ainda estão por determinar (são incógnitos).

2. Escrevemos formalmente a série de Taylor de $t \mapsto f(t, \varphi(t))$ em torno de t_0 como $\sum_{n=0}^{\infty} c_n (t - t_0)^n$, onde cada c_n é função (conhecida) dos coeficientes $b_0, \ldots, b_n$.

3. A unicidade dos coeficientes de Taylor nos permite transformar nossa EDO original em uma infinidade enumerável de equações algébricas:

$$\begin{cases} \frac{dx}{dt} = f(t, x) \\ x(t_0) = x_0 \end{cases} \Leftrightarrow b_0 = x_0; b_{n+1} = c_n/(n+1), \forall n \in \mathbb{N}$$

Em mais detalhes, descrevamos o caso particular em que $f(t,x) = g(x)$, com $g(x) = \sum_{n=0}^{\infty} a_n x^n$ a série de Taylor de g em torno da origem, conhecida. Escrevendo $\varphi(t) = \sum_{n=0}^{\infty} b_n (t-t_0)^n$ temos que φ é solução (analítica) de $\dot{x} = g(x); x(t_0) = x_0$ se, e só se, (localmente, para t suficientemente próximo a t_0) vale

$$\sum_{n=0}^{\infty}(n+1)b_{n+1}(t-t_0)^n = \sum_{n=0}^{\infty} a_n (\sum_{m=0}^{\infty} b_m (t-t_0)^m)^n,$$

o qual ocorre se, e somente se

$$b_0 = x_0; b_1 = \sum_{k=0}^{\infty} a_k b_0^k; \quad b_{n+1} = (\sum_{k=1}^{\infty} a_k (\sum_{m_1+\cdots+m_k=n} b_{m_1}\cdot\cdots\cdot b_{m_k}))/(n+1), \forall n \geq 1.$$

1.4 Exercícios

1. Seja $V \subset \mathbb{R}^n$, $x_0 \in V$ e $g : V \to \mathbb{R}^n$ uma aplicação contínua, com $g(x_0) = 0$. Calcule uma solução para o problema de Cauchy

$$\frac{dx}{dt} = g(x); x(t_0) = x_0.$$

   Dê um exemplo particular mostrando que a solução apresentada pode não ser a única se g for apenas contínua.

2. (Equação da Gravitação de Newton). Desprezando as influências da Lua e dos demais planetas, e considerando o Sol como imóvel na origem dos eixos cartesianos, uma boa aproximação para a posição x (e velocidade) da Terra é dada por:

$$\frac{d^2x}{dt^2} = \frac{GM_{Terra}}{\|x\|^2} \cdot \frac{-x}{\|x\|} = \frac{-GM_{Terra} \cdot x}{\|x\|^3}.$$

   (Claro, a derivada segunda da posição x é a conhecida aceleração). Transforme a equação de ordem 2 acima na equação de ordem 1 correspondente. Escreva o domínio da equação de ordem 1. Note que como Valor Inicial para esta equação, devemos dar a posição e a velocidade da Terra em um dado instante.

3. Calcule a solução das equações diferenciais ordinárias na reta:

$$\begin{cases} \frac{dx}{dt} = x^n, n \in \mathbb{N}, n \geq 1 \\ x(0) = 1 \end{cases}$$

Mostre que a solução é única. Note que somente para $n = 1$ a solução pode ser definida para todo o tempo, embora o domínio de $f(t, x) = x^n$ seja todo o $\mathbb{R}^2$. Exatamente para $n = 1$ o crescimento da derivada é linear.

Capítulo 2

Teoremas de existência e unicidade de soluções

Seja E um espaço de Banach. Devido ao teorema fundamental do cálculo, qualquer equação da forma $\frac{dx}{dt} = f(t,x); x(t_0) = x_0$, com $f : U \to E$ contínua no aberto $U \subset \mathbb{R} \times E$. tem uma correspondente equação integral, isto é , φ é solução do problema de Cauchy

$$\begin{cases} \frac{dx}{dt} = f(t,x) \\ x(t_0) = x_0, \end{cases}$$

definida em um certo intervalo $[t_0 - \alpha, t_0 + \alpha]$, se e só se o gráfico de φ está contido no domínio de f e

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))ds, \forall t \in [t_0 - \alpha, t_0 + \alpha].$$

Como se sabe, ao contrário do operador de derivação, o operador integral é contínuo no espaço $C^0 = C^0[t_0 - \alpha, t_0 + \alpha]$ das funções contínuas $\psi : [t_0 - \alpha, t_0 + \alpha] \to E$, dotado com a norma uniforme $\|\psi\| = \sup_{t \in [t_0 - \alpha, t_0 + \alpha]}\{|\psi(t)|\}$. Assim, tomaremos vantagem da redução de nosso problema diferencial a um problema integral, considerando o operador

$$F(\psi)(t) := x_0 + \int_{t_0}^{t} f(s, \psi(s))ds.$$

Tal operador terá como domínio um subconjunto fechado do espaço das funções contínuas. Note que os pontos fixos de F são justamente as soluções de nossa equação diferencial ordinária.

Com hipóteses bastante gerais sobre f, podemos provar que F ou um seu iterado é uma contração. Pelo teorema do ponto fixo em espaços métricos completos, isso garantirá a existência de um único ponto fixo de F, e portanto de uma única solução (local) para a equação original. Este é o teor do teorema de existência e unicidade de soluções de Picard. Tal ponto fixo será obtido como o limite (em C^0) da sequência de funções $F^n(\psi)$, onde $\psi \equiv x_0$.

Quando a função f que dá a equação é apenas contínua, o teorema de Picard e o fato de que o operador F não piora a regularidade/continuidade das funções em que é aplicado, nos permitirão usar o teorema de Ascoli-Arzelá para mostrar que $F^n(\psi)$ possui subsequências convergentes a algum ponto fixo (Teorema de existência de Peano).

Para provarmos os teoremas deste capítulo relembraremos na próxima seção o *teorema do ponto fixo para contrações* e o *teorema de aproximação de Weierstrass*.

2.1 O Teorema de Picard

Definição 2.1.1. (Aplicação Lipschitz em relação à segunda variável). Sejam E_1, E_2, E_3 espaços de Banach. Uma aplicação $f : U \subset E_1 \times E_2 \to E_3$ (não necessariamente contínua) é *lipschitziana* (ou *Lipschitz*) em relação à segunda variável se $\exists c > 0$ tal que $|f(z, y_1) - f(z, y_2)| \leq c \cdot |y_1 - y_2|$, $\forall (z, y_1), (z, y_2) \in U$. (Observe que c é o mesmo para todo z).

Uma aplicação $f : U \subset E_1 \times E_2 \to E_3$ (não necessariamente contínua) é dita *localmente lipschitziana* (ou *localmente Lipschitz*) em relação à segunda variável se todo ponto de U possui uma vizinhança restrita à qual f é lipschitziana com respeito à segunda variável.

Nosso primeiro objetivo dessa seção é provar o seguinte teorema de existência e unicidade de soluções:

Teorema 2.1.2. *(Picard). Seja E um espaço de Banach e $f : [t_0 - a, t_0 + a] \times \overline{B(x_0, b)} \subset \mathbb{R} \times E \to E$ uma aplicação limitada, contínua, lipschitziana em relação à segunda variável (note que se E tem dimensão finita, a condição de limitada é redundante). Então, existe uma única solução do problema de Cauchy correspondente a f com valores iniciais $x(t_0) = x_0$, definida no intervalo $[t_0 - \alpha, t_0 + \alpha]$, onde $\alpha = \min\{a, b/M\}$ e $M = \sup\{|f(t, x)|, (t, x) \in [t_0 - a, t_0 + a] \times \overline{B(x_0, b)}\}$.*

Prova: Considere o espaço vetorial de aplicações $C^0 = C^0([t_0 - \alpha, t_0 + \alpha]; E)$ dotado da norma uniforme ($\|\psi\| := \sup\{|\psi(x)|, x \in [t_0 - \alpha, t_0 + \alpha]\}$). É fato bem conhecido da análise que C^0 é um espaço de Banach (isto é, normado completo). Seja $\mathcal{C}$ o subconjunto de C^0 dado por $\mathcal{C} := C^0([t_0 - \alpha, t_0 + \alpha], \overline{B(x_0, b)})$, isto é, o conjunto das aplicações contínuas cujo domínio é $[t_0 - \alpha, t_0 + \alpha]$ e a imagem está contida em $\overline{B(x_0, b)}$. É fácil ver que $\mathcal{C}$ é um fechado de C^0. Em particular, como C^0 é espaço de Banach, vem que $\mathcal{C}$ é um espaço métrico completo.

Definamos portanto a aplicação $F : \mathcal{C} \to \mathcal{C}$ dada por

$$F(\psi)(t) := x_0 + \int_{t_0}^{t} f(s, \psi(s))ds.$$

Temos portanto:

1) F está bem definida. De fato, $F(\psi)$ é uma aplicação contínua se ψ o é. Além disso, se $\psi \in \mathcal{C}$, $|F(\psi(t)) - x_0| = |\int_{t_0}^{t} f(s, \psi(s))ds| \leq M \cdot |t - t_0| \leq b$, o que quer dizer que a imagem da aplicação $F(\psi)$ está contida em $\overline{B(x_0, b)}$, se $\psi \in \mathcal{C}$. Logo, F leva aplicações em $\mathcal{C}$ em aplicações em $\mathcal{C}$.

2) Os eventuais pontos fixos de F são soluções do problema de Cauchy com domínio $[t_0 - \alpha, t_0 + \alpha]$.

3) Existe $n_0 \in \mathbb{N}$ tal que F^m é contração, $\forall m > n_0$. De fato, seja c a constante de Lipschitz de f em relação à segunda variável. Por indução, provaremos que $\forall m \in \mathbb{N}$,

$$|F^m(\varphi_1)(t) - F^m(\varphi_2)(t)| \leq \frac{c^m}{m!}|t - t_0|^m d(\varphi_1, \varphi_2), \forall t \in [t_0 - \alpha, t_0 + \alpha], \forall \varphi_1, \varphi_2 \in \mathcal{C}.$$

Realmente, para $m = 1$, temos:

$$|F(\varphi_1)(t) - F(\varphi_2)(t)| \leq \int_{t_0}^{t} |f(s, \varphi_1(s)) - f(s, \varphi_2(s))|ds \leq$$

$$\int_{t_0}^{t} c \cdot d(\varphi_1, \varphi_2)ds \leq c \cdot |t - t_0| d(\varphi_1, \varphi_2).$$

Assumindo a hipótese de indução válida para um certo $m \in \mathbb{N}$, temos:

$$|F^{m+1}(\varphi_1)(t) - F^{m+1}(\varphi_2)(t)| = |F(F^m(\varphi_1)(t)) - F(F^m(\varphi_2)(t))| \leq$$

$$\int_{t_0}^{t} |f(s, F^m(\varphi_1(s))) - f(s, F^m(\varphi_2(s)))| ds \leq$$

(como f é Lipschitz em relação à segunda variável)

$$\int_{t_0}^{t} c|F^m(\varphi_1(s)) - F^m(\varphi_2(s))| ds \leq c \int_{t_0}^{t} \frac{c^m}{m!} |s - t_0|^m \cdot d(\varphi_1, \varphi_2) ds \leq$$

$$\frac{c^{m+1}}{m!} \cdot \int_{t_0}^{t} |s - t_0|^m ds \cdot d(\varphi_1, \varphi_2) \leq$$

$$\frac{c^{m+1}}{m!} \cdot \frac{t - {t_0}^{m+1}}{m+1} d(\varphi_1, \varphi_2) \leq \frac{c^{m+1}}{(m+1)!} |t - t_0|^{m+1} d(\varphi_1, \varphi_2),$$

o que conclui a indução.

Visto que $|t - t_0| \leq \alpha$, temos que

$$d(F^m(\varphi_1), F^m(\varphi_2)) \leq \frac{c^m \alpha^m}{m!} d(\varphi_1, \varphi_2).$$

Como o fatorial domina qualquer exponencial, temos que fixado $0 < \eta < 1$, existe n_0 tal que $\frac{c^m \alpha^m}{m!} < \eta, \forall m \geq n_0$. Portanto, para todo $m \geq n_0$, F^m é uma contração, como queríamos mostrar. Como $\mathcal{C}$ é métrico completo e $F : \mathcal{C} \to \mathcal{C}$, segue-se do Teorema do Ponto fixo que F^m possui um único ponto fixo. Mas se p é o único ponto fixo de F^m, ele também é o único ponto fixo de F, pois

$$F^m(p) = p \Rightarrow F(F^m(p)) = F(p) \Leftrightarrow F^m(F(p)) = F(p),$$

daí $F(p)$ é também ponto fixo de F^m e como este é único temos de ter $F(p) = p$. Como todo ponto fixo de F é também de F^m, segue-se que F só possui este ponto fixo. Como vimos no início desta demonstração isto equivale a existência de uma única solução para o problema de Cauchy.

□

Corolário 2.1.3. *Seja $f : I \times E \to E$ (onde $I \subset \mathbb{R}$ é um intervalo não degenerado, E é um espaço de Banach) uma aplicação contínua, lipschitziana em relação à segunda variável. Então, existe uma única solução do problema de Cauchy correspondente a f com valores iniciais $x(t_0) = x_0$, definida no intervalo $I \subset \mathbb{R}$.*

Prova:

Seja $\mathcal{C}$ o espaço de Banach de funções contínuas dado por $\mathcal{C} := C^0(J, E)$, onde $J \ni t_0$ é um intervalo compacto contido em I. Note que se $J \subset I$ é compacto arbitrário contendo t_0, e mostramos que uma única solução φ_J está definida em J, então está provado nosso corolário: como $I = \cup_{J \subset I, J \ni t_0} J$, apenas colocamos $\varphi : I \to E$ por $\varphi(x) = \varphi_J(x)$, para algum $J \ni x$.

Definamos portanto a aplicação $F : \mathcal{C} \to \mathcal{C}$ dada por

$$F(\psi)(t) := x_0 + \int_{t_0}^{t} f(s, \psi(s))ds$$

É imediato que $F : \mathcal{C} \to \mathcal{C}$ está bem definido.

Exatamente como acima, existe $n_0 \in \mathbb{N}$ tal que F^m é contração, $\forall m > n_0$. De fato, seja c a constante de Lipschitz de f em relação à segunda variável (note que tal constante pode ser restrita a $J \times E$, não precisando ser uniforme em $I \times E$), S o comprimento de J. Por indução, provaremos que $\forall m \in \mathbb{N}$,

$$|F^m(\varphi_1)(t) - F^m(\varphi_2)(t)| \leq \frac{c^m}{m!}|t - t_0|^m d(\varphi_1, \varphi_2), \forall t \in J, \forall \varphi_1, \varphi_2 \in \mathcal{C}.$$

Realmente, para $m = 1$, temos:

$$|F(\varphi_1)(t) - F(\varphi_2)(t)| \leq \int_{t_0}^{t} |f(s, \varphi_1(s)) - f(s, \varphi_2(s))|ds \leq$$

$$\int_{t_0}^{t} c \cdot d(\varphi_1, \varphi_2)ds \leq c \cdot |t - t_0| d(\varphi_1, \varphi_2).$$

Assumindo a hipótese de indução válida para um certo $m \in \mathbb{N}$, temos:

$$|F^{m+1}(\varphi_1)(t) - F^{m+1}(\varphi_2)(t)| = |F(F^m(\varphi_1)(t)) - F(F^m(\varphi_2)(t))| \leq$$

$$\int_{t_0}^{t} |f(s, F^m(\varphi_1(s))) - f(s, F^m(\varphi_2(s)))|ds \leq$$

(como f é Lipschitz em relação à segunda variável)

$$\int_{t_0}^{t} c|F^m(\varphi_1(s)) - F^m(\varphi_2(s))|ds \leq c \int_{t_0}^{t} \frac{c^m}{m!}|s - t_0|^m \cdot d(\varphi_1, \varphi_2)ds \leq$$

$$\frac{c^{m+1}}{m!} \cdot \int_{t_0}^{t} |s - t_0|^m ds \cdot d(\varphi_1, \varphi_2) \leq$$

$$\frac{c^{m+1}}{n!} \cdot \frac{t - {t_0}^{n+1}}{n+1} d(\varphi_1, \varphi_2) \leq \frac{c^{m+1}}{(m+1)!} |t - t_0|^{m+1} d(\varphi_1, \varphi_2),$$

o que conclui a indução.

Visto que $|t - t_0| \leq S$, temos que

$$d(F^m(\varphi_1), F^m(\varphi_2)) \leq \frac{c^m S^m}{m!} d(\varphi_1, \varphi_2).$$

Como o fatorial domina qualquer exponencial, segue-se que para todo $m \geq n_0$, F^m é contração, e portanto o corolário resulta da mesma maneira que no teorema de Picard.

□

A próxima proposição nos dá uma condição suficiente para que uma aplicação seja lipschitziana em relação à segunda variável:

Proposição 2.1.4. *Sejam E_1, E_2, E_3 espaços vetoriais normados e $U \subset E_1 \times E_2$ um aberto convexo. Suponha que $f : U \to E_3$ seja uma aplicação (não necessariamente contínua) com derivada em relação à segunda variável limitada em U. Então f é lipschitziana em relação à segunda variável.*

Prova: Sejam $(w_0, \hat{x}), (w_0, \tilde{x}) \in U$ e seja $c > 0$ uma cota para a norma de $\partial_2 f$ em U. Lembramos que $(\partial_2 f)|_{\{w_0\} \times E_2}$ é igual á derivada total de $f|_{(\{w_0\} \times E_2) \cap U} : (\{w_0\} \times E_2) \cap U \to E_3$. Desse modo, vale a desigualdade do Valor Médio (para tal, necessitamos também da convexidade de U):

$$\|f(w_0, \hat{x}) - f(w_0, \tilde{x})\|_{E_3} \leq c \cdot \|(w_0, \hat{x}) - (w_0, \tilde{x})\|_{E_1 \times E_2} = c \cdot \|\hat{x} - \tilde{x}\|_{E_2},$$

resultando em que f é lipschitziana em relação à segunda variável.

□

2.2 O Teorema de Peano

Teorema 2.2.1. *(Peano). Seja E um espaço de Banach de dimensão finita. Considere $f : [t_0 - a, t_0 + a] \times \overline{B(x_0, b)} \subset \mathbb{R} \times E \to E$ uma aplicação contínua (como E neste enunciado tem dimensão finita, a condição "limitada" presente no teorema de Picard é redundante). Então, existe solução do problema*

de Cauchy correspondente a f com valores iniciais $x(t_0) = x_0$, definida no intervalo $[t_0 - \alpha, t_0 + \alpha]$, onde $\alpha = \min\{a, b/M\}$ e $M = \sup\{|f(t,x)|, (t,x) \in [t_0 - a, t_0 + a] \times \overline{B(x_0, b)}\}$.

Prova: Pelo teorema de aproximação de Weierstrass, existe uma sequência de polinômios $f_j : [t_0 - a, t_0 + a] \times \overline{B(x_0, b)} \to E$ convergindo uniformemente para f. Não há perda em supor que

$$M_j := \sup_{(t,x) \in [t_0 - a, t_0 + a] \times \overline{B(x_0, b)}} \{\|f_j(t,x)\|\} \leq M.$$

De fato, defina a sequência de polinômios (p_j) como $p_j := c_j \cdot f_j$, onde (c_j) é a sequência de reais dada por $c_j := \frac{M}{M_j}$ se $M_j > M$, $c_j := 1$, caso contrário. Daí, temos

$$\|p_j(t,x)\| \leq c_j \cdot M_j \leq M, \forall (t,x) \in [t_0 - a, t_0 + a] \times \overline{B(x_0, b)}.$$

Como $f_j \to f$ uniformemente, temos que $|f_j| \to |f|$ também uniformemente, ou seja, $M_j \to M$. Logo o limite uniforme

$$\lim_{j \to +\infty} p_j = \lim_{j \to +\infty} c_j f_j = f.$$

Por conseguinte, de ora em diante assumiremos sem perda que

$$\|f_j(t,x)\| \leq M, \forall (t,x) \in [t_0 - a, t_0 + a] \times \overline{B(x_0, b)}.$$

Como cada f_j é C^∞ é definida em um compacto, é Lipschitz. Assim, pelo teorema de Picard, existe uma única solução $\varphi_j : [t_0 - \alpha_0, t_0 + \alpha_0] \to \overline{B(x_0, b)}$ do problema de Cauchy

$$\begin{cases} \frac{dx}{dt} = f_j(t,x) \\ x(t_0) = x_0. \end{cases}$$

Lembramos que a família $\{\varphi_j\}$ é equicontínua e equilimitada, pois:

$$|\varphi_j(t) - \varphi_j(\hat{t})| = |\int_{\hat{t}}^{t} f_j(s, \varphi_j(s)) ds| \leq M_1 \cdot |t - \hat{t}|, \forall t, \hat{t} \in [t_0 - \alpha_0, t_0 + \alpha_0],$$

o que também implica equilimitação, porque

$$|\varphi_j(t) - x_0| = |\varphi_j(t) - \varphi_j(t_0)| \leq M \cdot b/M = b.$$

Desse modo, temos pelo Teorema de Ascoli-Arzelá que (φ_j) possui uma subsequência φ_{j_l} uniformemente convergente a uma certa $\psi : [t_0 - \alpha, t_0 + \alpha] \to \overline{B(x_0, b)}$. Para simplificar a notação, chamemos φ_{j_l} simplesmente de φ_l, f_{j_l} de f_l. Mostremos que tal ψ é solução local do problema

$$\begin{cases} \frac{dx}{dt} = f(t,x) \\ x(t_0) = x_0, \end{cases}$$

De fato,

$$|x_0 + \int_{t_0}^{t} f(s, \psi(s))ds - \varphi_l(t)| =$$

$$|x_0 + \int_{t_0}^{t} f(s, \psi(s))ds - x_0 - \int_{t_0}^{t} f_l(s, \varphi_l(s))ds| \leq$$

$$\int_{t_0}^{t} |f_l(s, \varphi_l(s)) - f(s, \psi(s))|ds \leq$$

$$\int_{t_0}^{t} |f_l(s, \varphi_l(s)) - f(s, \varphi_l(s))|ds + \int_{t_0}^{t} |f(s, \varphi_l(s)) - f(s, \psi(s))|ds.$$

A expressão acima vai a zero uniformemente quando fazemos l tender a 0, a primeira parcela porque f_l converge uniformente a f e a segunda porque φ_l converge uniformemente a ψ e f é uniformemente contínua. Em ambas as parcelas, usamos ainda o fato de que o intervalo de integração é limitado. Logo, a sequência φ_l que converge a ψ também converge a $x_0 + \int_{t_0}^{t} f(s, \psi(s))ds$, o que pela unicidade do limite acaba por implicar que $\psi(t) = x_0 + \int_{t_0}^{t} f(s, \psi(s))ds$ e por conseguinte, ψ é solução de nosso problema de Cauchy original.

□

2.3 Intervalo maximal

Os teoremas de existência e unicidade nos permitem (por continuação local) estender soluções locais até um intervalo maximal de tempo. Mais precisamente:

Definição 2.3.1. (Solução maximal). Dada uma EDO $\dot{x} = f(t,x)$, uma solução $\varphi : J \to E$ é dita *maximal* se para toda solução $\psi : J_1 \to E$ com $J_1 \supset J$ e $\psi|_J = \varphi$, tem-se $J_1 = J$. Nesse caso, J é dito *intervalo maximal* de φ.

Proposição 2.3.2. *Considere E um espaço de Banach. Seja $f : U \subset \mathbb{R} \times E \to E$ contínua e tal que $\forall (t_0, x_0) \in U$, existe localmente uma única solução do problema de Cauchy $\dot{x} = f(t, x); x(t_0) = x_0$, definida em um intervalo I_{t_0,x_0}. Então, para todo par $(t_0, x_0) \in U$, existe uma única solução maximal $\varphi(t, (t_0, x_0))$ do problema de Cauchy acima indicado.*

Prova: Definamos o intervalo $I^{\max} = I^{\max}_{t_0,x_0}$ como a sendo a união $\cup_\varphi I_{t_0,x_0}(\varphi)$, de todos os domínios de soluções φ do problema de Cauchy $\dot{x} = f(t, x); x(t_0) = x_0$ do enunciado. Pomos então $\varphi_{\max}(t, (t_0, x_0)) := \psi(t)$, se $t \in I_{t_0,x_0}(\psi)$. Pela hipótese de unicidade local, $\varphi_{\max}(t, (t_0, x_0))$ está bem definida. De fato, sejam φ_1 e φ_2 duas soluções do nosso problema de Cauchy, definidas em intervalos I_1 e I_2, respectivamente, e considere $I_0 := I_1 \cap I_2$. Devemos mostrar que $\varphi(t, t_0, x_0) = \varphi_1(t) = \varphi_2(t), \forall t \in I_0$. Lembramos que um conjunto na reta é conexo se e só se é um intervalo, portanto convexo. Assim, $I_0 = I_1 \cap I_2$ é conexo (e intervalo), pois é a intersecção de dois intervalos (portanto convexos) da reta. Como tanto I_1 como I_2 contém t_0, temos ademais que $I_0 \ni t_0$ é um intervalo não vazio. Considere o conjunto $A \subset I_0$ dado por $A = \{t \in I_0; \varphi_1(t) = \varphi_2(t)\}$. Note que $A \neq \emptyset$, pois $t_0 \in A$. Além disso, se $t_A \in A$, então $\eta := \varphi_1(t_A) = \varphi_2(t_A)$, o que implica que φ_1 e φ_2 são soluções para o problema de Cauchy $\dot{x} = f(t, x); x(t_A) = \eta$, existindo uma vizinhança V de t_A em I_0 pela hipótese de unicidade local onde $\varphi_1|_V = \varphi_2|_V$. Isto implica que A é aberto em I_0. Por outro lado, o conjunto $I_0 \setminus A = \{t \in I_0; \varphi_1(t) \neq \varphi_2\}$ também é aberto em I_0, pois é a pré-imagem do aberto $E \setminus \{0\}$ pela aplicação contínua $\rho := \varphi_1 - \varphi_2$. Como I_0 é intervalo, é conexo, só admitindo a cisão trivial. Logo, $A = I_0$ e $\varphi_1 = \varphi_2$ em I_0, como queríamos mostrar.

Finalmente, como $I^{\max}$ é definida como a união de intervalos (conexos, portanto) com um ponto em comum (t_0), segue-se que $I^{\max}$ é conexo.

De sua definição, é imediato que $\varphi_{\max}$ é maximal: se outra solução ψ coincidir com $\varphi_{\max}$ em $I^{\max}$, satisfará $\psi(t_0) = x_0$ e portanto terá seu domínio $I_{t_0,x_0}(\psi) \subset I^{\max}$, o que implicará que $I_{t_0,x_0}(\psi) = I^{\max}$.

Note ainda mais que qualquer outra solução ψ que coincidir com $\varphi_{\max}$ em um ponto t_1, terá seu domínio $I(\psi)$ contido em $I^{\max}$ e teremos $\psi(t) = \varphi_{\max}(t), \forall t \in I(\psi)$. De fato, pelos mesmos argumentos de conexidade acima, temos que $\psi(t) = \varphi_{\max}(t), \forall t \in I(\psi) \cap I^{\max}$. Assim, podemos supor que $t_0 \notin I(\psi)$ (caso contrário, nada teremos a mostrar). Também não há perda de generalidade em supormos $t_1 > t_0$ (o caso $t_1 < t_0$ é análogo). Se, por absurdo, $I(\psi)$ não estiver contido em $I^{\max}$ então existe $t_2 > t_1$ tal que $t_2 \in I(\psi) \setminus I^{\max}$.

Neste caso, podemos definir a aplicação

$$\tilde{\psi}(t) := \begin{cases} \varphi_{\max}(t), & \text{se } t \in I^{\max} \setminus (t_1, +\infty) \\ \psi(t), & \text{se } t \in [t_1, t_2]), \end{cases}$$

Note que o domínio de $\tilde{\psi}$ contém propriamente $I^{\max}$ e que $\tilde{\psi}(t_0) = \varphi(t_0) = x_0$. Ademais, temos que os limites laterais

$$\tilde{\psi}(t_1) = \overset{+}{\lim_{t \to t_1}} \tilde{\psi}(t) = \psi(t_1) = \overset{-}{\lim_{t \to t_1}} \varphi_{\max}(t) = \overset{-}{\lim_{t \to t_1}} \tilde{\psi}(t),$$

o que implica a continuidade de $\tilde{\psi}$ em t_1. Do mesmo modo, prova-se a diferenciabilidade, e que $\tilde{\psi}$ é solução do problema de Cauchy $\frac{dx}{dt} = f(t, x), x(t_0) = x_0$:

$$\overset{+}{\lim_{t \to t_1}} \frac{d\tilde{\psi}}{dt}(t) = \overset{+}{\lim_{t \to t_1}} \frac{d\psi}{dt}(t) = f(t_1, \psi(t_1)) = f(t_1, \varphi_{\max}(t_1)) = \overset{-}{\lim_{t \to t_1}} \varphi(t) = \overset{-}{\lim_{t \to t_1}} \tilde{\psi}(t).$$

Para os demais valores de $t \in I^{\max} \cup [t_1, t_2]$, a continuidade e a diferenciabilidade de $\tilde{\psi}$ são triviais, assim como o fato de que $\tilde{\psi}$ é solução da equação diferencial em questão. Portanto, $\tilde{\psi}$ estende propriamente a solução maximal $\varphi_{\max}$, o que contradiz a a maximalidade desta última.

□

Dado um intervalo aberto I diremos que W é uma *vizinhança de seus extremos* se $W = I \setminus J$, onde J é um intervalo compacto contido em I.

Proposição 2.3.3. *Suponha E um espaço de Banach de dimensão finita. Seja $f : U \subset \mathbb{R} \times E \to E$, uma aplicação contínua tendo como domínio um aberto U e suponha que todo problema de Cauchy definido por f tenha solução local (portanto, maximal) única. Seja $\varphi : I \to E$ uma solução maximal de um problema de Cauchy da forma $\frac{dx}{dt} = f(t, x)$, $x(t_0) = x_0$ com $(t_0, x_0) \in U$ e $I = (\omega_-, \omega_+) \subset \mathbb{R}$ seu intervalo (maximal) de definição. Então o gráfico $\Phi : I \to U$ dado por $t \to (t, \varphi(t))$ da solução maximal satisfaz à seguinte propriedade de convergência: para todo compacto $K \subset U$, existe uma* vizinhança W dos extremos de I *tal que $t \in W \Rightarrow \Phi(t) \notin K$. Escrevemos esta condição como $\Phi(t) \to \partial U$ quando $t \to \omega_\pm$.*

Prova: Se um dos extremos de I for infinito, nada temos a demonstrar. Assim, vamos supor que $\omega_\pm$ são finitos. Seja $K \subset U$ um compacto e sem perda de generalidade suponhamos por absurdo que exista uma sequência

$t_j \in I$ com $t_j \to \omega_+$, tal que $\Phi(t_j) \in K, \forall j \in \mathbb{N}$. Podemos então supor que $\Phi(t_j)$ converge em K, digamos $\Phi(t_j) \to (\omega_+, x_1) \in K$.

Aplicamos o teorema de Peano a $\dot{x} = f(t,x)$, $x(\omega_+) = x_1$ e obtemos uma vizinhança $V = [\omega_+ - \alpha, \omega_+ + \alpha] \times \overline{B(x_1, b)}$, $V \subset U$, na qual o gráfico da solução do problema acima está definido. Restringindo V a $\tilde{V} = [\omega_+ - \alpha/3, \omega_+ + \alpha/3] \times \overline{B(x_1, b/2)}$, temos que para quaisquer dados iniciais $(\tilde{t}, \tilde{x}) \in \tilde{V}$, existe solução para o problema de Cauchy definida em $[\tilde{t} - \alpha/2, \tilde{t} + \alpha/2]$. De fato, seja $\alpha = b/M$, onde $|f| < M$ em V; aplicando-se Peano a valores iniciais $(\tilde{t}, \tilde{x}) \in \tilde{V}$ com o domínio da equação restrito a vizinhança $V' = [\tilde{t} - \alpha/2, \tilde{t} + \alpha/2] \times \overline{B(\tilde{x}, \alpha M/2)}$, temos que $V' \subset V$ e que Peano nos garante uma solução definida para o $\min\{\alpha/2, (\alpha M/2)/M\} = \alpha/2$.

Agora, para todo j suficientemente grande, temos $(t_j, \varphi(t_j)) \in \tilde{V}$, daí, fixemos um tal j, assumindo sem perda de generalidade que $t_j > t_0$. Temos que $\omega_+ - t_j < \alpha/3$. Contudo, vimos que a solução local ψ de $\dot{x} = f(t,x); x(t_j) = \varphi(t_j)$ está definida até $t_j + \alpha/2 \geq \omega_+ - \alpha/3 + \alpha/2 > \omega_+$. Mas ψ e φ são soluções de um mesmo problema de Cauchy, o que implica que o domínio da solução maximal φ precisa englobar o domínio (da solução maximal) de ψ, excedendo (ω_-, ω_+), o que é um absurdo. Mais precisamente, definindo-se

$$\tilde{\psi}(t) := \begin{cases} \varphi(t), & \text{se } t \in (\omega_-, t_j) \\ \psi(t), & \text{se } t \in [t_j, \omega_+ - \alpha/3 + \alpha/2)\,, \end{cases}$$

Do mesmo modo que no fim da proposição 2.3.2, temos que $\tilde{\psi}$ estende propriamente a solução maximal φ, o que é absurdo.

□

Corolário 2.3.4. *Seja $V \subset \mathbb{R}^n$ e $f : \mathbb{R} \times V \to \mathbb{R}^n$ uma aplicação contínua tal que qualquer problema de Cauchy definido por f possua solução única. Seja $\varphi : I \to V$ a solução maximal do problema de Cauchy $\dot{x} = f(t,x), x(t_0) = x_0$, com $(t_0, x_0) \in \mathbb{R} \times V$. Se $\varphi(I)$ está contido em um subconjunto compacto C de V, então $I = \mathbb{R}$.*

Prova: Suponha que $I = (\omega_-, \omega_+)$ e sem perda de generalidade, que tivéssemos $\omega_+ < +\infty$. Mas tal implicaria que $\{(t, \varphi(t)), t \in (t_0, \omega_+)\}$ estaria contido no compacto $[t_0, \omega_+] \times C =: K \subset \mathbb{R} \times V$. Tal é absurdo, pois vimos na prova da proposição que dado um tal compacto K e qualquer sequência (t_j), com $t_j \in I$ e $t_j \to \omega_+$, existe $j_0 \in \mathbb{N}$ tal que $(t_j, \varphi(t_j)) \notin K, \forall j \geq j_0$.

□

Proposição 2.3.5. *Suponha E um espaço de Banach com dimensão finita. Seja $f : U \subset \mathbb{R} \times E \to E$ contínua e limitada no aberto U. Se ω_+ (respectivamente, ω_-) é finito, então existe $\lim_{t\to\omega_+} \varphi(t)$ (respectivamente, $\lim_{t\to\omega_-} \varphi(t)$).*

Prova: Suponha sem perda de generalidade que $\omega_+ < +\infty$. Então, $\forall t, s \in (\omega_-, \omega_+)$ vale

$$\|\varphi(t) - \varphi(s)\| = \| \int_s^t f(u, \varphi(u))du\| \leq M \cdot |t - s|,$$

onde $\|f\| \leq M$ em U.

Portanto, se $t_j \to \omega_+$, temos $\|\varphi(t_j) - \varphi(t_m)\| < M|t_j - t_m| \to 0$, quando $j, m \to \infty$, logo $(\varphi(t_j))$ é de Cauchy; como t_j é arbitrária, segue-se que existe o limite do enunciado.

□

Corolário 2.3.6. *Suponha E um espaço de Banach de dimensão finita. Seja $f : \mathbb{R} \times E \to E$ contínua e limitada, tal que qualquer solução de problema de Cauchy dado por f seja (localmente) única. Então, as soluções maximais estão definidas para todo tempo.*

Prova: Seja φ uma solução maximal de

$$\dot{x} = f(t, x), x(t_0) = x_0$$

definida no intervalo maximal (ω_-, ω_+). Suponha sem perda de generalidade que $\omega_+ < +\infty$.

Seja ψ como a solução local do problema de Cauchy com valores iniciais $x(\omega_+, \lim_{t\to\omega_+} \varphi(t))$. Daí, defina $\tilde{\varphi}$ por

$$\tilde{\varphi}(t) := \begin{cases} \varphi(t), \text{ se } t \in (\omega_-, \omega_+) \\ \psi(t), \text{ se } t \in [\omega_+, \omega_+ + \alpha] \end{cases}$$

Da proposição anterior, $\tilde{\varphi}$ é contínua em ω_+ (portanto, é contínua). Além disso, similarmente ao fim da proposição 2.3.2, temos:

$$\overline{\lim_{t\to\omega_+}} \frac{d\tilde{\varphi}(t)}{dt} = \overline{\lim_{t\to\omega_+}} \frac{d\varphi(t)}{dt} = \overline{\lim_{t\to\omega_+}} f(t, \varphi(t)) = f(\omega_+, \psi(\omega_+)) =$$

$$\overline{\lim_{t\to\omega_+}} f(t, \psi(t)) = \overline{\lim_{t\to\omega_+}} \frac{d\psi(t)}{dt} = \overline{\lim_{t\to\omega_+}} \frac{d\tilde{\varphi}(t)}{dt},$$

o que implica pelo teorema do valor intermediário (para derivadas) que a derivada de $\tilde{\varphi}$ existe em ω_+ e que $\tilde{\varphi}$ é solução da EDO, estendendo propriamente uma solução maximal, o que é uma contradição.

□

O corolário 2.3.6 possui uma prova alternativa, aplicando-se o Teorema de Peano à caracterização de soluções maximais presente no início da demonstração da proposição 2.3.2. Basicamente, vimos em 2.3.2 que o domínio da solução maximal do problema de Cauchy $\dot{x} = f(t,x), x(t_0) = x_0$ contém o domínio de qualquer outra solução do mesmo problema. Como f ĺimitada, digamos, por $M > 0$, se $\omega_+ < +\infty$, podemos tomar$b := M \cdot (|t_0| + |\omega_+| + 1)$ e $a := |t_0| + |\omega_+| + 1$. Aplicando o teorema de Peano ao problema de Cauchy

$$\dot{x} = f|_{[t_0-a,t_0+a]\times\overline{B(x_0,b)}}(t,x), x(t_0) = x_0,$$

obtemos uma solução do problema de Cauchy cujo domínio é um intervalo cujo extremo superior ultrapassa o da solução maximal, o que é absurdo. Os detalhes desta prova são deixados ao leitor (exercício 3).

2.4 Exercícios

1. Seja $I \subset \mathbb{R}$ um intervalo não degenerado, e seja E um espaço de Banach. Defina a integral de Riemann no espaço formado pelas funções contínuas $f : I \to E$.

2. Enuncie e prove o teorema Fundamental do Cálculo no contexto de funções $f : I \to E$ de classe C^1, onde E é um espaço de Banach.

3. Dê uma prova alternativa do corolário 2.3.6 usando do Teorema de Peano.

4. Sejam $f, g : U \subset \mathbb{R}^2 \to \mathbb{R}$ duas funções contínuas tais que para qualquer valor inicial $(t_0, x_0) \in U$ a solução de cada um dos problemas de Cauchy $\dot{x} = f(t,x), x(t_0) = x_0$, $\dot{x} = g(t,x), x(t_0) = x_0$ é única. Suponha que $g(t,x) > f(t,x)$, $\forall (t,x) \in U$. Mostre que dado $(t_1, x_1) \in U$, se $\psi : [t_1, t_2] \to \mathbb{R}$ é solução de $\dot{x} = g(t,x), x(t_1) = x_1$ e $\varphi : [t_1, t_2] \to \mathbb{R}$ é solução maximal de $\dot{x} = f(t,x), x(t_1) = x_1$, então $\psi(t) \geq \varphi(t)$, $\forall t \in [t_1, t_2]$. Mostre que o mesmo vale se supusermos que $g(t,x) \geq f(t,x)$.

5. Dizemos que uma aplicação $\varphi = (\varphi_1, \ldots, \varphi_m) : V \subset \mathbb{C} \to \mathbb{C}^m$ é holomorfa no aberto V se para todo $z_0 \in U$ existe o limite:

$$\frac{d\varphi(z_0)}{dz} := \lim_{z-z_0} \frac{\varphi(z) - \varphi(z_0)}{z - z_0}.$$

Isto é obviamente equivalente às funções componentes $\varphi_1, \ldots, \varphi_m$ serem funções holomorfas em U, conceito estudado em cursos elementares de funções analíticas. O conceito de soluções de EDO pode ser transposto sem mudança essencial para o contexto holomorfo, trocando a derivada ordinária real por uma derivada holomorfa ordinária. Seja então $f : U \subset \mathbb{C}^{m+1} \to \mathbb{C}^m$, U aberto, tal que para toda aplicação holomorfa $\varphi : V \subset \mathbb{C} \to \mathbb{C}^m$, com $graf(\varphi) \subset U$, a composta

$$z \mapsto f(z, \varphi(z))$$

é também uma aplicação holomorfa. Mostre que existe, é única e holomorfa a solução local do problema de Cauchy:

$$\frac{dw}{dz} = f(z, w) w(z_0) = w_0.$$

(Sugestão: Seja $\overline{B(z_0, a)} \times \overline{B(w_0, b)}$ e $\alpha := \min\{a, b/(\sup_{z \in \overline{B(z_0,a)}}\{\|f(z, w)\|\})\}$. Defina

$$\mathcal{C} := \{\psi : \overline{B(z_0, \alpha)} \to \overline{B(w_0, b)}, \psi \text{ é holomorfa.}\}$$

e o operador $F : \mathcal{C} \to \mathcal{C}$ dado por

$$(F\psi)(z) := \int_0^1 f(z_0 + s(z - z_0), \psi(z_0 + s(z - z_0))) * (z - z_0) ds,$$

onde $*$ denota o produto por escalar complexo. Mostre que o operador acima está bem definido, e que converge uniformemente para um ponto fixo, o qual é solução do problema de Cauchy.

6. (Teorema de Peano usando poligonais). Seja $E = \mathbb{R}^n$, dotado com a norma do máximo (cujas bolas são quadrados). Considere $f : [t_0 - a, t_0 + a] \times \overline{B(x_0, b)} \subset \mathbb{R} \times E \to E$ uma aplicação contínua. Seja $\alpha = \min\{a, b/M\}$ e $M = \sup\{|f(t, x)|, (t, x) \in [t_0 - a, t_0 + a] \times \overline{B(x_0, b)}\}$. Mostre que:

(a) Dado $\delta > 0$, seja $\Upsilon := \{[t_j, t_j + 1), j = 0 \dots m-1\} \cup \{[t_m, t_0 + \alpha]\}$ uma partição de Riemann com diâmetro $< \delta$ do intervalo $[t_0, t_0 + \alpha]$. Mostre que a aplicação poligonal $\varphi_\Upsilon : [t_0, t_0 + \alpha] \to \overline{B(x_0, b)}$ dada por $\varphi(t_0) = x_0$ e

$$\varphi|_{[t_j, t_{j+1}]}(t) := \varphi(t_j) + f(t_j, \varphi(t_j)) \cdot (t - t_j)$$

está bem definida. Em particular, conclua que

$$\{\varphi_\Upsilon, \Upsilon \text{ é partição de Riemann de } [t_0, t_0 + \alpha]\}$$

é uma coleção equilimitada de curvas poligonais.

(b) Dada uma curva φ_Υ definida no item anterior, tome $\mathcal{P}$ uma partição de Riemann de $\overline{B(x_0, b)}$ com diâmetro $< \delta$, e considere a partição $\wp := \{I_j \times P_k, I_j \in \sigma, P_k \in \mathcal{P}\}$ de $[t_0, t_0 + \alpha] \times \overline{B(x_0, b)}$. Defina $f_\Upsilon : [t_0, t_0 + \alpha] \times \overline{B(x_0, b)} \to E$, constante em cada elemento $I_j \times P_k$ da partição $\wp$ por:

$$f_\Upsilon|_{I_j \times P_k} \equiv \begin{cases} f(t_j, \varphi_\Upsilon(t_j)), & \text{se} \quad \varphi_\Upsilon([t_0, t_0 + \alpha]) \cap I_j \times P_k \neq \emptyset; \\ f(\hat{t}, \hat{x}), & \text{caso contrário, onde } (\hat{t}, \hat{x}) \text{ é qualquer} \\ & \text{ponto escolhido em } I_j \times P_k. \end{cases}$$

Mostre que φ_Υ satisfaz a equação integral:

$$\varphi_\Upsilon(t) = x_0 + \int_{t_0}^{t_0 + \alpha} f_\Upsilon(s, \varphi_\Upsilon(s)) ds.$$

Conclua, em particular, que a coleção $\{\varphi_\Upsilon\}$ é equilipschitz, com constante de Lipschitz M. Em particular, $\{\varphi_\Upsilon\}$ é equicontínua.

(c) Mostre que $f_\Upsilon \to f$ quando $\delta \to 0$. Isso que dizer que, dado $\epsilon > 0$, existe $\delta_0 > 0$ tal que se $0 < \delta < \delta_0$ é o δ que acota o diâmetro das partições dos itens acima, então $\|f_\Upsilon(t, x) - f(t, x)\| < \epsilon, \forall (t, x) \in [t_0, t_0 + \alpha] \times \overline{B(x_0, b)}$. (Sugestão: use a continuidade uniforme da f, tomando $\delta_1 > 0$ tal que

$$\|(t, x) - (\hat{t}, \hat{x})\| < \delta_1 \Rightarrow \|f(t, x) - f(\hat{t}, \hat{x})\| < \epsilon,$$

para quaisquer $(t, x), (\hat{t}, \hat{x}) \in [t_0, t_0 + \alpha] \times \overline{B(x_0, b)}$.
Depois, tome $\delta_0 = \delta_1/(M+1)$ e verifique que $\|f_\Upsilon(t, x) - f(t, x)\| < \epsilon$ como afirmado.)

(d) Aplique o Teorema de Ascoli-Arzelá à coleção

$$\mathfrak{F} = \{\varphi_\Upsilon; \Upsilon \text{ é partição de } [t_0, t_0 + \alpha] \text{ com diâmetro} < \delta_0\}$$

para obter uma sequência uniformemente convergente $(\varphi_k), \varphi_k = \varphi_{\Upsilon_k} \in \mathfrak{F}$, com $\text{diam}(\Upsilon_k) \to 0$ quando $k \to +\infty$. Seja $\varphi = \lim_{k\to+\infty} \varphi_k$, a qual é contínua, pois é limite uniforme de curvas contínuas. Seja $f_k = f_{\Upsilon_k}$. Use os itens anteriores para mostrar que

$$\varphi_k(t) = x_0 + \int_{t_0}^{t} f_k(s, \varphi_k(s))ds \to x_0 + \int_{t_0}^{t} f(s, \varphi(s))ds,$$

quando $k \to +\infty$, uniformemente em $t \in [t_0, t_0 + \alpha]$. Conclua que φ satisfaz a equação integral

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))ds, \forall t \in [t_0, t_0 + \alpha],$$

donde obtemos que $\varphi : [t_0, t_0 + \alpha] \to E$ é solução da EDO

$$\begin{cases} \frac{dx}{dt} = f(t, x) \\ x(t_0) = x_0 \end{cases}$$

Note que apenas por razões didáticas (simplicidade de notação) consideramos $[t_0, t_0 + \alpha]$ no lugar de $[t_0 - \alpha, t_0 + \alpha]$. Assim, acabamos de obter uma nova prova do Teorema de Peano.

Observação 2.4.1. Por conseguinte, vimos que existe solução do problema de Cauchy correspondente a f com valores iniciais $x(t_0) = x_0$, definida no intervalo $[t_0 - \alpha, t_0 + \alpha]$, e que esta pode ser obtida pelo limite de uma sequência de poligonais da questão acima. Estas poligonais aparecem no conhecido método numérico de Euler de solução de EDO's.

Capítulo 3

Dependência das soluções em relação às condições iniciais e parâmetros.

3.1 Dependência contínua

Lema 3.1.1. *Considere $f_j : U \to \mathbb{R}^n$ uma sequência de aplicações contínuas no aberto $U \subset \mathbb{R} \times \mathbb{R}^n$, com $f_j \to f_0$ uniformemente em partes compactas de U. Dada uma sequência $(t_j, x_j) \in U$, com $(t_j, x_j) \to (t_0, x_0)$, se o problema de Cauchy*

$$\begin{cases} \frac{dx}{dt} = f_k(t,x) \\ x(t_k) = x_k, \quad k \geq 0 \end{cases}$$

possui solução maximal única φ_k em J_k, então $\forall [a,b] \subset J_0$, existe $k_0 = k_0(a,b) \geq 1$ tal que $j \geq k_0 \Rightarrow J_j \supset [a,b]$ e $\varphi_j|_{[a,b]} \to \varphi_0|_{[a,b]}$ uniformemente.

Prova: Sejam $\{(t, \varphi_0(t)), t \in [a,b]\} \subset \text{int}(\tilde{K}) \subset \tilde{K} \subset \text{int}(K) \subset K \subset U$, onde $\tilde{K}$ e K são compactos. Daí, existe constante M tal que $|f_k| < M, k \geq 0$ em K, pela convergência uniforme de f_j em partes compactas.

Por Peano, $\exists \alpha > 0$ tal que para todo $(\tilde{t}, \tilde{x}) \in \tilde{K}$, o problema de Cauchy $\dot{x} = f_k(t,x); x(\tilde{t}) = \tilde{x}$ possui uma solução definida em $[\tilde{t} - \alpha, \tilde{t} + \alpha]$, com gráfico contido em K. (Por exemplo, tome $\alpha = d(\tilde{K}, K^c)/(M+1)$).

Seja $\delta = \alpha/3$. Então, existe $k_\delta \in \mathbb{N}$ tal que $\forall j \geq k_\delta$, temos $(t_j, x_j) \in \tilde{K}$ e $|t_j - t_0| < \delta = \alpha/3$.

Não há perda de generalidade em supor que $t_0 \in [a,b]$. Comecemos então

por mostrar que a sequência (com $j \geq k_\delta$) $\varphi_j|_{[t_0-\delta,t_0+\delta]}$ converge uniformemente para $\varphi_0|_{[t_0-\delta,t_0+\delta]}$.

De fato, basta ver que a família de funções $\mathcal{F} = \{\varphi_j|_{[t_0-\delta,t_0+\delta]}, j \geq k_\delta\}$ é equicontínua e equilimitada. Ela é equilimitada porque seus gráficos estão todos contidos no compacto K e é equicontínua porque possuem constante de Lipschitz menor ou igual a M:

$$|\varphi_j(t) - \varphi_j(s)| = |\int_s^t f(u, \varphi_j(u)du| \leq \int_s^t |f(u, \varphi_j(u))|du \leq M|t-s|.$$

Por Ascoli-Arzelá, toda sequência em $\mathcal{F}$ possui subsequência convergente. Logo, para mostrar que $\varphi_j|_{[t_0-\delta,t_0+\delta]}$ converge uniformemente para $\varphi_0|_{[t_0-\delta,t_0+\delta]}$, basta mostrar que toda subsequência convergente $\varphi_{j_l}|_{[t_0-\delta,t_0+\delta]}$ tem como limite $\varphi_0|_{[t_0-\delta,t_0+\delta]}$. Como estamos sob a hipótese de unicidade de soluções dos problemas de Cauchy, basta mostrarmos que a aplicação limite $\psi := \lim_{l\to\infty} \varphi_{j_l}|_{[t_0-\delta,t_0+\delta]}$ é solução do problema

$$\begin{cases} \frac{dx}{dt} = f_0(t,x) \\ x(t_0) = x_0 \end{cases}$$

Realmente,

$$|x_0 + \int_{t_0}^t f_0(s,\psi(s))ds - \varphi_{j_l}(t)| =$$

$$\|x_0 + \int_{t_0}^t f_0(s,\psi(s))ds - x_{j_l} - \int_{t_{j_l}}^t f_{j_l}(s,\varphi_{j_l}(s))ds\| \leq$$

$$\|x_j - x_0\| + \int_{t_0}^{t_{j_l}} \|f_0(s,\psi(s))\|ds + \int_{t_{j_l}}^t \|f_{j_l}(s,\varphi_{j_l}(s)) - f_0(s,\psi(s))\|ds \leq$$

$$\|x_{j_l} - x_0\| + |t_{j_l} - t_0| \cdot M +$$

$$\int_{t_{j_l}}^t \|f_{j_l}(s,\varphi_{j_l}(s)) - f_0(s,\varphi_{j_l}(s))\|ds + \int_{t_j}^t \|f_0(s,\varphi_{j_l}(s)) - f_0(s,\psi(s))\|ds.$$

Vemos que $\int_{t_{j_l}}^t |f_{j_l}(s,\varphi_{j_l}(s)) - f_0(s,\varphi_j(s))|ds$ vai a zero quando $l \to \infty$ porque $f_{j_l} \to f$ uniformemente, e que $\int_{t_{j_l}}^t |f_0(s,\varphi_{j_l}(s)) - f_0(s,\psi(s))|ds$ vai a zero porque f_0 é uniformemente contínua e $\varphi_{j_l} \to \psi$ uniformemente.

Portanto, $\varphi_{j_l}(t) \to x_0 + \int_{t_0}^t f_0(s, \psi(s))ds$, e como $\varphi_{j_l}(t) \to \psi(t)$ por hipótese, segue-se que

$$\psi(t) = x_0 + \int_{t_0}^t f_0(s, \psi(s))ds,$$

sendo então ψ solução do problema de Cauchy como dissemos acima. Como tal problema tem solução única, e a subsequência convergente tomada é arbitrária, temos $\varphi_0|_{[t_0-\delta,t_0+\delta]} = \psi = \lim_{j\to\infty} \varphi_j|_{[t_0-\delta,t_0+\delta]}$.

Isto finda a prova do lema para $t; |t - t_0| \leq \delta$. Para estender o resultado ao intervalo $[a, b]$, aplicamos o mesmo raciocínio acima **sucessivamente** a intervalos $[t_0, t_0+2\delta]$, $[t_0+\delta, t_0+3\delta], \ldots [t_0+v\delta, b]$, com $t_0+w\delta, \varphi_0(t_0+w\delta)$ e $t_k+w\delta, \varphi_k(t_k+w\delta)$ em lugar de t_0, x_0 e t_j, x_j respectivamente, com $1 \leq w \leq v$ e $t_0+(v+1)\delta \geq b$. Daí, para $j \geq k_v \geq k_\delta$, obtemos que o gráfico de $\varphi_j|_{[t_0-\delta,b]}$ está contido em K e que $\varphi_j|_{[t_0-\delta,b]} \to \varphi_0|_{[t_0-\delta,b]}$ uniformemente. Aplicando novamente o mesmo raciocínio acima para $t < t_0$ e w com valores negativos, obtemos que para $j \geq k_0 \geq k_v \geq k_\delta$, o gráfico de $\varphi_j|_{[a,b]}$ está contido em K e que $\varphi_j|_{[a,b]} \to \varphi_0|_{[a,b]}$.

□

Usaremos o lema acima para a prova do seguinte

Teorema 3.1.2. *(Dependência contínua com respeito às condições iniciais e parâmetros). Seja $f : U \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ uma aplicação contínua no aberto U. Suponhamos que $\forall (t_0, x_0, z) \in U$, o problema de Cauchy correspondente admita uma única solução $\varphi = \varphi(t, t_0, x_0, z)$, definida em um intervalo maximal denotado por (ω_-, ω_+), onde $\omega_\pm = \omega_\pm(t_0, x_0, z)$. Então:*

(i) O conjunto

$$D := \{(t, t_0, x_0, z); (t_0, x_0, z) \in U, t \in (\omega_-(t_0, x_0, z), \omega_+(t_0, x_0, z))\}$$

é aberto em $\mathbb{R} \times U$.

(ii) $\varphi : D \to \mathbb{R}^n$ é contínua.

Observação 3.1.3. Note que com a mudança de variáveis $y = (x, z)$, o sistema

$$\begin{cases} \frac{dx}{dt} = f(t, x, z) \\ x(t_0) = x_0 \end{cases}$$

fica:

$$\begin{cases} \frac{dy}{dt} = (f(t, y), 0) \\ y(t_0) = (x_0, z) = y_0 \end{cases}$$

As soluções do primeiro problema são a projeção das n primeiras coordenadas das soluções do segundo.

Assim, o teorema 3.1.2 é consequência imediata do

Teorema 3.1.4. *(Dependência contínua com respeito às condições iniciais). Seja $f : U \subset \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ uma aplicação contínua no aberto U. Suponhamos que $\forall(t_0, x_0) \in U$, o problema de Cauchy correspondente admita uma única solução $\varphi = \varphi(t, t_0, x_0)$, definida em um intervalo maximal denotado por (ω_-, ω_+), onde $\omega_\pm = \omega_\pm(t_0, x_0)$. Então:*

(i) O conjunto

$$D := \{(t, t_0, x_0); (t_0, x_0) \in U, t \in (\omega_-(t_0, x_0), \omega_+(t_0, x_0))\}$$

é aberto em $\mathbb{R} \times U$.

(ii) $\varphi : D \to U$ é contínua.

Prova:

(i) Seja $(t, t_0, x_0) \in D$ fixado. Daí, existe $[a, b] \subset (\omega_-(t_0, x_0), \omega_+(t_0, x_0))$ tal que $t \in (a, b)$. Aplicando o lema com $f_j = f_0 = f$, obtemos que existe uma vizinhança $\tilde{U} \subset U$ de (t_0, x_0) tal que $\forall(\tilde{t}_0, \tilde{x}_0) \in \tilde{U}$, temos $[a, b] \subset (\omega_-(\tilde{t}_0, \tilde{x}_0), \omega_+(\tilde{t}_0, \tilde{x}_0))$. Em particular, o aberto $(a, b) \times \tilde{U} \subset D$ contém o ponto fixado (arbitrário) (t, t_0, x_0), donde D é aberto.

(ii) Sejam (t, t_0, x_0) e (s, s_0, y_0) em D. Temos:

$$|\varphi(t, t_0, x_0)-\varphi(s, s_0, y_0)| \leq |\varphi(t, t_0, x_0)-\varphi(s, t_0, x_0)|+|\varphi(s, t_0, x_0)-\varphi(s, s_0, y_0)|$$

Dado $\epsilon > 0$, pela continuidade de $\varphi(\cdot, t_0, x_0)$, temos que existe $\delta_0 > 0$ tal que $|t - s| \leq \delta_0 \Rightarrow |\varphi(t, t_0, x_0) - \varphi(s, t_0, x_0)| < \epsilon/2$. Por outro lado, novamente do lema, temos que existe δ_1 tal que se $|(t_0, x_0) - (s_0, y_0)| < \delta_1$, então $(\omega_-(s_0, y_0), \omega_+(s_0, y_0)) \supset [t - \delta_0, t + \delta_0]$. Existe ainda $0 < \delta_2 < \delta_1$ tal que, da convergência uniforme dada pelo lema, $|(t_0, x_0) - (s_0, y_0)| < \delta_2$ implica que $|\varphi(s, t_0, x_0) - \varphi(s, s_0, y_0)| < \epsilon/2$, $\forall s \in [t - \delta_0, t + \delta_0]$.

Seja $\delta = \min\{\delta_0, \delta_2\}$, Tomando a norma do máximo em nossos espaços, temos que, se $|(t, t_0, x_0) - (s, s_0, y_0)| < \delta$ então $|t - s| \leq \delta_0$ e $|(t_0, x_0) - (s_0, y_0)| < \delta_2$, implicando que

$$|\varphi(t, t_0, x_0) - \varphi(s, s_0, y_0)| \leq$$

$$|\varphi(t,t_0,x_0)-\varphi(s,t_0,x_0)|+|\varphi(s,t_0,x_0)-\varphi(s,s_0,y_0)|<\epsilon,$$

isto é, $\varphi : D \to U$ é contínua.

□

Apesar da relevância de se saber da continuidade das soluções em relação às condições iniciais e parâmetros, convém também obter cotas para a velocidade com que duas soluções possam se separar ao longo do tempo. Isso é obtido através de hipóteses adicionais sobre a aplicação f que expressa a equação diferencial. Basicamente, é possível mostrar que se $f = f(t,x)$ é Lipschitz em relação a x, então para intervalos compactos de tempo, as soluções são Lipschitz em relação a x_0. Para obtermos tal resultado, necessitamos do seguinte lema:

Lema 3.1.5. *(Desigualdade de Gronwall). Sejam $\alpha \geq 0$ uma constante, $u, v : I \to \mathbb{R}$ duas funções positivas contínuas tais que*

$$u(t) \leq \alpha + \int_{t_0}^{t} u(s)v(s)ds, \forall t \in I.$$

Então vale

$$u(t) \leq \alpha \cdot e^{\int_{t_0}^{t} v(s)ds}.$$

Em particular, se $\alpha = 0$, $u \equiv 0$.

Prova: Comecemos com $\alpha > 0$. Defina φ por $\varphi(t) = \alpha + \int_{t_0}^{t} u(s)v(s)ds$. Daí, pelo teorema fundamental do cálculo, $\frac{d\varphi}{dt} = u(t) \cdot v(t)$. Mas por hipótese, $u(t) \leq \varphi(t)$, logo $\frac{d\varphi}{dt} = u(t) \cdot v(t) \leq \varphi(t)v(t)$.

