

2

Conjuntos críticos de funções admissíveis

Como se verá, os conjuntos críticos de funções admissíveis são muito estruturados geometricamente. Nesse capítulo apresentam-se os primeiros resultados: os capítulos subsequentes usam ainda outras cotas para os conjuntos críticos, que serão obtidas mais tarde.

2.1

Os três tipos de não linearidades relevantes

Uma função $g \in C^1(\mathbb{R})$ é *superlinear* se for estritamente convexa e

$$\lim_{|x| \rightarrow +\infty} \frac{g(x)}{|x|} = +\infty$$

e é *assintoticamente linear* se

$$\lim_{x \rightarrow -\infty} \frac{g(x)}{-x} = a, \quad \lim_{x \rightarrow \infty} \frac{g(x)}{x} = b, \quad a < b.$$

Finalmente, $g \in C^0(\mathbb{R})$ é *linear por partes* se $g(x) = bx^+ - ax^-$, onde $x^+ = \max(x, 0)$ e $x^- = \max(-x, 0)$. Os valores a e b são os *parâmetros assintóticos* nos casos acima. Uma aplicação $f : \mathbb{R} \rightarrow \mathbb{R}^n$ é convexa, *superlinear*, *assintoticamente linear* ou *linear por partes*, respectivamente, se todas as coordenadas f_i apresentam tal propriedade.

Dadas uma matriz A real, simétrica de ordem n e uma família de funções $f_i : \mathbb{R} \rightarrow \mathbb{R}$ de classe C^0 , $i = 1, \dots, n$, consideremos a *aplicação admissível* F associada a A e f ,

$$\begin{aligned} F: \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ u &\mapsto Au - f(u) \end{aligned} \tag{2-1}$$

onde $f(u) = (f_1(u_1), f_2(u_2), \dots, f_n(u_n))^T$. Aplicações admissíveis F_{sl} , F_{al} e F_{pl} são do tipo *superlinear*, *assintoticamente linear* ou *linear por partes*, respectivamente, se todas as coordenadas f_i da não linearidade f apresentam tal propriedade. Por brevidade, vamos falar de aplicações *sl*-admissíveis, *al*-admissíveis e *pl*-admissíveis. Para aplicações *al*-admissíveis e *pl*-admissíveis,

supomos que os parâmetros assintóticos a e b são iguais para cada uma das funções coordenadas da não linearidade.

A definição de função admissível acomoda as discretizações por diferenças finitas de operadores elípticos, como se vê no Apêndice 6.4. Parece razoável supor, no caso $a\ell$ -admissível, que cada coordenada f_i tenha comportamentos assintóticos diferentes, a_i e b_i . Isso não basta para certos resultados deduzidos a seguir: é importante que os valores assintóticos sejam independentes de i . Por outro lado, as discretizações sugerem que a hipótese não é forçada.

Outra hipótese conveniente e natural no contexto de discretizações, que não será usada até o Capítulo 4, é supor que a matriz A , além de simétrica e positiva, tenha posições não positivas fora da diagonal, isto é, A é uma matriz de *Stieltjes*.

O lema a seguir é um exercício simples de cálculo.

Lema 2.1 *Se g é superlinear então*

$$\lim_{x \rightarrow \pm\infty} g'(x) = \pm\infty$$

e se g é assintoticamente linear convexa, então a imagem de sua derivada é o intervalo (a, b) .

