

Funções do plano no plano

Iaci P. Malta
Nicolau C. Saldanha
Carlos Tomei

Introdução

No primeiro curso de cálculo, um problema fundamental é o traçado do gráfico de uma função — é quase impossível sentir-se à vontade com uma função de \mathbb{R} em \mathbb{R} sem conhecer seu gráfico. Nos cursos de variável complexa, gráficos de funções analíticas não aparecem, escondidos em \mathbb{R}^4 , mas há outras representações visuais do comportamento de uma função que são igualmente relevantes: um aluno deve saber como a exponencial complexa age sobre o plano, ou como tantas outras funções agem sobre subconjuntos especiais do plano — os primeiros problemas de mapeamento conforme.

Entender globalmente o comportamento de uma função é o primeiro passo para perguntas mais difíceis como, por exemplo, calcular todas as imagens inversas de um ponto. Assim, o teorema fundamental da álgebra diz que um polinômio de grau n tem exatamente n raízes complexas. O passo seguinte, se quiséssemos calculá-las com precisão, seria ter uma idéia de onde elas se encontram, para depois refinar seu cálculo usando métodos locais.

O objetivo deste texto é o estudo *global* de uma grande classe de funções F de \mathbb{R}^2 em \mathbb{R}^2 — seu comportamento local é bem conhecido, tanto teórica quanto computacionalmente. Mais precisamente, vamos descrever uma representação visual muito conveniente de F e mostraremos como encontrar numericamente todas as imagens inversas de um ponto q , isto é, resolveremos sistemas não lineares com duas equações e duas incógnitas $F(p) = q$. Note que inicialmente, não sabemos sequer quantas soluções esperar para o sistema: nosso método conta soluções, localiza-as aproximadamente e converte seu cálculo a um problema local. O método é implementado num programa de desempenho bastante satisfatório. Os resultados deste texto foram apresentados originalmente em [MST].

No Capítulo 1, descreveremos os resultados clássicos para o comportamento local de uma função em pontos regulares e em pontos críticos menos degenerados: dobras e cúspides. No Capítulo 2, introduzimos vários ingredientes topológicos empregados no texto (por exemplo, espaços de recobrimento, grau). O método não se aplica a todas as funções do plano no plano, assim como mesmo depois de Cálculo I não desenhamos o gráfico de qualquer função real. No Capítulo 3, definimos as funções para as quais o método se aplica mais diretamente, as funções cordatas, que, apesar das exigências feitas na definição, formam um subconjunto abundante do conjunto de todas as funções suaves do plano no plano e estudamos seu comportamento global. No Capítulo 4, caracterizamos conjuntos críticos de funções cordatas. Para isso, empregamos o teorema de Blank que trata da extensão de imersões de S^1 no plano a imersões do disco no plano, juntamente com o teorema de Troyer, que trata do mesmo problema para o disco com k buracos. Finalmente no Capítulo 5 descrevemos a implementação em máquina do método resultante da nossa abordagem teórica, além de apresentarmos exemplos de funções estudadas em detalhe pelo programa.

Nossos agradecimentos vão a Dan Marchesin, que nos levou a estudar funções cordatas (em [MT]) ao apresentar o primeiro sistema que nos interessou resolver (equações de Rankine-Hugoniot provenientes de uma lei de conservação hiperbólica [SSMP]), a W. Massey, de Yale University, que prestou atenção aos primeiros diagramas, ao Departamento de Informática da PUC-Rio, que nos ofereceu todo o material de computação necessário, ao CNPq e SCT pelo apoio financeiro à pesquisa.

Rio de Janeiro, março de 1993

Iaci P. Malta

Nicolau C. Saldanha

Carlos Tomei

Departamento de Matemática, PUC-Rio

Rua Marquês de São Vicente 225

Gávea, Rio de Janeiro

RJ 22453-900, Brasil

malta@mat.puc-rio.br

nicolau@mat.puc-rio.br, <http://www.mat.puc-rio.br/~nicolau/>

tomei@mat.puc-rio.br

Capítulo 1

Os resultados mais simples no estudo de funções do plano no plano são de natureza local: eles descrevem o comportamento de uma função em uma vizinhança suficientemente pequena de um ponto do domínio. Neste capítulo descreveremos os comportamentos locais mais freqüentes: as formas locais próximas a pontos regulares, dobras e cúspides.

Seção 1.1: Curvas

Vamos recordar algumas definições habituais. Uma função de um aberto de \mathbb{R}^n para \mathbb{R}^m é *suave* sempre que todas as suas derivadas parciais de qualquer ordem estão definidas e são contínuas. Uma função de um aberto de \mathbb{R} para \mathbb{R}^m é *suave por partes* se ela é contínua em todo o domínio, suave se de seu domínio omitirmos um número finito de pontos e em cada um destes pontos existirem as derivadas laterais de qualquer ordem. Um *homeomorfismo* é uma bijeção contínua cuja inversa também é contínua; um *difeomorfismo* é um homeomorfismo suave com inversa suave.

Curvas são objetos fundamentais neste texto e vamos precisar de todo um vocabulário para falar delas. Uma função $\gamma : (a, b) \rightarrow \mathbb{R}^2$ ((a, b) é um intervalo aberto) é *mansa* quando:

- γ é suave por partes,
- exceto num número finito de pontos suaves de γ , a derivada não é nula, e nestes pontos, a derivada segunda γ'' e a derivada terceira γ''' são linearmente independentes (em particular, não nulas),
- nos pontos não suaves, as derivadas laterais são linearmente independentes.

Definimos analogamente funções mansas com domínios $[a, b]$, \mathbb{R} ou \mathbb{S}^1 . As funções $t \mapsto (\cos t, \sin t)$, $t \mapsto (t^2, t^3)$ e $t \mapsto (t, |t|)$ são mansas, mas $t \mapsto (0, |t|)$, $t \mapsto (t^2, t^5)$ e $t \mapsto (t^2, t^2|t|)$, não o são — desenhar as imagens destas funções é um exercício interessante.

Lema 1.1: *Se γ é mansa, então é localmente injetora (isto é, em torno de cada ponto do domínio de γ , existe uma vizinhança na qual γ é injetora).*

Demonstração: Seja t um ponto de (a, b) . Sem perda, $t = 0$.

Suponhamos inicialmente que 0 é um ponto onde γ é suave e no qual, compondo na imagem com uma transformação linear inversível, temos $\gamma'(0) = (1, 0)$. Pelo teorema do valor médio, para $t_1 < t_2$ em uma pequena vizinhança de 0 , a coordenada x de $\gamma(t_1)$ é menor que a coordenada x de $\gamma(t_2)$.

Suponha agora que 0 é um ponto onde γ é suave, mas $\gamma'(0) = 0$. Sem perda, como acima, $\gamma''(0) = (1, 0)$ e $\gamma'''(0) = (0, 1)$. Seja $y(t)$ a segunda coordenada de $\gamma(t)$. Temos $y'(t) = t^2 h(t)$, para alguma função suave $h(t)$ tal que $h(0) = \frac{1}{2}$, por Taylor. Assim, $y(t)$ é estritamente crescente numa vizinhança de $t = 0$.

Suponha finalmente que 0 é um ponto onde γ não é suave. Sem perda, as derivadas laterais são $(1, -1)$ e $(1, 1)$. Novamente, a coordenada $x(t)$ é estritamente crescente perto de 0 . ■

Uma curva *aberta normal* é um subconjunto Γ do plano que é a imagem de uma função mansa $\gamma : (a, b) \rightarrow \mathbb{R}^2$ com as seguintes propriedades:

- se θ_1 e θ_2 são números distintos para os quais $\gamma(\theta_1) = \gamma(\theta_2)$, então γ é suave nestes pontos e as derivadas são vetores linearmente independentes; neste caso, a imagem $\gamma(\theta_1) = \gamma(\theta_2)$ é chamada de um *ponto de auto-interseção* da curva Γ ,
- todo ponto de auto-interseção de $\gamma((a, b))$ tem exatamente duas pré-imagens.

A função mansa γ é uma *parametrização* da curva. Definimos de maneira semelhante curvas normais *com extremos* e curvas normais *fechadas*, cujos domínios naturais de parametrização são um intervalo fechado e o círculo unitário. A função $z \mapsto z^2$, em notação complexa, restrita ao círculo unitário, é mansa e sua imagem é o círculo unitário, sabidamente uma curva normal, mas esta função *não* é uma parametrização

do círculo, por desrespeitar o primeiro item da definição. As funções $z \mapsto z^2 + \bar{z}$ (novamente em notação complexa) nos círculos $|z| = \frac{3}{4}$ e $|z| = 1$ são mansas; a imagem da primeira na Figura 1.1(a) é normal mas a imagem da segunda não: como se vê na Figura 1.1(b), a origem tem três pré-imagens, as raízes cúbicas de -1 . Outro exemplo de função mansa cuja imagem não é uma curva normal é $\gamma(t) = (1 - t^2, t(t+1)^2(t-1)^2)$: temos $\gamma(1) = \gamma(-1) = (0, 0)$, $\gamma'(1) = (-2, 0)$ e $\gamma'(-1) = (2, 0)$ — as derivadas são colineares, como se vê na Figura 1.1(c).

Figura 1.1 (umas curvas)

Um ponto de uma curva é *regular* quando não é de auto-interseção e a parametrização é suave com derivada não nula neste ponto. Um ponto da curva onde a parametrização tem derivada zero é uma *cúspide* e um ponto onde γ não é suave é um *vértice*. Estes conceitos independem da escolha de parametrização. Em pontos regulares a curva tem uma reta tangente bem definida. Uma curva é *regular* quando admite uma parametrização suave com derivada sempre não nula e é *simples* quando não tem pontos de auto-interseção.

Uma *família normal* de curvas é um conjunto finito de curvas normais com as seguintes propriedades:

- nunca existem três curvas com um mesmo ponto em comum,
- curvas só se interceptam em pontos regulares,
- as retas tangentes em pontos de interseção são sempre distintas.

Mais geralmente, interseções entre curvas ocorrendo em pontos regulares, com retas tangentes distintas, são chamadas *transversais*. Neste texto, curvas e famílias de curvas são sempre normais. Folheando o Capítulo 5, você vai encontrar vários exemplos de famílias normais de curvas.

Seção 1.2: Pontos regulares

A função típica neste texto será $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, levando um ponto (x, y) do plano \mathbb{R}^2 a um ponto $(u(x, y), v(x, y))$. A *jacobiana* de F no ponto (x, y) é a matriz

$$DF(x, y) = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}.$$

Toda função F do texto será suave. Uma função $F : A \rightarrow \mathbb{R}^2$ (onde $A \subseteq \mathbb{R}^2$ é um aberto) é uma *imersão* se sua jacobiana é sempre inversível e *preserva orientação* se $\det(DF) > 0$. Um ponto p do domínio de F é *regular* se $DF(p)$ é inversível, caso contrário ele é dito *crítico*.

Teorema 1.2 (da Função Inversa [L1]): *Seja p um ponto regular de $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Então existem abertos U e V contendo respectivamente p e $F(p)$ tais que F é um difeomorfismo entre U e V .*

Um ponto p do domínio é uma *pré-imagem* de um ponto q da imagem se $F(p) = q$. Um ponto na imagem de F é um *valor regular* se todas as suas pré-imagens são pontos regulares. Os conceitos dos dois parágrafos acima se estendem naturalmente a funções de \mathbb{R}^n para \mathbb{R}^n .

Teorema 1.3 (Sard, forma fraca [S]): *Seja $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função suave. Dentro de qualquer aberto não vazio de \mathbb{R}^n existe sempre pelo menos um valor regular de F .*

Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ uma função suave com *gradiente* $\text{grad}f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$. Um número r é um *valor regular* de f se $(\text{grad}f)(p) \neq 0$, para todo ponto p tal que $f(p) = r$. O conjunto das pré-imagens por f de um real s é chamado de um *conjunto de nível*. Frequentemente, conjuntos de nível são curvas regulares, o que segue do teorema da função implícita.

Teorema 1.4 (da Função Implícita [L1]): *Seja (x_0, y_0) um ponto para o qual a segunda coordenada de $\text{grad}f(x_0, y_0)$ não é nula. Então existem abertos $D \subseteq \mathbb{R}$ e $U \subseteq \mathbb{R}^2$ contendo x_0 e (x_0, y_0) respectivamente tais que a interseção do conjunto de nível de f contendo (x_0, y_0) com U é o gráfico de uma função suave de D em \mathbb{R} .*

Uma consequência (não trivial) do teorema acima é a seguinte. Se o gradiente de f é não nulo em todo um conjunto de nível, então este conjunto na verdade é a união de uma família de curvas simples regulares disjuntas (abertas ou fechadas). A dificuldade nesta demonstração está em compatibilizar as várias parametrizações locais, que são fáceis de obter pelo teorema da função implícita, em uma única parametrização para toda a curva: isto é feito, por exemplo, em [Mi]. É fácil ver que o determinante muda de sinal de um lado para o outro da curva.

Seção 1.3: Pontos críticos

Como é o conjunto de pontos críticos de uma função $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$? Pela afirmação acima, ele será uma união de curvas simples regulares disjuntas desde que o gradiente do determinante da jacobiana de F nunca se anule em um ponto crítico de F . Isso de fato acontece para quase todas as funções em um sentido que pode ser tornado preciso ([W]). Assim, neste texto, vamos supor que todas as funções F tenham essa propriedade, em outras palavras, que 0 é um valor regular de $\det(DF)$. Esta hipótese faz com que a jacobiana de F nunca seja a matriz zero: se a jacobiana é nula em algum ponto, o determinante neste ponto tem que ter gradiente nulo, o que segue simplesmente derivando a função $\det(DF)$. Além disto, o conjunto crítico passa a ser um conjunto de curvas regulares, as *curvas críticas de F* — a elas dedicaremos grande parte de nossa atenção.

O teorema da função inversa descreve o comportamento de uma função numa vizinhança de um ponto regular. O que ocorre perto de um ponto crítico? Não tentaremos dar uma resposta que trate de qualquer ponto crítico — em vez disso, vamos estudar apenas os casos menos degenerados (e abundantes!). Estamos ainda nas condições do parágrafo anterior: o gradiente do determinante da jacobiana não se anula nos pontos críticos de F .

Vamos procurar entender o comportamento de F perto de um ponto crítico p do tipo mais abundante. Como 0 é valor regular de $\det(DF)$, a matriz jacobiana em p tem posto um e existe um vetor não nulo tangente à curva crítica. Em geral este vetor tangente não deve estar no núcleo de DF . Esta condição define nosso primeiro tipo de ponto crítico: a *dobra*. Mais precisamente, o ponto crítico p é uma *dobra* se e somente se:

- (1) $\text{grad}(\det DF)(p) \neq 0$,
- (2) o núcleo K de DF e a reta tangente T ao conjunto crítico em p não coincidem.

O item (1) equivale a dizer que zero é um valor regular de $\det(DF)$, para F restrita a algum aberto contendo p . Note que se p satisfaz as condições (1) e (2), outros pontos críticos perto de p também satisfazem as mesmas condições, isto é, dobras formam um aberto do conjunto crítico. Demonstramos no Apêndice 1.A o seguinte teorema, que afirma que todas as dobras são equivalentes a uma dobra padrão.

Teorema 1.5 (forma normal da dobra [W]): *Sejam $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma função suave e p um ponto crítico de F . Seja $F_0(x, y) = (x, y^2)$. As seguintes condições são equivalentes:*

- (a) p é uma dobra para F ,
- (b) existem bolas abertas U e V em torno de p e $F(p)$ respectivamente e difeomorfismos preservando orientação ϕ e ψ de U e V para U_0 e V_0 abertos em \mathbb{R}^2 , com $\phi(p) = 0$ e $\psi(F(p)) = 0$, tais que $F_0 \circ \phi = \psi \circ (F|_U)$.

Usando a forma normal, podemos concluir que, próximo a uma dobra, F se comporta como indicado na Figura 1.2: pontos perto de $F(p)$ de um lado da imagem da curva crítica têm duas pré-imagens próximas de p , enquanto pontos do outro lado não têm nenhuma. Nas figuras, imagens e pré-imagens são designados pela mesma letra.

Figura 1.2 (forma local da dobra)

E se o ponto crítico p for tal que o núcleo da jacobiana nele contiver o vetor tangente á curva crítica em p ? Lembre-se que o gradiente do determinante da jacobiana não se anula sobre a curva crítica e que a jacobiana tem posto um. Parametrizamos regularmente a curva crítica perto de p dando origem a uma função $c(t)$ tomando valores no plano para a qual $c(0) = p$. Definimos $\theta(t)$ como o ângulo entre o vetor

tangente à curva e um vetor no núcleo de DF no ponto $c(t)$, onde o vetor é escolhido de forma que $\theta(0) = 0$. Mais uma vez, é razoável esperar que a derivada $\theta'(0) \neq 0$. Um ponto crítico satisfazendo essas propriedades é uma *cúspide*. Mais precisamente, o ponto crítico p é uma *cúspide* se e somente se:

- (1) $\text{grad}(\det DF)(p) \neq 0$,
- (2) o núcleo K e a reta tangente T ao conjunto crítico em p coincidem,
- (3) se $c : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^2$ é uma parametrização regular de uma vizinhança da curva crítica passando por p com $c(0) = p$ e $\theta(t)$ é o ângulo entre o núcleo $K(t)$ e a reta tangente $T(t)$ ao conjunto crítico em $c(t)$ então $\theta'(0) \neq 0$.

O Teorema abaixo nos dá caracterizações alternativas da cúspide.

Teorema 1.6 (forma normal da cúspide [W]): *Sejam $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma função suave e p um ponto crítico de F . Sejam $F_{\pm}(x, y) = (x, y^3 \pm xy)$. As seguintes condições são equivalentes:*

- (a) p é uma cúspide para F ,
- (b) p satisfaz os itens (1) e (2) da definição acima e temos $(F \circ c)''(0) \neq 0$, onde c é como acima,
- (c) existem bolas abertas U e V em torno de p e $F(p)$ respectivamente e difeomorfismos respeitando orientação ϕ e ψ de U e V para U_0 e V_0 abertos em \mathbb{R}^2 , com $\phi(p) = 0$ e $\psi(F(p)) = 0$, tais que $F_- \circ \phi = \psi \circ (F|_U)$ ou $F_+ \circ \phi = \psi \circ (F|_U)$.

A demonstração deste fato é conceitualmente semelhante àquela do Teorema 1.1 mas, por ser muito mais longa, não será dada integralmente (mas veja o Apêndice 1.A). Usando a forma normal perto de uma cúspide, a função F_- se comporta como na Figura 1.3(a): pontos na região estreita entre os arcos da imagem da curva crítica têm três pré-imagens perto de p , enquanto pontos fora desta região só tem uma. Novamente, nas figuras, imagens e pré-imagens são designados pela mesma letra. Os pontos $F_-(c(t)), t \neq 0$, têm duas pré-imagens, como indicado na figura. Note que $\det(DF_-)$ é positivo nas pré-imagens não críticas de valores críticos, enquanto para F_+ este sinal é negativo. Note também na Figura 1.3(b) o comportamento das imagens das duas curvas suficientemente próximas correndo paralelas à curva crítica.

Figura 1.3(a,b) (forma local das cúspides)

A repetição do nome cúspide para curvas e funções é proposital.

Lema 1.7: *Seja $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma função cujos pontos críticos são todos dobras e cúspides. Então a composição de uma parametrização das curvas críticas (regulares) com F é uma função mansa. A imagem de pequenas vizinhanças de um ponto numa curva crítica é uma curva normal. A imagem de uma dobra é um ponto regular e a imagem de uma cúspide de F é uma cúspide da imagem da curva crítica.*

Demonstração: O resultado segue da forma normal de dobras e cúspides (mas falta provar a invariância da independência linear de $\gamma''(t)$ com $\gamma'''(t)$ por composição com difeomorfismos, dado que $\gamma'(t) = 0$). ■

Decorre do lema, que a imagem de uma curva crítica só deixa de ser uma curva normal por poder ter auto-interseções não transversais.

Vamos classificar todos os pontos críticos da função $F(x, y) = (x^2 - y^2 + x, 2xy - y)$, ou, em notação complexa, $F(z) = z^2 + \bar{z}$. A jacobiana no ponto (x, y) é $\begin{pmatrix} 2x + 1 & -2y \\ 2y & 2x - 1 \end{pmatrix}$, com determinante $4(x^2 + y^2) - 1$, logo o conjunto crítico é um círculo de raio $\frac{1}{2}$ centrado na origem. Note aliás que o gradiente do determinante, $(8x, 8y)$, não se anula sobre o conjunto crítico. O núcleo da jacobiana em um ponto crítico $(\frac{1}{2} \cos t, \frac{1}{2} \sin t)$ é gerado pelo vetor $(\sin(\frac{t}{2}), \cos(\frac{t}{2}))$ (um pouco de trigonometria). Assim os únicos pontos que não são dobras (isto é, aqueles nos quais o núcleo da jacobiana não é tangente ao círculo crítico) são os três pontos $(\frac{1}{2}, 0)$, $(-\frac{1}{4}, \pm \frac{\sqrt{3}}{4})$. Vamos usar a definição original de cúspide para caracterizar estes três pontos. Para a nossa parametrização, o argumento da reta tangente à curva crítica é $\frac{\pi}{2} + t$, enquanto o argumento do núcleo da matriz jacobiana é $\frac{\pi}{2} - \frac{t}{2}$, donde $\theta(t) = \frac{3t}{2} + k\pi$ (onde k é um inteiro que será escolhido de forma diferente em cada candidato a cúspide), com derivada sempre não nula. Assim, os três pontos críticos que não são

dobras são de fato cúspides. O conjunto crítico C de F , sua imagem $F(C)$ e o conjunto $F^{-1}(F(C))$ estão ilustrados na Figura 1.4.

Figura 1.4 ($z^2 + \bar{z}$)

Vejam um outro exemplo, em que utilizaremos o item (b) do Teorema 1.2 para estudar pontos críticos que não são dobras. Seja $F(x, y) = (x^2 + ay^2 + 2by, cy^2 + 2xy)$, $b \neq 0$ e a e c reais arbitrários. Derivando, temos

$$DF(x, y) = 2 \begin{pmatrix} x & ay + b \\ y & cy + x \end{pmatrix}$$

e $\det(DF(x, y)) = 2(x^2 - ay^2 + cxy - by)$, de forma que a origem é um ponto crítico de F . O núcleo de $DF(0, 0)$ é o eixo x e $\text{grad}(\det DF)(0, 0) = (0, -b)$ de forma que a reta tangente à curva crítica também é o eixo x ; assim, a origem não é uma dobra de F . Para verificar se a origem é uma cúspide, use o teorema da função implícita para parametrizar a curva crítica por $c(t) = (t, g(t))$. Derivando implicitamente, temos $2t - 2ag(t)g'(t) + cg(t) + ctg'(t) - bg'(t) = 0$, donde $g'(0) = 0$ e $2 - bg''(0) = 0$. Por outro lado, $(F \circ c)(t) = (t^2 + ag^2(t) + 2bg(t), cg^2(t) + 2tg(t))$, donde $(F \circ c)'(t) = (2t + 2ag'(t)g(t) + 2bg'(t), 2cg(t)g'(t) + 2g(t) + 2tg'(t))$ e $(F \circ c)''(0) = (2 + 2bg''(0), 0) = (6, 0) \neq (0, 0)$; assim, pelo item (b) do teorema da forma normal da cúspide, p é uma cúspide.

As duas *orientações* possíveis para uma curva são indicadas por setas, da maneira óbvia. Uma parametrização determina uma orientação: por exemplo, a orientação determinada pela parametrização $\gamma(t) = (\cos t, \text{sen} t)$ é a *positiva* ou *anti-horária*. Quando falarmos de uma parametrização de uma curva orientada, estaremos supondo que ela é compatível com a orientação.

Dada $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tal que 0 é um valor regular de $\det(DF)$, definimos o *sentido de dobra* de uma curva crítica Γ no domínio como sendo a orientação que deixa os pontos p com $\det(DF(p)) > 0$ imediatamente à esquerda de Γ , como indicado na Figura 1.5. A composição de F com uma parametrização da curva crítica, orientada com sentido de dobra, induz uma orientação para $F(\Gamma)$, o *sentido de dobra* de $F(\Gamma)$, com a seguinte propriedade geométrica: a imagem de uma pequena vizinhança de um ponto de dobra $p \in \Gamma$ está inteiramente à esquerda de $F(\Gamma)$, como na Figura 1.5. Isso segue da forma normal da dobra.