Como $\alpha > 0$ e u, v são maiores ou iguais a zero, segue-se que φ é estritamente maior que zero. Daí, podemos escrever:

$$\frac{\dot{\varphi}(t)}{\varphi(t)} \leq v(t) \Rightarrow \frac{d(\log \varphi(t))}{dt} \leq v(t) \Rightarrow \int_{t_0}^{t} \frac{d(\log \varphi(s))}{dt} ds \leq \int_{t_0}^{t} v(s)ds,$$

o que implica, pelo teorema fundamental do cálculo que

$$\log(\varphi(t)) - \log(\varphi(t_0)) \leq \int_{t_0}^{t} v(s)ds \Rightarrow \log(\frac{\varphi(t)}{\varphi(t_0)}) \leq \int_{t_0}^{t} v(s)ds \Rightarrow$$

$$\varphi(t) \leq \varphi(t_0) \cdot e^{\int_{t_0}^{t} v(s)ds}.$$

Como $\varphi(t_0) = \alpha$ e $u(t) \leq \varphi(t)$ por hipótese, segue-se o resultado para $\alpha > 0$.

Para o caso $\alpha = 0$, basta considerar $\alpha = \epsilon > 0$ no caso já provado e tomar o limite quando $\epsilon \to 0$.

□

Proposição 3.1.6. *Seja $f : U \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ contínua, e lipschitziana com respeito à segunda variável, digamos $|f(t,x) - f(t,y)| \leq K|x-y|$. Então, $\forall t \in I_{t_0,x_0} \cap I_{t_0,y_0}$ temos*

$$|\varphi(t,t_0,x_0) - \varphi(t,t_0,y_0)| \leq e^{K|t-t_0|}|x_0 - y_0|$$

Prova: De fato, se $u_1(t) = \varphi(t,t_0,x_0)$ e $u_2(t) = \varphi(t,t_0,y_0)$, temos

$$|u_1(t) - u_2(t)| = |x_0 - y_0 + \int_{t_0}^{t} f(s,u_1(s)) - f(s,u_2(s))ds| \Rightarrow$$

$$|u_1(t) - u_2(t)| \leq |x_0 - y_0| + \int_{t_0}^{t} K \cdot |(u_1 - u_2)(s)|ds,$$

o que implica, pela desigualdade de Gronwall (lema 3.1.5) com $\alpha = |x_0 - y_0|$, $u(t) = |u_1(t) - u_2(t)|$ e $v(t) = K$, supondo $t \geq t_0$, que

$$|\varphi(t,t_0,x_0) - \varphi(t,t_0,y_0)| \leq e^{K|t-t_0|}|x_0 - y_0|.$$

Se $t \leq t_0$, temos que

$$|u_1(t) - u_2(t)| \leq |x_0 - y_0| + \int_{t}^{t_0} K \cdot |(u_1 - u_2)(s)|ds,$$

valendo o mesmo resultado.

□

3.2 Dependência diferenciável

Para a prova do principal resultado dessa seção, necessitamos do seguinte lema:

Lema 3.2.1. *Seja $f : (a,b) \times A \to \mathbb{R}^n$, contínua, onde A é um aberto convexo de $\mathbb{R}^n$. Se f admite derivada parcial em relação à segunda variável $D_2 f$, contínua em $(a,b) \times A$, então existe uma função $g : (a,b) \times A \times A \to \mathcal{L}(\mathbb{R}^n)$, contínua, tal que*

$$g(t,x,x) = D_2 f(t,x), \forall (t,x) \in (a,b) \times A \text{ e}$$

$$f(t,x_2) - f(t,x_1) = g(t,x_1,x_2) \cdot (x_2 - x_1).$$

Prova: Simplesmente defina $g(t, x_1, x_2) := \int_0^1 D_2 f(t, sx_2 + (1-s)x_1)ds$. g é claramente contínua.

Aplicando o teorema fundamental do cálculo a $h(s) := f(t, sx_2 + (1-s)x_1)$, temos que

$$h(1) - h(0) = f(t, x_2) - f(t, x_1) = \int_0^1 \frac{d(f(t, sx_2 + (1-s)x_1))}{ds} ds.$$

Aplicando a regra da cadeia à derivada dentro da integral acima obtemos:

$$f(t, x_2) - f(t, x_1) = \int_0^1 D_2 f(t, sx_2 + (1-s)x_1) \cdot (x_2 - x_1) ds =$$

$$\int_0^1 D_2 f(t, sx_2 + (1-s)x_1) ds \cdot (x_2 - x_1) = g(t, x_1, x_2) \cdot (x_2 - x_1).$$

□

Teorema 3.2.2. *(Dependência diferenciável com respeito às condições iniciais). Seja $f : U \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ uma função contínua no aberto U, diferenciável com relação à variável x, sendo $\frac{\partial f}{\partial x}$ contínua em U. Então, como consequência do teorema de Picard, vimos que $\forall (t_0, x_0) \in U$, o problema de Cauchy correspondente admite uma única solução maximal $\varphi = \varphi(t, t_0, x_0)$, com t tomando valores em um intervalo maximal I_{t_0,x_0}. Seja D o aberto dado no teorema 3.1.4, tal que $\varphi : D \to U$. Então, existe e é contínua a derivada $\partial_{x_0}\varphi(t, t_0, x_0)$, $\partial_{x_0}\varphi : D \to \mathcal{L}(\mathbb{R}^n)$. Além disso, tal derivada é a solução da seguinte equação diferencial ordinária matricial linear:*

$$\begin{cases} \dot{Z} = \frac{\partial f}{\partial x}(t, \varphi(t, t_0, x_0)) \cdot Z \\ Z(t_0) = I_n \leftarrow \text{ matriz identidade } n \times n \end{cases}$$

Prova: Seja $[a, b] \subset I_{t_0,x_0}$. Como D é aberto e $[a, b] \times \{(t_0, x_0)\}$ é compacto, temos que existe vizinhança convexa $V_0 \subset U$ de (t_0, x_0) tal que $[a, b] \times \overline{V}_0 \subset D$. Em particular, temos $(a, b) \times V_0 \subset D$. Ainda estamos supondo que a distância entre $[a, b] \times \overline{V}_0$ e ∂U seja maior que uma constante δ_0.

Restrinjamos então φ a $(a, b) \times V_0$. Para mostrarmos que φ tem derivada contínua em relação a x_0, nada mais natural do que considerar os quocientes de Newton

$$z_h = z_h(t) = \frac{\varphi(t, t_0, x_0 + he_k) - \varphi(t, t_0, x_0)}{h}, k = 1 \ldots n, 0 < |h| < \delta_0.$$

Fixemos $k \in \{1 \dots n\}$. Cada função z_h acima é contínua. Assim, se provarmos que z_h converge uniformemente (em t) para uma aplicação z_0, esta será contínua. Além disso, se $z_h(t)$ converge a $z(t)$, quando $h \to 0$, como z_h é o quociente de Newton, $z(t)$ não passa da $k-$ésima derivada parcial no espaço de x_0.

Seja $g : (a, b) \times V_0 \times V_0 \to \mathcal{L}(\mathbb{R}^n)$ tal que $g(t, x_1, x_2) \cdot (x_2 - x_1) = f(t, x_2) - f(t, x_1)$, g dada pelo lemma 3.2.1 acima.

Note que z_h é solução (única) da equação $\dot{z} = g(t, \varphi(t, t_0, x_0), \varphi(t, t_0, x_0 + he_k) \cdot z$, com valor inicial $z(t_0) = e_k$. Fazendo $h \to 0$, temos da dependência contínua em relação às condições iniciais e *parâmetros* que as soluções z_h convergem uniformemente em $[a, b]$ para a solução z_0 da equação limite

$$\dot{z} = D_2 f(t, \varphi(t, t_0, x_0)) \cdot z, z(t_0) = e_k.$$

Mas como vimos, uma vez que z_h converge, por definição seu limite é a derivada parcial $\partial_{x_{0,k}} \varphi(t, t_0, x_0)$. Este limite $z_0 = z_0(t, t_0, x_0)$ é contínuo (não apenas em t) pelo teorema de dependência contínua em relação a condições iniciais e parâmetros. Variando k, obtemos um sistema de equações lineares que pode ser expresso matricialmente por:

$$\dot{Z} = \partial_2 f(t, \varphi(t, t_0, x_0)) \cdot Z, Z(t_0) = I_n,$$

cuja solução é, como mostramos acima, $\partial_{x_0} \varphi(t, t_0, x_0)$, a qual existe e é contínua, devido à continuidade de $\partial_{x_{0,k}} \varphi(t, t_0, x_0)$; e isto finda a prova do teorema.

□

3.3 Exercícios

1. Seja $U \subset \mathbb{R}^n$ um aberto e considere $f : U \to \mathbb{R}$ uma aplicação de classe C^2. Seja $a = (a_1, \dots, a_n) \in U$ tal que $f(a) = b$, com $\partial_1 f(a) \neq 0$. Comece por considerar a equação com $n-2$ parâmetros $x_3, \dots, x_n$ dada por
$$\frac{dx_2}{dx_1} = -\partial_2 f(x_1, x_2, \dots, x_n)/\partial_1 f(x_1, x_2, \dots x_n),$$
inicialmente com condições iniciais $x_2(a_1) = a_2$ e parâmetros $(x_3, \dots, x_n) = (a_3, \dots, a_n)$. Enuncie e demonstre uma versão C^2 do teorema da função

implícita para f em uma vizinhança de a, usando os teoremas de Picard e de dependência diferenciável em relação às condições iniciais e parâmetros.

2. Seja $U \subset \mathbb{R}^n$ um aberto e considere $f : U \to \mathbb{R}$ uma aplicação de classe C^1. Seja $a = (a_1, \ldots, a_n) \in U$ tal que $f(a) = b$, com $\partial_1 f(a) \neq 0$. Mostre, usando compacidade, que existe uma vizinhança $I \times W \ni a$, onde $I \subset \mathbb{R}$ e $W \subset \mathbb{R}^{n-1}$ são abertos convexos, tal que existe uma única $\varphi : W \to I$ contínua com a propriedade de que

$$(x_1, \ldots, x_n) \in I \times W, f(x_1, \ldots, x_n) = c \Leftrightarrow (x_1, \ldots, x_n) \in graf(\varphi).$$

Use esse fato para provar o exercício anterior com $f \in C^1$, agora usando o teorema de Peano.

Capítulo 4

Campos de Vetores

Definição 4.0.1. (Campo de vetores). Um *campo de vetores* em um aberto $U \subset \mathbb{R}^n$ é uma aplicação contínua $X : U \to \mathbb{R}^n$. Tal campo define uma EDO autônoma (isto é, que não depende de t): $\dot{x} = X(x)$.

Definição 4.0.2. (Nomenclatura geral). Seja $X : U \to \mathbb{R}^n$ um campo de vetores. Um ponto $p \in U$ tal que $X(p) = 0$ é dito *ponto singular* ou *singularidade* de X. Analogamente, se $X(p) \neq 0$, p é dito *ponto regular* de X. As soluções $\dot{x} = X(x)$ são chamadas de *soluções, trajetórias* ou *curvas integrais* de X. O conjunto imagem de uma solução maximal φ é dito *órbita* de φ (às vezes denotada com a mesma letra φ). Uma órbita φ de X é dita *periódica* se a solução $\varphi(t)$ é periódica.

Observação 4.0.3. Lembramos que qualquer equação do tipo

$$\dot{x} = f(t, x), x(t_0) = x_0$$

pode ser transformada em uma equação autônoma pela substituição

$$y = (t, x) \Rightarrow \dot{y} = (1, f(y)), y(t_0) = (t_0, x_0).$$

A solução da equação autônoma assim obtida é o gráfico da solução da equação original. Note que o campo $(1, f(y))$ é sem singularidades.

Do teorema de Picard, se X é localmente lipschitziano, $\forall x_0 \in X$ existe uma única solução da EDO dada por $\dot{x} = X(x)$, $x(t_0) = x_0$, definida em $[t_0 - \alpha, t_0 + \alpha]$.

De ora em diante até o fim do livro, suporemos sempre X de classe C^k, $k \geq 1$. Isto nos assegurará não apenas a existência e unicidade de soluções,

como também a diferenciabilidade em relação às condições iniciais vista na seção anterior.

O objetivo principal da presente seção é classificar o comportamento local das soluções dos campos, próximo a um ponto regular.

Comecemos entretanto com o seguinte resultado:

Teorema 4.0.4. *(Teorema do Fluxo Local). Seja $X : U \to \mathbb{R}^n$ um campo de vetores $C^k, k \geq 1$ no aberto U. Então:*

(i) $\forall x_0 \in U$, existe a solução maximal $\varphi(\cdot, x_0)$ de $\dot{x} = X(x), x(0) = x_0$, definida no intervalo maximal I_{x_0}.

(ii) O conjunto $D = \{(t, x_0) \in \mathbb{R} \times U, t \in I_{x_0}\}$ é aberto.

(iii) A aplicação $\varphi : D \to \mathbb{R}^n$, definida por $(t, x_0) \mapsto \varphi(t, x_0)$ é de classe C^k. Além disso, temos a relação

$$\frac{d(D_{x_0}\varphi(t, x_0))}{dt} = DX(\varphi(t, x_0)) \cdot D_{x_0}\varphi(t, x_0), \text{ com } (D_{x_0}\varphi)|_{(0,x_0)} = I_n.$$

(iv) (Propriedade de fluxo). Se $y_0 = \varphi(t_0, x_0)$, com $t_0 \in I_{x_0}$, então $I_{y_0} = I_{x_0} - t_0$ e $\varphi(s, y_0) = \varphi(s, \varphi(t_0, x_0)) = \varphi(t_0 + s, x_0)$.

Prova: Apenas (iv) é novidade. (i), (ii), (iii) são adaptações dos resultados de Picard e de dependência contínua e diferenciável para equações dadas por campos de vetores.

Assim, mostremos que $\varphi(s, y_0) = \varphi(t_0 + s, x_0), \forall s \in I_{y_0}$. Começaremos supondo $s \in I_{y_0} \cap (I_{x_0} - t_0)$ (note que tal intersecção é um aberto não vazio, porque intervalos maximais são abertos e porque 0 pertence a tal intersecção).

Para mostrar que as curvas acima são as mesmas, basta ver que satisfazem o seguinte problema de Cauchy, que sabemos ter solução única por Picard (teorema 2.1.2):

$$\dot{x} = X(x), x(0) = y_0.$$

De fato, $\varphi(0, y_0) = y_0 = \varphi(t_0 + 0, x_0)$ e derivando temos:

1) $\frac{d\varphi(s,y_0)}{ds} = X(\varphi(s, y_0))$

2) $\frac{d\varphi(t_0+s,x_0)}{ds} = X(\varphi(t_0 + s, x_0))$.

Como X é de classe C^1, segue-se do teorema de Picard que $\varphi(s, y_0)$ e $\varphi(t_0 + s, x_0)$ são a mesma aplicação. Em particular, I_{y_0} coincide com $I_{x_0} - t_0$: por um lado, I_{y_0} contém $I_{x_0} - t_0$, pois $\varphi(t_0 + s, x_0)$ (ou uma sua justaposição com $\varphi(s, y_0)$) estende $\varphi(s, y_0)$ como solução e I_{y_0} é maximal. Por outro lado,

está contida, Para ver essa outra inclusão, defina $\psi(t) := \varphi(t - t_0, y_0)$. O domínio de ψ é $I_{y_0} + t_0$, e note que tanto ψ como $\varphi(\cdot, x_0)$ são soluções de

$$\dot{x} = X(x), x(t_0) = y_0.$$

Como $\varphi(\cdot, x_0)$ é maximal, concluímos que

$$I_{x_0} \supset I_{y_0} + t_0 \Rightarrow I_{x_0} - t_0 \supset I_{y_0}.$$

□

Observação 4.0.5. Note que os itens (iii) e (iv) do teorema do Fluxo local implicam que se $(t, x_0) \in D$, existe uma vizinhança V de x_0 em U tal que $\varphi_t|_V : V \to \varphi_t(V) \subset U$ é um difeomorfismo. De fato, se $t \in I_{x_0}$, como D é aberto, temos que existe uma vizinhança V de x_0 em U tal que $t \in I_x, \forall x \in V$. O item (iii) nos diz que φ_t é C^k, e o item (iv) nos diz que $\varphi_{-t}|_{\varphi_t(V)} : \varphi_t(V) \to V$ está definido, sendo a inversa de $\varphi_t|_V$.

Definição 4.0.6. (Fluxo local). O *fluxo local* ou *grupo a um parâmetro* gerado por X é a aplicação $\varphi : D \to \mathbb{R}^n$. O campo X é dito *completo* se $I_x = \mathbb{R}; \forall x \in U$, ou seja, se $D = \mathbb{R} \times U$.

Observação 4.0.7. Dos teoremas sobre intervalo maximal temos que se $X : \mathbb{R}^n \to \mathbb{R}^n$ é limitado ($\exists M > 0$ tal que $X(x) \leq M, \forall x \in \mathbb{R}^n$), então X é completo. Outro caso em que um campo $X : \mathbb{R}^n \to \mathbb{R}^n$ é completo é se ele for (globalmente) Lipschitz.

Denotemos por $Diff^k(U)$ o grupo dos difeomorfismos de classe C^k do aberto $U \subset \mathbb{R}^n$, dotado da operação de composição.

O próximo corolário justifica a terminologia de grupo a um parâmetro da última definição:

Corolário 4.0.8. *Seja $X : U \subset \mathbb{R}^n \to \mathbb{R}^n$ um campo de vetores completo de classe $C^k, k \geq 1$. Então o fluxo de X, $\varphi : \mathbb{R} \times U \to U$ satisfaz:*

(i) $\frac{\partial(\varphi(t,x))}{\partial t} = X(\varphi(t,x)); \varphi(0,x) = x$. Em particular, $X(x) = \frac{\partial(\varphi)}{\partial t}|_{(0,x)}$.

(ii) $\varphi(s, \varphi(t,x)) = \varphi(s + t, x)$.

(iii) Fixado t, a aplicação $\varphi_t(x) = \varphi(t,x)$ define um difeomorfismo C^k. Além disso, a aplicação $\Phi : \mathbb{R} \to Diff^k(U)$, dada por $\Phi(t) = \varphi_t$ define um homomorfismo de grupos.

Proposição 4.0.9. *Seja $X : U \subset \mathbb{R}^n \to \mathbb{R}^n$ um campo de classe C^k, $k \geq 1$. Dada uma solução máxima φ de X, definida em I, temos três alternativas:*

(i) φ é constante e $I = \mathbb{R}$;

(ii) φ é injetiva;

(iii) $I = \mathbb{R}$ e φ é periódica.

Em particular, se $\gamma = \varphi(I)$ é a órbita correspondente, então temos as seguintes possibilidades para γ:

(i) $\gamma = \{p\}$;

(ii) γ é difeomorfa à reta.

(iii) γ é difeomorfa a um círculo.

Prova: Tomamos o caso em que φ não é injetiva. Então, existe $s > t_1$ tais que $\varphi(s) = \varphi(t_1)$. Podemos supor que $t_1 = 0$; caso este não seja o caso, basta que provemos a proposição para $\tilde{\varphi}(t) := \varphi(t + t_1)$, pois $\tilde{\varphi}$ e φ têm a mesma imagem e se uma das propriedades ((i),(ii),(iii)) vale para $\tilde{\varphi}$, também vale para φ. Assim, suponha que $t_1 = 0$. Seja $t \in I$. Então, existe $n \in \mathbb{Z}$ tal que $h = t - n \cdot s \in [0, s]$. Do teorema do fluxo local temos:

$$\varphi_t(x_0) = \varphi_{h+n\cdot s}(x_0) = \varphi_h \varphi_{n\cdot s}(x_0) = \varphi_h(x_0),$$

o que implica que $\gamma = \varphi(I) \subset \varphi([0, s])$. Como a outra inclusão é imediata, segue-se que $\varphi(I) = \varphi([0, s])$ é compacta, já que é imagem do compacto $[0, s]$ pela função contínua φ. Da proposição 2.3.3 e corolários, acerca de soluções maximais, segue-se que $I = \mathbb{R}$. Ademais, seja C o conjunto dos períodos de φ dado por $C := \{c \in \mathbb{R}, \varphi(t + c) = \varphi(t), \forall t \in \mathbb{R}\}$. Como φ é contínua, temos que C é fechado: se $C \ni c_j \to c$, então para cada t temos $\varphi(t) = \varphi(t + c_j) \to \varphi(t + c)$, o que implica $\varphi(t) = \varphi(t + c)$, isto é, $c \in C$. Finalmente, C é um subgrupo aditivo da reta: dados dois períodos $c_1, c_2 \in C$, $c_1 + c_2$ também é período de φ, logo pertence a C. É resultado de análise na reta que os únicos grupos fechados aditivos da reta são a própria reta ou grupos da forma $\tau \cdot \mathbb{Z}$, para algum $\tau > 0$.

Se $C = \mathbb{R}$, então $\varphi(t) = \varphi(t + c), \forall c \in \mathbb{R}$, o que implica que $X(\varphi(t)) = X(\varphi(0)) = \frac{\varphi(t)}{dt} = 0$, isto é, $\gamma = \{\varphi(0)\}$ corresponde a uma singularidade de X.

Se C é da forma $\tau \cdot \mathbb{Z}$, então para $0 \leq t < \tau$, φ é injetiva (caso contrário, existiria um período de φ menor que τ, o que seria comprovável usando-se a propriedade de grupo, como já fizemos mais acima). Em vista do parágrafo acima, é claro que $\varphi'(t)$ não pode se anular. Daí, dado $x = e^{i2\pi\theta} \in S^1, 0 <$

$\theta < 1$ defina o difeomorfismo $f : S^1 \to \gamma$ por $f(x) = \varphi(\theta(x) \cdot \tau), f(1) = \varphi(0)$. Daí, temos:

1) f é difeomorfismo entre $S^1 \setminus \{1\}$ e $\gamma \setminus \varphi(0)$, já que $\theta|(0,1)$ e $\varphi|(0,\tau)$ são difeomorfismos (pois são difeomorfismos locais e são globalmente injetivos); f é uma bijeção entre S^1 e γ.

2) f é contínua em 1: se $S^1 \ni x \to 1$, temos que θ se aproxima de 0 ou de 1; $f(1) = \varphi(0 \cdot \tau) = \varphi(1 \cdot \tau)$, segue-se que f é contínua. Em particular, como é bijeção entre compactos, segue-se que f é homeomorfismo. Para provarmos com rigor a continuidade da f em $1 \in S^1$, tomemos uma parametrização h de uma vizinhança de 1 em S^1, por exemplo, $h : (-\delta, \delta) \to S^1$ dada por $h(t) = e^{i2\pi t}$. Escrevendo

$$f|_{h((-\delta,\delta))} = (f \circ h) \circ h^{-1},$$

como h^{-1} é contínua, vemos que f será contínua em 1 se $f \circ h$ o for em zero. Mas temos que:

$$f(h(t)) = \varphi((\theta(h(t)) \cdot \tau) = \begin{cases} \varphi(t \cdot \tau), 0 \leq t < \delta \\ \varphi((1+t) \cdot \tau), 0 > t > -\delta \end{cases}$$

Como φ é periódica de período τ, segue-se que $\varphi((1+t)\cdot\tau) = \varphi(\tau + t \cdot \tau) = \varphi(t \cdot \tau)$.

Quando $t \to 0^+$, vale que $f(h(t)) = \varphi(t \cdot \tau) \to \varphi(0) = f(h(0))$. Por outro lado, se $t \to 0^-$, obtemos $f(h(t)) = \varphi((1+t) \cdot \tau) \to \varphi(\tau) = \varphi(0) = f(h(0))$.

3) Para ver que f é difeomorfismo, só resta mostrar que existe $f'(1)$ e este é não nulo. De fato,

$$\lim_{t \to 0^+} \frac{f \circ h(t) - f \circ h(0)}{t} = \lim_{t \to 0^+} \frac{\varphi(t \cdot \tau) - \varphi(0)}{t} = \tau \cdot \varphi'(0);$$

por outro lado,

$$\lim_{t \to 0^-} \frac{f \circ h(t) - f \circ h(0)}{t} = \lim_{t \to 0^-} \frac{\varphi((1+t) \cdot \tau) - \varphi(0)}{t} = \tau \cdot \tau\varphi'(\tau) = \tau \cdot \varphi'(0),$$

onde usamos a periodicidade de φ' na última igualdade. Concluímos que $(f \circ h)$ é derivável em 0, o mesmo valendo (por definição de derivadas em variedades) para f em 1.

É claro que $X(\varphi(t)) = \varphi'(t) \neq 0$, pois caso contrário, teríamos $\varphi \equiv 0$ (substituindo na equação e usando a existência e unicidade garantidas por

Picard), e estamos supondo φ não constante. Também usando a equação, e repetindo um raciocínio análogo ao que fizemos acima, é imediato ver que $f \in C^k$.

Logo f é difeomorfismo local (entre curvas) de mesma classe que φ e como é homeomorfismo, é difeomorfismo.

Para encerrar a prova, mostremos que no caso em que φ é injetiva, então γ é difeomorfa à reta. Por difeomorfismo, no contexto desta proposição, entendemos imersão injetiva (não necessariamente um mergulho). Mas se φ é injetiva, $\frac{d\varphi}{dt}(t) \neq 0, \forall t \in I$, pois se existisse t_0 tal que $\frac{d\varphi}{dt}(t_0) = 0$, então $\psi(t) := \varphi(t_0), \forall t \in \mathbb{R}$ seria solução do mesmo problema de Cauchy que φ e aíφ não seria injetiva. Portanto φ é uma imersão injetiva do intervalo maximal (portanto, aberto) I em U. Mas I é difeomorfo à reta, pois é aberto, e compondo φ com tal difeomorfismo, obtemos a imersão desejada.

□

Note que se duas órbitas γ, $\hat{\gamma}$ de um campo $X : U \to \mathbb{R}^m$ se intersectam em um ponto $p \in U$, então elas são a mesma. De fato, supondo que $\varphi(\cdot, x_0) : I_{x_0} \to \gamma$ e $\varphi(\cdot, y_0) : I_{y_0} \to \hat{\gamma}$, são as soluções maximais cujas imagens são respectivamente $\gamma, \hat{\gamma}$ então existem t_0, $\hat{t}_0$ tais que $\varphi(t_0, x_0) = p = \varphi(\hat{t}_0, y_0)$. Sem perda de generalidade, suponha $t_0 \geq \hat{t}_0 \geq 0$. Usando o teorema do fluxo local, definimos $\psi : I_{x_0} - t_0 + \hat{t}_0 \to \hat{\gamma}$ por

$$\psi(t) := \varphi(t + t_0 - \hat{t}_0, x_0);$$

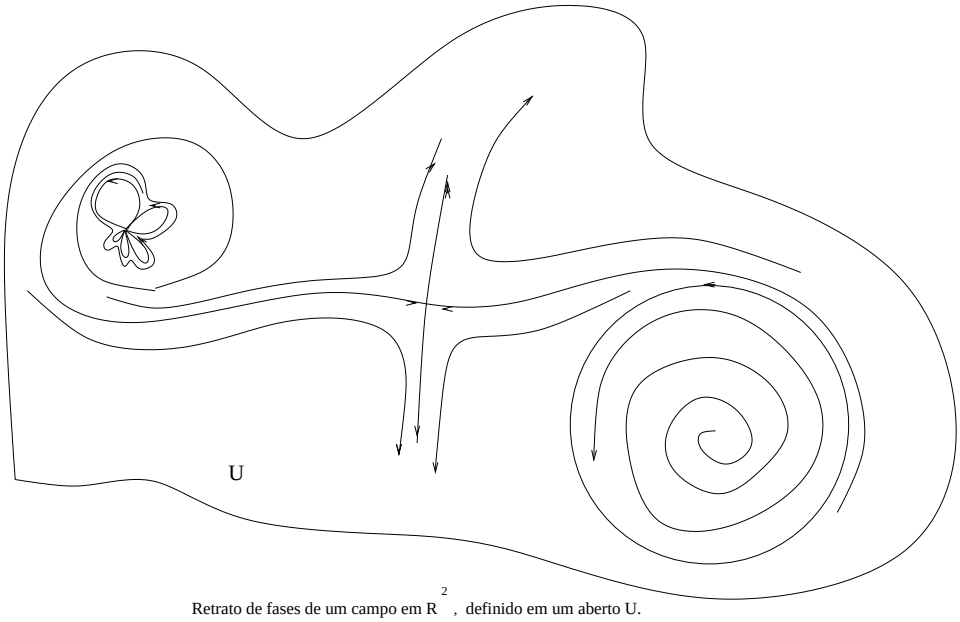
Daí, ψ e $\varphi(\cdot, x_0)$ são soluções maximais com a mesma imagem, γ. Como ψ e $\varphi(\cdot, y_0)$ são ambas soluções maximais do problema de Cauchy

$$\frac{dx}{dt} = X(x); x(\hat{t}_0) = p,$$

concluímos por Picard (e a maximalidade das soluções) que $\psi \equiv \varphi(\cdot, y_0)$ e portanto que $\gamma = \hat{\gamma}$. Como por cada ponto $x \in U$ passa uma órbita de X, concluímos que a coleção das órbitas de X é uma partição de U. Observe ainda que nos casos (ii) e (iii), as órbitas de X possuem uma orientação natural proveniente das soluções do campo: se $\gamma = \varphi(I)$ é uma órbita do campo, para cada $x \in \gamma$, pondo $v(x) := \dot{\varphi}(t)/\|\dot{\varphi}(t)\|$, temos que v é uma escolha contínua de uma base ortonormal em cada espaço (reta) tangente $T_x\gamma$ de γ. Ou seja, v é uma orientação de γ, dita a orientação positiva de γ.

Isto nos enseja a fazer a seguinte

Definição 4.0.10. (Retrato de fase). O conjunto aberto U, munido da decomposição em órbitas de X, chama-se *retrato de fase* de X. As órbitas são orientadas no sentido das curvas integrais do campo X; os pontos singulares são munidos da orientação trivial. A orientação positiva- sentido positivo do percurso- é indicada por meio de setas.



Retrato de fases de um campo em R^2, definido em um aberto U.

4.1 Equivalência e conjugação de campos vetoriais

Definição 4.1.1. (Equivalência de campos). Sejam X_1 e X_2 campos vetoriais de R^n, $X_1 : U_1 \to \mathbb{R}^n$ e $X_2 : U_2 \to \mathbb{R}^n$. Diz-se que X_1 é topologicamente equivalente a X_2 (resp. C^k-equivalente) quando existe um homeomorfismo (resp., um difeomorfismo C^k) $h : U_1 \to U_2$ que leva órbita de X_1 em órbita de X_2 preservando orientação. Mais precisamente, se $p \in U_1$ e $\gamma_1(p)$ é a órbita orientada de X_1 passando por p, então $h(\gamma_1(p))$ é a órbita orientada $\gamma_2(h(p))$ de X_2 passando por $h(p)$.

Definição 4.1.2. (Conjugação de campos). Sejam X e Y campos vetoriais de $\mathbb{R}^n$, $X : U \to \mathbb{R}^n$ e $X : V \to \mathbb{R}^n$, e sejam $\varphi : D \to U$ e $psi : \hat{D} \to V$ os fluxos gerados respectivamente por X e Y. Diz-se que X é topologicamente

conjugado a Y (resp. C^k-conjugado) quando existe um homeomorfismo (resp., um difeomorfismo C^k) $h : U \to V$ tal que $h(\varphi(t,x)) = \psi(t, h(x))$, $\forall (t,x) \in D$. (Em particular, $I^{max}(x) = I^{max}(h(x))$.) h é dita uma conjugação topológica (resp. C^k-conjugação) entre X e Y.

Observação 4.1.3. Suponha que o campo X tenha uma órbita periódica γ, de período τ. Seja $x \in \gamma$. Daí, temos:

$$h(x) = h(\varphi(0,x)) = h(\varphi(\tau,x)) = \psi(\tau, h(x)),$$

portanto, $\hat{\gamma} = h(\gamma)$ é uma órbita periódica de Y com mesmo período τ. O conceito de conjugação de campos é muito forte! Em geral, só são obtidas conjugações locais entre campos por essa razão. Isso explica a existência do conceito (mais fraco) de equivalência de campos.

Observação 4.1.4. Tanto a equivalência de campos como a conjugação são relações de equivalência. Para a equivalência de campos, isso é trivial. Para a conjugação, também não é difícil de provar:

1) Todo campo X é conjugado a si mesmo. Basta tomar $h = Identidade$ (reflexividade da conjugação).

2) Se um campo Xé conjugado a Y via h, então Y é conjugado a X via h^{-1}. De fato, se $h(\varphi(t,x)) = \psi(t, h(x))$, $\forall (t,x) \in D$, então, dado $y \in V$, $y = h(x)$, logo $\varphi(t, h^{-1}(y)) = h^{-1} \circ \psi(t,y)$, portanto h^{-1} é uma conjugação entre Y e X (simetria da conjugação).

3) Se h conjuga os campos X e Y e $\hat{h}$ conjuga os campos Y e Z, então $\tilde{h} = \hat{h} \circ h$ conjuga X e Z. (transitividade da conjugação).

Exemplo 4.1.5. Seja $X : U \to E$ um campo de classe C^1, completo. Daí, fixado $t \in \mathbb{R}$, o difeomorfismo tempo t, dado por $h(\cdot) := \varphi_t(\cdot) = \varphi(t, \cdot)$ conjuga X com ele mesmo. De fato, pelo teorema do Fluxo Local, vale

$$h(\varphi_s(x)) = \varphi_t(\varphi_s(x)) = \varphi_{t+s}(x) =$$

$$\varphi_{s+t}(x) = \varphi_s(\varphi_t(x)) = \varphi_s(h(x)), \forall s \in \mathbb{R}, \forall x \in U.$$

No caso em que X não é completo, tal difeomorfismo tempo t está definido somente em um subconjunto aberto $V_t \subset U$, conjugando então $X|_{V_t}$ e $X|_{\varphi_t(V_t)}$.

Observação 4.1.6. Podemos usar o exemplo anterior como uma forma de estender conjugações em muitos casos. De fato, suponha $X : U \to E$, $Y : V \to E$ dois campos de classe C^1 e suponha que existam $\hat{U} \subset U$ e $\hat{V} \subset V$ ($\hat{U}$, $\hat{V}$ não precisam ser abertos) e um homeomorfismo $\hat{h} : \hat{U} \to \hat{V}$ conjugando $X|_{\hat{U}}$ com $Y|_{\hat{V}}$. Denote por φ o fluxo de X e por ψ o fluxo de Y, e suponha que exista $t\ in \mathbb{R}$ tal que $\varphi(t, \hat{U}) \cap \hat{U} = \emptyset$ e $\psi(t, \hat{V}) \cap \hat{V} = \emptyset$. Chamemos de $\tilde{U} := \hat{U} \cup \varphi(t, \hat{U})$ e $\tilde{V} := \hat{V} \cup \psi(t, \hat{V})$. Daípodemos definir $\tilde{h} : \tilde{U} \to \tilde{V}$ por $\tilde{h}|_{\hat{U}} := \hat{h}$, enquanto que $\tilde{h}|_{\varphi(t,\hat{U})}$ é dado por:

$$\tilde{h}(x) = \psi(t, \hat{h}(\varphi(-t, x))) = \psi_t \circ \hat{h} \circ \varphi_{-t}(x).$$

Deixamos ao leitor a tarefa de verificar que $\tilde{h}$ é uma conjugação entre $X|_{\tilde{U}}$ e $Y|_{\tilde{V}}$.

O próximo lema estabelece uma caracterização útil e comumente verificável de quando um difeomorfismo é ou não uma conjugação entre campos vetoriais.

Lema 4.1.7. *Sejam $U \subset E$, $V \subset \hat{E}$ abertos em espaços de Banach E, $\hat{E}$ e sejam $X : U \to E$ e $Y : V \to \hat{E}$ campos de classe $C^k, k \geq 1$ e $h : U \to V$ um homeomorfismo de classe C^1 (com a inversa h^{-1} não necessariamente diferenciável). Então, h é uma conjugação entre X e Y se e só se*

$$Dh_p \cdot X(p) = Y(h(p)); \forall p \in U.$$

Em particular, se h for um difeomorfismo de classe $C^k, k \geq 1$, então h será uma C^k conjugação se e só se satisfizer a fórmula acima.

Prova: ($\Leftarrow$) Sejam φ, ψ respectivamente os fluxos de X e Y. Dado $p \in U$, defina $\hat{\psi}(t) := h(\varphi_1(t, p))$. Então $\hat{\psi}$ é solução de $\dot{x} = Y(x), x(0) = h(p)$, pois

$$\dot{\hat{\psi}}(t) = Dh(\varphi(t,p)) \cdot \frac{d\varphi(t,p)}{dt} = Dh(\varphi(t,p)) \cdot X(\varphi(t,p)) = Y(h(\varphi(t,p))) = Y(\hat{\psi}(t)).$$

Portanto, $h(\varphi(t,p)) = \psi(t, h(p)) \Rightarrow h$ é conjugação.

($\Rightarrow$) Se h é conjugação, dado $p \in U$, tem-se $h(\varphi(t,p)) = \psi(t, h(p))$, derivando-se em relação a t, obtemos:

$$Dh(\varphi(t,p)) \cdot X(\varphi(t,p)) = Y(\psi(t, h(p)))(\text{ avaliando-se em } |_{t=0}) \Rightarrow$$

$$Dh(p) \cdot X(p) = Y(h(p)).$$

$\square$

Observação 4.1.8. (Pull-back de um campo). Notamos que o lema 4.1.7 acima tem várias aplicações. A mais óbvia, expressa imediatamente em seu enunciado, e que será frequente em nosso texto, é o de estabelecer um critério necessário e suficiente para que um homeomorfismo seja uma conjugação entre dois campos dados. Contudo, há uma outra aplicação, mais sutil. Caso $h : U \to \tilde{U}$ seja um difeomorfismo entre abertos U e $\tilde{U}$ contidos em espaços de Banach E, $\tilde{E}$ respectivamente e tenhamos definido a priori um campo $Y : \tilde{U} \to \tilde{E}$, podemos definir de maneira única um campo $X : U \to E$ que é conjugado a Y via h. Claramente, pelo lema 4.1.7 um tal campo há de satisfazer à fórmula $Dh(p) \cdot X(p) = Y(h(p))$. Uma vez que h seja um difeomorfismo, para cada $p \in U$, $Dh(p)$ é um isomorfismo linear e portanto a fórmula acima determina de maneira única um campo $X : U \to E$. Nesse caso, o campo X é denominado o *pull-back de Y via h*.

4.2 O Teorema do Fluxo Tubular

Usaremos o lema 4.1.7 para classificar qualquer campo $C^k, k \geq 1$ na vizinhança de um ponto regular p. Veremos que um tal campo é localmente (em uma vizinhança de p) C^k conjugado a um campo constante (Teorema do Fluxo Tubular). Para provar tal resultado central, necessitamos primeiramente de uma definição:

Definição 4.2.1. (Secção transversal a um campo). Sejam $X : U \to \mathbb{R}^n$ um campo de classe $C^k, k \geq 1$, $U \subset \mathbb{R}^n$ um aberto e $A \subset \mathbb{R}^{n-1}$ também aberto. Uma aplicação diferenciável $f : A \to U$ de classe C^k chama-se *secção transversal local* a X quando $\forall a \in A$, $Df(a) \cdot \mathbb{R}^{n-1}$ e $X(f(a))$ geram o espaço $\mathbb{R}^n$.

Seja $\Sigma = f(A)$ munido da topologia induzida por U. Se $f : A \to \Sigma$ for um homeomorfismo, diz-se que Σ é uma *seção transversal* a X.

Exemplo 4.2.2. Seja $X : U \to \mathbb{R}^n$ um campo $C^k, k \geq 1$, $p \in U$ com $X(p) \neq 0$ e $\{v_1, \ldots, v_{n-1}, X(p)\}$ uma base de $\mathbb{R}^n$. Para $\delta > 0$ suficientemente pequeno, devido a continuidade do campo X, a aplicação $f : B(0, \delta) \to U$ dada por $f(x_1, \ldots, x_{n-1}) := p + \sum_{i=1}^{n-1} x_i \cdot v_i$ é uma secção transversal de X.

De fato, para ver isso em detalhes basta notar que a matriz Jacobiana J_f de f é justamente a matriz cujas colunas são $v_1, \ldots, v_{n-1}$ ou, o que dá na mesma coisa, que $v_1, \ldots, v_{n-1}$ formam uma base do espaço vetorial $Df(x) \cdot$

$\mathbb{R}^{n-1}$, qualquer que seja o ponto x fixado em $B(0,\delta)$. Considerando a função $g : B(0,\delta) \to \mathbb{R}$ dada por

$$g(x) := \det\Big(v_1, \dots v_{n-1}, X(f(x))\Big) = \det\Big(J_f(x), X(f(x))\Big),$$

temos que g é contínua, pois é composta de aplicações contínuas. Além disso, temos que $g(0) \neq 0$, o que pela continuidade de g implica que existe $\delta > 0$ tal que $g(x) \neq 0$, $\forall x \in B(0,\delta)$. Mas isso implica que para tal δ, os espaços $Df(x)\cdot\mathbb{R}^{n-1}$ e $X(f(x))$, $x \in B(0,\delta)$ geram o $\mathbb{R}^n$, e logo f é secção transversal (uma vez que é imediato de sua definição que f é um homeomorfismo).

Teorema 4.2.3. *(Teorema do Fluxo Tubular). Seja p um ponto não singular de um campo $X : U \to \mathbb{R}^n$ de classe C^k. Então, existe uma vizinhança V de p em U e um difeomorfismo C^k $F : (-\epsilon,\epsilon) \times B \to V$, onde $\epsilon > 0$ e B é uma bola em $\mathbb{R}^{n-1}$ tal que F é uma C^k-conjugação entre o campo constante $Y : (-\epsilon,\epsilon) \times B \to \mathbb{R}^n$ dado por $Y \equiv (1,0,\dots,0) \in \mathbb{R}^n$ e o campo $X|_V$.*

Prova: Seja $\varphi : D \to U$ o fluxo de X, e considere $f : A \to U$ uma secção transversal local com A um aberto de $\mathbb{R}^{n-1}$ contendo a origem, e $f(0) = p$. Defina $D_A \subset \mathbb{R} \times A \subset \mathbb{R}^n$ como $D_A := \{(t,u); (t,f(u)) \in D\}$. Daí, D_A é aberto de $\mathbb{R}^n$, já que é pré-imagem de um aberto de $\mathbb{R}^n$ por uma função contínua possuidora de domínio também aberto.

Definamos portanto $\tilde{F} : D_A \to U$ por $\tilde{F}(t,u) := \varphi(t,f(u))$. Note que $\tilde{F}$ aplica linhas paralelas (de fato, as curvas integrais do campo Y) em curvas integrais de X.

Vamos mostrar que $\tilde{F}$ é um difeomorfismo local em uma vizinhança de $0 = (0,0) \in \mathbb{R} \times \mathbb{R}^{n-1}$. Pelo teorema da função inversa, é suficiente provar que $D\tilde{F}(0)$ é um isomorfismo. De fato,

$$\partial_t \tilde{F}(0) = \frac{d\varphi(t,f(0))}{dt}|_{t=0} = X(\varphi(0,p)) = X(p),$$

e

$$\partial_u \tilde{F}(0) = (D_u\varphi(0,f(u)))|_{u=0} = D_u f(u)|_{u=0} = Df(0).$$

Portanto, a matriz Jacobiana de $\tilde{F}$ em 0 é

$$\begin{pmatrix} X(p) & Df(0) \end{pmatrix}$$

Como f é secção transversal local, tal matriz é invertível, pois suas colunas geram o $\mathbb{R}^n$. Portanto, $D\tilde{F}(0)$ é um isomorfismo linear, e pelo teorema

da função inversa, existem $\epsilon > 0$ e uma bola B em $\mathbb{R}^{n-1}$ com centro na origem tais que $F := \tilde{F}|_{(-\epsilon,\epsilon)\times B}$ é um difeomorfismo sobre o aberto $V = \tilde{F}((-\epsilon,\epsilon)\times B)$.

Só falta ver que F é conjugação entre Y e $X|_V$. Para tal, usaremos do lema 4.1.7. De fato:

$$DF(t,u)\cdot Y(t,u) = DF(t,u)\cdot(1,0,\ldots,0) = \partial_1 F(t,u) = \frac{d\varphi(t,f(u))}{dt} =$$

$$X(\varphi(t,f(u))) = X(F(t,u)) \underbrace{\Rightarrow}_{\text{lema 4.1.7}} F \text{ é conjugação de } Y \text{ e } X|_V.$$

□

Observação 4.2.4. Observamos que restringindo f a B (se necessário) na última demonstração, temos $F(\{0\}\times B) = f(B) = \Sigma\cap V$ e $F(0,0) = p$.

Observação 4.2.5. Note que é também fácil provar diretamente que a aplicação F definida no teorema do fluxo tubular conjuga (localmente) os campos Y e X. Para tal, chamemos de $\psi(t,x) = x + (t,0,\ldots,0)$ o fluxo de Y. Sejam $x = (t_0,u_0)\in(-\epsilon,\epsilon)\times B$ e $t\in(-\epsilon-t_0,\epsilon-t_0)$. Dizer que Y e $X|_V$ são conjugados significa o mesmo que dizer que para tal par (t,x) vale

$$F(\psi(t,x)) = \varphi(t,F(x)) \Leftrightarrow \varphi(-t,F(\psi(t,x))) = F(x),$$

ou sinteticamente, F conjuga Y e X se e só se

$$\varphi_{-t}\circ F\circ\psi_t(x) = F(x).$$

Portanto, a verificação de que F conjuga Y e X é assaz simples:

$$\varphi_{-t}\circ F\circ\psi_t(x) = \varphi_{-t}\circ F\circ\psi_t(t_0,u_0) = \varphi_{-t}\circ F\circ(t+t_0,u_0) =$$

(por definição de F)

$$\varphi_{-t}\circ\varphi(t+t_0,f(u_0)) = \varphi(t_0,f(u_0)) = F(t_0,u_0) = F(x).$$

Corolário 4.2.6. *Seja Σ uma secção transversal de um campo X e $p\in\Sigma$. Então existem uma vizinhança V de p, $\epsilon = \epsilon(p) > 0$ e uma função $\tau : V\to(-\epsilon,\epsilon)$ tal que $\tau(V\cap\Sigma) = 0$ e*

a) $\forall q\in V$, a curva integral $\varphi(t,q)$ de $X|_V$ é definida e biunívoca em $J_q = (-\epsilon+\tau(q),\epsilon+\tau(q))$;

b) $\eta(q) := \varphi(\tau(q),q)$ é o único ponto onde $\varphi(\cdot,q)|_{J_q}$ intercepta a Σ;

c) $\eta : V\to\Sigma$ é de classe C^k e $D\eta(q) : \mathbb{R}^n\to T_{\eta(q)}\Sigma$ é sobrejetiva $\forall q\in V$. Mais ainda, $D\eta_q\cdot v = 0 \Leftrightarrow v = \alpha X(q), \exists\alpha\in\mathbb{R}$.

Prova:

a) Temos que $F : (-\epsilon, \epsilon) \times B \to V$ dado por $F(t, u) = \varphi(t, f(u))$, é conjugação C^k de Y e $X|_V$. Chamemos de $\varphi_X = \varphi|_V$ e de φ_Y o fluxo do campo Y. Considere então a inversa $F^{-1} : V \to (-\epsilon, \epsilon) \times B$, onde dado $q \in V$ escrevemos $F^{-1}(q) = (-\tau(q), \beta(q))$. Como $-\tau$ é função coordenada de F^{-1}, que é C^k, vem que $\tau \in C^k$. Note que

$$F^{-1}(\varphi_X(t, q)) = \varphi_Y(t, F^{-1}(q)) = \varphi_Y(t, (-\tau(q), \beta(q))) =$$

$$\varphi_Y(t, \varphi_Y(-\tau(q), (0, \beta(q))) = \varphi_Y(t - \tau(q), (0, \beta(q)) \Rightarrow$$
$$\varphi_X(t, q) = F \circ \varphi_Y(t, F^{-1}(q)),$$

cujo domínio é J_q. Esta última igualdade, visto que $t \mapsto \varphi_Y(t, F^{-1}(q))$ é biunívoca e F é difeomorfismo (portanto, biunívoco) implica que $\varphi_X(t, q)$ é biunívoca em J_q.

b) Note que $\eta(q) = \varphi_X(\tau(q), q), q \in V$ está portanto bem definida e é C^k. De $\varphi_X(t, q) = F \circ \varphi_Y(t - \tau(q), (0, \beta(q))), \beta(q) \in B$, obtemos que se $t = \tau(q)$, $\eta(q) = \varphi_X(\tau(q), q) = F \circ \varphi_Y(0, (0, \beta(q))) = F(0, \beta(q)) \in \Sigma$, pelo teorema do fluxo tubular e a observação acima deste corolário. Note que $\eta(q)$ é o único ponto da órbita local de q que intercepta a Σ, pois se $t_1 \in J_q$ é tal que $\varphi_X(t_1, q) \in \Sigma$, teríamos

$$(0, B) \ni F^{-1} \circ \varphi_X(t_1, q) = \varphi_Y(t_1, (-\tau(q), \beta(q))) =$$

$$\varphi_Y(t_1, \varphi_Y(-\tau(q), (0, \beta(q))) = \varphi_Y(t_1 - \tau(q), (0, \beta(q))) = (t_1 - \tau(q), \beta(q)).$$

Portanto,

$$0 = t_1 - \tau(q) \Rightarrow t_1 = \tau(q).$$

c) Temos que

$$\begin{aligned}
\eta(q) = \quad \varphi_X(\tau(q), q) = \quad & F\varphi_Y(\tau(q), F^{-1}(q)) = \\
& F \circ \varphi_Y(\tau(q), (-\tau(q), \beta(q))) = \\
& F \circ \varphi_Y(\tau(q), \varphi_Y(-\tau(q), (0, \beta(q)))) = \\
& F \circ \varphi_Y(\tau(q) - \tau(q), (0, \beta(q))) = \\
& F \circ \varphi_Y(0, (0, \beta(q))) = F(0, \beta(q)) = \\
& F \circ (0, \pi_2) \circ F^{-1},
\end{aligned}$$

onde $\pi_2 : \mathbb{R}^n \to \mathbb{R}^{n-1}$ é a projeção canônica na segunda coordenada. Portanto, $\eta = F \circ (0, \pi_2) \circ F^{-1}$ é submersão de V em Σ.

Do lema 4.1.7, da página 69, vale:

$$DF^{-1}(q)\alpha \cdot X(q) = \alpha \cdot Y(F^{-1}(q)) = \alpha \cdot (1, 0, \ldots, 0)$$

De

$$\eta(q) = F \circ (0, \pi_2) \circ F^{-1},$$

temos

$$D\eta(q) = DF \cdot (0, \pi_2) \cdot DF^{-1}(q)$$

o que implica

$$D\eta(q)\alpha \cdot X(q) = DF \cdot (0, \pi_2) \cdot \alpha(1, 0, \ldots 0) = 0.$$

Como o núcleo de $D\eta(q)$ tem dimensão 1, segue-se que $D\eta(q) \cdot v = 0 \Leftrightarrow v = \alpha \cdot X(q)$.

□

Considere agora uma órbita periódica γ de um campo $X : U \to \mathbb{R}^n$ de classe C^k, $k \geq 1$, e uma secção transversal Σ passando por $p \in \gamma$. Sendo $T > 0$ o período de γ e $\varphi(t, x)$ o fluxo de X, podemos definir uma transformação de primeiro retorno ou *transformação de Poincaré* de uma secção transversal contendo p $\Sigma_0 \subset \Sigma$ em Σ usando o último corolário da seguinte forma:

1) Em primeiro lugar, tome uma vizinhança V de p em que esteja definida a conjugação do teorema do fluxo tubular. Em particular, para todo $x \in V$, está definido $\eta(x)$, onde $\eta : V \to \Sigma$ é a projeção vista no corolário acima. Se necessário, restrinja Σ, de modo a que $\Sigma = \Sigma \cap V$.

2) Pela continuidade do fluxo, como $\varphi(T, p) = p$, existe uma vizinhança (tomêmo-la compacta) $W \subset V$ de p tal que $\varphi(T, W) \subset (V \cap F((-\epsilon/2, \epsilon/2) \times B))$. Ainda pelos teoremas de continuidade, como a trajetória passando por p é periódica, seu intervalo maximal é $\mathbb{R}$, logo podemos supor ademais que o intervalo maximal de toda solução com valores iniciais $(0, x), x \in W$ contém ao menos $[-T - \epsilon, T + \epsilon]$.

3) Defina $\Sigma_0 = \Sigma \cap W$. A transformação de Poincaré $\pi : \Sigma_0 \to \Sigma$ é dada pela expressão:

$$\pi(q) := \eta \circ \varphi(T, q).$$

Corolário 4.2.7. *A transformação $\pi : \Sigma_0 \to \Sigma$ é um difeomorfismo sobre sua imagem.*

Prova: De fato, π é claramente contínua como composta de aplicações contínuas. Além disso, é injetiva, pois se $\pi(q_1) = \pi(q_2)$ então

$$\varphi(T + \tau(\varphi(T, q_1)), q_1) = \varphi(T + \tau(\varphi(T, q_2)), q_2) \underbrace{\Rightarrow}_{\circ\varphi_{-T}}$$

$$\varphi(\tau(\varphi(T, q_2)), q_2) = \varphi(\tau(\varphi(T, q_1), q_1).$$

Note que da nossa definição de W, $|\tau(\varphi(T, q_1))|$ e $|\tau(\varphi(T, q_2))|$ são ambos menores que $\epsilon/2$ e portanto $|\tau(\varphi(T, q_2)) - \tau(\varphi(T, q_1))| < \epsilon$. Assim, temos que

$$\varphi(\tau(\varphi(T, q_2)) - \tau(\varphi(T, q_1), q_2) = \varphi(0, q_1) = q_1 \in \Sigma_0 \Rightarrow$$

$$\tau(\varphi(T, q_2)) - \tau(\varphi(T, q_1) = 0 \text{ e } q_2 = q_1,$$

pois o único tempo t com $|t| < \epsilon$ tal que a trajetória de um ponto $q \in \Sigma$ intercepta a Σ é $t = 0$.

Não há perda alguma em supor que W (respectivamente Σ_0), é uma vizinhança *compacta* de p (no caso de Σ_0, em Σ). A injetividade e a continuidade de π implicam que π é um homeomorfismo sobre sua imagem.

Para ver que π é um difeomorfismo local (portanto global em Σ_0, já que é injetivo), começamos por observar que

$$\varphi_T : \cup_{s\in(-\epsilon,+\epsilon)}\varphi_s(W) \to \cup_{s\in(-\epsilon,+\epsilon)}\varphi_{T+s}(W)$$

é uma conjugação local de $X|_W$ e $X|_{\varphi_T(W)}$. De fato, se $s \in (-\epsilon, \epsilon)$ e $q \in W$, do teorema do fluxo local, vale:

$$\varphi_T \circ \varphi_s(q) = \varphi_s \circ \varphi_T(q).$$

Em particular, do lema 4.1.7 vem que

$$D(\varphi_T)(q) \cdot X(q) = X(\varphi_T(q)).$$

Se além do mais $q \in \Sigma_0$, por ser φ_T um difeomorfismo local, sua derivada $D(\varphi_T)(q)$ é uma aplicação linear invertível, levando a soma direta $X(q) \oplus T_q\Sigma_0$ em uma outra soma direta, que de acima vem a ser:

$$X(\varphi_T(q)) \oplus D(\varphi_T)(q) \cdot T_q\Sigma_0.$$

Pelo corolário 4.2.6, $X(\varphi_T(q))$ gera o núcleo de $D\eta(\varphi_T(q))$, o que implica que $D\eta(\varphi_T(q))|_{D(\varphi_T)(q)\cdot T_q\Sigma_0}$ é injetiva. Mas será também sobrejetiva, pois as

dimensões de domínio e imagem são a mesma. Por conseguinte, aplicando a derivada de $\pi = \eta \circ \varphi_T|_{\Sigma_0}$ ao espaço tangente $T_q\Sigma_0$ vem:

$$D\pi(q) \cdot T_q\Sigma_0 = D\eta(\varphi_T(q)) \cdot D(\varphi_T)(q) \cdot T_q\Sigma_0 = T_{\pi(q)}\Sigma,$$

o que implica que $D\pi(q)$ é invertível e que π é difeomorfismo, pelo teorema da função inversa.

□

Definição 4.2.8. (Integral primeira). Uma função contínua (não necessariamente diferenciável) $f : U \subset \mathbb{R}^n \to \mathbb{R}$ é dita uma *integral primeira* de um campo X de classe C^1 se f é constante ao longo das órbitas de X, e não é constante em nenhum aberto de U.

Temos ainda o seguinte corolário do teorema do fluxo tubular:

Corolário 4.2.9. *Dado um ponto p regular de um campo $X : U \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe C^k, existe uma vizinhança $V \ni p$ onde estão definidas $n-1$ integrais primeiras de classe C^k de $X|_V$. Além do mais, tais integrais primeiras são funcionalmente independentes, o que quer dizer que seus gradientes em cada ponto de V formam um conjunto linearmente independente.*

Prova: De fato, temos que

$$F^{-1}\varphi_X(t,q) = \varphi_Y(t, F^{-1}(q)) = (t, F^{-1}(q)).$$

Denotando F^{-1} por h pelo lema 4.1.7 da página 69, temos que isso implica que

$$Dh(\varphi_X(t,q)) \cdot X(\varphi_X(t,q)) = (1, 0, \ldots, 0),$$

isto é, se $h = (h_1, h_2, \ldots, h_n)$ o acima é o mesmo que

$$\begin{pmatrix} \nabla h_1(\varphi_X(t,q)) \\ \nabla h_2(\varphi_X(t,q)) \\ \vdots \\ \nabla h_n(\varphi_X(t,q)) \end{pmatrix} \cdot X(\varphi_X(t,q)) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Em outras palavras temos que

$$\begin{gathered} < \nabla h_2(\varphi_X(t,q)), X(\varphi_X(t,q)) >= 0 \\ \vdots \\ < \nabla h_n(\varphi_X(t,q)), X(\varphi_X(t,q)) >= 0 \end{gathered}$$

Portanto
$$\frac{d(h_i \circ \varphi_X(t,q))}{dt} = 0, \forall i = 2, \ldots n,$$
o que implica que $h_2, \ldots, h_n$ são integrais primeiras de X, uma vez que, do fato de serem as funções coordenadas de um difeomorfismo, são funcionalmente independentes, não se anulando em nenhum aberto.

□

4.3 Exercícios

1. Sejam Σ_1, Σ_2 hiperfícies transversais a um campo $X : U \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe $C^k, k \geq 1$. Se $p_i = \varphi(t_i) \in \Sigma_i$, com $t_1 < t_2$, mostre que existe uma vizinhança V_i de p_i e uma função $\tilde{\tau} : V_1 \cap \Sigma_1 \to \mathbb{R}$ tal que a aplicação $h : V_1 \cap \Sigma_1 \to V_2 \cap \Sigma_2$ dada por $h(q) = \varphi(\tilde{\tau}(q), q)$ é um difeomorfismo.

2. Seja $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ uma aplicação de classe C^1, isto é, f é uma família contínua (na topologia C^1) de campos também de classe C^1 da forma $f_\lambda(x) = f(x, \lambda)$. Suponha que
$$\dot{x} = f(x, 0)$$
possui uma única solução periódica $p(t)$ não constante. Se w é o período desta solução, suponhamos ainda que as únicas soluções de
$$\dot{y} = D_1 f(p(t), 0) \cdot y; \quad y(0) = y(w)$$
são as funções da forma $a \cdot p'(t)$, com $a \in \mathbb{R}$.

   Prove que existe $\delta > 0$ e uma única função $\tau(\lambda)$ de classe C^1 em $|\lambda| < 1$ tal que $\tau(0) = w$ e
$$\dot{x} = f(x, \lambda)$$
tem uma única solução $p(t, \lambda)$ de classe C^1 periódica de período $\tau(\lambda)$ com $p(t, 0) = p(t)$.

   (Sugestão: Sem perda de generalidade, suponha $p(0) = 0$ e $p'(0) = (1, 0, \ldots, 0)$. Isto facilitará os cálculos de derivadas no final da questão. Seja H o hiperplano normal à curva $p(t)$ em 0. Construa então uma transformação de Poincaré $\pi_\lambda : \Sigma \subset H \to H$, provando ainda que

$\pi(h,\lambda) = \pi_\lambda(h), h \in \Sigma$ é de classe C^1. Para cada λ, qualquer ponto fixo de π_λ corresponde a uma órbita periódica de $\dot{x} = f(x,\lambda)$. Para encontrar tais pontos fixos, únicos para cada λ, aplique o teorema da função implícita a $\pi_\lambda(h) - h$.)

3. Seja $X : U \to \mathbb{R}^n$ um campo de vetores, e $f : U \to \mathbb{R}$ uma integral primeira (não necessariamente diferenciável) de X. Então dado $c \in \mathbb{R}$, $M_c = f^{-1}(c)$ é invariante por X (isto é, toda órbita γ de X que intersecta M_c está contida em M_c). No caso particular em que $f \in C^1$ e $c \in f(U)$ é um valor regular de f, podemos considerar as órbitas de M_c como um subsistema com dimensão $n-1$.