2.2

Geometria do conjunto crítico para não linearidades suaves

Seja $\text{Diag}(v)$ a matriz diagonal cujas entradas diagonais são as coordenadas do vetor v . Derivando (2-1), obtemos

$$DF(u) \cdot v = [A - \text{Diag}(f'(u))]v,$$

ou seja,

$$DF(u) = A - \text{Diag}(f'(u)) = A - \begin{bmatrix} f'_1(u_1) & 0 & \cdots & 0 \\ 0 & f'_2(u_2) & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & f'_n(u_n) \end{bmatrix}.$$

O conjunto de autovalores de uma matriz A será denotado por $\sigma(A)$, e será sempre indexado em ordem não decrescente,

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Se F é de classe $C^1(\mathbb{R}^n)$, o conjunto crítico de F , denotado por $\mathcal{C}(F)$ ou simplesmente \mathcal{C} quando não houver ambiguidade, é o conjunto

$$\mathcal{C}(F) = \{u \in \mathbb{R}^n \mid 0 \in \sigma(DF(u))\} = \{u \in \mathbb{R}^n \mid \det(DF(u)) = 0\}.$$

Pra um vetor $w \in \mathbb{R}^n$, usaremos as notações $w > 0$ e $w < 0$ pra indicar que todas as coordenadas de w são estritamente positivas e estritamente negativas, respectivamente. Nos próximos resultados, mostramos que, para certas funções admissíveis, $\mathcal{C}(F)$ é composto por n subconjuntos

$$\mathcal{C}_i(F) = \{u \in \mathbb{R}^n \mid DF(u) \text{ tem seu } i\text{-ésimo autovalor } \lambda_i(u) = 0\}$$

que são gráficos de funções contínuas $\gamma_i : p^\perp \rightarrow \mathbb{R}$, $p > 0$. Os argumentos a seguir são empregados na demonstração de propriedades análogas de certos operadores diferenciais em [BT].

Lema 2.2 *Sejam $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ um aplicação admissível suave e $u, w \in \mathbb{R}^n$. Suponha que $\tilde{\lambda}(u)$ seja um autovalor simples de $DF(u)$ com $v(u) = (v_1(u), v_2(u), \dots, v_n(u))$ como autovetor unitário associado. Então*

$$\frac{\partial \tilde{\lambda}}{\partial w}(u) = \left. \frac{d\tilde{\lambda}}{dt}(u + tw) \right|_{t=0} = - \sum_i f_i''(u_i) w_i v_i^2(u).$$

Demonstração:

Um autovalor simples $\tilde{\lambda}$ de uma matriz simétrica S e seu autovetor v associado, com $\langle v, v \rangle = 1$, variam suavemente com S ([LAX], Teoremas 7 e 8 do Capítulo 9). Sejam então \mathcal{S}_n o conjunto das matrizes simétricas de ordem n e uma curva suave $t \mapsto S(t) \in \mathcal{S}_n$. Seja $\tilde{\lambda}(t)$ um autovalor de $S(t)$ associado ao autovetor normalizado $v(t)$ e suponhamos $\tilde{\lambda}(0)$ simples. Derivando $S(t)v(t) = \tilde{\lambda}(t)v(t)$ obtemos, para t próximo de 0,

$$\dot{\tilde{\lambda}}(t) = \langle \dot{S}(t)v(t), v(t) \rangle \tag{2-2}$$

Logo por (2-2),

$$\frac{\partial \tilde{\lambda}}{\partial w}(u) = \left. \frac{d\tilde{\lambda}}{dt}(u + tw) \right|_{t=0} = \langle \dot{S}(u + tw)v(u + tw), v(u + tw) \rangle \Big|_{t=0} \tag{2-3}$$

Estamos interessados em

$$S(t) = DF(u + tw) = A - \text{Diag}(f'((u + tw)_i)) =$$

$$= A - \begin{bmatrix} f'_1(u_1 + tw_1) & 0 & \cdots & 0 \\ 0 & f'_2(u_2 + tw_2) & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & f'_n(u_n + tw_n) \end{bmatrix}$$

e para essa curva

$$\left. \frac{dS}{dt}(u + tw) \right|_{t=0} = -\text{Diag}(f''_i(u_i)w_i)$$

Substituindo em (2-3),

$$\frac{\partial \tilde{\lambda}}{\partial w}(u) = -\langle \text{Diag}(f''_i(u_i)w_i)v(u), v(u) \rangle = -\sum_i f''_i(u_i)w_i v_i^2(u). \quad (2-4)$$

□

Por conveniência, dizemos que os *parâmetros assintóticos* de uma aplicação sl -admissível são $a = -\infty$ e $b = +\infty$.