Figura 1.5 (sentidos de dobra)

Exercício: Considere a função $F(r \cos \theta, r \text{sen} \theta) = (g(r) \cos \theta, g(r) \text{sen} \theta)$, onde g é definida pelo gráfico na Figura 1.6. Mostre que o conjunto crítico de F , sua imagem e os sentidos de dobra em cada curva são como indicado na Figura 1.7. Aliás, quantas pré-imagens tem cada ponto da imagem?

Figura 1.6 (a parte radial do cogumelo)

Figura 1.7 (C e $F(C)$ do cogumelo)

Vamos considerar agora um exemplo mais sofisticado. Seja

$$F(x, y) = (-6x^4 - 6x^2y^2 + xy^3 + 6y^4 - x, \frac{25}{24}x^4 + x^3y + x^2y^2 + \frac{1}{6}xy^3 - y^4 - y).$$

O conjunto crítico C de F consiste de uma única curva, esboçada junto com sua imagem na Figura 1.8, onde mostramos ainda o conjunto $F^{-1}(F(C))$. A função tem cinco cúspides, a, b, c, d e e . O sinal da jacobiana é negativo fora de C e positivo dentro. A cúspide e é a única cuja forma local é F_- .

Figura 1.8 (raposinha)

Este exemplo sugere que se interprete a classificação das cúspides pela forma normal de maneira mais geométrica. Informalmente, a presença de cada cúspide é sentida pela ação da função de um lado só da curva: uma pequena vizinhança de p é dividida em dois pela curva crítica — de um lado F é um homeomorfismo e do outro não. Assim, poderíamos chamar as cúspides de tipo F_+ (resp., F_-) de cúspides à direita (resp., esquerda) em relação ao sentido de dobra. Neste exemplo, a, b, c e d são cúspides à *direita* e e à *esquerda* de C . Como C é uma curva fechada, é expressivo entretanto chamar a, b, c e d de cúspides *externas* e e de cúspide *interna*.

É interessante considerar o comportamento da imagem de duas curvas suficientemente próximas e correndo paralelas à curva crítica (por dentro e por fora) em dois exemplos, a função $F_0(z) = z^2 + \bar{z}$ e

F , a função do exemplo acima. As imagens são mostradas nas Figuras 1.9 para F_0 e 1.10 para F ((a) é a imagem da curva de dentro e (b) é a da de fora). Para F_0 , todas as cúspides são externas.

Figura 1.9(a,b)(curvas perto de C, z^2)

Figura 1.10(a,b)(idem, raposinha)

Apêndice 1.A: Caracterizações de dobras e cúspides

No texto, apresentamos duas caracterizações de pontos de dobra: a *caracterização geométrica* corresponde a nossa definição original e a *caracterização por forma normal* corresponde ao item (b) do Teorema 1.5. O objetivo principal deste apêndice é demonstrar o Teorema 1.5, isto é, a equivalência das duas caracterizações. Para cúspides, apresentamos três caracterizações, correspondentes aos itens (a), (b) e (c) do Teorema 1.6. Mostraremos a equivalência entre (a), (b) e uma quarta caracterização mais técnica.

Como sempre, $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ é uma função suave; seja p um ponto crítico do domínio.

Lema 1.8: *As duas caracterizações para dobras e as três para cúspides são invariantes por composições com difeomorfismos locais no domínio e na imagem. Mais precisamente, sejam α e β difeomorfismos locais de abertos em torno de p e $F(p)$: p satisfaz uma caracterização para F se e somente se $\alpha^{-1}(p)$ satisfaz a mesma caracterização para $\beta \circ F \circ \alpha$.*

Demonstração: Para a caracterização por forma normal quer da dobra quanto da cúspide, o resultado é evidente. Vamos mostrar agora que se p é tal que $\text{grad}(\det DF)(p) \neq 0$ então $\alpha^{-1}(p)$ satisfaz a mesma condição para $\beta \circ F \circ \alpha$: por simetria, isto verifica a invariância do item (1) de todas as caracterizações restantes.

Seja $A = (\text{grad det } DF)(p) \neq 0$; note que $(\det DF)(p) = 0$. Vamos ver que $B = (\text{grad det } D(\beta \circ F \circ \alpha))(\alpha^{-1}(p)) \neq 0$. Para calcular B , usamos a regra da cadeia e o fato que o determinante de um produto de matrizes é o produto de seus determinantes:

$$\det D(\beta \circ F \circ \alpha)(\alpha^{-1}(p)) = \det D\beta(*) \det DF(p) \det D\alpha(*),$$

onde os asteriscos denotam expressões dependendo de p . Tomando o gradiente em p ,

$$B = \det D\beta(*) \cdot (D\alpha(*))^T A \cdot \det D\alpha(*),$$

já que $(\det DF)(p) = 0$ (a notação $(D\alpha(*))^T$ representa a transposta de $D\alpha(*)$). Temos então que B é diferente de zero, por ser o produto de dois números não nulos (lembre que α e β são difeomorfismos) pelo vetor $(D\alpha(*))^T A$, também não nulo pois A é diferente de zero e $D\alpha(*)$ é inversível.

Contas análogas mostram que $D(\beta \circ F \circ \alpha)$ é um produto de matrizes $X \cdot DF \cdot Y$, onde X e Y são inversíveis. É fácil ver que α^{-1} leva o conjunto crítico de F ao de $\beta \circ F \circ \alpha$ e que $D(\alpha^{-1})$ leva as retas K e T para F em seus análogos para $\beta \circ F \circ \alpha$, mostrando assim a invariância do item (2) de todas as caracterizações.

Vamos mostrar agora a invariância do item (3) da definição original de cúspide. Sejam $c(t)$ uma parametrização regular de uma vizinhança da curva crítica perto de p com $c(0) = p$, e $K(t)$ e $T(t)$ o núcleo da jacobiana e o espaço tangente à curva em $c(t)$. Sejam também $v(t) \in K(t)$ e $w(t) \in T(t)$ não nulos. Note que o determinante $A(t)$ da matriz cujas colunas são $w(t)$ e $v(t)$ é igual a $|w(t)||v(t)|\text{sen}(\theta(t))$, cuja derivada para $t = 0$ (e aí $\theta(0) = 0$, pela definição de cúspide) vale $A'(0) = |w(0)||v(0)|\theta'(0)$, implicando $A'(0) \neq 0$ se e somente se $\theta'(0) \neq 0$ — em particular, a escolha dos vetores $v(t)$ e $w(t)$ (desde que não nulos) não é relevante. A condição $A'(0) \neq 0$ é certamente invariante por composição por difeomorfismos à direita e à esquerda.

Finalmente, a invariância do item (3') é uma simples aplicação da regra da cadeia e fica como um exercício. ■

Nas demonstrações, F vai sendo composta com difeomorfismos à direita e à esquerda (sem perda das propriedades de cada caracterização, pelo lema), visando representações mais simples. As considerações a seguir são comuns ao tratamento de dobras e cúspides.

Seja p um ponto crítico de uma função suave F e suponha que p satisfaz $\text{grad}(\det DF)(p) \neq 0$. Sem perda, $p = F(p) = (0, 0)$ e $DF(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$: isto se faz compondo F à direita e à esquerda com transformações afins inversíveis (lembre-se que $DF(p)$ tem posto 1). Escreva $F(x, y) = (u(x, y), v(x, y))$. Considere $\zeta(x, y) = (u(x, y), y)$: sua derivada na origem é a identidade, logo, pelo teorema da função inversa, ζ é um difeomorfismo numa vizinhança da origem, e podemos substituir F por $\tilde{F} = F \circ \zeta^{-1}$, para a qual vale $\tilde{F}(x, y) = (x, h(x, y))$, com o mesmo comportamento de F na origem.

O conjunto crítico de \tilde{F} é descrito pela equação $h_y(x, y) = 0$: em particular, $h_y(0, 0) = 0$. Lembre-se que $\text{grad}(\det DF) \neq 0$: assim, $(h_{xy}(0, 0), h_{yy}(0, 0)) \neq 0$. O núcleo K em qualquer ponto crítico é o eixo vertical e a reta tangente T ao conjunto crítico é gerada pelo vetor $(h_{yy}(x, y), -h_{xy}(x, y))$. Assim, p é uma dobra se e somente se $h_{yy}(0, 0) \neq 0$.

Vejamos como ficam as caracterizações de cúspide com estas coordenadas. Usando (2) temos $h_{yy}(0, 0) = 0$ e portanto $h_{xy}(0, 0) \neq 0$. Pelo Teorema da Função Implícita existe uma função suave η com $h_y(\eta(y), y) = 0$. Daí, $c(t) = (\eta(t), t)$ é uma parametrização regular do conjunto crítico. Note que $\eta'(0) = 0$, derivando em y a equação acima.

Com esta parametrização, temos $\tan(\theta(t)) = \eta'(t)$ (lembre-se que o núcleo é sempre uma reta vertical) e $\theta'(0) = \eta''(0)$. Assim, a origem é uma cúspide segundo nossa definição original se e somente se $\eta''(0) \neq 0$. Por outro lado, $(F \circ c)(t) = (\eta(t), h(\eta(t), t))$, $(F \circ c)'(t) = (\eta'(t), \eta'(t)h_x(\eta(t), t) + h_y(\eta(t), t))$ e $(F \circ c)''(0) = (\eta''(0), 0)$. Assim, também a condição (3') é equivalente a $\eta''(0) \neq 0$, e demonstramos a equivalência entre (a) e (b) no Teorema 1.2.

Derivando implicitamente, temos $\eta'(y)h_{xy}(\eta(y), y) + h_{yy}(\eta(y), y) = 0$ e $\eta''(0)h_{xy}(0, 0) + h_{yyy}(0, 0) = 0$. Assim, a origem é uma cúspide se e somente se $h_{yyy}(0, 0) \neq 0$.

Resumindo, para uma função da forma $(x, y) \mapsto (x, h(x, y))$ a origem é uma dobra se e somente se $h_{yy}(0, 0) \neq 0$ e é uma cúspide se e somente se $h_{xy}(0, 0) \neq 0$, $h_{yy}(0, 0) = 0$ e $h_{yyy}(0, 0) \neq 0$. Em [W] é dada a demonstração de que estas condições implicam as respectivas formas normais; daremos aqui a demonstração para dobras.

Demonstração do Teorema 1.5: Pelo lema, um ponto que satisfaz a caracterização por forma normal da dobra (isto é, o item (b)) necessariamente satisfaz a caracterização geométrica (ou seja, o item (a)). Falta agora apenas demonstrar que (a) implica (b).

Pelo teorema da função implícita, então, podemos definir uma função η em uma vizinhança da origem tal que $h_y(x, \eta(x)) = 0$: em outras palavras, o gráfico de η é o conjunto crítico de H . Vamos compor H com o difeomorfismo $(x, y) \mapsto (x, y + \eta(x))$ no domínio e $(X, Y) \mapsto (X, Y - h(X, \eta(X)))$ na imagem, obtendo assim uma nova função \tilde{H} para a qual

- a origem é uma dobra de \tilde{H} segundo a caracterização geométrica,
- $\tilde{H}(0, 0) = (0, 0)$, $D\tilde{H}(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$,
- $\tilde{H}(x, y) = (x, \tilde{h}(x, y))$,
- $\tilde{h}(x, 0) = 0$ e $\tilde{h}_y(x, 0) = 0$,
- $\tilde{h}_{yy}(x, y) \neq 0$, numa vizinhança da origem,
- o conjunto crítico de \tilde{H} é o eixo horizontal e sua imagem também é o eixo horizontal.

Compondo se necessário com $-\text{Id}$ no domínio e imagem, podemos supor que $\tilde{h}_{yy}(x, y) > 0$ em uma vizinhança da origem. Note então que retas verticais vão em verticais por \tilde{H} . Falta agora endireitar \tilde{H}

em cada vertical: uma família unidimensional de problemas unidimensionais. Pelo comportamento de \tilde{h} em $(x, 0)$, podemos escrever

$$\tilde{h}(x, y) = y^2 g(x, y) = \left(y \sqrt{g(x, y)} \right)^2,$$

onde g é suave (pela fórmula de Taylor) e positiva. Agora, basta compor \tilde{H} no domínio com o difeomorfismo $(x, y) \mapsto (x, y\sqrt{g(x, y)})$. ■

Capítulo 2

Neste capítulo, apresentamos as ferramentas topológicas básicas que serão empregadas no texto. Procuramos tornar o texto razoavelmente auto-suficiente, mas utilizamos alguns resultados clássicos que apresentamos sem demonstração.

Seção 2.1: Recobrimentos e grau topológico em \mathbb{S}^1

Um subconjunto X de \mathbb{R}^n é *conexo por caminhos* se, dados dois pontos \mathbf{x} e \mathbf{y} de X , existe $\gamma : [0, 1] \rightarrow X$ contínua com $\gamma(0) = \mathbf{x}$ e $\gamma(1) = \mathbf{y}$ — γ é um *caminho* entre \mathbf{x} e \mathbf{y} . Um subconjunto X de \mathbb{R}^n é *conexo* se, dados abertos disjuntos A e B com $X \subseteq A \cup B$ tem-se que $X \subseteq A$ ou $X \subseteq B$. Para subconjuntos abertos de \mathbb{R}^n , as duas definições acima são equivalentes ([L1]).

Sejam E e X subconjuntos conexos por caminhos de \mathbb{R}^n com a topologia induzida, e $\pi : E \rightarrow X$ uma sobrejeção contínua. Dizemos que E é um *espaço de recobrimento* de X por π , ou que $\pi : E \rightarrow X$ é um *recobrimento*, se para todo ponto $x \in X$ existir uma vizinhança V de x tal que π restrita a cada componente conexa de $\pi^{-1}(V)$ é um homeomorfismo. Referências para espaços de recobrimentos são [L2] e [Ma], onde podem ser encontradas todas as demonstrações omitidas nesta seção.

Um primeiro exemplo unidimensional é o recobrimento do círculo pela reta. Mais precisamente, seja \mathbb{S}^1 o círculo unitário, isto é, o conjunto de pontos (x, y) no plano para os quais $x^2 + y^2 = 1$. Seja $\Pi : \mathbb{R} \rightarrow \mathbb{S}^1$ o *recobrimento canônico* de \mathbb{S}^1 por \mathbb{R} definido por $t \mapsto \exp 2\pi it$, onde i é o número complexo habitual e estamos fazendo a identificação costumeira entre os números complexos e o plano. Note que se $\Pi(t) = p$, as pré-imagens de p são exatamente os pontos da forma $t + k$, k inteiro.

Um outro exemplo é a função $F : \mathbb{C} - \{0\} \rightarrow \mathbb{C} - \{0\}$ dada por $F(z) = z^n$, onde n é um inteiro positivo (\mathbb{C} é o plano complexo). Neste caso, as pré-imagens de cada ponto são suas n n -ésimas raízes. Um outro exemplo é a função $\exp : \mathbb{C} \rightarrow \mathbb{C} - \{0\}$, onde cada ponto da imagem tem infinitas pré-imagens.

Se E recobre X via π , e temos uma função contínua $g : Z \rightarrow X$, chamaremos de *levantamento* de g a uma função contínua $\hat{g} : Z \rightarrow E$ satisfazendo $g = \pi \circ \hat{g}$.

Teorema 2.1 (do levantamento de caminhos): *Sejam $Z = [0, 1]$ ou $Z = \mathbb{R}$ e E, X, π e g como acima. Sejam $z_0 \in Z, e_0 \in E$ tais que $\pi(e_0) = g(z_0)$. Então existe um único levantamento \hat{g} de g com $\hat{g}(z_0) = e_0$.*

Este resultado se generaliza para qualquer Z simplesmente conexo (esse conceito vai ser definido mais adiante).

Teorema 2.2 (da ordem de um recobrimento): *O número de pré-imagens de um ponto qualquer $x \in X$ pelo recobrimento $\pi : E \rightarrow X$ independe de x .*

Demonstração: Sejam x e x' em X . Ligue os dois pontos por um caminho α , e utilize o Teorema 2.1 para construir caminhos ligando pré-imagens de x a pré-imagens de x' , obtendo assim uma bijeção entre os dois conjuntos de pré-imagens. ■

Este número de pré-imagens é a *ordem* do recobrimento. A técnica de levantar caminhos, que utilizamos pela primeira vez na demonstração acima, será usada muitas vezes no texto, para fins teóricos e numéricos.

Sejam X e Y subconjuntos de \mathbb{R}^n com a topologia induzida. Uma *homotopia* entre duas funções contínuas f e g de X em Y é uma função contínua $H : X \times [0, 1] \rightarrow Y$ tal que $H(\mathbf{x}, 0) = f(\mathbf{x})$ e $H(\mathbf{x}, 1) = g(\mathbf{x})$, para todo $\mathbf{x} \in X$. Um conjunto X é *simplesmente conexo* se toda função contínua $f : \mathbb{S}^1 \rightarrow X$ é homotópica (em X !) a uma função constante. Por exemplo, o disco unitário em \mathbb{R}^2 é simplesmente conexo, já que toda função tomando valores no disco pode ter seus valores multiplicados por escalares cada vez menores até se tornar constante igual a zero. Analogamente, $\mathbb{R}, \mathbb{R} \times [0, 1]$ e \mathbb{R}^2 são simplesmente conexos. É um pouco mais difícil ver que a esfera unitária em \mathbb{R}^3 também é simplesmente conexa.

Teorema 2.3: *Seja E um recobrimento de um conjunto simplesmente conexo X por uma função π . Então π é um homeomorfismo.*

Seja $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ uma função contínua. Dizemos que uma função $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ é uma *suspensão* de f se \tilde{f} é contínua e $\Pi \circ \tilde{f} = f \circ \Pi$, onde Π é o recobrimento canônico de \mathbb{S}^1 por \mathbb{R} . Ou seja, \tilde{f} é um levantamento de $f \circ \Pi$. Observe que duas suspensões diferem por um número inteiro.

A única dificuldade para calcular o valor de uma suspensão \tilde{f} em um ponto vem da existência de várias pré-imagens por Π . Entretanto, por continuidade, uma vez que sabemos o valor de \tilde{f} em um ponto, os valores numa vizinhança pequena desse ponto estão determinados. Quão pequena deve ser essa vizinhança? Considere pontos x_0 e x_1 de \mathbb{R} contidos num intervalo de tamanho 1, com imagens por Π dadas pelos pontos p_0 e p_1 tais que a imagem por f do arco p_0p_1 subentende um ângulo entre $-\pi$ e π . Se α é o ângulo subentendido pelo arco $f(p_0)f(p_1)$ contido em $f(p_0p_1)$, tomamos α entre $-\pi$ e π e para uma suspensão qualquer \tilde{f} temos $\tilde{f}(x_0) - \tilde{f}(x_1) = \alpha/2\pi$. Em particular, levando em conta as distâncias entre as pré-imagens por Π de um ponto $f(p_1)$, é fácil ver que se o arco p_0p_1 satisfaz a propriedade acima, então a partir de $\tilde{f}(x_0)$, fica determinado $\tilde{f}(x_1)$ na pré-imagem x_1 de p_1 que dista menos do que um de x_0 . Em resumo, \tilde{f} num ponto x_1 pode ser calculado a partir de um valor inicial $\tilde{f}(x_0)$, obtendo os valores do levantamento para vários pontos intermediários entre x_0 e x_1 suficientemente próximos. Essa construção é praticamente uma demonstração da existência de um levantamento da função $f \circ \Pi$.

Note que uma suspensão \tilde{f} satisfaz $\tilde{f}(x+1) = \tilde{f}(x) + k$ para algum inteiro k que independe de $x \in \mathbb{R}$. Pelo teorema acima, esse inteiro só depende de f e não da suspensão utilizada, e é chamado de *grau topológico* de f , ou $\text{grau}(f)$. Do parágrafo anterior, obtemos uma maneira concreta de calcular o grau: escolha um p qualquer no domínio de f , considere uma pré-imagem x de p por Π , e a partir do valor de $\tilde{f}(x)$ (por sua vez, uma pré-imagem qualquer de $f(p)$ por Π), obtemos o valor de $\tilde{f}(x+1)$. Uma maneira alternativa de calcular o grau de uma função é a seguinte. Distribua pontos $p_0, p_1, \dots, p_n = p_0$ no sentido anti-horário pelo domínio de f de modo que a imagem do arco $p_i p_{i+1}$ por f é um arco de comprimento menor que π . A variação de argumento dos vetores do plano $f(p_i)$ e $f(p_{i+1})$ determina então um ângulo estritamente entre $-\pi$ e π . O grau de f é a soma total das variações dividida por 2π : verifique! Diremos que o grau topológico calculado dessa maneira é obtido por *aproximação discreta* de f .

Vamos considerar um exemplo importante e simples, a função f de \mathbb{S}^1 em \mathbb{S}^1 , que leva $(\cos 2\pi t, \text{sen} 2\pi t)$ a $(\cos 2\pi nt, \text{sen} 2\pi nt)$. É fácil ver que sua suspensão $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ é dada por $\tilde{f}(t) = nt$, e assim seu grau é igual a n .

Para funções suaves, o grau topológico pode ser definido por uma *caracterização diferencial*. Seja q um valor regular de $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$, isto é, um ponto da imagem de f com um número finito de pré-imagens p_1, \dots, p_n nas quais a derivada de f não se anula (isto é, a derivada de \tilde{f} não se anula nas pré-imagens por Π de p_1, \dots, p_n). Então temos

$$\text{grau}(f) = \sum_{i=1, \dots, n} \text{sgn}(f'(p_i))$$

(de novo, o sinal na expressão acima é o da derivada de \tilde{f} em qualquer pré-imagem por Π de p_i). Este fato é evidente ao considerarmos a Figura 2.1. A existência de um valor regular para funções suaves segue do Teorema 1.3 ($n = 1$).

Figura 2.1 (um levantamento unidimensional)

Na Figura 2.1, damos o gráfico de uma suspensão \tilde{f} de f restrita ao domínio $[x, x+1]$. O ponto y é tal que $\Pi(y) = q$, de forma que as outras pré-imagens por Π de q são da forma $y + \text{inteiro}$. Os pontos x_i são tais que $\Pi(x_i) = p_i$. Considere a variação de $[\tilde{f}(t) - y]$, onde os colchetes denotam a parte inteira de um número. Este valor muda exatamente nos pontos x_i com uma variação de $\text{sgn}(f'(p_i))$: o resultado agora é imediato. Uma conseqüência da equivalência das duas definições é que a segunda independe da escolha de q .

Exercício: Use a fórmula acima para calcular novamente o grau da função $f(\cos 2\pi t, \text{sen} 2\pi t) = (\cos 2\pi nt, \text{sen} 2\pi nt)$.

Sejam f e g funções de \mathbb{S}^1 a \mathbb{S}^1 homotópicas. Seja $H : \mathbb{S}^1 \times [0, 1] \rightarrow \mathbb{S}^1$ uma homotopia correspondente. Seja $\mathcal{H} : \mathbb{R} \times [0, 1] \rightarrow \mathbb{S}^1$, dada por $\mathcal{H}(s, t) = H(\Pi(s), t)$. Como já mencionamos após o Teorema 2.1, qualquer

função com domínio simplesmente conexo admite um levantamento. Seja portanto \hat{H} um levantamento de \mathcal{H} , que chamaremos de \tilde{H} , pois para cada t fixo, $\tilde{H}(\cdot, t)$ é uma suspensão de $H(\cdot, t)$. O grau topológico das funções $H(\cdot, t)$ é dado por $\tilde{H}(s+1, t) - \tilde{H}(s, t)$, que é inteiro, logo constante. Demonstramos assim a proposição abaixo.

Proposição 2.4:

O grau topológico é invariante por homotopias, isto é, duas funções homotópicas f e g de \mathbb{S}^1 em \mathbb{S}^1 têm o mesmo grau.

Uma conseqüência desta proposição é que \mathbb{S}^1 não é simplesmente conexo: a identidade tem grau um, logo não pode ser homotópica a uma função constante, que tem grau zero.