4. Se $X : U \to \mathbb{R}^n$ de classe $C^k, k \geq 1$ possui r integrais primeiras (também de classe C^k) funcionalmente independentes em um ponto $p \in U$, então existe uma vizinhança V de p tal que $X|_V$ é C^k-conjugado a um campo Y da forma

$$Y = (Y_1, Y_2, \ldots, Y_{n-r}, 0, \ldots, 0).$$

5. Se dois campos são topologicamente conjugados e um deles possui uma integral primeira, o mesmo vale para o outro.

6. (Campos Gradientes). Seja $U \subset \mathbb{R}^n$. Dizemos que um campo de classe C^1 $X : U \to \mathbb{R}^n$ é um *campo gradiente* se existe uma função $f : U \to \mathbb{R}$ tal que $X(x) = \nabla f(x), \forall x \in X$, ou seja,

$$df(x) \cdot v =< X(x), v >, \forall x \in U, \forall v \in \mathbb{R}^n.$$

Mostre que se X é campo gradiente (de $f : U \to \mathbb{R}$), a função f é crescente ao longo das trajetórias regulares de X, ou seja, dada uma solução não constante $\varphi : I \to U$ de X, a função $f \circ \varphi : I \to \mathbb{R}$ é crescente. Conclua daíque todo campo gradiente não possui órbitas periódicas.

7. Seja $X : U \to \mathbb{R}^n$ um campo. Dados dois pontos x e $y \in U$ e $\gamma : [a,b] \to \Gamma \subset U$, uma curva de classe C^1 tal que $\gamma(a) = x, \gamma(b) = y$, definimos a integral de X ao longo de γ por:

$$\int_\Gamma X := \int_a^b < X(\gamma(s)), \gamma'(s) > ds.$$

Essa definição se estende de maneira natural para curvas C^1 por partes. Prove que a integral acima não depende da parametrização de Γ escolhida.

8. Mostre que se $X : U \to \mathbb{R}^n$ é campo gradiente, e $f : U \to \mathbb{R}$ uma função C^1 da qual X é gradiente, então para todo caminho $\gamma : [a, b] \to \Gamma$ C^1 por partes unindo x e y, vale:

$$\int_\Gamma X = f(y) - f(x).$$

9. Mostre que vale uma recíproca do item anterior. Seja U um aberto conexo de $\mathbb{R}^n$ e $X : U \to \mathbb{R}^n$ um campo C^1 tal que para algum $x_0 \in U$, dado $x \in U$ o valor de $\int_\Gamma X$ é o mesmo para qualquer curva Γ unindo x_0 a x, então X é um campo gradiente. Podemos supor X apenas C^0?

10. Mostre que um campo de classe C^1 $X : U \to \mathbb{R}^n$ é localmente um campo gradiente (isto é, que dado $x \in U$ existe uma vizinhança $B_x \ni x$ tal que $X|B_x$ é gradiente de alguma função C^2 $f : B_x \to \mathbb{R}$) se e só se, a matriz Jacobiana de X é auto-adjunta.

Capítulo 5

Os conjuntos de $\omega-$limite e $\alpha-$limite

Seja $X : U \subset \mathbb{R}^m \to \mathbb{R}^m$ um campo de classe C^k, $k \geq 1$. Seja $\varphi(t) = \varphi(t, p)$ a curva integral passando pelo ponto p em $t = 0$. Suponha que o intervalo maximal de φ seja a reta. Neste capítulo, examinaremos o conjunto de acumulação da órbita de φ quando $t \to +\infty$ (respectivamente, $t \to -\infty$). Mais precisamente, temos a seguinte definição.

Definição 5.0.1. ($\omega-$limite e $\alpha-$limite). O conjunto de $\omega-$limite de p é definido por:

$$\omega(p) := \{q \in U, \exists(t_n) \text{ com } t_n \to +\infty \text{ quando } n \to +\infty \text{ e } \varphi(t_n) \to q\}.$$

Analogamente, definimos o conjunto de $\alpha-$limite de p por:

$$\alpha(p) := \{q \in U, \exists(t_n) \text{ com } t_n \to -\infty \text{ quando } n \to +\infty \text{ e } \varphi(t_n) \to q\}.$$

Proposição 5.0.2. *Seja $X : U \subset \mathbb{R}^m \to \mathbb{R}^m$ um campo de classe C^k, $k \geq 1$, definido em um aberto $U \subset \mathbb{R}^m$, e seja $\gamma^+(p) := \{\varphi(t, p), t \geq 0\}$. Então $\omega(p)$ é fechado em U e invariante por X. Por invariante por X, queremos dizer que se $q \in \omega(p)$, a órbita de X que passa por q está contida em $\omega(p)$. Se, ademais, $\gamma^+(p)$ está contida em um compacto $K \subset U$, então $\omega(p)$ é compacto não vazio e conexo.*

Prova: Comecemos por mostrar que $\omega(p)$ é fechado em U. Note que é possível que $\omega(p)$ seja vazio, mas nesse caso, o que afirmamos no enunciado já está provado. Assim, suponha que exista $q_n \in \omega(p)$, com $q_n \to q \in U$ quando

$n \to +\infty$. Como cada $q_n \in \omega(p)$, para cada q_n podemos tomar $\varphi(t_n, p)$ tal que

$$d(\varphi(t_n, p), q_n) < 1/n, \text{ com } t_n \to +\infty \text{ quando } n \to +\infty,$$

onde $d(\cdot, \cdot)$ designa a distância em U. Portanto,

$$d(\varphi(t_n, p), q) < \underbrace{d(\varphi(t_n, p), q_n)}_{<1/n} + d(q_n, q) \Rightarrow$$

$$\varphi(t_n, p) \to q \text{ quando } n \to +\infty \Rightarrow q \in \omega(p),$$

donde concluímos que $\omega(p)$ é fechado em U.

Mostremos agora que $\omega(p)$ é invariante por X. Seja $q \in \omega(p)$ e seja $q_0 = \varphi(t_0, q)$. Como $q \in \omega(p)$, existe $\varphi(t_n, p) \to q$ quando $t_n \to +\infty$. Considere a sequência $\varphi(t_n + t_0, p)$. Daí,

$$q_0 = \varphi(t_0, q) = \varphi(t_0, \lim_{n \to +\infty} \varphi(t_n, p)) =$$

(pela continuidade de φ)

$$\lim_{n \to +\infty} \varphi(t_0, \varphi(t_n, p)) = \lim_{n \to +\infty} \varphi(t_n + t_0, p),$$

o que implica que $q_0 \in \omega(p)$, e como t_0 é arbitrário dentro do intervalo maximal da órbita de q, segue-se que esta órbita está contida em $\omega(p)$.

A partir de agora, acrescentemos a hipótese de que $\gamma^+(p)$ esteja contida em um compacto $K \subset U$. Daí, tomando $t_n = n \in \mathbb{N}$, como $(\varphi(n)) \in K$, temos que existe uma subsequência convergente $\varphi(n_r) \to q \in K \subset U$, quando $r \to +\infty$, o que por definição, implica que $q \in \omega(p) \Rightarrow \omega(p) \neq \emptyset$. Como $\omega(p) \subset \overline{\gamma^+(p)} \subset K$, e como já provamos que $\omega(p)$ é fechado em U (e daí, em K), isto implica que $\omega(p)$ é compacto.

Resta apenas ver que $\gamma^+(p)$ é conexo. Veremos que tal é consequência do fato de que (conforme provamos acima) $\omega^+(p)$ é compacto. Para tanto, suponha por absurdo que $\omega(p)$ não seja conexo. Então existe uma cisão por abertos A_1 e A_2 não trivial de $\omega(p)$. Como $\omega(p)$ é compacto, existe um aberto limitado B tal que $\omega(p) \subset B \subset \overline{B} \subset U$. Em particular, também $A_1 \cap B$ e $A_2 \cap B$ formam uma cisão não trivial de $\omega(p)$. Como esta cisão é não trivial, $\exists x_1 \in \omega(p), x_1 \in A_1 \cap B$ e existe $x_2 \in \omega(p), x_2 \in A_2 \cap B$. Daí, podemos compor uma sequência $t_1 < \cdots < t_j < \ldots$ tal que $\varphi(t_{2n}, p) \in A_1 \cap B$, $\varphi(t_{2n+1}, p) \in A_2 \cap B$, $\forall n \in \mathbb{N}$. Como $(A_1 \cap B) \cap (A_2 \cap B) = \emptyset$, pelo teorema

da Alfândega, para cada natural n existe $t_{2n} < \tau_n < t_{2n+1}$ tal que $\varphi(\tau_n, p) \in \partial(A_1 \cap B)$. Como $\partial(A_1 \cap B)$ é compacta, $\exists n_i$ tal que $\varphi(\tau_{n_i}, p) \to q \in \partial(A_1 \cap B)$. Portanto, $q \in \omega(p)$, o que é absurdo, pois $\omega(p) \cap (A_1 \cap B) \cup (A_2 \cap B)$ e logo $\partial(A_1 \cap B) \cap \omega(p) = \emptyset$. Donde concluímos que $\omega(p)$ é conexo.

□

Corolário 5.0.3. *Seja $X : U \to \mathbb{R}^m$ um campo definido em um aberto $U \subset \mathbb{R}^m$. Seja $p \in U \subset \mathbb{R}^m$ tal que $\gamma^+(p)$ esteja contida em um compacto $K \subset U$. Se $q \in \omega(p)$, então a trajetória de q está definida para todo tempo $t \in \mathbb{R}$.*

Prova: Pela última proposição, a órbita de q está contida em $\omega(p)$, o qual é compacto. Dos resultados acerca de intervalos maximais, segue-se que a trajetória $\varphi(t, q)$ está definida para todo tempo $t \in \mathbb{R}$. (Em particular, temos ainda que $\omega(q) \subset \omega(p)$).

□

Para a próxima seção, necessitaremos do seguinte lema de caráter geral acerca de pontos regulares contidos em $\omega(p)$:

Lema 5.0.4. *Seja $X : U \to \mathbb{R}^m$ um campo de classe C^1 definido em um aberto $U \subset \mathbb{R}^m$. Seja $\gamma = \gamma(p)$ uma órbita de X correspondente a uma curva integral $\varphi(\cdot, p)$ e $q \in U$ um ponto regular contido em $\omega(\gamma)$. Dadas secções transversais $\Sigma_0 \subset \Sigma$, com Σ_0 compacta passando por q, então $\Upsilon^+ = \{t > 0; \varphi(t, p) \in \Sigma_0\}$ é um conjunto não vazio e discreto de pontos $\{t_n > 0, n \in \mathbb{N}\}$, e existe $\epsilon > 0$ tal que que $t_n + \epsilon \leq t_{n+1}, \forall n \in \mathbb{N}$.*

Prova: Primeiro mostremos que não apenas Υ^+ é não vazio, como é ilimitado. Lembramos que por Σ_0 passar por q queremos dizer que $q \in int(\Sigma_0)$. Usando-se de $\Sigma \cap V$ a secção transversal a X do enunciado tome $q \in V$ a vizinhança de q dada pelo teorema do Fluxo Tubular (cf. 4.2.3, 71), na qual $X|_V$ é conjugado a um campo constante $Y \equiv (1, 0, \ldots, 0)$, definido em uma vizinhança $(-\epsilon, \epsilon) \times B$ da origem do $\mathbb{R}^m$.

Como $q \in \omega(\gamma)$, existe $\varphi(\tilde{t}_n, p) \to q$, quando $\tilde{t}_n \to +\infty$ com

$$\varphi(\tilde{t}_n, p) \in V, \forall n \in \mathbb{N}.$$

Como $\tilde{t}_n \to +\infty$, passando a uma subsequência se necessário, podemos supor $\tilde{t}_{n+1} - \tilde{t}_n > \epsilon$. Façamos

$$\hat{t}_n := \tilde{t}_n + \tau(\varphi(\tilde{t}_n, p)).$$

Daí, pelo teorema do Fluxo Local temos

$$\varphi(\hat{t}_n, p) = \varphi(\tilde{t}_n + \tau(\varphi(\tilde{t}_n, p)), p) = \varphi(\tau(\varphi(\tilde{t}_n), p), \varphi(\tilde{t}_n, p)),$$

implicando que $\varphi(\hat{t}_n, p) \in \Sigma$, por definição de τ. Da continuidade de τ, obtemos que

$$\lim_{n\to\infty} \varphi(t_n, p) = \lim_{n\to+\infty} \varphi(\tau(\varphi(\tilde{t}_n, p)), \varphi(\tilde{t}_n, p)) = \varphi(\tau(q), q) = \varphi(0, q) = q,$$

como queríamos provar.

Considere agora $\Upsilon^+ := \{t \in \mathbb{R}^+; \varphi(t, p) \in \Sigma_0\}$, que vimos ser não vazio. Este conjunto é discreto. De fato, devido ao teorema do Fluxo Tubular, dado $x \in \Sigma_0$ existem $epsilon_x > 0$ e uma vizinhança $V_x \ni x$ tais que se $s \in \Upsilon^+$, e $\varphi(s, p) \in V_x$, então $\varphi(t, p) \notin \Sigma \cap V_x, \forall t \in (s - \epsilon_x, s + \epsilon_x) \setminus \{s\}$. Realmente, se $\varphi(t, p) = \varphi(t - s, \varphi(s, p))$ pertencesse a Σ_0, então $t - s = \tau(\varphi(s, p)) = 0$, pois $\varphi(s, p) \in \Sigma_0$. Por conseguinte, cada $s \in \Upsilon^+$ é isolado. Como Σ_0 é tomada compacta, podemos tomar uma subcobertura finita $V_{x_1}, \ldots, V_{x_j}$ de Σ_0, e $\epsilon = \min\{\epsilon_{x_1}, \ldots, \epsilon_{x_j}\}$. Daí, dado $s \in \Upsilon^+$, concluímos que $\varphi(t, p) \notin \Sigma \forall t \in (s - \epsilon, s + \epsilon) \setminus \{s\}$, e portanto que $(s - \epsilon, s + \epsilon) \cap \Upsilon^+ = \{s\}$.

Como qualquer conjunto discreto da reta é enumerável, segue-se que podemos escrever

$$\Upsilon^+ = \{t_1, t_2, t_3, \ldots, \text{ com } t_1 < t_1 + \epsilon \leq t_2 < \ldots\},$$

e que qualquer sequência $(s_n) : \mathbb{N} \to \Upsilon$ de termos dois a dois distintos tende a $+\infty$ quando $n \to +\infty$.

□

5.1 O teorema de Poincaré-Bendixson

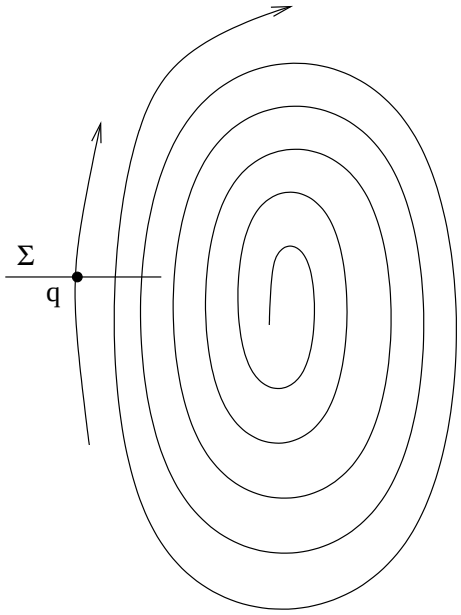
Começamos essa seção lembrando a definição de curva fechada e simples:

Definição 5.1.1. (Curva fechada e simples). Seja $p : [a, b] \to \gamma$ um caminho contínuo. Dizemos que γ é uma *curva fechada e simples* se $p|_{[a,b)}$ é injetivo e $p(a) = p(b)$.

Um resultado importante devido à Jordan (Teorema da Curva de Jordan) afirma que toda curva fechada e simples γ contida em $\mathbb{R}^2$ divide o espaço em duas componentes abertas conexas das quais γ é fronteira comum. Uma

sentença análoga vale para uma curva fechada e simples contida na esfera S^2. Designamos então uma curva fechada e simples nesses espaços como *uma curva de Jordan.*

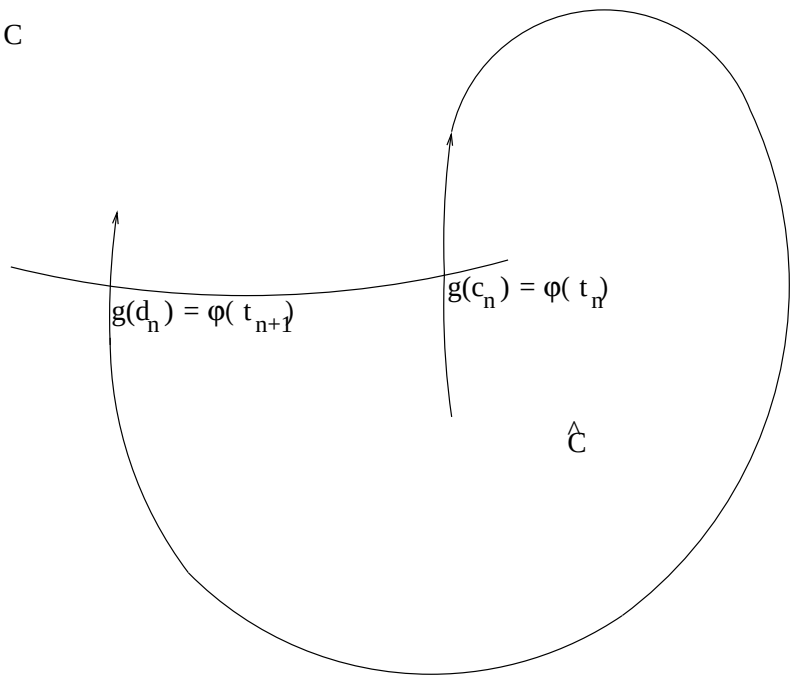
O principal teorema dessa seção é um resultado ímpar de classificação geral de conjuntos de $\omega-$limite. Ele possui a limitação de só ser válido em contextos bidimensionais onde valha o Teorema da curva de Jordan (basicamente, superfícies homeomorfas a S^2 ou a $\mathbb{R}^2$).



Teorema 5.1.2. *(Poincaré-Bendixson). Seja $X : U \to \mathbb{R}^m$ um campo $C^k, k \geq 1$, definido em um aberto $U \subset \mathbb{R}^m$. Suponha que X possui um número finito de singularidades. Se a semi-órbita positiva $\gamma^+(p)$ de um ponto $p \in U$ está contida em um compacto $K \subset U$, então temos as seguintes possibilidades para $\omega(p)$:*

1. *$\omega(p)$ é constituído de uma única órbita singular.*
2. *$\omega(p)$ é constituído de uma única órbita periódica.*
3. *$\omega(p)$ é constituído de um número finito de órbitas singulares, e talvez infinito, de órbitas regulares. Cada órbita regular tem como ω e α-limites uma dessas singularidades.*

Em dimensão dois, toda secção transversal é localmente difeomorfa a um intervalo compacto da reta. Desse modo, podemos muní-la com a ordenação oriunda da própria reta.



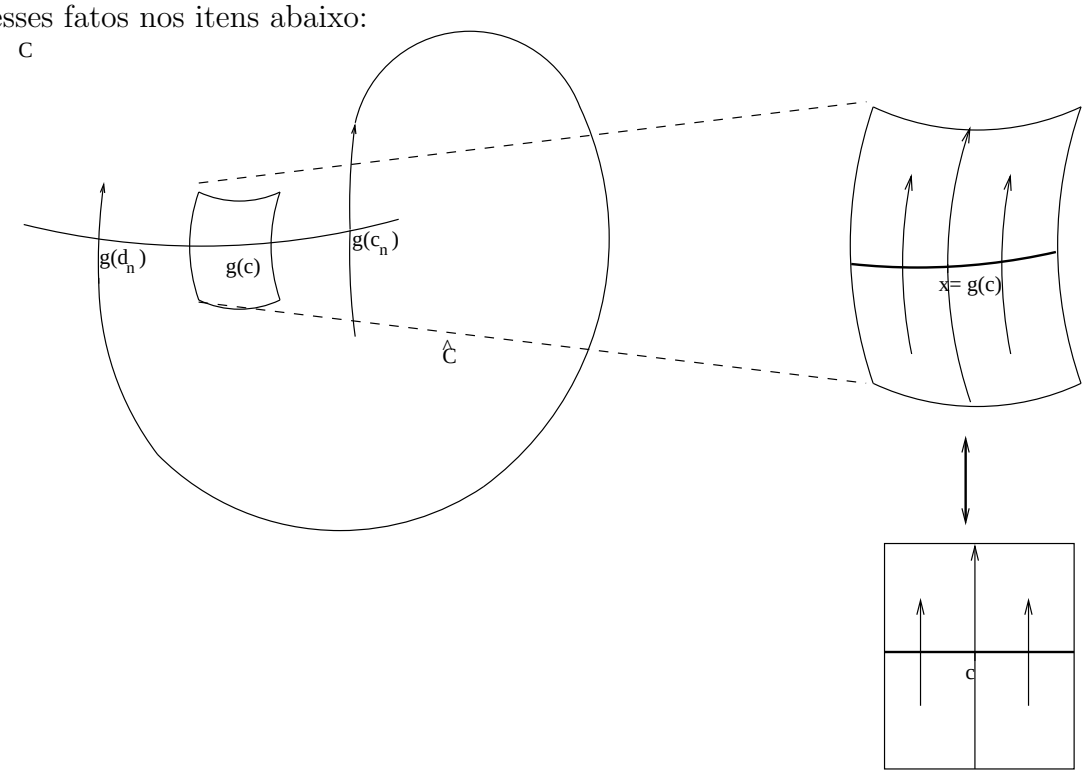
Lema 5.1.3. *Seja $X : U \to \mathbb{R}^2$ um campo de classe C^1 definido em um aberto $U \subset \mathbb{R}^2$. Seja γ uma órbita de X correspondente a uma curva integral $\varphi(\cdot, p)$ cujo domínio contém $\mathbb{R}^+$ e $q \in U$ um ponto regular contido em $\omega(\gamma)$. Então dada uma secção transversal $g : [a, b] \to \Sigma$ passando por q, a órbita de φ intersecta Σ em uma sequência monótona de pontos.*

Prova: Seja $\epsilon_0 > 0$ dado como no teorema do Fluxo Tubular. Conforme vimos no lema anterior, $\gamma \cap \Sigma = \{\varphi(t_n),$ com $t_n \to \infty$ e $t_{n+1} > t_n + \epsilon_0\}$. Se $\varphi(t_2, p) = \varphi(t_1, p)$ a órbita será periódica e nada temos a demonstrar. Desse modo, sem perda, ordenemos Σ de maneira a que $\varphi(t_2) > \varphi(t_1)$. Mostremos que $\varphi(t_{n+1}, p) > \varphi(t_n, p), \forall n \in \mathbb{N}$. Bem, o caso $n = 1$ já vale pela ordenação escolhida. Mostremos que se $\varphi(t_{n+1}, p) > \varphi(t_n, p)$, então $\varphi(t_{n+2}, p) > \varphi(t_{n+1}, p)$. Seja então $\psi : [t_n, t_{n+1} + (t_{n+1} - t_n)] \to \Gamma$ a curva C^1 por partes dada por:

$$\psi(t) = \begin{cases} \varphi(t, p), \text{ se } t \in [t_n, t_{n+1}] \\ g(c_n + (t - t_{n+1}) \cdot \frac{d_n - c_n}{t_{n+1} - t_n}), set \in [t_{n+1}, t_{n+1} + (t_{n+1} - t_n)] \end{cases}$$

onde $g(c_n) = \varphi(t_n)$ e $g(d_n) = \varphi(t_{n+1})$ e $d_n > c_n$. Claramente ψ é a parametrização de uma curva de Jordan C^1 por partes Γ, a qual é fronteira comum de dois abertos conexos disjuntos C e $\hat{C}$, com $C \cup \Gamma \cup \hat{C} = \mathbb{R}^2$.

Dado $x = g(c), c_n < c < d_n$, (trocando os nomes de C e $\hat{C}$ se necessário) começaremos por mostrar que $\varphi(t, x) \in C$ e $\varphi(-t, x) \in \hat{C}$, $\forall t > 0$. Depois, mostraremos que $\varphi(t, g(d_n)) \in C$ e $\varphi(-t, g(c_n)) \in \hat{C}$, $\forall t > 0$. Para completar o quadro, mostraremos finalmente que $g(d) \in C, \forall d > d_n$ e $g(c) \in \hat{C}, \forall C < c_n$ Isso nos permitirá concluir o lema. Para fins didáticos dividimos a prova desses fatos nos itens abaixo:



1. Tomamos uma vizinhança V_x dada pelo teorema do Fluxo Tubular em torno de x, de modo a que $V_x \cap \{g(c_n), g(d_n)\} = \emptyset$. Devido ao difeomorfismo existente com o cilindro $(-\epsilon_x, +\epsilon_x) \times B(0, r)$, segue-se que V_x pode ser particionada em uma união disjunta

$$V_x = A_x \cup \underbrace{(g((c_n, d_n)) \cap V_x)}_{:=S_x} \cup \hat{A}_x,$$

com

$$A_x := \varphi((0, +\epsilon_x) \times S_x), \quad \hat{A}_x := \varphi((-\epsilon_x, 0) \times S_x).$$

Note que $A_x \subset C$ e $\hat{A}_x \subset \hat{C}$. De fato, se existisse ponto de $\hat{C}$ em A_x, pelo teorema da Alfândega, haveria também pontos de Γ em A_x, o que

é absurdo, pelo modo como tomamos a vizinhança V. Por outro lado, se tanto $A_x \subset C$ como $\hat{A}_x \subset \hat{C}$, x não seria um ponto da fronteira comum de C e $\hat{C}$.

2. Também por argumento de conexidade, vê-se que $A := \cup_{x \in g((c_n,d_n))} A_x \subset C$ e $\hat{A} := \cup_{x \in g((c_n,d_n))} \hat{A}_x \subset \hat{C}$. O argumento se dá da seguinte maneira: digamos sem perda que para um certo $x_0 \in g((c_n, d_n))$, $A_{x_0} \subset C$. Agora seja
$$S := \{x \in p((c_n, d_n)); A_x \subset C\}.$$
Ora, S é aberto: se $x \in S$, para todo $y \in V_x$, temos $A_x \cap A_y \subset C$, o que implica (uma vez que ou $A_y \subset C$ ou $A_y \subset \hat{C}$ e $A_x \cap A_y \neq \emptyset$) que $A_y \subset C$. Por outro lado, o complementar S^c de S é expresso por uma sentença análoga a de S:
$$S^c = \{x \in p((c_n, d_n)); A_x \subset \hat{C}\},$$
sendo também aberto em $p((c_n, d_n))$. Desse modo, S é aberto e fechado no conexo $p((c_n, d_n))$ e como $x_0 \in S$, segue-se que $p((c_n, d_n)) = S$

3. Dado $x \in g((c_n, d_n))$, temos portanto que existe $\epsilon_x > 0$ tal que $\varphi(t, x) \in C, \forall 0 < t < \epsilon_x$. Se tivéssemos $\varphi(\mathbb{R}^+, x) \not\subset C$, então pela tricotomia existiria $y = \varphi(\hat{t}, x) \in \Gamma$, sendo que podemos tomar $\hat{t} > 0$ mínimo tal que isso ocorra. Ora, nesse caso, y não poderia pertencer a $g((c_n, d_n))$, porque senão $\varphi(y, -s) = \varphi(\hat{t} - s, x) \in \hat{C}$, para s suficientemente pequeno, o que é absurdo, pelo primeiro item acima. Outra possibilidade é que $\varphi(\hat{t}, x) \in \varphi([t_n, t_{n+1}], p)$. Da minimalidade de $\hat{t}$, concluímos que $\varphi(\hat{t}, x) = \varphi(t_n, p)$. Mas isso implica que $\varphi(t_n - s, p) \in C$ para todo valor de $s > 0$ suficientemente pequeno. Se tal ocorresse, colocando uma caixa $V_{\varphi(t_n,p)}$ dada pelo teorema do Fluxo Tubular com centro em $\varphi(t_n, p)$, concluiríamos pelo mesmo argumento do primeiro item que haveria pontos $y \in g((c_n, d_n))$ e $\epsilon_y > 0$, tal que $\varphi(-s, y) \in C$, $\forall 0 < s < \epsilon_y$, absurdo.

4. Como $\varphi(g(d_n), \mathbb{R}^+)$ é acumulado por pontos contidos em C do tipo $\varphi(g(c), t), t > 0$, concluímos que também $\varphi(g(d_n), \mathbb{R}^+) \subset C$. Analogamente, $\varphi(g(c), \mathbb{R}^-) \subset \hat{C}, \forall c_n \leq c < d_n$. Concluímos então que $C \ni \varphi(t_{n+2}, p) = \varphi(t_{n+2} - t_{n+1}, \varphi(t_{n+1}, p)$ não pertence a $g([c_n, d_n))$. Colocando uma caixa do Teorema do Fluxo Tubular centrada em $g(d_n)$ concluímos ainda que existem $\epsilon_n > 0$ e $\hat{d} > d_n$ tal que $\varphi(g(d), s) \in C$,

$\forall 0 < s < \epsilon_n$, $\forall d_n < d < \hat{d}$. Logo $g(d)$ é acumulado por pontos de C, donde concluímos (já que $g(d) \notin \partial C$ que $g(d) \in C, \forall d_n < d < \hat{d}$. Usando da conexidade da secção transversal, concluímos que $g(d) \in C, \forall d > d_n$. Um raciocínio análogo implica que $g(c) \in \hat{C}, \forall c < c_n$. Logo, só podemos ter $\varphi(t_{n+2}, p) > \varphi(t_{n+1}, p)$, como queríamos provar.

□

Lema 5.1.4. *Seja $X : U \to \mathbb{R}^2$ um campo de classe C^1 definido em um aberto $U \subset \mathbb{R}^2$. Seja γ_p uma órbita de X correspondente a uma curva integral $\varphi(\cdot, p)$, cuja semi-órbita positiva está contida em um compacto de U. Seja ainda $q \in U$ um ponto regular contido em $\omega(\gamma)$ e γ_q a órbita de q. Suponha que exista um ponto regular $\hat{q} \in \omega(q)$. Então γ_q é uma órbita periódica, e $\omega(p) = \gamma_q = \omega(q)$.*

Prova: Seja $g : [a, b] \to \Sigma$ uma secção transversal a X passando por $\hat{q}$. Como $\hat{q} \in \omega(q) \subset \omega(p)$, temos em particular que $x_n = \varphi(t_n, p)$, $0 < t_n < t_{n+1}$ é uma sequência monótona de Σ, onde $\{\varphi(t_n), p\} = \gamma(p) \cap \Sigma$. Portanto, esta sequência possui um único ponto de acumulação, logo $\omega(p) \cap \Sigma = \hat{q} = \gamma_q \cap \Sigma$. Também concluímos daíque γ_q é periódica, já que seu único ponto de intersecção com a secção transversal Σ que passa por $\hat{q} \in \omega(\gamma_q)$ é $\hat{q}$. Para vermos que $\omega(p) = \gamma_q$ usemos da conexidade de $\omega(p)$. Basta que mostremos que γ_q é aberta e fechada em $\omega(p)$ que isto implicará que $\gamma_q = \omega(p)$. Por um lado, é óbvio que γ_q é fechada, pois ela é uma órbita periódica, homeomorfa a S^1 que é compacta. Para vermos que γ_q é aberta em $\omega(p)$, basta vermos que para cada $\tilde{q} \in \gamma_q$ existe um aberto $V_{\tilde{q}} \ni q$ tal que

$$V_{\tilde{q}} \cap \omega(p) = V_{\tilde{q}} \cap \gamma_q.$$

Seja então $\Sigma_{\tilde{q}}$ uma secção transversal passando por $\tilde{q}$ e $V_{\tilde{q}} \ni \tilde{q}$ uma correspondente caixa dada pelo teorema do Fluxo Tubular. Pelo que vimos no parágrafo anterior, $\tilde{q} = \omega(p) \cap \tilde{\Sigma}$, o que também implica que $\tilde{q} = \gamma_q \cap \tilde{\Sigma}$. Ademais, dado algum elemento $\tilde{p} \in \omega(p) \cap V_{\tilde{q}}$, temos pela invariância de $\omega(p)$ que $\eta(\tilde{p}) \in \tilde{\Sigma} \cap \omega(p)$, donde concluímos que $\eta(\tilde{p}) = \tilde{q}$ e portanto que $\tilde{p} \in \gamma_q$.

□

Procedamos agora à prova do Teorema de Poincaré-Bendixson:

Prova: Temos três casos a considerar:

1. Se $\omega(p)$ é constituído apenas de pontos singulares, como ele é conexo (pois está contido no compacto $K \subset U$, vide a caracterização dos conjuntos de $\omega-$limite da seção anterior), e o conjunto das singularidades

de X está sendo suposto (em particular) enumerável, segue-se que $\omega(p)$ reduz-se a um único ponto singular.

2. Se $\omega(p)$ é constituído apenas de pontos regulares, então dado $q \in \omega(p)$, temos pelo lema 5.1.4 que $\omega(q) = \omega(p) = \gamma_q$, e que γ_q é uma órbita periódica.

3. Se $\omega(p)$ possui tanto pontos regulares como singulares, tomemos então $q \in \omega(p)$ um ponto regular. Novamente pelo lema 5.1.4, temos que $\omega(q)$ não pode possuir pontos regulares, pois caso isso ocorresse, $\omega(q) = \omega(p)$ seria uma órbita periódica, o que contradiz estarmos supondo neste item que $\omega(p)$ possui pontos singulares. Portanto, $\omega(q)$ é constituído de pontos singulares. Concluímos aplicando o item 1 acima a $\omega(q)$ que este é constituído de um único ponto singular. O mesmo vale também para o α-limite da órbita de q.

□

Corolário 5.1.5. *Seja U um aberto simplesmente conexo de $\mathbb{R}^2$ e seja $X : U \to \mathbb{R}^2$ um campo de classe C^k, $k \geq 1$, exibindo uma órbita periódica $\gamma \subset U$. Então X possui singularidade p pertencente a componente conexa de U limitada e com γ como fronteira.*

Prova: Chamemos de γ^0 a região homeomorfa a um disco que possui γ como sua fronteira. Suponha, por absurdo, que $X|_{\gamma^0}$ não possua singularidade. Então, em $\overline{\gamma^0}$, escrevendo $X(x) := (X_1(x), X_2(x))$ podemos definir o campo sem singularidades

$$X^{\perp}(x) := (X(x))^{\perp} = (-X_2(x), X_1(x)), \forall x \in \overline{\gamma^0}.$$

Claramente, γ é uma secção transversal para o campo $X^{\perp}$, e podemos supor (trocando $X^{\perp}$ por $-X^{\perp}$, se necessário) que $X^{\perp}$ aponta para dentro de γ. Essa última assertiva significa que para cada $x \in \gamma$, existe $\delta = \delta_x > 0$ tal que $x + s \cdot X^{\perp}(x) \in \gamma^0, \forall 0 < s < \delta_x$. De fato, se supusermos que para algum $x_0 \in \gamma$ existirem sequências $\hat{s}_n \to 0, \tilde{s}_n \to 0$ tais que $x_0 + \hat{s}_n X^{\perp}(x_0) \in \gamma^0$, $x_0 + \tilde{s}_n X^{\perp}(x_0) \in (\gamma^0)^c$, pelo teorema da Alfândega, existirá uma sequência (tomemo-la monótona) $s_n \to 0, s_n > 0$ tal que $x_n = x_0 + s_n X^{\perp}(x_0) \in \partial\gamma_0 = \gamma$. Mas nesse caso, tomando uma parametrização de γ, por exemplo, uma restrição da própria solução φ que corresponde a γ,

escrevendo $\varphi(t_n) = x_n \to \varphi(t_0) = x_0$, obtemos que $t_n \to t_0$ (pois toda parametrização é homeomorfismo)

$$\lim_{n\to\infty} \frac{\varphi(t_n) - \varphi(t_0)}{s_n} = \lim_{n\to\infty} \frac{\varphi(t_n) - \varphi(t_0)}{t_n} \cdot \frac{t_n}{s_n} = X^{\perp}(x_0),$$

o que implica que existe $\lim_{n\to\infty} \frac{t_n}{s_n}$ e que $X^{\perp}(x_0)$ é um múltiplo de $\varphi'(t_0) = X(x_0)$, absurdo. Concluímos portanto que para cada $x \in \gamma$, existe $\delta_x > 0$ tal que $x + sX^{\perp}(x) \notin \gamma$, $\forall 0 < s < \delta_x$. Só falta ver que se para um $x_0 \in \gamma$ vale $x_0 + sX^{\perp}(x_0) \in \gamma^0, \forall 0 < s < \delta_{x_0}$, o mesmo vale para todo $x \in \gamma$. Chamando de γ^1 o outro aberto conexo que γ delimita como fronteira, se supuséssemos que existem $x_0, y_0 \in \gamma$ tais que

$$x_0 + sX^{\perp}(x_0) \in \gamma^0, \forall 0 < s < \delta_{x_0}; \quad y_0 + rX^{\perp}(y_0) \in \gamma^1, \forall 0 < r < \delta_{y_0},$$

então tomando $\delta_0 < \min\{\delta_{x_0}, \delta_{y_0}\}$ e tomando $s_n \to 0$, com $0 < s_n < \delta_0$, concluímos do teorema da Alfândega e da continuidade de $x \mapsto x + s_n X^{\perp}(x)$ que existem sequências x_n e $q_n \to 0$ tais que $x_n + q_n X^{\perp}(x_n) \in \gamma, \forall n$, e pela compacidade de γ podemos supor $x_n \to \hat{x}$, para algum $\hat{x} \in \gamma$. Mas novamente usando da parametrização de γ em torno de $\hat{x}$ dada por φ, vemos que

$$X^{\perp}(\hat{x}) = \lim_{n\to\infty} \frac{x_n - \hat{x}}{q_n} = \lim_{n\to\infty} \frac{\varphi(t_n) - \varphi(\hat{t})}{t_n} \cdot \frac{t_n}{q_n},$$

implicando que existe o limite $\lim_{n\to\infty} \frac{t_n}{q_n}$ e que $X^{\perp}(\hat{x})$ é um múltiplo de $\varphi'(\hat{t}) = X(\hat{x})$, absurdo.

Agora, precisamos ver que definindo $\delta_x := \sup\{t > 0; x + sX^{\perp}(x) \in \gamma^0, \forall 0 < s < t\}$ e $\delta := \sup_{x\in\gamma}\{\delta_x\}$, então $\delta > 0$. Caso contrário, poderíamos tomar (usando da compacidade de γ) alguma sequência $x_n \to x_0 \in \gamma$ com $\delta_{x_n} \to 0$. Note que da definição de δ_x, se o mesmo é finito, então $x + \delta_x X^{\perp}(x) \in \gamma$. Logo, $\gamma \ni x_n + \delta_{x_n} X^{\perp}(x_n) \to x_0$ e o mesmo raciocínio que usamos no parágrafo anterior nos dá que $X^{\perp}(x_0)$ é múltiplo de $X(x_0)$, absurdo. Logo, $\delta > 0$.

Dado $0 < t < \delta$, defina $f_t : \overline{\gamma^0} \to \mathbb{R}^2$ por $f_t(x) := x + tX^{\perp}(x)$. Como $X^{\perp}$ é Lipschitz, vale que $t \cdot X^{\perp}$ é uma contração para t suficientemente pequeno, segue-se do teorema de perturbação da identidade que f_t é um homeomorfismo para todo t pequeno. Em particular, é uma aplicação aberta, que leva fronteira em fronteira, logo, leva $\gamma = \partial\gamma^0$ dentro de γ^0. Tal implica que $f_t(\overline{\gamma^0}) \subset \gamma^0$, para t pequeno. Pelo teorema do ponto fixo de Brower,

concluímos que f_t possui um ponto fixo $p \in \gamma^0$. Logo,

$$p = p + t \cdot X^{\perp}(p) \Rightarrow t \cdot X^{\perp}(p) = 0 \underbrace{\Rightarrow}_{t>0} X^{\perp}(p) = 0,$$

absurdo, pois $X^{\perp}(p) = 0 \Leftrightarrow X(p) = 0$, e supusemos X sem singularidades.

$\square$

5.2 Exercícios

1. Seja $U \subset \mathbb{R}^n$ um aberto e $X : U \to \mathbb{R}^n$ um campo gradiente com uma quantidade (no máximo enumerável de singularidades). Dada uma trajetória $\varphi(\cdot, p)$ cujo domínio contém $\mathbb{R}^+$, mostre que $\omega(p)$ ou é vazio ou é um conjunto unitário.

2. Dois campos completos $X : U \to \mathbb{R}^n$ e $Y : U \to \mathbb{R}^n$ são chamados de *campos comutativos* se seus respectivos fluxos, digamos respectivamente, φ e ψ, comutam. Isto quer dizer que $\varphi_t(\psi_s(x)) = \psi_s(\varphi_t(x)), \forall s, t \in \mathbb{R}, \forall x \in U$. Mostre o Teorema de Elon: Dois campos comutativos $X : S^2 \to TS^2$ e $Y \to TS^2$ possuem uma singularidade em comum.

Capítulo 6

Equações lineares

Seja E um espaço de Banach, e $\mathcal{L}(E)$ o espaço vetorial dos operadores lineares contínuos $A : E \to E$.

Definição 6.0.1. (Equações lineares homogêneas e não homogêneas). Seja $A : I \to \mathcal{L}(E)$ uma aplicação contínua com domínio no intervalo não degenerado $I \subset \mathbb{R}$. A *equação linear homogênea* definida por A é a equação diferencial ordinária:

$$\dot{x} = A(t) \cdot x.$$

Dado um caminho contínuo $b : I \to E$ a *equação linear não homogênea* definida por A e b é a equação diferencial ordinária:

$$\dot{x} = A(t) \cdot x + b(t).$$

6.1 Caracterização das soluções

Proposição 6.1.1. *Seja $\dot{x} = A(t) \cdot x + b(t)$ uma equação linear não homogênea, com $A : I \to \mathcal{L}(E)$ e $b : I \to E$. Então a solução $\varphi(t, t_0, x_0)$ do problema de Cauchy correspondente à equação acima com condição $x(t_0) = x_0$ está definida para todo $t \in I$.*

Prova: Para provarmos o resultado em questão, basta mostramos que $\varphi(\cdot, t_0, x_0)$ está definida em qualquer que seja o intervalo compacto $J \subset I$.

Basta observar que $f(t, x) = A(t) \cdot x + b(t)$ é de Lipschitz em relação à segunda coordenada em $J \times E$, sendo $J \subset I$ um intervalo compacto qualquer, fixado.

De fato, como A é contínua, sua norma $\|A(t)\|$ atinge um valor máximo em J compacto, digamos, C.

Desse modo, temos:

$$\|A(t)\cdot x+b(t)-A(t)\cdot y-b(t)\| \leq \|A(t)\cdot(x-y)\| \leq \|A(t)\|\cdot\|x-y\| \leq C\cdot\|x-y\|$$

Pelo corolário 2.1.3 da página 39, segue-se a proposição.

□

Teorema 6.1.2. *Seja $A : I \to \mathcal{L}(E)$ uma aplicação contínua no intervalo I e considere a equação linear homogênea (isto é, $b(t) \equiv 0$) definida por $A(t)$: $\dot{x} = A(t) \cdot x \quad (*)$. Então, o conjunto $\mathcal{S}$ das soluções de (*) é um espaço vetorial e, fixado $t_0 \in I$, a aplicação $\Phi_{t_0} : E \to \mathcal{S}$ dada por $\Phi_{t_0}(x_0) = \varphi(\cdot, t_0, x_0)$ é isomorfismo linear. Em particular, se $E = \mathbb{R}^n$, $\mathcal{S}$ é um espaço vetorial $n-$dimensional.*

Prova: Pela última proposição, sabemos que todas as soluções de (*) têm o intervalo I como domínio. Dessa maneira, faz sentido tentar provar que $\mathcal{S}$ forma um espaço vetorial de funções. Em primeiro lugar, é imediato verificar que $\varphi \equiv 0$ é solução de (*). Sejam portanto $\varphi_1 : I \to E$ e $\varphi_2 : I \to E$ elementos de $\mathcal{S}$ e $c \in \mathbb{R}$. Daí, temos

$$\frac{d(c \cdot \varphi_1 + \varphi_2)}{dt} = c \cdot \frac{d\varphi_1}{dt} + \frac{d\varphi_2}{dt} = c \cdot A(t) \cdot \varphi_1 + A(t) \cdot \varphi_2 =$$

$$A(t) \cdot (c\varphi_1 + \varphi_2) \Rightarrow c \cdot \varphi_1 + \varphi_2 \in \mathcal{S},$$

portanto temos que o conjunto $\mathcal{S}$ das soluções de (*) é um espaço vetorial.

Mostremos que a aplicação Φ_{t_0} definida no enunciado é isomorfismo.

1) Φ_{t_0} é homomorfismo de espaços vetoriais.

De fato, dados x_0 e y_0 em E, e c um escalar temos

$$\Phi_{t_0}(x_0 + cy_0) = \varphi(\cdot, t_0, x_0 + cy_0),$$

isto é, $\Phi_{t_0}(x_0+cy_0)$ é a solução de (*) com valor inicial (em t_0) de x_0+cy_0. Por outro lado, como $\mathcal{S}$ é espaço vetorial, temos que $\Phi_{t_0}(x_0)+c\Phi_{t_0}(y_0) \in \mathcal{S}$. Mas $\Phi_{t_0}(x_0)+c\Phi_{t_0}(y_0)$ tem como valor em t_0 justamente x_0+cy_0, logo pelo teorema de existência e *unicidade* de soluções $\Phi_{t_0}(x_0 + cy_0) = \Phi_{t_0}(x_0) + c\Phi_{t_0}(y_0)$, e portanto Φ_{t_0} é homomorfismo.

2) Sobrejetividade de $\Phi_{t_0} : E \to \mathcal{S}$.

Seja $\tilde{\varphi} \in \mathcal{S}$, pela proposição acima temos que o domínio de $\tilde{\varphi} = I$. Portanto, dado $t_0 \in I$, faz sentido considerar o ponto $\tilde{x}_0 = \tilde{\varphi}(t_0)$. Pela unicidade de soluções, vale que $\Phi_{t_0}(\tilde{x}_0) = \varphi(\cdot, t_0, \tilde{x}_0) = \tilde{\varphi}(\cdot)$. Portanto, Φ_{t_0} é sobre $\mathcal{S}$.

3) Injetividade de Φ_{t_0}.

Sejam $x_0, x_1 \in E$, com $\Phi_{t_0}(x_0) = \Phi_{t_0}(x_1)$. Daí, $\varphi(t, t_0, x_0) = \varphi(t, t_0, x_1), \forall t \in I \Rightarrow x_0 = \varphi(t_0, t_0, x_0) = \varphi(t_0, t_0, x_1) = x_1$, o que prova a injetividade de Φ_{t_0}. $\square$

Observação 6.1.3. Devido ao teorema acima, no caso em que E tem dimensão finita, nada mais natural para encontrar uma base do espaço de soluções da equação $\dot{x} = A(t) \cdot x$ (*) do que escolher uma base β de $\mathbb{R}^n$ e resolver a dita equação n vezes, tomando por valor inicial a cada vez um certo t_0 fixado e um diferente vetor de β. Em forma compacta, isso é o mesmo que resolver a equação linear matricial (**) dada por

$$\dot{X} = A(t) \cdot X, X(t_0) = X_0;$$

onde X_0 é a matriz $n \times n$ cujas colunas são os vetores da base β.

Definição 6.1.4. (Matriz fundamental). Suponha que E seja um espaço de Banach de dimensão finita (ou seja, $E \simeq \mathbb{R}^n$). Uma matriz $\psi(t)$ de ordem $n \times n$ cujas colunas formam uma base do espaço de soluções de (*) chama-se *matriz fundamental* de (*).

Observação 6.1.5. Se $\psi(t)$ é matriz fundamental, então $\psi(t)$ é solução de (**) com $X_0 = \psi(t_0)$, X_0 matriz invertível. Isto porque se aplicamos o isomorfismo inverso $(\Phi_{t_0})^{-1} : \mathcal{S} \to \mathbb{R}^n$ às colunas $\psi_1(\dot{)}, \ldots, \psi_n(\dot{)}$, suas respectivas imagens $\psi_1(t_0), \ldots, \psi_n(t_0)$ constituem uma base de $\mathbb{R}^n$, logo a matriz cujas colunas são esses vetores é invertível. Além disso, também pelo teorema 6.1.2, se $\psi(t)$ é solução de (**) e é invertível para algum $t_1 \in I$, então é invertível para todo $t \in I$. De fato, dada uma coluna $\varphi_j(t)$ de $\psi(t)$, temos $\varphi_j(t) = (\Phi_t)^{-1}(\Phi_{t_1}(\varphi_j(t_1)))$. Como Φ_t e Φ_{t_1} são isomorfismos, levam base em base, logo (fixado t) os vetores $\varphi_j(t)$ formam uma base de $\mathbb{R}^n$, o que implica que $\psi(t)$ é invertível.

A definição de matriz fundamental é caso particular da seguinte:

Definição 6.1.6. (Solução fundamental). Seja E um espaço de Banach de dimensão qualquer (possivelmente infinita). Dada a equação linear homogênea $\dot{x} = A(t) \cdot x \quad (*)$, com $A : I \to \mathcal{L}(E)$ contínua, podemos definir

uma nova equação (**), com espaço de fases em $\mathcal{L}(E)$, dada por:

$$\dot{X} = A(t) \cdot X, X(t_0) = X_0 \in \mathcal{L}(E). \qquad (**)$$

Dizemos que $\psi : I \to \mathcal{L}(E)$ é uma *solução fundamental* (relacionada ao problema (*)) se ela é solução de (**), com $\psi(t_0) = X_0 \in \mathcal{L}(E)$ a ser um isomorfismo linear de E (injetivo e sobrejetivo).

Mesmo em dimensão infinita, as soluções fundamentais são extremamente úteis, pois conhecendo uma solução fundamental, podemos facilmente calcular qualquer solução do problema linear homogêneo original (*), ou mesmo de uma variante não homogênea deste problema. Para podermos demonstrar este fato, precisamos provar a proposição seguinte que é uma versão em dimensão qualquer da observação 6.1.5:

Proposição 6.1.7. *Seja E um espaço de Banach qualquer e $\psi : I \to \mathcal{L}(E)$ uma solução fundamental, isto é, ψ é solução da equação*

$$\dot{X} = A(t) \cdot X; X(t_0) = X_0 \in \mathcal{L}(E),$$

com X_0 invertível. Então, para cada $t_1 \in I$ fixado, a aplicação linear $\psi(t_1) : E \to E$ é invertível.

Prova: Comecemos pela injetividade. Seja $t_1 \in I$ fixado qualquer e sejam $y_1, z_1 \in E$ tais que

$$\psi(t_1) \cdot y_1 = \psi(t_1) \cdot z_1.$$

Defina $\hat{y}(t) := \psi(t) \cdot y_1$ e $\hat{z}(t) := \psi(t) \cdot z_1$. Daí, afirmamos que tanto y como z são soluções do problema de Cauchy

$$\dot{x} = A(t) \cdot x; x(t_1) = \psi(t_1) \cdot y_1 = \psi(t_1) \cdot z_1.$$

De fato,

$$\frac{d\hat{y}(t)}{dt} = \frac{d\psi(t) \cdot y_1}{dt} = \frac{d\psi(t)}{dt} \cdot y_1 = A(t) \cdot \psi(t) \cdot y_1 = A(t) \cdot \hat{y}(t);$$

cálculos análogos aplicam-se a $\hat{z}$. A condição inicial é imediatamente satisfeita.

Então, pela unicidade de soluções, segue-se que $\hat{y}(t) = \psi(t) \cdot y_1 = \hat{z}(t) = \psi(t) \cdot z_1$, $\forall t \in I$. Em particular, para $t = t_0$, temos:

$$\hat{y}(t_0) = X_0 \cdot y_1 = X_0 \cdot z_1 = \hat{z}(t_0) \underbrace{\Rightarrow}_{X_0^{-1}\circ} y_1 = z_1,$$

pois por hipótese X_0 é invertível. Isso implica a injetividade de $\psi(t_1) : E \to E$.

Demonstremos agora a sobrejetividade de $\psi(t_1)$. Tome $y_1 \in E$. Mostremos que existe $x_1 \in E$ tal que

$$\psi(t_1) \cdot x_1 = y_1.$$

Seja $y : I \to E$ a solução do problema de Cauchy

$$\dot{x} = A(t) \cdot x; x(t_1) = y_1.$$

Daí, seja $y_0 = y(t_0)$. Defina $x_1 := X_0^{-1} \cdot y_0$. Afirmamos que $y(t) = \psi(t) \cdot x_1, \forall t \in I$. Realmente, tal é obtido pelo mesmo tipo de conta que fizemos no caso da injetividade. Temos que

$$\frac{d\psi(t) \cdot x_1}{dt} = \frac{d\psi(t)}{dt} \cdot x_1 = A(t) \cdot \psi(t) \cdot x_1,$$

com $\psi(t_0) \cdot x_1 = X_0 \cdot X_0^{-1} \cdot y_0 = y(t_0)$, o que implica nossa afirmação, por unicidade da solução do problema de Cauchy

$$\dot{x} = A(t) \cdot x; x(t_0) = y_0.$$

Mas então $y_1 = y(t_1) = \psi(t_1) \cdot x_1$, como queríamos mostrar.

□

Proposição 6.1.8. *Seja E um espaço de Banach. Sejam $\psi : I \to \mathcal{L}(E)$ e $\psi_1 : I \to \mathcal{L}(E)$ soluções de $\dot{X} = A(t) \cdot X$, sendo ψ fundamental. Então, existe uma única aplicação $C \in \mathcal{L}(E)$ tal que $\psi_1(t) = \psi(t) \cdot C$.*

Prova: Considere

$$\frac{d}{dt}(\psi^{-1}(t) \cdot \psi_1(t)) = \frac{d}{dt}(\psi^{-1}(t)) \cdot \psi_1(t) + \psi^{-1}(t) \cdot \frac{d}{dt}(\psi_1(t)) =$$

$$-\psi^{-1}(t) \cdot \frac{d}{dt}(\psi(t)) \cdot \psi^{-1}(t) \cdot \psi_1(t) + \psi^{-1}(t) \cdot A(t) \cdot \psi_1(t) =$$

$$-\psi^{-1}(t) \cdot A(t) \cdot \psi(t) \cdot \psi^{-1}(t) \cdot \psi_1(t) + \psi^{-1}(t) \cdot A(t) \cdot \psi_1(t) = 0 \Rightarrow$$

$$\psi^{-1}(t) \cdot \psi_1(t) = C.$$

□

Corolário 6.1.9. *C é invertível se, e somente se, $\psi_1(t)$ é fundamental.*

Prova: Imediata da proposição e da definição de solução fundamental. De fato, tomando $t = t_0$, temos que

$$C = \psi^{-1}(t_0) \cdot \psi_1(t_0) \text{ é invertível} \Leftrightarrow$$

$$\psi_1(t_0)\text{é invertível} \underbrace{\Leftrightarrow}_{\text{prop.6.1.7}}$$

$$\psi_1 \text{ é solução fundamental.}$$

□

Proposição 6.1.10. *(Solução da EDO (***) dada por $\dot{x} = A(t) \cdot x + b(t)$, com $x(t_0) = x_0$, conhecendo-se uma solução fundamental ψ). Sejam E um espaço de Banach qualquer, $I \subset \mathbb{R}$ um intervalo não degenerado, $A : I \to \mathcal{L}(E)$ e $b : I \to E$ aplicações contínuas. Se ψ é uma solução fundamental de $\cdot x = A(t) \cdot x$, então a solução φ de (***) é dada por*

$$\varphi(t, t_0, x_0) = \psi(t) \cdot [\psi^{-1}(t_0) \cdot x_0 + \int_{t_0}^{t} \psi^{-1}(s) \cdot b(s) ds]$$

Prova:

A fórmula é obtida com a seguinte substituição: defina a aplicação auxiliar $c : I \to \mathbb{R}^n$ dada por

$$c(t) := \psi^{-1}(t) \cdot \varphi(t) \Rightarrow c(t_0) = \psi^{-1}(t_0) \cdot x_0.$$

Temos portanto que

$$\dot{c}(t) = (\dot{\psi^{-1}}(t)) \cdot \varphi(t) + \psi^{-1}(t) \cdot (A(t) \cdot \varphi(t) + b(t)) =$$

$$-\psi^{-1}(t) \cdot \dot{\psi}(t) \cdot \psi^{-1}(t) \cdot \varphi(t) + \psi^{-1}(t) A(t) \varphi(t) + \psi^{-1}(t) \cdot b(t) =$$

$$-\psi^{-1} \cdot A(t) \cdot \psi(t) \cdot \psi^{-1}(t) \varphi(t) + \psi^{-1}(t) \cdot A(t) \varphi(t) + \psi^{-1}(t) \cdot b(t) = \psi^{-1}(t) \cdot b(t).$$

Logo, $\dot{c} = \psi^{-1}(t) \cdot b(t)$ e pelo teorema fundamental do cálculo vale

$$c(t) = c(t_0) + \int_{t_0}^{t} \psi^{-1}(s) \cdot b(s) ds = \psi^{-1}(t_0) \cdot x_0 + \int_{t_0}^{t} \psi^{-1}(s) \cdot b(s) ds \Rightarrow$$

$$\varphi(t) = \psi(t) \cdot c(t) = \psi(t) \cdot (\psi^{-1}(t_0) \cdot x_0 + \int_{t_0}^{t} \psi^{-1}(s) \cdot b(s) ds).$$

□

Proposição 6.1.11. *(Fórmula de Liouville). Suponha $E \simeq \mathbb{R}^n$ e seja ψ uma matriz solução de $\dot{X} = A(t) \cdot X$. Então vale*

$$\det \psi(t) = \det \psi(t_0) \cdot e^{\int_{t_0}^{t} \textit{traço}(A(s))ds}$$

Prova:

Se ψ não é matriz fundamental, a igualdade acima é trivial ($\det \psi(t) = 0, \forall t$).Assim, suponha ψ fundamental. Note que a igualdade acima (fórmula de Liouville) nos diz que $\det \psi(t)$ é solução de

$$\begin{cases} \dot{y} = \text{traço}(A(t)) \cdot y \\ y(t_0) = det(\psi(t_0)) \end{cases}$$

Portanto, seja $\varphi(t) = \det(\psi(t))$. Daí, denotando a derivada total de det por $\det'$, temos

$$\frac{d\varphi}{dt}(t) = \det'(\psi(t)) \cdot \dot{\psi}(t) = \det'(\psi(t)) \cdot (A(t) \cdot \psi(t)) =$$

$$\det(A(t) \cdot \psi_1(t)\, \psi_2(t)\, \ldots\, \psi_n(t)) + \det(\psi_1(t)\, A(t) \cdot \psi_2(t)\, \ldots\, \psi_n(t)) + \ldots$$

$$\det(\psi_1(t)\, \psi_2(t),\, \ldots\, A(t) \cdot \psi_n(t)) = \sum_{j=1}^{n} \det(\psi_1(t) \ldots \underbrace{A(t) \cdot \psi_j(t)}_{\text{posição } j} \ldots \psi_n(t)) \quad (*)$$

Observamos que nem traço nem determinante dependem da base adotada em $\mathbb{R}^n$. Escrevamos $A(t) \cdot \psi_j(t)$ na base $\psi_1(t) \ldots \psi_n(t)$. Isto é,

$$A(t) \cdot \psi_j(t) = \sum_{i=1}^{n} \alpha_{ij} \psi_i(t).$$

Substituindo em (*), fica:

$$\sum_{j=1}^{n} \det(\psi_1(t) \ldots \underbrace{A(t) \cdot \psi_j(t)}_{\text{posição } j} \ldots \psi_n(t)) =$$

$$\sum_{j=1}^{n} \det(\psi_1(t) \ldots \underbrace{\sum_{i=1}^{n} \alpha_{ij} \psi_i(t)}_{\text{posição} j} \ldots \psi_n(t)) =$$

$$\sum_{j=1}^{n}(\sum_{i=1}^{n}\det(\psi_1(t)\ldots\underbrace{\alpha_{ij}\psi_i(t)}_{\text{posição}j}\ldots\psi_n(t))).$$

Se $i \neq j$, temos que

$$\det(\psi_1(t)\ldots\underbrace{\alpha_{ij}\psi_i(t)}_{\text{posição}j}\ldots\psi_n(t)) = 0,$$

pois o determinante é forma multilinear alternada e portanto

$$\sum_{j=1}^{n}(\sum_{i=1}^{n}\det(\psi_1(t)\ldots\underbrace{\alpha_{ij}\psi_i(t)}_{\text{posição}j}\ldots\psi_n(t))) =$$

$$\sum_{j=1}^{n}\alpha_{jj}\det(\psi_1(t)\ldots\psi_n(t)) = \text{traço}(A(t))\cdot\det\psi(t),$$

como queríamos demonstrar.