Proposição 2.3 *Seja $F = A - f$ uma aplicação sl -admissível ou $p\ell$ -admissível, onde f possui coordenadas f_i suaves, convexas, com parâmetros assintóticos a e b e $f''_i = 0$ num conjunto de pontos isolados. Então cada autovalor $\tilde{\lambda}_k(t)$ de $DF(u + tp)$ é estritamente decrescente ao longo de retas $r(t) = u + tp$, $p > 0$. Mais ainda,*

$$\lim_{t \rightarrow -\infty} \tilde{\lambda}_k(t) = \lambda_k - a \quad e \quad \lim_{t \rightarrow +\infty} \tilde{\lambda}_k(t) = \lambda_k - b.$$

O conjunto crítico $\mathcal{C}(F)$ é localmente uma subvariedade de codimensão 1 numa vizinhança $V_u \subset \mathbb{R}^n$ de pontos u para os quais $\tilde{\lambda}_k(u)$ é simples.

Assim, ao longo de cada reta $r(t) = u + tp$ existe no máximo um ponto no qual $\tilde{\lambda}_k(t) = 0$. Mais, os subconjuntos críticos $\mathcal{C}_k(F)$ são não vazios exatamente quando $\lambda_k \in (a, b)$.

Demonstração:

Para pontos $u + t^*p$ onde $\tilde{\lambda}_k$ é simples, pelo lema anterior,

$$\left. \frac{d\tilde{\lambda}_k}{dt}(u + tp) \right|_{t=t^*} = -\sum_i f''_i(u_i + t^*p_i)p_i v_i^2(u + t^*p) \leq 0, \quad (2-5)$$

devido à convexidade de f_i e o fato que $p > 0$. A desigualdade estrita vale em quase todos os pontos da reta $u + tp$, ou seja, nos pontos onde $f'' > 0$. Assim,

cada $\tilde{\lambda}_k$ é estritamente decrescente ao longo da reta $u + tp$. Logo, existe no máximo um único ponto em que $\tilde{\lambda}_k(t) = 0$.

Os elementos da diagonal de $DF(u + tp)$ são $A_{ii} - f'_i(u_i + tp_i)$ e pela admissibilidade

$$\lim_{t \rightarrow -\infty} [f'_i(u_i + tp_i)] = a \quad \text{e} \quad \lim_{t \rightarrow +\infty} [f'_i(u_i + tp_i)] = b.$$

Daí, quando $t \rightarrow -\infty$ e $t \rightarrow +\infty$, $DF(u + tp)$ converge respectivamente para $A - aI$ e $A - bI$, o que implica em

$$\lim_{t \rightarrow -\infty} \tilde{\lambda}_k(t) = \lambda_k - a \quad \text{e} \quad \lim_{t \rightarrow +\infty} \tilde{\lambda}_k(t) = \lambda_k - b.$$

Assim, $\lambda_k - b \leq \tilde{\lambda}_k \leq \lambda_k - a$ e $\tilde{\lambda}_k$ se torna zero (isto é, $\mathcal{C}_k(F)$ é não vazio) exatamente quando $\lambda_k \in (a, b)$. Esses argumentos são independentes de u : para um índice k , $\lambda_k \in (a, b)$ e $p > 0$, cada linha $u + tp$ encontra $\mathcal{C}_k(F)$ somente uma única vez. Portanto, para $u \in p^\perp$, podemos definir $\gamma_k(u) = t_k$, onde $u + t_k p$ é o único ponto sobre a reta $u + tp$ em $\mathcal{C}_k(F)$.