Seja $f : [a, b] \rightarrow \mathbb{S}^1$ uma função contínua com $f(a) = f(b)$. Obviamente, existe uma única função contínua $\check{f} : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ satisfazendo $\check{f}(\Pi((x-a)/(b-a))) = f(x)$. Isto nos diz que podemos interpretar f como uma função \check{f} de \mathbb{S}^1 em \mathbb{S}^1 . Em particular, podemos nos referir ao grau de f , $\text{grau}(f) = \text{grau}(\check{f})$.

Quando $f(a) = f(b) = f(c)$, já vimos como interpretar f como uma função de \mathbb{S}^1 em \mathbb{S}^1 . Podemos também, entretanto, interpretar $f|_{[a,c]}$ ou $f|_{[c,b]}$ como funções de \mathbb{S}^1 em \mathbb{S}^1 .

A partir de agora, não distinguiremos as duas funções f e \check{f} , e passaremos livremente de um para outro ponto de vista.

Um homeomorfismo de \mathbb{S}^1 em \mathbb{S}^1 *preserva orientação* se ao percorrermos o domínio em um sentido, a imagem é percorrida no mesmo sentido; caso contrário, diremos que o homeomorfismo *reverte orientação*.

Na proposição a seguir estão listadas algumas propriedades do grau que usaremos neste texto.

Proposição 2.5:

- (a) *Sejam f e g funções de \mathbb{S}^1 em \mathbb{S}^1 . Então $\text{grau}(f \circ g) = \text{grau}(f)\text{grau}(g)$.*
- (b) *Seja $f : [a, b] \rightarrow \mathbb{S}^1$ com $f(a) = f(c) = f(b)$. Então $\text{grau}(f|_{[a,c]}) + \text{grau}(f|_{[c,b]}) = \text{grau}(f)$.*
- (c) *Todo homeomorfismo que preserva orientação de \mathbb{S}^1 em \mathbb{S}^1 tem grau 1. Homeomorfismos que revertem orientação têm grau -1 .*

Demonstração: Para o item (a), observe que a suspensão de $f \circ g$ é a composição das suspensões de f e de g . Se f tem grau n e g tem grau m , temos $\tilde{g}(x+1) = \tilde{g}(x) + m$ e $\tilde{f}(y+k) = \tilde{f}(y) + kn$. Portanto, $\tilde{f} \circ \tilde{g}(x+1) - \tilde{f} \circ \tilde{g}(x) = mn$, que é o grau de $f \circ g$.

O item (b) decorre trivialmente do fato que o gráfico de uma suspensão de f é obtida justapondo os gráficos de suspensões de $f|_{[a,c]}$ e $f|_{[c,b]}$, como na Figura 2.2.

Figura 2.2 (justapondo suspensões)

Para o item (c), considere uma suspensão \tilde{f} de um homeomorfismo f que preserve orientação. A função \tilde{f} é estritamente crescente e satisfaz $\tilde{f}(x+1) = \tilde{f}(x) + n$, onde n é o grau de f . Se $n > 1$, existe y entre x e $x+1$ com $\tilde{f}(y) = \tilde{f}(x) + 1$, contrariando a injetividade de f . O outro caso é análogo. ■

Decorre trivialmente desta proposição que se $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ é contínua e $\eta : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ é um homeomorfismo que preserva orientação, então $\text{grau}(f \circ \eta) = \text{grau}(f)$.

Seção 2.2: Voltas e gições

Pelo famoso Teorema da curva de Jordan ([Mu]), uma curva fechada simples decompõe o plano em dois conjuntos, o lado de dentro e o lado de fora da curva. Um *disco* é um subconjunto fechado do plano limitado por uma curva normal fechada simples e um *anel* é um subconjunto fechado do plano limitado por duas curvas normais fechadas simples, uma dentro da outra. O disco D' é *equivalente* a D se existir um homeomorfismo $\phi : D' \rightarrow D$, que é um difeomorfismo entre os interiores de D' e D . De forma um pouco mais geral, um *disco com k buracos* é um subconjunto do plano limitado por $k+1$ curvas normais fechadas simples $\Gamma_0, \dots, \Gamma_k$, onde todas as curvas $\Gamma_i, i \neq 0$ estão dentro de Γ_0 , mas nunca dentro uma da outra. Note

que um disco e um anel são casos particulares desta definição. Dois discos com k buracos são equivalentes se existir um homeomorfismo entre os dois que é um difeomorfismo em seus interiores. Vamos precisar do seguinte resultado técnico.

Teorema 2.6: *Todos os discos com k buracos ($k = 0, 1, \dots$) são equivalentes.*

Em [N], é dada uma classificação de discos com k buracos por um tipo especial de difeomorfismo, as transformações conformes. É fácil a partir deste resultado demonstrar o teorema acima. Em particular, qualquer disco com um buraco é equivalente ao anel entre dois círculos concêntricos.

Vamos definir o conceito de *número de voltas* $w(\gamma, q)$ de uma função contínua $\gamma : \mathbb{S}^1 \rightarrow \mathbb{R}^2$ em torno de um ponto q não pertencente à imagem de γ : $w(\gamma, q)$ é o grau da função (contínua) $\vartheta(\theta) = (\gamma(\theta) - q)/|\gamma(\theta) - q|$ de \mathbb{S}^1 em \mathbb{S}^1 . Seja $\eta : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ um homeomorfismo de grau 1. Então ϑ para $\gamma \circ \eta$ é obtida compondo a função ϑ para γ com η , logo pela observação no final da seção anterior, $w(\gamma \circ \eta, q) = w(\gamma, q)$. Em particular, se Γ é uma curva normal fechada, o número de voltas em torno de um ponto de uma parametrização γ de Γ só depende da orientação determinada por γ . Por isso, definimos o número de voltas de uma curva orientada em torno de um ponto como o número de voltas de uma parametrização da curva orientada.

O número de voltas das curvas em torno dos pontos A, B, C, D e E na Figura 2.3 são respectivamente 0, 1, 0, 1 e 2.

Figura 2.3 (triângulo e pentágono estrelado, cinco letras)

Proposição 2.7: *O número de voltas satisfaz as propriedades abaixo.*

(a) *Seja $\gamma : [a, b] \rightarrow \mathbb{R}^2 - \{q\}$ uma função contínua e $c \in (a, b)$ tal que $\gamma(a) = \gamma(c) = \gamma(b)$. Temos*

$$w(\gamma, q) = w(\gamma|_{[a, c]}, q) + w(\gamma|_{[c, b]}, q).$$

(b) *Seja A um disco com k buracos com fronteira externa Γ_0 e fronteiras internas Γ_i , todas orientadas no sentido anti-horário. Sejam γ_i uma parametrização de Γ_i para cada i . Seja $\phi : A \rightarrow \mathbb{R}^2 - \{q\}$ uma função contínua. Então*

$$w(\phi \circ \gamma_0, q) = \sum_{1 \leq i \leq k} w(\phi \circ \gamma_i, q).$$

(c) *Sejam dois pontos q e q' na mesma componente conexa do complemento da imagem de γ . Então $w(\gamma, q) = w(\gamma, q')$.*

Demonstração: O item (a) segue diretamente do item (b) da Proposição 2.5.

Considere duas funções γ_1 e γ_2 que admitem uma homotopia H tal que o ponto q não está na imagem de H . Então $w(\gamma_1, q) = w(\gamma_2, q)$, pela Proposição 2.4 aplicada à função ϑ acima. Da mesma forma, sejam D um disco, γ uma parametrização de ∂D e $\phi : D \rightarrow \mathbb{R}^2 - \{q\}$ contínua. Use o Teorema 2.6 (com $k = 0$) para obter um difeomorfismo $\psi : D \rightarrow D_0$ preservando orientação, onde D_0 é o disco unitário e seja $\phi_0 : D_0 \rightarrow \mathbb{R}^2 - \{q\}$ com $\phi = \phi_0 \circ \psi$. A função $\phi_0 \circ \gamma$ é um homeomorfismo de grau 1 de \mathbb{S}^1 em \mathbb{S}^1 , e assim temos $w(\phi \circ \gamma, q) = w(\phi_0 \circ \psi \circ \gamma, q) = w(\phi_0|_{\mathbb{S}^1}, q)$. Seja $H : \mathbb{S}^1 \times [0, 1] \rightarrow \mathbb{R}^2 - \{q\}$ definida por $H(p, t) = \phi_0(tp)$: H é uma homotopia de $\phi_0|_{\mathbb{S}^1}$ a uma função constante, logo $w(\phi \circ \gamma, q) = 0$.

Estamos agora prontos para demonstrar (b). Use novamente o Teorema 2.6 para supor, sem perda, que A é como indicado na Figura 2.3 e que existe uma homotopia entre a curva exterior Γ_0 e a curva Γ na Figura 2.3. Pela invariância de w por homotopia, $w(\phi \circ \gamma, q) = w(\phi \circ \gamma_0, q)$, onde γ é uma parametrização de Γ . Por (a), $w(\phi \circ \gamma, q) = \sum_{1 \leq i \leq k} w(\phi \circ \gamma_i, q) + \sum_{1 \leq j \leq k-1} w(\phi \circ \beta_j, q)$, onde β_j são parametrizações das curvas fechadas B_j mostradas na Figura 2.4. Pelo parágrafo anterior, $w(\phi \circ \beta_j, q) = 0$, já que B_j é o bordo de um disco contido em A .

O item (c) também segue da Proposição 2.4 observando que ao variarmos continuamente de q a q' , sem passar por Γ , o ponto em relação ao qual w é calculado, a função ϑ descreve uma homotopia, mantendo portanto seu grau. ■

Figura 2.4 (um disco esburacado com uma curva complicada)

Das proposições acima, o número de voltas de uma curva normal fechada simples Γ ao redor de um ponto p do lado de dentro dela é ± 1 , dependendo da curva estar orientada no sentido anti-horário ou horário. Para ver isto, traçamos ao redor de p um pequeno círculo Γ' , orientado no mesmo sentido que Γ . Pelo item (b) da proposição acima ($k = 1$), $w(\Gamma, p) = w(\Gamma', p)$ (note que estamos fazendo o abuso já anunciado de, nesta situação, confundir uma curva orientada com sua parametrização) — para Γ' , o cálculo do número de voltas é trivial. O número de voltas da curva ao redor de um ponto do lado de fora dela é zero: use o mesmo item (b) com $k = 0$.

Vamos agora definir a *giração* de uma função mansa $\gamma : \mathbb{S}^1 \rightarrow \mathbb{R}^2$. Para uma γ suave com derivada não nula, a giração será o número de voltas do vetor unitário tangente à imagem de γ , isto é, o grau da função de \mathbb{S}^1 em \mathbb{S}^1 dada por $\gamma'(t)/|\gamma'(t)|$. Veremos no Lema 2.8 que quando γ é a parametrização de uma curva normal Γ , sua giração só depende da orientação de Γ : isto nos permite definir a giração de uma curva orientada. Assim, percorrer um círculo uma única vez no sentido anti-horário vai corresponder a uma giração igual a um, e no sentido horário a uma giração igual a menos um, e percorrer uma figura oito (em qualquer sentido) corresponderá a uma giração igual a zero. Para as curvas (não regulares) na Figura 2.2, percorridas no sentido indicado, a giração é respectivamente um e dois.

Em geral, seja $\gamma : \mathbb{S}^1 \rightarrow \mathbb{R}^2$ uma função mansa. Pelo Lema 1.1, para cada ponto $p \in \mathbb{S}^1$ existe um δ positivo para o qual γ é injetora no intervalo $(p - \delta, p + \delta)$. Pela compacidade de \mathbb{S}^1 , podemos escolher um δ comum a todos os pontos θ . Seja

$$\rho(p) = (\gamma(p + \delta) - \gamma(p))/|\gamma(p + \delta) - \gamma(p)|.$$

A *giração* $\tau(\gamma)$ da função γ é o grau topológico da aplicação ρ .

Lema 2.8: *A giração está bem definida, isto é, independe da escolha de δ , para δ suficientemente pequeno. Além disso, se $\gamma : \mathbb{S}^1 \rightarrow \mathbb{R}^2$ é mansa e $\eta : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ é um difeomorfismo preservando orientação, temos $\tau(\gamma \circ \eta) = \tau(\gamma)$.*

Demonstração: Seja $\Delta : \mathbb{S}^1 \rightarrow \mathbb{R}^2$ contínua com $0 < \Delta(p) < \delta$, para todo $p \in \mathbb{S}^1$: temos $\gamma(p + \Delta(p)) \neq \gamma(p)$. Podemos então considerar a função contínua $\rho_\Delta : \mathbb{S}^1 \rightarrow \mathbb{S}^1$,

$$\rho_\Delta(p) = (\gamma(p + \Delta(p)) - \gamma(p))/|\gamma(p + \Delta(p)) - \gamma(p)|.$$

Temos $\text{grau}(\rho) = \text{grau}(\rho_\Delta)$: de fato, ρ e ρ_Δ são homotópicas por

$$H(p, t) = (\gamma(p + (1 - t)\delta + t\Delta(p)) - \gamma(p))/|\gamma(p + (1 - t)\delta + t\Delta(p)) - \gamma(p)|.$$

Vemos assim em particular, que a giração independe da escolha de δ (ou de Δ).

Para a segunda afirmação, vamos comparar ρ nos dois casos. Seja $\delta > 0$ suficientemente pequeno para podermos usá-lo no cálculo da giração de γ e seja $\rho_{\delta, \gamma}$ definida como acima. Seja $\Delta(p) = \eta_{-1}(\eta(p) + \delta) - p$. Temos então que $\rho_{\delta, \gamma} \circ \eta = \rho_{\Delta, \gamma \circ \eta}$, o que demonstra a igualdade das girações. ■

Proposição 2.9: *A giração satisfaz as propriedades de aditividade abaixo.*

- (a) *Sejam $\gamma : [a, b] \rightarrow \mathbb{R}^2$ uma função satisfazendo $\gamma(a) = \gamma(c) = \gamma(b)$, para $c \in (a, b)$. Suponha que γ é mansa quando interpretada como função de \mathbb{S}^1 em \mathbb{R}^2 e suave com derivada não nula em $a \cong b$ e c . Suponha ainda que $\gamma'(a)$ e $\gamma'(c)$ são linearmente independentes, isto é, que a auto-interseção em questão é transversal. Temos*

$$\tau(\gamma) = \tau(\gamma|_{[a, c]}) + \tau(\gamma|_{[c, b]}).$$

- (b) *Seja A um disco com k buracos, com fronteira exterior Γ_0 e fronteiras interiores Γ_i , todas orientadas no sentido anti-horário. Sejam γ_i uma parametrização de Γ_i para cada i e $\phi : A \rightarrow \mathbb{R}^2$ imersão (suave). Então*

$$\tau(\phi \circ \gamma_0) = \sum_{1 \leq i \leq k} \tau(\phi \circ \Gamma_i) - (k - 1) \text{sgn}(\det(D\phi)).$$

Demonstração: Seja $\gamma(t) = (u(t), v(t))$. Suponha sem perda de generalidade, que $\gamma(a) = \gamma(c) = \gamma(b) = (0, 0)$ e, dada a transversalidade da auto-interseção, que $u'(a) = u'(b)$ e $u'(c)$ são positivas, como na Figura 2.5. Pelo Teorema da função implícita, existe uma vizinhança em torno da origem na qual podemos definir duas funções suaves Y_0 e Y_1 tais que $Y_0(u(t)) = v(t)$ se t está próximo de a ou b , e $Y_1(u(t)) = v(t)$ se t está próximo de c . Sejam A, ℓ, ϵ e δ_0 reais positivos tais que $2\arctan(A) + \epsilon < \pi$, para $|t - a| < \delta_0$, $|t - c| < \delta_0$ ou $|t - b| < \delta_0$, vale $|u(t)| < \ell$ e para $|x_0|, |x_1| < \ell$, valem $|Y_0'(x_0)| < A$, $|Y_1'(x_0)| < A$, $|Y_0'(x_0) - Y_0'(x_1)| < \epsilon$ e $|Y_1'(x_0) - Y_1'(x_1)| < \epsilon$.

Figura 2.5 (a curva perto da auto-interseção)

Seja $\delta_1 > 0$ pequeno o suficiente para podermos usá-lo no cálculo da giração de γ e $\delta = \min\{\delta_0, \delta_1\}$. Usando este δ , definimos $\rho_{[a,b]}$, $\rho_{[a,c]}$ e $\rho_{[c,b]}$ como acima para γ , $\gamma|_{[a,c]}$ e $\gamma|_{[c,b]}$. Vamos considerar seus levantamentos $\hat{\rho}_{[a,b]}$, $\hat{\rho}_{[a,c]}$ e $\hat{\rho}_{[c,b]}$. A função $\hat{\rho}_{[a,b]}$ tem como domínio o intervalo $[a, b]$ e $\tau(\gamma) = \hat{\rho}_{[a,b]}(b) - \hat{\rho}_{[a,b]}(a)$. Fórmulas análogas valem para os outros dois casos. Note que, como na Figura 2.6, $\rho_{[a,b]}$ coincide, a menos de uma constante aditiva inteira, com $\rho_{[a,c]}$ no intervalo $[a, c - \delta]$. Analogamente, $\rho_{[a,b]}$ coincide, novamente a menos de constante aditiva inteira, com $\rho_{[c,b]}$ em $[c, b - \delta]$. Falta assim provar que

$$\hat{\rho}_{[a,b]}(b) - \hat{\rho}_{[a,b]}(b - \delta) + \hat{\rho}_{[a,b]}(c) - \hat{\rho}_{[a,b]}(c - \delta) = \hat{\rho}_{[c,b]}(b) - \hat{\rho}_{[c,b]}(b - \delta) + \hat{\rho}_{[a,c]}(c) - \hat{\rho}_{[a,c]}(c - \delta).$$

Provaremos isto subtraindo um termo do outro e provando que a resposta é um inteiro de módulo menor do que 1, logo zero. Esta diferença é um inteiro: ela é obtida somando e subtraindo as três girações envolvidas (cancelando termos).

Figura 2.6 (levantamentos $\hat{\rho}$ fragmentados)

Vamos mostrar agora que $|\hat{\rho}_{[a,b]}(b) - \hat{\rho}_{[a,b]}(b - \delta)| < \frac{\epsilon}{2\pi}$. Seja $b - \delta \leq t \leq b$. Temos $\tan(2\pi\rho_{[a,b]}(t)) = (v(t + \delta - b + a) - v(t)) / (u(t + \delta - b + a) - u(t)) = Y_0'(x)$, para algum x , $-\ell < x < \ell$, pelo teorema do valor médio. Assim, para valores de t neste intervalo, $\tan(2\pi\rho_{[a,b]}(t))$ assume valores apenas em um intervalo de tamanho menor que ϵ . Como a função tangente tem derivada sempre maior do que 1, $\rho_{[a,b]}(t)$ novamente assume valores apenas em um intervalo de tamanho menor do que $\frac{\epsilon}{2\pi}$, demonstrando a afirmação. Analogamente, $|\hat{\rho}_{[a,b]}(c) - \hat{\rho}_{[a,b]}(c - \delta)| < \frac{\epsilon}{2\pi}$.

Mostremos agora que $|\hat{\rho}_{[c,b]}(b) - \hat{\rho}_{[c,b]}(b - \delta)| < \frac{\arctan(A)}{\pi}$. De fato, para $b - \delta < t < b$, temos $\tan(2\pi\rho_{[c,b]}(t)) = (v(t + \delta - b + c) - v(t)) / (u(t + \delta - b + c) - u(t))$, donde $|\tan(2\pi\rho_{[c,b]}(t))| < A$, novamente pelo teorema do valor médio, implicando trivialmente a desigualdade acima. Finalmente, de forma análoga, $|\hat{\rho}_{[a,c]}(c) - \hat{\rho}_{[a,c]}(c - \delta)| < \frac{\arctan(A)}{\pi}$. O item (a) está demonstrado, fazendo uso da relação entre A e ϵ .

Seja A o anel limitado por C_1 e C_2 , círculos centrados na origem, com raios $r_1 < r_2$, orientados no mesmo sentido. Seja $H : A \rightarrow \mathbb{R}^2$ uma função suave sem pontos críticos. Então as girações de $H(C_1)$ e $H(C_2)$ são iguais. Para cada círculo centrado na origem de raio r , considere a função $\rho_r : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ levando cada ponto ao vetor tangente (normalizado) à imagem do círculo. A invariância do grau por homotopia nos dá o resultado desejado.

Seja D um disco. Suponha $F : D \rightarrow \mathbb{R}^2$ suave, sem ponto críticos em D . Então $\tau(F(\mathbb{S}^1)) = \text{sgn det}(DF)$, onde $F(\mathbb{S}^1)$ é uma curva com orientação determinada pela orientação anti-horária de \mathbb{S}^1 por F e o sinal de $\text{det}(DF)$ é constante em D . Pelo Teorema 2.5, podemos supor sem perda que estamos tratando do disco unitário. Use então o teorema da função inversa para concluir a existência de um raio r_1 suficientemente pequeno para o qual $\tau(F(C_1)) = \text{sgn det}(DF)$, para C_1 centrado na origem, de raio r_1 e orientação anti-horária. Pelo parágrafo anterior, $\tau(F(\mathbb{S}^1)) = \tau(F(C_1))$.

Para o item (b), suponha sem perda que nosso disco com k buracos é como mostrado na Figura 2.3. Como na proposição anterior, existe uma homotopia regular levando a curva Γ_0 para a curva auxiliar Γ indicada na Figura 2.3. Use agora o item (a) para decompor a giração de $\phi(\Gamma)$ na soma das girações de $2k - 1$ curvas. Pelo parágrafo anterior, a giração das curvas adicionais é igual a $-\text{sgn det}(D\phi)$. ■

A giração de uma curva normal fechada simples é ± 1 , dependendo de sua orientação — isto é exatamente o item (b) para $k = 0$.

Exercício: Seja A um disco com k buracos, com bordo externo Γ_0 orientado positivamente, e bordos internos Γ_i orientados *negativamente*. Sejam γ_i parametrizações de Γ_i , respeitando as orientações.

(a) Seja $\phi : A \rightarrow \mathbb{R}^2 - \{p\}$ contínua. Mostre que

$$\sum_{i=0,1,\dots,k} w(\phi \circ \gamma_i, p) = 0.$$

(b) Seja $\phi : A \rightarrow \mathbb{R}^2$ uma imersão suave. Mostre que

$$\sum_{i=0,1,\dots,k} \tau(\phi \circ \gamma_i) = \pm(1 - k),$$

e determine o sinal correto.

Vamos apresentar ainda outra maneira de calcular a giração: a *decomposição de Seifert*. Dada uma família normal de curvas fechadas orientadas, obtemos uma família de curvas simples fechadas pelo seguinte processo: cada arco entre duas interseções continua a existir com a mesma orientação mas sempre que chegarmos a uma interseção dobramos à direita ou esquerda, de forma a respeitar as orientações. As curvas simples assim obtidas só se interceptam em pontos isolados (as antigas interseções) mas aí nunca se cruzam.

Proposição 2.10: *Suponha que a decomposição de Seifert de uma família normal de curvas fechadas e orientadas dê origem a m curvas simples orientadas no sentido anti-horário e n outras curvas simples orientadas no sentido horário. Então a soma das girações das curvas da família é $m - n$.*

Demonstração: Lembre que a giração de uma curva simples é ± 1 , dependendo de sua orientação. Devemos assim mostrar que a decomposição de Seifert não altera a giração total da família. Isto é inteiramente análogo ao item (a) da Proposição 2.9: devemos fazer as mesmas estimativas de como os \hat{p} se modificam próximo a cada auto-interseção. ■

Por exemplo, a curva na Figura 2.7(a) se decompõe em curvas como em 2.7(b), e tem giração igual a 3.

Figura 2.7(a,b) (uma decomposição de Seifert de curva com giração igual a 3)

A seguir, algumas propriedades simples da decomposição de Seifert.