□

6.2 Campos lineares a coeficientes constantes

Seja E um espaço de Banach e $A : E \to E$ um operador linear contínuo (E não necessariamente de dimensão finita). Estudaremos nessa seção a solução explícita da equação

$$\begin{cases}\dot{x} = A\cdot x;\\ x(0) = x_0.\end{cases}$$

Começamos observando que como A é contínua, é automaticamente de classe C^∞ e mesmo analítica. Como vimos na primeira seção deste capítulo, a equação acima tem solução única, definida em toda reta. Pela proposição 6.1.10 tal solução é da forma

$$x(t) = \psi(t)\cdot x_0,$$

onde ψ é a (única) solução fundamental (no caso em que E tem dimensão infinita, ψ não é uma matriz, mas um elemento de $\mathcal{L}(E)$), solução do problema em $\mathcal{L}(E)$ dado por $\dot{X} = A\cdot X$, com $\psi(0) = I$. Calcularemos explicitamente

essa matriz fundamental. Comecemos por calcular as derivadas de $x(t)$ em todas as ordens, avaliadas em $t = 0$:

$$x^{(0)} = x \Rightarrow x^{(0)}(0) = x_0;$$

$$x^{(1)} = \dot{x} \Rightarrow x^{(1)}(0) = A \cdot x_0;$$

$$x^{(2)} = \dot{\dot{x}} = (A \dot{\cdot} x) = \underbrace{\dot{A}}_{=0} \cdot x + A \cdot \dot{x} = A \cdot A \cdot x \Rightarrow x^{(2)}(0) = A^2 \cdot x_0;$$

Mostremos por indução sobre k que $x^{(k)}(t) = A^k \cdot x(t)$:

1. Já provamos a validade da fórmula para $k = 0, 1, 2$.

2. Supondo a fórmula válida para k, temos:

$$x^{(k+1)} = \dot{x^{(k)}} = \dot{A^k} \cdot x = \underbrace{\dot{A}^k}_{=0} \cdot x + A^k \cdot \dot{x} = A^k \cdot A \cdot x = A^{k+1} \cdot x.$$

Concluímos, em particular, que $x^{(k)}(0) = A^k \cdot x_0$.

Note que a série $\sum_{k=0}^{\infty} \|t^k A^k\|/k!$ é absolutamente convergente, pois tem seu termo $\|t^k A^k\|/k! \leq |t|^k \|A\|^k/k!$, e converge (uniformemente em intervalos compactos e absolutamente) a série $\sum_{k=0}^{\infty} |t|^k \|A\|^k/k! = e^{|t|\|A\|}$.

Como $\mathcal{L}(E)$ é um espaço de Banach se E o é, temos que a série $\sum_{k=0}^{\infty} t^k A^k/k!$ converge a um elemento $\Psi(t) \in \mathcal{L}(E)$. Mostrando que Ψ é solução fundamental, temos que $\Psi = \psi$. De fato, $\Psi(0) = I$ e dos teoremas de derivação de séries de potências temos que

$$\dot{\Psi}(t) = \sum_{k=1}^{\infty} k \cdot t^{k-1} A^k/k! = \sum_{k=1}^{\infty} A \cdot t^{k-1} A^{k-1}/(k-1)! =$$

$$A \cdot \sum_{k=0}^{\infty} t^k A^k/k! = A \cdot \Psi(t).$$

Portanto, $\Psi = \psi$ é a solução fundamental do problema $\dot{x} = A \cdot x$. Isso enseja a próxima definição:

Definição 6.2.1. (Exponencial de um operador linear) Dado um operador $A \in \mathcal{L}(E)$, a *exponencial* de A é o elemento $e^A \in \mathcal{L}(E)$ definido por:

$$e^A := \sum_{k=0}^{\infty} A^k/k!$$

Proposição 6.2.2. *(Propriedades da exponencial). Dados operadores $A, B \in \mathcal{L}(E)$, temos as seguintes propriedades da exponencial:*

1. *Se A é um operador linear equivalente a B, isto é, se existe um isomorfismo linear $P \in \mathcal{L}(E)$ com $B = P \cdot A \cdot P^{-1}$, então e^A é equivalente a e^B, com $e^B = P \cdot e^A \cdot P^{-1}$.*

2. *$A \cdot B = B \cdot A \Rightarrow B \cdot e^A = e^A \cdot B$; em particular, todo operador A comuta com sua exponencial e^A.*

3. *$e^{t(A+B)} = e^{tA} \cdot e^{tB} \Leftrightarrow A \cdot B = B \cdot A$; em particular, $e^{-A} \cdot e^A = e^{A-A} = e^A \cdot e^{-A} = e^0 = I$, o que implica que e^A é sempre um isomorfismo (sobrejetivo) de E.*

Prova: Daremos apenas um esboço geral da prova da proposição, que é assaz simples. Os detalhes são deixados para o leitor.

1. Se $B = P \cdot A \cdot P^{-1}$, então para qualquer iterado

$$B^m = \underbrace{P \cdot A \cdot P^{-1} P \cdot A \cdot P^{-1} \ldots P \cdot A \cdot P^{-1} P \cdot A \cdot P^{-1}}_{m \text{ vezes}} = P \cdot A^m \cdot P^{-1}.$$

   Similarmente, mostra-se que se f é um polinômio, $f(B) = P \cdot f(A) \cdot P^{-1}$. Por argumento de passagem ao limite, o mesmo vale para séries de potências absolutamente convergentes, como é o caso da exponencial.

2. O argumento é similar ao do ítem anterior, mostrando-se que se B comuta com A, então comuta com as potências e polinômios de A e, passando-se ao limite, com as séries de potências absolutamente convergentes de A.

3. ($\Leftarrow$) A prova é deixada como exercício, seguindo o padrão de demonstrações similares vistas em Análise complexa.

   ($\Rightarrow$) Supondo $e^{t(A+B)} = e^{tA} \cdot e^{tB}$, derivando-se em ambos os membros desta equação em relação a t, obtemos:

$$(A + B) \cdot e^{t(A+B)} = A \cdot e^{tA} \cdot e^{tB} + e^{tA} \cdot B \cdot e^{tB} \Leftrightarrow$$

$$A \cdot e^{tA} \cdot e^{tB} + B \cdot e^{tA} \cdot e^{tB} = A \cdot e^{tA} \cdot e^{tB} + e^{tA} \cdot B \cdot e^{tB} \underbrace{\Leftrightarrow}_{\circ e^{-tB}}$$

$$B \cdot e^{tA} = e^{tA} \cdot B, \forall t \in \mathbb{R}.$$

Derivando-se novamente em relação a t a equação acima obtida, encontramos

$$B \cdot A \cdot e^{tA} = A \cdot e^{tA} \cdot B;$$

avaliando-se a última igualdade em $t = 0$, segue-se que

$$B \cdot A = A \cdot B.$$

□

6.3 C^1-Conjugação de campos lineares a coeficientes constantes

Lema 6.3.1. *Seja E um espaço de Banach e sejam $A \in \mathcal{L}(E)$, $B \in \mathcal{L}(E)$. Dadas as equações $\dot{x} = A \cdot x$ (1) e $\dot{x} = B \cdot x$ (2), então (1) e (2) são linearmente conjugados se e só se $\exists C \in \mathcal{L}(E)$ tal que $C \cdot A = B \cdot C$, com C invertível.*

Prova: ($\Leftarrow$)

$$C \cdot A = B \cdot C \Rightarrow C \cdot tAC^{-1} = tB \Rightarrow C \cdot e^{tA}C^{-1} = e^{tB} \Rightarrow C \cdot e^{tA} = e^{tB} \cdot C.$$

Dado $x_0 \in E$, os fluxos φ_A de (1) e φ_B de (2) tem a forma:

$$\left.\begin{array}{l} \varphi_A(t, x_0) = e^{tA} \cdot x_0 \\ \varphi_B(t, x_0) = e^{tB} \cdot x_0 \end{array}\right\} \Rightarrow C \cdot \varphi_A(t, x_0) = \varphi_B(t, C \cdot x_0).$$

Logo, (1) e (2) são linearmente conjugados (conjugados por isomorfismo linear).

($\Rightarrow$) Temos que

$$C \cdot \varphi_A(t, x_0) = \varphi_B(t, C \cdot x_0) \Leftrightarrow C \cdot e^{tA} \cdot x_0 = e^{tB} \cdot C \cdot x_0 \Leftrightarrow$$

$$C \cdot e^{tA} \cdot C^{-1} = e^{tB} \Leftrightarrow e^{tC \cdot A \cdot C^{-1}} = e^{tB}.$$

Derivando a última expressão em t e avaliando o resultado em $t = 0$, vem:

$$C \cdot A \cdot C^{-1} e^{tC \cdot A \cdot C^{-1}} = B \cdot e^{tB} \quad (|_{t=0}) \Rightarrow C \cdot A \cdot C^{-1} = B.$$

□

Proposição 6.3.2. *Os campos lineares $\dot{x} = A \cdot x$ (1) e $\dot{x} = B \cdot x$ (2) são C^1-conjugados se e só se A e B são similares; portanto, se e só se são linearmente conjugados.*

Prova: ($\Leftarrow$) Imediato do lema 6.3.1

($\Rightarrow$) Inicialmente, suponha que o difeomorfismo h que conjuga (1) e (2) seja tal que $h(0) = 0$. Daí, temos que $h(e^{tA} \cdot x) = e^{tB} \cdot h(x)$ implica, derivando em relação a t:

$$Dh(e^{tA} \cdot x) \cdot A \cdot e^{tA} \cdot x = B \cdot e^{tB} \cdot h(x).$$

Substituindo em $t = 0$, fica:

$$Dh(x) \cdot A \cdot x = B \cdot h(x), \forall x \in E.$$

Dado $\lambda \in \mathbb{R} \setminus \{0\}$, escrevemos:

$$Dh(\lambda x) \cdot A \cdot \lambda x = B \cdot h(\lambda x) \Rightarrow Dh(\lambda x) \cdot A \cdot x = B \cdot h(\lambda x)/\lambda \Rightarrow$$

(fazendo $\lambda \to 0$)

$$Dh(0) \cdot A \cdot x = B \cdot \lim_{\lambda \to 0} \left(\frac{h(\lambda x)}{\lambda} \right) =$$

$$B \cdot \lim_{\lambda \to 0} \frac{h(\lambda x) - h(0)}{\lambda} = B \cdot Dh(0) \cdot x,$$

o que implica que $Dh(0)$ conjuga A e B, e portanto pelo lema 6.3.1, (1) e (2) são linearmente conjugados por $Dh(0)$.

Suponha agora que $h(0) = c \neq 0$. Defina $H(x) := h(x) - c$. Claramente H é homeomorfismo como composição de homeomorfismos. Note que

$$h(e^{tA} \cdot x) = e^{tB} \cdot h(x) \Rightarrow c = h(0) = e^{tB} \cdot h(0) = e^{tB} \cdot c.$$

Dessa forma, temos:

$$H(e^{tA} \cdot x) = h(e^{tA} \cdot x) - c = e^{tB} \cdot h(x) - c =$$

$$e^{tB} \cdot h(x) - e^{tB} \cdot c = e^{tB} \cdot (h(x) - c) = e^{tB} \cdot H(x).$$

Logo $H \in C^1$ conjuga (1) e (2), com $H(0) = 0$, e recaímos no caso já provado. $\square$

6.4 Revisão de álgebra linear

Teorema 6.4.1. *(Teorema da decomposição em autoespaços generalizados). Sejam $A : \mathbb{C}^n \to \mathbb{C}^n$ um operador linear complexo e $Sp(A)$ o conjunto dos autovalores de A. Então existe decomposição $\mathbb{C}^n = \oplus_{\lambda \in Sp(A)} E(\lambda)$ onde:*

- $A \cdot E(\lambda) \subset E(\lambda)$.
- $(A - \lambda I)|_{E(\lambda)}$ *é nilpotente, isto é,* $(A - \lambda I)^k|_{E(\lambda)} \equiv 0$, *para algum* $k \leq dim(E(\lambda))$.

Para a prova desse teorema, precisamos do seguinte lema:

Lema 6.4.2. *Seja E um espaço vetorial,* $\dim(E) = n < +\infty$. *Seja $T : E \to E$ um operador linear. Então, existe uma decomposição em soma direta $E = E_0 \oplus E_1$ tal que*

- $T \cdot E_0 \subset E_0$ *e* $T|_{E_0}$ *é nilpotente, com nulidade menor ou igual à dimensão de* E_0.
- $T \cdot E_1 = E_1$.

Prova: Note que se T fosse tal que $T \cdot E \cap \mathrm{Ker}(T) = \{0\}$, nada mais teríamos a mostrar (bastaria tomar $E_0 = \mathrm{Ker}(T)$ e $E_1 = T \cdot E$). Isso não ocorre em geral. Entretanto, podemos mostrar que ocorre para algum T^m, $1 \leq m \leq n$. De fato, as sequências abaixo se estabilizam (em certo $m \leq n$):

$$\mathrm{Ker}(T) \subset \mathrm{Ker}(T^2) \subset \cdots \subset \mathrm{Ker}(T^n) \subset E,$$

$$E \supset T \cdot E \supset T^2 \cdot E \supset \cdots \supset T^n \cdot E.$$

A estabilização de tais sequências ocorre porque a dimensão de E é finita.

Note que se $\mathrm{Ker}(T^i) = \mathrm{Ker}(T^{i+1})$, então $\mathrm{Ker}(T^{i+2}) = \mathrm{Ker}(T^{i+1})$, pois se

$$v \in \mathrm{Ker}(T^{i+2}) \Rightarrow T^{i+2} \cdot v = 0 \Rightarrow$$

$$\underbrace{T^{i+1}(T \cdot v) = T^i(T \cdot v) = 0}_{T \cdot v \in \mathrm{Ker}(T^{i+1}) = \mathrm{Ker}(T^i)} \Rightarrow$$

$$v \in \mathrm{Ker}(T^{i+1}).$$

Logo, por indução, temos nesse caso $\mathrm{Ker}(T^j) = \mathrm{Ker}(T^i), \forall j \geq i$.

De um modo análogo, se $T^i(E) = T^{i+1}(E)$ então

$$T \cdot T^i(E) = T \cdot T^{i+1}(E) \Rightarrow T^{i+1}(E) = T^{i+2}(E)$$

Logo, $T^i(E) = T^j(E)$, $\forall j \geq i$.

Tal implica que as sequências acima realmente se estabilizam até, no máximo seu n-ésimo termo. Além disso, são estritamente monótonas (respectivamente, crescente e decrescente) até um índice a partir dos quais elas se tornam constante.

Mostremos que esse índice é o mesmo para ambas as sequências. Suponha que a sequência de imagens de E estabiliza para $m \leq n$. Isso implica que

$$T^j \cdot T^m(E) = E_1 := T^m(E), \forall j \geq 0 \Rightarrow T(E_1) = E_1.$$

Daí, pondo $E_0 := \mathrm{Ker}(T^m)$, temos que dado $v \in \mathrm{Ker}(T^{m+1})$, como $T^{m+1} \cdot v = 0$ se por absurdo $v \notin \mathrm{Ker}(T^m)$, então

$$T^m \cdot v \neq 0 \in E_1 \quad \underbrace{\Rightarrow}_{T|_{E_1} \text{ é isomorfismo}} \quad T \cdot (T^m \cdot v) \neq 0$$

(absurdo, pois $v \in \mathrm{Ker}(T^{m+1})$). Observamos ademais que se $\hat{m}$ é o primeiro índice em que a sequência de núcleos se estabiliza, então se supomos $T^{\hat{m}} \cdot E \supset\neq T^{\hat{m}+1} \cdot E = T(T^{\hat{m}} \cdot E)$, segue-se que existe $0 \neq v \in T^{\hat{m}}(E)$ tal que $T \cdot v = 0$. Seja portanto w tal que $T^{\hat{m}} \cdot w = v$. Então $w \in \mathrm{Ker}(T^{\hat{m}+1}) \setminus \mathrm{Ker}(T^{\hat{m}})$, absurdo. Concluímos dos parágrafos acima que $m = \hat{m}$, isto é, ambas as sequências se estabilizam exatamente para um mesmo índice. Até o índice m, as inclusões dos espaços dessas sequências são estritas. Em particular, concluímos que a dimensão de $E_0 = \mathrm{Ker}(T^m)$ é maior ou igual a m, ou por outra, que a nulidade (menor número de iterações que anula um operador nilpotente) de $T|_{E_0}$ é menor ou igual a $\dim(E_0)$.

Pelo teorema do núcleo e da imagem, temos que $\dim(E_0) + \dim(E_1) = n$. Para mostrar que $E = E_0 \oplus E_1$ basta ver então que $E_0 + E_1$ gera o espaço E. De fato, seja $x \in E$. Tomando $T^m(x) \in E_1 = T^m(E) = T^{2m}(E) \Rightarrow \exists y \in E; T^m(x) = T^{2m}(y) \Rightarrow T^m(x - T^m(y)) = 0$. Logo

$$x = \underbrace{(x - T^m(y))}_{\in \mathrm{Ker}(T^m)} + T^m(y),$$

o que implica que $E_0 + E_1$ geram E e dadas as dimensões desses espaços, E_0 e E_1 estão em soma direta.

□

Podemos agora proceder à prova do teorema de decomposição em autoespaços generalizados:

Prova: Seja $\lambda_1 \in Sp(A)$. A existência de um tal λ_1 é devida ao teorema fundamental da álgebra aplicado ao polinômio característico de A dado por $p(\lambda) := \det(A - \lambda \cdot I)$ (os autovalores de A são as raízes desse polinômio). Aplicando o lema a $T := A - \lambda_1 \cdot I$, obtemos que $\mathbb{C}^n$ se escreve como $\mathbb{C}^n = E(\lambda_1) \oplus E_1$, com $T|_{E(\lambda_1)}$ nilpotente e $T|_{E_1}$ isomorfismo. Como sabemos que dado um autovalor (por exemplo, λ_1), existe pelo menos um autovetor v_1 que lhe corresponde, temos que $v_1 \in \mathrm{Ker}(A - \lambda_1 \cdot I) \subset \mathrm{Ker}((A - \lambda_1 \cdot I)^m) = E(\lambda_1)$, o que implica que $E(\lambda_1)$ é não trivial. Como já dissemos, $(A - \lambda_1)|_{E(\lambda_1)}$ é nilpotente, com nulidade $k = m \leq \dim(E(\lambda_1))$. Como $(A - \lambda_1 \cdot I)(E(\lambda_1)) \subset E(\lambda_1)$, vale ainda que

$$A(E(\lambda_1)) = (A - \lambda_1 \cdot I)(E(\lambda_1)) + \lambda_1 \cdot I(E(\lambda_1)) \subset E(\lambda_1).$$

Note ainda que $T(E_1) = E_1$, portanto, $(A - \lambda_1 \cdot I)(E_1) = E_1$, e se tomamos $v \in E_1$, então $A \cdot v - \lambda_1 \cdot v \in E_1 \Rightarrow A \cdot v \in E_1$. Donde obtemos que $A(E_1) \subset E_1$. Observamos ainda que:

- $A|_{E_0} : E_0 \to E_0$ *não contém autovetor de A que não seja do autovalor λ_1*. De fato, se $\lambda \neq \lambda_1$ é um autovalor de A, se por absurdo existisse um autovetor v de λ contido em E_0, obteríamos:

  $$(A - \lambda_1) \cdot v = A \cdot v - \lambda_1 \cdot v = (\lambda - \lambda_1) \cdot v \Rightarrow T^j \cdot v = (\lambda - \lambda_1)^j \cdot v \neq 0, \forall j \in \mathbb{N}$$

  o que é uma contradição com o fato de que $T|_{E_0}$ é nilpotente.

- *Todos os outros possíveis autovetores de A, referentes aos autovalores distintos de λ_1 estão contidos em E_1.* Realmente, suponha por absurdo que existe um autovetor $v \in E$ de um autovalor $\lambda \in \mathbb{C}$ mas $v \notin E_1$ e $v \notin E_0 = E(\lambda_1)$. Então podemos escrever $v = v_0 + v_1$, com $v_0 \in E_0$ e $v_1 \in E_1$ não nulos. Supondo que k_1 seja a nulidade de $(A - \lambda_1)|_{E_0}$, obteríamos:

  $$E_1 \not\ni (\lambda - \lambda_1)^{k_1} \cdot v = (A - \lambda_1)^{k_1} \cdot v = (A - \lambda_1)^{k_1} \cdot v_0 + (A - \lambda_1)^{k_1} \cdot v_1 =$$

  $((A - \lambda_1)|_{E_0}$ é nilpotente)

  $$(A - \lambda_1)^{k_1} \cdot v_1 \in E_1,$$

  absurdo.

Logo, $A|_{E_1} : E_1 \to E_1$, e podemos reaplicar o lema, dessa vez tomando um autovalor λ_2 de $A|_{E_1}$. Aí obtemos $\mathbb{C}^n = E(\lambda_1) \oplus \underbrace{E(\lambda_2) \oplus E_2}_{E_1}$; continuando nesse procedimento até que $E_j = \{0\}$ (e por conseguinte, sejam exauridos todos os autovalores de A, que são em número finito pois o espaço tem dimensão finita) segue-se o teorema.

□

Corolário 6.4.3. *Todo operador linear $A : \mathbb{C}^n \to \mathbb{C}^n$ se escreve como $A = D + N$, com $D \cdot N = N \cdot D$, onde D é um operador diagonalizável e N é nilpotente. Além disso, tal decomposição é única.*

Prova: Definamos o operador linear D em $\mathbb{C}^n$ definindo-o em cada $E(\lambda_i)$ da decomposição em soma direta $\mathbb{C}^n = \oplus_{\lambda \in Sp(A)} E(\lambda) = \oplus_{j=1}^{r} E(\lambda_j)$. De fato, definimos $D|_{E(\lambda_j)} := \lambda_j \cdot I|_{E(\lambda_j)}$, o que implica definirmos $N|_{E(\lambda_j)} := (A - \lambda_j \cdot I)|_{E(\lambda_j)}$.

Seja

$$\beta = \{v_{11}, \ldots, v_{1d_1}, \ldots, v_{r1}, \ldots, v_{rd_r}\},$$

onde

$$d_j = \dim(E(\lambda_j)) \text{ e } \{v_{j1}, \ldots, v_{jd_j}\}$$

constitui uma base de $E(\lambda_j)$. Como os $E(\lambda_j)$ estão em soma direta, temos que β é base de $\mathbb{C}^n$. Daí,

$$\begin{array}{l} D \cdot v_1 = \lambda_1 \cdot v_1, \quad \forall v_1 \in E(\lambda_1) \\ \quad\vdots \\ D \cdot v_r = \lambda_r \cdot v_r, \quad \forall v_r \in E(\lambda_r) \end{array} \Rightarrow D_\beta = \begin{pmatrix} \lambda_1 & 0 & \ldots & & & 0 \\ 0 & \ddots & \ddots & \ldots & & \vdots \\ 0 & \ddots & \lambda_1 & 0 & \ldots & \\ 0 & \ldots & 0 & \lambda_2 & \ddots & \\ 0 & \ldots & & \ddots & \ddots & 0 \\ 0 & \ldots & & & 0 & \lambda_r \end{pmatrix}$$

Portanto, D é diagonalizável. Que N é nilpotente, já mostramos (imediato do teorema de decomposição em autoespaços generalizados). Note que

$$A|_{E(\lambda_j)} = D|_{E(\lambda_j)} + N|_{E(\lambda_j)} \Rightarrow A = D + N.$$

Vejamos que vale $D \cdot N = N \cdot D$. Para tal, basta mostrarmos que, dado $v \in E(\lambda)$, para $E(\lambda)$ qualquer, vale $D \cdot N \cdot v = N \cdot D \cdot v$. E de fato, neste

caso temos:

$$D \cdot N \cdot v = D \cdot N|_{E(\lambda)} \cdot v = D \cdot \underbrace{(A - \lambda I)|_{E(\lambda)} \cdot v}_{\subset E(\lambda)} =$$

$$D|_{E(\lambda)} \cdot (A - \lambda I)|_{E(\lambda)} \cdot v = (\lambda I)|_{E(\lambda)} \cdot (A - \lambda I)|_{E(\lambda)} \cdot v =$$
$$(A - \lambda I)|_{E(\lambda)} \cdot (\lambda I)|_{E(\lambda)} \cdot v = N|_{E(\lambda)} \cdot D|_{E(\lambda)} \cdot v = N \cdot D \cdot v.$$

Só nos resta agora mostrar que a decomposição acima ($A = D + N$, com D diagonalizável, N nilpotente e $D \cdot N = N \cdot D$) é única.

De fato, se $D+N = A = N'+D'$, como sempre, basta que nos restrinjamos a mostrar que $N = N'$ e $D = D'$ se restritos a um $E(\lambda)$ fixado arbitrário.

Restritos a tal $E(\lambda)$, temos:

$$\lambda I + (A - \lambda I) = N' + D' \Rightarrow \lambda I - D' = N' - (A - \lambda I).$$

Note que todos os operadores comutam com A, e, do acima, vemos que comutam entre si. Lembramos que se dois operadores diagonalizáveis comutam, existe uma base de autovetores comum a ambos, isto é, eles são simultaneamente diagonalizáveis (a recíproca também é obviamente válida). Um esboço de prova desse fato é o seguinte: fixado um autoespaço $E(\lambda_j)$ de D, com $v_j \in E(\lambda_j)$, temos:

$$D \cdot D' \cdot v_j = D' \cdot D \cdot v_j = D' \cdot \lambda_j \cdot v_j = \lambda_j \cdot D' \cdot v_j \Rightarrow D' \cdot v_j \in E(\lambda_j),$$

ou seja, os autoespaços de D são invariantes por D' (e vice-versa). Tal também implica (permutando os papéis de D e D') que cada autoespaço de D é soma de autoespaços de D' ou vice-versa (está contido em um autoespaço de D'). Desse modo, é possível decompor o espaço E em uma soma direta de autoespaços de D ou D' com a propriedade de que cada subespaço dessa soma não contém propriamente nenhum outro autoespaço de D ou D'. Em qualquer base obtida reunindo bases dos autoespaços dessa soma direta, ambos os operadores D, D' são diagonais. Em particular, temos que $D - D'$ é diagonalizável, e diagonal naquela base.

Seja k a nulidade de N e k' a nulidade de N'. Considerando que $D - D' = N' - N$, elevando ambos os membros desta equação a $(k + k')$, temos, usando o binômio de Newton (veja que N' comuta com $N|_{E(\lambda)}$) que o segundo membro é zero. Isto implica que $(D - D')^{k+k'} = 0$, o que para um operador diagonalizável implica que $(D - D') = 0$, isto é, $D = D'$, e daí, $N = N'$.

$\square$

Corolário 6.4.4. *(Teorema de Cayley-Hamilton). Existe um polinômio p de grau menor ou igual a n tal que $p(A) = 0 \in \mathcal{L}(\mathbb{C}^n)$.*

Prova: Tome como polinômio $p(x) = (x-\lambda_1)^{k_1} \cdot \cdots \cdot (x-\lambda_r)^{k_r}$. Considere então a matriz $Z = p(A) = (A - \lambda_1 I)^{k_1} \cdot \cdots \cdot (A - \lambda_r I)^{k_r}$ (lembramos que k_j é a nulidade do operador $(A - \lambda_j)|_{E(\lambda_j)}$). Para mostrar que $Z = 0$, basta mostrar que $Z|_{E(\lambda)} = 0$, com $\lambda = \lambda_1, \ldots, \lambda_r$. Seja $v \in E(\lambda)$. Daí, como A comuta consigo mesma e com $\lambda_j I$, temos que

$$Z \cdot v = (A - \lambda_1)^{k_1} \cdot \cdots \cdot (A - \lambda)^{k(\lambda)} \cdot \cdots \cdot (A - \lambda_r)^{k_r} \cdot v =$$

$$(A - \lambda_1)^{k_1} \cdot \cdots \cdot (A - \lambda_r)^{k_r} \cdot (A - \lambda)^{k(\lambda)} \cdot v = 0,$$

pois $(A - \lambda)^{k(\lambda)} \cdot v = 0$, para todo $v \in E(\lambda)$.

□

Lema 6.4.5. *(A ser usado no Teorema da forma de Jordan). Seja E um espaço vetorial, $\dim(E) < +\infty$ e seja $T : E \to E$ um operador linear nilpotente, isto é, existe um primeiro $k \in \mathbb{N}$ tal que $T^k \equiv 0$. Então existe uma base de E formada por um número finito de sequências (também finitas) linearmente independentes $\{v_{1,1}, \ldots, v_{1,j_1}\}, \ldots \{v_{q,1}, \ldots, v_{q,j_q}\}$ tais que $T \cdot v_{s,j_s} = v_{s,j_s-1} \ldots T \cdot v_{s,1} = 0$, com $s = 1 \ldots q$.*

Prova: Vimos do lema anterior que

$$\{0\} = \mathrm{Ker}(T^0) \subset_{\neq} \mathrm{Ker}(T) \subset_{\neq} \cdots \subset_{\neq} \mathrm{Ker}(T^k) = E.$$

Comecemos nosso algoritmo por E^{k-1}, um espaço complementar de $\mathrm{Ker}(T^{k-1})$ dentro de $\mathrm{Ker}(T^k) = E$. Note que

$$T^s(E^{k-1}) \cap \mathrm{Ker}(T^j) = \{0\}, \forall 1 \leq s \leq k-1 \text{ e } \forall 0 \leq j \leq k-s-1.$$

Em particular, $T^i(E^{k-1})$ é imagem isomorfa de E^{k-1}. Fixe $v_{1,k-1} \ldots v_{q',k-1}$ uma base de E^{k-1} e considere seus iterados $T^{k-s}(v_{r,k-1}) := v_{r,s}$, com $k \geq s \geq 1$ e $1 \leq r \leq q'$, o que já nos dá se não todas, algumas das sequências do enunciado. De fato, para ver que os espaços

$$E^{k-1}, T \cdot E^{k-1}, \ldots, T^{k-1} \cdot E^{k-1},$$

estão em soma direta, observamos inicialmente que todo vetor não nulo em E^{k-1} precisa ser iterado exatamente (no mínimo) k vezes por T para ser

levado no zero. Isso implica que cada vetor não nulo de $T \cdot E^{k-1}$ precisa ser iterado $k-1$ vezes por T para ser levado no zero, e assim por diante. Vemos deste raciocínio que os espaços considerados têm intesecção dois a dois igual a $\{0\}$. Para vermos que estão em soma direta (embora esta soma não perfaça necessariamente o espaço E), seja $v_s \neq 0$ pertencente a um dos espaços acima, digamos $v_s \in T^{k-s} \cdot E^{k-1}$. Daí, $T^s \cdot v_s = 0$, e $T^j(v_s) \neq 0, \forall 0 \leq j < s$. Mostremos que v_s não pode ser expresso como combinação linear de vetores nos demais espaços, do tipo:

$$v_s = \sum_{j \neq s} \alpha_j v_j, v_j \in T^{k-j} E^{k-1}, \alpha_j \text{ não todos nulos.}$$

Realmente, se pudesse, teríamos, podemos mostrar que todos os α_j são nulos. Procedamos pelo princípio da Boa Ordenação. Seja $B = \{j > s; \alpha_j \neq 0\}$. Mostremos que B é vazio. De fato, suponha que não. Seja r o máximo de B. Daí,

$$0 = T^{r-1} v_s = \sum_{j \neq s} \alpha_j T^{r-1} v_j = \alpha_r T^{r-1} v_r \Rightarrow \alpha_r = 0; \text{ (absurdo).}$$

Assim, todos os α_j com $j > s$ são nulos. Por outro lado, daí obtemos que

$$0 \neq T^{s-1} \cdot v_s = \sum_{j < s} \alpha_j T^{s-1} v_j = 0,$$

o que implica que v_s não pode ser expresso segundo uma tal combinação de vetores.

Agora, tome $E^{k-2} \supset T(E^{k-1})$ um espaço complementar de $\mathrm{Ker}(T^{k-2})$ dentro de $\mathrm{Ker}(T^{k-1})$. Repetimos o mesmo raciocínio de antes, a E^{k-2}, descartando as sequências de vetores já contidas nas sequências de E^{k-1}. Como o espaço tem dimensão finita, em um número finito de passos o lema está provado.

□

Teorema 6.4.6. *(Forma de Jordan- caso complexo). Seja $A : \mathbb{C}^n \to \mathbb{C}^n$ um operador linear com autovalores complexos distintos $\lambda_1 \ldots \lambda_r, 1 \leq r \leq n$.*

Então, existe uma base β de $\mathbb{C}^n$ na qual o operador é representado pela matriz

$$A_\beta = \begin{pmatrix} \lambda_1 & 0 \ ou \ 1 & 0 & \dots & & 0 & 0 \\ 0 & \ddots & \ddots & \dots & & & \vdots \\ 0 & \ddots & \lambda_1 & 0 & \dots & & \\ 0 & \dots & 0 & \lambda_2 & 0 \ ou \ 1 & & \\ 0 & \dots & & \ddots & \ddots & & 0 \\ 0 & \dots & & & 0 & \lambda_r & 0 \ ou \ 1 \\ 0 & \dots & & & 0 & & \lambda_r \end{pmatrix}$$

Prova: Aplicamos o último lema a $(A - \lambda_k \cdot I)|_{E(\lambda_k)}$. Pelo lema, existe uma base β_k de $E(\lambda_k)$ em que $(A - \lambda_k \cdot I)|_{E(\lambda_k)}$ é representada pela matriz

$$\begin{pmatrix} 0 & 0 \text{ ou } 1 & 0 & \dots 0 \\ 0 & 0 & \ddots & 0 \dots \\ 0 & \dots & 0 & 0 \text{ ou } 1 \\ 0 & \dots & & 0 \end{pmatrix}$$

Note que nessa base, como em qualquer outra base, $(\lambda_k \cdot I)|_{E(\lambda_k)}$ se escreve como:

$$\begin{pmatrix} \lambda_k & 0 \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots 0 & \lambda_k \end{pmatrix}$$

Como $A|_{E(\lambda_k)} = (\lambda_k \cdot I)|_{E(\lambda_k)} + (A - \lambda_k \cdot I)|_{E(\lambda_k)}$ segue-se que $A|_{E(\lambda_k)}$ se escreve na base β_k como:

$$\begin{pmatrix} \lambda_k & 0 \text{ ou } 1 & 0 & \dots 0 \\ 0 & \lambda_k & \ddots & 0 \dots \\ 0 & \dots & \ddots & 0 \text{ ou } 1 \\ 0 & \dots & & \lambda_k \end{pmatrix}$$

Tomando a base ordenada β formada pelos vetores em $\beta_1, \dots, \beta_r$, da invariância de cada $E(\lambda_k)$ obtemos que o operador A na base β se escreve

como:

$$A_\beta = \begin{pmatrix} \lambda_1 & 0 \text{ ou } 1 & 0 & \dots 0 & & & & \dots & 0 \\ 0 & \lambda_1 & \ddots & 0 \dots & & & & & \vdots \\ 0 & \dots & \ddots & 0 \text{ ou } 1 & & & & & \\ 0 & \dots & & \lambda_1 & 0 & & & & \\ 0 & \dots & & & \ddots & 0 & & & 0 \\ \vdots & & & & & \lambda_r & 0 \text{ ou } 1 & 0 & \dots 0 \\ & & & & & 0 & \lambda_r & \ddots & 0 \dots \\ & & & & & 0 & \dots & \ddots & 0 \text{ ou } 1 \\ 0 & \dots & & & & 0 & \dots & & \lambda_r \end{pmatrix}$$

□

Definição 6.4.7. (Complexificado de um operador real). Considere um operador linear $A : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbb{C}^n = \mathbb{R}^n \oplus \mathbb{R}^n = (\mathbb{R}^n)_1 \oplus (\mathbb{R}^n)_2$, onde $(\mathbb{R}^n)_1 := (\mathbb{R}^n, 0)$ e $(\mathbb{R}^n)_2 := (0, \mathbb{R}^n)$. Se $v = (v_1, v_2) \in \mathbb{R}^n \oplus \mathbb{R}^n$, então definimos o *complexificado* $\tilde{A} : \mathbb{C}^n \to \mathbb{C}^n$ o operador estendendo A dado por

$$\tilde{A} \cdot v := (A \cdot v_1, A \cdot v_2) = A \cdot v_1 + iA \cdot v_2.$$

Definição 6.4.8. (A aplicação conjugação $\overline{} : \mathbb{C}^n \to \mathbb{C}^n$). Dado $v \in \mathbb{C}^n = \mathbb{R}^n \oplus \mathbb{R}^n$, $v = (v_1, v_2)$, a aplicação *conjugação* $\overline{} : \mathbb{C}^n \to \mathbb{C}^n$ é o isomorfismo linear dado por

$$\overline{v} = \overline{(v_1, v_2)} := (v_1, -v_2).$$

Proposição 6.4.9. *Seja $A : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear real. Então o complexificado $\tilde{A}$ de A comuta com a aplicação de conjugação, isto é, $\overline{\tilde{A} \cdot v} = \tilde{A} \cdot \overline{v}$, $\forall v \in \mathbb{C}^n$.*

Prova: A prova é direta:

$$\overline{\tilde{A} \cdot v} = (A \cdot v_1, -A \cdot v_2) = (A \cdot v_1, A \cdot -v_2) = \tilde{A} \cdot \overline{v}.$$

□

Teorema 6.4.10. *(Forma de Jordan- caso real). Seja $A : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear com autovalores reais $\lambda_1 \dots \lambda_r$ e autovalores complexos não*

reais $a_1+ib_1, \dots a_s+ib_s$ *. Então, existe uma base* β *de* $\mathbb{R}^n$ *na qual o operador é representado pela matriz em blocos na diagonal*

$$A_\beta = \begin{pmatrix} J_1 & 0 & & \dots & & 0 \\ 0 & \ddots & & & & \vdots \\ & 0 & J_r & & & \\ \vdots & 0 & & \tilde{J}_1 & & 0 \\ & & & & \ddots & \\ & & & & & \tilde{J}_s \end{pmatrix},$$

onde cada $J_k, 1 \leq k \leq r$ *é da forma:*

$$\begin{pmatrix} \lambda_k & 0 \ ou \ 1 & 0 & \dots 0 \\ 0 & \lambda_k & \ddots & 0 \dots \\ 0 & \dots & \ddots & 0 \ ou \ 1 \\ 0 & \dots & & \lambda_k \end{pmatrix},$$

e cada $\tilde{J}_l, 1 \leq l \leq s$ *é da forma:*

$$\begin{pmatrix} a_l & b_l & c_1 & 0 & 0 \dots & 0 \\ -b_l & a_l & 0 & c_1 & \ddots & 0 \\ 0 & & & \ddots & c_d & 0 \\ \vdots & & & \ddots & 0 & c_d \\ 0 & & & & a_l & b_l \\ 0 & \dots & & 0 & -b_l & a_l \end{pmatrix},$$

onde cada $c_e = 1$ *ou* $c_e = 0$*,* $e = 1 \dots d$*.*

Prova: Note que identificamos A com $\tilde{A}|_{(\mathbb{R}^n)_1}$. Dividiremos a prova em vários passos, por razões didáticas:

1. Como $\tilde{A}$ provem de um operador real, se λ é autovalor de $\tilde{A}$, o mesmo vale para $\overline{\lambda}$, e se v é autovetor correspondente a λ, $\overline{v}$ é autovetor associado a $\overline{\lambda}$. De fato, como $\tilde{A}$ é o complexificado de um operador real, a decomposição $\mathbb{C}^n = (\mathbb{R}^n)_1 \oplus (\mathbb{R}^n)_2$ é invariante por $\tilde{A}$, isto é, $\tilde{A}|_{(\mathbb{R}^n)_1} \cdot (\mathbb{R}^n)_1 \subset (\mathbb{R}^n)_1$ e $\tilde{A}|_{(\mathbb{R}^n)_2} \cdot (\mathbb{R}^n)_2 \subset (\mathbb{R}^n)_2$. Daí,

$$\overline{\tilde{A} \cdot v} = \overline{\lambda \cdot v} = \overline{\lambda} \cdot \overline{v} \Rightarrow \tilde{A} \cdot \overline{v} = \overline{\lambda} \cdot \overline{v}.$$

2. Como A é operador real, se λ é um autovalor qualquer de $\tilde{A}$, então $\overline{E(\lambda)} \supset \overline{\tilde{A} \cdot E(\lambda)} = \tilde{A} \cdot \overline{E(\lambda)}$, o que implica que $\overline{E(\lambda)}$ é deixado invariante por $\tilde{A}$. Ademais,

$$(\tilde{A} - \lambda)^{k_j}(E(\lambda)) = 0 \Leftrightarrow \overline{(\tilde{A} - \lambda)^{k_\lambda}(E(\lambda))} = 0 \Leftrightarrow (\tilde{A} - \overline{\lambda})^{k_\lambda}(\overline{E(\lambda)}) = 0.$$

Isso significa que $(\tilde{A} - \lambda)|_{E(\lambda)}$ e $(\tilde{A} - \overline{\lambda})|_{\overline{E(\lambda)}}$ são operadores nilpotentes de mesma nulidade. Daí, $\overline{\lambda}$ é o único autovalor de $\tilde{A}$ em $\overline{E(\lambda)}$. Além do mais, lembramos que $E(\overline{\lambda}) = \mathrm{Ker}((\tilde{A} - \overline{\lambda})^{d_{\overline{\lambda}}} \supset \mathrm{Ker}((\tilde{A} - \overline{\lambda})^{k_\lambda})$, conforme o lema 6.4.2. Em particular, $\overline{E(\lambda)} \subset E(\overline{\lambda})$. Trocando λ com $\overline{\lambda}$, obtemos que $\overline{E(\overline{\lambda})} \subset E(\lambda)$, donde tiramos, já que a conjugação é um isomorfismo (sesquilinear), que $\dim(E(\lambda)) = \dim(E(\overline{\lambda}))$ e que $\overline{E(\lambda)} = E(\overline{\lambda})$.

3. Como já observamos, $\tilde{A}|_{(\mathbb{R}^n)_1}$ é (identificado com) nosso A original. Note que se λ_j é um autovalor real, do item anterior temos $E(\lambda_j) = \overline{E(\lambda_j)} = E(\overline{\lambda}_j)$. Tal implica que tomando $w_1, \ldots, w_{d_j}$ uma base de $E(\lambda_j)$ e a base formada pelos conjugados $\overline{w_1}, \ldots, \overline{w_{d_j}}$ então as partes reais $(w_1 + \overline{w_1})/2, \ldots, (w_{d_j} + \overline{w_{d_j}})/2$ e imaginárias $(w_1 - \overline{w_1})/2i, \ldots, (w_{d_j} - \overline{w_{d_j}})/2i$ pertencem a $E(\lambda_j)$. Ademais, tais vetores (que são reais) geram $E(\lambda_j)$ enquanto espaço complexo, já que por exemplo, geram $w_1, \ldots, w_{d_j}$. Em particular, do conjunto dessas partes reais e imaginárias, podemos extrair uma base de vetores reais de $E(\lambda_j)$. Os vetores desta base são linearmente independentes sobre $\mathbb{C}$, o que quer dizer que são linearmente independentes enquanto vetores reais, sobre $\mathbb{R}$. Isso significa que esses vetores são uma base do espaço real $E(\lambda_j) \cap (\mathbb{R}^n)_1$, já que tal espaço tem como dimensão real máxima igual à dimensão complexa de $E(\lambda_j)$. Pelo teorema da decomposição em autoespaços generalizados, $\tilde{A}|_{E(\lambda_j)} = \lambda_j \cdot I|_{E(\lambda_j)} + (\tilde{A} - \lambda_j \cdot I)|_{E(\lambda_j)}$. Note que tais parcelas deixam invariante $(\mathbb{R}^n)_1$, pois $\lambda_j \in \mathbb{R}$. Como $(\tilde{A} - \lambda_j \cdot I)|_{E(\lambda_j)}$ é nilpotente, e deixa $E(\lambda_j) \cap (\mathbb{R}^n)_1$ invariante, podemos aplicar à mesma o lema 6.4.5, obtendo uma base de vetores (reais) na qual $(\tilde{A} - \lambda_j \cdot I)|_{E(\lambda_j) \cap (\mathbb{R}^n)_1}$ se escreve como

$$\begin{pmatrix} 0 & 0 \text{ ou } 1 & 0 & \ldots 0 \\ 0 & & \ddots & \vdots \\ 0 & \ddots & & 0 \text{ ou } 1 \\ 0 & \ldots & & 0 \end{pmatrix}$$

Definindo $E := \oplus_{j=1}^{r} E(\lambda_j)$, e justapondo as bases de vetores reais encontradas acima para diferentes valores de j, em uma base γ de do espaço $E \cap (\mathbb{R}^n)_1$, seguindo a prova do teorema da forma de Jordan, versão complexa, temos que:

$$(A|_{E\cap(\mathbb{R}^n)_1})_\gamma = \begin{pmatrix} \lambda_1 & 0 \text{ ou } 1 & 0 & \dots 0 & & & & \dots & 0 \\ 0 & \lambda_1 & \ddots & 0 \dots & & & & & \vdots \\ 0 & \dots & \ddots & 0 \text{ ou } 1 & & & & & \\ 0 & \dots & & \lambda_1 & 0 & & & & \\ 0 & \dots & & & \ddots & 0 & & & 0 \\ \vdots & & & & & \lambda_r & 0 \text{ ou } 1 & 0 & \dots 0 \\ & & & & & 0 & \lambda_r & \ddots & 0 \dots \\ & & & & & 0 & \dots & \ddots & 0 \text{ ou } 1 \\ 0 & \dots & & & & 0 & \dots & & \lambda_r \end{pmatrix}$$

4. No caso dos autoespaços generalizados de autovalores complexos com parte imaginária não nula, a situação é uma pouco diversa. Comecemos por fixar um autovalor λ complexo (e com parte imaginária não nula) de $\tilde{A}$. Observamos que nesse caso, $\dim(E(\lambda) \cap (\mathbb{R}_n)_1) = 0$. De fato, nesse caso $\lambda \neq \overline{\lambda}$, e como vimos $\overline{E(\lambda)} = E(\overline{\lambda})$. Logo

$$E(\lambda) \cap E(\overline{\lambda}) = \{0\} \Rightarrow E(\lambda) \cap \overline{E(\lambda)} = \{0\},$$

o que significa que $E(\lambda)$ (assim como $E(\overline{\lambda})$) não possui vetores reais.

5. Por outro lado, o espaço $\hat{E} = E(\lambda) \oplus E(\overline{\lambda})$ possui uma intersecção não trivial com $(\mathbb{R}^n)^1$. De fato, dado um vetor $v = (v_1, v_2) = v_1 + i \cdot v_2 \in E(\lambda)$ sua parte real v_1 pertence a $\hat{E}$, bem como sua parte imaginária v_2:

$$v_1 = (v + \overline{v})/2\,;\ v_2 = (v - \overline{v})/(2 \cdot i),$$

o que em outras palavras quer dizer que $(v_1, 0) \in \hat{E} \cap (\mathbb{R}^n)_1$ e que também $(v_2, 0) \in \hat{E} \cap (\mathbb{R}^n)_1$.

6. Observe que se $w_1, \dots, w_{d_\lambda}$ constituem uma base que deixa $\tilde{A}|_{E(\lambda)}$ na forma de Jordan (complexa), o mesmo pode ser dito de $\overline{w_1}, \dots, \overline{w_{d_\lambda}}$ com respeito a $\tilde{A}_{\overline{E(\lambda)}}$. Dado $v \in \hat{E}$, designemos sua parte real por v' e sua

parte imaginária por v'' que, como vimos acima, pertencem também a $\hat{E} \cap (\mathbb{R}^n)_1$.

Portanto, dada a base η de $\hat{E}$ dada por $w_1, \ldots, w_{d_\lambda}, \overline{w_1}, \ldots, \overline{w_{d_\lambda}}$ os vetores $w'_1, w''_1, \ldots, w'_{d_\lambda}, w''_{d_\lambda}$ constituem uma base γ_λ de $\hat{E}$ como espaço complexo, bem como de $\hat{E} \cap (\mathbb{R}^n)_1$, como espaço sobre $\mathbb{R}$. De fato, para ver isso, basta observar que o conjunto $\{w'_1, w''_1, \ldots, w'_{d_\lambda}, w''_{d_\lambda}\}$ gera a base η acima, e tem a cardinalidade da dimensão (complexa) de $\hat{E}$, logo tais vetores são linearmente independentes (olhando-os como vetores complexos). Ou seja, tais vetores constituem uma base do espaço complexo $\hat{E}$. Mas se são linearmente independentes sobre o corpo dos complexos, (sendo também vetores reais), também o são sobre o corpo dos reais. Como a dimensão real de $\hat{E} \cap \mathbb{R}^n$ é (no máximo) $2 \cdot d_\lambda$, isso implica a afirmação de que $\{w'_1, w''_1, \ldots, w'_{d_\lambda}, w''_{d_\lambda}\}$ são uma base de $\hat{E} \cap (\mathbb{R}^n)_1$, como espaço real.

7. Agora só falta mostrar que na base γ_λ $\tilde{A}|_{\hat{E} \cap (\mathbb{R}^n)_1}$ tem a forma de J_λ do enunciado. Isto é obtido por cálculo direto, pois sabemos qual a representação de $\tilde{A}$ na base η de $\hat{E}$ e como ela se relaciona com a base (como espaço sobre $\mathbb{R}$) γ_λ. Realmente, temos que

$$(\tilde{A}|_{\hat{E}})_\eta = \begin{pmatrix} \lambda & c_1 & 0 & & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ldots & \lambda & & 0 & \\ & & & \overline{\lambda} & c_1 & 0 \\ & & & & \ddots & \ddots \\ 0 & \ldots & & & 0 & \overline{\lambda} \end{pmatrix},$$

onde $c_1, \ldots, c_{d_\lambda - 1}$ são constantes que podem ser igual a zero ou 1. A $j-$ésima coluna $(1 \leq j \leq d_\lambda)$ acima é a representação de $\tilde{A}|_{\hat{E}} \cdot w_j$ na base η. Do mesmo modo, a $(d_\lambda + j)$- ésima coluna $(1 \leq j \leq d_\lambda)$ acima é a representação de $\tilde{A}|_{\hat{E}} \cdot \overline{w_j}$ na base η.

Temos, por exemplo, que:

$$\tilde{A} \cdot w'_1 \frac{\tilde{A} \cdot w_1 + \tilde{A}\overline{w_1}}{2} = \frac{\lambda \cdot w_1 + \overline{\lambda} \cdot \overline{w_1}}{2} =$$

(escrevendo $\lambda = a + bi$)

$$\frac{(a+bi) \cdot (w'_1 + i \cdot w''_1) + (a - bi) \cdot (w'_1 - i \cdot w''_1)}{2} = a \cdot w'_1 - b \cdot w''_1.$$

Similarmente, calculamos que $\tilde{A} \cdot w_1'' = b \cdot w_1' + a \cdot w_1''$. Só com essas contas, já obtivemos que

$$(\tilde{A}|_{\hat{E}\cap(\mathbb{R}^n)_1})_{\gamma_\lambda} = \begin{pmatrix} a & b & ? & \dots \\ -b & a & ? & \\ 0 & 0 & ? & \\ \vdots & \vdots & ? & \end{pmatrix}.$$

Temos, atuando $\tilde{A}$ em w_2' e w_2'':

$$\tilde{A} \cdot w_2' = \frac{\tilde{A} \cdot w_2 + \tilde{A} \cdot \overline{w_2}}{2} = \frac{c_1 \cdot w_1 + \lambda \cdot w_2 + c_1 \cdot \overline{w_1} + \overline{\lambda}\overline{w_2}}{2} =$$

$$c_1 \cdot w_1' + a \cdot w_2' - b \cdot w_2'';$$

$$\tilde{A} \cdot w_2'' = \frac{\tilde{A} \cdot w_2 - \tilde{A} \cdot \overline{w_2}}{2i} = \frac{c_1 \cdot w_1 + \lambda \cdot w_2 - c_1 \cdot \overline{w_1} - \overline{\lambda}\overline{w_2}}{2i} =$$

$$c_1 \cdot w_1'' + b \cdot w_2' + a \cdot w_2''.$$

Tais computações já nos dão a forma:

$$(\tilde{A}|_{\hat{E}\cap(\mathbb{R}^n)_1})_{\gamma_\lambda} = \begin{pmatrix} a & b & c_1 & 0 & ? & \dots \\ -b & a & 0 & c_1 & ? & \\ 0 & 0 & a & b & ? & \\ \vdots & \vdots & -b & a & ? & \\ 0 & 0 & 0 & 0 & ? & \end{pmatrix}.$$

Prosseguindo nessas mesmas contas, obtemos a forma desejada, justapondo as (sub)bases γ e as diversas γ_λ de modo a obter uma base de $(\mathbb{R}^n)_1$.

Observação 6.4.11. Quando tratarmos de operadores reais, designaremos por $E(\lambda)$ o autoespaçøgeneralizado real associado a λ, se λ for real. Caso contrário, abusando um pouco da notação, designaremos por $E(\lambda)$ a soma dos espaços complexos associados a λ e $\overline{\lambda}$, intersectada com $(\mathbb{R}^n)_1 \simeq \mathbb{R}^n$.

□

6.5 Aplicações da Forma de Jordan

Nesta seção veremos que as matrizes na Forma de Jordan possuem uma exponencial bastante simples. Usaremos este fato para duas aplicações bem distintas: classificar topologicamente as equações lineares a coeficientes constantes e estudar as equações lineares de ordem superior a coeficientes constantes.

Comecemos por analisar uma equação do tipo

$$\begin{cases} \dot{x} = A \cdot x, \\ x(0) = x_0 \end{cases} \Rightarrow x(t) = e^{tA} \cdot x_0,$$

onde A é uma matriz $n \times n$ constante, real ou complexa. Vimos que no caso em que A é uma matriz complexa, $A = D + N$, com D diagonalizável e N nilpotente (digamos, com $N^m \equiv 0$), $D \cdot N = N \cdot D$. Tal comutatividade implica que

$$e^{tA} = e^{t(D+N)} = e^{tD} \cdot e^{tN} = e^{tD} \cdot (I + tN + \frac{t^2 N^2}{2} + \cdots + \frac{t^{m-1} N^{m-1}}{(m-1)!}).$$

No caso de A já estar na forma de Jordan (complexa), digamos

$$A = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & \\ 0 & \dots & & J_r \end{pmatrix},$$

onde cada J_q, $q = 1 \dots r$ é um bloco do tipo

$$J_q = \begin{pmatrix} \lambda_q & c_{q,1} & 0 \dots & 0 \\ 0 & \ddots & \ddots & \\ \vdots & & & c_{q,d_q-1} \\ 0 & \dots & 0 & \lambda_q \end{pmatrix},$$

e as constantes $c_{q,1} \dots c_{q,d_q-1}$ assumem certos valores de zero ou 1. Redividimos cada bloco J_q em blocos menores $B_{q,l}, l = 1 \dots \tilde{l} = \tilde{l}(q)$, da forma

$$B_{q,l} = \begin{pmatrix} \lambda_q & 1 & 0 \dots & 0 \\ 0 & \ddots & \ddots & \\ \vdots & & & 1 \\ 0 & \dots & 0 & \lambda_q \end{pmatrix},$$

Temos então que a exponencial de A é:

$$e^{tA} = \begin{pmatrix} e^{tB_{1,1}} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & \\ 0 & \dots & & e^{tB_{r,\tilde{l}(r)}} \end{pmatrix},$$

onde cada $e^{tB_{q,b}}, q = 1 \dots r, b = 1 \dots \hat{b}(q)$ é igual a:

$$\begin{pmatrix} e^{t\lambda_q} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & \\ & & & 0 \\ 0 & \dots & & e^{t\lambda_q} \end{pmatrix} \cdot \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{\tilde{l}(q)-1}}{(\tilde{l}(q)-1)!} \\ 0 & \ddots & t & \frac{t^2}{2} \dots & \\ 0 & \dots & & \ddots & \vdots \\ 0 & \dots & & & t \\ 0 & \dots & & & 1 \end{pmatrix} =$$

$$\begin{pmatrix} e^{t\lambda_q} & te^{t\lambda_q} & \frac{t^2}{2}e^{t\lambda_q} & \dots & \frac{t^{\tilde{l}(q)-1}}{(\tilde{l}(q)-1)!}e^{t\lambda_q} \\ 0 & \ddots & te^{t\lambda_q} & \frac{t^2}{2}e^{t\lambda_q} \dots & \\ 0 & \dots & & \ddots & \vdots \\ 0 & \dots & & & te^{t\lambda_q} \\ 0 & \dots & & & e^{t\lambda_q} \end{pmatrix}$$

O caso da forma de Jordan real é inteiramente análogo. Neste caso, a forma de Jordan de A se escreve como

$$\begin{pmatrix} J_1 & 0 & & \dots & 0 \\ 0 & \ddots & & & \vdots \\ & 0 & J_r & & \\ \vdots & 0 & & \tilde{J}_1 & 0 \\ & & & & \ddots \\ & & & & \tilde{J}_s \end{pmatrix},$$

onde cada $J_k, 1 \leq k \leq r$ é da forma:

$$\begin{pmatrix} \lambda_k & 0 \text{ ou } 1 & 0 & \dots 0 \\ 0 & \lambda_k & \ddots & 0 \dots \\ 0 & \dots & \ddots & 0 \text{ ou } 1 \\ 0 & \dots & & \lambda_k \end{pmatrix},$$

e cada $\tilde{J}_l, 1 \leq \tilde{l} \leq s$ é da forma:

$$\begin{pmatrix} a_l & b_l & c_1 & 0 & 0 \dots & 0 \\ -b_l & a_l & 0 & c_1 & \ddots & 0 \\ 0 & & & \ddots & c_d & 0 \\ 0 & & & \ddots & 0 & c_d \\ 0 & & & & a_l & b_l \\ 0 & & & & -b_l & a_l \end{pmatrix},$$

onde cada $c_e = 1$ ou $c_e = 0$, $e = 1 \dots d$. Os blocos correspondentes a autovalores reais têm a exponencial vista mais acima. Já os blocos do tipo $\tilde{J}_l, 1 \leq l \leq s$ podem ser subdivididos em blocos do tipo

$$\tilde{B}_{l,\tilde{b}} := \begin{pmatrix} a_l & b_l & 1 & 0 & 0 \dots & 0 \\ -b_l & a_l & 0 & 1 & \ddots & 0 \\ 0 & & & \ddots & 1 & 0 \\ 0 & & & \ddots & 0 & 1 \\ 0 & & & & a_l & b_l \\ 0 & & & & -b_l & a_l \end{pmatrix} =$$

$$\underbrace{\begin{pmatrix} a_l & b_l & 0 & 0 & 0 \dots & 0 \\ -b_l & a_l & 0 & 0 & \ddots & 0 \\ 0 & & & \ddots & 0 & 0 \\ 0 & & & \ddots & 0 & 0 \\ 0 & & & & a_l & b_l \\ 0 & & & & -b_l & a_l \end{pmatrix}}_{:=\tilde{D}_{l,\tilde{b}}} + \underbrace{\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \dots & 0 \\ 0 & 0 & 0 & 1 & \ddots & 0 \\ 0 & & & \ddots & 1 & 0 \\ 0 & & & \ddots & 0 & 1 \\ 0 & & & & 0 & 0 \\ 0 & & & & 0 & 0 \end{pmatrix}}_{:=\tilde{N}_{l,\tilde{b}}},$$

onde é fácil ver que $\tilde{D}_{l,\tilde{b}} \cdot \tilde{N}_{l,\tilde{b}} = \tilde{N}_{l,\tilde{b}} \cdot \tilde{D}_{l,\tilde{b}}$ e que $N_{l,\tilde{b}}$ é nilpotente. Exatamente como antes, a exponencial e^{tA} é formada pela matriz constituída pelos blocos diagonais da forma $e^{tB_{q,b}}$ e $e^{\tilde{B}_{l,\tilde{b}}}$. Como já calculamos exponenciais do tipo $e^{tB_{q,b}}$, só nos resta calcular

$$e^{t\tilde{B}_{l,\tilde{b}}} = e^{tD_{l,\tilde{b}}} \cdot e^{tN_{l,\tilde{b}}} =$$

$$e^{ta_l}\cdot\begin{pmatrix} \cos(tb_l) & \sin(tb_l) & 0 & 0 & 0\ldots & 0 \\ -\sin(tb_l) & \cos(tb_l) & 0 & 0 & \ddots & 0 \\ 0 & & & \ddots & 0 & 0 \\ 0 & & & \ddots & 0 & 0 \\ 0 & & & & \cos(tb_l) & \sin(tb_l) \\ 0 & & & & -\sin(tb_l) & \cos(tb_l) \end{pmatrix}\cdot\begin{pmatrix} 1 & 0 & t & 0 & \frac{t^2}{2}\ldots & 0 \\ 0 & 1 & 0 & t & \ddots & 0 \\ 0 & & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 0 & t \\ 0 & & & & 1 & 0 \\ 0 & & & & 0 & 1 \end{pmatrix} =$$

$$e^{ta_l}\cdot\begin{pmatrix} \cos(tb_l) & \sin(tb_l) & t\cos(tb_l) & t\sin(tb_l) & \frac{t^2}{2}\cos(tb_l) & \frac{t^2}{2}\sin(tb_l) & \ldots \\ -\sin(tb_l) & \cos(tb_l) & -t\sin(tb_l) & t\cos(tb_l) & -\frac{t^2}{2}\sin(tb_l) & \frac{t^2}{2}\cos(tb_l) & \ldots \\ 0 & 0 & & & & & \\ & & \ddots & & & & \\ \vdots & & & & & t\cos(tb_l) & t\sin(tb_l) \\ 0 & & & & & -t\sin(tb_l) & t\cos(tb_l) \\ & & & & & \cos(tb_l) & \sin(tb_l) \\ 0 & & & & & -\sin(tb_l) & \cos(tb_l) \end{pmatrix}$$

6.5.1 Classificação dos campos lineares hiperbólicos

Começamos com duas definições:

Definição 6.5.1. (Índice de estabilidade de um campo linear). O *índice de estabilidade* de um campo linear $\dot{x} = A \cdot x$ é a dimensão da soma dos autoespaços generalizados associados a autovalores de A com parte real negativa (veja observação 6.4.11, uma vez que A é real).