Em pontos críticos $u + t_k p$ nos quais $\tilde{\lambda}_k(u + t_k p) = 0$, se a desigualdade (2-5) vale estritamente, pelo teorema de função implícita, a componente crítica $\mathcal{C}_k(F)$ é gráfico de uma função suave $\tilde{\gamma}_k : p^\perp \cap \mathcal{V}(u) \rightarrow \mathbb{R}$ para uma vizinhança $\mathcal{V}(u)$ de $u \in p^\perp$. Obviamente, $\gamma_k = \tilde{\gamma}_k$.

□

Existe uma classe de matrizes especialmente conveniente para ilustrar os resultados acima.

Proposição 2.4 *Seja A uma matriz $n \times n$ simétrica tridiagonal com posições $A_{i,i+1}, A_{i+1,i}, i = 1, \dots, n - 1$ não nulas. Então $\sigma(A)$ é simples.*

Demonstração:

Suponha que existam u e v autovetores linearmente independentes associados ao autovalor λ . Considere w uma combinação linear de u e v cuja primeira coordenada é nula. A equação $(A - \lambda I)w = 0$ obriga sucessivamente que as demais coordenadas de w também sejam nulas, o que implica no alinhamento de u e v .

□

Exemplo 2.5: Sejam

$$A_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix},$$

e não linearidades $f_{sl} = (g, g, g)$ e $f_{al} = (h, h, h)$ onde

$$g(x) = x^2 \quad \text{e} \quad h(x) = \frac{2.8(2x \arctan(x) - \ln(1 + x^2) + x\pi)}{2\pi}.$$

Obviamente, $F_{sl} = A_3 - f_{sl}$ é sl -admissível, $F_{al} = A_3 - f_{al}$ é al -admissível e $\sigma(A_3) = \{\lambda_1 = 2 - \sqrt{2} \approx 0.58, \lambda_2 = 2, \lambda_3 = 2 + \sqrt{2} \approx 3.41\}$. Mais ainda, a imagem de h' é o intervalo aberto $(0.0, 2.8)$, que contém apenas λ_1 e λ_2 , que garante que $\mathcal{C}(F_{al})$ possui exatamente duas componentes críticas. Nas Figuras 2.1 e 2.2 exibimos seus conjuntos críticos. Na figura, mostramos um vetor $-p$, $p > 0$, para sugerir que as superfícies críticas de fato são gráficos sobre p^\perp .

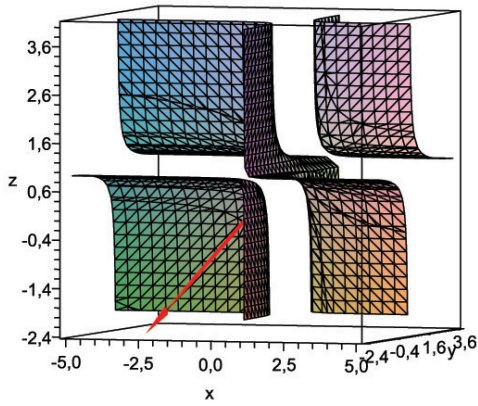


Figura 2.1: $\mathcal{C}(F_{sl})$

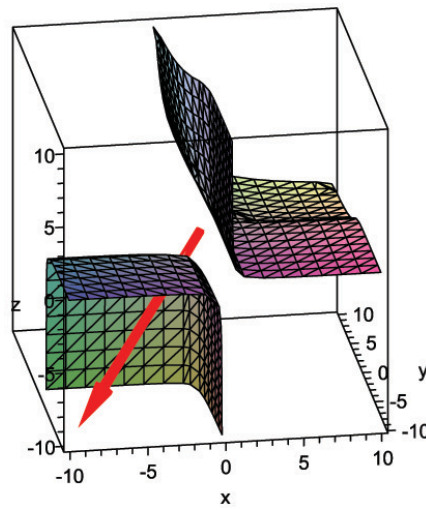


Figura 2.2: $\mathcal{C}(F_{al})$

□

Exemplo 2.6: Sejam

$$A = \begin{bmatrix} 7 & 1 & -2 \\ 1 & 7 & -2 \\ -2 & -2 & 10 \end{bmatrix}$$

para a qual $\sigma(A) = \{6, 6, 12\}$, $f_{sl} = (g, g, \dots, g)$ onde $g(x) = x^2$ e $F_{sl} = A - f_{sl}$. Portanto,