Proposição 2.9: *Seja Γ uma curva normal fechada orientada.*

- (a) *Se Γ não é simples, então tem mais de um ciclo de Seifert.*
- (b) *Se Γ só tem ciclos de Seifert orientados positivamente, então sua giração é estritamente positiva.*
- (c) *Se Γ só tem ciclos de Seifert orientados positivamente e sua giração é igual a 1, então Γ é simples.*

Demonstração: Se existirem ciclos com as duas orientações, (a) é imediato. Se todos os ciclos têm a mesma orientação, a fronteira da componente ilimitada do complemento de Γ é um ciclo de Seifert, que não pode ser Γ toda, a menos que a curva seja simples. Os itens (b) e (c) agora são imediatos. ■

Seção 2.3: Grau topológico de funções próprias no plano

Uma função $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ contínua é *própria* se a imagem inversa $F^{-1}(K)$ de todo conjunto compacto K do plano também é compacta. É fácil ver (exercício!) que F é própria se e somente se $\lim_{|p| \rightarrow \infty} |F(p)| = \infty$.

Seja F uma função contínua própria do plano no plano e p um ponto do plano. Note que, como F é própria, as pré-imagens de p formam um conjunto compacto, logo limitado, do plano.

Nosso objetivo agora é definir o grau topológico de uma função própria no plano. Para funções de \mathbb{S}^1 em \mathbb{S}^1 , vimos na seção 1 mais de uma caracterização do grau: por meio de suspensões (que chamaremos de uma caracterização *homotópica*) e por meio da contagem (com sinais) de pré-imagens (que chamaremos de uma caracterização *diferencial*).

O resultado a seguir relaciona um ingrediente homotópico com um diferencial no caso de funções do plano no plano.

Teorema 2.10: *Sejam $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma função suave própria, $q \in \mathbb{R}^2$ e γ uma parametrização de uma curva fechada normal simples Γ tal que todas as pré-imagens de q estão do lado de dentro de Γ .*

(a) *O inteiro $w(F \circ \gamma, q)$ independe de γ e q .*

(b) *Se q é um valor regular de F então*

$$w(F \circ \gamma, q) = \sum_{F(p)=q} \text{sgn det}(DF(p)).$$

Demonstração: Para o item (a), sejam q um ponto qualquer do plano e Γ_1 e Γ_2 curvas fechadas normais e simples envolvendo todas as pré-imagens de q . Seja Γ_3 um grande círculo envolvendo Γ_1 e Γ_2 . Sejam γ_i parametrizações de Γ_i . Pelo item (b) da Proposição 2.6, (para $k = 1$) $w(F \circ \gamma_1, q) = w(F \circ \gamma_3, q)$ e, analogamente, $w(F \circ \gamma_2, q) = w(F \circ \gamma_3, q)$, e concluímos assim a independência na escolha de γ .

Seja então $gr(q) = w(F \circ \gamma, q)$, onde γ é uma parametrização de uma curva fechada normal simples envolvendo todas as pré-imagens de q . Vamos provar que gr é localmente constante no plano, logo constante. Seja q um ponto qualquer do plano e seja V uma bolinha fechada centrada em q . A pré-imagem de V por F é compacta, já que F é própria. Escolhemos uma γ envolvendo todo o conjunto $F^{-1}(V)$. Assim, $F \circ \gamma$ não passa por V , logo $w(F \circ \gamma, q) = w(F \circ \gamma, q')$, onde q' é qualquer ponto de V , pelo item (c) da Proposição 2.6.

Para o item (b), seja D o disco limitado por Γ . Pelo teorema da função inversa, podemos tomar pequenas bolas B_p ao redor de cada pré-imagem p de q tais que, restrita a cada uma delas, a função F é um difeomorfismo. Sejam β_p parametrizações de B_p , orientadas positivamente. O conjunto $A = D - \bigcup_{F(p)=q} B_p$ é um disco com buracos e a ele aplicaremos o item (b) da Proposição 2.6 — note que $w(F \circ \beta_p, q) = \text{sgn det}(DF(p))$. Assim, provamos a igualdade do enunciado do teorema. ■

Definiremos o *grau* de F , $\text{grau}(F)$, como sendo o inteiro descrito de duas formas no teorema. Existem valores regulares q de F pelo teorema de Sard e a caracterização homotópica do grau (e os argumentos de independência acima) não fazem uso da suavidade de F . Note que o número de pré-imagens de um valor regular q é pelo menos igual a $|\text{grau}(F)|$ e tem a mesma paridade.

Capítulo 3

Neste capítulo finalmente definimos as funções que nos interessam, começamos o estudo de seu comportamento global e vemos alguns exemplos.

Seção 3.1: Funções cordatas

Neste texto, consideraremos apenas funções $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ suaves satisfazendo as propriedades abaixo:

- (1) F é própria,
- (2) o conjunto crítico C de F é limitado,
- (3) todo ponto crítico é uma dobra ou uma cúspide,
- (4) $F(C)$ é uma família normal de curvas.

Tais funções serão chamadas de *cordatas*. Lembre que (3) implica que zero é valor regular de $\det(DF)$. Funções que satisfazem (3) são *funções de Whitney*, que demonstrou em [W] sua abundância no conjunto de todas as funções suaves do plano no plano. Para elas, a existência de valores regulares é trivial: eles são o complemento de $F(C)$, um conjunto compacto.

Como já vimos, C é formado por curvas regulares. Por (2), estas curvas são fechadas e em número finito (note que curvas diferentes não podem acumular em um ponto do plano, por (3)). Por (2) e (3) também, o número de cúspides é finito. O número de pré-imagens de qualquer ponto também é finito: a imagem inversa de um ponto tem que ser compacta por (1), e só contém pontos isolados, pelo comportamento de F perto de pontos regulares, dobras e cúspides.

Como veremos, o comportamento de uma função cordata F é bastante estruturado. Para descrevê-lo, consideraremos o conjunto $\mathcal{F} = F^{-1}(F(C))$, a *flor* de F . Por (1) e (2), $F(C)$ e \mathcal{F} são limitados. Se S é um subconjunto de \mathbb{R}^2 , as componentes conexas de $\mathbb{R}^2 - S$ serão chamadas de *placas de S* . Assim falaremos de placas de C , de $F(C)$ ou de \mathcal{F} . Nos três casos, existe uma única placa ilimitada. A imagem de uma placa de \mathcal{F} está sempre contida em uma única placa de $F(C)$.

Teorema 3.1: *Seja F cordata, E uma placa de \mathcal{F} e X a placa de $F(C)$ que contém $F(E)$. Então E é um recobrimento de ordem finita de X por F .*

Demonstração: Vamos começar vendo que $F(E) = X$. Como E se compõe de pontos regulares, $F(E)$ é um aberto, pelo teorema da função inversa. Agora, basta provar que $F(E)$ é fechado em X , e a sobrejetividade de F segue por conexidade. Considere uma seqüência de pontos $x_n = F(e_n)$ convergindo a $x \in X$. Como F é própria, existe uma subsequência dos e_n convergindo para um ponto e no fecho de E , tal que $F(e) = x$. Pela definição de placas, e tem que estar em E , logo $x \in F(E)$.

Agora, dado um ponto $x \in X$, devemos encontrar um aberto V contendo x tal que F restrito a cada componente conexa de $F^{-1}(V)$ é um homeomorfismo. O número de pré-imagens de x em E é finito, logo, pelo teorema da função inversa (E só contém pontos regulares), é fácil obter V conexa contendo x tal que a afirmação acima vale para cada componente conexa de $F^{-1}(V)$ contendo alguma pré-imagem de x . Falta apenas provar que não existem componentes conexas de $F^{-1}(V)$ que não contenham uma pré-imagem de x . Seja W uma componente conexa de $F^{-1}(V)$ — provaremos que $F(W) = V$, mostrando que $F(W)$ é aberto e fechado em V . Pelo teorema da função inversa, novamente, $F(W)$ é aberto. Considere uma seqüência de pontos $x_n = F(e_n)$, onde $e_n \in W$, e x_n converge a um ponto x_∞ em V . Tomando uma subsequência dos e_n convergente a e_∞ , temos que necessariamente $F(e_\infty) = x_\infty$, donde $e_\infty \in F^{-1}(V)$, portanto $e_\infty \in W$, mostrando que $F(W)$ é fechado em V . ■

Assim, se X é uma placa de $F(C)$, $F^{-1}(X)$ é uma união disjunta de placas de \mathcal{F} : chamaremos cada uma dessas placas de \mathcal{F} de uma *pré-imagem* de X . Mais geralmente, para qualquer $S \subseteq \mathbb{R}^2$, componentes conexas de $F^{-1}(S)$ são *pré-imagens* de S . Não provamos ainda que $F^{-1}(X)$ é não vazio — isto será demonstrado no

Corolário 3.4. Note que o número de pré-imagens de uma placa X não é necessariamente igual ao número de pré-imagens de um ponto em X .

Corolário 3.2: *Seja F cordata.*

- (a) *Se uma placa de $F(C)$ é um disco, suas pré-imagens também são discos.*
- (b) *A única pré-imagem da placa ilimitada de $F(C)$ é a placa ilimitada de \mathcal{F} .*
- (c) *O grau de F não é nulo e F é sobrejetora.*

Demonstração: O item (a) segue do teorema acima e do Teorema 2.2, já que um disco é simplesmente conexo. Para o item (b), como já vimos, as pré-imagens da placa ilimitada de $F(C)$ são placas de \mathcal{F} : não podemos ter uma pré-imagem limitada por continuidade de F , e a placa ilimitada de \mathcal{F} deve ir na placa ilimitada de $F(C)$ porque F é própria. Para o item (c), seja q um ponto qualquer da placa ilimitada de $F(C)$. Por (b), q tem pelo menos uma pré-imagem e todas estão na placa ilimitada de \mathcal{F} . Usando a caracterização diferencial do grau, como $\det(DF)$ tem o mesmo sinal em todas as pré-imagens de q , $\text{grau}(F) \neq 0$. Quanto à sobrejetividade, obviamente todos os valores críticos (isto é, pontos de $F(C)$) estão na imagem de F . Como já vimos no final do capítulo anterior, cada valor regular tem pelo menos $|\text{grau}(F)|$ pré-imagens, implicando a sobrejetividade. ■

Corolário 3.3: *Seja F cordata, p um ponto regular no domínio, e α uma função contínua de $[0, 1]$ a uma placa de $F(C)$, com $F(p) = \alpha(0)$. Então existe uma única função contínua $\beta : [0, 1] \rightarrow \mathbb{R}^2$ com $\beta(0) = p$ e $F \circ \beta = \alpha$.*

Demonstração: O resultado segue diretamente combinando o Teorema 3.1 com o Teorema da existência do levantamento da Seção 2.1. ■

Teorema 3.3 (forma normal em infinito): *Seja $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ uma função cordata e $n = |\text{grau}(F)|$. Então existem abertos U e V no domínio e na imagem respectivamente, e existe um difeomorfismo $\psi : U \rightarrow U'$ com as seguintes propriedades:*

- $F = F_0 \circ \psi$, onde $F_0(z) = z^n$ (em notação complexa),
- V é o complemento de um disco fechado centrado na origem, U é o complemento de um disco fechado (de fronteira não necessariamente circular) e U' é a imagem inversa de V por F_0 ,
- ψ preserva orientação se e somente se $\text{grau}(F) > 0$.

Demonstração: Seja D um disco fechado na imagem centrado na origem cujo interior contém toda a imagem do conjunto crítico de F , e defina V como sendo o complemento de D . Pela proposição anterior, o número de pré-imagens de pontos de V é constante. Seja U a imagem inversa de V por F : afirmamos que U é conexa. O complemento de U é compacto, logo temos apenas uma componente ilimitada. Por outro lado, não podem existir componentes limitadas: um ponto w em uma componente limitada é o ponto de partida para uma pré-imagem de um raio ligando $F(w)$ a qualquer ponto em V , fazendo uso do Corolário 3.3, o que é um absurdo pois esta pré-imagem deve ser limitada.

Seja p_0 um ponto qualquer de U : tracemos por $F(p_0)$ um círculo C_0 centrado na origem. Como C_0 só contém valores regulares, a imagem inversa de C_0 em U , B_0 , será uma união de curvas regulares, das quais uma passa por p_0 . Afirmamos que B_0 é conexa, isto é, que é formada por uma única curva regular.

De fato, suponha por absurdo que $B_0 = B_1 \cup B_2$ com B_1 e B_2 fechados disjuntos em B_0 . Vamos definir abertos disjuntos U_1 e U_2 com $U = U_1 \cup U_2$. Dado um ponto p em U , trace uma semi-reta a partir da origem passando por $F(p)$ e considere o segmento contido nesta semi-reta que liga $F(p)$ a algum ponto de C_0 . Invertendo este segmento, novamente pelo Corolário 2.9, temos, entre outras pré-imagens que não nos interessam, um caminho em U ligando p a B_0 : dizemos que p está em U_1 ou U_2 conforme tiver sido ligado a um ponto de B_1 ou B_2 . Os conjuntos U_1 e U_2 são obviamente abertos, contradizendo a conexidade de U .

Sabemos que B_0 é uma curva regular fechada sobre a qual $\det(DF)$ tem sinal constante e contém todas as n pré-imagens de $F(p_0)$, pelo Corolário 3.3, já que $F(p_0)$ está na placa ilimitada de $F(C)$. Pela construção acima, B_0 é um recobrimento de ordem n de C_0 . Note, aliás, que o sinal de $\text{grau}(F)$ é o de $\det(DF)$ em U .

Definiremos $F_0(z) = z^n$ e $\psi(p_0)$ como sendo qualquer uma das n soluções de $z^n = F(p_0)$. Devemos mostrar que podemos construir um homeomorfismo ψ entre B_0 e o círculo B'_0 centrado na origem passando por $\psi(p_0)$, respeitando a lei de composição $F = F_0 \circ \psi$. Sejam p_0, \dots, p_{n-1} as pré-imagens de $F(p_0)$ seqüenciadas sobre B_0 no sentido anti-horário. A função F leva cada arco $p_i p_{i+1}$ de B_0 em C_0 todo: compondo com o ramo apropriado da n -ésima raiz, este arco é levado a um n -ésimo de B'_0 e por justaposição contruímos nosso homeomorfismo. O processo de compor com a n -ésima raiz na verdade corresponde a tomar um levantamento pelo recobrimento $z \mapsto z^n$, em notação complexa, sendo portanto justificado pelo Teorema da existência de levantamento.

Vamos agora estender ψ a U . Seja p um ponto qualquer de U : liguemos $F(p)$ a C_0 por um segmento radial e invertamos este segmento. Temos assim, além de outras pré-imagens que não nos interessam, um arco ligando p a um ponto \tilde{p} de B_0 , para o qual ψ já foi definida. Extraia a raiz n -ésima do segmento radial ligando $F(\tilde{p})$ a $F(p)$ para obter, além de outros segmentos que não nos interessam, um segmento radial ligando $\psi(\tilde{p})$ ao que definiremos como $\psi(p)$. A verificação de que ψ de fato satisfaz todas as propriedades do enunciado fica como um exercício. ■

O próximo corolário nos dá uma terceira caracterização do grau de uma função cordata.

Corolário 3.5: *Seja F cordata, com conjunto crítico C . Seja Γ uma curva normal fechada simples, orientada no sentido anti-horário, envolvendo C , com parametrização γ . Então $\tau(F \circ \gamma) = \text{grau}(F)$.*

Demonstração: Pela Proposição 2.5, item (b) com $k = 1$, podemos supor $\Gamma \subseteq U$, onde U é como no enunciado do Teorema 3.3.

Seja $F_0(z) = z^{|\text{grau}F|}$ e α_0 a parametrização habitual do círculo unitário: é fácil ver que $\tau(F_0 \circ \alpha_0) = |\text{grau}F|$. Novamente pela Proposição 2.5 (de novo, item (b) com $k = 1$), se α é uma parametrização de uma curva normal simples fechada contornando a origem, $\tau(F_0 \circ \alpha) = \pm |\text{grau}F|$, conforme α induza a orientação/positiva ou negativa em sua imagem.

Voltando à situação do corolário, $\tau(F \circ \gamma) = \tau(F_0 \circ (\psi \circ \gamma))$, e podemos aplicar nossas conclusões a $\alpha = \psi \circ \gamma$: o resultado decorre do fato que a orientação induzida por α é positiva se $\det D\psi > 0$ e negativa caso contrário. ■

Para o próximo teorema, precisamos de uma definição: um *arco maximal* é uma componente conexa do conjunto das imagens de dobras menos os pontos de interseção de imagens de curvas críticas.

Teorema 3.6: *Seja F cordata. O número de pré-imagens é constante dentro de uma placa de $F(C)$ ou de um arco maximal. Numa vizinhança de uma imagem q de uma dobra, na imagem q' de uma cúspides ou de um ponto de interseção q'' com k pré-imagens, o número de pré-imagens muda de acordo com os diagramas na Figura 3.1, onde as setas indicam o sentido de dobra.*

Figura 3.1 (como variam as pré-imagens)

Demonstração: Pelo Teorema da ordem de um recobrimento, o número de pré-imagens é constante dentro de uma placa de $F(C)$. Vamos demonstrar a afirmação para imagens q de dobras. Seja p_0 a dobra tal que $F(p_0) = q$, e sejam p_1, \dots, p_{k-1} as demais pré-imagens regulares de q . Aplique a forma normal em torno da dobra p_0 e o teorema da função inversa em torno das pré-imagens regulares para construir uma vizinhança V de q com as seguintes propriedades:

- V e $V \cap F(C)$ são conexos,
- V não contém imagens de interseções,
- os pontos p_i estão em componentes conexas distintas de $F^{-1}(V)$,
- a componente conexa U_0 de $F^{-1}(V)$ contendo p_0 é um domínio para a forma normal da dobra p_0 ,
- F é um difeomorfismo restrito à componente conexa U_i de $F^{-1}(V)$ contendo p_i , para cada $i \neq 0$,

Assim, já sabemos que pontos na imagem têm pelo menos o número de pré-imagens indicado no diagrama. Falta provar que não existem outras componentes conexas de $F^{-1}(V)$ além dos U_i . Seja W uma componente conexa de $F^{-1}(V)$, diferente de U_0 — provaremos que $F(W) = V$, mostrando que $F(W)$

é aberto e fechado em V , imitando a demonstração do Teorema 3.1. Note que W é formado por pontos regulares, pelas duas primeiras propriedades da construção de V e o fato que F é cordata. Pelo teorema da função inversa, novamente, $F(W)$ é aberto. A demonstração de que $F(W)$ é fechado pode ser copiada textualmente da demonstração do Teorema 3.1. As demonstrações associadas aos diagramas (b) e (c) ficam como um exercício. ■

Seção 3.2: Exemplos

Os teoremas 3.1, 3.3 e 3.6 fornecem uma descrição topológica completa da ação de uma função cordata. Vamos descrever a geometria global de algumas funções cordatas simples.

Inicialmente, considere $F(x, y) = (x^2 - y^2 + x, 2xy - y)$, ou, em notação complexa, $F(z) = z^2 + \bar{z}$. Alguns fatos sobre esta função são evidentes. Para pontos longe da origem, a função se comporta essencialmente como $z \mapsto z^2$. Além disso, $F(0) = 0$ e perto da origem a função se comporta como $z \mapsto \bar{z}$. Dito de uma forma um pouco diferente, pequenos círculos orientados positivamente em torno da origem são levados bijetivamente a curvas quase redondas orientadas negativamente também próximas da origem. Círculos enormes centrados em torno da origem orientados positivamente são levados a curvas que dão duas voltas em torno da origem no sentido anti-horário: assim, $\text{grau}F = 2$. Não é evidente o que acontece a círculos centrados na origem de tamanho intermediário — em particular, não é óbvio como F passa de forma contínua de um comportamento para outro. Para entender a geometria global de F , a primeira coisa que faremos é calcular seu conjunto crítico C , o que já foi feito na seção 1.2: C é o círculo de raio $\frac{1}{2}$, onde foram encontrados três cúspides, os pontos $(\frac{1}{2}, 0), (-\frac{1}{4}, \pm\frac{\sqrt{3}}{4})$. O passo seguinte é ter uma idéia de $F(C)$. Por meio de contas explícitas, vamos ver que $F(C)$ é uma curva fechada que não tem auto-interseções. Considere dois pontos (x, y) e (a, b) em C com a mesma imagem. Algebricamente, isto corresponde ao sistema

$$\begin{aligned}x^2 + y^2 &= \frac{1}{4}, \\a^2 + b^2 &= \frac{1}{4}, \\x^2 - y^2 + x &= a^2 - b^2 + a, \\2xy - y &= 2ab - b.\end{aligned}\tag{1}$$

Tire y^2 e b^2 das duas primeiras equações e eleve ao quadrado (2) para obter

$$2x^2 + x - 2a^2 - a = 0,\tag{3}$$

$$(2x - 1)^2\left(\frac{1}{4} - x^2\right) = (2a - 1)^2\left(\frac{1}{4} - a^2\right).\tag{4}$$

Resolvendo (3) para x e usando (2), obtemos $x = a$ e $y = b$ ou $x = -(\frac{1+2a}{2})$ e $y^2 = -a(1+a)$. No segundo caso, (3) se reduz a

$$16a^3 + 12a^2 + 3a + \frac{1}{4} = 0,$$

cujas únicas soluções reais são $a = -\frac{1}{4}$, obrigando $x = a$ e $y = b$. As curvas C e $F(C)$ estão desenhadas na Figura 1.4. Como F é própria — é só ver que $\lim_{|p| \rightarrow \infty} |F(p)| = \infty$ — temos que F é cordata. Como $\text{grau}F = 2$, todos os pontos na placa ilimitada de $F(C)$ são alcançados duas vezes, e pontos na outra placa quatro vezes, pelo Teorema 3.6. Dito de maneira um pouco diferente, se quiséssemos resolver o sistema (duas equações, duas incógnitas reais) $F(x, y) = q$, para algum ponto dentro da imagem de $F(C)$, esse sistema teria quatro soluções. A flor \mathcal{F} também está na Figura 1.4, e deixa muito claro vários fatos interessantes a respeito de F :

- não só os pontos na placa Q envolvida por $F(C)$ têm quatro pré-imagens, como cada uma delas deve estar dentro de uma placa limitada de \mathcal{F} : F restrita a cada uma destas placas é um recobrimento de Q , pelo Teorema 3.1 — mais do que isso: como Q é simplesmente conexo, F restrita a cada placa é um difeomorfismo sobre Q , pelos Teoremas 2.2 e 3.1,

- a placa ilimitada de \mathcal{F} é um recobrimento de ordem dois da placa ilimitada de $F(C)$ — note que anéis não são simplesmente conexos,
- imagens de dobras são alcançadas três vezes, e imagens de cúspides, duas — isto mais uma vez segue do Teorema 3.6, mas está evidente na figura, onde um ponto e suas pré-imagens são denotados pela mesma letra.

Exercício: A Figura 1.4 já fornece uma idéia muito boa de F . Tente entender como variam as imagens de círculos cada vez maiores centrados na origem — é bom lembrar o comportamento local de F perto de pontos críticos.

No próximo Capítulo, vamos ver uma maneira ainda mais direta de calcular o grau topológico de uma função cordata a partir de informação sobre seu conjunto crítico e a imagem dele (Corolário 4.2). Assim, o cálculo do número de pré-imagens para qualquer ponto da imagem se tornará completamente elementar uma vez conhecidos estes conjuntos.

O exemplo seguinte é a função F cuja flor é dada na Figura 1.5. Mais uma vez as placas simplesmente conexas de $F(C)$ estão sendo recobertas (difeomorficamente) por placas de \mathcal{F} : a placa T_0 tem quatro pré-imagens conexas, e T_1 , seis. O grau desta função é -2 , como se tornará claro no próximo capítulo, mas isso se deduz da figura: uma curva um pouco para fora da fronteira externa de \mathcal{F} , é mandada com a orientação trocada para uma curva que dá duas voltas próxima à fronteira externa de $F(C)$ — é só seguir as letras para convencer-se disso. Mais uma vez, o anel correspondente à placa ilimitada de $F(C)$ está sendo recoberto duas vezes pela placa ilimitada de \mathcal{F} . Sistemas da forma $F(x, y) = q$, para esta função, podem ter de duas a seis soluções, dependendo da posição de q na imagem. Informações sobre C e $F(C)$ vão ser empregadas para resolver numericamente sistemas deste tipo no Capítulo 5. Aliás, note que a imagem da curva crítica tem uma auto-interseção e giração igual a dois.