Definição 6.5.2. (Campo linear hiperbólico). O campo linear $A : \mathbb{R}^n \to \mathbb{R}^n$ é dito *hiperbólico* se os autovalores de A têm parte real não nula. Nesse caso, também se diz que a equação $\dot{x} = A \cdot x$ é *hiperbólica*.

Nesta seção, construiremos conjugações topológicas entre campos lineares hiperbólicos em $\mathbb{R}^m$ que possuam o mesmo índice de estabilidade. Para tanto começaremos por estabelecer conjugações entre campos onde esse índice é máximo (campos lineares atratores). Todavia, começemos com um exemplo que resolve a questão na reta ($m = 1$).

Exemplo 6.5.3. (Campos lineares atratores da reta são conjugados). Sejam $a < 0$, $b < 0$, e considere os campos $\dot{x} = ax$ e $\dot{x} = bx$. Mostremos que ambos os campos são conjugados, exibindo uma conjugação entre ambos. Note que

se um campo é conjugado a outro via um homeomorfismo h, e φ e ψ são seus respectivos fluxos, então $h \circ \varphi_t = \psi_t \circ h \Leftrightarrow \psi_{-t} \circ h \circ \varphi_t = h, \forall t$. Essa última fórmula nos dá uma certa rigidez para que um homeomorfismo seja uma conjugação, a qual podemos usar em nosso favor. Note ainda que no caso de nossos campos lineares da reta, para todo $x \neq 0$, existe um único tempo $\tau(x)$ tal que $|\varphi(\tau(x), x)| = 1$ (respectivamente, existe um único $\hat{\tau}(x) \in \mathbb{R}$ tal que $|\psi(\hat{\tau}(x), x)| = 1$). Se h conjuga os campos acima, necessariamente satisfaz

$$\psi_{-\tau(x)} \circ h|_{\{1,-1\}} \circ \varphi_{\tau(x)}(x) = h(x).$$

Note que apenas o valor de $h|_{\{1,-1\}}$ está livre na expressão acima! Colocando $h|_{\{1,-1\}}$ como a identidade em $\{1, -1\}$, mostremos que de fato h é conjugação, seguindo os passos abaixo:

1. Calculemos $\tau(x)$:

$$|\varphi(\tau(x), x)| = 1 \Leftrightarrow e^{a\tau(x)}|x| = 1 \Leftrightarrow \tau(x) = \log(|x|^{-1})/a.$$

2. Explicitemos h. Para que h seja conjugação, precisa levar singularidade de um campo em singularidade do outro campo, logo pomos $h(0) = 0$. Para $x \neq 0$, a rigidez observada acima, aliada à escolha que fizemos de $h|_{\{-1,1\}}$ nos dá:

$$h(x) = e^{-b\tau(x)} \cdot e^{a\tau(x)} \cdot x = |x|^{b/a} \cdot |x|^{-1} \cdot x = |x|^{b/a-1} \cdot x.$$

   Claramente h, assim definido, é um homeomorfismo.

3. É fácil verificar diretamente que h é uma conjugação:

$$h \circ e^{at} \cdot x = e^{a(b/a-1)t} \cdot |x|^{b/a-1} \cdot e^{at} \cdot x = e^{bt} \cdot |x|^{b/a-1} \cdot x = e^{bt} \cdot h(x).$$

   Veremos adiante que a razão para isso decorre de modo simples da fórmula de rigidez e da definição de τ acima, desde que h assim obtida seja contínua.

Observação 6.5.4. No último exemplo, notamos que a conjugação h obtida é de classe C^1 se $b > a$. Contudo, neste caso, a derivada de h na origem é obrigatoriamente 0, dada a proposição 6.3.2.

Com o fim de generalizar para dimensões mais altas o resultado do exemplo 6.5.3, vamos rever alguns resultados da geometria diferencial.

Definição 6.5.5. (Hiperfície mergulhada). Um conjunto $\mathcal{M} \subset \mathbb{R}^m$, dotado da topologia de $\mathbb{R}^m$ é dito uma *hiperfície mergulhada* de $\mathbb{R}^m$, ou simplesmente, uma *hiperfície* de $\mathbb{R}^m$ se ele é coberto pela união das imagens de uma família de aplicações $\{p_\beta\}$, $p_\beta : U_\beta \subset \mathbb{R}^{m-1} \to \mathbb{R}^m$ tais que

- Cada U_β é um aberto de $\mathbb{R}^{m-1}$.
- A imagem de cada p_β está contida em $\mathcal{M}$.
- Cada p_β é um homeomorfismo (sobre um aberto de $\mathcal{M}$ na topologia induzida por $\mathbb{R}^m$).
- Cada p_β é uma imersão, ou seja, em cada ponto u de U_β, a derivada $Dp_\beta(u)$ possui posto máximo.

Lema 6.5.6. *Seja $\mathcal{M}$ uma hiperfície conexa mergulhada em $\mathbb{R}^m$, dividindo o espaço ambiente em duas componentes conexas $\mathcal{M}_0$ e $\mathcal{M}_1$ que a possuem como fronteira. Seja $N : \mathcal{M} \to \mathbb{R}^m$ um campo contínuo normal a $\mathcal{M}$. Então para alguma dessas componentes, digamos $\hat{\mathcal{M}}$ para cada $x \in \mathcal{M}$ existe $\epsilon_x > 0$ tal que $x + sN(x)$ pertence a $\hat{\mathcal{M}}, \forall 0 < s < \epsilon_x$. Nesse caso, dizemos que N aponta para $\hat{\mathcal{M}}$. Se $X(x)$ é um campo de vetores tal que $< N(x), X(x) >> 0, \forall x \in \mathcal{M}$, então X também aponta para a mesma componente $\hat{\mathcal{M}}$ que N.*

Prova: Seja $x \in \mathcal{M}$, e suponha por absurdo que não existisse o ϵ_x do enunciado. Podemos então tomar uma sequência $s_n \to 0$ tal que $x+s_nN(x) \in \mathcal{M}$. Caso não pudéssemos tomar tal sequência, existiriam sequências $0 < s_n^0 \to 0$ e $0 < s_n^1 \to 0$ tais que $x + s_n^0 N(x) \in \mathcal{M}_0, \forall n \in \mathbb{N}$ e $x + s_n^1 N(x) \in \mathcal{M}_1, \forall n \in \mathbb{N}$. Pelo teorema da Alfândega, existiria $0 < s_n \to 0$ tal que $x + s_n N(x) \in \mathcal{M}, \forall n \in \mathbb{N}$, como queríamos.

Como a hiperfície é mergulhada, podemos tomar uma parametrização $p : B(0,r) \subset \mathbb{R}^{m-1} \to \mathbb{R}^m$ de uma vizinhança de x em $\mathcal{M}$ (a qual será uma vizinhança na topologia de $\mathcal{M}$ induzida por $\mathbb{R}^m$) e escrever $p(0) = x$, $p(h_n) = x + s_n N(x)$, com $h_n \to 0$ quando $n \to +\infty$. Sem perda, usando da compacidade de S^{m-1}, suponha que $h_n/\|h_n\| \to h \in S^{m-1}$. Daí,

$$x + s_n N(x) - x = p(h_n) - p(0) \Rightarrow \lim_{n\to+\infty} \frac{s_n}{\|h_n\|} N(x) =$$

$$\lim_{n\to+\infty} \frac{p(h_n) - p(0)}{\|h_n\|} = p'(0) \cdot h \in T_x\mathcal{M}.$$

$p'(0)\cdot h \neq 0$ porque p é parametrização. Mas note que $\frac{s_n}{\|h_n\|}N(x)$ converge para algum elemento do espaço normal a $\mathcal{M}$, do que acabamos de ver não nulo. Tal é absurdo, porque o único vetor a pertencer simultaneamente a $T_x\mathcal{M}$ e a um espaço em soma direta com $T_x\mathcal{M}$ é o vetor nulo. Sem perda, podemos então considerar que existe $\epsilon_x > 0$ tal que $x + sN(x) \in \mathcal{M}_0, \forall 0 < s < \epsilon$. Se existisse outro $y \in \mathcal{M}$ tal que para $0 < t < \epsilon_y$ $y + tN(y) \in \mathcal{M}_1$. Tomando $0 < \hat{t} < \min\{\epsilon_x, \epsilon_y\}$, e $\gamma_{xy} \subset \mathcal{M}$ uma curva compacta unindo x e y, definimos a aplicação contínua $\alpha_{\hat{t}} : \gamma_{xy} \to \mathbb{R}^m$ dada por

$$\alpha_{\hat{t}}(z) = z + \hat{t} \cdot N(z)$$

Pelo teorema da Alfândega, como $\alpha_{\hat{t}}(x) \in \mathcal{M}_0$ e $\alpha_{\hat{t}}(y) \in \mathcal{M}_1$, existe $w \in \gamma_{xy}$ tal que $\alpha_{\hat{t}}(w) \in \mathcal{M}$. Fazendo $\hat{t} = 1/n$, o argumento acima nos dá uma sequência de pontos $w_n \in \gamma_{xy}$ tais que $w_n + 1/nN(w_n) \in \mathcal{M}$. Dada a compacidade de γ_{xy}, podemos supor sem perda que $w_n \to w \in \gamma_{xy}$ quando $n \to +\infty$. Daí, passando a uma parametrização p tal que $p(0) = w$, $p(h_n) = w_n + 1/nN(w_n)$, $p(l_n) = w_b$ com $h_n \to 0$, $l_n \to 0$ e $\lim_{n\to+\infty}(h_n - l_n)/\|h_n - l_n\| = h \in S^{m-1}$, obtemos como antes que

$$0 \neq \lim_{n\to+\infty} \frac{p(h_n) - p(l_n)}{\|h_n - l_n\|} = p'(0) \cdot h = \lim_{n\to+\infty} \frac{N(w_n)}{n \cdot \|h_n - l_n\|}.$$

A primeira igualdade decorre de corolário da desigualdade do Valor Médio quando a aplicação é de classe C^1, que é o caso de p. Observe que como $N(w_n) \to N(w)$, necessariamente $n \cdot \|h_n - l_n\|$ converge na reta para algum número positivo, donde a exemplo do que tivemos acima teríamos $N(w) \in T_w\mathcal{M}$, o que é absurdo.

Concluímos portanto que para cada $x \in \mathcal{M}$, existe ϵ_x tal que $x + sN(x)$ pertence a uma das componentes conexas do enunciado, para todo $0 < s < \epsilon_x$. Sem perda, podemos supor que $\mathcal{M}_0$ é essa componente conexa. Agora considere um campo X tal que $< X(x), N(x) >> 0$. Então podemos repetir o mesmo raciocínio acima e concluir que X aponta para uma das componentes conexas do enunciado. Mostremos que essa componente é a mesma de N , isto é, $\mathcal{M}_0$. Para provarmos isso, basta nos atermos a olharmos o que ocorre em um ponto $x \in \mathcal{M}$ fixado. Procedendo por absurdo, tome então $\epsilon > 0$ tal que $x + sN(x) \in \mathcal{M}_0, \forall 0 < s < \epsilon$ e $x + sX(x) \in \mathcal{M}_1, \forall 0 < s < \epsilon$. Sem perda, podemos supor $\|X(x)\| = 1$ (não há diferença substancial na prova em trocar $X(x)$ por um outro múltiplo não nulo com mesmo sentido). Novamente para cada $s \in (0, \epsilon)$, podemos definir o caminho contínuo $\alpha_s : [0, 1] \to \mathbb{R}^m$ dado

por

$$\alpha_s(t) = x + s((1-t)N(x) + tX(x)).$$

Claramente α_s é contínua, com $\alpha_s(0) \in \mathcal{M}_0$ e $\alpha_s(1) \in \mathcal{M}_1$. Logo, pelo para cada $s \in (0, \epsilon)$, existe $t_s \in (0,1)$ tal que $\alpha_s(t_s) \in \mathcal{M}$. Ademais, pondo $v(t_s) := ((1-t_s)N(x) + t_s X(x))$, do fato que $< N(x), X(x) >> 0$, existe $c > 0$ tal que para todo $s \in (0, \epsilon)$ temos $\|v(t_s)\| \geq 1/2$, e $< N(x), v(t_s) >> c > 0$. Passando a uma subsequência convergente $v_n = v(t_{s_n})$ de $v(t_s)$ com $v_n \to v$ quando $n \to +\infty$, passando a uma parametrização tal que $p(0) = x$, $p(h_n) = x + s_n v_n$, supondo sem perda que $h_n/\|h_n\| \to h \in S^{m-1}$, como antes obtemos:

$$0 \neq \frac{p(h_n) - p(0)}{\|h_n\|} = p'(0) \cdot h = \lim_{n \to +\infty} \frac{s_n v_n}{\|h_n\|}.$$

Como antes, teríamos que $\lim_{n \to +\infty} \frac{s_n v_n}{\|h_n\|}$ seria ao mesmo tempo um múltiplo não nulo de $v \notin T_x\mathcal{M}$ e $p'(0) \cdot h \in T_x\mathcal{M}$. Absurdo.

□

Corolário 6.5.7. *Seja $\varphi : I \to \mathbb{R}^m$ uma curva C^1 transversal a uma hiperfície compacta conexa $\mathcal{M}$ de classe C^1 contida em $\mathbb{R}^m$. Suponha que $\varphi(0) \in \mathcal{M}$. Considere $\mathcal{M}_0$ e $\mathcal{M}_1$, respectivamente as regiões abertas limitada e ilimitada que possuem $\mathcal{M}$ como fronteira. Se $\varphi'(0)$ aponta para $\mathcal{M}_0$, então existe $\delta > 0$ tal que $\varphi(t) \in \mathcal{M}_0$, $\forall 0 < t < \delta$.*

Prova: Usando o último lema, temos que existe δ_0 tal que $\delta_0 > t > 0$ implica $\varphi(0) + \varphi'(0)t \in \mathcal{M}_0$. Suponha por absurdo, que exista uma sequência $t_n \searrow 0, t_n < \delta_0$ tal que $\varphi(t_n) \in (\mathcal{M}_0)^c$. Ora, tal sequência não poderia pertencer a $\mathcal{M}$, caso contrário, considerando uma parametrização p de $\mathcal{M}$ sobre uma vizinhança de $\varphi(0)$, com $p(0) = \varphi(0), p(x_n) = \varphi(t_n)$ teríamos

$$\frac{\varphi(t_n) - \varphi(0)}{t_n - 0} = \frac{p(x_n) - p(0)}{t_n - 0} = \frac{p(x_n) - p(0)}{\|x_n - 0\|} \frac{\|x_n - 0\|}{t_n - 0}$$

Note que a primeira expresso das igualdades acima converge para $\varphi'(0) \neq 0$ quando $n \to \infty$. Por outro lado, passando a uma subsequência se necessário, podemos supor que $x_n/\|x_n\|$ converge a $v \in S^1$. Desse modo,

$$\frac{p(x_n) - p(0)}{\|x_n - 0\|} = \frac{p'(0)(x_n - 0) + r(x_n - 0)}{\|x_n - 0\|},$$

que converge a $p'(0) \cdot v \in T_{\varphi(0)}\mathcal{M}$. Isso implica ainda que $\frac{\|x_n - 0\|}{t_n - 0}$ converge e a contradição. Portanto, proseguindo em nosso argumento por absurdo,

podemos supor que $\varphi(t_n) \in \mathcal{M}_1$, $\forall n \in \mathbb{N}$. Neste caso, pelo teorema da Alfândega, o segmento de reta que une $\varphi(t_n) \in \mathcal{M}_1$ a $\varphi(0) + vr'(0)t_n \in \mathcal{M}_0$ necessariamente intersecta $\mathcal{M}$. Ou seja, existe $u_n \in (0,1)$ tal que

$$u_n \cdot \varphi(t_n) + (1 - u_n) \cdot (\varphi(0) + \varphi'(0)t_n) \in \mathcal{M}.$$

Mas tal implica que:

$$u_n \cdot (\varphi(0) + \varphi'(0)t_n + r(t_n)) + (1 - u_n) \cdot (\varphi(0) + \varphi'(0)t_n) = \varphi(0) + \varphi'(0)t_n + u_n r(t_n) \in \mathcal{M},$$

com $r(t_n)/\|t_n\| \to 0$ quando $t_n \to 0$. Daí, por um lado,

$$\frac{\varphi(0) + \varphi'(0)t_n + u_n r(t_n)}{t_n} \to \varphi'(0) \neq 0,$$

e por outro, escrevendo $p(x_n) = \varphi(0) + \varphi'(0)t_n + u_n r(t_n), p(0) = \varphi(0)$ como fizemos mais acima, supondo sem perda $x_n/\|x_n\| \to v$, temos que

$$\frac{p(x_n) - p(0)}{\|x_n\|} \to p'(0)v \in T_{\varphi(0)}\mathcal{M}.$$

Concluímos com isso que

$$\frac{\varphi(0) + \varphi'(0)t_n + u_n r(t_n)}{t_n} = \frac{p(x_n) - p(0)}{\|x_n\|} \cdot \frac{\|x_n\|}{t_n}$$

converge simultaneamente a $\varphi'(0)$ e a um múltiplo de $p'(0) \cdot v \in T_{\varphi(0)}\mathcal{M}$, absurdo.

□

Observação 6.5.8. Observamos que se um campo de vetores X aponta para $\mathcal{M}_0$, necessariamente seu simétrico aponta para $\mathcal{M}_1$. De fato, como a hiperfície é mergulhada, fixado $y \in M$, existe uma bola $B(y,r) \subset \mathbb{R}^m$ tal que $\mathcal{M} \cap B(0,r)$ é o gráfico de uma aplicação C^1 do plano tangente $y + T_yM \cap B(0,r)$ em $y + span(N(y))$. Claramente um tal gráfico divide a bola $B(0,r)$ em duas componentes conexas, as quais pertencem de maneira exclusiva a uma das componentes conexas disjuntas, caso contrário $\mathcal{M}$ não seria fronteira comum a ambas. Se $N(y)$, por exemplo, aponta para $\mathcal{M}_0 \cap B(0,r)$, obviamente, $-N(y)$ aponta para $\mathcal{M}_1 \cap B(0,r)$. Do mesmo modo, se $< X(y), N(y) >> 0$, temos $< -X(y), -N(y) >> 0$, donde concluímos que $-X$ aponta para $\mathcal{M}_1$.

Lema 6.5.9. *Seja $\mathcal{M} \subset \mathbb{R}^m$ uma hiperfície (portanto, com dimensão $m-1$), compacta, conexa (no caso $m = 1$, esta hipótese deve ser substituída por $\mathcal{M} = \{p_1, p_2\}$). Denote por $\mathcal{M}_0$ a região aberta limitada que $\mathcal{M}$ determina como fronteira, e suponha que $0 \in \mathcal{M}_0$. Sejam $X : \mathbb{R}^m \to \mathbb{R}^m$ e $Y : \mathbb{R}^m \to \mathbb{R}^m$ campos completos transversais a $\mathcal{M}$ e tais que seus fluxos*

$$X_t(x) \to 0, \text{ se } t \to +\infty, \quad \forall x \in \mathbb{R}^m Y_t(x) \to 0, \text{ se } t \to +\infty, \quad \forall x \in \mathbb{R}^m$$

e todas as trajetórias de X e Y, exceto 0, intersectam $\mathcal{M}$. Então X é conjugado a Y.

Prova: Por razões didáticas, mais uma vez dividiremos a demonstração em vários itens.

1. **Tanto X como Y apontam "para dentro" de $\mathcal{M}$.** De fato, suponha por absurdo que X aponte para fora de $\mathcal{M}$. Então dado $x \in \mathcal{M}$, existe $\delta_x > 0$ tal que $X_t(x) \in \mathcal{M}_1$, $\forall 0 < t < \delta_x$. Como $X_t(x) \to 0 \in \mathcal{M}_0$, se $t \to +\infty$, existe $t_0 > 0$ tal que $X_t(x) \in \mathcal{M}_0$, $\forall t \geq t_0$. Seja $t_1 = \inf\{t \leq t_0; X_s(x) \in \mathcal{M}_0, \forall t \leq s \leq t_0\}$. É claro que o conjunto do qual consideramos o ínfimo na última sentença está contido em $(\delta_x, t_0]$ e contém t_0. Também é claro que, como ínfimo, $X_{t_1}(x)$ não pode pertencer a $\mathcal{M}_0$, pois esta é aberta e $X_t(x)$ é contínua em t (se $X_{t_1}(x) \in \mathcal{M}_0$, existe uma vizinhança de t_1 que também pertence). Por razões análogas, não podemos ter $X_{t_1}(x) \in \mathcal{M}_1$, donde concluímos que $t_1 \in \mathcal{M}$. Mas como por absurdo supusemos que X aponta para fora, isto implica que existe $\delta_{X_{t_1}(x)} > 0$ tal que $X_{t_1+s} \in \mathcal{M}_1, \forall 0 < s < \delta_{X_{t_1}(x)}$, o que contradiz a definição de t_1. Por conseguinte, X aponta para dentro. O mesmo raciocínio aplicado a Y prova que Y também aponta para dentro.

2. Definimos $h : \mathbb{R}^m \to \mathbb{R}^m$ como:

$$h(0) := 0, h|_{\mathcal{M}} := I$$
$$h(x) = Y_{-\hat{\tau}(x)} \circ h \circ X_{\hat{\tau}(x)}(x), \text{ se } x \neq 0,$$

   onde $\hat{\tau} : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}$ **é o <u>único</u> tempo tal que** $X_{\hat{\tau}(x)} \in \mathcal{M}$. Para vermos que $\hat{\tau}$ está bem definido (é mesmo único, já que existe, por hipótese), comecemos por considerar o caso em $x \in \overline{\mathcal{M}_0}$. Neste caso, como X aponta para dentro e $\mathcal{M}_0$ é aberto, segue-se que existe $\delta_x > 0$ tal que $\forall t \in (0, \delta_x)$, vale $X_t(x) \in \mathcal{M}_0$. Mostremos que $X_t(x) \in$

$\mathcal{M}_0, \forall t > 0$. Para tal seja $t_1 = \sup\{t > 0; X_s(x) \in \mathcal{M}_0, \forall s \in (0,t)\}$. Note que o conjunto cujo sup consideramos não é vazio, e que $t_1 \geq \delta_x > 0$. Se $t_1 = +\infty$, nada temos a provar. Suponha que $t_1 < +\infty$. Neste caso, como $\mathcal{M}_0$ é aberto e $X_t(x)$ é contínuo, temos que $X_{t_1}(x) =: y$ necessariamente pertence a $\mathcal{M}$. Mas como $-X$ aponta para fora (vide observação 6.5.8), segue-se que existe $\delta_y > 0$ tal que $X_{-t}(y) \in \mathcal{M}_1, \forall -\delta_y < -t < 0$. Ora, mas isso é o mesmo que dizer que $X_{t_1-t}(x) \in \mathcal{M}_1$, $\forall -\delta_y < -t < 0$, o que contradiz a definição de t_1, a qual implica que $X_{t_1}(x)$ é acumulado por pontos $X_s(x) \in \mathcal{M}_0$, com $s \nearrow t_1$.

Um raciocínio inteiramente análogo prova que

$$x \in \mathcal{M}_1 \Rightarrow X_t(x) \in \mathcal{M}_1, \forall t < 0,$$

ou seja, se $x \in \mathcal{M}_1$, não existe tempo negativo t tal que $X_t(x) \in \mathcal{M}$.

Concluímos então que:

- Se $x \in \mathcal{M}$, vale que $X_t(x) \in \mathcal{M}_0, \forall t > 0$ e $X_t(x) \in \mathcal{M}_1, \forall t < 0$. Nesse caso, $\hat{\tau}(x) = 0$ é o único tempo que a órbita de x intersecta $\mathcal{M}$.
- Se $x \in \mathcal{M}_0$, defina $\hat{\tau}(x) := \sup\{t < 0; X_t(x) \in \mathcal{M}\}$. Claramente $X_{\hat{\tau}(x)}(x) \in \mathcal{M}$, e em particular $\hat{\tau}(x) < 0$. Chamando $y = X_{\hat{\tau}(x)}$, temos do item anterior aplicado a y que $X_{t+\hat{\tau}(x)}(x) = X_t(y) \notin \mathcal{M}, \forall t \neq 0$, logo $-\hat{\tau}(x)$ é novamente o único tempo em que a órbita de x intersecta $\mathcal{M}$.
- Se $x \in \mathcal{M}_1$, definindo $\hat{\tau}(x) := \inf\{t > 0; X_t(x) \in \mathcal{M}\}$, pelo mesmo argumento do item acima, vemos que neste último caso também vale que se $X_t(x) \in \mathcal{M} \Leftrightarrow t = \hat{\tau}(x)$.

3. $\hat{\tau}$ **é** C^1 **devido ao teorema da função implícita** (ver resultados acerca de transformação de Poincaré, no capítulo sobre campos). Para vermos isso, fixado $x \in \mathbb{R}^m \setminus 0$ seja $T = \hat{\tau}(x)$. Seja ainda V uma vizinhança em torno de $\varphi_T(x)$ dada pelo teorema do fluxo tubular, e consideremos $\tau : V \to (-\epsilon, \epsilon)$ definida no contexto de corolário do Teorema do Fluxo Tubular. Da continuidade de φ_T, temos que existe uma vizinhança $W \ni x$ tal que $\varphi_T(W) \subset V$. Por conseguinte, da unicidade de $\hat{\tau}$ provada no item anterior, é imediato que $\hat{\tau}|_W = T + \tau$, concluindo que $\hat{\tau}$ é de classe C^1.

4. Seja $t_1 > t_2$. ***Mostremos que*** $X_{t_1}(\overline{\mathcal{M}}_0) \subset X_{t_2}(\mathcal{M}_0)$. Ora, mas

$$X_{t_1}(\overline{\mathcal{M}}_0) \subset X_{t_2}(\mathcal{M}_0) \Leftrightarrow X_{t_1-t_2}(\overline{\mathcal{M}}_0) \subset \mathcal{M}_0,$$

   o que é verdade como vimos nos primeiros itens, pois se $x \in \overline{\mathcal{M}}_0$, então $X_t(x) \in \mathcal{M}_0, \forall t > 0$.

   Concluímos que $X_{t_1}(\overline{\mathcal{M}}_0) \subset X_{t_2}(\mathcal{M}_0)$, se $t_1 > t_2$. Resultado idêntico vale para Y_t no lugar de X_t.

5. Pelo item 3, o problema de continuidade de h é apenas na origem. Demonstremos esta continuidade. Seja $x_n \to 0$. Primeiramente, mostremos que $-\hat{\tau}(x_n) \to +\infty$. Como $\mathcal{M}$ é compacta e X_j é contínua, e $0 \notin X_j(\mathcal{M})$, para cada $j \in \mathbb{N}$, existe $\epsilon_j > 0$ tal que $d(X_j(\mathcal{M}), 0) \geq \epsilon_j$, com $\epsilon_j \searrow 0$.

   Afirmamos que temos também que $d(X_t(\mathcal{M}), 0) \geq \epsilon_j, \forall t \leq j$. De fato, visto que X_t e X_j são difeomorfismos, levam abertos em abertos e fronteiras em fronteiras. Como pelo item 2 $X_j(\overline{\mathcal{M}}_0) \subset X_t(\mathcal{M}_0)$ se $t < j$, segue-se que $\partial X_t(\mathcal{M}_0) = X_t(\mathcal{M})$ tem intersecção vazia com $X_j(\overline{\mathcal{M}}_0)$. Considere então $\hat{x} \in X_t(x)$ tal que $d(\hat{x}, 0) = \inf_{x \in \mathcal{M}}\{d(X_t(x), 0)\}$. Escrevendo $[0, \hat{x}]$ para o segmento de reta unindo 0 e $\hat{x}$, como $0 \in X_j(\mathcal{M}_0)$ e $\hat{x} \in (X_j(\hat{\mathcal{M}}_0))^c$ o teorema da Alfândega implica que existe $\tilde{x} \in X_j(\mathcal{M}) \cap (0, \hat{x})$. Logo

$$\inf_{x \in \mathcal{M}}\{d(X_t(x), 0)\} = d(\hat{x}, 0) \geq d(\tilde{x}, 0) \geq \inf_{x \in \mathcal{M}}\{d(X_j(x), 0)\} \geq \epsilon_j$$

   o que prova nossa afirmação.

   Para ver que $-\hat{\tau}(x_n) \to +\infty$ quando $x_n \to 0$, seja $j \in \mathbb{N}$ dado. Portanto, como $x_n \to 0$, dado ϵ_j da sequência acima, existe n_j tal que para todo $n > n_j$, $|x_n| < \epsilon_j$, o que implica que $-\hat{\tau}(x_n) > j$, $\forall n > n_j$. Ou seja, temos que $-\hat{\tau}(x_n) \to +\infty$ se $n \to +\infty$.

   Note que $\{0\} = \cap_{j \in \mathbb{N}} Y_j(\overline{\mathcal{M}_0}) = \cap_{t \in \mathbb{R}^+} Y_t(\overline{\mathcal{M}_0})$ e que $Y_t(\mathcal{M}) \to 0$ uniformemente quando $t \to \infty$. De fato, se $y \in \cap_{j \in \mathbb{N}} Y_j(\overline{\mathcal{M}_0}) \neq 0$, tal implica que existe $(y_j), y_j \in \overline{\mathcal{M}_0}$ tal que $Y_j(y_j) = y, \forall j \in \mathbb{N} \Leftrightarrow y_j = Y_{-j}(y) \in \overline{\mathcal{M}_0}, \forall j \in \mathbb{N}$. Mas definindo $y_0 := Y_{\tilde{\tau}(y)}(y) \in \mathcal{M}$, da unicidade de $\tilde{\tau}(y)$ como único tempo em que a trajetória de Y com valor inicial em y intersecta a $\mathcal{M}$, temos que $Y_{-t}(y) \in \mathcal{M}_1$, $\forall t > -\tilde{\tau}(y)$. Em particular, $y_j \notin \overline{\mathcal{M}_0}, \forall j > -\tilde{\tau}(y)$, absurdo. Donde concluímos que

$y = 0$. Agora, para cada $j \in \mathbb{N}$, seja $y_j \in Y_j(\overline{\mathcal{M}_0})$ um ponto de maior distância a zero. Como tais $Y_j(\overline{\mathcal{M}_0})$ são encaixantes, segue-se que a sequência de $|y_j|$ é monótona e converge. Se converge para zero, isto equivale a que $Y_t(\overline{\mathcal{M}_0}) \to 0$ uniformemente quando $t \to \infty$. Sem perda de generalidade, podemos, passando a uma subsequência se necessário, supor que $y_j \to y \in \overline{\mathcal{M}_0}$. Como $y_{\tilde{j}} \in Y_j(\overline{\mathcal{M}_0}), \forall \tilde{j} \geq j$, segue-se que $y \in Y_j(\overline{\mathcal{M}_0})$, $\forall j$. Portanto,

$$y \in \cap_{j\in\mathbb{N}} Y_j(\overline{\mathcal{M}_0}) = \cap_{t\in\mathbb{R}^+} = \{0\},$$

isto é, $y = 0$.

Em resumo, se $x_n \to 0$, temos que $-\hat{\tau}(x_n) \to +\infty$ o que implica:

$$|h(x_n)| = |Y_{-\hat{\tau}(x_n)} \circ \underbrace{h \circ X_{\hat{\tau}(x_n)}(x_n)}_{\in\mathcal{M}}| \leq \operatorname{diam}(Y_{-\hat{\tau}(x_n)}(\mathcal{M})) \to 0,$$

pois vimos que $Y_t(\mathcal{M}) \to 0$ uniformemente, se $t \to +\infty$. Isso prova a continuidade de h na origem.

6. h é conjugação. De fato, dados $t \in \mathbb{R}$ e $x \in \mathbb{R}^m$, temos:

$$\begin{gathered} Y_t \circ h \circ X_{-t}(x) = Y_t \circ Y_{-\hat{\tau}(X_{-t}(x))} h X_{\hat{\tau}(X_{-t}(x))} \circ X_{-t}(x) = \\ Y_t \circ Y_{-\hat{\tau}(x)-t} h X_{t+\hat{\tau}(x)} \circ X_{-t}(x) = \\ Y_{t-t} \circ Y_{-\hat{\tau}(x)} h X_{-t+t} \circ X_{\hat{\tau}(x)}(x) = \\ Y_{-\hat{\tau}(x)} \circ h \circ X_{\hat{\tau}(x)}(x) = h(x). \end{gathered}$$

Note que usamos nas igualdades intermediárias acima o fato de que $\hat{\tau}(X_{-t}(x)) = \hat{\tau}(x) + t$, além da propriedade de grupo dos fluxos. Portanto h é uma semiconjugação. Com uma fórmula similar a de h, definimos uma aplicação $H : \mathbb{R}^m \to \mathbb{R}^m$, que mostraremos ser igual a h^{-1}, o que implica que h^{-1} é contínua:

$$H(y) := X_{-\tilde{\tau}(y)} \circ h^{-1} \circ Y_{\tilde{\tau}(y)}(y),$$

onde $\tilde{\tau}(y)$ é o único tempo tal que $Y_{\tilde{\tau}(y)}(y) \in \mathcal{M}$. É simples ver que $H = h^{-1}$. Por exemplo, vejamos que $H \circ h(x) = x, \forall x \in \mathbb{R}^m$:

$$H \circ h(x) = X_{-\tilde{\tau}(y)} \circ h^{-1} \circ Y_{\tilde{\tau}(y)} \circ \underbrace{Y_{-\hat{\tau}(x)} \circ h \circ X_{\hat{\tau}(x)}(x)}_{y} = x,$$

pois $h \circ X_{\hat{\tau}(x)}(x) \in \mathcal{M}$, implicando que $\tilde{\tau}(y) = \hat{\tau}(x)$.

Logo, h é um homeomorfismo e conjuga os fluxos de Y e X.

□

O escólio (consequência da demonstração) a seguir nos diz que podemos enfraquecer a hipótese de que ambos os campos (X e Y) tenham de ter o mesmo domínio e serem transversais à mesma hiperfície:

Escólio 6.5.10. *Sejam $E, \hat{E}$ dois espaços de Banach reais de mesma dimensão finita m, e $X : E \to E$, $Y : \hat{E} \to \hat{E}$ dois campos completos. Suponha que existam $\mathcal{M} \subset E$, $\hat{\mathcal{M}} \subset \hat{E}$ hiperfícies (portanto, com dimensão $m-1$), compactas, conexas (no caso $m = 1$, esta hipótese deve ser substituída por $\mathcal{M}$, $\hat{\mathcal{M}}$ da forma $\{p_1, p_2\}$) e homeomorfas. Denote por $\mathcal{M}_0$, $\hat{\mathcal{M}}_0$ as regiões abertas limitadas que $\mathcal{M}$ e $\hat{\mathcal{M}}$ respectivamente determinam como fronteira, e suponha que existam $p \in \mathcal{M}_0$, $\hat{p} \in \mathcal{M}_0$ tais que:*

1. $$\begin{aligned} X_t(x) &\to p, \ \text{se } t \to +\infty, \quad \forall x \in E \\ Y_t(y) &\to \hat{p}, \ \text{se } t \to +\infty, \quad \forall y \in \hat{E}; \end{aligned}$$

2. *X é transversal a $\mathcal{M}$ e Y é tranversal a $\hat{\mathcal{M}}$;*

3. *Todas as órbitas de X, exceto p intersectam $\mathcal{M}$ e todas as órbitas de Y, exceto $\hat{p}$, intersectam $\hat{\mathcal{M}}$.*

Então os campos X e Y são topologicamente conjugados.

Prova: Simples adaptação da prova do último lema, em que a conjugação $h|_{\mathcal{M}}$ em vez de ser a identidade em $\mathcal{M}$, é substituída pelo homeomorfismo que leva $\mathcal{M}$ em $\hat{\mathcal{M}}$.

□

Observação 6.5.11. Lembramos que em qualquer espaço E de dimensão finita m, o espaço das funções $m-$lineares alternadas em E é um espaço vetorial de dimensão 1. Por conseguinte, dois geradores quaisquer são múltiplos não nulos um do outro. Seja $\hat{\det}$ um tal gerador. Dada uma aplicação linear $A : E \to E$, esta induz uma forma $m-$linear dada por:

$$A^{\Phi}(v_1, \ldots, v_m) := \hat{\det}(A \cdot v_1, \ldots, A \cdot v_m), \forall (v_1, \ldots, v_m) \in E^m.$$

Dado $A \in \mathcal{L}(E)$, definimos então o determinante de A como o único escalar $det(A)$ tal que:

$$A^{\Phi} = \det(A) \cdot \hat{\det};$$

em outras palavras,

$$\det(A) = \hat{\det}(A \cdot v_1, \ldots, A \cdot v_m)/\hat{\det}(v_1, \ldots, v_m),$$

onde $(v_1, \ldots, v_m)$ é uma base qualquer de E. Note que se fixarmos uma base qualquer de E, e escrevermos a matriz de A nessa base, a noção acima coincide com a noção tradicional de determinante. Vemos assim que a noção de determinante, bem como de polinômio característico independe de base ou do isomorfismo que consideremos em um dado espaço $E \simeq \mathbb{R}^m$.

Proposição 6.5.12. *Seja $\dot{x} = A \cdot x$ um campo linear a coeficientes constantes de $E \simeq \mathbb{R}^m$ tal que a parte real de todos os autovalores de A é negativa. Então $\dot{x} = A \cdot x$ é topologicamente conjugado a $\dot{x} = -x$.*

Prova: Pela observação anterior, basta considerarmos $E = \mathbb{R}^m$. 'E fácil de constatar que $e^{tA} \cdot x \to 0$, $e^{-t} \cdot x \to 0$ quando $t \to +\infty$. Além disso, quando $t \to -\infty$, tais valores tendem a ∞. Deste modo, toda órbita (à exceção de $\{0\}$) destes campos intersectará uma hiperfície $\mathcal{M}$ (de codimensão 1) compacta que contenha 0 na região aberta limitada cuja fronteira é $\mathcal{M}$. Encontremos uma tal hiperfície $\mathcal{M}$ que seja transversal a ambos os campos lineares do enunciado.

Tome $q(x) := \int_0^{+\infty} < e^{sA} \cdot x, e^{sA} \cdot x > ds$. Tal integral converge, porque $\|e^{sA}\|^2 \to 0$ exponencialmente rápido quando $s \to +\infty$ (deixamos ao leitor os detalhes). Portanto, $q(x)$ é positiva definida. Além disso,

$$\frac{d(q(e^{tA} \cdot x)}{dt} = - < e^{tA} \cdot x, e^{tA} \cdot x >,$$

pois

$$q(e^{tA}(x)) = \int_t^{+\infty} < e^{vA} \cdot x, e^{vA} \cdot x > dv.$$

Em particular, como

$$0 > \frac{d(q(e^{tA} \cdot x)}{dt} = Dq(e^{tA} \cdot x) \cdot \frac{d(e^{tA} \cdot x)}{dt} = Dq(e^{tA} \cdot x) \cdot A \cdot e^{tA} \cdot x,$$

avaliando a expressão acima em $t = 0$, obtemos

$$Dq(x) \cdot (A \cdot x) =< Dq(x), A \cdot x >< 0,$$

o que implica que $A \cdot x$ é transversal a qualquer elipsóide

$$\{x; q(x) = c; c \text{ é constante positiva.}\},$$

pois $Dq(x)$ é ortogonal ao espaço tangente a esse elipsóide, que é pré-imagem por q do valor regular c.

Do mesmo modo, vemos que

$$\frac{d(q(e^{-t} \cdot x))}{dt} = \frac{d(e^{-2t})}{dt} \cdot \int_0^{+\infty} < e^{sA} \cdot x, e^{sA} \cdot x > ds < 0,$$

implicando que o campo $-x$ é transversal a qualquer elipsóide

$$\{x; q(x) = c; c \text{ é constante positiva.}\}$$

Fixando uma constante c positiva qualquer e tomando $\mathcal{M} := \{x; q(x) = c\}$, segue-se do lema 6.5.9 que os campos $\dot{x} = A \cdot x$ e $\dot{x} = -x$ são conjugados.

□

Corolário 6.5.13. *Dois campos lineares $\dot{x} = A \cdot x$ e $\dot{x} = B \cdot x$, hiperbólicos a coeficientes constantes de $\mathbb{R}^m$ com o mesmo índice de estabilidade são topologicamente conjugados.*

Prova: Seja s o índice de estabilidade de A. A última proposição implica que $\cdot x = A|_{E^s} x$ é conjugado ao campo $\dot{x} = -x$ em $\mathbb{R}^s$. Analogamente, é fácil adaptar a última proposição de modo a provar que $\dot{x} = A|_{E^u} x$ é conjugada ao campo em $\mathbb{R}^{m-s}$ dado por $\dot{x} = x$. Estendemos h^s e h^u ao $\mathbb{R}^m$ da seguinte forma, dotando as extensões com a mesma notação:

$$h^s(x) = h^s \circ \pi_s(x);$$
$$h^u(x) = h^u \circ \pi_u(x),$$

onde $\pi_u : \mathbb{R}^m \to E^u$ e $\pi_s : \mathbb{R}^m \to E^s$ são as projeções naturais associadas à decomposição $\mathbb{R}^m = E^s \oplus E^u$.

Seja portanto $g : \mathbb{R}^m \to \mathbb{R}^m$ definida por

$$g(x) := h^s(x) + h^u(x).$$

É fácil ver que g é um homeomorfismo. Daí, dado $x = x_s + x_u \in \mathbb{R}^m$, com $x_s \in E^s$ e $x_u \in E^u$, temos:

$$g \circ A(x) = h^s \circ A(x) + h^u \circ A(x) = h^s \circ A(x_s + x_u) + h^u \circ A(x_s + x_u) =$$

$$h^s \circ \underbrace{A(x_s)}_{\in E^s} + h^s \circ \underbrace{A(x_u)}_{\in E^u} + h^u \circ \underbrace{A(x_s)}_{\in E^s} + h^u \circ \underbrace{A(x_u)}_{\in E^u} =$$

$$h^s \circ A(x_s) + h^u \circ A(x_u) = B \circ h^s(x_s) + B \circ h^u(x_u) = B(g(x)),$$

donde concluímos que g conjuga A e B em $\mathbb{R}^m$.

□

6.5.2 Equações lineares de ordem superior na Reta

Podemos aplicar a teoria vista anteriormente na resolução de equações lineares de ordem superior a coeficientes constantes na reta:

Definição 6.5.14. (Equação linear de ordem superior a coeficientes constantes na reta). Uma equação linear de ordem n a coeficientes constantes na reta é uma equação do tipo

$$x^{(n)} = f(t, x, \ldots, x^{(n-1)}) = -a_0 \cdot x - \cdots - a_{n-1} \cdots x^{(n-1)},$$

onde $x^{(j)} = \frac{d^j}{dt^j} x$, $f : \mathbb{R}^n \to \mathbb{R}$ e $a_0, \ldots, a_{n-1} \in \mathbb{R}$.

Como é bem sabido, utilizando-nos de variáveis auxiliares

$$y_0 = x, y_1 = \dot{x}, y_2 = x^{(2)}, \ldots, y_{n-1} = x^{(n-1)},$$

podemos transformar a equação de ordem superior da definição acima na seguinte equação linear de ordem 1 no $\mathbb{R}^n$:

$$\begin{pmatrix} \dot{y}_0 \\ y_1 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ -a_0 \cdot y_0 - \cdots - a_{n-1} \cdot y_{n-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 \ldots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \\ -a_0 & \ldots & & a_{n-1} \end{pmatrix}}_{:=A} \cdot \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix}$$

Nosso objetivo aqui será encontrar uma base do espaço de soluções da equação linear de ordem superior a coeficientes constantes. Como vimos acima, para tal basta que encontremos uma matriz fundamental de seu sistema equivalente. De fato, basta que encontremos a primeira linha de uma matriz fundamental do sistema equivalente, já que as demais linhas serão simplesmente as derivadas de ordem superior (até ordem $n-1$) desta primeira linha.

Sabemos (da proposição 6.1.8) que qualquer matriz fundamental do sistema em questão é da forma $e^{tA} \cdot P$, com P invertível. Por outro lado, se J é a forma de Jordan (real) de A, então existe uma matriz invertível P tal que

$$J = P^{-1} \cdot A \cdot P \Rightarrow e^{tJ} = P^{-1} \cdot e^{tA} \cdot P \Rightarrow P \cdot e^{tJ} = \underbrace{e^{tA} \cdot P}_{\text{matriz fundamental}}$$

Sendo A a transposta de uma matriz companheira, seu polinômio característico será igual ao minimal, sendo dado por:

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

O fato de que os polinômios minimal e característico de A coincidem implica que a parte nilpotente da forma de Jordan de A será máxima. Ou seja, se $A : \mathbb{R}^n \to \mathbb{R}^n$ possui autovalores reais $\lambda_1 \ldots \lambda_r$ e autovalores complexos não reais $a_1 + ib_1, \ldots a_s + ib_s$, então sua forma de Jordan real é do tipo

$$J = \begin{pmatrix} J_1 & 0 & & \ldots & & 0 \\ 0 & \ddots & & & & \vdots \\ & 0 & J_r & & & \\ \vdots & 0 & & \tilde{J}_1 & & 0 \\ & & & & \ddots & \\ & & & & & \tilde{J}_s \end{pmatrix},$$

onde cada $J_k, 1 \leq k \leq r$ é da forma:

$$\begin{pmatrix} \lambda_k & 1 & 0 & \ldots 0 \\ 0 & \lambda_k & \ddots & 0 \ldots \\ 0 & \ldots & \ddots & 1 \\ 0 & \ldots & & \lambda_k \end{pmatrix},$$

e cada $\tilde{J}_l, 1 \leq l \leq s$ é da forma:

$$\begin{pmatrix} a_l & b_l & 1 & 0 & 0 \ldots & 0 \\ -b_l & a_l & 0 & 1 & \ddots & 0 \\ 0 & & & \ddots & 1 & 0 \\ 0 & & & \ddots & 0 & 1 \\ 0 & & & & a_l & b_l \\ 0 & & & & -b_l & a_l \end{pmatrix}.$$

Concluímos que uma matriz fundamental para o sistema dado por $\dot{X} = A \cdot X$ é

$$P \cdot e^{tJ} = \begin{pmatrix} p_{11} & \dots & p_{1m_1} & \dots & p_{1m} \\ p_{21} & \dots & & & p_{2m} \\ \vdots & & & & \vdots \\ & & & & \\ p_{m1} & \dots & & & p_{mm} \end{pmatrix} \cdot \begin{pmatrix} e^{tJ_1} & 0 & & \dots & & 0 \\ 0 & \ddots & & & & \vdots \\ & 0 & e^{tJ_r} & & & \\ \vdots & 0 & & e^{t\tilde{J}_1} & & 0 \\ & & & & \ddots & \\ & & & & & e^{t\tilde{J}_s} \end{pmatrix}.$$

Como já dissemos, para obtermos uma base de soluções da equação linear de ordem superior, basta calcularmos a primeira linha da matriz fundamental acima. E esta é obtida pelo produto da primeira linha de P pelos blocos diagonais de e^{tJ} acima. Note que a primeira linha de P está dividida em várias partes $p_1, \dots p_r, \tilde{p}_1, \dots, \tilde{p}_s$ compostas de componentes contíguas, cada qual atuando em seu respectivo bloco $e^{tJ_k}, k = 1 \dots r$ ou $e^{t\tilde{J}_l}, l = 1 \dots s$, resultando na parte correspondente da primeira linha da matriz fundamental. Assim, podemos tomar sem perda de generalidade um desses produtos, por exemplo, no caso dos blocos correspondentes aos autovalores reais

$$p_k \cdot e^{tJ_k} = (q_1, \dots, q_{m_k}) \cdot \begin{pmatrix} e^{t\lambda_k} & te^{t\lambda_k} & \frac{t^2}{2}e^{t\lambda_k} & \dots & \frac{t^{m_k-1}}{(m_k-1)!}e^{t\lambda_k} \\ 0 & \ddots & te^{t\lambda_k} & \frac{t^2}{2}e^{t\lambda_k} \dots & \\ 0 & \dots & & \ddots & \vdots \\ 0 & \dots & & & te^{t\lambda_k} \\ 0 & \dots & & & e^{t\lambda_k} \end{pmatrix} =$$

$$\begin{pmatrix} q_1 e^{t\lambda_k} & q_1 te^{t\lambda_k} + q_2 e^{t\lambda_k} & \dots & q_1 t^{m_k-1} e^{t\lambda_k} + \dots + q_{m_k} e^{t\lambda_k} \end{pmatrix}.$$

Temos que $q_1 \neq 0$ (caso contrário, teríamos uma coluna nula na matriz fundamental, o que é absurdo). Portanto, podemos dividir a matriz fundamental por q_1 (obtendo uma outra matriz fundamental). Além disso, note que $e^{t\lambda_k}, \dots, t^{m_k-1}e^{t\lambda_k}$ são m_k funções que geram o mesmo espaço que $e^{t\lambda_k}, te^{t\lambda_k} + \frac{q_2}{q_1}e^{t\lambda_k}, \dots t^{m_k-1}e^{t\lambda_k} + \dots + \frac{q_{m_k}}{q_1}e^{t\lambda_k}$, implicando que são linearmente independentes e constituem parte da base do espaço de soluções que procuramos.

Analogamente, o caso dos blocos correspondentes aos autovalores complexos não reais fica:

$$p_l \cdot e^{t\tilde{J}_l} = (\tilde{q}_1, \dots, \tilde{q}_{m_l}) \cdot$$

$$e^{ta_l}\begin{pmatrix} \cos(tb_l) & \sin(tb_l) & t\cos(tb_l) & t\sin(tb_l) & \frac{t^2}{2}\cos(tb_l) & \frac{t^2}{2}\sin(tb_l) & \dots \\ -\sin(tb_l) & \cos(tb_l) & -t\sin(tb_l) & t\cos(tb_l) & -\frac{t^2}{2}\sin(tb_l) & \frac{t^2}{2}\cos(tb_l) & \dots \\ 0 & 0 & & & & & \\ & & & \ddots & & & \\ \vdots & & & & & t\cos(tb_l) & t\sin(tb_l) \\ 0 & & & & & -t\sin(tb_l) & t\cos(tb_l) \\ & & & & & \cos(tb_l) & \sin(tb_l) \\ 0 & & & & & -\sin(tb_l) & \cos(tb_l) \end{pmatrix} =$$

$$\begin{gathered} \big(\tilde{q}_1 e^{ta_l}\cos(tb_l) - \tilde{q}_2 e^{ta_l}\sin(tb_l) \\ \tilde{q}_1 e^{ta_l}\sin(tb_l) + \tilde{q}_2 e^{ta_l}\cos(tb_l) \\ \tilde{q}_1 t e^{ta_l}\cos(tb_l) - \tilde{q}_2 t e^{ta_l}\sin(tb_l) + \tilde{q}_3 e^{ta_l}\cos(tb_l) - \tilde{q}_4 e^{ta_l}\sin(tb_l) \\ \tilde{q}_1 t e^{ta_l}\sin(tb_l) + \tilde{q}_2 t e^{ta_l}\cos(tb_l) + \tilde{q}_3 e^{ta_l}\sin(tb_l) + \tilde{q}_4 e^{ta_l}\cos(tb_l) \quad \dots\big). \end{gathered}$$

Aqui, pode ocorrer que $\tilde{q}_1 = 0$, mas então devemos ter $q_2 \neq 0$ e similarmente ao caso real, o par $e^{ta_l}\cos(tb_l), e^{ta_l}\sin(tb_l)$ gera o mesmo que as duas componentes acima.

Portanto, podemos dividir a matriz fundamental por q_1 (obtendo uma outra matriz fundamental). Além disso, note que $e^{t\lambda_k}, \dots, t^{m_k-1}e^{t\lambda_k}$ são k funções que geram o mesmo espaço que $e^{t\lambda_k}, te^{t\lambda_k} + \frac{q_2}{q_1}e^{t\lambda_k}, \dots t^{m_k-1}e^{t\lambda_k} + \dots + \frac{q_{m_k}}{q_1}e^{t\lambda_k}$, implicando que são linearmente independentes e constituem parte da base do espaço de soluções que procuramos.

6.6 Exercícios

1. Seja $X : U \to \mathbb{R}^m$ um campo completo de classe C^1, associado a um fluxo $\varphi(t,x) = \varphi_t(x)$. Prove que para todo aberto limitado B, $v(t) :=$ volume$[\varphi_t(B)]$ satisfaz a equação

$$\frac{dv}{dt}(t) = \int_{\varphi_t(B)} divX.$$

Em particular, conclua que se $divX$ não muda de sinal, então φ_t diminui, preserva ou expande o volume, conforme respectivamente $divX < 0$, $divX \equiv 0$ ou $divX > 0$. Lembramos que $divX$ é definido como o traço de DX.

Capítulo 7

Noções Básicas de Teoria Espectral

Neste capítulo, em continuação ao que já fizemos no capítulo anterior acerca de operadores lineares em dimensão finita, nos aprofundaremos no estudo de operadores lineares em dimensão qualquer. Aplicaremos este estudo a dois alvos. O primeiro, a caracterização espectral dos chamados isomorfismos lineares hiperbólicos, que são operadores que aparecem no enunciado do Teorema de Grobman-Hartman (no próximo capítulo) e em outros importantes teoremas da área de Sistemas Dinâmicos. A segunda aplicação deste estudo é a caracterização do conjunto das soluções de problemas de Contorno lineares.

Demo-nos o trabalho de ser minuciosos, especialmente nos conceitos de adjunto de um operador. A razão disso é que a maior parte dos livros de Análise Funcional não se prolonga em maiores comentários acerca da razão pela qual tal operador adjunto está bem definido. A falta de familiaridade com o conceito de operador adjunto torna-se ainda mais crítica nestes textos quando se trabalha posteriormente com o conceito de adjunto para operadores descontínuos. Ademais, os textos clássicos de Análise Funcional nem sempre se aplicam diretamente a EDO. Isto porque lá os operadores lineares (pelo menos, os contínuos) atuam invariavelmente em espaços completos (de Banach ou de Hilbert). Em EDO, nem sempre temos essa escolha: o espaço que nos é dado para buscarmos a solução de uma equação nem sempre é completo, sendo necessário adaptar alguns aspectos da teoria de Análise Funcional.

Ainda assim, estruturamos este capítulo e os posteriores de modo a que

se possa fazer uma leitura *quase* independente dos próximos capítulos em relação a este. Desse modo, algumas definições poderão aparecer em redundância nos capítulos ulteriores.

Como é de praxe, comecemos com uma definição:

Definição 7.0.1. (Espectro de um operador linear). Seja E um espaço vetorial complexo e seja $A : E \to E$ um operador linear. O espectro de A é o conjunto

$$sp(A) := \{\lambda \in \mathbb{C}, (\lambda I - A) \text{ não é invertível.}\}$$

Note que a definição acima não faz menção à norma de E, muito menos a qualquer métrica no espaço de operadores lineares de E em E, já que o fato de uma aplicação ser invertível (injetiva e sobre) ou não, depende apenas de seu gráfico e dos conjuntos de domínio e contradomínio, e não de qualquer métrica nos mesmos.

Estudaremos nessa seção uma caracterização do espectro de $A : E \to E$ quando $A \in \mathcal{L}(E)$, com E um espaço de Banach complexo e $\mathcal{L}(E)$ sendo o espaço de aplicações lineares contínuas de E em E. Primeiramente, estamos interessados em obter cotas para a norma de qualquer elemento em $sp(A)$.

A idéia para isso será estudarmos $res(A) := sp(A)^c$, também conhecido como o conjunto resolvente de A. Ora, para $z \in res(A)$, sabemos que é um isomorfismo linear (contínuo) o operador $(zI - A)$. Lembramos que se E é um espaço de Banach, $\mathcal{L}(E)$ também é um espaço de Banach com a conhecida norma do Operador. Para $T \in \mathcal{L}(E)$, sua norma é:

$$\|T\|_{op} = \sup_{v \in E; \|v\|=1} \{\|T(v)\|\} = \mathrm{Lip}(T).$$

Consideraremos então a *aplicação resolvente* $\rho : res(A) \to \mathcal{L}(E)$ dada por $\rho(z) := (zI - A)^{-1}$. Mostraremos que esta aplicação é analítica, e adaptaremos o que conhecemos sobre raio de convergência de série de potências para obter a cota desejada para $sp(A)$.

Antes de tudo, observemos que se A é contínuo, o conjunto resolvente de A é não vazio, e que o espectro é limitado. Se $|z| > \|A\|_{op}$, então $(zI - A) = z(I - A/z)$ é invertível, se e só se, $(I - A/z)$ é invertível. Mas $(I - A/z)$ é invertível pelo Teorema da Perturbação da Identidade, já que

$$Lip(-A/z) = \| - A/z\|_{op} = \|A\|_{op}/|z| < 1.$$

Isto implica que $sp(A) \subset \overline{B(0, \|A\|_o p)}$ e que $res(A) \neq \emptyset$.

Portanto, faz sentido considerarmos a aplicação resolvente $\rho : res(A) \to \mathcal{L}(E)$ definida acima. Também do que vimos acima, está claro que para $z \neq 0$ para que $(zI - A)$ seja invertível é necessário e suficiente que $(I - A/z)$ seja invertível. Inspirados na série geométrica, para $z, |z| > 0$, estudemos a convergência absoluta da série $\sum_{n\geq 0}(A/z)^n$, a qual esperamos que convirja a $(I - A/z)^{-1}$. Ora, tal série converge absolutamente se, e só se, a série $\sum_{n\geq 0} \|A^n\|_{op}/|z|^n$ converge na reta. Chamando de $a_n := \|A^n\|_{op}$ do critério de comparação (com a série geométrica) que esta última série converge para z tal que

$$\limsup \sqrt[n]{a_n}/|z| < 1 \Rightarrow |z| > \limsup \sqrt[n]{a_n} = \limsup \sqrt[n]{\|A^n\|}.$$

Notamos que a composição de aplicações lineares com A é contínua em $\mathcal{L}(E)$. Por exemplo, para a composição com A á esquerda, temos:

$$\|A \circ B - A \circ C\|_{op} = \|A \circ (B - C)\|_{op} \leq \|A\|_{op} \cdot \|B - C\|_{op}, \forall B, C \in \mathcal{L}(E),$$

mostrando que tal aplicação de composição é Lipschitz. Temos assim da continuidade da composição que para $|z| > \limsup \sqrt[n]{\|A^n\|}$, vale

$$(I - A/z)(\lim_{n\to\infty} \sum_{j=0}^{n} (A/z)^n = \lim_{n\to\infty} (I - A/z) \sum_{j=0}^{n} (A/z)^n =$$

(os limites acima são na norma do operador em $\mathcal{L}$)

$$\lim_{n\to\infty} \sum_{j=0}^{n} (A/z)^n - \sum_{j=1}^{n+1} (A/z)^n = \lim_{n\to\infty} I - (A/z)^n = I.$$

Efetuando contas similares, só que com a composição á direita com $(I - A/z)$, concluímos que para $|z| > \limsup \sqrt[n]{\|A^n\|}$, existe $(zI - A)^{-1} = z \cdot \sum_{n=0}^{\infty} (A/z)^n$. Isso nos dá uma cota mais fina para o raio da bola fechada onde se encontra $sp(A)$.

Para mostrarmos que $\sup\{|x|; x \in sp(A)\} = \limsup \sqrt[n]{\|A^n\|}$, basta que adaptemos a teoria de Funções Holomorfas de $\mathbb{C}$ em $\mathbb{C}$, para curvas holomorfas em espaços de Banach. Mais particularmente, basta que adaptemos a teoria de Cauchy-Goursat para este contexto.

Definição 7.0.2. (Aplicação Holomorfa). Seja $U \subset \mathbb{C}$ um conjunto aberto e $f : U \to E$, onde E é um espaço de Banach. Dizemos que f é *holomorfa*

em $z_0 \in U$ se existe o limite

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} = f'(z_0).$$

Neste caso, $f'(z_0)$ é chamada de *derivada holomorfa de* f *em* z_0. Se f é holomorfa em cada ponto de U, dizemos que f é *holomorfa em* U ou, simplesmente, que f é *holomorfa*.

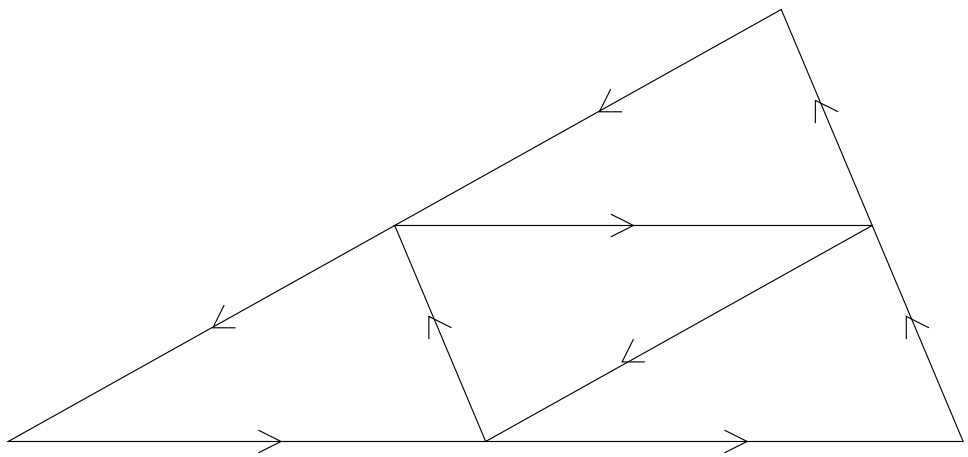
Lembramos aqui a prova do Teorema de Cauchy-Goursat para regiões triangulares, adaptando-o ao contexto de espaços de Banach.