$$DF_{sl}(u) = \begin{bmatrix} 7 - 2u_1 & 1 & -2 \\ 1 & 7 - 2u_2 & -2 \\ -2 & -2 & 10 - 2u_3 \end{bmatrix}.$$

O conjunto crítico $\mathcal{C}(F_{sl})$ é exibido na Figura 2.3 onde vemos que os conjuntos

$$\mathcal{C}_1 = \{u \in \mathbb{R}^3 \mid \tilde{\lambda}_1(u) = 0\} \quad \text{e} \quad \mathcal{C}_2 = \{u \in \mathbb{R}^3 \mid \tilde{\lambda}_2(u) = 0\}$$

se interceptam no ponto $u_0 = (3, 3, 3)$, logo $\mathcal{C}(F_{sl})$ não é localmente uma subvariedade numa vizinhança de u_0 .

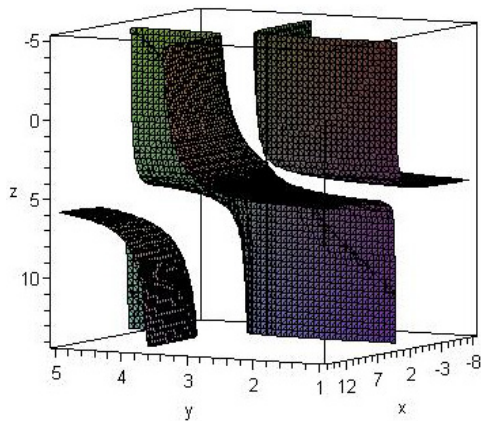


Figura 2.3: $\mathcal{C}(F_{sl})$: não é uma subvariedade

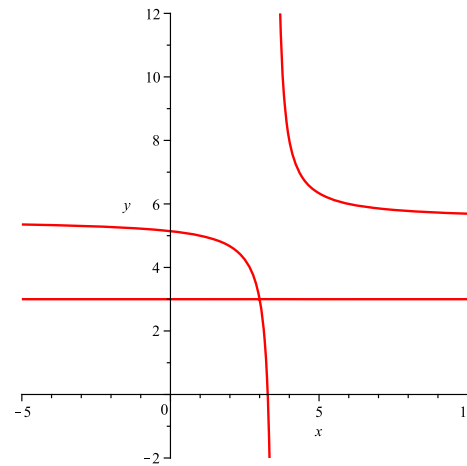


Figura 2.4: Intersecção de $\mathcal{C}(F_{sl})$ com o plano $y = z$

□

Quando ocorre multiplicidade de autovalores, os conjuntos $\mathcal{C}_i(F)$ não formam componentes disjuntas de $\mathcal{C}(F)$. A figura à esquerda não deixa isso tão claro: a triangulação das superfícies elimina o ponto comum. Isso entretanto é evidente no corte desenhado ao lado. Em outras palavras, $\mathcal{C}(F)$ não é uma subvariedade, perdendo regularidade nos pontos de espectro múltiplo.

2.3

O conjunto crítico para funções pl -admissíveis

No espaço \mathbb{R}^n , denominamos de *ortante* cada um dos 2^n subconjuntos caracterizados pelas desigualdades

$$s_1x_1 > 0, s_2x_2 > 0, \dots, s_nx_n > 0$$

onde $s_i = -1$ ou $s_i = +1$.