A figura do exemplo anterior é esquemática, por razões de apresentação. As duas flores na Figura 5.12 foram obtidas em máquina, e mostram um outro aspecto métrico freqüente em aplicações: a diversidade de escalas relevantes na mesma função. Os desenhos são flores de duas funções cordatas F e G cujas expressões são dadas no Capítulo 5. As duas funções têm duas curvas críticas, uma com três, a outra com quatro cúspides: seus conjuntos críticos são topologicamente indistinguíveis. As diferenças surgem nas imagens dos conjuntos críticos: para as duas funções, as imagens de suas curvas críticas são curvas simples, mas em um caso (para F) uma está por dentro da outra. Mais uma vez, fazendo uso do Corolário 4.2, calculamos $\text{grau}F = \text{grau}G = 4$, e passamos a saber, por exemplo, que o pequeno triângulo na imagem de F tem oito pré-imagens. Quatro delas estão em volta da própria curva crítica que gerou o triângulo, uma situação topologicamente idêntica à que gerou a flor de $z \mapsto z^2 + \bar{z}$, nosso primeiro exemplo. Duas das outras quatro são fáceis de ver na escala da figura, mas as últimas duas só são visíveis com uma ampliação. Há outro fato interessante nesta flor: a placa limitada entre as imagens das duas curvas críticas é um anel, e está sendo recoberta duas vezes por uma de suas pré-imagens (aquela que envolve a curva crítica com três cúspides). Assim, este anel tem cinco pré-imagens, uma das quais o recobre duplamente. A flor de G tem menos surpresas: as placas limitadas na imagem têm seis pré-imagens cada uma, e elas são facilmente identificáveis no desenho. Note que a placa ilimitada de \mathcal{G} é difeomorfa a um disco com cinco buracos, que recobre quatro vezes a placa ilimitada de $F(C)$, por sua vez, um disco com dois buracos.

Seção 3.3: Invertendo funções cordatas

Vamos mostrar como resolver sistemas $F(x, y) = q$ para funções cordatas. Nesta seção, faremos considerações teóricas diretamente ligadas à possibilidade de implementar um algoritmo de resolução que será discutido no Capítulo 5.

Vamos supor aqui que *continuações* são efetuáveis: esta suposição vai ser justificada no Capítulo 5. Mais precisamente, pelo Corolário 3.2, para um caminho β ligando q_0 a q_1 sem passar por $F(C)$, e para p_0 uma pré-imagem de q_0 , existe um único caminho γ começando em p_0 (e terminando em p_1 , uma pré-imagem de q_1) para o qual $\beta = F \circ \gamma$: vamos supor que γ (e em particular p_1) pode ser efetivamente calculado a partir

de β e p_0 . De maneira mais geral, suponha que o caminho β tem apenas um extremo em $F(C)$ e que este extremo é a imagem de um ponto de dobra. Se este extremo for q_1 , a existência e unicidade de γ continua assegurada: vamos supor que também neste caso γ pode ser efetivamente calculado. Se este extremo for q_0 , a imagem da dobra p_0 , então existem zero ou duas γ satisfazendo $\beta = F \circ \gamma$, dependendo do sentido de dobra (isto segue da forma normal da dobra mais considerações análogas às dos casos anteriores): no caso em que existem duas γ , vamos supor que elas podem ser efetivamente calculadas.

De novo, vamos considerar o exemplo simples dado por $F(z) = z^2 + \bar{z}$ em notação complexa, cujo comportamento global descrevemos em detalhe na seção anterior. O primeiro passo na solução do sistema é resolvê-lo para algum valor remoto q_0 , o que pode ser feito como passamos a descrever. Escolha um ponto p_0 suficientemente longe da origem (isto é, com $F(p_0)$ na placa ilimitada de $F(C)$), e faça $q_0 = F(p_0)$: uma das pré-imagens de q_0 é obviamente p_0 . Para obter a outra, seja β um círculo centrado na origem indo de q_0 de volta a q_0 , e por continuação, obtemos a imagem inversa de β saindo de p_0 . Como na demonstração da forma normal em infinito, esta imagem inversa nos levará à outra pré-imagem p'_0 de q_0 .

Seja agora q_1 um outro ponto na placa ilimitada de $F(C)$. Para resolver $F(p) = q_1$, basta traçar uma curva β ligando q_0 a q_1 sem passar por $F(C)$ e invertê-la por continuação a partir das duas condições iniciais p_0 e p'_0 .

Seja finalmente q_2 um ponto na placa limitada de $F(C)$, logo com quatro pré-imagens. Para resolver $F(p) = q_2$, ligue q_0 a $q_3 = F(p_3)$ (sem esbarrar em $F(C)$), a imagem de um ponto de dobra. O ponto q_3 tem três pré-imagens: uma delas é o próprio p_3 , portanto já conhecido. As outras duas podem ser obtidas por continuação, como indicado na Figura 3.2. Finalmente, ligue q_3 a q_2 : todas as quatro pré-imagens de q_2 agora podem ser obtidas por continuação. Note que duas delas são obtidas a partir de continuações originárias de um ponto de dobra.

Figura 3.2 (invertendo $z^2 + \bar{z}$)

Na verdade, tudo o que fizemos neste exemplo são construções gerais, adequadas para a inversão de qualquer função cordata. Na Figura 3.2, a flor nos ajuda a compreender o comportamento das pré-imagens dos segmentos envolvidos, mas o processo que descrevemos não requer conhecimento da flor: os objetos essenciais são C e $F(C)$. Suponha, assim, que ao estudar uma função cordata, encontramos algumas curvas críticas. Se estas forem todas as curvas críticas, estamos equipados para inverter F . Mas como saber se não existem outras curvas críticas ainda não encontradas? O próximo capítulo trata deste problema: as ferramentas necessárias são apresentadas nas seções 4.1 a 4.4, e seu modo de emprego, na seção 4.5.

Capítulo 4

O objetivo deste Capítulo é caracterizar os conjuntos C e $F(C)$ para uma função cordata. Isto é, queremos determinar quando dois conjuntos C_\bullet e \tilde{C}_\bullet são, respectivamente, o conjunto crítico de alguma função cordata F e sua imagem $F(C_\bullet)$. Pelos resultados dos capítulos anteriores já conhecemos algumas condições necessárias: C_\bullet deve ser uma união disjunta de curvas regulares simples fechadas e \tilde{C}_\bullet deve ser uma família normal de curvas sem vértices. Mais precisamente, seja f_\bullet definida em um conjunto C_\bullet de curvas regulares fechadas disjuntas no plano, parametrizando uma família normal $f_\bullet(C_\bullet)$ de curvas fechadas (sem vértices). Vamos obter condições necessárias e suficientes para decidir se existe uma F_\bullet cordata estendendo f_\bullet ao plano todo, com conjunto crítico igual a C_\bullet , sendo dados o grau, sentidos de dobra e classificação das cúspides (em internas e externas).

Na verdade, a pergunta que nos interessa para fins computacionais é mais específica. Seja F uma função cordata e $C_\bullet \subseteq C$ um conjunto de curvas críticas: desejamos saber se existe uma função cordata F_\bullet com mesmo grau que F cujo conjunto crítico seja C_\bullet e que coincida com F numa vizinhança de C_\bullet . Como vimos na seção 3.3, para inverter F queremos calcular seu conjunto crítico. Assim, se já tivermos encontrado um conjunto C_\bullet de curvas críticas, e soubermos que não existe uma tal F_\bullet , concluímos que devemos procurar novas curvas críticas.

Seção 4.1: Os teoremas de contagem

Para uma curva crítica Γ de F orientada com o sentido de dobra, definimos $\kappa_{\text{int}}(\Gamma)$ e $\kappa_{\text{ext}}(\Gamma)$ como sendo o número de cúspides internas e externas respectivamente em Γ , e $\mathbf{v}(\Gamma) = \tau(F(\Gamma))$.

Teorema 4.1: *Seja F cordata, com conjunto crítico C . Para uma placa de C , com fronteira exterior Γ_0 (se houver) e fronteiras interiores $\Gamma_1, \dots, \Gamma_k$ (de novo, se houver),*

$$\mathbf{v}(\Gamma_0) - \kappa_{\text{int}}(\Gamma_0) + \sum_{i=1, \dots, k} (\mathbf{v}(\Gamma_i) - \kappa_{\text{ext}}(\Gamma_i)) = 1 - k.$$

Se não existir fronteira exterior, $\mathbf{v}(\Gamma_0)$ deve ser interpretado como $|\text{grau}(F)|$.

Demonstração: Consideremos inicialmente uma placa limitada. Seja A o disco com k buracos indicado na Figura 4.1, cujas curvas na fronteira $\tilde{\Gamma}_i$ estão todas orientadas no sentido anti-horário. De maneira mais precisa, os argumentos a seguir valem para uma escolha de curvas $\tilde{\Gamma}_i$ suficientemente próximas de Γ_i . Seja $s = \text{sgn det}(DF)(p)$, onde p é um ponto qualquer de A .

Figura 4.1

O teorema segue das identidades abaixo:

$$\begin{aligned} \tau(F(\tilde{\Gamma}_0)) &= -s(\kappa_{\text{int}}(\Gamma_0) - \mathbf{v}(\Gamma_0)), \\ \tau(F(\tilde{\Gamma}_i)) &= s(\kappa_{\text{ext}}(\Gamma_i) - \mathbf{v}(\Gamma_i)), \quad i \neq 0. \end{aligned} \tag{*}$$

De fato, usando a Proposição 2.5, item (b), temos

$$\tau(F(\tilde{\Gamma}_0)) - \sum_{i=1, \dots, k} \tau(F(\tilde{\Gamma}_i)) = s(1 - k).$$

Por (*), o Teorema segue. Falta portanto apenas demonstrar (*) — vamos mostrar apenas a primeira identidade: a segunda é análoga. O resultado no caso em que o sentido de dobra em Γ_0 é o anti-horário deve ficar claro considerando a Figura 4.2: o outro sentido de dobra é tratado exatamente da mesma maneira.

Figura 4.2

Lembre-se que, neste caso, pela definição de sentido de dobra, $s = +1$. A forma normal das cúspides explica porque em cada cúspide interna perdemos +1 ao passarmos da giração de $F(\Gamma_0)$ à de $F(\tilde{\Gamma}_0)$.

Falta ainda demonstrar o caso da componente ilimitada, ou seja, a situação em que a placa de C não tem fronteira exterior. A única alteração no argumento acima é que tomamos A como o disco com buracos cuja fronteira exterior é uma curva regular fechada simples $\tilde{\Gamma}_0$ contornando C , que, pelo Corolário 3.5 satisfaz $\tau(F(\tilde{\Gamma}_0)) = \text{grau}(F)$. ■

Corolário 4.2: *Seja F cordata com conjunto crítico C consistindo de curvas $\Gamma_1, \dots, \Gamma_n$, com um total de κ cúspides. Então*

$$|\text{grau}(F)| = 1 + \kappa - 2 \sum_{1 \leq i \leq n} \mathbf{v}(\Gamma_i).$$

Demonstração: Sejam k_i o número de buracos da placa de C com fronteira externa Γ_i e k_0 o número de buracos da placa ilimitada de C . Some a identidade do teorema anterior para obter:

$$|\text{grau}(F)| + \sum_{1 \leq i \leq n} \mathbf{v}(\Gamma_i) = \sum_{0 \leq i \leq n} (1 - k_i) + \sum_{1 \leq i \leq n} \kappa_{\text{int}}(\Gamma_i) + \sum_{1 \leq i \leq n} \kappa_{\text{ext}}(\Gamma_i) - \sum_{1 \leq i \leq n} \mathbf{v}(\Gamma_i)$$

ou, equivalentemente,

$$|\text{deg}(F)| = \kappa - 2 \sum_{1 \leq i \leq n} \mathbf{v}(\Gamma_i) + \sum_{0 \leq i \leq n} (1 - k_i).$$

Falta assim apenas mostrar que o último somatório é sempre igual a 1 para qualquer conjunto de n curvas. Para $n = 0$ ou $n = 1$ isto é trivial. Por outro lado, se acrescentarmos a um conjunto de curvas uma pequena curva simples sem nenhuma das antigas dentro, o somatório não se altera. De fato, ganhamos mais um termo igual a 1 mas outro termo cai de 1: aquele correspondente à placa onde entrou a nova curva, que agora tem um buraco a mais. ■

O Teorema 4.1 dá um conjunto de condições necessárias, uma em cada placa de C , que devem ser satisfeitas para conjuntos de curvas candidatas a formar o conjunto crítico de uma função cordata, como descrito no começo do capítulo. Note, aliás, que, dados C e $F(C)$ de uma função cordata F , o Corolário 4.2 calcula seu grau e, a partir disso, podemos obter o número de pré-imagens de qualquer ponto usando os Teoremas 3.4 e 3.6.

Considere a função $F(z) = z^7 + (5 + 10i)\bar{z}^4 + 4z$ (em notação complexa). Os conjuntos C e $F(C)$ estão desenhados na Figura 4.3. O número total de cúspides é 16 e as girotões das imagens das curvas críticas são 1 e 4. Pelo Corolário 4.2, $\text{grau}(F) = 7$, como deveria ser (por quê?).

Exercício: Confira o Teorema 4.1 nas quatro placas de C deste exemplo. Aliás, quantas pré-imagens tem a origem?

Figura 4.3

Suponha dada uma função cordata F , para a qual encontramos duas curvas críticas sem cúspides, como indicadas na Figura 4.4, junto com suas imagens. O Teorema 4.1 é satisfeito nas três componentes do plano menos estas duas curvas, entretanto C tem que conter uma curva crítica adicional. De fato, pelo Corolário 4.2, $|\text{grau}(F)| = 1$, e, pelo Teorema 3.6, existiriam pontos com um número negativo de pré-imagens, um absurdo.

Figura 4.4

Exercício: Encontre um desenho possível para C e $F(C)$ na situação acima, acrescentando curvas críticas.

A moral é que convém encontrar condições suficientes (e necessárias) completando o Teorema 4.1 — este é o conteúdo da próxima seção.

Seção 4.2: O Teorema de Blank

Seja Γ um círculo no plano orientado positivamente e $f : \Gamma \rightarrow \mathbb{R}^2$ uma imersão normal. Considere na imagem $f(\Gamma)$ a orientação induzida por f e pela orientação de Γ . O Teorema de Blank dá uma condição

necessária e suficiente para a existência de uma imersão F que preserva orientação, indo do disco D limitado por Γ para o plano, estendendo f . Começamos com um exemplo simples. Considere as curvas Γ e $f(\Gamma)$ na Figura 4.5(a). Como poderíamos decidir se existe uma imersão F estendendo f ao disco limitado por Γ ? Neste caso existe tal F , e ela pode ser construída justapondo três homeomorfismos levando três sub-discos no domínio em três discos na imagem, como indicado na Figura 4.5(b), coincidindo nas partes comuns das fronteiras.

Figura 4.5(a,b,c)

O Teorema de Blank dá uma descrição combinatória da busca de uma imersão a partir de uma decomposição do domínio em discos, dentro dos quais a F existe. As definições a seguir estão exemplificadas na Figura 4.5(c). Uma *família de raios* para f é um conjunto de imersões $r_i : [0, +\infty) \rightarrow \mathbb{R}^2$, os raios, com as seguintes propriedades:

- cada $r_i(0)$ está em alguma placa limitada de $f(\Gamma)$,
- para cada placa limitada de $f(\gamma)$ existe um único i tal que $r_i(0)$ é um ponto da placa,
- $\lim_{t \rightarrow +\infty} r_i(t) = \infty$,
- os r_i são injetores, e suas imagens são disjuntas,
- os r_i nunca passam por interseções de $f(\Gamma)$,
- as interseções de r_i com $f(\Gamma)$ são transversais.

Faremos o abuso de notação habitual de chamar de raios também às curvas orientadas de $r_i(0)$ ao infinito, parametrizadas pelos r_i . Como $f(\Gamma)$ e cada r_i são orientados, as interseções naturalmente ganham uma orientação: quando a curva atravessa o raio da direita para a esquerda, dizemos que a orientação da interseção é positiva, e caso contrário, negativa. Cada interseção tem também uma *altura* associada a ela: o número de outras interseções no mesmo raio e mais perto de sua origem. Assim, a primeira interseção do raio com a curva tem altura 0. A *palavra de Blank* é obtida percorrendo $f(\Gamma)$ uma única vez seguindo sua orientação, e, a cada interseção, escrevendo uma letra correspondente ao raio, o sinal da interseção e sua altura. Na verdade, a altura não é necessária para o enunciado do Teorema de Blank mas é necessária na versão que demonstraremos bem como no Teorema de Troyer: quando desnecessárias, alturas serão omitidas da palavra. A palavra está definida a menos de uma permutação cíclica, isto é, qualquer interseção pode ser tomada como começo da palavra. No exemplo da Figura 4.5(c), a palavra de Blank é

$$P = a_0^+ b_2^+ c_1^+ d_1^+ e_0^+ f_1^+ b_1^+ c_0^+ f_0^+ g_0^+ b_0^- d_0^- e_1^+ f_2^+ g_1^+.$$

Uma palavra de Blank admite uma *simplificação* se existe um par de letras z^+ , z^- , tal que, depois de uma permutação cíclica se necessário, não existem letras com expoente negativo de z^+ a z^- . Neste caso, a simplificação é a palavra obtida eliminando o trecho de palavra entre z^+ e z^- , e diremos que z^- foi *cancelada* com z^+ . Uma simplificação é *positiva* se a altura de z^- é menor que a de z^+ . Dizemos que uma palavra de Blank *grupa* (ou *admite um grupamento*) se podemos simplificá-la sucessivamente até chegarmos a uma palavra sem expoente negativo. Um grupamento é *positivo* quando todas as suas simplificações são positivas. No exemplo acima, um grupamento positivo da palavra é dado por

$$P \rightarrow a_0^+ b_2^+ c_1^+ d_1^+ e_0^+ f_1^+ d_0^- e_1^+ f_2^+ g_1^+ \rightarrow a_0^+ b_2^+ c_1^+ e_1^+ f_2^+ g_1^+.$$

Indicaremos as simplificações ligando as duas letras de sinais opostos por colchetes como no exemplo abaixo:

$$P = a^+ b^+ c^+ d^+ e^+ f^+ b^+ c^+ f^+ g^+ b^- d^- e^+ f^+ g^+.$$

Observe que em um grupamento, colchetes não podem se interceptar. Como a palavra é circular, para cada simplificação podemos escolher entre os dois colchetes que cobrem trechos complementares da palavra: estes colchetes são equivalentes quanto à possibilidade de gruparmos a palavra a partir desta simplificação. Uma vez fixado uma representação linear da palavra, podemos supor então sem perda que todos os colchetes

cobrem um trecho conexo da palavra. Informalmente, simplificações descrevem como dividir o disco no domínio em subdiscos dentro dos quais definimos homeomorfismos que serão colados para construir F .

Teorema 4.3 (Blank,[B]): *Seja Γ regular, fechada simples, orientada positivamente, $f : \Gamma \rightarrow \mathbb{R}^2$ e uma família de raios como acima. Então existe uma imersão F estendendo f ao disco limitado por Γ se e somente se*

- (a) a *giração* de $f(\Gamma)$ é igual a 1,
- (b) a *palavra* de Blank grupa.

Não demonstraremos este teorema, mas sim outra versão, por duas razões: a demonstração de nossa versão é mais simples e seu enunciado, mais apropriado para implementação algorítmica.

Teorema 4.4: *Seja Γ regular, fechada simples, orientada positivamente, $f : \Gamma \rightarrow \mathbb{R}^2$ e uma família de raios como acima. Então existe uma imersão F estendendo f ao disco limitado por Γ se e somente se*

- (a) a *giração* de $f(\Gamma)$ é igual a 1,
- (b) a *palavra* de Blank grupa positivamente.

Vamos precisar de alguns lemas técnicos.

Lema 4.5: *Sejam Γ , f e uma família de raios como acima. Se a palavra de Blank não tem expoente negativo, a *giração* de $f(\Gamma)$ é estritamente positiva. Se além disso, a *giração* é igual a 1, a curva $f(\Gamma)$ é simples.*

Demonstração: Considere a decomposição de Seifert de $f(\Gamma)$. Se existir algum ciclo com orientação negativa, existe um raio começando dentro deste ciclo, e que, ao sair dele, gera uma interseção com sinal negativo. Logo, não existem ciclos com orientação negativa. A afirmação agora segue da Proposição 2.7. ■

Lema 4.6: *Sejam Γ , f e uma família de raios como acima. Se a palavra de Blank grupa positivamente, então a *giração* de $f(\Gamma)$ é maior ou igual a um.*

Demonstração: Definimos a *profundidade* de uma interseção como o número de outras interseções no mesmo raio mais próximas de infinito — assim, profundidade mais alta é o número de interseções no raio menos um. Para cada f e família de raios para f , associamos o polinômio de coeficientes naturais $\sum a_n X^n$, onde a_n é o número de interseções negativas de profundidade n . Definimos uma ordem para os polinômios: $P < Q$ se e somente se o coeficiente de mais alto grau de $Q - P$ é positivo. Precisaremos do seguinte fato, que fica como exercício: qualquer conjunto não vazio de polinômios de coeficientes naturais tem um mínimo.

Suponha por absurdo que o lema seja falso. O conjunto de polinômios associados a contra-exemplos tem então um mínimo. Suponha, a partir de agora, que estamos trabalhando com um tal contra-exemplo mínimo.

Considere a primeira simplificação (positiva por hipótese) associando $f(p_-)$ e $f(p_+)$, interseções da curva $f(\Gamma)$ com um raio. Ligue os pontos p_- e p_+ de Γ por uma curva simples regular contida totalmente em D . Defina novas curvas Γ_1 e Γ_2 e discos D_1 e D_2 como indicados na Figura 4.6. Sejam $f_1 : \Gamma_1 \rightarrow \mathbb{R}^2$ e $f_2 : \Gamma_2 \rightarrow \mathbb{R}^2$ tais que, sobre Γ , elas coincidem com f , e sobre o novo segmento, elas seguem o raio. De forma similar ao item (a) da Proposição 2.5, $\tau(f) = \tau(f_1) + \tau(f_2) - 1$.

O problema original então deu origem a dois problemas análogos para duas curvas. Para f_2 e Γ_2 , vamos agora descrever uma família de raios para f_2 . Tome os raios do problema original, movendo ligeiramente para a esquerda o raio sobre o qual foi feita a simplificação (acompanhe o argumento na Figura 4.6): as duas razões pelas quais este conjunto de raios pode não servir para este novo problema são a existência de dois ou mais raios partindo da mesma placa de $f_2(\Gamma_2)$ e de placas de $f_2(\Gamma_2)$ das quais não parte nenhum raio. A primeira dificuldade se resolve simplesmente omitindo alguns raios. As únicas placas nas quais pode ocorrer a segunda dificuldade são aquelas que vêm imediatamente à direita do segmento do raio original indo de $f(p_-)$ a $f(p_+)$. Em qualquer caso, podemos construir raios para estas placas atravessando o raio original e acompanhando-o até infinito. Assim, a palavra para f_2 , Γ_2 e a família de raios descrita só tem

interseções positivas: isto segue do fato que estamos trabalhando com uma simplificação, de modo que não há interseções negativas no trecho da curva original entre $f(p_-)$ e $f(p_+)$. Pelo lema anterior, $\tau(f_2) \geq 1$.

Figura 4.6(dom e im)

Para f_1 e Γ_1 , tome os raios do problema original, movendo ligeiramente para a direita o raio sobre o qual foi feita a simplificação (Figura 4.7). Omita e acrescente raios exatamente como acima. Afirmamos que a palavra associada a esta família de raios também grupa positivamente. De fato, para cada interseção negativa sobre um raio antigo, a simplificação que cancelava esta interseção na palavra para f continua dando conta da interseção. Para os novos raios, podemos imitar os grupamentos ao longo do raio que eles acompanham.

Figura 4.7

Note que o polinômio associado a f_1 é estritamente menor que aquele associado a f . De fato, ao passar de f para f_1 , eliminamos uma interseção negativa e qualquer nova interseção negativa que surja deve ter profundidade menor. Note também que eventuais eliminações de raios só podem diminuir o número de interseções e a passagem de Γ para Γ_1 só pode diminuir profundidades de interseções.