Teorema 7.0.3. *(Teorema de Cauchy-Goursat para regiões triangulares). Sejam* $U \subset \mathbb{C}$ *um aberto,* E *um espaço de Banach,* $f : U \to E$ *uma aplicação holomorfa e seja* Δ *um triângulo compacto contido (inclusive o seu interior) em* U*. Então*

$$\int_\Delta f(z)dz = 0.$$

Prova: Realizemos uma construção indutiva para a prova do teorema. Escrevamos $\Delta = \Delta_0$ e subdividamos este triângulo em quatro triângulos $(\Delta_0^1, \Delta_0^2, \Delta_0^3, \Delta_0^4)$ a ele semelhantes, cujos lados têm metade do comprimento de seus correspondentes no triângulo original. Ademais, orientamos os bordos de cada um dos triângulos no sentido horário. Daí,

$$\int_{\Delta_0} f(z)dz = \int_{\Delta_0^1} f(z)dz + \int_{\Delta_0^2} f(z)dz + \int_{\Delta_0^3} f(z)dz + \int_{\Delta_0^4} f(z)dz.$$



O triangulo Δ subdividido em quatro triangulos semelhantes, com metade do lado e 1/4 de sua area.

Vejamos como se dá o passo de indução: supondo que temos construído um triângulo Δ_n para um certo $n \in \mathbb{N}$ (por exemplo, já definimos, para $n = 0$, $\Delta_0 := \Delta$). Daí, dividimos Δ_n em 4 triângulos $\Delta_n^1, \Delta_n^2, \Delta_n^3, \Delta_n^4$ semelhantes como explicado acima. Definimos $\Delta_{n+1} := \Delta_n^j$, onde

$$|\int_{\Delta_n^j} f(z)dz| = \max\{|\int_{\Delta_n^1} f(z)dz|, |\int_{\Delta_n^2} f(z)dz|, |\int_{\Delta_n^3} f(z)dz|, |\int_{\Delta_n^4} f(z)dz|\}$$

Daí,

$$|\int_{\Delta_n} f(z)dz| \leq 4 \cdot |\int_{\Delta_{n+1}} f(z)dz|$$

Ademais, se δ_n é o comprimento do maior lado do triângulo Δ_n, é claro que

$$\delta_{n+1} = \delta_n/2 = \delta_0/(2^n),$$

$$\ell(\Delta_{n+1}) = \ell(\Delta_n)/2 = \ell(\Delta_0)/(2^n).$$

Como os triângulos $\Delta_n, n \in \mathbb{N}$ formam uma família encaixante de compactos não vazios, podemos tomar $z_0 \in \cap_{n\in\mathbb{N}}\Delta_n$. Como f é holomorfa, dado $\epsilon > 0$, $\exists \tau > 0$ tal que

$$|z - z_0| < \tau \Rightarrow |f(z) - f(z_0) - f'(z_0) * (z - z_0)| \leq \frac{\epsilon}{\delta_0 \cdot \ell(\Delta)} \cdot |z - z_0|.$$

Daí,

$$\int_{\Delta_n} f(z)-f(z_0)-f'(z_0)*(z-z_0)dz = \int_{\Delta_n} f(z)dz-(\int_{\Delta_n} f(z_0)+f'(z_0)*(z-z_0)dz) =$$

(pois o Teorema de Cauchy-Goursat claramente vale para aplicações holomorfas afins)

$$\int_{\Delta_n} f(z)dz.$$

Por conseguinte,

$$|\int_{\Delta} f(z)dz \leq 4^n|\int_{\Delta_n} f(z)dz| = 4^n|\int_{\Delta_n} f(z)-f(z_0)-f'(z_0)*(z-z_0)dz| \leq$$

(Supondo n suficientemente grande de modo a que $\delta_n < \tau$)

$$4^n \cdot \frac{\epsilon}{\delta_0 \cdot \ell(\Delta)} \cdot \sup\{|z-z_0|\} \cdot \ell(\Delta_n) \leq 4^n \cdot \frac{\epsilon}{\delta_0 \cdot \ell(\Delta)} \cdot \frac{\delta_0}{2^n} \cdot \frac{\ell(\Delta)}{2^n} \leq \epsilon.$$

Como $\epsilon > 0$ é arbitrário, segue-se que

$$\int_{\Delta} f(z)dz = 0.$$

□

A partir da versão acima, é bastante fácil de provar uma versão similar para círculos (e curvas convexas) no lugar de triângulo.

Usando da definição de integral curvilínea complexa, sabemos que $\int_\gamma \frac{1}{z-z_0}dz = 2\pi i$, para qualquer curva fechada de simples γ contendo z_0 na região aberta limitada de $\mathbb{C}$ que possui γ como fronteira. O resultado mais importante na teoria de aplicação analíticas é o seguinte:

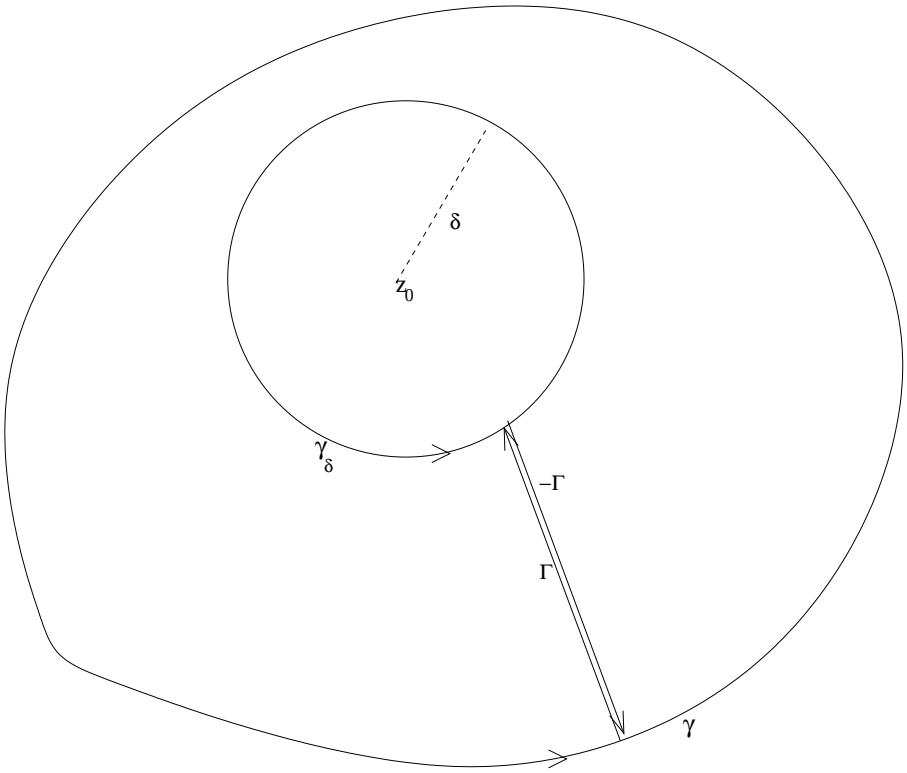
Teorema 7.0.4. *(Fórmula Integral de Cauchy). Seja E um espaço de Banach sobre $\mathbb{C}$, $U \subset \mathbb{C}$ um aberto simplesmente conexo e $f : U \to E$ uma aplicação holomorfa. Seja $\gamma_0 \subset U$ uma região compacta cuja fronteira é uma curva de Jordan γ. Então, dado $z_0 \in int(\gamma_0)$, vale:*

$$f(z_0) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z-z_0}dz.$$

Prova: Dado $\epsilon > 0$, seja $\delta > 0$ da continuidade uniforme de f em γ_0 tal que

$$\|z - z_0\| \leq \delta \Rightarrow \|f(z) - f(z_0)\| < \frac{\epsilon}{2\pi}.$$

Obviamente, podemos supor $\delta > 0$ suficientemente pequeno de modo a que $\overline{B(0,\delta)} \subset int(\gamma_0)$. Chamemos de γ_δ a curva qu é o círculo de centro z_0 e raio δ.



Justapondo as curvas $\Gamma, -\gamma_\delta, -\Gamma, \gamma$, e aplicando o Teorema de Cauchy−Goursat, obtemos que a integral no circulo de raio delta em torno de z_0 e´ zero.

Ligando γ a γ_δ por meio de uma curva auxiliar Γ difeomorfa a um intervalo compacto, conforme mostra a figura, usando da propriedade de que uma integral de linha muda de sinal se trocamos a orientação e aplicando o teorema de Cauchy-Goursat, obtemos que

$$\int_\gamma \frac{f(z)}{z - z_0} dz = \int_{\gamma_\delta} \frac{f(z)}{z - z_0} dz.$$

Mas

$$\| \int_{\gamma_\delta} \frac{f(z)}{z-z_0} dz - f(z_0) 2\pi i \| = \| \int_{\gamma_\delta} \frac{f(z)}{z-z_0} dz - f(z_0) \int_{\gamma_\delta} \frac{1}{z-z_0} dz \| =$$

$$\int_{\gamma_\delta} \frac{f(z)-f(z_0)}{z-z_0} dz.$$

Como para z sobre a curva γ_δ, temos $\|f(z)-f(z_0)\| < \epsilon/2\pi$ e $\|z-z_0\| = \delta$, obtemos

$$\| \int_{\gamma_\delta} \frac{f(z)-f(z_0)}{z-z_0} dz \| < \frac{\epsilon}{2\pi\delta} \cdot \ell(\gamma_\delta) = \epsilon.$$

Concluímos que

$$\| \int_{\gamma} \frac{f(z)}{z-z_0} dz - 2\pi i f(z_0) \| = \| \int_{\gamma_\delta} \frac{f(z)-f(z_0)}{z-z_0} dz \| < \epsilon, \forall \epsilon > 0,$$

logo

$$\int_{\gamma} \frac{f(z)}{z-z_0} = 2\pi i f(z_0).$$

□

Dizemos que $\mathcal{N} \subset \mathbb{C}$ é um anel centrado em $a \in \mathbb{C}$, se $\mathcal{N}$ é da forma

$$\mathcal{N} = \mathcal{N}(a, r_1, r_2) := \{z \in \mathbb{C}, r_1 \leq |z-a| \leq r_2, \text{ com } r_1, r_2 > 0, a \in \mathbb{C}\}.$$

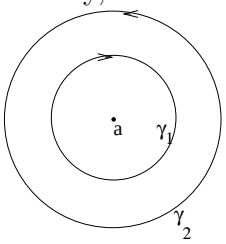
Isto nos permite demonstrar o seguinte teorema sobre aplicações Holomorfas em um anel:

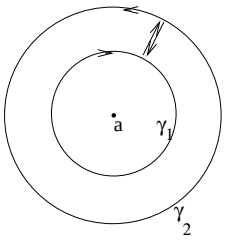
Teorema 7.0.5. *(Séries de Laurent em Espaços de Banach). Sejam $\mathcal{N} \subset \mathbb{C}$ um anel centrado em $a \in \mathbb{C}$, $V \subset \mathbb{C}$ uma vizinhança de $\mathcal{N}$, e $f : V \to E$ uma aplicação holomorfa tomando valores em um espaço de Banach E. Então existem únicos $A_n \in E, n \in \mathbb{Z}$ tais que*

$$f(z) = \sum_{n=-\infty}^{+\infty} A_n (z-a)^n, \forall z \in \mathcal{N},$$

a convergência do limite acima sendo absoluta e uniforme em $\mathcal{N}$.

Sendo $\mathcal{N}$ um anel centrado em $a \in \mathbb{C}$ e $f : V \to E$, e orientando a fronteira de $\mathcal{N}$ conforme a figura, dado $z \in \mathcal{N} \setminus \partial\mathcal{N}$, pela fórmula integral de Cauchy, temos:





$$f(z) = \frac{1}{2\pi i}\int_{\partial\mathcal{N}} \frac{f(w)}{w-z}dw = \frac{1}{2\pi i}(\int_{\gamma_2} \frac{f(w)}{w-z}dw - \int_{\gamma_1} \frac{f(w)}{w-z}dw) =$$

$$\frac{1}{2\pi i}(\int_{\gamma_2} \frac{f(w)}{w-a-(z-a)}dw + \int_{\gamma_1} \frac{f(w)}{z-a-(w-a)}dw) =$$

$$\frac{1}{2\pi i}(\int_{\gamma_2} \frac{f(w)}{(w-a)*(1-\frac{z-a}{w-a})}dw + \int_{\gamma_1} \frac{f(w)}{(z-a)*(1-\frac{w-a}{z-a})}dw) =$$

(note que para $w \in \gamma_2$ vale $|w-a| > |z-a|$, $\forall z \in \mathcal{N}$; já para $w \in \gamma_1$ vale $|w-a| < |z-a|$)

$$\frac{1}{2\pi i}\Big(\int_{\gamma_2} \frac{f(w)}{(w-a)} * \sum_{j=0}^{\infty}(\frac{z-a}{w-a})^j dw + \int_{\gamma_1} \frac{f(w)}{z-a} * \sum_{j=0}^{\infty}(\frac{w-a}{z-a})^j dw\Big).$$

As somas geométricas dentro das integrais convergem absolutamente e uniformemente em partes compactas de $int(\mathcal{N})$, logo podemos permutar seus limites com as integrais, e usando a linearidade das integrais, obtemos:

$$f(z) = \frac{1}{2\pi i}\sum_{j=0}^{\infty}\int_{\gamma_2} \frac{f(w)}{(w-a)^{j+1}}dw(z-a)^j + \sum_{j=1}^{\infty}\int_{\gamma_1} f(w)(w-a)^{j-1}dw*(z-a)^{-j}dw,$$

também chamada de *Série de Laurent* de f no anel $\mathcal{N}$.

Para vermos a unicidade dos coeficientes de Laurent, basta notarmos que se $f(z) = \sum_{n=-\infty}^{+\infty} A_n(z-a)^n$, então dado $k \in \mathbb{Z}$, e para qualquer círculo ga com centro em a e contido em $\mathcal{N}$, temos

$$\frac{1}{2\pi i}\int_{\gamma} f(z)\cdot(z-a)^{k+1}dz = \frac{1}{2\pi i}\int_{\gamma}\sum_{n=-\infty}^{+\infty} A_n(z-a)^{n+k+1}dz = A_k,$$

uma vez que $\int_\gamma (z-a)^{n+k+1}dz = 0$, se $n+k+1 \neq -1$, e é igual a $2\pi i$, se $n+k+1 = -1$.

□

Voltemos agora à aplicação resolvente $\rho : res(A) \subset \mathbb{C} \to \mathcal{L}(E)$. É muito fácil ver que $res(A)$ é aberto, e que ela é analítica (holomorfa) em $res(A)$. Realmente, pelo Teorema da Perturbação do Isomorfismo (veja corolário 0.2.16, na página 10), se $\lambda \in res(A)$ e μ tal que $|\mu| < \|\rho(\lambda)\|^{-1}$, então $(\lambda I - A) + \mu I$ é invertível. Isso implica que $B(\lambda, \|\rho(\lambda)\|^{-1}) \subset res(A)$ e, por conseguinte, que $res(A) \subset \mathbb{C}$ é aberto. Ademais,

$$((\mu+\lambda)I - A)^{-1} = (\mu I + \lambda I - A)^{-1} \cdot ((\lambda I - A)^{-1})^{-1} \cdot (\lambda I - A)^{-1} =$$

$$((\lambda I - A)^{-1} \cdot (\mu I + \lambda I + A))^{-1} \cdot (\lambda I - A)^{-1} = (\mu I \cdot (\lambda I - A)^{-1} + I)^{-1} \cdot (\lambda I - A)^{-1}.$$

Por analogia com a série geométrica vemos que a inversa de $(\lambda+\mu)I - A$ deve ser dada formalmente pela série:

$$S(\mu) = \sum_{n=0}^{\infty} (-1)^n ((\lambda I - A)^{-1})^{n+1} \mu^n = \sum_{n=0}^{\infty} (-1)^n \rho(\lambda)^{n+1} \mu^n.$$

Por tomarmos $\|\mu\rho(\lambda)\| < 1$, a série acima converge absolutamente, e uniformemente em partes compactas de $B(0, \|\rho(\lambda)\|^{-1})$. Em particular, segue-se (da teoria de séries de potências) que ρ é holomorfa, com derivada holomorfa em λ igual a $-\rho(\lambda)^2$.

Note que a série de Laurent de ρ em torno de zero é

$$\rho(z) = \sum_{j=0}^{+\infty} (A/z)^n.$$

Concluímos então a partir do teorema 7.0.5 que tal série converge para todo $z \in \mathbb{C}$ tal que $|z| > \sup sp(A)$ e, é claro, não converge para $|z| < \sup sp(A)$. Logo, $\sup sp(A) = \limsup \sqrt[n]{\|A^n\|}$. Uma consequência imediata, e bastante importante disso, é que se o espectro de A está contido na bola unitária aberta $B(0,1)$, automaticamente todo iterado suficientemente grande de A será uma contração.

Veremos na próxima seção algo mais: que se o espectro de A não intersecta S^1, então o espaçøE admite uma decomposição em soma direta $E = E^s \oplus E^u$ tal que $A(E^s) \subset E^s$, $A(E^u) = E^u$, sendo $A|_{E^s}$ e $[A|_{E^u}]^{-1}$ contrações.

7.1 Funções de um Operador

Na seção anterior, adaptamos a Teoria clássica de Análise Complexa com a finalidade de estudar a aplicação resolvente ρ de um operador linear $A : E \to E$ fixado, onde E é um espaço de Banach. Usamos o fato de que ρ é uma aplicação holomorfa de um aberto de $\mathbb{C}$ em $\mathcal{L}(E)$. A idéia desta nova seção é estudar o espectro sob um foco diferente, cuja motivação é a seguinte. Dado um polinômio $p(z) = \sum_{n=0}^{m} c_n z^n$, com $c_n \in \mathbb{C}, \forall n \in \{0, \ldots, m\}$, podemos avaliá-lo em $\mathcal{L}(E)$ (no lugar de avaliá-lo em $\mathbb{C}$) pela fórmula:

$$\mathcal{L}(E) \ni p(A) = \sum_{n=0}^{m} c_n A^n.$$

Dizemos que $p(A)$ é uma função polinomial do operador A. Como as funções holomorfas são localmente limite uniforme de polinomiais, claro está que dada uma função $f : U \subset \mathbb{C} \to \mathbb{C}$ deve ser possível estender o conceito de função de operador para funções analíticas quaisquer, obtendo-se $f(A)$.

Uma questão natural é saber que relação existe entre o espectro de $f(A)$ e o espectro de A (veremos mais adiante que essa relação é: $sp(f(A)) = f(sp(A))$). A definição precisa de $f(A)$, das relações entre seu espectro e o espectro de A e suas consequências são o objetivo da presente seção.

Definição 7.1.1. (Função de operador). Seja $A \in \mathcal{L}(E)$ um operador linear em um espaço de Banach E e $f : U \to \mathbb{C}$ uma função holomorfa definida uma vizinhança U não necessariamente conexa de $sp(A)$. Suponha que $\partial U = C$ é composta de curvas fechadas, C^1 por partes, orientadas com a orientação induzida no bordo. Definimos a função do operador A dada por f como

$$f(A) := \frac{1}{2\pi} \int_C f(\lambda)\rho(\lambda)d\lambda$$

Observação 7.1.2. Denotamos por $\mathfrak{F}(A)$ à coleção de todas as funções holomorfas em alguma vizinhança com fronteira C^1 por partes de $sp(A)$. Note que se duas dessas funções coincidem em uma tal vizinhança, automaticamente (mesmo que possuam domínios diferentes) a mesma função de operador. De fato, suponha que $f : U \to \mathbb{C}$, $\hat{f} : \hat{U} \to \mathbb{C}$ pertençam a $\mathfrak{F}(A)$.

Teorema 7.1.3. *(Cálculo Funcional). Dadas $f, g \in \mathfrak{F}(A)$, $c \in \mathbb{C}$, valem:*

1. $c \cdot f + g \in \mathfrak{F}(A)$ *e* $(c \cdot f + g)(A) = c \cdot f(A) + g(A)$.

2. *$f \cdot g \in \mathfrak{F}(A)$ e $(f \cdot g)(A) = f(A) \cdot g(A)$.*

3. *Se f possui expansão em série de Taylor $f(\lambda) = \sum_{k=0}^{\infty} a_n \lambda^n$, absolutamente convergente em uma vizinhança de $sp(A)$, então $f(A) = \sum_{n=0}^{\infty} a_n A^n$.*

Prova:

Para o item 1, devemos esclarecer que por $h = c \cdot f + g$ entendemos a função obtida somando-se na intersecção dos domínios de f e g. consequência óbvia da linearidade da integral.

Para mostrarmos o item 2, observamos que vale a seguinte identidade, também conhecida com equação do resolvente:

$$\rho(\lambda) - \rho(\mu) = (\mu - \lambda)\rho(\lambda)\rho(\mu)$$

De fato,

$$\rho(\lambda) - \rho(\mu) = \rho(\lambda)\rho(\mu)(\mu I - A)(\lambda I - A)(\rho(\lambda) - \rho(\mu)) =$$

$$\rho(\lambda)\rho(\mu)(\mu I - A)(I - \lambda\rho(\mu) + A\rho(\mu)) =$$

$$\rho(\lambda)\rho(\mu)(\mu I - A - \lambda I + A) =$$

$$(\mu - \lambda)\rho(\lambda)\rho(\mu).$$

$$f(A) \cdot g(A) = -\frac{1}{4\pi^2} \int_{C_1} f(\lambda)\rho(\lambda)d\lambda \int_{C_2} g(\mu)\rho(\mu)d\mu =$$

$$-\frac{1}{4\pi^2} \int_{C_1} (\int_{C_2} f(\lambda)g(\mu)\rho(\lambda)\rho(\mu)d\mu)d\lambda =$$

(aplicando a equação de resolvente e lembrando que C_2 e C_1 são disjuntos)

$$-\frac{1}{4\pi^2} \int_{C_1} (\int_{C_2} f(\lambda)g(\mu)\frac{\rho(\lambda) - \rho(\mu)}{\mu - \lambda}d\mu)d\lambda =$$

$$-\frac{1}{4\pi^2} \int_{C_1} f(\lambda)(\int_{C_2} \frac{g(\mu)}{\mu - \lambda}d\mu)\rho(\lambda)d\lambda + \frac{1}{4\pi^2} \int_{C_2} g(\mu)(\int_{C_1} \frac{f(\lambda)}{\mu - \lambda}d\lambda)\rho(\mu)d\mu =$$

(pois tomamos C_2 exterior a C_1)

$$\frac{1}{2\pi i} \int_{C_1} f(\lambda)g(\lambda)\rho(\lambda)d\lambda = (f \cdot g)(A).$$

Quanto ao item 3, sabemos do curso elementar de Análise Complexa que qualquer série de potências converge absolutamente em bolas abertas em torno de um centro, logo, se a série $\sum a_n\lambda^n$ converge em uma vizinhança de $sp(A)$, estão existe ϵ_0 tal que existe o limite (uniforme) $\sum_{n=0}^{\infty} a_n\lambda_n, \forall\lambda \leq \sup sp(A) + \epsilon_0 = r$. Em particular, denotando por S_r^1 a esfera unitária de centro 0 e raio r, obtemos:

$$f(A) = \frac{1}{2\pi i}\int_{S_r^1}(\sum_{n=0}^{\infty} a_n\lambda_n)\rho(\lambda)d\lambda = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\int_{S_r^1} a_n\lambda_n\rho(\lambda)d\lambda =$$

$$\frac{1}{2\pi i}\sum_{n=0}^{\infty} a_n\int_{S_r^1}\lambda_n(\sum_{j=0}^{\infty}\frac{T^j}{\lambda^{j+1}})d\lambda = \sum_{n=0}^{\infty} a_nT^n.$$

□

O próximo teorema (junto com o anterior) pode ser considerado o proto-teorema Espectral, isto é, uma versão não lapidada (e portanto, mais geral) do teorema Espectral.

Teorema 7.1.4. *(Mapeamento espectral). Se $f \in \mathfrak{F}(A)$, então $sp(f(A)) = f(sp(A))$. Em particular, se A é invertível, então $sp(A^{-1}) = (sp(A))^{-1} := \{\mu^{-1}, \mu \in sp(A)\}$.*

Prova:
($\Rightarrow$) Seja $\lambda \in sp(A)$. A idéia é tentar escrever

$$f(\lambda)I - f(A) = (\lambda I - A)\cdot g(A),$$

com $g \in \mathfrak{F}(A)$. Daí, como os operadores de A comutam, fica claro que se $f(\lambda)$ não estivesse em $sp(f(A))$, então $g(A)\cdot(f(\lambda) - f(A))^{-1}$ seria inversa de $(\lambda I - A)$, absurdo. A própria fórmula acima nos indica como definir g em uma vizinhança de $sp(A)$:

$$g(z) = \begin{cases} \frac{f(\lambda)-f(z)}{\lambda - z}, \text{ se } z \neq \lambda \\ f'(\lambda), \text{ caso } z = \lambda. \end{cases}$$

Como g é holomorfa em um disco furado com centro em λ e é contínua em λ (pois f é holomorfa em λ), segue-se que g é holomorfa inclusive em λ, possuindo assim o mesmo domínio que f. Do Teorema do Cálculo Funcional, segue-se que $g(A)$ satisfaz a fórmula desejada relacionando $\lambda I - A$ e $f(\lambda)I - f(A)$.

($\Leftarrow$) Agora seja $\mu \in sp(f(A))$ e suponha por absurdo que $\mu \notin f({}^{(}A))$. Neste caso, $f(\lambda) - \mu \neq 0, \forall \lambda \in sp(A)$ e portanto $h(z) = (f(z) - \mu)^{-1}$ está definida (e é holomorfa) em uma vizinhança de $sp(A)$. Ora, do Teorema do Cálculo Funcional, segue-se que

$$h(A) \cdot (f(A) - \mu I) = I,$$

o que implica que $\mu \notin sp(f(A))$, absurdo.

Se A é invertível, então $0 \notin sp(A)$, logo $f(z) = 1/z$ é uma função holomorfa definida na vizinhança $\mathbb{C} \setminus \{0\}$ de $sp(A)$. Ora, do teorema do Cálculo Funcional, de $f(z) \cdot z = z \cdot f(z) = 1$, concluímos que $f(A) \cdot A = A \cdot f(A) = I$, ou seja, que $f(A) = A^{-1}$. Da parte provada acima do Mapeamento Espectral, concluímos que $sp(A^{-1}) = sp(f(A)) = f(sp(A)) = (sp(A))^{-1}$.

□

Definição 7.1.5. (Conjunto espectral). Seja $A : E \to E$ um operador linear definido em um espaço de Banach E. Um conjunto $X \subset sp(A)$ é dito ser um *conjunto espectral* se ele é aberto e fechado em sp(A).

Note que como $sp(A)$ é compacto, todo conjunto espectral também o é. Note ainda que se X é um conjunto espectral, o mesmo vale para X^c (=complementar de X em $sp(A)$).

Definição 7.1.6. (Projeção espectral) Seja X um conjunto espectral do espectro de um operador linear A. Seja $P_X : V \to \mathbb{C}$ definida em uma vizinhança não conexa $V = V_X \cap V_{X^c}$ de $sp(A)$, tal que $V_X \supset X$, tal que

$$P_X(z) = 1, \forall z \in V_X; P_X(z) = 0, \forall z \in V_{X^c}.$$

Então a aplicação $\Pi_X := P_X(A) \in \mathcal{L}(E)$ é dita *projeção espectral* associada a X.

Teorema 7.1.7. *Seja $A \in \mathcal{L}(E)$ um operador linear em um espaço de Banach, e seja $X \subset sp(A)$ um conjunto espectral. Então existe uma decomposição $A-$invariante $\hat{E} \oplus \tilde{E} = E$ tal que $sp(A|_{\hat{E}}) = X$ e $sp(A|_{\tilde{E}}) = X^c$.*

Prova: Pelo teorema do Cálculo Funcional, vale que Π_X e Π_{X^c} comutam com A e entre si (todos os operadores de A comutam entre si), que $I = \Pi_X + \Pi_{X^c}$, e $0 = \Pi_X \cdot \Pi_{X^c}$ (pois $P_X(z) \cdot P_{X^c}(z) = 0$. Ademais, notamos que $P_X(z) = P_X(z) \cdot P_X(z)$ (resp. $P_{X^c}(z) = P_{X^c}(z) \cdot P_{X^c}(z)$) vale que $\Pi_X = \Pi_X \cdot \Pi_X$ (resp. $\Pi_{X^c} = \Pi_{X^c} \cdot \Pi_{X^c}$.

Em particular, vale ainda que $A = \Pi_X \cdot A + \Pi_{X^c} \cdot A$. Definindo $\hat{E} := \Pi_X(E)$ e $\tilde{E} := \Pi_{X^c}(E)$, temos que $\hat{E} + \tilde{E} = I(E) = E$ e se $v \in \hat{E} \cap \tilde{E}$, então

$$\Pi_X(v) = v = \Pi_{X^c}(v) \Rightarrow \Pi_X \cdot \Pi_{X^c}(v) = v \Rightarrow v = 0,$$

o que implica que $\hat{E}$ e $\tilde{E}$ estão em soma direta.

Finalmente, da comutatividade existente entre A e as projeções espectrais, concluímos abaixo a $A-$invariância dos espaços $\hat{E}$ e $\tilde{E}$:

$$A(\hat{E}) = A(\Pi_X(E)) = \Pi_X(A(E)) \subset \Pi_X(E) = \hat{E};$$

$$A(\tilde{E}) = A(\Pi_{X^c}(E)) = \Pi_{X^c}(A(E)) \subset \Pi_{X^c}(E) = \tilde{E}.$$

Agora, mostremos que $sp(A|_{\hat{E}}) = X$ e que $sp(A|_{\tilde{E}}) = X^c$.

Primeiramente, observe que como $\hat{E}$ e $\tilde{E}$ são invariantes por A, também o são por $A - \lambda I$. Desse modo,

$$A - \lambda I \text{ é invertível} \Leftrightarrow$$

$$(A - \lambda I)|_{\hat{E}} \text{ é invertível e } (A - \lambda I)|_{\tilde{E}} \text{ é invertível.}$$

Em outras palavras, $res(A) = res(A|_{\hat{E}}) \cap res(A|_{\tilde{E}})$, o que equivale a dizer que

$$sp(A) = sp(A|_{\hat{E}}) \cup sp(A|_{\tilde{E}}).$$

Seja $r \notin sp(A)$, e defina $g : V_X \cup V_{X^c} \to \mathbb{C}$ por $g(z) = P_X(z) * z + r * P_{X^c}$. Isso implica que $g(A) = \Pi_X \cdot A + r\Pi_{X^c}$. Ou seja, $g(A) = (A|_{\hat{E}}, I|_{\tilde{E}})$

Ora, o mapeamento espectral, junto com o mesmo raciocínio acima (baseado na invariância dos espaços $\hat{E}$, $\tilde{E}$) aplicado a g no lugar de A nos dão:

$$X \cup \{r\} = sp(g(A)) = sp(A|_{\hat{E}}) \cup \{r\};$$

e analogamente, poderí amos concluir que

$$X^c \cup \{r\} = sp(A|_{\tilde{E}}) \cup \{r\}.$$

Como r não pertence a $sp(A)$, não pertence a nenhum dos subconjuntos $sp(A|_{hatE}$, $sp(A|_{tildeE}$, X e X^c, donde concluímos que $sp(A|_{\hat{E}}) = X$ e $sp(A|_{\tilde{E}} = X^c$.

$\square$

Definição 7.1.8. (Automorfismo linear hiperbólico). Um operador (ou automorsfismo) linear $A \in \mathcal{L}(E)$ é dito *hiperbólico* se o espectro de A não intersecta a esfera S^1. Se E tem dimensão finita, isso é o mesmo que dizer que nenhum autovalor de A tem norma 1.

Corolário 7.1.9. *Seja E um espaço de Banach (complexo), e $A \in \mathcal{L}(E)$ um automorfismo linear hiperbólico. Então existem $C > 1$, $0 < \lambda < 1$ e uma decomposição $E = E^s \oplus E^u$ tal que*

- *A decomposição é $A-$invariante, isto é, $A(E^s) \subset E^s$ e $A(E^u) \subset E^u$.*
- $\|A^n|_{E^s}\| \leq C\lambda^n$ *e* $\|A^n|_{E^u} \cdot v\| \geq C^{-1}\lambda^{-n}$, $\forall n \in \mathbb{N}$, $\forall v \in E$.

Prova: Dado um automorfismo linear hiperbólico A, vemos que seu espectro se decompõe em $sp(A) = X^s \dot{\cup} X^u$, onde X^s, X^u são os conjuntos espectrais definidos por:

$$X^s := \{\mu \in sp(A); |\mu| < 1\}; X^u := \{\mu \in sp(A); |\mu| > 1\}.$$

Note como X^s e X^u são fechados no compacto $sp(A)$, eles também são compactos. Devido ao fato de A ser hiperbólico, nem X^s, nem X^u intersectam S^1. Portanto, existe $0 < \lambda < 1$ tal que

$$X^s \subset B(0, \lambda) \text{ e } X^u \subset \overline{B(0, \lambda^{-1})}^c.$$

Pelo teorema 7.1.7, existem espaços $A-$invariantes E^s e E^u tais que $sp(A|_{E^s}) = X^s$ e $sp(A|_{E^u}) = X^u$.

Como consequência do final da última seção, temos que o raio espectral de $A|_{E^s} = \limsup \sqrt[n]{\|A^n_{E^s}\|} = \sup\{|\mu|, \mu \in X^s\} < \lambda$. Por conseguinte, existe $n_0 \in \mathbb{N}$ tal que

$$\|A^n_{E^s}\| \leq \lambda^n, \forall n \geq n_0.$$

Tomando $C_s > \max\{1, \|A|_{E^s}{}^j\|/\lambda^j, j = 1, \ldots, n_0 - 1\}$, obtemos que

$$\|A^n|_{E^s}\| \leq C_s\lambda^n.$$

Por outro lado, como $sp(A|_{E^u}) = X^u$, em particular, $0 \notin sp(A|_{E^u})$, e por conseguinte, $A|_{E^u}$ é invertível. Do teorema do mapeamento espectral, temos que $sp([A|_{E^u}]^{-1}) = (sp(A|_{E^u}))^{-1} = (X^u)^{-1}$.

Em particular, temos que o raio espectral de $[A|_{E^u}]^{-1} = \limsup \sqrt[n]{\|A^n_{E^u}\|} = \sup\{|\mu|^{-1}, \mu \in X^u\} < \lambda$, o que, como antes, implica que existe $n_1 \in \mathbb{N}$ tal que

$$\|A^{-n}_{E^u}\| \leq \lambda^n, \forall n \geq n_1.$$

Tomando $C_u > \max\{1, \|A|_{E^u}{}^{-j}\|/\lambda^j, j = 1, \ldots, n_1 - 1\}$, obtemos que

$$\|A^{-n}|_{E^u}\| \leq C_u \lambda^n.$$

Definindo $C := \max\{C_s, C_u\}$, segue-se o resultado.

□

7.2 O Operador Adjunto e seu espectro

Seja E um espaço vetorial normado. O *espaço dual* de E, denotado por E^*, é o espaço vetorial dado por

$$E^* := \{\imath : E \to \mathbb{R}_{\mathbb{C}}; \imath \text{ é funcional linear contínuo.}\}$$

É claro que devido às completudes de $\mathbb{R}$ e $\mathbb{C}$, E^* é sempre um espaço de Banach com a norma do operador:

$$\|\imath\|_{op} := \sup_{x \in E, \|x\|=1} \{|\imath(x)|\}$$

Se $\hat{E}$ é um outro espaço normado, e $A \in \mathcal{L}(E, \hat{E})$, então dado $\jmath \in \hat{E}^*$, podemos definir um funcional linear $A^*(\jmath) \in E^*$ por:

$$A^*(\jmath)(x) = \jmath \circ A(x), \forall x \in E.$$

Note que a aplicação $A^* : \hat{E}^* \to E^*$ dada por $\jmath \mapsto A^*(\jmath)$ é, ela mesma, linear, denominada a *adjunta de* A.

Embora a definição acima seja bastante geral, nos restringiremos nessa seção a estudar operdores definidos em espaços vetoriais normados cuja norma $\|\cdot\|$ provém de um produto interno $< \cdot, \cdot >$, via a fórmula usual $\|v\| = \sqrt{< v, v >}, \forall v \in E$. Lembramos a seguir algumas definições e fatos referentes a tais espaços:

Definição 7.2.1. (Espaço de Hilbert). Um espaço vetorial normado E é dito um *espaço de Hilbert* se sua norma provém de um produto interno e se ele é completo (o que quer dizer que toda sequência de Cauchy na norma de E possui limite em E).

Definição 7.2.2. (Espaço Ortogonal) Seja E um espaço dotado de um produto interno e $\hat{E}$ um subespaço vetorial de E. O espaço ortogonal a $\hat{E}$, denotado por $\hat{E}^{\perp}$ é definido como:

$$\hat{E}^{\perp} := \{v \in E; < \hat{x}, v >= 0, \forall \hat{x} \in \hat{E}\}.$$

Claramente $\hat{E}^{\perp}$ é um subespaço vetorial fechado de E e temos $E = \hat{E} \oplus \hat{E}^{\perp}$.

Definição 7.2.3. (Base Ortonormal) Seja E um espaço vetorial dotado de produto interno. Uma base ortonormal é um conjunto $\beta \subset E$ tal que valem $\|v\| = 1, \forall v \in \beta$, $< v, w >= 0, \forall v, w \in \beta$, com $v \neq w$ e finalmente, dado $x \in E$ existem escalares não nulos $\alpha_1, \ldots, \alpha_n, \ldots$ e $v_1, \ldots, v_n, \cdots \in E$ satisfazendo

$$x = \sum_{j=1}^{\infty} \alpha_j v_j.$$

Observamos que uma base ortonormal não precisa ser enumerável. Por outro lado, dado $x \in E$, e β uma base ortonormal de E, os escalares não nulos α_j, e os vetores v_j que entram na expressão de x da definição acima são (a menos de reenumeração dos pares (α_j, v_j)) unicamente determinados por x e β. Para vermos isso, basta tomarmos o produto interno de x com um elemento arbitrário $v \in \beta$ e usarmos da continuidade do produto interno:

$$< x, v >=< \sum_{j=1}^{\infty} \alpha_j v_j, v >= \sum_{j=1}^{\infty} \alpha_j < v_j, v >= \begin{cases} \alpha_j, \text{ se } v_j = v; \\ 0, \text{caso contrário.} \end{cases}$$

Outra definição útil em espaços dotados de produto interno é a de subespaço ortogonal:

Definição 7.2.4. (Subespaço ortogonal). Seja E um espaço vetorial munido de um produto interno e $\hat{E} \subset E$ um seu subespaço vetorial. O *espaço ortogonal* de $\hat{E}$ é o conjunto:

$$\hat{E}^{\perp} := \{x \in E, < x, \hat{v} >= 0, \forall \hat{v} \in \hat{E}\},$$

o qual claramente é um subespaço vetorial de E.

Exemplo 7.2.5. Seja $E = (C^0([0,1]; \mathbb{R}), < \cdot, \cdot >)$ o espaço das funções contínuas com domínio no intervalo $[0,1]$, dotado do produto interno $<$

$f, g >:= \int_0^1 f(t) \cdot g(t)dt$. Seja $(g_n), g_n \in E$ uma sequência de Cauchy em E sem limite em E. Por exemplo, tome

$$g_n := \begin{cases} 0, \text{ para } t \in [0, 1/2 - 1/(n+1)]; \\ 1/2 + (t - 1/2) \cdot (n+1)/2, \text{ se } t \in (1/2 - 1/(n+1), 1/2 + 1/(n+1)); \\ 1, \text{ para } t \in [1/2 + 1/(n+1), 1]. \end{cases}$$

Daí, defina o funcional linear $\hat{g} : E \to \mathbb{R}$ por:

$$\hat{g}(f) := \lim_{n\to\infty} < f, g_n >, \forall f \in E.$$

É fácil de verificar que $\hat{g}$ é contínuo. De fato, se $f_j \in E, f_j \to 0$, temos:

$$\lim_{j\to\infty} |\hat{g}(f_j)| \leq \lim_{j\to\infty} \lim_{n\to\infty} \|f_j\| \|g_n\| \leq \lim_{j\to\infty} \|f_j\| = 0,$$

sendo a primeira desigualdade devido a Cauchy-Schwarz, e a seguinte porque a sequência (g_n) é limitada (com norma menor do que 1, em nosso caso específico). Considere $\hat{E} = ker(\hat{g})$. Como $\hat{g}$ é contínuo, segue-se que $\hat{E}$ é fechado em E. Note que qualquer função contínua $f : [0, 1] \to \mathbb{R}$ que se anule em $[1/2, 1]$ pertence a $\hat{E}$, o que mostra que esse espaço não é trivial. Por outro lado, $\hat{E} \neq E$, uma vez que qualquer função contínua $f : [0, 1] \to \mathbb{R}$ tal que $f(t) > 0, \forall t \in (1/2, 1)$ não está contida em $\hat{E}$. Contudo, $\hat{E}^\perp = \{0\}$. Tal é demonstrado, em grande generalidade, na próxima proposição.

Proposição 7.2.6. *Seja E um espaço vetorial dotado de um produto interno, $(g_n), g_n \in E$ uma sequência de Cauchy não convergente em E e $\hat{g} : E \to \mathbb{R}$ o funcional linear dado por*

$$\hat{g}(x) = \lim_{n\to\infty} < x, g_n > .$$

Então:

- *$\hat{g}$ é contínuo;*
- *$\hat{E} = \ker(\hat{g})$ é um subespaço fechado (em E) próprio de E;*
- *$\hat{E}^\perp = \{0\}$.*

Prova: A prova dos dois primeiros itens é análoga aos argumentos já vistos no exemplo 7.2.5 acima. Para o último item, procedamos por absurdo.

De fato, se um vetor $w \neq 0$ pertencesse a $\hat{E}^{\perp}$, poderíamos escrever qualquer vetor v em E como $v = v- < v, w > /\|w\|^2 w+ < v, w > /\|w\|^2 w$. Ora,

$$< v- < v, w > /\|w\|^2 w, w >=< v, w > - < v, w >< w, w > / < w, w >= 0,$$

o que implica que $v- < v, w > /\|w\|^2 w \in (\hat{E}^{\perp})^{\perp} = \hat{E}$, pois $\hat{E}$ é fechado em E. Não há perda em normalizar w, isto é, supor que $\|w\| = 1$. Afirmamos que w realiza a norma de $\hat{g}$. De fato, se $v \in E$ é outro vetor de norma 1, não colinear a w, vimos acima que $v = v - \hat{v}+ < v, w > w$, com $\hat{v} \in \ker(\hat{g})$. Daí,

$$|\hat{g}(v)| = |\hat{g}(\hat{v})+ < v, w > \hat{g}(w)| = | < v, w > |\|\hat{g}(w)\| <$$

(aplicando Cauchy-Schwarz em sua forma estrita, e supondo sem perda $\hat{g} \not\equiv 0$)

$$\|v\|\|w\|\|\hat{g}(w)\| = \|hatg(w)\|,$$

o que implica que $\|\hat{g}\| = \|\hat{g}(w)\|$, como afirmamos. Observe ainda que $\|\hat{g}\| = \lim_{n\to\infty} \|g_n\|$. De fato,

$$\lim_{n\to\infty} \|g_n\| = \lim_{n\to\infty} \frac{< g_n, g_n >}{\sqrt{< g_n, g_n >}} = \lim_{n\to\infty} < \frac{g_n}{\sqrt{< g_n, g_n >}}, g_n >\geq$$

(novamente, por Cauchy-Schwarz)

$$\lim_{n\to\infty} < w, g_n >= |\hat{g}(w)| = \|\hat{g}\|.$$

Para a outra desigualdade, começamos por observar que para cada $j \in \mathbb{N}$, vale $\hat{g}(g_j/\|g_j\|) = \lim_{n\to\infty} < g_j/\|g_j\|, g_n >\leq \|hatg\|$. Por outro lado, como g_n é de Cauchy, ela é limitada, digamos, com norma acotada por $M > 0$ e ainda como $g_n \not\to 0$, dado $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que

$$\|\frac{g_j}{\|g_j\|} - \frac{g_n}{\|g_n\|}\| < \epsilon/M, \forall j, n \geq n_0.$$

Isso implica que

$$| < \frac{g_j}{\|g_j\|}, g_n > - < \frac{g_n}{\|g_n\|}, g_n > | \leq \|\frac{g_j}{\|g_j\|} - \frac{g_n}{\|g_n\|}\|M < \epsilon, \forall j, n \geq n_0,$$

e por conseguinte, $\lim_{n\to\infty} \sqrt{< g_n, g_n >} = \lim_{j\to\infty} \hat{g}(g_j) \leq \|\hat{g}\|$.

Desse modo,

$\lim_{n\to\infty} < g_n - < g_n, w > /\|w\|^2 w, g_n - < g_n, w > /\|w\|^2 w >= \lim_{n\to\infty} < g_n - < g_n, w > w, g_n >=$

$\lim_{n\to\infty} < g_n, g_n > - << g_n, w > w, g_n >= \lim_{n\to\infty} < g_n, g_n > - < g_n, w >^2 = \|\hat{g}\|^2 - \|\hat{g}(w)\|^2 = 0.$

Daí, concluímos que existe $\lim_{n\to\infty} g_n$, e este seria um múltiplo não nulo de w, o que contradiz a hipótese de que a sequência g_n não converge em E.

□

É bastante fácil ver que dado um espaço vetorial E munido com um produto interno e um seu subespaço vetorial $\hat{E} \subset E$, temos $E = \overline{\hat{E}} \oplus \hat{E}^{\perp}$. Deixamos isso como exercício, que usaremos no próximo

Lema 7.2.7. *(Versão fraca da Identidade de Parseval, ou Teorema de Pitágoras). Seja E um espaço vetorial com produto interno e seja $\hat{E} \subset E$ um subespaço vetorial de dimensão finita, o qual dotamos do produo interno oriundo de E. Suponha que $\hat{\beta} = \{\hat{v}_1, \ldots, \hat{v}_n\}$ seja uma base ortonormal de $\hat{E}$. Então, dado $v \in E$, este se escreve de maneira única como $v = \alpha_1 \hat{v}_1 + \cdots + \alpha_n \hat{v}_n + \hat{v}^{\perp}$, onde $\hat{\alpha}_1 =< v, \hat{v}_1 >, \ldots, \hat{\alpha}_n =< v, \hat{v}_n >$ e $\hat{v}^{\perp} \in \hat{E}^{\perp}$, valendo*

$$\|v\|^2 = (\sum_{j=1}^{n} |\alpha_j|^2) + \|\hat{v}^{\perp}\|^2.$$

Em particular, vale $\|v\|^2 \geq \sum_{j=1}^{\infty} |\alpha_j|^2$.

Prova: Como $\hat{E}$ tem dimensão finita, em particular é fechado em E, implicando que $E = \hat{E} \oplus \hat{E}^{\perp}$. Assim, dado $v \in E$, podemos escrever $v = \hat{v} + \hat{v}^{\perp}$, com $\hat{v} \in \hat{E}$ e $\hat{v}^{\perp} \in \hat{E}^{\perp}$. Ademais, $\hat{v} = \sum_{j=1}^{n} \alpha_j \hat{v}_j$, com

$$\alpha_j =< \hat{v}, \hat{v}_j >=< \hat{v} + \hat{v}^{\perp}, \hat{v}_j >=< v, \hat{v}_j >,$$

devido à ortogonalidade existente entre $\hat{v}^{\perp}$ e v_j.

Finalmente, temos

$$< v, v >=< \sum_{j=1}^{n} \alpha_j \hat{v}_j + \hat{v}^{\perp}, \sum_{j=1}^{n} \alpha_j \hat{v}_j + \hat{v}^{\perp} >=$$

(devido às relações de ortogonalidade existentes entre os diversos vetores $\hat{v}_1, \ldots \hat{v}_n$ e $\hat{v}^{\perp}$)

$$\sum_{j=1}^{n} |\alpha_j|^2 < \hat{v}_j, \hat{v}_j > + < \hat{v}^{\perp}, \hat{v}^{\perp} >= (\sum_{j=1}^{n} |\alpha_j|^2) + \|\hat{v}^{\perp}\|^2.$$

□

Lema 7.2.8. *Seja E um espaço de Hilbert e seja $\hat{E} \subset E$ um subespaço fechado próprio. Então, dado $v \notin \hat{E}$, existe $\hat{v} \in \hat{E}$ tal que*

$$\inf_{\hat{x}\in\hat{E}}\{\|v-\hat{x}\|\} = \|v-\hat{v}\|.$$

Ademais, $v - \hat{v} = w \in \hat{E}^{\perp}$, o que implica que $\hat{E}^{\perp} \neq \{0\}$.

Prova: Seja $\delta = \inf_{\hat{x}\in\hat{E}}\{\|v-\hat{x}\|\}$. Seja $(\hat{x}_j), \hat{x}_j \in \hat{E}$ uma sequência que minimiza a distância entre v e $\hat{E}$. Mostremos que (x_j) é de Cauchy. De fato,

$$0 \leq \|x_j - x_m\|^2 =< x_j - x_m, x_j - x_m >=< x_j - v + v - x_m, x_j - v + v - x_m >=$$

$$< x_j - v, x_j - v > + < v - x_m, v - x_m > + < x_j - v, v - x_m > + < v - x_m, x_j - v >=$$

$$\|x_j - v\|^2 + \|x_m - v\|^2 - < x_j - v, x_m - v > - < x_m - v, x_j - v > \leq$$

(pela desigualdade de Cauchy-Schwarz)

$$\|x_j - v|^2 + \|x_m - v\|^2 - 2\|x_j - v\|\|x_m - v\|,$$

que converge a zero, uma vez que $\|x_j - v\| \to \delta, \|x_m - v\| \to \delta$, quando $j \to \infty$, $m \to \infty$. Concluímos que (x_j) é de Cauchy, e como a sequência minimizante tomada é arbitrária, concluímos (por argumento canônico de Análise) toda sequência minimizante possui o mesmo limite, digamos $\hat{v} \in \hat{E}$. Como $v \notin \hat{E}$, segue-se que $w = v \setminus \hat{v} \neq 0$. Mostremos que $w \in \hat{E}^{\perp}$. Para $\hat{x} \in \hat{E}$ e $\alpha \in \mathbb{R}$ (ou $\mathbb{C}$, se o espaço for complexo), temos:

$$\delta^2 \leq < \underbrace{v - \hat{v}}_{=w} + \alpha\hat{x}, \underbrace{v - \hat{v}}_{=w} + \alpha\hat{x} > \Leftrightarrow$$

$$\delta^2 \leq < w, w > + < w, \alpha\hat{x} > + < \alpha\hat{x}, w > + |\alpha^2| < \hat{x}, \hat{x} > \Leftrightarrow$$

$$0 \leq \overline{\alpha} < w, \hat{x} > + \alpha < \hat{x}, w > + |\alpha|^2 < \hat{x}, \hat{x} > .$$

Agora, fazendo $\alpha = r < w, \hat{x} >$, onde $r \in \mathbb{R}$ é qualquer, obtemos:

$$0 \leq 2r| < w, \hat{x} > |^2 + r^2| < w, \hat{x} > |^2\|\hat{x}\|^2, \forall r \in \mathbb{R},$$

o que só é possível se $< w, \hat{x} >= 0$. Como $\hat{x} \in \hat{E}$ é arbitrário, segue-se que $w \in \hat{E}^{\perp}$.

□

Note que é imediato do lema acima que se E é um espaço de Hilbert e $\hat{E}$ é um seu subespaço fechado, então $E = \hat{E} \oplus \hat{E}^{\perp}$.

Teorema 7.2.9. *(Representação de Riesz). Seja E um espaço de Hilbert. Então, dado um funcional linear contínuo $f \in E^*$, existe um único $w \in E$ tal que $f(x) =< x, w >$, $\forall x \in E$.*

Prova:

Suponha que $f \neq 0$, pois este caso é imediato. Seja $\hat{E} = \ker(f)$. Como f é contínuo, $\hat{E}$ é fechado em E. Pelo lema anterior, $\ker(f)^\perp \neq \{0\}$. Seja $\tilde{w} \neq 0$ um vetor em $\ker(f)^\perp$ tal que $f(\tilde{w}) = 1$, e seja $w := \tilde{w}/ < \tilde{w}, \tilde{w} >$. Daí, dado $v \in E$, escrevendo $v = (v- < v, \tilde{w} > \tilde{w}/ < \tilde{w}, \tilde{w} >)+ < v, \tilde{w} > \tilde{w}$, é claro que $\hat{v} := (v- < v, \tilde{w} > \tilde{w}/ < \tilde{w}, \tilde{w} >) \in \ker(f)$, temos:

$$f(v) = f(\hat{v}) + f(< v, \tilde{w} > \tilde{w}/ < \tilde{w}, \tilde{w} >) = \frac{< v, \tilde{w} >}{< \tilde{w}, \tilde{w} >} f(\tilde{w}) =< v, w > .$$

Finalmente, para vermos a unicidade, basta aplicarmos mais uma vez o lema: w e z são tais que $f(v) =< v, w >=< v, z >, \forall v \in E$, então vale:

$$< v, w >=< v, z >, \forall v \in E \Leftrightarrow < v, z-w >= 0, \forall v \in E \Leftrightarrow z-w \in E^\perp = \{0\},$$

implicando que $z = w$.

□

Corolário 7.2.10. *Seja E um espaço de Hilbert. Então a aplicação $F : E \to E^*$ dada por*

$$F(w) =< \cdot, w >,$$

é um isomorfismo (sesqui)linear isométrico de E em E^*.

Prova:

Da proposição anterior, fica claro que F é sobrejetivo: se $f \in E^*$, vimos que existe w tal que $f(\cdot) =< \cdot, w >= F(w)$. Também é óbvia a injetividade, pela unicidade vista no teorema de Representação de Riesz. Se o espaço for real, o isomorfismo acima é claramente linear. No caso complexo, é sesquilinear.

Mostremos que $\|F(w)\| = \|w\|, \forall w \in E$.

De fato,

$$\|F(w)\| = \sup_{\|v\|=1} | < v, w > |,$$

supremo que sabemos, da desigualdade de Cauchy-Schwarz, ser atingido para $v = w/\|w\|$. Por conseguinte, $\|F(w)\| = | < w, w > /\|w\|| = \|w\|$, ou seja, F é isometria.

□

Observação 7.2.11. (Representação dos funcionais lineares em E^*, quando E é espaço vetorial com produto interno, não necessariamente completo). Seja E um espaço vetorial dotado de um produto interno, e $f : E \to \mathbb{R}_{\mathbb{C}}$ um funcional linear contínuo. Então, pelo teorema de extenso de operadores lineares (o conhecido B.L.T.), o funcional linear f possui uma única extensão contínua $\tilde{f} : \tilde{E} \to \mathbb{R}_{\mathbb{C}}$, onde $\tilde{E}$ é o completamento de E. Analogamente, dado $\tilde{f} : \tilde{E} \to \mathbb{R}_{\mathbb{C}}$ um funcional linear contínuo, sua restrição a E determina um único funcional contínuo $f : E \to \mathbb{R}_{\mathbb{C}}$. Em ambos os casos, como E é denso em $\tilde{E}$, obtemos que $\|f\| = \|\tilde{f}\|$. Isso implica que E^* é isometricamente isomorfo a $\tilde{E}^*$, via aplicação $\tilde{F} : E^* \to \tilde{E}^*$ dada por $\tilde{F}(f) = \tilde{f}$, em que $\tilde{f}$ é a única extensão contínua de um funcional f com domínio em E ao completamento $\tilde{E}$. Ora, do Teorema de Representação de Riesz, temos que qualquer funcional linear contínuo $\tilde{f}$ (definido no espaço de Hilbert $\tilde{E}$, completamento de E) é da forma:

$$\tilde{f}(\tilde{x}) =< \tilde{x}, \tilde{w} >, \forall \tilde{x} \in \tilde{E},$$

onde $\tilde{w} \in \tilde{E}$ é um vetor constante, unicamente determinado por $\tilde{f}$. Ora, se $f = \tilde{F}^{-1}(\tilde{f})$, então $f = \tilde{f}|_E$. Em particular, tomando-se uma sequência $w_n \to \tilde{w}$, onde $w_n \in E$, é claro que para $x \in E$ vale

$$f(x) = \tilde{f}(x) =< x, \tilde{w} >= \lim_{n\to\infty} < x, w_n >,$$

o que nos fornece uma representação (não única) para os funcionais lineares em E, simplesmente em termos de sequências em E.

A mais importante conclusão a que chegamos a partir da observação acima é que embora nem todo funcional linear em E^* (quando E não é completo) possa ter uma representação do tipo $f(x) =< x, w >, \forall x \in E$, com $w \in E$, vetor constante, mesmo assim, os funcionais desse tipo podem ser usados para aproximar qualquer funcional em E^*, pois formam um subconjunto denso de E^*. Desse modo, estamos aptos a fazer a seguinte:

Definição 7.2.12. (Operador Adjunto em Espaços vetoriais com produto interno). Seja $A : E \to E$ um operador linear contínuo, definido no espaço vetorial E, dotado de produto interno $< \cdot, \cdot >$. O adjunto de A é o único operador linear $A^* : E \to E$ dado por:

$$< A \cdot x, y >=< x, A^* \cdot y >, \forall x, y \in E.$$

Compare a definição acima com a definição de operador adjunto em espaços normados. Claro está que no caso de espaços vetoriais dotados com produto interno, identificamos E com seu mergulho em E^*. Com isso, temos que em espaços dotados de produto interno, tanto o operador como seu adjunto atuam no mesmo domínio, E.

Uma proposição básica sobre adjuntos de um operador é a seguinte:

Proposição 7.2.13. *Seja E um espaço de Hilbert e $A \in \mathcal{L}(E)$. Então*

$$\ker(A)^\perp = \overline{A^*(E)} \text{ e } A(E)^\perp = \ker(A^*).$$

Prova:

Note que o espaço ortogonal a qualquer subespaço é sempre fechado, devido á continuidade do produto interno. Seja $v \in \ker(A)$, e $\hat{w} = A(w) \in A^*(E)$. Daí,

$$< v, \hat{w} >=< v, A^*(w) >=< A(v), w >=< 0, w >= 0,$$

o que quer dizer que $A^*(E) \subset \ker(A)^\perp$. Por outro lado, se $\overline{A^*(E)} \neq \ker(A)^\perp$ então existiria um vetor não nulo $\hat{v} \in \overline{A^*(E)}^\perp \cap \ker(A)^\perp$. Daí,

$$< \hat{v}, \hat{w} >= 0, \forall \hat{w} \in \ker(A) + \overline{A^*(E)} \Rightarrow$$

(em particular, para $\hat{w} \in A^*(E)$)

$$< \hat{v}, A^*(w) >= 0, \forall w \in E \Rightarrow < A(\hat{v}), w >= 0, \forall w \in E \Rightarrow$$

$$A(\hat{v}) = 0,$$

o que é absurdo, pois $\hat{v}$ foi suposto não nulo e pertencente ao $\ker(A)^\perp$. Isso mostra que $\overline{A^*(E)} = \ker(A)^\perp$.

Isso também implica que $A^*(E)^\perp \supset (\ker(A)^\perp)^\perp = \ker(A)$.