A aplicação $F_{pl}(u) = Au - f_{pl}(u)$ não é suave em \mathbb{R}^n . De fato, F_{pl} é apenas suave quando restrita a cada ortante, e contínua globalmente. No ortante \mathcal{O} , cujos os vetores possuem coordenadas com o padrão de sinais $s_1s_2 \dots s_n$, temos

$F_{p\ell}(u) = (A - D^{\mathcal{O}})u$, onde $D_{ii}^{\mathcal{O}} = a$ quando s_i é negativo e $D_{ii}^{\mathcal{O}} = b$ quando s_i é positivo. Imporemos que $F_{p\ell}$ seja *não ressonante*, isto é, sua restrição a cada ortante é inversível. Em outras palavras, a não ressonância de $F_{p\ell}$ implica que $\det DF_{p\ell}$ no ortante \mathcal{O} , ou seja, $\det(A - D^{\mathcal{O}})$, tenha um sinal bem definido. Assim, a perda de suavidade de $F_{p\ell}$ ocorre apenas na fronteira dos ortantes, que é formada pelos planos coordenados. Dizemos que dois ortantes são *adjacentes* se seus padrões de sinais diferem apenas numa localização.

Definimos então o conjunto crítico $\mathcal{C}(F_{p\ell})$ como o conjunto dos pontos dos planos coordenados que é fronteira de ortantes adjacentes possuindo $\det(A - D^{\mathcal{O}})$ com sinais opostos.

As figuras apresentadas no Exemplo 2.7 ilustram algumas possibilidades em dimensão 2: à esquerda, os conjuntos críticos de aplicações $p\ell$ -admissíveis estão sobre os eixos coordenados. As notações $0+$ e $0-$ representam os semi-eixos verticais positivo e negativo, respectivamente. Para os semi-eixos horizontais, analogamente, usamos as notações -0 e $+0$. À direita, as figuras exibem a imagem dos conjuntos críticos pela correspondente aplicação.

Perto de um ponto crítico genérico u (mais especificamente, aqueles u críticos com uma única coordenada nula), $F_{p\ell}$ é uma dobra. Em outras palavras, depois de compor com homeomorfismos locais, $F_{p\ell}$ perto de u tem forma normal $(x, y_1, \dots, y_{n-1}) \rightarrow (x^2, y_1, \dots, y_{n-1})$.

Exemplo 2.7: Considere a matriz

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 5 \end{bmatrix}$$

cujos espectro é $\{\approx 2.44, \approx 6.56\}$. Seja ainda $f_{p\ell} = bu^+ - au^-$ para os pares $(a, b) \in \{(0, 3), (6.5, 7), (0, 7)\}$, todos dando origem a funções $p\ell$ -admissíveis não ressonantes. As figuras abaixo exibem os conjuntos críticos de $F_{p\ell}$ para cada um desses pares. Em cada quadrante está indicado o sinal do determinante de $DF_{p\ell}$.