Como por hipótese f corresponde a um contra-exemplo mínimo, temos $\tau(f_1) \geq 1$. Assim, sabemos que $\tau(f_1) \geq 1$ e $\tau(f_2) \geq 1$, donde $\tau(f) = \tau(f_1) + \tau(f_2) - 1 \geq 1$, contradizendo a hipótese de absurdo. ■

Demonstração do Teorema 4.4: Vamos mostrar inicialmente como a existência de uma extensão F induz um grupamento positivo na palavra de Blank associada a uma família de raios qualquer. Para cada interseção de $f(\Gamma)$ com a família de raios, vamos indicar sua (única) pré-imagem pela letra correspondente na palavra de Blank. As imagens inversas dos raios por F (mostradas em um exemplo na Figura 4.8) satisfazem as seguintes propriedades, cuja verificação fica como um exercício:

- cada componente conexa é uma curva regular com extremos,
- estas curvas são disjuntas,
- as curvas têm uma orientação natural induzida pela orientação dos raios,
- as curvas sempre começam ou em uma interseção negativa ou em uma pré-imagem da origem do raio,
- as curvas sempre terminam em uma interseção positiva,
- toda interseção negativa é o ponto de partida de uma curva,
- curvas que começam e terminam em Γ associam interseções negativas a interseções positivas de altura maior.

Figura 4.8

O grupamento positivo é definido associando as interseções ligadas pelas curvas que começam e terminam em Γ .

Vamos usar os lemas acima (e suas demonstrações) para mostrar o outro sentido do teorema. Novamente, associe a cada situação o mesmo polinômio com coeficientes naturais da demonstração do lema e suponha que estamos trabalhando com um contra-exemplo mínimo.

Dada uma simplificação envolvendo p_- e p_+ , efetue a construção na demonstração do Lema 4.6 para obter f_1 , f_2 e famílias de raios como no Lema. Como $\tau(f) = 1$, $\tau(f_1) \geq 1$ e $\tau(f_2) \geq 1$, temos $\tau(f_1) = \tau(f_2) = 1$. Assim, pelo Lema 4.5, $f_2(\Gamma_2)$ é o bordo de um disco, já que a palavra associada não tem expoente negativo e f_2 se estende a uma imersão F_2 em D_2 . Pela hipótese de minimalidade, já que o polinômio para f_1 é menor que o polinômio para f , f_1 também se estende a uma imersão F_1 em D_1 . A justaposição das duas imersões nos dá um homeomorfismo local \tilde{F} (isto é, para qualquer ponto do domínio, existe uma vizinhança aberta restrita à qual \tilde{F} é um homeomorfismo sobre a imagem) estendendo f a D , já que F_1 e F_2 coincidem em sua fronteira comum.

É possível por técnicas clássicas alterar \tilde{F} em uma vizinhança do segmento ligando p_- e p_+ de modo a obter uma imersão F como desejada, mas não vamos dar detalhes. Mais uma vez, isto contradiz a hipótese de absurdo. ■

O teorema acima trata de extensões que preservam orientação. Uma pergunta natural é o que acontece se queremos extensões que revertem orientação.

Corolário 4.7: *Seja Γ regular fechada simples, orientada negativamente, $f : \Gamma \rightarrow \mathbb{R}^2$ e raios como antes. Então existe uma imersão F revertendo orientação e estendendo f ao disco limitado por Γ se e somente se*

- (a) a giração de $f(\Gamma)$ é igual a 1,
- (b) a palavra de Blank grupa positivamente.

Demonstração: Inverta a orientação do domínio, compondo com $R(x, y) = (x, -y)$. Note que Γ é levada em uma curva positivamente orientada $\bar{\Gamma} = R(\Gamma)$, sobre a qual está definida $\bar{f} = f \circ R$. Aplique Blank a \bar{f} : a giração de \bar{f} é igual à de f e as palavras de Blank para as duas funções coincidem. De fato, em $f(\Gamma) = \bar{f}(\bar{\Gamma})$, as orientações induzidas por f e \bar{f} coincidem. Uma imersão \bar{F} , estendendo \bar{f} e preservando orientação corresponde a $F = \bar{F} \circ R$, uma imersão que reverte orientação e estende f . ■

Vamos terminar esta seção com um exemplo devido a Milnor de uma função $f : \mathbb{S}^1 \rightarrow \mathbb{R}^2$ admitindo duas imersões F_1 e F_2 que a estendem para o disco unitário de formas completamente diferentes. Da Figura 4.9(a), a palavra de Blank é $f_0^+ d_1^+ b_2^+ a_0^- e_0^+ d_0^+ b_1^+ g_0^+ a_1^+ c_0^+ e_1^+ b_0^- a_2^+ c_1^+ e_2^+$. Há dois grupamentos diferentes: a imersão indicada na Figura 4.9(b) corresponde ao grupamento

$$f_0^+ d_1^+ b_2^+ a_0^- e_0^+ d_0^+ b_1^+ g_0^+ a_1^+ c_0^+ e_1^+ b_0^- a_2^+ c_1^+ e_2^+,$$

enquanto a imersão em 4.9(c) corresponde a

$$f_0^+ d_1^+ b_2^+ a_0^- e_0^+ d_0^+ b_1^+ g_0^+ a_1^+ c_0^+ e_1^+ b_0^- a_2^+ c_1^+ e_2^+.$$

Blank mostrou que, em um certo sentido, existe uma correspondência entre extensões e grupamentos. Na verdade, a demonstração do Teorema 4.4 dá a construção de uma bijeção entre a classe de grupamentos positivos e a de imersões.

Figura 4.9(a,b,c) (o exemplo de Milnor)

Seção 4.3: O Teorema de Troyer

O Teorema de Troyer é uma generalização do Teorema de Blank para discos com k buracos. Seja A um disco com k buracos no plano, com fronteira exterior Γ_0 orientada positivamente e fronteiras interiores $\Gamma_1, \dots, \Gamma_k$ orientadas negativamente. Sejam $f_i : \Gamma_i \rightarrow \mathbb{R}^2$ parametrizações de uma família normal de curvas regulares. O Teorema de Troyer diz quando existe uma imersão $F : A \rightarrow \mathbb{R}^2$ estendendo as f_i .

Já temos uma condição necessária, dada pelo item (b) do exercício após a Proposição 2.5. Veremos agora uma condição necessária e suficiente envolvendo raios e palavras como no Teorema de Blank. Famílias de raios são definidos exatamente como antes, para placas de $\bigcup_{0 \leq i \leq k} f_i(\Gamma_i)$. Considere nas curvas $f_i(\Gamma_i)$ as orientações induzidas pelas f_i e orientações de Γ_i . Interseções com os raios são ditas positivas ou negativas também como antes. Assim, percorrendo cada $f_i(\Gamma_i)$, temos uma palavra de Blank para cada Γ_i . Uma *concatenação* de duas palavras de Blank a partir de um par de interseções a^- e a^+ , uma em cada palavra, é obtida permutando ciclicamente as duas palavras de modo a colocar a^- o extremo direito e a^+ no extremo esquerdo das respectivas palavras, justapondo as duas palavras assim obtidas e eliminando o par $a^- a^+$ que aparece no meio da justaposição. Uma concatenação é *positiva* quando a altura de a^- é menor que a de a^+ . Uma família de palavras de Blank *grupa positivamente* se existem concatenações positivas dando origem a uma palavra só que por sua vez grupa positivamente.

Teorema 4.8 (Troyer [T]): *Sejam A, Γ_i, f_i , família de raios e palavras de Blank como acima. Então existe uma imersão $F : A \rightarrow \mathbb{R}^2$ preservando orientação que estende as f_i a A se e somente se*

- (a) a soma das girações de f_i é $1 - k$,

(b) as palavras de Blank grupam positivamente.

Note que o Teorema de Troyer para $k = 0$ é o Teorema 4.4. Mais uma vez, vamos precisar de lemas técnicos.

Lema 4.9: *Sejam $g_i : \mathbb{S}^1 \rightarrow \mathbb{R}^2$, $i = 0, \dots, k$, parametrizações de uma família normal de curvas. Construa raios e palavras de Blank como acima para as placas de $\bigcup_i g_i(\mathbb{S}^1)$. Se as palavras assim obtidas grupam positivamente, então a soma das gições das g_i é maior ou igual a $1 - k$.*

Demonstração: Vamos induzir sobre k : o caso $k = 0$ é o Lema 4.6. Suponha $k \neq 0$. Suponha que a primeira concatenação foi feita a partir de $g_m(p_-)$ a $g_n(p_+)$ — é claro que $m \neq n$. Nosso objetivo agora é substituir as funções g_m e g_n por uma nova função $g : \mathbb{S}^1 \rightarrow \mathbb{R}^2$, que junto com as antigas $g_i, i \neq m, n$, forma um conjunto de $k - 1$ curvas parametrizadas para o qual temos uma família de raios e palavras de Blank que grupam positivamente: o resultado segue então por indução já que $\tau(g) = \tau(g_m) + \tau(g_n) + 1$. A construção de g é a inversa da construção de f_1 e f_2 na demonstração do Lema 4.6. Mais precisamente, desenhe os domínios de g_m e g_n como dois círculos concêntricos de orientações opostas e seja A o anel limitado por eles. Ligue p_- a p_+ por um segmento de reta e seja A' o disco fechado obtido ao remover de A uma vizinhança tubular aberta deste segmento, como na Figura 4.10. Seja Γ a fronteira de A' e γ uma parametrização de Γ por \mathbb{S}^1 . Defina g da seguinte forma: nos trechos correspondentes aos círculos originais, g é obtida compondo γ com g_m ou g_n e nos trechos correspondentes ao segmento g acompanha de perto o trecho de raio entre $g_m(p_-)$ e $g_n(p_+)$.

Figura 4.10 (juntando círculos)

Mais uma vez, omitimos raios supérfluos e acrescentamos se necessário novos raios acompanhando o raio da concatenação. Como acima, as novas palavras de Blank grupam positivamente. ■

Demonstração do Teorema 4.8: De novo, um dos sentidos é mais fácil. O item (a) segue do item (b) da Proposição 2.5 (na verdade, ele é o exercício logo abaixo da proposição). Para o item (b), dada F , tome a imagem inversa dos raios sobre A . Novamente, as pré-imagens dos raios satisfazem todas as propriedades listadas na demonstração do Teorema de Blank. Pré-imagens de raios indo da fronteira para a fronteira indicam concatenações ou grupamentos. Falta apenas ver que existem pré-imagens de raios ligando as várias curvas de tal forma que podemos escolher k pré-imagens conexas de raios para serem os k grupamentos necessários para concatenar as $k + 1$ palavras. Assim, falta apenas provar que o conjunto de fronteiras unido às pré-imagens dos raios é conexo. Isto segue do lema acima: se este conjunto fosse desconexo, teríamos $\ell \geq 2$ conjuntos disjuntos de curvas, com $1 + k_j$ curvas, $1 \leq j \leq \ell$, cada um grupando positivamente de maneira independente. Pelo Lema 4.8, a soma das gições em cada grupo seria maior ou igual a $1 - k_j$ e a soma total seria portanto maior ou igual a $2\ell - 1 - k$ (lembre que $\sum(1 + k_j) = 1 + k$), o que contradiz o item (a), já demonstrado, segundo o qual a gição é $1 - k$.

Vamos demonstrar a recíproca, outra vez por indução sobre k . Novamente, o caso $k = 0$ é o Teorema 4.4. Considere a primeira concatenação feita a partir das interseções $f_m(p_-)$ e $f_n(p_+)$, com $m \neq n$. Ligue p_- a p_+ por uma curva simples completamente contida em A e seja A' o disco fechado com $k - 1$ buracos obtido ao remover uma vizinhança tubular aberta desta curva de A , como indicado na Figura 4.11. Seja Γ a componente de fronteira de A' que não coincide com uma componente de fronteira de A .

Figura 4.11 (abrindo um buraco)

Agora, imite a construção no Lema 4.8 que define em Γ uma função f : nos trechos correspondentes a antigas fronteiras, f coincide com as antigas f_m e f_n e nos trechos correspondentes à curva ligando p_- a p_+ , f acompanha de perto o trecho de raio entre $f_m(p_-)$ e $f_n(p_+)$. De novo, omitindo raios supérfluos e acrescentando novos raios que acompanham o raio da concatenação, obtemos uma nova família de raios para um novo problema de existência de imersão para um valor menor de k — o que, por indução, prova o teorema. ■

Um corolário do Teorema de Troyer necessário para a próxima seção é a situação em que o disco A é substituído por uma região ilimitada. Sejam $\Gamma_i, i = 1, \dots, k$ curvas simples fechadas que são as fronteiras internas de uma região (fechada) ilimitada A_∞ , orientadas negativamente, como no caso anterior. Uma função própria $F : A_\infty \rightarrow \mathbb{R}^2$ tem grau em infinito $n = \text{grau}F$ se uma função contínua que a estender ao

plano tiver grau n — é claro que toda extensão contínua de F ao plano é própria e tem o mesmo grau. Para uma imersão própria $F : A \rightarrow \mathbb{R}^2$ vale a mesma forma normal em infinito que demonstramos para funções cordatas com a mesma demonstração.

Para este corolário vamos considerar apenas funções com grau positivo. Precisamos de mais um ingrediente, a *palavra no infinito*, que desempenha o papel da palavra de Blank associada à fronteira externa (agora inexistente). Dada uma família de raios para as f_i , construa inicialmente uma subpalavra anotando as letras dos raios na ordem em que eles chegam em infinito no sentido anti-horário. Por definição, todas as letras na subpalavra recebem um sinal de interseção positivo e uma altura igual a infinito. A palavra no infinito é obtida repetindo esta subpalavra n vezes, onde n é o grau em infinito da imersão procurada. O primeiro exemplo da seção 5.3 descreve a palavra em infinito de uma função de grau 5 (aliás, verifique, olhando para a Figura 5.14, que $\text{grau}F = 5$).

Corolário 4.10: *Sejam Γ_i , A_∞ e uma família de raios como acima e $f_i : \Gamma_i \rightarrow \mathbb{R}^2$ parametrizações de uma família normal de curvas regulares. Então existe uma imersão $F : A_\infty \rightarrow \mathbb{R}^2$ de grau em infinito igual a n ($n > 0$) se e somente se*

- (a) *a soma das rotações das f_i é $1 - k - n$,*
- (b) *as palavras associadas às curvas Γ_i , juntamente com a palavra no infinito, grupam positivamente.*

Demonstração: Para demonstrar cada sentido deste corolário, aplicamos o sentido correspondente do Teorema de Troyer a uma região limitada $S \subseteq S_\infty$ com as mesmas fronteiras interiores que S_∞ e com fronteira externa uma curva simples Γ_0 que passamos a construir.

Sem perda, podemos supor que os raios a partir de certo ponto são linhas retas. Desenhe então uma curva fechada regular $\tilde{\Gamma}_0$ na imagem, parametrizada por $\tilde{\gamma}_0 : [0, 1] \rightarrow \mathbb{R}^2$ com $\tilde{\gamma}_0(0) = \tilde{\gamma}_0(1)$ satisfazendo as seguintes propriedades (veja a Figura 4.12).

- A curva $\tilde{\Gamma}_0$ dá n voltas ao redor da origem e todas as imagens de Γ_i por f_i estão na mesma placa de $\tilde{\Gamma}_0$ que a origem.
- A curva $\tilde{\Gamma}_0$ intercepta cada raio transversalmente exatamente n vezes.
- A curva $\tilde{\Gamma}_0$ tem exatamente $n - 1$ auto-interseções, todas transversais.
- O argumento de $\tilde{\gamma}_0(t)$ é $2\pi nt$

Figura 4.12 (a curva em infinito)

Provamos inicialmente que se existe uma imersão, então valem (a) e (b). Pela forma normal em infinito e a existência de levantamentos, existe $\gamma_0 : [0, 1] \rightarrow \mathbb{R}^2$ com $\tilde{\gamma}_0 = F \circ \gamma_0$. Queremos mostrar que $\gamma_0(0) = \gamma_0(1)$: de fato, novamente pela forma normal em infinito, podemos tomar o argumento de $(\psi \circ \gamma)(t)$ como $2\pi t$, onde ψ é o difeomorfismo que nos dá a forma normal.

Observamos que, a menos de raios que partam das placas criadas por $\tilde{\Gamma}_0$, os raios que já temos compõem uma família de raios para $F|_{\partial A}$. Por outro lado, é evidente que podemos construir raios para essas placas que só tenham interseções positivas e portanto não interferem na propriedade de grupamento positivo da palavra. Assim, a palavra de Blank para $\tilde{\Gamma}_0$ é, a menos de letras inócuas, a palavra em infinito. As condições (a) e (b) do Corolário seguem das correspondentes para o Teorema de Troyer: note que $\tau(\tilde{\gamma}_0) = n$.

Para demonstrar a recíproca, seja Γ_0 qualquer curva simples fechada tendo todas as Γ_i do seu lado de dentro. Defina $f_0 : \Gamma_0 \rightarrow \mathbb{R}^2$ uma parametrização regular de $\tilde{\Gamma}_0$, levando o sentido anti-horário no sentido anti-horário. Seja A como acima e a ele apliquemos o Teorema de Troyer. É claro que as alturas para as letras da palavra de Γ_0 não são infinitas, mas são maiores do que a altura de qualquer outra interseção no mesmo raio em outra curva, de tal forma que as simplificações que, para A_∞ cancelam letras (positivas) na palavra em infinito, continuam correspondendo a simplificações positivas em A . Portanto, as condições (a) e (b) do Corolário (para A_∞) implicam nas condições (a) e (b) do Teorema de Troyer (para A). Existe então uma F estendendo as f_i a A . A existência de uma F estendendo f_0 a $A_\infty - A$ é trivial e isto nos dá uma F sobre A_∞ (que pode não ser suave, mas pode ser tornada suave). ■

Analogamente ao Corolário 4.7, que generaliza o Teorema 4.4 para imersões que revertem orientação, temos generalizações do Teorema 4.8 e do Corolário 4.10 com a mesma demonstração.

Exercício: Enuncie essas generalizações.

Seção 4.4 Caracterização de C e $F(C)$

Voltamos à nossa pergunta original: dados um conjunto de curvas C_\bullet no domínio e $f_\bullet : C_\bullet \rightarrow \mathbb{R}^2$, quando é que podemos estender f_\bullet a uma função cordata F_\bullet de grau n dado?

Vamos tentar montar a extensão F_\bullet definindo-a em cada placa de C_\bullet . Para isso, as principais ferramentas são os Teoremas de Blank e Troyer (mais precisamente, 4.4 e 4.8), mas não podemos entretanto aplicá-los diretamente a placas de C_\bullet pela presença de pontos críticos no bordo das placas. Para contornar esta dificuldade, vamos aplicar os teoremas a curvas próximas a C_\bullet .

Como na introdução deste capítulo, sejam C_\bullet uma união disjunta (não vazia) de curvas regulares orientadas fechadas no plano e f_\bullet definida em C_\bullet parametrizando uma família normal $f_\bullet(C_\bullet)$ de curvas fechadas, onde a cada cúspide de $f_\bullet(C_\bullet)$, está associado um índice com valor int ou ext. Finalmente, seja n um inteiro não nulo. Suponha que estão satisfeitas as seguintes condições.

- (a) Se $\Gamma_0, \Gamma_1 \subseteq C_\bullet$ são o bordo exterior e uma das componentes do bordo interior de uma mesma placa de C_\bullet , então Γ_0 e Γ_1 têm orientações opostas; se Γ_1 é um bordo interior da placa ilimitada, a orientação de Γ_1 deve ser oposta ao sinal de n .
- (b) Se γ é uma parametrização regular de uma curva $\Gamma_i \subseteq C_\bullet$ respeitando orientação então $f_\bullet \circ \gamma$ é uma função suave.
- (c) Se $(f_\bullet \circ \gamma)'(t) = 0$, então o determinante da matriz cujas colunas são $(f_\bullet \circ \gamma)'''(t)$ e $(f_\bullet \circ \gamma)''(t)$ é positivo.

Vamos obter condições necessárias e suficientes para decidir se existe uma F_\bullet cordata de grau n estendendo f_\bullet ao plano todo, com conjunto crítico igual a C_\bullet , sentidos de dobra correspondendo às orientações já dadas em C_\bullet , e tal que toda cúspide é externa se seu índice for ext e interna, se for int.

Observamos que os itens acima são condições necessárias para a existência de F_\bullet . De fato, o item (a) é trivial: basta considerar o sinal de $\det(DF)$ na placa em questão. O item (b) diz que as curvas $f_\bullet(C_\bullet)$ não têm vértices, o que sabemos ser o caso para a imagem do conjunto crítico de uma função cordata. O item (c) diz que em cada cúspide, a orientação induzida por f_\bullet em $f_\bullet(C_\bullet)$ é aquela mostrada na Figura 4.13. Por outro lado já vimos (Teorema 3.6) que, para uma função cordata, o sentido de dobra perto de uma cúspide deve ser novamente aquele da Figura 4.13.

Figura 4.13 (orientação de uma cúspide)

Seja A o fecho de uma placa de C_\bullet , isto é, um disco com k buracos, com fronteira externa Γ_0 e fronteiras internas $\Gamma_i, i = 1, \dots, k$. Diremos que cúspides com índice int sobre Γ_0 e com índice ext sobre as fronteiras internas são *efetivas* na placa correspondente a A , ou, por abuso de notação, em A . Como nos teoremas de Blank e Troyer definiremos famílias de raios, só que agora cada placa de C_\bullet terá sua própria família de raios. A família de raios para A agora incluirá, além de um raio começando em cada placa de $\bigcup_{i=0, \dots, k} f_i(\Gamma_i)$, um raio partindo da imagem de cada cúspide efetiva em A — fora isso, o conceito de raio permanece inalterado. Quando um raio começa em uma cúspide, a própria cúspide será contada como uma interseção (de altura 0), sempre negativa. Defina palavras de Blank, inclusive a palavra em infinito, como antes.

Assim, por exemplo, sejam $n = 2$ e C_\bullet e $f_\bullet(C_\bullet)$ orientadas como na Figura 4.14, onde todas as cúspides têm índice ext. Para a placa ilimitada, com os raios da figura, a palavra de Blank para a única curva em C_\bullet é $a_0^- b_0^- c_0^- d_0^+$ e a palavra em infinito é $a_\infty^+ b_\infty^+ c_\infty^+ d_\infty^+ a_\infty^+ b_\infty^+ c_\infty^+ d_\infty^+$. Verifique que as palavras grupam positivamente. Para a placa limitada, só há um raio, já que não há cúspides efetivas e podemos construir o raio interceptando $f_\bullet(C_\bullet)$ uma única vez. Neste caso, a palavra de Blank correspondente à única curva tem uma única letra positiva.

Figura 4.14 ($z^2 + \bar{z}$ com raios)

Para o exemplo ilustrado na Figura 4.15, com $n = 2$, onde e é a única cúspide com índice int, as palavras de Blank para a placa ilimitada (Figura 4.15(b)) são: $a_0^- g_0^+ d_0^- a_1^+ e_0^- f_0^+ g_1^+ b_0^-$ para a curva e $a_\infty^+ c_\infty^+ f_\infty^+ g_\infty^+ b_\infty^+ d_\infty^+ a_\infty^+ c_\infty^+ f_\infty^+ g_\infty^+ b_\infty^+ d_\infty^+$ em infinito (grupe positivamente!). Para a placa limitada (Figura 4.15(c)), a palavra para a curva é $g_0^+ e_0^- f_0^+ g_0^+ e_0^+$.

Figura 4.15(a,b,c) (raposinha com raios)

Passamos agora a construir uma F_\bullet satisfazendo as exigências deste exemplo. A compatibilidade das orientações (como em (a)) está satisfeita. A orientação negativa da curva em C_\bullet diz que F_\bullet deve preservar orientações na placa ilimitada de C_\bullet e reverter orientações na placa limitada. Inicialmente, estenda f_\bullet a uma vizinhança tubular de C_\bullet de tal forma que o conjunto crítico da extensão \tilde{f}_\bullet é C_\bullet , \tilde{f}_\bullet só tem dobras e cúspides, o sentido de dobra é o horário sendo e a única cúspide interna. As condições (b) e (c) acima, que estão satisfeitas neste exemplo, garantem que isto pode ser feito. Sejam $\tilde{\Gamma}_0$ e $\tilde{\Gamma}_1$ curvas regulares simples fechadas dentro da vizinhança tubular, acompanhando de perto C_\bullet , uma de cada lado, como na Figura 4.15(d). Oriente $\tilde{\Gamma}_0$ e $\tilde{\Gamma}_1$ no mesmo sentido que a curva em C_\bullet . As imagens por \tilde{f}_\bullet de $\tilde{\Gamma}_0$ e $\tilde{\Gamma}_1$ são então como indicadas nas Figuras 4.15(e) e 4.15(f).