□

Proposição 7.2.14. *Seja E um espaço dotado de produto interno e $A : E \to E$ um operador linear auto-adjunto contínuo. Então*

$$\|A\| = \sup_{\|x\|=1} \{\|A(x)\|\} = \sup_{\|x\|=1} \{| < A(x), x > |\}$$

Prova: Em primeiro lugar, se $x \in E$ com $\|x\| = 1$, claramente

$$| < A(x), x > | \leq \|A(x)\|\|x\| \leq \|A\|.$$

Por outro lado, note que

$$\|A(x)\| =< A(x)/\|A(x)\|, A(x) >=< A(A(x))/\|A(x)\|, x >\leq$$

(devido à desigualdade de Cauchy-Schwarz)

$$< A(A(x))/\|A(x)\|, A(x)/\|A(x)\| > \leq \sup_{\|y\|=1} | < A(y), y > |.$$

□

Estamos particularmente interessados em estudar operadores definidos em um espaço normado separável E cuja norma provém de um produto interno $< \cdot, \cdot >$. Por E ser *separável* entendemos que E possui um subconjunto enumerável denso. Conforme nos assevera a próxima proposição, isso equivale a existência de uma base ortonormal enumerável:

Proposição 7.2.15. *Seja E um espaço vetorial com produto interno separável. Então E possui uma base (analítica) de vetores ortogonais enumerável. Isso quer dizer que existem $v_1, \ldots, v_n, \cdots \in E$, vetores não nulos tais que $< v_j, v_k >= 0, \forall j \neq k$ e dado $v \in E$ este se escreve de maneira única como $v = \sum_{n=1}^{+\infty} \alpha_n v_n$, com $\alpha_1, \ldots, \alpha_n, \ldots$ escalares. Reciprocamente, se E é um espaço com produto interno exibindo uma base ortogonal enumerável, então E é separável.*

Prova:

($\Rightarrow$) Seja $D := \{w_j, j \in \mathbb{N}\}$ um subconjunto denso enumerável de E. Definamos indutivamente uma sequência de subconjuntos β_n encaixantes de D. Sem perda, podemos supor que $w_1 \neq 0$. Desse modo, definimos $\beta_1 := \{w_1/\|w_1\|\}$. Suponha que para $n \in \mathbb{N}$, obtivemos uma base ortonormal de $span(\{w_1, \ldots, w_n\}) := E_n$. Então colocamos $\beta_{n+1} := \beta_n$ se $span(\beta_n) = span(\{w_1, \ldots, w_{n+1}\}) := E_{n+1}$. Caso contrário, se $k_n = \#\beta_n$, escrevemos

$$v_{k_n} := \frac{w_{n+1} - \sum_{j=1}^{k_n} < w_{n+1}, v_j > v_j}{\|w_{n+1} - \sum_{j=1}^{k_n} < w_{n+1}, v_j > v_j\|},$$

e pomos $\beta_{n+1} = \beta_n \cup \{v_{k_n}\}$ (veja que a cardinalidade k_n de β_n pode não ser n). Tomando $\beta = \cup_{n\in\mathbb{N}}\beta_n$, é imediato que β é um conjunto ortonormal e que

$D \subset span(\beta)$. Em particular, $\overline{span(\beta)} \supset \overline{D} = E$. Seja $v \in E$. Mostremos que

$$v = \sum_{n=1}^{\infty} < v, v_n > v_n$$

Para tal, tome $w_j \to v$, tal que $w_j \in E_j$, $\forall j \in \mathbb{N}$ (para obter w_j com essas propriedades, alguns desses w_j talvez tenham de ser tomados repetidos, e tenhamos de lançar mão de que $E_j \subset E_{j+1}$). Como $w_j \in E_j$, temos que

$$w_j = \sum_{n=1}^{j} < w_j, v_n > v_n = \sum_{n=1}^{\infty} < w_j, v_n > v_n,$$

o que por Pitágoras nos dá

$$\|w_j\|^2 = \sum_{n=1}^{n_j} < w_j, v_n >^2 = \sum_{n=1}^{\infty} < w_j, v_n >^2 .$$

Prosseguindo, vale que

$$\|v - \sum_{n=1}^{j} < v, v_n > \| \leq \|v - w_j\| + \|(\sum_{n=1}^{j} < v, v_n > - \sum_{n=1}^{j} < w_j, v_n >)v_n\|$$

Note que tendo em vista os últimos parágrafos, a primeira parcela vai a zero quando $j \to \infty$. Já para a última parcela, temos:

$$\| \sum_{n=1}^{j} < w_j - v, v_n > v_n\|^2 = \sum_{n=1}^{j} < w_j - v, v_n >^2 \leq \|w_j - v\|^2,$$

implicando na ida da proposição.

($\Leftarrow$) A prova da recíproca é muito simples. Basta tomar

$$D := \{\sum_{j=1}^{n} \alpha_j v_j; n \in \mathbb{N}, \alpha_j \in \mathbb{Q}\},$$

onde $\{v_j, j \in \mathbb{N}\}$ é base ortogonal enumerável de E. Claramente, D é um subconjunto denso e enumrável de E.

$\square$

7.3 Operadores Compactos e Problemas de Contorno

É um fato bem conhecido em cursos básicos de Análise (vide o apêndice em [**?**]) que um operador linear, ou é Lipschitz, ou é descontínuo em todos os pontos de seu domínio. De maneira intuitiva, isso quer dizer que em qualquer vizinhança de qualquer ponto de seu domínio, existem pontos que são levados em pontos distantes na imagem. Ora, se um operador linear A descontínuo e invertível leva pontos próximos em pontos distantes, é de se esperar que sua inversa precise ser supercontínua, para tomar tais pontos distantes e aproximá-los novamente.

Essa é mais ou menos a idéia desta seção: em geral, operadores que envolvem a derivação em dimensão infinita são descontínuos. Contudo, em muitos casos tais operadores possuem como inversa operadores lineares supercontínuos (tecnicamente conhecidos como *operadores compactos*). Desse modo, podemos resolver problemas associados a operadores de Derivação simplesmente estudando suas inversas.

Definição 7.3.1. (Operador Compacto). Seja E um espaço de Hilbert. Um operador linear $A : E \to E$ é dito *compacto* se dada $x_n \in E$ limitada, então $A(x_n)$ possui subsequência convergente.

Claramente todo operador compacto é contínuo. De fato, se (x_n), $\|x_n\| = 1$ é tal que $\|A(x_n)\| \to \sup_{\|x\|=1}\{\|A(x)\|\}$ quando $n \to +\infty$, então $\|A(x_n)\| \not\to +\infty$, já que da compacidade de A, tal sequência possui subsequência convergente. Também é fácil ver que o conjunto dos operadores compactos forma um subespaço vetorial fechado de $\mathcal{L}(E)$ que também é um ideal em relação a esse espaço: o produto (composição) de um operador compacto qualquer por um operador contínuo é um operador compacto.

Para operadores compactos em espaços de Hilbert, vale o seguinte

Teorema 7.3.2. *(Teorema!espectral!para operadores compactos). Seja E um espaço de Hilbert, e $A : E \to E$ um operador compacto auto-adjunto. Então*

Prova:

□

Capítulo 8

O Teorema de Grobman-Hartman

O seguinte lema será importante por toda essa seção:

Lema 8.0.3. *Seja E um espaço de Banach, $f : \mathcal{N} \subset E \to E$ uma aplicação C^k, $k \geq 1$ de um aberto $\mathcal{N} \subset E$ contendo 0, com $f(0) = 0$ e seja $A = Df_0$. Dado $\epsilon > 0$, existe uma vizinhança $U = U(0)$ e existe uma extensão de $f|_U$ da forma $(A+\phi)$, onde $\phi \in C_b^0(E)$ é lipschitziana com constante de Lipschitz limitada por ϵ.*

Prova: Seja $\beta : \mathbb{R} \to [0,1]$ uma função C^∞ com as seguintes propriedades:

$$\begin{array}{ccc} \beta(t) = 0 & \text{se} & t \geq 1 \\ \beta(t) = 1 & \text{se} & t \leq 1/2 \\ |\beta'(t)| \leq K, & \forall t \in \mathbb{R}, & K > 2. \end{array}$$

Seja $f = A + \varphi$, com $\varphi(0) = 0$ e $D\varphi_0 = 0$. Considere B_r uma bola de centro na origem e raio $r > 0$ tal que $\|D\varphi_x\| < \epsilon/2K$, $\forall x \in B_r$. (Para a existência de tal bola, usamos apenas a continuidade de $D\varphi$, decorrente do fato de f ser C^1).

Tomemos

$$\phi(x) := \beta\Big(\frac{|x|}{r}\Big) \cdot \varphi(x).$$

Daí, $\phi(x) = 0$ se $|x| \geq r$, o que implica que ϕ é limitada em E, visto que $|\phi| \leq |\varphi|$, e devido a desigualdade do valor médio, a constante de Lipschitz de $\varphi|_{B_r}$ é menor que $\epsilon/(2K)$. Em resumo:

$$|\phi(x)| \leq |\varphi(x)| = |\varphi(x) - \varphi(0)| \leq \frac{\epsilon}{2K} \cdot |x - 0| \leq \frac{\epsilon}{2K} \cdot r, \forall x \in B_r.$$

Temos ainda que $\phi(x) = \varphi(x)$ se $|x| \leq r/2$, donde concluímos que $A + \phi$ é extensão de $f|_{B_{r/2}}$.

Mostremos que ϕ é lipschitziana e sua constante de Lipschitz pode ser tomada como menor ou igual a ϵ.

Realmente, se x_1 e x_2 pertencem a B_r, temos:

$$\Big|\phi(x_1) - \phi(x_2)\Big| = \Big|\beta\Big(\frac{|x_1|}{r}\Big) \cdot \varphi(x_1) - \beta\Big(\frac{|x_2|}{r}\Big) \cdot \varphi(x_2)\Big| =$$

$$\Big|\Big(\beta\Big(\frac{|x_1|}{r}\Big) - \beta\Big(\frac{|x_2|}{r}\Big)\Big) \cdot \varphi(x_1) - \beta\Big(\frac{|x_2|}{r}\Big) \cdot (\varphi(x_2) - \varphi(x_1))\Big| \leq$$

$$\Big|\Big(\beta\Big(\frac{|x_1|}{r}\Big) - \beta\Big(\frac{|x_2|}{r}\Big)\Big) \cdot \varphi(x_1)\Big| + \Big|\underbrace{\beta\Big(\frac{|x_2|}{r}\Big)}_{0 \leq \beta(\cdot) \leq 1} \cdot (\varphi(x_2) - \varphi(x_1))\Big| \leq$$

$$K \cdot \Big|\frac{|x_1| - |x_2|}{r}\Big| \cdot \frac{\epsilon}{2K} \cdot |x_1| + \frac{\epsilon}{2K} \cdot |x_1 - x_2| \leq$$

$$\frac{|x_1 - x_2|}{r} \cdot \frac{\epsilon}{2} \cdot r + \frac{\epsilon}{2} \cdot |x_1 - x_2| = \epsilon \cdot |x_1 - x_2|.$$

Se $x_1 \in B_r$ e $x_2 \notin B_r$, obtemos, a partir das mesmas contas:

$$\Big|\phi(x_1) - \phi(x_2)\Big| = \Big|\beta\Big(\frac{|x_1|}{r}\Big) \cdot \varphi(x_1) - \beta\Big(\frac{|x_2|}{r}\Big) \cdot \varphi(x_2)\Big| \leq$$

$$\Big|\Big(\beta\Big(\frac{|x_1|}{r}\Big) - \beta\Big(\frac{|x_2|}{r}\Big)\Big) \cdot \varphi(x_1)\Big| + \Big|\underbrace{\beta\Big(\frac{|x_2|}{r}\Big)}_{=0,\text{ pois } |x_2|/r \geq 1} \cdot (\varphi(x_2) - \varphi(x_1))\Big| \leq$$

$$K \cdot \Big|\frac{|x_1| - |x_2|}{r}\Big| \cdot \frac{\epsilon}{2K} \cdot |x_1| \leq \frac{|x_1 - x_2|}{r} \cdot \frac{\epsilon}{2} \cdot r \leq \epsilon \cdot |x_1 - x_2|.$$

Finalmente, se $x_1 \notin B_r$ e $x_2 \notin B_r$, temos que

$$|\phi(x_1) - \phi(x_2)| = 0 \leq \epsilon \cdot |x_1 - x_2|.$$

□

Observação 8.0.4. Apenas para título de informação, observamos que no caso em que E é um espaço de Hilbert (isto é, um espaço de Banach cuja norma provém de um produto interno), a aplicação ϕ é de classe C^k. Em tal contexto, a norma é C^∞, exceto na origem, mas em torno da origem, β é constante, implicando que $\beta \circ |\cdot|$ é C^∞ e consequentemente ϕ tem a mesma classe de diferenciabilidade que f.

8.1 O teorema de Grobman-Hartman para difeomorfismos

Definição 8.1.1. (Ponto fixo hiperbólico). Seja E um espaço de Banach. Um isomorfismo linear $A \in \mathcal{L}(E)$ é dito *hiperbólico* se o espectro de A não intersecta a esfera S^1. Se E tem dimensão finita, isso é o mesmo que dizer que nenhum autovalor de A tem norma 1. Dado um difeomorfismo C^k $f : U \subset E \to E$, um ponto fixo $p \in U$ de f é dito *hiperbólico* se Df(p) é um isomorfismo hiperbólico.

No texto abaixo, $\mathcal{M}$ designa uma variedade diferenciável (incluindo a possibilidade de ser um espaço de Banach), em dimensão qualquer.

Teorema 8.1.2. *(Grobman-Hartman para difeomorfismos) Sejam $f \in Diff^k(\mathcal{M})$ e $p \in \mathcal{M}$ um ponto fixo hiperbólico de f. Seja $A = Df_p : T\mathcal{M}_p \to T\mathcal{M}_p$. Então existem vizinhanças $V = V(p) \subset \mathcal{M}$ e $U = U(0)$ de $T\mathcal{M}_p$ e um homeomorfismo $h : U(0) \to V(p)$ tais que*

$$h \circ A = f \circ h$$

Se $A : E \to E$ é um isomorfismo linear hiperbólico do espaço de Banach E nele mesmo, existe uma decomposição invariante (por A) $E = E^s \oplus E^u$ e uma norma (equivalente a norma original de E) $|\cdot|$ em E segundo a qual

$$\begin{array}{ll} \|A^s\| \leq a < 1, & \text{onde } A^s := A|_{E^s} : E^s \to E^s \\ \|(A^u)^{-1}\| \leq a < 1, & \text{onde } A^u := A|_{E^u} : E^u \to E^u \end{array}$$

Seja $C_b^0(E)$ o espaço de aplicações contínuas e limitadas de E em E com a norma uniforme: $|v| := \sup\{|v(x)|, x \in E\}, \forall v \in C_b^0(E)$. Como $E = E^s \oplus E^u$, temos uma decomposição em soma direta

$$C_b^0(E) := C_b^0(E, E^s) \oplus C_b^0(E, E^u), \text{ com } v = v_s \oplus v_u, \forall v \in C_b^0(E),$$

onde $v_s := \pi_s \circ v$ e $v_u := \pi_u \circ v$ ($\pi_s : E^s \oplus E^u \to E^s$ e $\pi_u : E^s \oplus E^u \to E^u$ são as projeções naturais).

O Teorema de Grobman-Hartman para difeomorfismos será consequência do seguinte lema, que pode ser entendido como um proto-teorema de Grobman-Hartman:

Lema 8.1.3. *Se $A : E \to E$ é um isomorfismo hiperbólico do espaço de Banach E, então existe $\epsilon > 0$ tal que se ϕ_1 e $\phi_2 \in C_b^0(E)$ têm constante de Lipschitz menor ou igual a ϵ, então $(A + \phi_1)$ e $(A + \phi_2)$ são conjugados.*

Prova: Estamos à cata de um homeomorfismo $h : E \to E$ tal que

$$h \circ (A + \phi_1) = (A + \phi_2) \circ h.$$

Procuremos uma solução desta equação funcional da forma $h = I + w$, com $w \in C_b^0(E)$ (isto é, h está a distância finita da identidade).

Manipulando nossa equação funcional, obtemos:

$$(I + w) \circ (A + \phi_1) = (A + \phi_2) \circ (I + w) \Leftrightarrow$$

$$A + \phi_1 + w \circ (A + \phi_1) = A + A \circ w + \phi_2 \circ (I + w) \Leftrightarrow$$

$$A \circ w - w \circ (A + \phi_1) = \phi_1 - \phi_2 \circ (I + w). \tag{8.1}$$

Mostremos que $\exists! w \in C_b^0(E)$ satisfazendo a úlima expressão acima. Para tal, consideremos o operador linear no espaço de Banach $C_b^0(E)$ dado por

$$\hat{L}(y) := A \circ y - y \circ (A + \phi_1).$$

Provemos que $\hat{L}$ é inversível com

$$\|\hat{L}^{-1}\| \leq \frac{\|A^{-1}\|}{(1 - a)}.$$

De fato, $\hat{L} = \hat{A} \cdot L$, onde $L : C_b^0(E) \to C_b^0(E)$ é dado por $L(y) := y - A^{-1} \circ y \circ (A + \phi_1)$ e $\hat{A} : C_b^0(E) \to C_b^0(E)$ é dada simplesmente por $\hat{A}(y) = A \circ y$. Portanto, com $\hat{A}$ é inversível, basta provarmos que L também o é, que teremos $\hat{L}$ inversível, com $\hat{L}^{-1} = L^{-1} \cdot \hat{A}^{-1}$.

Observamos que $C_b^0(E, E^s)$ e $C_b^0(E, E^u)$ são invariantes por L, pois E^s e E^u são invariantes por A^{-1}. Se $y_s \in C_b^0(E, E^s)$ (resp. $y_u \in C_b^0(E, E^u)$), claramente $L(y_s) = y_s - A^{-1} \circ y_s \circ (A + \phi_1)$ é uma aplicação com imagem contida em E^s (resp. em E^u), logo pertence a $C_b^0(E, E^s)$ (resp. $C_b^0(E, E^u)$).

Daí, podemos escrever $L = L^s + L^u$, onde

$$L^s := L|_{C_b^0(E,E^s)},\ L^s : C_b^0(E, E^s) \to C_b^0(E, E^s);$$
$$L^u := L|_{C_b^0(E,E^u)},\ L^u : C_b^0(E, E^u) \to C_b^0(E, E^u).$$

Se ϵ é suficientemente pequeno, temos do lema de perturbação de isomorfismo, que $(A + \phi_1)$ é um homeomorfismo. Logo, o operador $y_s \mapsto A^{-1} \circ y_s \circ (A + \phi_1)$ é inversível, e sua inversa (que corresponde a compor à esquerda com A^s e à direita com $(A + \phi_1)^{-1}$)

$$y_s \mapsto A^s \circ y_s \circ (A + \phi_1)^{-1}$$

é uma contração com norma limitada por $a < 1$. Pelo lema 0.2.14, parte b), temos que L^s é inversível e $\|(L^s)^{-1}\| \leq a/(1-a)$. Pela parte a) do mesmo lema, L^u é inversível com $\|(L^u)^{-1}\| \leq 1/(1-a)$. Portanto, $\hat{L}$ é inversível, com norma

$$\|\hat{L}^{-1}\| = \|L^{-1} \cdot \hat{A}\| \leq \frac{\|A^{-1}\|}{1-a}.$$

Temos portanto que a equação 8.1 é equivalente a

$$w = \hat{L}^{-1}(\phi_1 - \phi_2 \circ (I + w)) \tag{8.2}$$

Mas w satisfará a equação 8.2 se e só se for ponto fixo do operador $T : C_b^0(E) \to C_b^0(E)$, cuja fórmula é dada por:

$$T(y) := \hat{L}^{-1}(\phi_1 - \phi_2 \circ (I + y)).$$

Dados y_1 e y_2 pertencentes a $C_b^0(E)$, temos:

$$\|T(y_1) - T(y_2)\| = \|\hat{L}^{-1}(\phi_1) - \hat{L}^{-1}(\phi_2 \circ (I + y_1)) - \hat{L}^{-1}(\phi_1) + \hat{L}^{-1}(\phi_2 \circ (I + y_2))\| =$$

$$\|\hat{L}^{-1}(\phi_2 \circ (I + y_2)) - \hat{L}^{-1}(\phi_2 \circ (I + y_1))\| \leq \|\hat{L}^{-1}\| \cdot |\phi_2 \circ (I + y_1) - \phi_2 \circ (I + y_2)| \leq$$

$$\frac{\|A^{-1}\|}{1-a} \cdot \epsilon \cdot |y_1 - y_2|.$$

Novamente, se ϵ é pequeno, T é contração e possui um único ponto fixo w que satisfaz 8.2 e logo

$$(I + w) \circ (A + \phi_1) = (A + \phi_2) \circ (I + w).$$

Só nos resta mostrar que $(I + w)$ é um homeomorfismo. Para tal, basta ver que com o mesmo raciocínio acima (permutando os papéis de ϕ_1 e ϕ_2), obtemos um único $v \in C_b^0(E)$ tal que

$$(I + v) \circ (A + \phi_2) = (A + \phi_1) \circ (I + v).$$

Mostremos que $(I + w) \circ (I + v) = (I + v) \circ (I + w) = I$. Realmente,

$$(I+w) \circ (I+v) \circ (A+\phi_2) = (I+w) \circ (A+\phi_1) \circ (I+v) = (A+\phi_2) \circ (I+w) \circ (I+v).$$

Temos por conseguinte que $(I + w) \circ (I + v)$ semiconjuga $(A + \phi_2)$ com ele mesmo. Mas $(I + w) \circ (I + v)$ está a uma distância finita da identidade:

$$(I + w) \circ (I + v) = I + \underbrace{v + w \circ (I + v)}_{\in C_b^0(E)}$$

Como a identidade I também semiconjuga $(A+\phi_2)$ consigo mesmo, da unicidade da construção feita, segue-se que $I = (I+w)\circ(I+v)$. Da mesma maneira, prova-se que $(I+v)\circ(I+w)$ semiconjuga $(A+\phi_1)$ consigo mesmo; o que implica que $h = (I+w)$ é homeomorfismo conjugando $(A+\phi_1)$ e $(A+\phi_2)$.

□

Prova: (Teorema de Grobman-Hartman para difeomorfismos)

Por tratar-se de resultado local, podemos, sem perda de generalidade, (usando cartas locais) supor f um difeomorfismo definido de uma vizinhança W para outra N de zero em $E = T_p\mathcal{M}$, com $f(0) = 0$. Seja $\epsilon_0 > 0$ tal que $A+\phi$ é globalmente conjugado a A em E, para todo ϕ limitado com constante de Lipschitz limitada por ϵ_0, conforme o lema 8.1.3. Para tal ϵ_0, pelo lema 8.0.3, podemos tomar uma vizinhança $B_r \subset W\cap N$ tal que $(A+\phi)|_{B_{r/2}} = f|_{B_{r/2}}$, $(A+\phi)|_{B_r{}^c} = A$, ϕ é limitada e tem constante de Lipschitz menor ou igual a ϵ_0. Pelo lema 8.1.3, vale que $(A+\phi)$ é globalmente conjugado a A em E: existe um homeomorfismo $h : E \to E$ a distância finita da identidade tal que $h\circ A = (A+\phi)\circ h$.

Note que como A é isomorfismo hiperbólico, não possui outro ponto fixo exceto o zero (pois outro ponto fixo diferente de zero seria um autovetor do autovalor 1). Tal implica que $(A+\phi)$ possui um único ponto fixo, pois se p é ponto fixo de $(A+\phi)$ temos

$$h\circ A(h^{-1}(p)) = (A+\phi)\circ h(h^{-1}(p)) \Rightarrow$$

$$h\circ A\circ h^{-1}(p) = (A+\phi)(p) = p \Rightarrow A\circ h^{-1}(p) = h^{-1}(p),$$

o que significa que $h^{-1}(p)$ é o único ponto fixo de A, e portanto é zero. No nosso contexto, temos que $(A+\phi)(0) = 0$, portanto temos de ter $h(0) = 0$.

Por conseguinte, podemos restringir h a uma vizinhança $U := U(0) \subset B_{r/2}$ tal que $V := h(U(0)) \subset W$. Temos então que para todo $x \in U\cap A^{-1}(U)$ vale

$$h(A(x)) = f(h(x)),$$

findando a prova do teorema.

□

Note que pelo teorema de Grobman-Hartman para difeomorfismos, obtemos como Corolário que todo ponto fixo hiperbólico de um difeomorfismo é isolado. A próxima proposição mostra que tal fato vale para pontos fixos um pouco mais gerais que os hiperbólicos:

Proposição 8.1.4. *Seja $f : U \subset E \to E$ uma aplicação de classe C^1 limitada e com derivada limitada no aberto U contido em um espaço de Banach E em E. Se $p = p_f \in U$ é um ponto fixo de f (isto é, se $f(p) = p$) e $1 \notin Sp(Df_p)$, então toda g C^1-próxima a f possui um ponto fixo p_g. Além disso, existe uma vizinhança $V \subset U$ de p tal que p_g é o único ponto fixo de g em V.*

Prova: Começamos por considerar o espaço vetorial $C^1_b(U, E)$ das aplicações de classe C^1 com domínio em U e imagem contida em E, limitadas e com respectivas derivadas limitadas. Tal espaço, dotado da norma $|\cdot|_1$ definida por

$$|g|_1 := \max\{\sup_{x\in U}\{|g(x)|\}, \sup_{x\in U}\{\|Dg(x)\|\}\}, \forall g \in C^1_b(U, E),$$

é um espaço de Banach. Duas aplicações em $C^1_b(U, E)$ estão próximas (na norma $|\cdot|_1$) se elas mesmas e suas derivadas são uniformemente próximas ponto a ponto.

Seja $\delta > 0$ por fixar mais adiante. Defina o operador

$$F : \underbrace{B(f, \delta)}_{\subset C^1_b(U,E)} \times \underbrace{B(p, \delta)}_{\subset U} \to E,$$

dado por

$$F(g, x) := g(x) - x.$$

Dotaremos $B(f, \delta) \times B(p, \delta)$ da topologia produto. Nesta topologia, F é contínua. De fato, dados $(g, x) \in B(f, \delta) \times B(p, \delta)$ e $(h, y) \in B(f, \delta) \times B(p, \delta)$ temos

$$|F(g, x) - F(h, y)| = |g(x) - x - h(y) + y| \leq$$

$$|g(x) - g(y)| + |g(y) - h(y)| + |x - y| <$$

$$\sup_{z\in B(p,\delta)} \{\|D(g(z)\|\}|x - y| + |g - h|_1 + |x - y| <$$

(pois $|g|_1 < |f|_1 + \delta$)

$$(1 + |f|_1 + \delta)|x - y| + |g - h|_1 \leq (1 + |f|_1 + \delta) \max\{|g - h|_1, |x - y|\} =$$

$$(1 + |f|_1 + \delta)d((g, x), (h, y))$$

implicando que F é lipschitziana e portanto contínua.

Observe que

$$F(g+w,x) := (g+w)(x) - x = g(x) - x + w(x) = F(g,x) + w(x).$$

Note que a aplicação $C_b^1(U,E) \ni w \mapsto w(x) \in E$ é uma aplicação linear contínua.

Portanto, dado $(g,x) \in B(f,\delta) \times B(p,\delta)$, F tem derivada parcial dada por $\partial_1 F(g,x) \cdot w = w(x)$. Ademais, essa aplicação derivada parcial $(g,x) \mapsto \partial_1 F(g,x)$ é contínua. Realmente

$$\sup_{|w|_1=1} |(\partial_1 F(g,x) - \partial_1 F(\hat{g},\hat{x})) \cdot w| = |w(x) - w(\hat{x})| \leq |w|_1 \cdot |x - \hat{x}| = |x - \hat{x}| \leq$$

$$\max\{|g - \hat{g}|_1, |x - \hat{x}|\} = d((g,x),(\hat{g},\hat{x})).$$

Já a derivada parcial em relação a x em cada par (g,x) é a aplicação linear contínua de E em E dada por

$$\partial_2 F(g,x) = Dg(x) - I.$$

Como função de g e x, tal derivada é claramente contínua. De fato, seja fixado $(g,x) \in\in B(f,\delta) \times B(p,\delta)$. Dado $\epsilon > 0$ seja $\hat{\delta} > 0$ tal que $\|Dg(x) - Dg(\hat{x})\| < \epsilon/2$ para todo $\hat{x}$ com $|\hat{x} - x| < \hat{\delta}$. Daí,

$$\|Dg(x) - I - (D\hat{g}(\hat{x}) - I\| \leq \|Dg(x) - D\hat{g}(\hat{x})\| + \|Dg(\hat{x}) - D\hat{g}(\hat{x})\| \leq \epsilon/2 + \|g - \hat{g}\|_1 < \epsilon,$$

desde que tomemos $d((g,x),(\hat{g},\hat{x})) < \min\{\hat{\delta}, \epsilon/2\}$. Isto mostra nossa afirmação de que $\partial_2 F(g,x)$ é contínua.

Como ambas as derivadas parciais $\partial_1 F$ e $\partial_2 F$ são contínuas, segue-se que F é de classe C^1.

Note que $F(f,p) = 0$ e $\partial_2 F(f,p) = Df_p - I : E \to E$ é invertível porque 1 não pertence ao espectro de Df_p.

Pelo teorema da função implícita, para $\delta \geq \delta_0 > 0$ pequeno, 0 é valor regular de F, e dado $g \in B(f,\delta_0)$, existe um único $p = p_g \in B(p,\delta)$ tal que

$$F(g,p_g) = 0 \Rightarrow g(p_g) - p_g = 0 \Rightarrow p_g \text{ é ponto fixo de } g.$$

□

Observação 8.1.5. Note que não há grande perda em se supor no resultado acima f limitada e com derivada limitada em U, já que toda aplicação C^1 é localmente limitada (bem como sua derivada), ou seja, devido à continuidade de f e sua derivada sempre podemos restringir o aberto U a um aberto $\hat{U} \subset U$ contendo p onde $f|_{\hat{U}} : \hat{U} \to E$ satisfaz as hipóteses da proposição.

8.2 O teorema de Grobman-Hartman para campos

Lema 8.2.1. *Seja $X : V \subset \mathbb{R}^m \to \mathbb{R}^m$ um campo de vetores de classe C^k, $(k \geq 1)$ com $X(0) = 0$. Seja $L = (DX)_0$. Dado $\epsilon > 0$, existe um campo $Y : \mathbb{R}^m \to \mathbb{R}^m$ com as seguintes propriedades:*

1. *O campo Y tem constante de Lipschitz limitada por K e, portanto, o fluxo induzido por Y está definido em $\mathbb{R} \times \mathbb{R}^m$;*
2. *$Y = L$ fora de uma bola $B(0, r)$;*
3. *Existe um aberto $U \subset V$ contendo zero tal que $Y = X$ em U;*
4. *Escrevendo $Y_t = L_t + \phi_t$, existe $M > 0$ tal que $|\phi_t| \leq M$ para todo $t \in [-2, 2]$ e ϕ_1 tem constante de Lipschitz menor ou igual a ϵ.*

Prova: Como $L = (DX)_0$, temos que $X = L + \psi$, onde $\psi : V \to \mathbb{R}^m$ é C^k tal que $\psi(0) = 0$ e $D\psi_0 = 0$. Seja $\beta : \mathbb{R} \to \mathbb{R}$ uma função C^∞ tal que $\beta(\mathbb{R}) \subset [0, 1]$, $\beta(t) = 1$ se $t \leq r/2$ e $\beta(t) = 0$ se $t \geq r$. Seja $\Psi : \mathbb{R}^m \to \mathbb{R}^m$ definida por $\Psi(x) = \beta(|x|) \cdot \psi(x)$ se $x \in V$ e $\Psi(x) = 0$ se $x \in \mathbb{R}^m \setminus V$. Dado $\delta > 0$, pelo lema 8.0.3 podemos escolher $r > 0$ de tal forma que Ψ seja C^k e seja $\delta-$Lipschitz. Daí, da definição de Ψ, $\Psi = \psi$ em $B(0, r/2)$ e $\Psi \equiv 0$ fora de $B(0, r)$. Seja $Y : \mathbb{R}^m \to \mathbb{R}^m$ o campo de vetores definido por $Y := L + \Psi$. Daí, $Y = X$ em $B(0, r/2)$, $Y = L$ fora de $B(0, l)$ e Y satisfaz (1).

Só falta verificar (4). De fato, como consequência da desigualdade de Gronwall, já vimos (dependência Lipschitz em relação às condições iniciais) que:

$$|Y_t(x) - Y_t(y)| \leq e^{K|t|} \cdot |x - y| \underbrace{\leq}_{|t| \leq 2} e^{2K} \cdot |x - y|.$$

Seja $\phi_t := Y_t - L_t$; então

$$\phi_t(x) - \phi_t(y) = \int_0^t [\Psi(Y_s(x)) - \Psi(Y_s(y))]ds + \int_0^t L(\phi_s(x) - \phi_s(y))ds \Rightarrow$$

$$|\phi_t(x) - \phi_t(y)| \leq \delta \cdot e^{2K} \cdot |x - y| \cdot 2 + |\int_0^t L(\phi_s(x) - \phi_s(y))ds| \leq$$

$$\underbrace{\delta \cdot e^{2K} \cdot |x - y| \cdot 2}_{:=\alpha} + \int_0^t \underbrace{\|L\|}_{:=v(s)} \cdot \underbrace{|\phi_s(x) - \phi_s(y)|}_{:=u(s)} ds.$$

Usando a desigualdade de Gronwall, obtemos

$$|\phi_t(x) - \phi_t(y)| \leq \delta \cdot e^{2K} \cdot |x - y| \cdot 2 \cdot e^{\|L\| \cdot \int_0^t ds} \leq \delta \cdot e^{2K} \cdot |x - y| \cdot 2 \cdot e^{\|L\| \cdot 2}.$$

Pelo lema 8.0.3, podemos tomar Ψ de modo que sua constante de Lipschitz $\delta > 0$ seja menor ou igual a $\epsilon/(e^{2K} \cdot 2 \cdot e^{\|L\| \cdot 2})$. Tal implica, em particular, que ϕ_1 tem constante de Lipschitz $\leq \epsilon$. Finalmente, $|\phi_t|, t \in [-2, 2]$ é limitada: se $x \in B(0, r)$

$$|\phi_t(x)| = |\phi_t(x) - \phi_t(0)| \leq \epsilon \cdot r.$$

Se $x \notin B(0, r)$, então

$$|\phi_t(x)| = |\phi_t(x) - \phi_t(0)| = |\int_0^t [\Psi(Y_s(x)) - \Psi(Y_s(0))]ds + \int_0^t L(\phi_s(x) - \phi_s(0))ds| \leq$$

(Note que $\Psi(Y_s(x)) = 0$, se $Y_s(x) \notin B(0, r)$)

$$\int_0^t \epsilon \cdot r ds + |\int_0^t L(\phi_s(x) - \phi_s(0))ds| \leq 2 \cdot \epsilon \cdot r + \|L\| \cdot \int_0^t |\phi_s(x) - \phi_s(0)|ds,$$

o que implica novamente pela desigualdade de Gronwall que existe $M > 0$ tal que $|\phi_t(x)| \leq M$, $\forall x \in \mathbb{R}^m$ e $\forall t \in [-2, 2]$.

□

Definição 8.2.2. (Singularidade hiperbólica). Dado um campo de vetores C^k $X : U \to \mathbb{R}^m$, uma singularidade $p \in U$ de X é dita *hiperbólica* se a equação determinada pela sua parte linear $DX(p) \in \mathcal{L}(\mathbb{R}^m)$ é hiperbólica (isto é, se os autovalores de $DX(p)$ têm parte real não nula).

O próximo lema relaciona as singularidades hiperbólicas de campos e pontos fixos hiperbólicos.

Lema 8.2.3. *Seja $X : U \to \mathbb{R}^m$ um campo de vetores, e φ_t o seu fluxo. Então p é singularidade hiperbólica de $X \Leftrightarrow p$ é ponto fixo hiperbólico do difeomorfismo φ_1, tempo 1 de X.*

Prova: ($\Leftarrow$) Se p é ponto fixo do tempo 1 de X e é hiperbólico, em particular, pela proposição 8.1, é isolado. Note que p não pode pertencer a uma órbita periódica de período 1, pois em tal situação, os outros pontos da órbita periódica seriam pontos fixos para f, e p não seria ponto fixo isolado. Logo, como $\varphi(n, p) = p, \forall n \in \mathbb{N}$ e $\varphi(\cdot, p)$ não é periódica regular, segue-se

da classificação das trajetórias de um campo que $\varphi(t,p) = p, \forall t \in \mathbb{R} \Rightarrow p$ é singularidade (isolada) de X.

Mostremos que p é singularidade hiperbólica, ou seja, que os elementos de $Sp(DX(p))$ têm parte real não nula. Da dependência diferenciável em relação às condições iniciais, temos que $\partial_x\varphi$ é solução de $\dot{Z} = DX(p) \cdot Z$; $Z_0 = I$. Portanto, $\partial_x\varphi(t,p) = e^{t \cdot DX(p)}$, o que nos dá

$$Df_p = \partial_x\varphi(1,p) = e^{DX(p)},$$

e portanto o espectro $Sp(Df_p) = e^{Sp(DX(p))}$, implicando que $|\lambda| \neq 1, \forall \lambda \in Sp(Df_p)$.

($\Rightarrow$) Se p é singularidade hiperbólica de X, é imediato que p é ponto fixo de $f = \varphi_1$. Como vimos acima, da dependência diferenciável em relação às condições iniciais, temos que $\partial_x\varphi$ é solução de $\dot{Z} = DX(p) \cdot Z$; $Z_0 = I$. Portanto, $\partial_x\varphi(t,p) = e^{t \cdot DX(p)}$, o que nos dá

$$Df_p = \partial_x\varphi(1,p) = e^{DX(p)},$$

e portanto o espectro $Sp(Df_p) = e^{Sp(DX(p))}$, implicando que $|\lambda| \neq 1$, $\forall \lambda \in Sp(Df_p)$, isto é, p é ponto fixo hiperbólico de f.

□

Teorema 8.2.4. *(Grobman-Hartman para campos) Seja $X : V \subset \mathbb{R}^m \to \mathbb{R}^m$ um campo C^k ($k \geq 1$) e p uma singularidade hiperbólica de X. Seja $L = DX_p$. Então X é localmente (topologicamente) conjugado (via um homeomorfismo h) a L, em vizinhanças de p e zero.*

Prova: Seja $Y : \mathbb{R}^m \to \mathbb{R}^m$ um campo C^k como no lema 8.2.1. Como $Y = X$ em U, vizinhança de zero, temos que a aplicação identidade conjuga localmente Y e X em U. Como a conjugação é uma relação de equivalência, portanto transitiva, só nos resta mostrar que (os fluxos) Y_t e L_t são conjugados, $\forall t \in \mathbb{R}$. singularidade de Y, como $(DY)_0 = L$, da dependência diferenciável em relação às condições iniciais, temos que a derivada $(DY_1)_0$ do difeomorfismo Y_1 na origem é $e^L = L_1$. De fato, escrevendo $\varphi(t,x) = Y_t(x)$, temos que $(DY_t)(x) = \frac{\partial\varphi(t,x)}{\partial x}$ é solução de

$$\begin{cases} \dot{Z} = DY(\varphi(t,x)) \cdot Z \\ Z(0) = I_m \leftarrow \text{ matriz identidade } m \times m \end{cases}$$

Por ser $x = 0$ uma singularidade, $\varphi(t, 0) \equiv 0$, e a equação acima fica:

$$\begin{cases} \dot{Z} = DY(0) \cdot Z = L \cdot Z \\ Z(0) = I_m \leftarrow \text{ matriz identidade } m \times m, \end{cases}$$

o que implica que $(DY_t)(0) = e^{t \cdot L} \Rightarrow_{t=1} (DY_1)(0) = e^L = L_1$.

Logo, o difeomorfismo $Y_1 = L_1 + \phi_1$ tem a origem como ponto fixo hiperbólico e ϕ_1 como o resto de sua derivada (L_1) na origem. Pelo lema 8.1.3 do teorema de Grobman-Hartman para difeomorfismos, existe um único homeomorfismo $h : \mathbb{R}^m \to \mathbb{R}^m$ a uma distância finita da identidade que satisfaz $h \circ Y_1 = L_1 \circ h$. Mostraremos que este mesmo h também conjuga todos os outros tempos de Y_1 e L_1, isto é, que $h \circ Y_t(x) = L_t \circ h(x)$, $\forall t \in \mathbb{R}$, $\forall x \in \mathbb{R}^m$, o que significa, por definição, Y ser topologicamente conjugado a L.

Definimos $H : \mathbb{R}^m \to \mathbb{R}^m$ (que, no final, veremos ser igual a h) por

$$H(x) := \int_0^1 L_{-t} \circ h \circ Y_t(x) dt.$$

Da expressão acima, H é obviamente contínua e da condição (4) do lema 8.2.1, temos que H está a distância finita da identidade. Mostremos agora que para todo $s \in \mathbb{R}$ vale

$$H \circ Y_s = L_s \circ H.$$

Para vermos isso, afimamos que basta mostrarmos que H é homeomorfismo e que a expressão acima vale $\forall s \in [0, 1]$, pois dado $q \in \mathbb{R}^+$, podemos escrever $q = n + s$, com $n \in \mathbb{N}$ e $s \in [0, 1]$. Daí,

$$H \circ Y_q = H \circ \overbrace{Y_1 \circ Y_1 \circ \cdots \circ Y_1}^{n\times} \circ Y_s =$$

$$L_1 \circ H \circ \overbrace{Y_1 \circ \cdots \circ Y_1}^{n-1\times} \circ Y_s = L_{n+s} \circ H = L_q \circ H.$$

Se $q < 0$ (e H é inversível), então

$$H \circ Y_q = (Y_{-q} \circ H^{-1})^{-1} = (H^{-1} \circ L_{-q})^{-1} = L_q \circ H,$$

o que comprova nossa afirmação.

Seja portanto $s \in [0, 1]$. Temos:

$$L_{-s} \circ H \circ Y_s = L_{-s} \circ \Big(\int_0^1 L_{-t} \circ h \circ Y_t \, dt \Big) \circ Y_s =$$

$$\int_0^1 L_{-s} \circ L_{-t} \circ h \circ Y_t \circ Y_s \, dt = \int_0^1 L_{-(s+t)} \circ h \circ Y_{t+s} \, dt.$$

Note que para escrevermos a penúltima igualdade, precisamos usar a linearidade de L_{-s} (s fixado), e que se estivéssemos compondo Y_s à direita e não à esquerda, não o poderíamos colocar "para dentro" da integral. Usamos ainda a continuidade de Y_s. Já na última igualdade usamos somente a propriedade de grupo dos fluxos de Y e L.

Tomando $u = t + s - 1$, obtemos:

$$\int_0^1 L_{-(s+t)} \circ h \circ Y_{t+s} \, dt = \int_{-1+s}^{s} L_{-(u+1)} \circ h \circ Y_{u+1} \, du =$$

$$\int_{-1+s}^{0} L_{-u+1} \circ h \circ Y_{u+1} \, du + \int_0^s L_{-(u+1)} \circ h \circ Y_{u+1} \, du.$$

Fazendo $v = u + 1$ na primeira parcela vem que:

$$L_{-s} \circ H \circ Y_s = \int_s^1 L_{-v} \circ h \circ Y_v \, dv + \int_0^s L_{-u} \circ \underbrace{(L_{-1} \circ h \circ Y_1)}_{=h} \circ Y_u \, du =$$

$$\int_0^1 L_u \circ h \circ Y_u \, du = H.$$

Temos portanto que H é contínua, e semiconjuga Y_1 e L_1. H está a distância finita da identidade: dado $t \in [0, 1]$,

$$L_{-t} \circ h \circ Y_t = L_{-t} \circ (I + w) \circ (L_t + \phi_t) = I + \underbrace{L_{-t} \circ \phi_t + L_{-t} \circ w \circ L_t + L_{-t} \circ w \circ \phi_t}_{:= \tilde{w}_t, \exists \tilde{M} > 0 / |\tilde{w}_t(x)| \leq \tilde{M}, \forall t \in [0,1], \forall x \in \mathbb{R}^m} \Rightarrow$$

$$H(x) := \int_0^1 L_{-t} \circ h \circ Y_t(x) dt = \int_0^1 I + \tilde{w}_t dt = I + \int_0^1 \tilde{w}_t dt,$$

com $|\int_0^1 \tilde{w}_t(x) dt| \leq \int_0^1 \tilde{M} dt = \tilde{M}$.

Da unicidade da tese do lema 8.1.3, segue-se que $H = h$, que é homeomorfismo. Como vimos acima, isto implica que h conjuga Y_s e L_s, $\forall s \in \mathbb{R}$.

$\square$

8.3 Apêndice: Classificação dos isomorfismos hiperbólicos

O teorema de Grobman-Hartman reduz o problema de classificar as conjugações locais de difeomorfismos em torno de pontos fixos hiperbólicos e de campos em torno de singularidades hiperbólicas ao de classificar as conjugações, respectivamente, de isomorfismos hiperbólicos e de campos lineares com singularidades hiperbólicas. Portanto, nesta seção, classificaremos os isomorfismos lineares de $\mathbb{R}^m$ segundo suas classes de conjugação via homeomorfismo. Na seção seguinte, obteremos resultados análogos para campos lineares.

Proposição 8.3.1. *Seja $A \in \mathcal{L}(\mathbb{R}^m)$ um isomorfismo linear hiperbólico. Existe $\epsilon > 0$ tal que, se $B \in \mathcal{L}(\mathbb{R}^m)$ e satisfaz $\|A - B\| < \epsilon$, então B é conjugado a A.*

Prova:

1. Pelo teorema de Grobman-Hartman para difeomorfismos, B é localmente conjugado a A, isto é, existe um homeomorfismo h tal que $h \circ A = B \circ h$ em vizinhanças (V e $\tilde{V}$) de 0.

2. Considere os espaços vetoriais estáveis E^s e $\tilde{E}^s$ respectivamente de A e B e também espaços instáveis E^u e $\tilde{E}^u$ dados por:

$$\begin{array}{ll} E^s = \oplus_{|\lambda_i|<1} E(\lambda_i), & \lambda_i \text{ é autovalor de } A; |\lambda_i| < 1 \\ \tilde{E}^s = \oplus_{|\tilde{\lambda}_i|<1} E(\tilde{\lambda}_i), & \tilde{\lambda}_i \text{ é autovalor de B}; |\tilde{\lambda}_i| < 1 \\ E^u = \oplus_{|\lambda_j|>1} E(\lambda_j), & \lambda_j \text{ é autovalor de } A; |\lambda_j| > 1 \\ \tilde{E}^u = \oplus_{|\tilde{\lambda}_j|>1} E(\tilde{\lambda}'_j), & \tilde{\lambda}_j \text{ é autovalor de } B; |\tilde{\lambda}_j| > 1. \end{array}$$

   Note que por serem definidos como soma direta de autoespaços generalizados, E^s e E^u são invariantes por A; $\tilde{E}^s$ e $\tilde{E}^u$ são invariantes por B.

   Como A e B estão próximos e os autovalores variam continuamente com a matriz, tomando $\epsilon > 0$ pequeno podemos supor $\dim E^s = \dim \tilde{E}^s$, $\dim E^u = \dim \tilde{E}^u$.

   Se $A|_V$ é conjugado a $B|_{\tilde{V}}$, como vimos, definamos

$$\begin{array}{ll} V^s := V \cap E^s; & \tilde{V}^s = \tilde{V} \cap \tilde{E}^s \\ V^u := V \cap E^u; & \tilde{V}^u = \tilde{V} \cap \tilde{E}^u \end{array}$$

3. Mostremos que

$$\begin{aligned} x \in E^s &\Leftrightarrow \quad A^n(x) \to 0, \text{ se } n \to \infty \\ x \in E^u &\Leftrightarrow \quad A^{-n}(x) \to 0, \text{ se } n \to \infty. \end{aligned}$$

De fato, basta mostrar que vale o comportamento acima para $x \in E(\lambda_i) \subset E^s$. Para tal x vale que $A^n(x) = ((A-\lambda_i I)|_{E(\lambda_i)} + (\lambda_i I)|_{E(\lambda_i)})^n(x)$. Sendo k a nulidade de $(A-\lambda_i)|_{E(\lambda_i)}$, temos que para $n > k$

$$\|A^n|_{E(\lambda_i)}\| = \|\lambda_i^{n-k+1} \cdot ((\begin{matrix} n \\ n-k+1 \end{matrix})(A-\lambda_i I)^{k-1} + \cdots + \lambda_i^{k-1} I)\| \leq$$

$$\lambda_i^{n-k+1} \cdot n \cdot \cdots \cdot (n-k) \cdot k \cdot \max_{l=1\ldots k}\{\|(A-\lambda_i I)^l\|, \|\lambda_i I\|\}.$$

Como $\lambda_i < 1$ e a exponencial domina qualquer polinômio, a última expressão converge a zero quando $n \to +\infty$. (Mutatis Mutandis para E^u).

Para ver que se $A^n(x)$ converge a zero quando $n \to +\infty$, então $x \in E^s$, basta escrever $x = x_s + x_u$, com $x_s \in E^s$ e $x_u \in E^u$ e observar que $A^n(x_u) \to \infty$, quando $n \to +\infty$. Esta última observação é verdadeira porque se $x_u \in E^u$, então x_u pertence ao espaço estável de A^{-1}. Como vimos, isto implica que $A^{-n}(x_u) = (A^{-1})^n(x_u) \to 0$, ou ainda, que $\|A^{-n}|_{E^u}\| \to 0$, quando $n \to +\infty$. Daí,

$$|x_u| = |A^{-n} \cdot A^n(x_u)| \leq \|A^{-n}_{E^u}\| \cdot |A^n(x_u)| \Rightarrow$$

$$|A^n(x_u)| \geq \frac{|x_u|}{\|A^{-n}_{E^u}\|} \to +\infty, \text{ quando } n \to +\infty,$$

logo se $A^n(x) \to 0, n \to +\infty$, com $x = x_s + x_u$ como acima, então $x_u = 0$.

4. Note que o último item implica que existe $n_0 \in \mathbb{N}$ tal que $A^n|_{E^s}$ é contração (ajustando n_0, é contração tão forte quanto se queira) para todo $n > n_0$. Em particular, podemos tomar n_0 tal que $A^n(V \cap E^s) = A^n(V^s) \subset V^s$, $\forall n \geq n_0$. Chamando $W^s := A^{n_0}(V^s)$ temos que $A^j(W^s) \subset V^s$, $\forall j \in \mathbb{N}$. Definimos ainda $\tilde{W}^s := h(W^s)$. Note que se $z \in \tilde{W}^s$, existe $y \in \tilde{V}^s$ tal que $h(y) = z$, logo

$$h \circ \underbrace{A^j(y)}_{\in V} = h \circ A \circ A^{j-1}(y) = B \circ hA^{j-1}(y) = B^j \circ h(y) \in \tilde{V},$$

isto é $B^j(h(y)) = B^j(z) \in, \forall j \in \mathbb{N}$, se $z \in \tilde{W}$. Ademais, fazendo $j \to \infty$, vemos que $B^j(z) \to 0$, logo $\tilde{W}^s \subset \tilde{E}^s$.

5. Note que
$$h(0) = h \circ A(0) = B \cdot h(0) \Rightarrow h(0) = 0.$$
Se $x \in W^s$, então
$$\begin{array}{ll} h \circ A^n(x) & \to h(0) = 0 \\ \quad \| & \\ B^n \cdot h(x) & \Rightarrow B^n \cdot h(x) \to 0 \quad \Rightarrow \quad h(x) \in \tilde{W}^s \subset \tilde{E}^s. \end{array}$$

6. Definimos uma conjugação h^s de $A|_{E^s}$ e $B|_{\tilde{E}^s}$ dada por:
$$h^s(x) := B^{-n_s(x)} \circ h \circ A^{n_s(x)} \cdot x,$$
onde $n_s(x)$ é o primeiro iterado não negativo de $x \in E^s$ tal que $A^{n_s(x)} \in W^s$.

   Tanto faz $n_s(x)$ ser o "primeiro" ou não, pois h^s é o mesmo se tomamos qualquer $n_1(x) > n_s(x)$ na sua definição. Realmente, seja $H^s(x) := B^{-n_1(x)} \circ h \circ A^{n_1(x)} \cdot x$, com $n_1(x) > n_s(x)$. Então
$$H^s(x) = B^{-n_s(x)} \circ \underbrace{B^{-n_1(x)+n_s(x)} \circ h \circ A^{n_1(x)-n_s(x)}}_{=h} \circ A^{n_s(x)} =$$
$$B^{-n_s(x)} \circ h \circ A^{n_s(x)} = h^s(x).$$

   Vejamos que h^s está bem definida e é contínua. Se $x \in E^s$ é tal que $n_s(x) = n_x$, então como A^{n_x} é uma aplicação contínua, existe uma vizinhança V_x de x em E^s tal que $A^{n_x}(V_x) \subset W^s$. Isto quer dizer que $n_s(y) \geq n_x, \forall y \in V_x$. Do parágrafo anterior, temos que
$$h^s(y) = B^{-n_x} \circ h \circ A^{n_x}(y), \forall y \in V_x,$$
   o que implica que $h^s|_{V_x}$ é contínua e portanto, h^s é contínua. Note que o acima mostra ainda que h^s é localmente aberta, portanto aberta, pois $h^s|_{V_x}$ é expressa como composta de aplicações abertas.

7. Mostremos que h^s é injetiva. De fato, se $B^{-n_s(x)} \circ h \circ A^{n_s(x)}(x) = B^{-n_s(y)} \circ h \circ A^{n_s(y)}(y)$, supondo sem perda de generalidade, $n_s(y) \geq n_s(x)$, temos que
$$B^{-n_s(x)} \circ h \circ A^{n_s(x)}(x) = B^{-n_s(y)} \circ h \circ A^{n_s(y)}(x) = B^{-n_s(y)} \circ h \circ A^{n_s(y)}(y).$$
   Aplicando à esquerda $A^{-n_s(y)} \circ h^{-1} \circ B^{n_s(y)}$ em ambos os termos da última igualdade, obtemos $x = y$.

8. $h^s : E^s \to \tilde{E}^s$ é sobrejetiva. De fato, basta definir $\tilde{h}^s : \tilde{E}^s \to E^s$ por

$$\tilde{h}^s(z) = A^{-\tilde{n}_s(z)} \circ h^{-1} \circ B^{\tilde{n}_s(z)} \cdot z,$$

onde $\tilde{n}_s(z)$ é o primeiro natural tal que $B^{\tilde{n}_s(z)} \cdot z \in \tilde{W}^s$. Como $\tilde{h}^s$ é definida de maneira análoga a h^s, segue-se que $\tilde{h}^s$ é contínua e injetiva. Mostremos que $I_{E^s} = \tilde{h}^s \circ h^s$ e que $I_{\tilde{E}^s} = h^s \circ \tilde{h}^s$. De fato, se tomamos por exemplo

$$\tilde{h}^s \circ h^s = A^{-\tilde{n}_s(z)} \circ h^{-1} \circ B^{\tilde{n}_s(z)} \circ \underbrace{B^{-n_s(x)} \circ h \circ A^{n_s(x)} \cdot x}_{z},$$

considerando $\hat{n} = \max\{n_s(x), \tilde{n}_s(z)\}$, temos que a expressão acima é o mesmo que

$$A^{-\hat{n}} \circ h^{-1} \circ B^{\hat{n}} \circ B^{-\hat{n}} \circ h \circ A^{\hat{n}}(x) \cdot x = x.$$

Similarmente, mostra-se que $h^s \circ \tilde{h}^s(z) = z, \forall z \in \tilde{E}^s$.

9. h^s conjuga $A|_{E^s}$ e $B|_{\tilde{E}^s}$. Realmente, se $x \in E^s \setminus W^s$ (o caso $x \in W^s$ é trivial), temos que $n_s(x) \geq n_s(A(x))$, logo

$$h^s \circ A|_{E^s}(x) = B^{-n_s(A(x))} \circ h \circ A^{n_s(A(x))} \circ A(x) = B^{-n_s(x)} \circ h \circ A^{n_s(x)} \circ A(x) =$$

$$B^{-n_s(x)} \circ h \circ A \circ \underbrace{A^{n_s(x)}(x)}_{\in W^s \subset V^s} = B^{-n_s(x)} \circ B \circ h \circ A^{n_s(x)}(x) = B \circ h^s(x).$$

10. Definindo $h^u : E^u \to \tilde{E}^u$ de maneira simétrica a h^s como

$$h^u(x) = B^{n_u(x)} \circ h \circ A^{-n_u(x)} \cdot x,$$

onde $-n_u(x)$ é o primeiro iterado negativo tal que $A^{-n_u(x)} \in W^u \subset V^u \subset E^u$, sendo W^u definido de modo análogo a W^s.

Fazendo para h^u o mesmo que fizemos para h^s, temos que h^u conjuga $A|_{E^u}$ e $B|_{\tilde{E}^u}$.

Estendemos h^s e h^u ao $\mathbb{R}^m$ da seguinte forma, dotando as extensões com a mesma notação:

$$h^s(x) = h^s \circ \pi_s(x);$$
$$h^u(x) = h^u \circ \pi_u(x),$$

onde $\pi_u : \mathbb{R}^m \to E^u$ e $\pi_s : \mathbb{R}^m \to E^s$ são as projeções naturais associadas à decomposição $\mathbb{R}^m = E^s \oplus E^u$.

Seja portanto $g : \mathbb{R}^m \to \mathbb{R}^m$ definida por

$$g(x) := h^s(x) + h^u(x).$$

É fácil ver que g é um homeomorfismo. Daí, dado $x = x_s + x_u \in \mathbb{R}^m$, com $x_s \in E^s$ e $x_u \in E^u$, temos:

$$g \circ A(x) = h^s \circ A(x) + h^u \circ A(x) = h^s \circ A(x_s + x_u) + h^u \circ A(x_s + x_u) =$$

$$h^s \circ \underbrace{A(x_s)}_{\in E^s} + h^s \circ \underbrace{A(x_u)}_{\in E^u} + h^u \circ \underbrace{A(x_s)}_{\in E^s} + h^u \circ \underbrace{A(x_u)}_{\in E^u} =$$

$$h^s \circ A(x_s) + h^u \circ A(x_u) = B \circ h^s(x_s) + B \circ h^u(x_u) = B(g(x)),$$

donde concluímos que g conjuga A e B em $\mathbb{R}^m$.

□

Corolário 8.3.2. *Se A e B são isomorfismos hiperbólicos de $\mathbb{R}^m$, então A e B são conjugados se e só se $A|_E^s$ é conjugado a $B|_{\tilde{E}^s}$ e $A|_E^u$ é conjugado a $B|_{\tilde{E}^u}$.*

Prova: Deixamos como exercício a prova deste corolário.

□

Lema 8.3.3. *Se dois isomorfismos lineares hiperbólicos A_0 e A_1 estão na mesma componente conexa do conjunto dos isomorfismos hiperbólicos (que é um aberto de $\mathcal{L}(\mathbb{R}^m)$), então A_0 e A_1 são (topologicamente) conjugados.*

Prova: De fato, uma tal componente conexa é aberta, portanto, é conexa por caminhos. Seja $\gamma : [0,1] \to \mathcal{L}(\mathbb{R}^m)$ um caminho contínuo com $\gamma(0) = A_0$, $\gamma(1) = A_1$ e com $\gamma(t)$ sendo isomorfismo hiperbólico para todo $t \in [0,1]$. Como $\gamma([0,1])$ é compacto, existe uma cobertura finita $\mathcal{B} = \cup_{j=1}^{l} B_j$ de bolas abertas B_j contidas no conjunto dos isomorfismos lineares hiperbólicos, onde dois quaisquer isomorfismos contidos em uma mesma bola pertencem a mesma classe de conjugação topológica, segundo a proposição 8.3.1. Seja ϵ_0 o número de Lebesgue da cobertura $\mathcal{B}$; isto é se dois elementos distam menos que ϵ_0, então eles pertencem a uma mesma bola da cobertura. Como $[0,1]$

é compacto, γ é uniformemente contínuo, o que implica que existe $\delta > 0$ tal que

$$|t - s| < \delta \Rightarrow \|\gamma(t) - \gamma(s)\| < \epsilon_0, \forall t, s \in [0, 1].$$

Dividamos portanto o intervalo $[0, 1]$ em um número finito de intervalos $[t_i, t_{i+1}], i = 0 \dots k$, com $t_0 = 0, t_k = 1$ e $|t_{i+1} - t_i| < \delta$. Segue-se que $\gamma(t_{i+1})$ é conjugado a $\gamma(t_i)$ e daí, pela transitividade da conjugação, $A_0 = \gamma(t_0)$ é conjugado a $\gamma(t_k) = A_1$.

□

Definição 8.3.4. (Índice de um isomorfismo linear). O *índice* de um isomorfismo linear $A \in \mathcal{L}(\mathbb{R}^m)$ é a dimensão do espaço estável de A. Tal *espaço estável* é a soma dos autoespaços generalizados com autovalores contrativos (de módulo menor que 1).

Proposição 8.3.5. *Sejam A_1 e A_2 isomorfismos de $\mathbb{R}^m$ com índice m que na base canônica são representados pelas seguintes matrizes:*

$$A_1 = \begin{pmatrix} 1/2 & & 0 \\ & \ddots & \\ 0 & & 1/2 \end{pmatrix}; \quad A_2 = \begin{pmatrix} -1/2 & & & 0 \\ & 1/2 & & \\ & & \ddots & \\ 0 & & & 1/2 \end{pmatrix}.$$

Se $A \in \mathcal{L}(\mathbb{R}^m)$ tem índice m e $\det(A) > 0$, então A é conjugado a A_1. Se A tem índice m e $det(A) < 0$, então A é conjugado a A_2.