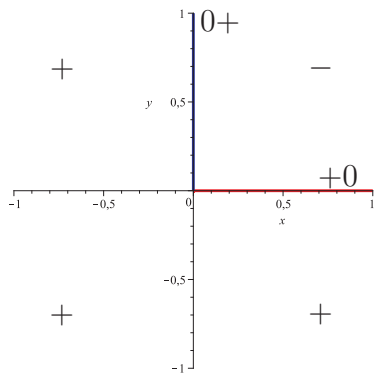


Figura 2.5: $\mathcal{C}(F_{pl})$: $a = 0$ e $b = 3$

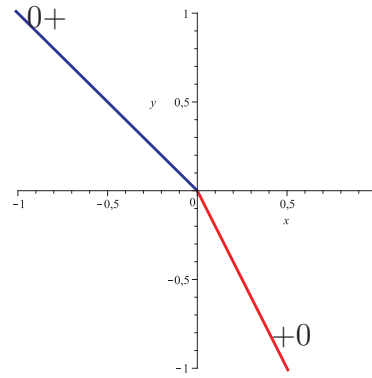


Figura 2.6: $F_{pl}(\mathcal{C}(F_{pl}))$

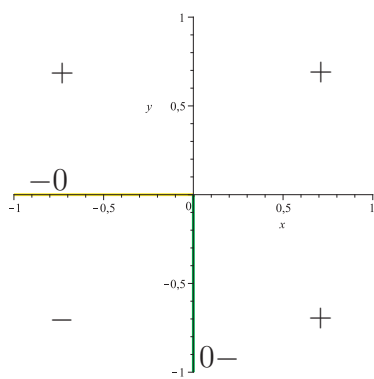


Figura 2.7: $\mathcal{C}(F_{pl})$: $a = 6.5$ e $b = 7$

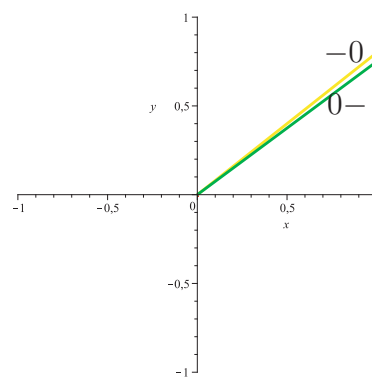


Figura 2.8: $F_{pl}(\mathcal{C}(F_{pl}))$

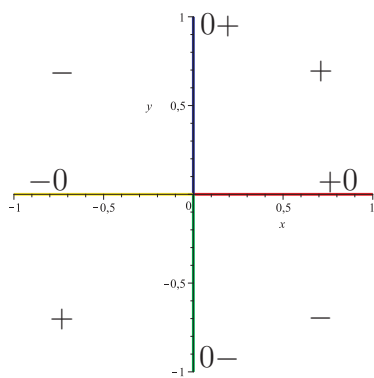


Figura 2.9: $\mathcal{C}(F_{pl})$: $a = 0$ e $b = 7$

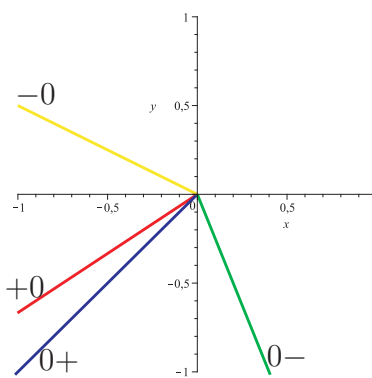


Figura 2.10: $F_{pl}(\mathcal{C}(F_{pl}))$

□

Proposição 2.8 *Seja F_{pl} uma aplicação pl -admissível tal que seus parâmetros assintóticos a e b não são autovalores de A . Suponha que os intervalo (a, b)*

contenha os autovalores $\lambda_i \leq \dots \leq \lambda_j$. Então uma reta $r(t) = y - tp$, $p > 0$, intercepta o conjunto crítico $\mathcal{C}(F_{p\ell})$ em $j - i + 1$ pontos. Em particular, se $i = 1$ e $j = n$, então $\mathcal{C}(F_{p\ell})$ é a união de todos os hiperplanos coordenados.

Demonstração:

Aproximando $f_{p\ell_i}$ por uma função suave e convexa f_{al_i} possuindo os mesmos parâmetros assintóticos de $f_{p\ell_i}$ temos, pela Proposição 2.3, que a reta $r(t) = y - tp$, $p > 0$, intercepta o conjunto crítico $\mathcal{C}(F_{al})$ em $j - i + 1$ pontos onde F_{al} é a aplicação al -admissível $F_{al}(u) = Au - f_{al}$. Para $|t|$ suficientemente grande, os sinais de $\det DF_{p\ell}(u)$ e $\det DF_{al}(u)$ são os mesmos. \square