Vamos aplicar o Corolário 4.7 para a restrição de \tilde{f}_\bullet a $\tilde{\Gamma}_0$ e o Corolário 4.10 para a restrição de \tilde{f}_\bullet a $\tilde{\Gamma}_1$. Nos dois casos valem as relações entre giros. Por outro lado, as famílias de raios construídas acima são também famílias de raios para as restrições de \tilde{f}_\bullet a $\tilde{\Gamma}_0$ e $\tilde{\Gamma}_1$ e também palavras de Blank para estas aplicações, inclusive a palavra em infinito, coincidem com as palavras construídas acima e portanto grupam positivamente em cada placa. Assim, em ambos os casos, existem as imersões desejadas e, colando-as, temos uma função que só não é a F_\bullet desejada pela possível perda de suavidade sobre $\tilde{\Gamma}_0$ e $\tilde{\Gamma}_1$. Como já mencionamos anteriormente, não trataremos deste problema. Construímos assim a função cordata F_\bullet desejada. Aliás, no Capítulo 1 já foi dito que a função

$$F(x, y) = (-6x^4 - 6x^2y^2 + xy^3 + 6y^4 - x, \frac{25}{24}x^4 + x^3y + x^2y^2 + \frac{1}{6}xy^3 - y^4 - y)$$

tem exatamente o comportamento qualitativo deste exemplo.

Já vimos na seção 4.1 como provar por argumentos de contagem de pré-imagens que não existe para a Figura 4.4 a F_\bullet pedida. Entretanto, se considerarmos o exemplo da Figura 4.16, sendo $n = 1$ e todas as cúspides com índice ext, os critérios de contagem de pré-imagens são inconclusivos. Por outro lado, se repetirmos o processo descrito acima obteremos para a placa ilimitada as palavras de Blank $b^- c^+ f^- g^+ a^+$, $h^- i^+ a^+ d^- e^+$ e palavra em infinito $a^+ b^+ c^+ d^+ e^+ f^+ g^+ h^+ i^+$, que não grupam. Pelo Corolário 4.10, não existe F_\bullet com esta descrição.

Figura 4.16 (os dois lábios, com raios)

O Teorema 4.11 a seguir é a formalização da discussão acima. Seu enunciado e demonstração correspondem ao processo descrito nos exemplos.

Teorema 4.11: *Sejam C_\bullet , f_\bullet , n , orientações e índices para cúspides como acima. Sejam uma família de raios e palavras de Blank para cada placa de C_\bullet , também como acima. Então existe uma função cordata F_\bullet de grau n estendendo f_\bullet e satisfazendo os requisitos dados se e somente se em cada placa de C_\bullet (com fronteira externa Γ_0 , se houver, e fronteiras internas $\Gamma_i, i = 1, \dots, k$, todas orientadas) valerem as seguintes condições:*

(a) Vale a relação

$$\tau(f_\bullet(\Gamma_0)) - \kappa_{\text{int}}(\Gamma_0) + \sum_{i=1, \dots, k} (\tau(f_\bullet(\Gamma_i)) - \kappa_{\text{ext}}(\Gamma_i)) = 1 - k,$$

onde, se não existir fronteira exterior, $\tau(f_\bullet(\Gamma_0))$ deve ser substituído por $|n|$, e $\kappa_{\text{int}}(\Gamma_0)$ ou $\kappa_{\text{ext}}(\Gamma_i)$ denotam o número de cúspides efetivas nas curvas da fronteira.

(b) As palavras de Blank para $\Gamma_i, i = 0, \dots, k$, grupam positivamente, onde, para a placa ilimitada, a palavra em infinito substitui a palavra para Γ_0 .

Demonstração: Para cada placa limitada, considere, como nos exemplos, um disco fechado A com k buracos contido na placa, com fronteiras externa $\tilde{\Gamma}_0$ e internas $\tilde{\Gamma}_i$, $i = 1, \dots, k$. Para a placa ilimitada seja A_∞ com bordos internos $\tilde{\Gamma}_i$, $i = 1, \dots, k$.

Como na demonstração do Teorema 4.1,

$$\begin{aligned}\mathbf{v}(\tilde{\Gamma}_0) &= \tau(f_\bullet(\Gamma_0)) + \kappa_{\text{int}}(\Gamma_0), \\ \mathbf{v}(\tilde{\Gamma}_i) &= \tau(f_\bullet(\Gamma_i)) + \kappa_{\text{ext}}(\Gamma_i), \quad i \neq 0.\end{aligned}$$

Como nos exemplos, as palavras de Blank para a placa coincidem com as palavras de Blank para as curvas $\tilde{\Gamma}_i$

Suponha que exista F_\bullet cordata satisfazendo as condições do enunciado; vamos demonstrar que em cada placa valem os itens (a) e (b). Para placas limitadas, se Γ_0 é orientada positivamente (isto é, $\det(DF) > 0$ na placa), aplicamos o Teorema 4.8 a A : pelo que vimos acima, cada item do Teorema 4.8 implica no item correspondente deste teorema. Se Γ_0 é orientada negativamente, estamos lidando com uma imersão que reverte a orientação: usamos então a formulação análoga ao Corolário 4.7 para o Teorema 4.8. Para placas ilimitadas, aplique o Corolário 4.10 (ou seu análogo para orientações revertidas).

Se estão satisfeitos os itens (a) e (b) para todas as placas, construiremos F_\bullet placa a placa como nos exemplos. Inicialmente, estendemos f_\bullet a uma vizinhança tubular de C_\bullet com sentido de dobra correspondente à orientação de C_\bullet e classificação das cúspides de acordo com os índices dados. Esta construção requer cuidado pelas exigências de que a extensão \tilde{f}_\bullet só tenha dobras e cúspides. Isto aliás só é possível por causa das exigências feitas acima sobre C_\bullet e f_\bullet , especificamente, os itens (b) e (c).

Feito isto, contruimos A com bordos dentro desta vizinhança tubular. Os itens (a) e (b) deste teorema implicam nos itens correspondentes do Teorema 4.8 (ou um de seus corolários). Aplicando então a A o Teorema 4.8 (ou um de seus corolários) podemos concluir que existe uma imersão em A .

Por este processo, construimos uma função que satisfaz nossas exigências para F_\bullet exceto pela questão da suavidade sobre as curvas auxiliares $\tilde{\Gamma}_i$. Suavizando esta função por uma construção que não faremos neste texto obtemos F_\bullet . ■

Seção 4.5: Como usar estes teoremas

Voltemos ao problema de obter o conjunto crítico C de uma função cordata F — como vimos na seção 3.3, conhecer C é fundamental para nosso processo de inversão de F .

Suponha que, ao estudar F , já obtivemos curvas críticas formando um conjunto C_\bullet . Como saber se $C_\bullet = C$? Se este é o caso, devem valer todas as relações descritas nos enunciados dos Teoremas 4.1 e 4.11 (note aliás que o Teorema 4.1 corresponde ao item (a) do Teorema 4.11). Então, se C_\bullet não passa em um dos testes impostos pelo Teorema 4.11, falta pelo menos uma curva crítica, e como os testes se aplicam em cada placa de C_\bullet , sabemos até em que placa procurar pela nova curva. Veremos na seção 5.4 vários exemplos de aplicação deste método.

O Teorema 4.11 dá o melhor teste desta natureza que se pode esperar: se toda a informação conhecida sobre F é seu comportamento numa vizinhança de C_\bullet e este conjunto passa nos testes impostos pelo Teorema 4.11, então é perfeitamente possível que $C_\bullet = C$: de fato, existe uma função cordata F_\bullet que coincide com F em uma vizinhança de C_\bullet e cujo conjunto crítico é C_\bullet .

Vamos descrever duas situações representativas onde não encontramos todas as curvas críticas, mas C_\bullet passa por todos os testes. Na Figura 4.17, o conhecimento da curva crítica Γ_0 e do grau de F não é suficiente para decidir se existem outras curvas críticas. De fato, se $\tilde{\alpha}$ é uma curva simples fechada circundando $F(\Gamma_1)$ e $F(\Gamma_2)$, a imagem das curvas críticas não detetadas, sua pré imagem contem uma curva simples fechada α que contorna Γ_1 e Γ_2 . Se nunca procurarmos curvas críticas dentro do disco limitado por α (o que pode facilmente acontecer se este disco for pequeno ou estiver longe da origem) não teremos razão alguma para

suspeitar que não encontramos todo o conjunto crítico: certamente existe uma função cordata F_\bullet que coincide com F fora do disco limitado por α e é um difeomorfismo neste disco.

Figura 4.17(a,b) (exemplo do artigo)

Por outro lado, poderíamos ter encontrado somente a curva Γ_1 , o que novamente não nos dá nenhuma indicação de que devam existir outras curvas críticas além desta. De fato, o comportamento local de Γ_1 coincide com o de Γ_0 e portanto é também compatível com o comportamento em infinito. No entanto, do ponto de vista global, os comportamentos destas duas curvas são bastante diferentes: por exemplo, o conjunto $F(\Gamma_0)$ tem uma pré-imagem mas $F(\Gamma_1)$ tem três (informalmente, Γ_0 é a curva responsável pelo fato do grau de F ser dois).

Capítulo 5

Vamos descrever agora um programa baseado na teoria desenvolvida nas seções anteriores para estudar funções F do plano no plano. Supomos que existem rotinas que calculam F e suas primeiras derivadas parciais em um ponto dado. O programa inicialmente obtém o conjunto crítico C de F e, com estes resultados, pode desenhar a flor $F^{-1}(F(C))$ ou inverter pontos da imagem (isto é, resolver sistemas $F(p) = q$, para q dado). O programa é feito para operar com funções cordatas: quando F não o é, recebemos um aviso de erro e a execução é interrompida.

Esperamos que fique claro que algumas das tarefas aparentemente simples que compõem esse algoritmo não são de implementação óbvia, por razões variadas (estabilidade numérica, velocidade de execução, ...). À luz disto e da complexidade do programa, o resultado final é satisfatório mas não tem pretensões de ser ideal: mesmo os aspectos globais do programa são passíveis de melhora substancial. Apesar de não podermos esperar que o programa funcione com qualquer função cordata, nossa experiência mostra que ele desempenha bastante bem mesmo em exemplos difíceis, como veremos mais adiante.

Seção 5.1: Problemas clássicos

No programa, encontramos as seguintes tarefas clássicas em análise numérica:

- o cálculo de uma raiz ou de um extremo de uma função real num intervalo fechado (sem uso de sua derivada),
- o cálculo (sem usar derivadas) de uma solução de um sistema (não linear) de n equações e n incógnitas a partir de uma boa aproximação em um pequeno aberto no qual sabemos que o sistema tem solução única,
- traçado de uma curva de nível de uma função (de novo, sem poder usar derivadas),
- a resolução de equações por meio de métodos de continuação.

Note que todas estas tarefas são de natureza local na escala de emprego do programa. Nosso interesse ao desenvolver o programa consistia em elaborar aspectos globais do algoritmo: as tarefas acima foram isoladas em rotinas que desempenham de forma aceitável, mas que podem facilmente ser substituídas por versões superiores. Referências para estes problemas são [BS], [F], [AG]. Passamos a descrever versões muito ingênuas de soluções para estes problemas.

Para o primeiro, podemos usar o método da secante (ou *regula falsi*): obtemos uma aproximação para a raiz da função, ligando dois pontos no gráfico por uma linha reta, como na Figura 5.1(a). Repita o processo para obter aproximações cada vez melhores. A deficiência mais evidente do método é que existe uma tendência a um dos extremos do segmento se manter fixo durante a iteração, como na Figura 5.1(b). Uma alternativa seria usar outras curvas em vez de retas.

Figura 5.1 (a,b)(regula falsi; porque não é bom)

Para (b), simulamos o método de Newton. Lembramos que o método de Newton para resolver a equação $F(p) = 0$ consiste em passar de uma solução aproximada p_n para uma aproximação supostamente melhor $p_{n+1} = p_n - (DF(p_n))^{-1}(F(p_n))$ (na Figura 5.2, ilustramos uma iteração no caso unidimensional). Para evitar o cálculo das derivadas, podemos estimá-las por aproximações de quocientes de Newton.

Figura 5.2 (Newton)

Suponha para (c) que dois pontos próximos da curva de nível já são conhecidos. Como na Figura 5.3, desenhe uma reta por eles, avance uma distância apropriada até um ponto p da reta, trace uma perpendicular por p e nela marque dois pontos, um de cada lado de p . Se os parâmetros tiverem sido bem ajustados, o intervalo entre estes dois pontos conterá um único ponto da curva de nível, que pode ser encontrado usando (a).

Figura 5.3 (continuando um nível)

Vamos tornar mais preciso o que é um método de continuação como em (d) e o que ele resolve. Suponha que queremos resolver $F(p_1) = q_1$, com q_1 dado, e já conhecemos uma solução p_0 de $F(p_0) = q_0$. Suponha ainda que temos um caminho β na imagem ligando $\beta(0) = q_0$ a $\beta(1) = q_1$. Queremos encontrar p_1 traçando um caminho γ com $\gamma(0) = p_0$ e $F \circ \gamma = \beta$. Para variações pequenas em t , o problema se reduz a resolver $F(\gamma(t_{n+1})) = \beta(t_{n+1})$ com aproximação para $\gamma(t_{n+1})$ dada por $\gamma(t_n)$ — um problema já discutido em (b). Este processo não funciona bem se β passa pela imagem do conjunto crítico.

Seção 5.2: Cálculo do conjunto crítico C

Em linhas muito gerais, o programa começa encontrando algumas curvas críticas, calculando suas cúspides, seus pontos de interseção e giros de suas imagens. A seguir, o conjunto das curvas críticas já obtidas é submetido aos testes descritos na seção 4.5. Se algum dos testes falha, é porque ainda falta calcular curvas críticas, e o programa parte em busca delas (ele até sabe onde procurá-las). Se todos os testes dão certo, o programa entende que já completou o cálculo do conjunto crítico e passa para a etapa seguinte, descrita na seção 5.3.

Para fins expositórios, cada etapa será dividida em passos.

Passo 0: Estudando F em infinito

Se F tem o comportamento esperado em infinito (descrito no Teorema 3.4), a imagem de um grande círculo M centrado na origem deve dar grau(F) voltas em torno da origem. Assim, o cálculo de F em alguns pontos de M pode ser usado para obter grau(F) (como veremos em detalhe no passo 3), além de servir para denunciar (em alguns casos) uma função não cordata.

Passo 1: Encontrando as primeiras curvas críticas

O programa calcula $\det(DF)$, o determinante da jacobiana de F nos pontos de uma grade retangular. A grade é descrita por parâmetros ajustáveis, sendo regular perto da origem e mais esparsa longe daí. Ao encontrar dois pontos vizinhos com determinantes de sinais opostos, uma rotina encontra um ponto crítico (ou seja, uma raiz do determinante) no segmento ligando os dois pontos: esta é uma das tarefas clássicas listadas acima. Uma outra rotina obtém uma lista de pontos sobre a curva por um método de continuação de passo variável que percorre a curva de modo a ter $\det(DF) > 0$ à esquerda. A variação do passo é necessária para levar em conta o fato de que frequentemente na mesma função existem curvas críticas de tamanhos muito diferentes, e que na mesma curva às vezes um pequeno trecho é de maior interesse e dificuldade numéricas. A curva provavelmente atravessa vários segmentos da grade, que são marcados para evitar repetições na busca de outras curvas críticas. Uma vez completada a curva, o programa continua a pesquisa pela grade. No final deste passo, o programa encontrou algumas curvas críticas, formando um conjunto C_\bullet .

Passo 2: Marcando pontos especiais

Para cada curva crítica, obtemos sua imagem e identificamos alguns pontos especiais: extremos (no domínio e na imagem), cúspides e interseções. Os extremos locais das coordenadas x e y , tanto na lista de pontos do domínio quanto na da imagem, são encontrados comparando pontos vizinhos. A curva crítica e sua imagem ficam assim divididas em *arcos bimonotônicos*: segmentos maximais das curvas em que cada coordenada é monótona. A seguir, obtemos retângulos minimais com lados paralelos aos eixos limitando cada curva crítica e sua imagem.

Também procuramos candidatos a vizinhos de cúspides, isto é, pontos p_1 , p_2 e p_3 no domínio para os quais há uma mudança súbita de direção entre os vetores $F(p_1) - F(p_2)$ e $F(p_2) - F(p_3)$, como na Figura 5.4.

Figura 5.4 (achando cúspides grosseiramente)

A seguir, uma rotina (veja (a)) encontra a cúspide neste arco (se é que existe uma) com precisão, juntamente com dois vizinhos, um de cada lado. Estes pontos estão mais perto da cúspide do que o espaçamento típico neste trecho da lista, mas não tão perto a ponto de suas imagens serem indistinguíveis em precisão

de máquina. Finalmente, o programa determina se as cúspides encontradas são internas ou externas, inspecionando o valor de F nos pontos vizinhos às cúspides.

Encontramos também pontos de interseção das curvas críticas (na verdade, isso é feito depois das curvas passarem pelos testes de contagem do passo 4). Por hipótese, as curvas críticas no domínio são disjuntas, mas suas imagens podem se interceptar: em particular, devemos procurar também pontos de auto-interseção. Para identificar nas listas do domínio e imagem os vizinhos de pontos de interseção procedemos da seguinte forma. Dadas duas imagens de curvas críticas, não necessariamente distintas, verificamos inicialmente se seus retângulos limitantes têm uma interseção não vazia. Neste caso, passamos a considerar todos os pares de arcos bimonotônicos, um em cada curva. Mais uma vez, se os dois retângulos limitantes não se interceptam, não há nada a fazer. Caso contrário, existem dois subcasos, ilustrados nas Figuras 5.5(a) e 5.5(b). No caso (a), uma pesquisa binária nas listas de pontos encontra o ponto de interseção (se existir). No caso (b), os pontos de interseção são encontrados pelo algoritmo sugerido na Figura 5.5. Os pontos de interseção são calculados com precisão, pelo método descrito em (b). Este parágrafo ilustra como é conveniente ter à disposição as partições das curvas críticas (e suas imagens) em arcos bimonotônicos.

Figura 5.5 (a,b) (encontrando interseções na lista)

(a) Encontrando cúspides com precisão

Começamos obtendo as direções aproximadas indicadas na Figura 5.6. A cúspide é o extremo da função da reta na reta sugerida pela figura: o problema se reduz então a encontrar o extremo de uma função de \mathbb{R} em \mathbb{R} sem fazer uso de sua derivada, uma das tarefas clássicas. Os novos vizinhos da cúspide são obtidos separadamente.

Figura 5.6 (encontrando cúspides com precisão)

(b) Encontrando interseções com precisão

A esta altura, temos valores aproximados para as interseções, dados pelas interseções das poligonais induzidas pelas listas de pontos. Estas interseções podem ser refinadas: isto corresponde a resolver um sistema de quatro incógnitas (as coordenadas de dois pontos do plano) e quatro equações (a caracterização de dois arcos de curvas críticas e as equações de interseção), mais uma tarefa clássica (veja a Figura 5.7(a)). Devemos, entretanto, ser capazes de reconhecer a situação delicada ilustrada em 5.7(b): neste caso, simplesmente não há interseção.

Figura 5.7(a,b)(interseções com precisão)

Passo 3: Estudo das propriedades topológicas das curvas

Nesse passo, determinamos orientações, números de voltas, giros e relações de inclusão entre as curvas críticas já calculadas. Esta informação é necessária para decidir em que placa (do domínio e da imagem) está um ponto. Vários destes números também são empregados nos testes de contagem do passo 4.

Lembramos que o sentido de dobra de uma curva crítica Γ no domínio é tal que $\det(DF)$ é positivo à esquerda de Γ , e que os pontos de Γ são listados nesse sentido. Existe uma outra orientação conveniente para Γ , aquela que dá a Γ o sentido anti-horário em \mathbb{R}^2 . Usamos um indicador para registrar se as duas orientações coincidem: a segunda orientação é obtida examinando a vizinhança de um ponto cuja imagem tenha coordenada x máxima (esses pontos já foram obtidos no passo anterior).

Vamos descrever agora como calculamos o número de voltas de uma curva em torno de um ponto. Novamente, a decomposição da curva em arcos bimonotônicos é muito conveniente. Suponha que o ponto em questão seja a origem. Freqüentemente os extremos de um arco bimonotônico estão no mesmo quadrante ou em quadrantes adjacentes. Mudanças para um quadrante adjacente somam ou subtraem um quarto de volta do número total de voltas de forma óbvia. Mudanças para um quadrante oposto somam ou subtraem meia volta: usamos então o procedimento a seguir para verificar se a origem está acima ou abaixo do arco.

Encontramos os pontos mais próximos do eixo vertical dentre os pontos da lista que descreve o arco por meio de uma *pesquisa binária*. Isto é, consideramos inicialmente os dois extremos do arco, e um terceiro ponto aproximadamente no meio da lista. Torna-se simples então verificar de que lado deste terceiro ponto está a interseção procurada. O processo se repete sempre dividindo ao meio o segmento de lista pesquisado, até termos dois pontos vizinhos. Se necessitarmos de mais precisão, devemos fazer um pouco de análise

numérica. Se a curva em questão é crítica, basta olhar o sinal de $\det(DF)$ na origem. Se a curva é a imagem de uma curva crítica, temos uma idéia aproximada da posição do ponto desejado dentro de um pequeno arco de imagem de curva crítica entre dois pontos vizinhos. Devemos agora resolver um sistema de duas equações e duas incógnitas a partir de uma boa aproximação da solução, mais uma vez uma tarefa clássica.

Para calcular gições, novamente tomamos arcos bimonotônicos. Vetores secantes com extremos no mesmo arco bimonotônico estão no mesmo quadrante: associamos o quadrante ao arco. Em particular, vetores secantes como mencionados na definição de gição também pertencem ao quadrante associado ao arco correspondente, desde que eles estejam inteiramente contidos neste arco. É fácil calcular a contribuição à gição correspondente à transição entre dois arcos bimonotônicos quando eles se encontram em um extremo suave: os vetores secantes para arcos consecutivos estão em quadrantes adjacentes e a contribuição à gição é mais ou menos um quarto de volta. Cúspides são ainda mais simples: basta somar meia volta à gição (por quê?).

A partir dos ingredientes acima, produzimos facilmente uma rotina que decide se um ponto está no interior do disco limitado por uma curva crítica, determinando assim as relações de inclusão entre curvas críticas no domínio.

Passo 4: Aplicando os testes

Neste passo, não há análise numérica, apenas manipulações combinatórias. Aplicaremos, seqüencialmente, três tipos de testes ao conjunto C_\bullet de curvas críticas já calculadas. Assim que um teste falha, o programa volta a procurar outras curvas críticas em uma placa de C_\bullet . No programa, cada placa recebe o índice de sua fronteira exterior e a componente ilimitada, por convenção, o índice -1 .

O primeiro teste é muito simples. Dadas duas curvas críticas conhecidas, vizinhas imediatas por inclusão, verificamos se os sinais de $\det(DF)$ imediatamente fora de cada curva são diferentes; isto é feito comparando o indicador de orientação associado a cada curva crítica no domínio. Estes sinais devem ser diferentes pois caso contrário deve haver outra curva crítica entre as duas. Já vimos que isto equivale à condição necessária relacionando sentidos de dobra mencionada na seção 4.4 (condição (a)), sem a qual não se aplica o Teorema 4.11.

O segundo teste interpreta o Teorema 4.1 como condição necessária para que já tenhamos encontrado todas as curvas críticas, sendo aplicado em cada placa de C_\bullet .

O terceiro teste é mais complicado: ele corresponde ao Teorema 4.11, da mesma forma que o teste anterior corresponde ao Teorema 4.1. Inicialmente construímos palavras de Blank, e depois procuramos simplificações positivas até uma palavra sem expoentes negativos.