Prova: Temos que $A = P \cdot J \cdot P^{-1}$, onde J está na forma de Jordan (real). Logo A é topologicamente (de fato, $C^\infty-$, linearmente conjugado) conjugado a J. Como a conjugação topológica é uma relação de equivalência, isso reduz o nosso problema a mostrar que J é conjugado topologicamente a A_1 (se $\det(A) = \det(J) > 0$) ou a A_2 (se $\det(A) = \det(J) < 0$). Nossa estratégia será construir um caminho contínuo com imagem contida no conjunto dos isomorfismos hiperbólicos unindo J ao isomorfismo que lhe corresponde (A_1 ou A_2). Pelo lema 8.3.3, J e seu correspondente serão conjugados. Escrevamos

portanto

$$J = \begin{pmatrix} \lambda_1 & 1 \text{ ou } 0 & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \lambda_{s'} & & & & & & \\ & & & \mu_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \mu_{s''} & & & \\ & & & & & & B_1 & C_1 & \\ & & & & & & & \ddots & \ddots \\ 0 & & & & & & & & B_{s'''} \end{pmatrix},$$

$$\begin{array}{cc} -1 < \lambda_j < 0, & j = 1 \dots s' \\ 0 < \mu_j < 1, & j = 1 \dots s'' \\ B_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, & j = 1 \dots s''' \\ \text{com } b_j \neq 0, & a_j^2 + b_j^2 < 1 \\ C_j = \begin{pmatrix} c_j & 0 \\ 0 & c_j \end{pmatrix}, \text{ com} & c_j = 0 \text{ ou } 1. \end{array}$$

Denominemos $\mathcal{H} \subset \mathcal{L}(\mathbb{R}^m)$ o conjunto dos isomorfismos hiperbólicos. Seja $\beta : [0,1] \to \mathcal{H}$ o caminho contínuo dado por:

$$\beta(t) = \begin{pmatrix} \lambda_1 & 1-t \text{ ou } 0 & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \lambda_{s'} & & & & & & \\ & & & \mu_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \mu_{s''} & & & \\ & & & & & & B_1 & C_{1,t} & \\ & & & & & & & \ddots & \ddots \\ 0 & & & & & & & & B_{s'''} \end{pmatrix},$$

$$\begin{array}{cc}
-1 < \lambda_j < 0, & j = 1 \dots s' \\
0 < \mu_j < 1, & j = 1 \dots s'' \\
B_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, & j = 1 \dots s''' \\
\text{com } b_j \neq 0, & a_j^2 + b_j^2 < 1 \\
C_{j,t} = \begin{pmatrix} (1-t)\cdot c_j & 0 \\ 0 & (1-t)\cdot c_j \end{pmatrix}, \text{ com} & c_j = 0 \text{ ou } 1,
\end{array}$$

portanto, J é equivalente à sua parte diagonal (eliminamos a parte nilpotente).

Seja agora $\gamma : [0,1] \to \mathcal{H}$, dada por

$$\gamma(t) = \begin{pmatrix}
(1-t)\lambda_1 + \frac{-t}{2} & & & & & & & 0 \\
& \ddots & & & & & & \\
& & (1-t)\lambda_{s'} + \frac{-t}{2} & & & & & \\
& & & (1-t)\mu_1 + \frac{t}{2} & & & & \\
& & & & \ddots & & & \\
& & & & & (1-t)\mu_{s''} + \frac{t}{2} & & \\
& & & & & & B_1 & \\
& & & & & & & \ddots \\
0 & & & & & & & B_{s'''}
\end{pmatrix},$$

$$\begin{array}{cc}
-1 < \lambda_j < 0, & j = 1 \dots s' \\
0 < \mu_j < 1, & j = 1 \dots s'' \\
B_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, & j = 1 \dots s''' \\
\text{com } b_j \neq 0, & a_j^2 + b_j^2 < 1,
\end{array}$$

o que implica que J é topologicamente conjugado (justapondo-se os caminhos β e γ) a:

$$\begin{pmatrix} -1/2 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & -1/2 & & & & & & & \\ & & & 1/2 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 1/2 & & & & \\ & & & & & & B_1 & & & \\ & & & & & & & \ddots & & \\ 0 & & & & & & & & & B_{s'''} \end{pmatrix}.$$

Defina agora o caminho $\omega : [0,1] \to \mathcal{H}$ dado por

$$\omega(t) = \begin{pmatrix} -1/2 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & -1/2 & & & & & & & \\ & & & 1/2 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 1/2 & & & & \\ & & & & & & B_{1,t} & & & \\ & & & & & & & \ddots & & \\ 0 & & & & & & & & & B_{s''',t} \end{pmatrix},$$

onde cada $B_{j,t}$ é definida para $t \in [0, 1/2]$ como

$$B_{j,t} := \begin{pmatrix} \cos(\theta_j t) & \sin(\theta_j t) \\ -\sin(\theta_j t) & \cos(\theta_j t) \end{pmatrix} \cdot \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix},$$

com $\cos(\theta_j/2) = a_j/(\sqrt{a_j^2 + b_j^2})$ e $\sin(\theta_j/2) = -b_j/(\sqrt{a_j^2 + b_j^2})$. Para $t \in [1/2, 1]$ definimos

$$B_{j,t} := \begin{pmatrix} 1/2(2t-1) + 2(1-t)\sqrt{a_j^2 + b_j^2} & 0 \\ 0 & 1/2(2t-1) + 2(1-t)\sqrt{a_j^2 + b_j^2} \end{pmatrix}.$$

Daí, justapondo os caminhos β, γ e ω, temos novamente pelo lema 8.3.3 que J é topologicamente conjugado a

$$\begin{pmatrix} -1/2 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & -1/2 & & & & & & & \\ & & & 1/2 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 1/2 & & & & \\ & & & & & & 1/2 & & & \\ & & & & & & & \ddots & & \\ 0 & & & & & & & & & 1/2 \end{pmatrix},$$

onde as primeiras s' linhas possuem $-1/2$ em sua única componente não nula e as demais linhas possuem $1/2$ em sua única componente não nula. Note que $\det(J) < 0$, s' é ímpar, caso contrário, s' é par. Sem perda de generalidade, vamos supor que $\det(J) < 0$ (o outro caso é análogo). A última curva $\sigma : [0,1] \to \mathcal{H}$ fica então:

$$\sigma(t) := \begin{pmatrix} -1/2 & & & & & & & & & 0 \\ & D_{1,t} & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & D_{(s'-1)/2,t} & & & & & & \\ & & & & 1/2 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1/2 & & & \\ & & & & & & & \ddots & & \\ 0 & & & & & & & & & 1/2 \end{pmatrix},$$

onde

$$D_{j,t} := \begin{pmatrix} \cos(\pi t) & \sin(\pi t) \\ -\sin(\pi t) & \cos(\pi t) \end{pmatrix} \cdot \frac{-1}{2},$$

o que conclui a prova. $\square$

Corolário 8.3.6. *Sejam $\hat{A}_1$ e $\hat{A}_2$ isomorfismos hiperbólicos de $\mathbb{R}^m$ com índice*

0 *que na base canônica são representados pelas seguintes matrizes:*

$$\hat{A}_1 = \begin{pmatrix} 2 & & 0 \\ & \ddots & \\ 0 & & 2 \end{pmatrix}; \quad \hat{A}_2 = \begin{pmatrix} -2 & & & 0 \\ & 2 & & \\ & & \ddots & \\ 0 & & & 2 \end{pmatrix}.$$

Se $\hat{A} \in \mathcal{L}(\mathbb{R}^m)$ *é isomorfismo hiperbólico com índice* 0 *e* $\det(\hat{A}) > 0$*, então* $\hat{A}$ *é conjugado a* $\hat{A}_1$*. Se* $\hat{A}$ *é isomorfismo hiperbólico de índice* 0 *e* $det(\hat{A}) < 0$*, então* $\hat{A}$ *é conjugado a* $\hat{A}_2$*.*

Prova: Inteiramente análoga à proposição anterior.

□

Teorema 8.3.7. *Seja* A *um isomorfismo linear hiperbólico qualquer de índice* s*. Então* A *é topologicamente conjugado a um dos seguintes isomorfismos hiperbólicos:*

$$A_{1,1} := \begin{pmatrix} 1/2 & & 0 & & & \\ & \ddots & & & & \\ 0 & & 1/2 & & & \\ & & & 2 & & 0 \\ & & & & \ddots & \\ & & & 0 & & 2 \end{pmatrix},$$

$$A_{1,2} := \begin{pmatrix} 1/2 & & 0 & & & \\ & \ddots & & & & \\ 0 & & 1/2 & & & \\ & & & -2 & & 0 \\ & & & & 2 & \\ & & & & & \ddots \\ & & & 0 & & 2 \end{pmatrix},$$

$$A_{2,1} := \begin{pmatrix} -1/2 & & & 0 & & \\ & 1/2 & & & & \\ & & \ddots & & & \\ 0 & & & 1/2 & & \\ & & & & 2 & 0 \\ & & & & & \ddots \\ & & & & 0 & 2 \end{pmatrix},$$

$$A_{2,2} := \begin{pmatrix} -1/2 & & & 0 & & & & \\ & 1/2 & & & & & & \\ & & \ddots & & & & & \\ 0 & & & 1/2 & & & & \\ & & & & -2 & & & 0 \\ & & & & & 2 & & \\ & & & & & & \ddots & \\ & & & & 0 & & & 2 \end{pmatrix}.$$

Prova: A exemplo da demonstração da proposição 8.3.5, não há perda de generalidade em supor que A esteja na forma de Jordan. Nesse caso, os espaços estáveis de A e dos isomorfismos do enunciado coincidem. O mesmo vale para os espaços instáveis. Da proposição 8.3.5, aplicada a $A|_{E^s} : E^s \to E^s$, obtemos que existe $h^s : E^s \to E^s$ um homeomorfismo conjugando $A|_{E^s}$ e $A_{j,l}|_{E^s}$, sendo $A_{j,l}$ um isomorfismo dos acima que inverte ou mantém a orientação nos espaços estável e instável da mesma maneira que A o faz. Usando o corolário 8.3.6 em $A|_{E^u} : E^u \to E^u$, obtemos $h^u : E^u \to E^u$ conjugando $A|_{E^u}$ e $A_{j,l}|_{E^u}$. Do mesmo modo que na proposição 8.3.1, estendemos h^s e h^u ao $\mathbb{R}^m$ da seguinte forma, dotando as extensões com a mesma notação:

$$\begin{aligned} h^s(x) &= h^s \circ \pi_s(x); \\ h^u(x) &= h^u \circ \pi_u(x), \end{aligned}$$

onde $\pi_u : \mathbb{R}^m \to E^u$ e $\pi_s : \mathbb{R}^m \to E^s$ são as projeções naturais associadas à decomposição $\mathbb{R}^m = E^s \oplus E^u$.

Seja portanto $h : \mathbb{R}^m \to \mathbb{R}^m$ definida por

$$h(x) := h^s(x) + h^u(x).$$

É fácil ver que h é um homeomorfismo. Daí, dado $x = x_s + x_u \in \mathbb{R}^m$, com $x_s \in E^s$ e $x_u \in E^u$, temos:

$$h \circ A(x) = h^s \circ A(x) + h^u \circ A(x) = h^s \circ A(x_s + x_u) + h^u \circ A(x_s + x_u) =$$

$$h^s \circ \underbrace{A(x_s)}_{\in E^s} + h^s \circ \underbrace{A(x_u)}_{\in E^u} + h^u \circ \underbrace{A(x_s)}_{\in E^s} + h^u \circ \underbrace{A(x_u)}_{\in E^u} =$$

$$h^s \circ A(x_s) + h^u \circ A(x_u) = A_{j,l} \circ h^s(x_s) + A_{j,l} \circ h^u(x_u) = A_{j,l}(h(x)),$$

donde concluímos que h conjuga A e $A_{j,l}$ em $\mathbb{R}^m$.

$\square$

Corolário 8.3.8. *Seja $f : U \to \begin{smallmatrix}\mathbb{R}^m \\ \mathcal{M}\end{smallmatrix}$ um difeomorfismo C^1 de um aberto U de $\mathbb{R}^m$ (ou de uma variedade $m-$dimensional $\mathcal{M}$). Então, existe uma vizinhança W de f em $Diff^1(U)$ tal que todo $g \in W$ possui ponto fixo hiperbólico p_g tal que g é conjugado a f em uma vizinhança de f.*

Prova: Consequência imediata do teorema de Grobman-Hartman para difeomorfismos e o último teorema. $\square$

Capítulo 9

O Teorema da Variedade Estável

Durante esse capítulo, consideraremos E um espaço de Banach.

Vimos no capítulo anterior que se $f : W \subset E \to E$ é um difeomorfismo com $p \in W$ como ponto fixo hiperbólico, então f é localmente conjugado a $A = Df_p : E \to E$ em vizinhanças de 0 e p. Em particular, vimos que existe um homeomorfismo $h : U(0) \subset W \to V(p)$, tal que $h(0) = p$ e que

$$h(A(x)) = f(h(x)), \text{ para } x \in A^{-1}(U(0)) \cap B(0, \beta).$$

Vimos ainda que existe um aberto de E^s denotado por $\tilde{U} \subset (U \cap E^s)$ tal que $A^m(\tilde{U}) \in U$, $\forall m \in \mathbb{N}$. Se consideramos a variedade topológica $\tilde{V} = h(\tilde{U})$, então para $y \in \tilde{V}$ tal que $y = h(x)$ podemos escrever

$$f^m(y) = f^m(h(x)) = h(A^m(x)) \to h(0) = p, \text{ quando } m \to +\infty,$$

ou seja, $f^m(y) \to p$ quando $m \to +\infty$. Portanto, concluímos que se $z \in \cup_{j\in\mathbb{N}} f^{-j}(\tilde{V})$, então $f^m(z) \to p$ quando $m \to +\infty$. Além do mais, se $z \in E$ é tal que $f^m(z) \to p$, então para m grande vale $h^{-1}(f^m(z)) \to 0$ quando $m \to \infty$, o que implica que $h^{-1}(f^m(z)) \in \tilde{U}$ para m grande, ou seja que $z \in \cup_{j\in\mathbb{N}} f^{-j}(\tilde{V})$. Em resumo,

$$f^m(z) \to p \text{ quando } m \to +\infty \Leftrightarrow z \in \cup_{j\in\mathbb{N}} f^{-j}(\tilde{V}).$$

Note que $\cup_{j\in\mathbb{N}} f^{-j}(\tilde{V})$ é uma variedade topológica. Em outras palavras, o teorema de Grobman-Hartman tem como consequência que o conjunto dos

pontos de E cujo $\omega-$limite é o ponto fixo hiperbólico p constitui uma variedade topológica, a chamada *variedade estável* de p.

No presente capítulo, daremos uma prova da existência da variedade estável que independe inteiramente do teorema de Grobman-Hartman. Ademais, provaremos que a variedade estável é, de fato, uma variedade diferenciável, da mesma classe de diferenciabilidade que o difeomorfismo f. Finalmente, demonstraremos resultados análogos para singularidades hiperbólicas de campos.

9.1 O Teorema da Variedade Estável para difeomorfismos

Comecemos com algumas definições.

Definição 9.1.1. (Conjunto estável de um ponto). Seja $f : W \subset M \to M$ um homeomorfismo de um subconjunto aberto W de um espaço métrico M (dotado da métrica $d(\cdot,\cdot)$). Dado $p \in W$, o *conjunto estável* de p é definido como

$$W^s(p) := \{x \in W, d(f^n(x), f^n(p)) \to 0, \text{ quando } n \to +\infty\}.$$

Se p é um ponto fixo de f, então seu conjunto estável é constituído dos pontos $x \in W$ tais que $f^n(x) \to p$, quando $n \to +\infty$.

Definição 9.1.2. Seja $p \in E$ e seja $f : U \subset E \to V$ um difeomorfismo, onde U, V são subconjuntos abertos de um espaço de Banach E.

Fixada uma vizinhança B da órbita de p, com $B \subset U \cap V$, definimos o conjunto estável local (a depender de B) de p como

$$W^s_{loc}(p) := \{q \in B, f^n(q) \in B, \forall n \geq 0 \text{ e } d(f^n(q), f^n(p)) \to 0, \text{ se } n \to +\infty\}$$

Analogamente, definimos o conjunto instável local (a depender de B) de p

$$W^u_{loc}(p) := \{q \in B, f^{-n}(q) \in B, \forall n \geq 0 \text{ e } d(f^{-n}(q), f^{-n}(p)) \to 0 \text{ se } n \to +\infty\}$$

Definição 9.1.3. (Conjuntos maximais invariantes de uma vizinhança). conjunto maximal positivamente invariante de B por: Seja $f : U \to V$ um difeomorfismo e $B \subset U$ uma vizinhança. O *conjunto maximal negativamente invariante em B* é definido por:

$$\Lambda^s(B) := \cap_{n=0}^{+\infty} f^n(B);$$

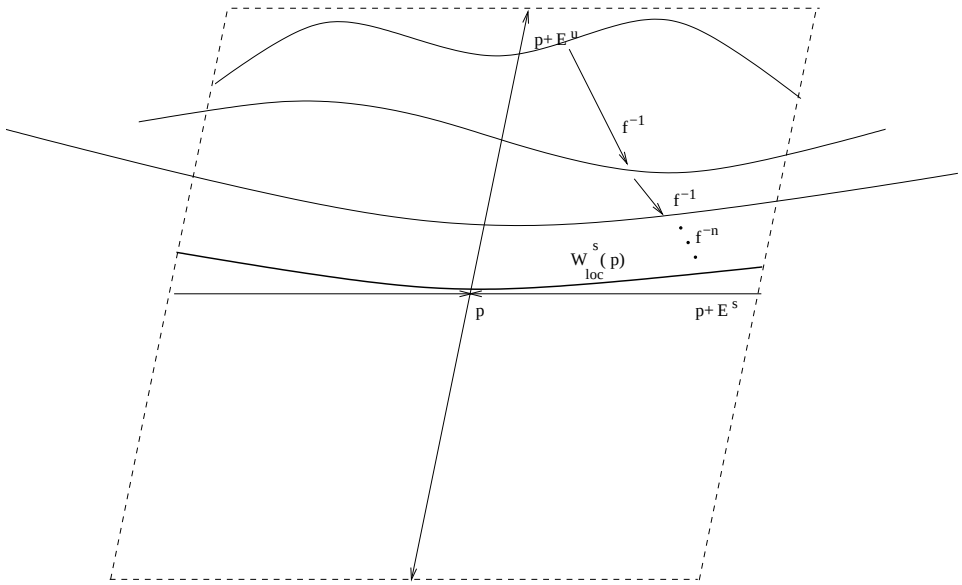


Figura 9.1: Fazendo a dinâmica trabalhar por nós: a iteração por f^{-1} de variedades próximas à variedade estável converge para a variedade estável de f.

ou seja, Λ^s é formado por aqueles pontos pertencentes a B tais que todas as suas pré-imagens também estão em B.

Simetricamente o *conjunto maximal positivamente invariante em* B é definido como

$$\Lambda^u(B) := \cap_{n=0}^{+\infty} f^{-n}(B).$$

Lembramos que se p é um ponto fixo hiperbólico, então existe $E = E^s \oplus E^u$ uma única decomposição em soma direta do espaço E em espaços $Df(p)-$invariantes E^s e E^u, tais que o espectro $Df|_{E^s} : E^s \to E^s$ está contido em $B(0,1) \subset \mathbb{C}$ e o espectro de $Df|_{E^u} : E^u \to E^u$ está contido em $\mathbb{C} \setminus \overline{B(0,1)}$. Em tal contexto, sempre é possível substituir a norma em E por uma norma equivalente do tipo adaptada, na qual

$$\|v\| = \|v_s + v_u\| = \max\{\|v_s\|, \|v_u\|\},$$

com $v_s \in E^s$, $v_u \in E^u$, e tal que $Df|_{E^s}$ e $Df^{-1}|_{E^u}$ são ambas λ-contrações, para um certo $0 < \lambda < 1$. De ora em diante, adotaremos em E a dita métrica adaptada, sendo as bolas consideradas nessa métrica.

Teorema 9.1.4. *(Variedade Estável). Sejam U e V abertos contidos em um espaço de Banach E e $f : U \to V$ um difeomorfismo de classe C^k, $k \geq 1$. Se $p \in E$ é um ponto fixo hiperbólico de f, então o conjunto estável $W^s(p)$ de p é uma variedade imersa de mesma dimensão que E^s, chamada a* variedade estável (global) *de p. Ademais, existem $r > 0$ e bolas $B^s = \overline{B(p,r)} \cap E^s$ e $B^u = \overline{B(p,r)} \cap E^u$ tais que o conjunto estável $W^s_{loc}(p)$ em $\overline{B(p,r)}$, neste caso chamado de* variedade estável local, *se escreve como o gráfico de uma aplicação $g : B^s \to B^u$ de classe C^k com as seguintes propriedades:*

1. *O gráfico de g é igual ao maximal (positivamente) invariante*

$$\cap_{j=0}^{+\infty} f^{-j}(B^s \times B^u),$$

   sendo portanto invariante por iterados positivos de f.

2. $W^s(p) = \cup_{n=0}^{+\infty} f^{-n}(W^s_{loc}(p))$.

3. *g possui derivada tal que $Lip(g) = \sup_{x \in B^s}\{\|Dg(x)\|\} < 1$.*

4. *A restrição de f ao gráfico de g é uma contração.*

5. *O gráfico de g é tangente a E^s em p.*

A parte crucial no enunciado do teorema é construir a variedade estável local, dada pelo gráfico de g. As propriedades da variedade estável global são obtidas iterando-se para trás a variedade estável local, conforme sugere o item 2 do teorema.

Como é usual em enunciados deste tipo, provaremos primeiro uma versão Lipschitz (e global) do Teorema, com $f : E \to E$ simplesmente suposta lipschitziana e com $\mathrm{Lip}(f - T)$ suficientemente pequena. Neste caso, a variedade estável global será obtida como um gráfico de uma aplicação Lipschitz $g : E^s \to E^u$, sendo portanto uma variedade mergulhada, e não apenas imersa (veja mais adiante a observação 9.1.12). O enunciado preciso da versão Lipschitz do Teorema da Variedade Estável é o seguinte:

Teorema 9.1.5. *(Variedade Estável- versão Lipschitz). Seja E um espaço de Banach e T um isomorfismo hiperbólico, e seja $0 < \lambda < 1$ tal que o raios espectrais de $T|_{E^s}$ e $[T|_{E^u}]^{-1}$ sejam ambos menores que λ. Então para toda aplicação $f : E \to E$ lipschitziana tal que $\mathrm{Lip}(f - T) < 1 - \lambda$ valem:*

1. *f possui um único ponto fixo $p \in E$;*

2. *Existe uma única aplicação $g : E^s \to E^u$, cujo gráfico é invariante por f, tal que $W^s(p) = \mathrm{graf}(g)$.*

3. $\mathrm{Lip}(g) \leq 1$.

4. *Dada qualquer bola $B = B^s \times B^u$ centrada em p,*

$$graf(g) \cap B = \cap_{j=0}^{\infty} f^{-j}(B^s \times B^u);$$

5. *A restrição de f ao gráfico de g é uma contração. Em particular, o ponto fixo p atrai todos os pontos no gráfico de g.*

A idéia da prova da versão Lipschitz é a seguinte, bastante simples. Do fato de que $\mathrm{Lip}(f - T)$ é pequena, segue-se que qualquer variedade mergulhada γ próxima a E^s será por f^{-1} fundamentalmente "expandida" na direção estável e "contraída" na direção instável. Iterando-se uma tal variedade γ por f^{-1}, $f^{-n}(\gamma)$ deverá convergir para a variedade estável de f. Mas aqui temos um problema: dar um sentido preciso à proximidade entre γ e E^s e também à convergência $f^{-n}(\gamma) \to W^s_{loc}(p)$. A idéia ingênua para definir tal sentido de convergência de sequência de variedades é o estabelecimento de uma distância (do tipo norma uniforme) entre as parametrizações de γ e de suas pré-imagens. Entretanto, o próximo exemplo muito simples nos mostra que mesmo duas parametrizações de uma mesma variedade, com mesmos domínio e imagem, podem estar distantes na métrica uniforme:

Exemplo 9.1.6. Sejam $f : (0, 2\pi) \to S^1 \setminus \{(1, 0)\}$, $\hat{f} : (0, 2\pi) \to S^1 \setminus \{(1, 0)\}$ as parametrizações bijetivas de $S^1 \setminus \{(1, 0)\}$ dadas por

$$f(\theta) = (\cos(\theta), \sin(\theta)); \qquad \hat{f}(\theta) = (\cos(-\theta), \sin(-\theta)).$$

Daí, é fácil ver que $\sup_{\theta \in (0,2\pi)}\{\|f(\theta) - \hat{f}(\theta)\|\} = \sin(\pi/2) - \sin(-\pi/2) = 2$.

Como duas parametrizações quaisquer de uma variedade podem estar, enquanto aplicações, distantes na norma uniforme, vê-se claramente que precisamos escolher um tipo particular, canônico, de parametrização para compararmos sua proximidade. Ora, sabemos que toda variedade mergulhada é localmente um gráfico de alguma aplicação. No nosso caso, esperamos que γ e suas (pré-)imagens possam ser expressas como gráficos de aplicações definidas em abertos de E^s com contradomínio em E^u. Desse modo, a

representação de γ (e suas pré-imagens) é localmente única como gráfico, uma vez que fixemos a decomposição $E = E^s \oplus E^u$, onde E^s conterá o domínio e E^u será o contradomínio. Fixemos portanto $r > 0$, e tomemos bolas $B^s = \overline{B^s(p_s, r)} \subset E^s$, $B^u = \overline{B^u(p_u, r)} \subset E^u$, com $p = p_s + p_u$, e uma $\sigma : B^s \to B^u$ tal que $\mathrm{Lip}(\sigma) \leq 1$ (o que fará com que σ não esteja muito longe da aplicação $\tilde{\sigma} \equiv 0$, cuja aplicação gráfico parametriza $E^s \cap \overline{B(p, r)} = B^s$). Seja $\gamma = graf(\sigma)$. Ao aplicarmos f^{-1} a γ, precisamos então encontrar uma aplicação $\Gamma_{f^{-1}}\sigma : B^s \to B^u$ cujo gráfico parametrize $f^{-1}(\gamma) \cap (B^s \oplus B^u)$. Procedamos primeiro formalmente, supondo que possamos mesmo escrever $f^{-1}(\gamma) \cap B(p, r)$ como o gráfico de uma aplicação Lipschitziana de B^s em E^u. Escrevendo um ponto $(x_s, x_u) \in f^{-1}(\gamma)$ por $(x_s, x_u) = (x_s, (\Gamma_{f^{-1}}\sigma)(x_s))$, escrevendo $(x_s, x_u) = f^{-1}(y_s, \sigma(y_s))$, obtemos que $x_s = \pi_s \circ f^{-1}(y_s, \sigma(y_s))$, ou seja, $y_s = (\pi_s \circ f^{-1}(id, \sigma))^{-1}(x_s)$. Donde concluímos que

$$\Gamma_{f^{-1}}(x_s) = \pi_u \circ f^{-1}(y_s, \sigma(y_s)) = [(\pi_u \circ f^{-1}) \circ (id, \sigma)](y_s) =$$
$$[(\pi_u \circ f^{-1}) \circ (id, \sigma)] \circ (\pi_s \circ f^{-1}(id, \sigma))^{-1}(x_s).$$

Claro está que boa parte do trabalho inicial consiste em mostrar que a aplicação acima, também conhecida como *transformação de gráfico*, está bem definida. Para isso, necessitamos de ser mais precisos, estabelecendo claramente o domínio e o contradomínio desta transformação. Note que a variedade estável da transformação T é justamente E^s, que é parametrizada como a aplicação gráfico de $g_T : E^s \to E^u$, dada por $g_T(x_s) = 0$, $\forall x_s \in E^s$. Como se espera que a variedade estável de f seja mesmo tangente a E^s em p (no caso $f \in C^1$), é razoável supor que, ao menos localmente, a variedade estável deve ser dada como o Lipschitz pequena. Assim, mostraremos que $\Gamma_{f^{-1}} : \mathrm{Lip}_1(B^s, B^u) \to \mathrm{Lip}_1(B^s, B^u)$. Desse modo, teremos que mostrar que quando tomamos a pré-imagem de uma γ, $f^{-1}(\gamma) \cap B(p, r)$ possa mesmo ser expressa como um tal gráfico de uma aplicação em $\mathrm{Lip}_1(B^s, B^u)$. Será extremamente conveniente termos dotado E com a norma do máximo dada por:

$$\|v\| := \max\{\|v_s\|, \|v_u\|\}, \text{ onde } v = v_s \oplus v_u, \text{ com } v_s \in E^s \text{ e } v_u \in E^u.$$

A conveniência desta norma é que dada qualquer $\sigma \in \mathrm{Lip}_1(B^s, B^u)$, a projeção natural $\pi_s : E \to E^s$, restrita ao gráfico de σ, é uma **isometria** entre o gráfico de σ e B^s. De fato, tomando dois pontos $q = (x_s, \sigma(x_s))$ e $\hat{q} = (\hat{x}_s, \sigma(\hat{x}_s))$ quaisquer no gráfico de σ temos:

$$d(q, \hat{q}) = \|(x_s, \sigma(x_s)) - (\hat{x}_s, \sigma(\hat{(x_s)}))\| = \max\{\|x_s - \hat{x}_s\|, \|\sigma(x_s) - \sigma(\hat{x}_s)\|\} \leq$$

(como σ tem 1 como constante de Lipschitz)

$$\max\{\|x_s - \hat{x}_s\|, \|x_s - \hat{x}_s\|\} = \|x_s - \hat{x}_s\| = d(\pi_s(q), \pi_s(\hat{q})).$$

Note que $\pi_s|_{graf(\sigma)}$ é a inversa da aplicação gráfico de σ dada por $x_s \mapsto (x_s, \sigma(x_s))$, a qual parametriza o gráfico de σ. Devido ao parágrafo anterior, isto quer dizer que na norma que fixamos em B, para toda $\sigma \in \mathrm{Lip}_1(B^s, B^u)$, a aplicação de gráfico de σ é uma isometria entre B^s e $graf(\sigma)$.

Note que $\mathrm{Lip}_1(B^s, B^u)$ é um subconjunto fechado do espaço de Banach das funções limitadas de B^s em E^u, dotado da norma do sup (norma uniforme). Portanto, $\mathrm{Lip}_1(B^s, B^u)$ é um espaço métrico completo. Uma vez demonstrado que a aplicação de gráfico está bem definida , usaremos da hiperbolicidade para mostrar que ela é uma contração em $\mathrm{Lip}_1(B^s, B^u)$, e se único ponto fixo nos dará a aplicação cujo gráfico é a variedade estável local.

Os próximos três lemas nos dão conta de que a transformação de gráfico está bem definida, se $\mathrm{Lip}(f - T)$ for suficientemente pequena:

Lema 9.1.7. *Seja $T : E \to E$ um isomorfismo hiperbólico em um espaço de Banach E, com $\|T|_{E^s}\|, \|[T|_{E^u}]^{-1}\|$ ambos menores que $\lambda < 1$. Se $f : E \to E$ é tal que $\mathrm{Lip}(f - T) < (1 - \lambda)$, então f possui um único ponto fixo $p \in E$.*

Prova: A primeira parte do lema é uma espécie de versão Lipschitz da proposição 8.1 da página 172. Ora, para que f possua um único ponto fixo, é suficiente que $F := f - I$ seja um homeomorfismo sobre E ($I : E \to E$ é a identidade). Temos então:

$$f - I = (f - T) + \underbrace{(T - I)}_{\text{isomorfismo, pois } 1 \notin sp(T)}$$

Pelo Teorema da perturbação do isomorfismo (Corolário 0.2.16 na página 10) aplicado a $T - I$, se $\mathrm{Lip}(f - T) < \|(T - I)^{-1}\|^{-1}$ então $f - I$ é homeomorfismo sobre E. Em particular, como $(1 - \lambda) \leq \|(T - I)^{-1}\|^{-1}$, temos que se $\mathrm{Lip}(f - T) < (1 - \lambda)$, então $f - I$ é homeomorfismo sobre E, e por conseguinte existe um único $p \in E$ tal que $f(p) - p = 0$.

□

Lembramos que se $(f - T) < \|T^{-1}\|^{-1} = \inf_{\|v\|=1} \|T(v)\|$, ainda pelo Teorema da perturbação do Isomorfismo, vale que $f = T + (f - T)$ é um homeomorfismo sobre E.

Lema 9.1.8. *Seja $T : E \to E$ um isomorfismo hiperbólico em um espaço de Banach $E = E^s \oplus E^u$, com $\|T|_{E^s}\|, \|[T|_{E^u}]^{-1}\| < \lambda < 1$. Então dado $0 < \epsilon < \lambda^{-1}$, existe $\delta = \delta(T, \epsilon) > 0$ tal que se $\mathrm{Lip}(f - T) < \delta$ então existe $f^{-1} : E \to E$, e temos que $\mathrm{Lip}(f^{-1} - T^{-1}) < \epsilon$ e que $\pi_s \circ f^{-1}(id, \sigma) : E^s \to E^s$ é um homeomorfismo bilipschitz, cuja inversa possui constante de Lipschitz*

$$\mathrm{Lip}([\pi_s \circ f^{-1}(id, \sigma)]^{-1}) \leq \frac{1}{\lambda^{-1} - \epsilon}.$$

Em particular, tomando $\epsilon < \lambda^{-1} - 1$, dado $r > 0$ se $\mathrm{Lip}(f - T) < \min\{\delta, 1 - \lambda\}$, onde p é o ponto fixo de f, então para toda $\sigma \in \mathrm{Lip}_1(\overline{B^s(p_s, r)}, \overline{B^u(p_u, r)})$ vale que $\pi_s \circ f^{-1}(id, \sigma)(\overline{B^s(p_s, r)}) \supset \overline{B^s(p_s, r)}$.

Prova:

Ainda sem fixar $\delta = \delta(T)$, vamos supô-lo menor ou igual a $\|T^{-1}\|^{-1}$. Isso já implica a existência de f^{-1}, como vimos no parágrafo que antecede este lema.

Como T deixa E^s invariante, podemos considerar a aplicação $T^s := T|_{E^s} : E^s \to E^s$. Como T é invertível, o mesmo ocorre com T^s. Pelo teorema da pertubação da aplicação bilipschitz (corolário 0.2.15 da página 9), para que $\pi_s \circ f^{-1}(id, \sigma)$ seja invertível, basta que tenhamos

$$\mathrm{Lip}(\pi_s \circ f^{-1}(id, \sigma) - [T^s]^{-1}) < \mathrm{Lip}(T^s)^{-1}.$$

Ora, como $\lambda^{-1} \leq \mathrm{Lip}(T^s)^{-1}$, é suficiente mostrarmos que

$$\mathrm{Lip}(\pi_s \circ f^{-1}(id, \sigma) - [T^s]^{-1}) < \lambda^{-1}$$

Observe que como T deixa E^s invariante, de fato vale

$$\pi_s \circ f^{-1}(x_s, \sigma(x_s))) - [T^s]^{-1}(x_s) =$$

$$(\pi_s \circ f^{-1}(X_s, \sigma(x_s))) - T^{-1}|_{E^s}(x_s) = \pi_s \circ (f^{-1} - T^{-1}) \circ (x_s, \sigma(x_s));$$

logo

$$\mathrm{Lip}(\pi_s \circ f^{-1}(id, \sigma) - T^{-1}|_{E^s}) \leq \mathrm{Lip}(\pi_s) \cdot \mathrm{Lip}(f^{-1} - T^{-1}) \cdot \mathrm{Lip}(id, \sigma) = \mathrm{Lip}(f^{-1} - T^{-1}).$$

Portanto, dado $0 < \epsilon < \lambda^{-1}$, tudo que temos de fazer é obter uma cota para $\mathrm{Lip}(f - T)$ de modo a que $\mathrm{Lip}(f^{-1} - T^{-1}) < \lambda^{-1}$

Ora,

$$f^{-1} - T^{-1} = (T + (f - T))^{-1} - T^{-1} = (T \cdot (I + T^{-1}(f - T))^{-1} - T^{-1} =$$

$$[I + T^{-1}(f - T)]^{-1} \circ T^{-1} - T^{-1} = ([I + T^{-1}(f - T)]^{-1} - I) \circ T^{-1} =$$
$$(I - [I + T^{-1}(f - T)]) \circ [I + T^{-1}(f - T)]^{-1} \circ T^{-1} =$$
$$(-T^{-1}(f - T)) \circ [I + T^{-1}(f - T)]^{-1} \circ T^{-1}.$$

Por conseguinte,

$$\mathrm{Lip}(f^{-1} - T^{-1}) \leq \mathrm{Lip}(T^{-1})^2 \cdot \mathrm{Lip}(f - T) \cdot \mathrm{Lip}([I + T^{-1}(f - T)]^{-1}) \leq$$
$$\mathrm{Lip}(T^{-1})^2 \cdot \mathrm{Lip}(f - T) \cdot \frac{1}{1 - \mathrm{Lip}(T^{-1})\,\mathrm{Lip}((f - T))}$$

Fazendo $\delta := \epsilon/2 \max\{\mathrm{Lip}(T^{-1})^2, \mathrm{Lip}(T^{-1})\}$, e $\mathrm{Lip}(f - T) < \delta$, segue-se a primeira parte do enunciado.

No caso em que $\epsilon < \lambda^{-1} - 1$, então

$$\mathrm{Lip}([\pi_s \circ f^{-1}(id, \sigma)]^{-1}) \leq \frac{1}{\lambda^{-1} - \epsilon} < 1,$$

o que implica que $\pi_s \circ f^{-1}(id, \sigma)$ expande uniformemente em todas as direções (mais precisamente, é a inversa de uma contração) e logo, como p é fixo, vale $\pi_s \circ f^{-1}(id, \sigma)(\overline{B^s(p_s, r)}) \supset \overline{B^s(p_s, r)}$.

□

Lema 9.1.9. *Seja $T : E \to E$ um isomorfismo hiperbólico em um espaço de Banach $E = E^s \oplus E^u$, com $\|T|_{E^s}\| \leq \lambda < 1$, $\|[T|_{E^u}]^{-1}\| \leq \lambda < 1$. Suponha $\epsilon < 1 - \lambda$ e considere o correspondente $\delta = \delta(T, \epsilon)$ dado no lema anterior. Se $\mathrm{Lip}(f - T) < \min\{\delta, 1 - \lambda\}$, então para qualquer $r > 0$, a transformação de gráfico $\Gamma_{f^{-1}} : \mathrm{Lip}_1(\overline{B^s(r)}, \overline{B^u(r)}) \to \mathrm{Lip}_1(\overline{B^s(r)}, \overline{B^u(r)})$ está bem definida.*

Prova: Note que se $0 < \lambda < 1$, então $(1 - \lambda) < \lambda^{-1} - 1$, logo estamos sob as hipóteses dos últimos lemas. Pelo lema anterior, já obtivemos que a fórmula abaixo (que define a transformação de gráfico avaliada em x_s)

$$\Gamma_{f^{-1}}(\sigma)(x_s) = [(\pi_u \circ f^{-1}) \circ (id, \sigma)] \circ (\pi_s \circ f^{-1}(id, \sigma))^{-1}(x_s),$$

faz sentido para $x_s \in \overline{B^s(p, r)}$ se $\sigma \in \mathrm{Lip}_1(\overline{B^s(p, r)}, \overline{B^u(p, r)})$, nos dando um valor em E^u. Para provarmos que $\Gamma_{f^{-1}}(\sigma)$ está bem definida, resta-nos verificar duas coisas:

- $\Gamma_{f^{-1}}(\sigma) \in \mathrm{Lip}_1(\overline{B^s(p,r)}, E^u)$, se $\sigma \in \mathrm{Lip}_1(\overline{B^s(p,r)}, \overline{B^u(p,r)})$. De fato,

$$\mathrm{Lip}(\Gamma_{f^{-1}}(\sigma)) \leq \mathrm{Lip}((\pi_u \circ f^{-1}) \circ (id, \sigma)) \cdot \mathrm{Lip}((\pi_s \circ f^{-1}(id, \sigma))^{-1}) \leq$$

$$\mathrm{Lip}((\pi_u \circ f^{-1}) \circ (id, \sigma)) \cdot \frac{1}{\lambda^{-1} - \epsilon} \leq$$

$$\mathrm{Lip}((\pi_u \circ f^{-1}) \circ (id, \sigma)) \leq \mathrm{Lip}((\pi_u \circ f^{-1}) \cdot \mathrm{Lip}((id, \sigma)) \leq$$

$$\mathrm{Lip}(\pi_u \circ f^{-1}) \leq \mathrm{Lip}(\pi_u T^{-1} + (\pi_u f^{-1} - \pi_u T^{-1})) \leq$$

$$\mathrm{Lip}([T^u]^{-1}) + \mathrm{Lip}(f^{-1} - T^{-1}) \leq \lambda + \epsilon \leq 1.$$

- $\Gamma_{f^{-1}}(\sigma)(x_s) \in \overline{B^u(p_u,r)}$, se $x_s \in \overline{B^s(p_s,r)}$ e $\sigma \in \mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)}$. Vimos no lema anterior que $[\pi_s \circ f^{-1}(id,\sigma)]^{-1} : \overline{B(p_s,r)} \to \overline{B(p_s,r)}$. Portanto para mostrarmos o atual item basta mostramos que se $x_s \in B(p_s,r)$, então $(\pi_u \circ f^{-1}) \circ (x_s, \sigma(x_s)) \in \overline{B(p_u,r)}$, onde $p = p_s + p_u$, com $p_s \in E^s$ e $p_u \in E^u$. Ora, no final do último item concluímos que

$$\mathrm{Lip}(\pi_u \circ f^{-1}) < \lambda + \epsilon.$$

Por conseguinte, como $f^{-1}(p_s, p_u) = (p_s, p_u)$, temos:

$$\|\pi_u f^{-1}(x_s, \sigma(x_s)) - p_u\| = \|\pi_u f^{-1}(x_s, \sigma(x_s)) - \pi_u f^{-1}(p_s, p_u)\| \leq$$

$$\mathrm{Lip}(\pi_u \circ f^{-1}) \cdot \|(x_s, \sigma(x_s)) - (p_s, p_u)\| \leq$$

$$(\lambda + \epsilon) \max\{\|x_s - p_s\|, \|\sigma(x_s) - p_u\|\} \leq (\lambda + \epsilon) r,$$

pois $(x_s, \sigma(x_s))$ foi assumido como pertencente a $\overline{B(p,r)}$.

□

De ora em diante, consideraremo-nos sob as hipóteses nas quais $\Gamma_{f^{-1}}$ está bem definida, fixando $0 < \epsilon < \lambda - 1$ e $\delta > 0$, de modo a que se $\mathrm{Lip}(f - T) < \delta$, então as teses dos lemas deste capítulo sejam todas satisfeitas.

Nosso próximo passo é mostrar que $\Gamma_{f^{-1}} : \mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_s,r)})$ é uma contração.

Lema 9.1.10. *Seja $(x_s, x_u) \in \overline{B(p,r)}$ tal que $\pi_s f^{-1}(x_s, x_u) \in B(p_s, r)$. Então para toda $\sigma \in \mathrm{Lip}_1(\overline{B(p_s,r)}, \overline{B(p_u,r)})$ vale a seguinte desigualdade:*

$$\|\pi_u f^{-1}(x_s, x_u) - (\Gamma_{f^{-1}}\sigma)(\pi_s(f^{-1}(x_s, x_u))\| \leq (\lambda + 2\epsilon)\|x_u - \sigma(x_s)\|.$$

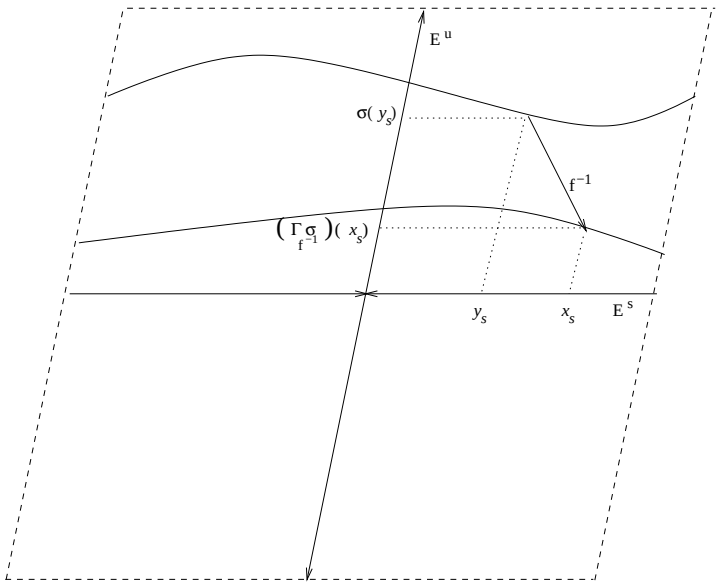


Figura 9.1: Transformação de gráfico.

Prova: A demonstração é bastante direta. O primeiro membro da inequação do enunciado é o mesmo que:

$$\|\pi_u f^{-1}(x_s, x_u) - [(\pi_u \circ f^{-1}) \circ (id, \sigma)] \circ (\pi_s \circ f^{-1}(id, \sigma))^{-1}(\pi_s(f^{-1}(x_s, x_u)))\| \leq$$

$$\|\pi_u f^{-1}(x_s, x_u) - [(\pi_u \circ f^{-1}) \circ (id, \sigma)](x_s)\| +$$
$$\|(\Gamma_{f^{-1}}\sigma)(\pi_s(f^{-1}(x_s, \sigma(x_s)) - (\Gamma_{f^{-1}}\sigma)(\pi_s(f^{-1}(x_s, x_u))\| \leq$$

(somando e subtraindo $(\Gamma_{f^{-1}}\sigma)(\pi_s(f^{-1}(x_s, \sigma(x_s))$ e aplicando a desigualdade triangular)

$$\mathrm{Lip}(\pi_u \circ f^{-1})\|(x_s, x_u) - (x_s, \sigma(x_s))\| +$$

$$\mathrm{Lip}(\Gamma_{f^{-1}}\sigma)\|\pi_s(f^{-1}(x_s, \sigma(x_s)) - \pi_s(f^{-1}(x_s, x_u)\| \leq$$

(pois vimos no lema anterior que $\mathrm{Lip}(\pi_u \circ f^{-1}) \leq \lambda + \epsilon$ e que $\mathrm{Lip}(\Gamma_{f^{-1}}\sigma) \leq 1$)

$$(\lambda + \epsilon)\|(x_s, x_u) - (x_s, \sigma(x_s))\| + \|\pi_s(f^{-1}(x_s, \sigma(x_s)) - \pi_s(f^{-1}(x_s, x_u)\| =$$

(observando que $\pi_s T^{-1}(x_s, x_u) = \pi_s T^{-1}(x_s, \sigma(x_s)$ e com mais um argumento de soma e subtração)

$$(\lambda + \epsilon)\|(x_s, x_u) - (x_s, \sigma(x_s))\| +$$

$$\|\pi_s(f^{-1}(x_s,\sigma(x_s)) - \pi_s T^{-1}(x_s,\sigma(x_s)) - \pi_s(f^{-1}(x_s,x_u) + \pi_s T^{-1}(x_s,x_u)\| \leq$$
$$(\lambda+\epsilon)\|x_u - \sigma(x_s)\| + \mathrm{Lip}(\pi_s f^{-1} - \pi_s T^{-1})\|x_u - \sigma(x_s)\| \leq (\lambda+2\epsilon)\|x_u - \sigma(x_s)\|.$$

□

Lema 9.1.11. *Seja $r > 0$ qualquer, tome $\epsilon < 1-2\epsilon$ arbitrário e $\delta = \delta(\epsilon, T) > 0$ correspondente (nos lemas anteriores) de modo a que $\mathrm{Lip}(f - T) < \delta$ implique em que $\Gamma_{f^{-1}} : \mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)}) \to \mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$ esteja bem definida e que $\mathrm{Lip}(f^{-1} - T^1) < \epsilon$. Considere ainda $Lip_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$ dotada da métrica uniforme. Então $\Gamma_{f^{-1}}$ é uma $\lambda + 2\epsilon$-contração em $\mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$. Em particular, $\Gamma_{f^{-1}}$ possui um único ponto fixo.*

Prova:

Sejam $\sigma, \hat{\sigma} \in Lip_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$. Dado $x_s \in \overline{B(p_s,r)}$, pela segunda parte do enunciado do lema 9.1.8, existe $y_s \in \overline{B(p_s,r)}$ tal que $x_s = \pi_s \circ f^{-1}(id,\sigma)(y_s)$

$$\|(\Gamma_{f^{-1}}\sigma)(x_s) - (\Gamma_{f^{-1}}\hat{\sigma})(x_s)\| =$$

$$\|(\pi_u \circ f^{-1})(id,\sigma) \circ (\pi_s \circ f^{-1}(id,\sigma))^{-1}(\pi_s(f^{-1}(y_s,\sigma(y_s))) - (\Gamma_{f^{-1}}\hat{\sigma})(\pi_s(f^{-1}(y_s,\sigma(y_s)))\| =$$
$$\|(\pi_u \circ f^{-1})(y_s,\sigma(y_s)) - (\Gamma_{f^{-1}}\hat{\sigma})(\pi_s(f^{-1}(y_s,\sigma(y_s)))\| \leq$$

(pelo lema anterior)

$$(\lambda + 2\epsilon) \cdot \|\sigma(y_s) - \hat{\sigma}(y_s)\| \leq (\lambda + \epsilon) \cdot \sup_{x \in \overline{B^s(p_s,r)}} \|\sigma(x) - \hat{\sigma}(x)\|.$$

Tomando o supremo em x_s na expressão acima, comcluímos que

$$\|\Gamma_{f^{-1}}\sigma - \Gamma_{f^{-1}}\hat{\sigma}\| \leq (\lambda + 2\epsilon) \cdot \|\sigma - \hat{\sigma}\|,$$

ou seja, $\Gamma_{f^{-1}}$ é uma contração para a métrica uniforme em $\mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$.

Como $\mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$ é um subconjunto fechado do espaço de Banach $C_b^0(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$ das aplicações contínuas e limitadas de $\overline{B^s(p_s,r)}$ em $\overline{B^u(p_u,r)}$, segue que é um espaço métrico completo. Desse modo, o Teorema do Ponto Fixo para Contrações (teorema 0.2) implica que $\Gamma_{f^{-1}}$ possui um único ponto fixo $g_r \in \mathrm{Lip}_1(\overline{B^s(p_s,r)}, \overline{B^u(p_u,r)})$.

□

Podemos agora arrematar a prova do Teorema da Variedade Estável, em sua versão Lipschitz:

Prova:

Fixemos $\epsilon > 0$, $\epsilon < 1 - \lambda$. Para cada $r > 0$, aplicamos os lemas acima de modo a obter $\Gamma_{f^{-1}}$ e seu correspondente (e único em $\mathrm{Lip}_1(\overline{B^s(p_s, r)}, \overline{B^u(p_u, r)})$) ponto fixo g_r. Definimos $g : E^s \to E^u$ como

$$g(x_s) := g_r(x_s), \text{ se } x_s \in B^s(p_s, r).$$

Pela unicidade local de cada g_r, segue-se que g está bem definida, e pertence a $\mathrm{Lip}_1(E^s, E^u)$: dados $x_s, y_s \in E^s$, existe $\hat{r} > 0$ tal que $x_s, y_s \in B^s(p_s, \hat{r})$. Portanto,

$$d(g(x_s), g(y_s)) = d(g_{\hat{r}}(x_s), g_{\hat{r}}(y_s)) \leq \mathrm{Lip}(g_{\hat{r}}) d(x_s, y_s) = d(x_s, y_s).$$

e o item 3 do enunciado está demonstrado.

O item 1 (existência e unicidade de ponto fixo de f) do teorema já foi provado no lema 9.1.7.

Por construção, dados $r > 0$, e $\sigma \in \mathrm{Lip}_1(B^s(p_s, r), B^u(p_u, r))$ temos que $\mathrm{graf}(\Gamma_{f^{-1}}(\sigma)) \subset f^{-1}(\mathrm{graf}(\sigma))$. Como g_r é ponto fixo de $\Gamma_{f^{-1}}$ temos então que $\mathrm{graf}(g_r) \subset f^{-1}(\mathrm{graf}(g_r))$, e logo $f(\mathrm{graf}(g_r)) \subset \mathrm{graf}(g_r)$, o que em particular nos dá a invariância do gráfico de g por f (o que é parte do item 2).

Note que já provamos no parágrafo anterior que fixado $r > 0$, então $\mathrm{graf}(g) \cap \overline{B(p, r)} = \mathrm{graf}(g_r) \subset \cap_{j=0}^{\infty} f^{-j}(B^s \times B^u)$. Mostremos reciprocamente que, fixado $\hat{r} > 0$ se um ponto x possui todas as suas imagens em $B(p, r)$, então $x \in \mathrm{graf}(g_r)$. De fato, se $x = (x_s, x_u) = f^{-1}(y)$, com $(y_s, y_u) = y \in B(p, r)$, pelo lema 9.1.10, temos que

$$\|\pi_u f^{-1}(y_s, y_u) - g_r(\pi_s(f^{-1}(y_s, y_u)))\| \leq (\lambda + 2\epsilon)\|y_u - g_r(y_s)\|.$$

Usando de indução, temos que se $y = (y_s, y_u)$ é tal que $y, f^{-1}(y), \ldots, f^{-n}(y) = x = (x_s, x_u)$ pertencem a $B(p, r)$, temos que

$$\|\pi_u f^{-n}(y_s, y_u) - g_r(\pi_s(f^{-n}(y_s, y_u)))\| \leq$$

(caso $n = 1$, provado acima)

$$(\lambda + 2\epsilon)\|\pi_u f^{-n+1}(y_s, y_u) - g_r(\pi_s f^{-n+1}(y_s, y_u)\| \leq$$

(hipótese de indução)

$$(\lambda + 2\epsilon) \cdot (\lambda + 2\epsilon)^{n-1}\|y_u - g_r(y_s)\|.$$

Em particular, vale que $\|x_u - g_r(x_s)\| \leq (\lambda + 2\epsilon)^n \cdot r$. Por conseguinte, se $(x_s, x_u) = x \in B(p,r)$ possui toda a sua semi-órbita positiva contida em $B(p,r)$, então $x_u = g_r(x_s)$, ou seja, $x \in \text{graf}(g_r)$. Isso é o mesmo que dizer que

$$\cap_{j=0}^{\infty} f^{-j}(B^s \times B^u) \subset \text{graf}(g_r).$$

Isso conclui o item 4.

Mostremos que $f|_{\text{graf}(g)}$ é uma contração. Para isso basta vermos que $f|_{\text{graf}(g_r)}$ é uma contração, para $r > 0$ arbitrário. Lembramos que pelo item 2, $f(\text{graf}(g_r)) \subset \text{graf}(g_r)$. Ora, vimos em nossa digressão anterior aos lemas que a norma adotada faz da projeção $\pi_s|_{\text{graf}(g_r)} : \text{graf}(g_r) \to B^s(p_s, r)$ uma isometria, cuja inversa é simplesmente a aplicação gráfico $x_s \mapsto (x_s, g_r(x_s))$. Esta última aplicação é a nossa parametrização canônica de $\text{graf}(g_r)$, daí temos (pelo fato de $\pi_s|_{\text{graf}(g)}$ e sua inversa serem isometrias) que tanto $f|_{\text{graf}(g_r)}$ como sua expressão em carta bilipschitz $\pi_s|_{graf(g_r)} \circ f|_{\text{graf}(g_r)} \circ (id, g) : B^s(p_s, r) \to B^s(p_s, r)$ possuem a mesma constante de Lipschitz.

Ora, mas como o gráfico de g_r é $f-$invariante,

$$\pi_s|_{\text{graf}(g_r)} \circ f \circ (id, g_r) = (id, g_r)^{-1} \circ [f^{-1}]^{-1} \circ [\pi_s|_{\text{graf}(g_r)}]^{-1} = [\pi_s \circ f^{-1}(id, g_r)]^{-1}.$$

Portanto, segue-se que

$$\text{Lip}(f|_{\text{graf}(g_r)}) = \text{Lip}(\pi_s \circ f \circ (id, g_r)) =$$

(pelo lema 9.1.8)

$$\text{Lip}([\pi_s \circ f^{-1}(id, g_r)]^{-1}) \leq \frac{1}{\lambda^{-1} - \epsilon} < 1.$$

Veja que a constante de Lipschitz acima não depende de r. Concluímos que nesse caso Lipschitz global, $f|_{\text{graf}(g)}$ é uma contração, concluindo o item 5 que restava.

□

Observação 9.1.12. Note que, em geral, a hipótese global $\text{Lip}(f - T)$ suficientemente pequeno é forte. Comumente tal só ocorre para restrições de uma aplicação a vizinhança de algum ponto fixo hiperbólico $p = f(p)$. Em tal situação, $\text{Lip}([f - Df(p)]|_{B(p,r)})$ pode ser tomado (positivo) tão pequeno quanto se queira, bastando para tal tomar $r > 0$ adequadamente. Por esta razão a Variedade Estável Global no contexto diferenciável (com hipóteses locais) geralmente não é mergulhada, mas apenas injetivamente imersa, como ilustra a figura 9.1.

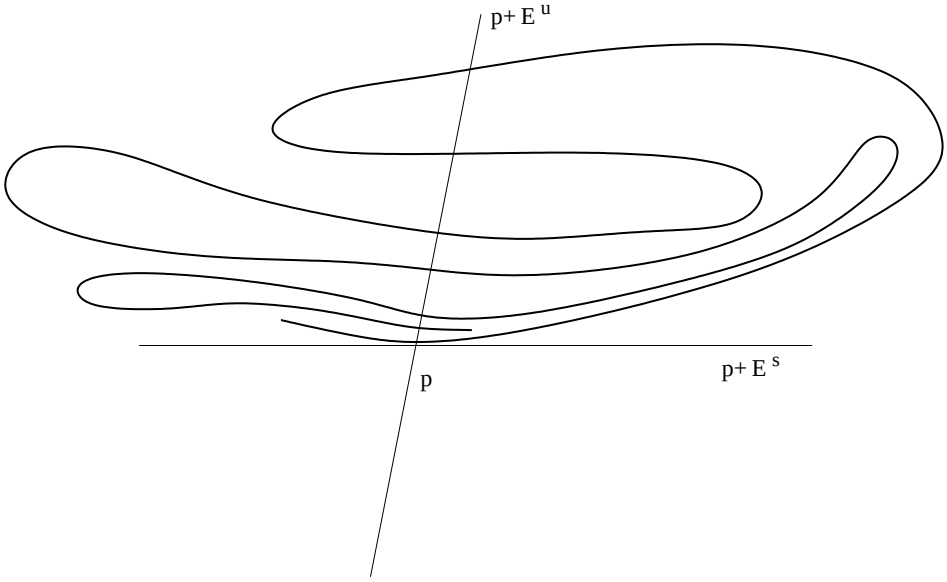


Figura 9.1: Representação de uma variedade estável global. Note que a variedade estável global pode exibir auto-acumulação. Em particular, a intersecção da variedade estável global com qualquer vizinhança B do ponto fixo hiperbólico p pode ser diferente da variedade estável local $W^s_{loc}(p)$ correspondente.

9.2 O Teorema da Variedade estável para campos

Índice Remissivo

Índice
de estabilidade de um campo linear, 121
de um isomorfismo linear, 184

Adjunta
de uma aplicação linear, 154
Anel
centrado em $a \in \mathbb{C}$, 145
Aplicação
Holomorfa, 140
Resolvente, 139

Base
de uma topologia, 22
ortonormal, 155
Bola aberta, 2

Cálculo Funcional, 149
Caminho
integrável à Riemann, 20
Campo
de Vetores, 27, 61
gradiente, 78
linear
hiperbólico, 121
Complexificado de um operador real, 112
Conjugação
de campos, 67
Conjunto
aberto, 2
convexo, 23
de $\alpha-$limite, 80
de $\omega-$limite, 80
estável, 193
fechado, 2
Resolvente de um operador, 139
Conjuntos
maximais invariantes, 193
Contração, 5
Curva
de Jordan, 84
fechada e simples, 83

Desigualdade
de Gronwall, 56
Diâmetro
de uma partição, 20

Equação
a variáveis separáveis, 30
autônoma, 27
Equações
lineares
homogêneas e não homogêneas, 92
Equicontinuidade, 12
Equivalência de campos, 67
Espaço
de Banach, 4
de Hilbert, 154

dual, 154
estável, 184
métrico, 1
completo, 4
ortogonal, 155
topológico, 2
vetorial
normado, 2
Espectro de um operador, 139
Exponencial
de um operador, 100

Fórmula
de Liouville, 98
Fórmula Integral de Cauchy, 144
Fluxo, 62
local, 63
tubular, 71

Hiperfície, 123

Integral
de Riemann, 20
Integral primeira, 76
Intervalo
maximal, 43
Isomorfismo
linear hiperbólico, 153

Métrica, 1
Matriz
fundamental, 94

Norma, 2

Operador
compacto, 165

Partição
de um intervalo, 19
Polinômio
característico, 106
Ponto fixo
hiperbólico, 168
Problema
de Cauchy, 28
Propriedade
da intersecção finita, 11
Pull-Back
de um campo, 70

Refinamento
de uma partição, 20
Retrato de fase, 67

Série
de Laurent, 146
Secção transversal
a um campo, 70
Sequência, 3
convergente, 3
de Cauchy, 3
equicontínua de funções, 12
Singularidade
hiperbólica, 175
Solução
da equação linear, 97
fundamental, 94
maximal, 43
Soma
de Riemann, 20
Subespaço
ortogonal, 155
Subsequência, 3
convergente, 3

Teorema
da Curva de Jordan, 83

da decomposição em autoespaços generalizados, 104
da Fórmula Integral de Cauchy, 144
da Forma de Jordan
caso complexo, 110
caso real, 112
da perturbação
da aplicação bilipschitz, 9
da identidade, 7
do isomorfismo, 10
da Variedade Estável, 195
de Aproximação de Weierstrass, 18
de Ascoli-Arzelá, 15
de Cantor-Tychonov, 14
de Cauchy-Goursat, 141
de Cayley-Hamilton, 109
de dependência contínua, 54
de dependência diferenciável, 58
de Dini, 15
de Grobman-Hartman
para campos, 176
para difeomorfismos, 168
de Peano, 41
de Picard, 37
de Pitágoras, 158
de Poincaré-Bendixson, 84
do Fluxo Local, 62
do Fluxo Tubular, 71
do mapeamento espectral, 150
do ponto fixo
para contrações, 5
Topologia, 2
Transformação de Poincaré, 74

Valor
Inicial, 28

Variedade
Estável Local, 195