Seja S uma placa de C_\bullet ; de agora em diante, apenas consideraremos curvas críticas na fronteira de F e cúspides que são efetivas dentro de S , isto é, cúspides interiores da fronteira exterior e cúspides exteriores para as outras componentes da fronteira. Primeiro consideraremos o caso onde S é limitado. Neste passo, chamaremos de *arco* a parte da imagem de uma curva crítica entre duas interseções consecutivas dotada de uma orientação (ambas podem ser escolhidas), ou a imagem toda de uma curva crítica que não tem interseções, novamente com uma orientação. Assim, cada pedaço da imagem de uma curva crítica entre duas interseções dá origem a dois arcos com orientações opostas. Dobrando sempre à direita nas interseções, vamos de um arco para seu *sucessor*. Um *ciclo* é uma seqüência fechada de sucessores, como na Figura 5.8.

Figura 5.8 (ciclos e arcos)

Fronteiras de placas de $F(\partial S)$ são ciclos. Um ciclo é a fronteira exterior de uma placa de $F(\partial S)$ se e somente se ele tem gição igual a -1 . Consideraremos um ciclo imaginário em infinito com gição -1 , de tal forma que todas as placas têm uma fronteira exterior. É fácil verificar inclusões entre ciclos calculando números de voltas: isso nos diz quais fronteiras internas correspondem a uma dada fronteira externa.

Para cada placa de $F(\partial S)$, ou equivalentemente, para cada fronteira exterior, vamos definir indutivamente a *profundidade* e um *arco de fuga*. A profundidade da placa ilimitada é zero. Repetimos o processo a seguir até que todas as fronteiras externas próprias (isto é, diferentes da fronteira em infinito) tenham sido levadas em conta. Para cada placa limitada P , verificamos se um dos arcos α de sua fronteira exterior,

quando visto com orientação invertida, é parte da fronteira de uma placa Q cuja profundidade já tenha sido definida. Neste caso, definimos a profundidade de P como 1 mais a profundidade de Q e tomamos α como o arco de fuga de P . Arcos de fuga são indicados por setas na Figura 5.9, e são as posições por onde os raios que construiremos sairão da placa. Pela nossa construção, começando em qualquer placa limitada e saindo através de seu arco de fuga, chegaremos à placa ilimitada.

Os raios de Blank agora são fáceis de construir. Na verdade, a única informação necessária sobre os raios é a combinatória de suas interseções com arcos (por exemplo, na Figura 5.9, ligue as interseções associadas a cada raio de maneira que os raios não se encontrem). Comece pelas placas mais fundas de $F(\partial S)$ e encaminhe-se para as mais rasas recursivamente. Em um passo de indução típico, há raios entrando já construídos em passos anteriores saindo de placas mais fundas para a placa de $F(\partial S)$ sendo considerada. Neste ponto, temos uma lista de raios (incompletos) e uma lista de interseções com arcos de fuga de placas mais profundas, já tratadas pela indução, onde cada interseção guarda sua altura (isto é, a posição da interseção no raio). A cada grupo de raios vindo de uma dada componente da fronteira, adicionamos um raio para cada cuspide neste ciclo que deixa a placa de $F(\partial S)$ em questão à direita (na Figura 5.9, os arcos associados a cúspides são b, c, d, f, g, h e l). A ordem dos raios em um grupo é dada, dentro de cada arco, por passos indutivos anteriores e, mais globalmente, pela ordem dos arcos no ciclos. Se estivermos tratando da fronteira externa, devemos considerar a posição dos vários arcos, inclusive o de fuga, na fronteira. Forçamos todos os raios a saírem através do arco de fuga, preservando a ordem dos raios em cada grupo: a ordem entre os grupos é arbitrária. Finalmente, acrescentamos um último raio para a própria placa de $F(\partial S)$ (na Figura 5.9 os arcos associados a placas são a, e, i, j, k e m). Este processo dá uma descrição completa do padrão combinatório mencionado acima para uma escolha satisfatória de raios de Blank. A palavra de Blank para cada curva é obtida simplesmente seguindo a curva e registrando as interseções de raios com seus arcos, incluindo as interseções formais nas cúspides definidas no Capítulo 4. Já observamos que o sinal nas cúspides é sempre negativo: os outros sinais das interseções são obtidos comparando a orientação do arco de fuga relevante com o sentido de dobra da curva. Na Figura 5.9, as palavras obtidas são as seguintes. Para o lábio (isto é, a curva simples com duas cúspides), a palavra é $j_2^+ k_1^+ l_1^- h_1^-$, para a curva com três cúspides, $a_1^+ b_1^- c_1^- d_1^-$, para a curva com duas voltas, $a_3^+ b_3^+ c_3^+ d_3^+ e_2^+ f_2^+ g_2^+ h_2^+ i_1^+ j_1^+ a_2^+ b_2^+ c_2^+ d_2^+ e_1^+ f_1^- g_1^-$ e para a curva de fora, $a_4^+ b_4^+ c_4^+ d_4^+ e_3^+ f_3^+ g_3^+ h_3^+ i_2^+ j_3^+ k_2^+ l_2^+ m_1^+$.

Figura 5.9 (raios e setonas de fuga)

Para terminar, verificamos por exaustão se as várias palavras de Blank podem ser grupadas positivamente, como descrito no Capítulo 4. No exemplo da Figura 5.9, começamos descartando os raios a, e, i, j, k e m , por só terem interseções positivas, sendo portanto inócuos. As palavras então se reduzem, depois de permutações circulares, às expressões $P_1 = l_1^- h_1^-$, $P_2 = c_1^- d_1^- b_1^-$, $P_3 = g_1^- b_3^+ c_3^+ d_3^+ f_2^+ g_2^+ h_2^+ b_2^+ c_2^+ d_2^+ f_1^-$ e $P_4 = b_4^+ c_4^+ d_4^+ f_3^+ g_3^+ h_3^+ i_2^+$. O Diagrama 5.10 exhibe uma concatenação e grupamento.

Diagrama 5.10 (concatenando e grupando uma palavra)

O algoritmo acima trata das placas limitadas S de C_\bullet . A placa ilimitada S_∞ é tratada da forma descrita no Teorema 4.11: a palavra no infinito é de fácil construção.

Passo 5: Procurando mais curvas críticas

Suponhamos que uma placa S de C_\bullet não passou em algum dos testes anteriores: procuramos então um ponto crítico em S , buscando dois pontos p_0 e p_1 em S nos quais $\det(DF)$ tem sinais opostos e tal que o segmento $p_0 p_1$ está contido em S : em outras palavras, o segmento não intercepta nenhuma curva crítica conhecida. Isso é necessário para garantir que, ao final do processo, não encontremos de novo uma curva crítica já conhecida.

Uma violação do primeiro teste é fácil de tratar. Suponha S limitada. Curvas na fronteira de S se dividem em duas classes de equivalência. Fronteiras interiores são *boas* se sua orientação dada pelo sentido de dobra é oposta à da fronteira exterior; as demais são *más*. A fronteira exterior é boa por definição e a violação do primeiro teste significa que existe pelo menos uma curva má. Seja p' o ponto nas curvas más com coordenada y mínima. Dentre os pontos em curvas boas com coordenada y menor ou igual à de p' , tome p'' como o mais próximo de p' com a distância definida por $d((x_0, y_0), (x_1, y_1)) = \max\{|x_0 - x_1|, |y_0 - y_1|\}$,

obtido como em (a) abaixo. Os pontos p_0 e p_1 são obtidos aproximando ligeiramente p' e p'' um do outro ao longo do segmento $p'p''$. O caso S ilimitado é semelhante e deixado como exercício.

A partir de agora, supomos que a região passou pelo primeiro teste. Um ponto p é *fácil* se $\det(DF)(p)$ tem mesmo sinal que $\det(DF)$ para pontos de S perto da fronteira exterior, ou se esta for a placa ilimitada, o mesmo sinal que o grau topológico de F . Pela nossa suposição, todo ponto perto da fronteira da placa é fácil. Seja p um ponto aleatório em S (para obter esse ponto, empregue a rotina que já descrevemos para decidir se um ponto está em S). Se p for um ponto difícil (isto é, não fácil), podemos tomar $p_0 = p$ e p' o ponto na fronteira de S que está mais perto de p_0 , usando a noção de distância definida no parágrafo anterior. Obtemos p_1 aproximando ligeiramente p' de p_0 ao longo do do segmento p_0p' . Se p for um ponto fácil, usamos um método gradiente modificado, a ser descrito em (b), para a função $\det(DF)$ de forma a encontrar um novo ponto que esperamos ser difícil. Iteramos o método gradiente modificado até obtermos um ponto difícil ou desistirmos e aí recomeçamos a busca com outro ponto aleatório.

Sempre que uma nova curva crítica é encontrada, o programa repete as partes necessárias dos passos 2 a 5 até que todos os testes sejam satisfeitos. O programa então entende que encontrou todas as curvas críticas de F .

(a) Obtendo o ponto mais próximo

Fazemos isto em cada arco bimonotônico, e depois escolhemos o melhor dos resultados. Para a noção de distância que estamos usando, é fácil ver que o ponto mais próximo é um extremo do arco ou uma interseção do arco com uma linha reta passando por p com coeficiente angular ± 1 . Um simples teste dependendo da posição do arco em relação a p nos diz em qual dos casos nos encontramos. No segundo caso, uma pesquisa binária obtém o ponto mais próximo.

(b) O método gradiente modificado

Dado um ponto fácil p , obtemos uma aproximação \mathbf{u} para o gradiente de $\det(DF)$ em p por diferenças finitas. A seguir, sobre a linha que passa por p na direção de \mathbf{u} , resolvemos para p' a equação para a linearização de $\det(DF)$

$$\det(DF)(p) + \mathbf{u} \cdot (p' - p) = -\alpha \det(DF)(p).$$

O parâmetro α é uma constante entre zero e um. Devemos verificar se p' está em S .

Seção 5.3: Invertendo a função e calculando a flor

No começo do processo de inversão, temos a nosso dispor o número de curvas críticas, uma lista de pontos em cada curva crítica Γ_i , incluindo cúspides e pontos de interseção, calculados virtualmente com precisão de máquina. A lista inclui os valores das imagens de todos os pontos (a lista de pontos em $F(\Gamma_i)$), junto com uma coleção de indicadores para as seguintes propriedades de cada ponto:

- (a) se o ponto é uma cúspide, e se a cúspide é interna ou externa,
- (b) se o ponto é um extremo local da coordenada x ou y nas listas Γ_i ou $F(\Gamma_i)$ (há quatro indicadores deste tipo),
- (c) se o ponto é uma interseção na imagem, junto com a orientação desta interseção.

Queremos, então, calcular os valores de p que satisfazem a equação

$$F(p) = q, \quad p, q \in \mathbb{R}^2. \quad (*)$$

Uma operação frequente será a inversão por um método de continuação de F ao longo de um segmento de reta q_0q_1 na imagem para o qual uma pré-imagem p_0 de q_0 é conhecida — mais uma das tarefas clássicas. Observe que não estamos interessados em inverter pontos intermediários, exceto na medida em que isto for necessário para obter a pré-imagem desejada de q_1 . Mais precisamente, teremos uma rotina que tem como entrada os pontos q_0 , q_1 e p_0 e como saída pontos p_1 , p_* e um indicador de desempenho para cada um dos três resultados possíveis a seguir:

- A rotina foi bem sucedida e a pré-imagem desejada é p_1 .
- A rotina descobriu que o processo de continuação não vai além de um ponto p_* . Isto só pode ocorrer quando o segmento q_0q_1 encontra uma imagem de uma curva crítica: p_* será esta interseção.
- Ocorrência de um erro numérico na rotina.

Novamente, quebraremos a descrição do algoritmo em passos.

Passo 0: Calculando todas as pré-imagens de alguns pontos

Tomamos um ponto p do domínio fora do conjunto crítico C , e tal que $F(p)$ também esteja fora de $F(C)$ (isto se faz a partir de informação coletada nos indicadores de tipo (b)). Como na Figura 5.11, ligamos sua imagem $F(p)$ por um segmento horizontal ou vertical I a um ponto próximo q_a em uma das diagonais $y = x$ ou $y = -x$. Sejam q_b, q_c e q_d os pontos na imagem obtidos através de rotações de noventa graus de q_a ao redor da origem. A seguir, tentamos inverter ao longo do segmento I com condição inicial p .

Figura 5.11 (invertendo ao longo do quadrado)

Se tudo correr bem, obtemos assim a primeira pré-imagem p_a do ponto q_a . Usando p_a como uma condição inicial para inverter ao longo do lado q_aq_b , obtemos uma pré-imagem de q_b . Repetindo o processo ao longo dos lados q_bq_c, q_cq_d e q_dq_a do quadrado $q_aq_bq_cq_d$, obtemos (se tudo correr bem) pré-imagens de q_c e q_d e uma outra pré-imagem de q_a . Repetimos então o ciclo de inversões ao longo dos lados do quadrado até obtermos o número total de pré-imagens de cada um destes quatro pontos, número este que deve ser igual a $|\text{grau}(F)|$, que já foi calculado. Este processo é justificado pelo Teorema 3.4. Se o processo de inversão falhar, começamos de novo com p mais longe da origem. Novamente pelo Teorema 3.4, este processo deve funcionar para p suficientemente longe da origem. Em conclusão, temos quatro pontos q_a, q_b, q_c e q_d nas diagonais cujas pré-imagens são todas conhecidas. Isto equivale a termos resolvido o sistema (*) para o lado direito igual a cada um destes quatro pontos. Pontos cujas pré-imagens são todas conhecidas serão chamados de *pontos resolvidos*. Alguns dos pontos resolvidos serão guardados permanentemente na memória em um *banco de pontos resolvidos*.

Passo 1: Achando as pré-imagens de um ponto arbitrário

Seja q um ponto da imagem de F . Ligue q a um ponto q_0 do banco de pontos resolvidos, por um caminho em forma de L, isto é, um caminho formado por dois segmentos, um horizontal e outro vertical (a ordem não é importante), orientado de q_0 a q . Na Figura 5.12, esboçamos um caminho e sua pré-imagem para a função $z \mapsto z^2 + \bar{z}$, em notação complexa. Existem muitos caminhos assim, e escolhemos um favorecendo as propriedades abaixo.

Figura 5.12 (nosso primeiro L)

- (1) Os dois segmentos devem passar longe de imagens de cúspides (que são pontos possivelmente problemáticos para o processo de inversão) e de extremos locais das coordenadas x ou y em $F(C)$ (o contrário geraria interseções quase não transversais, novamente com risco de instabilidade numérica).
- (2) Ambos os segmentos devem ter poucas interseções com $F(C)$.
- (3) Os segmentos devem ser curtos.

A propriedade 1 nos dá a proibição mais forte, e caminhos passando perto demais de uma cúspide ou extremo local são em princípio inaceitáveis; entretanto, se q está próximo de uma cúspide ou extremo local, isso é inevitável. Dentre os caminhos satisfazendo a propriedade 1, procuramos aqueles que tornam mínimo o número total de interseções (propriedade 2) e, havendo empate, tomamos o mais curto (propriedade 3). Não consideramos todos os pontos do banco, mas apenas um pequeno número fixo, e para este fim tomamos aqueles que estão mais próximos de q (novamente, a propriedade 3).

Uma vez escolhido o caminho, executamos um processo de inversão começando no ponto resolvido q_0 ao longo dos dois segmentos. Fazendo isto para cada pré-imagem de q_0 , obtemos algumas pré-imagens de q , apesar de sabermos que às vezes este processo será interrompido em uma interseção com $F(C)$. Dependendo do sentido das interseções com $F(C)$, podem surgir ou desaparecer duas pré-imagens perto de C , como indicado na Figura 5.13.

Figura 5.13 (nascimento e morte de pré-imagens)

Quando duas pré-imagens desaparecem, o processo ‘falha’ como descrito anteriormente. Devemos apenas verificar se estas falhas ocorrem nos lugares e quantidades corretas. Infelizmente, a partir da informação disponível, é impossível saber a priori quais pré-imagens estão fadadas a desaparecer.

Consideremos agora o caso em que duas pré-imagens aparecem. Seja q_* o ponto correspondente no caminho. Calculamos facilmente o ponto p_* no conjunto crítico C cuja imagem é q_* e o núcleo K de $DF(p_*)$. Escolhemos então dois pontos p_1 e p_2 perto de C na reta $p_* + K$, um de cada lado da curva crítica, como na Figura 5.14. As imagens $q_1 = F(p_1)$ e $q_2 = F(p_2)$ provavelmente não pertencem ao segmento que está sendo invertido, mas estarão muito perto dele: para isso utilizamos K . Conectando q_1 e q_2 ao segmento que queremos inverter por pequenos segmentos de reta I_1 e I_2 , invertemos com as condições iniciais evidentes e continuamos o processo de inversão até q ao longo do caminho original.

Figure 5.14 (fazendo nascer pré-imagens perto de dobra)

Se este processo como um todo é bem sucedido, isto é, se as únicas falhas do processo de inversão correspondem ao desaparecimento de pré-imagens, isto nos dá todas as pré-imagens de q . O que pode dar errado? Fazer aparecer um par de novas pré-imagens é uma construção delicada, por envolver métodos de continuação perto do conjunto crítico. A experiência nos mostra entretanto que esta não é uma causa frequente de problemas. Um outro problema é reconhecer que as falhas do processo de inversão ocorreram nos lugares certos, ou seja, numa das interseções conhecidas, onde pré-imagens desaparecem. Um terceiro problema possível surge quando o próprio ponto q está muito próximo da imagem do conjunto crítico: isto não é sério para os métodos de continuação empregados. Finalmente, uma dificuldade mais básica é que existem casos onde nenhum caminho em forma de L é aceitável no sentido da propriedade (1) acima. Uma possibilidade seria procurar caminhos mais complicados ligando q a algum ponto do banco, desviando-nos de regiões perigosas para a inversão. Uma outra possibilidade é esperar até que o banco de pontos resolvidos seja rico o suficiente para conter um ponto que possa ser ligado a q por um caminho em forma de L aceitável. Esta discussão nos leva ao próximo passo.

Passo 2: Gerenciando o banco de pontos resolvidos

É conveniente manter um banco consistindo de alguns pontos da imagem e seus conjuntos completos de pré-imagens, inicializado com q_a, q_b, q_c e q_d já obtidos no Passo 0. Ocasionalmente, acrescentaremos a este banco um novo ponto resolvido, invertido ou por pedido explícito ou por ser o vértice de um caminho em forma de L usado. Em outras palavras, um novo ponto resolvido é sempre obtido a partir de um velho por continuação ao longo de um segmento horizontal ou vertical. Isto permitirá melhoras na escolha de futuros caminhos em forma de L, assim tornando mais pontos acessíveis ao processo de inversão. Existem vários critérios razoáveis para inserção de um novo ponto resolvido no banco. O processo de inversão ao longo de um segmento sem interseções com $F(C)$ é geralmente considerado de baixo custo. Uma alternativa é acrescentar um novo ponto resolvido ao banco se e somente se o segmento (horizontal ou vertical) usado no seu processo de inversão interceptar $F(C)$. Outra possibilidade é acrescentar ao banco todos os pontos resolvidos até atingirmos um máximo previamente estabelecido. As duas estratégias mencionadas são bastante exageradas, mas a segunda foi satisfatória nos testes descritos a seguir.

Passo 3: Calculando a flor

Para calcular a flor, devemos inverter cada ponto nas listas descrevendo $F(C)$. Um ponto está sempre próximo de seus vizinhos na mesma curva, de modo que quase sempre temos boas condições iniciais para sua inversão. Mas as coisas não funcionam tão bem perto de uma cúspide ou de uma interseção. Trabalhamos portanto arco a arco, onde um arco é uma porção de uma das curvas em $F(C)$ entre dois tais pontos problemáticos. A situação para um ponto típico dentro de um arco é muito simples de se tratar numericamente: todas as pré-imagens, exceto uma, estão longe do conjunto crítico C , e esta última, que está em C , já é conhecida explicitamente (é o ponto em C que é levado ao ponto em $F(C)$ que está sendo invertido). Trabalhando arco a arco, podemos então usar o processo previamente discutido de inversão para calcular todas as pré-imagens desconhecidas de um ponto e inverter o resto do arco pelos métodos de continuação habituais. Dificuldades podem aparecer quando as pré-imagens regulares (isto é, pré-imagens que não pertencem a C) se aproximam muito de C . Isto acontece perto dos extremos de um arco, o que, se

o arco for extraordinariamente curto, pode significar todo o arco. Isto é precisamente o que ocorre quando F tem um pequeno rabo de andorinha, ou seja, um par de cúspides próximas, uma para dentro e outra para fora, na mesma curva crítica (há um rabo de andorinha na Figura 1.8). Além disso, duas curvas em $F(C)$ podem estar muito próximas sem se interceptar. Esta última possibilidade pode ser incômoda para o programa, já que ela pode implicar na má inversão de todo um arco (possivelmente longo); este caso não é comum. A atual versão do programa simplesmente abandona a inversão de um arco se o método de continuação falha: isto geralmente faz pouca diferença para a informação visual dada pela flor.

Seção 5.4: Exemplos

Descreveremos algumas execuções do programa. Estes experimentos foram feitos em uma estação Sparc Station SLC (diskless, 8 Mbytes). Inicialmente consideraremos a função (de grau 5)

$$F(x, y) = (x^5 - 10x^3y^2 + 5xy^4 + 6x^2 + 6xy + x, 5x^4y - 10x^2y^3 + y^5 - y).$$

Depois de percorrer uma grade de $61 \times 61 = 3721$ pontos, o programa encontrou três curvas críticas, que foram traçadas com 57, 61 e 157 pontos. Estas curvas e suas imagens são mostradas na Figura 5.15. O programa verificou que os dois primeiros testes eram satisfeitos pelas curvas críticas e a seguir calculou as palavras de Blank associadas à região não limitada no domínio (neste exemplo, as regiões dentro das curvas críticas trivialmente satisfazem o critério da palavra). Seguindo o algoritmo descrito anteriormente, ele encontrou as expressões $a_0^- b_0^- c_0^-$ para a curva 0, $d_0^- e_0^- f_0^-$ para a curva 1 e $b_1^+ c_1^+ d_1^+ e_1^+ f_1^+ g_0^- h_0^- i_0^- j_0^-$ para a curva 2, e construiu a palavra em infinito $b_2^+ c_2^+ d_2^+ e_2^+ f_2^+ g_1^+ h_1^+ i_1^+ j_1^+ a_1^+ b_3^+ c_3^+ d_3^+ e_3^+ f_3^+ g_2^+ h_2^+ i_2^+ j_2^+ a_2^+ b_4^+ c_4^+ d_4^+ e_4^+ f_4^+ g_3^+ h_3^+ i_3^+ j_3^+ a_3^+ b_5^+ c_5^+ d_5^+ e_5^+ f_5^+ g_4^+ h_4^+ i_4^+ j_4^+ a_4^+ b_6^+ c_6^+ d_6^+ e_6^+ f_6^+ g_5^+ h_5^+ i_5^+ j_5^+ a_5^+$ (lembre-se que índices de altura são irrelevantes para a palavra em infinito). O programa verificou que estas palavras grupam positivamente. Neste ponto, o programa considerou que o conjunto crítico era completamente conhecido. Esta parte do programa levou 4.5 segundos: 1.5 s correndo pela grade (exagerada), 1.44 s foram gastos traçando as curvas críticas com detalhe maior do que o necessário, o cálculo preciso das interseções levou 0.22 s e o teste da palavra, 0.23 s.

Figura 5.15 (quintp)

Fizemos então duas experiências para avaliar o desempenho das rotinas de inversão. Produzimos 199 pontos distribuídos aleatoriamente no quadrado $[-1, 1] \times [-1, 1]$, e calculamos todas as suas pré-imagens com nosso algoritmo. Isto levou 13.65 segundos, mais 1.84 s para ler e escrever as respostas em um arquivo. Observe que o quadrado contém uma porção não trivial da imagem do conjunto crítico: há pontos com cinco, sete e nove pré-imagens: o número médio de pré-imagens por ponto resultou ser 5.82. Tomamos então outros 199 pontos distribuídos aleatoriamente no quadrado $[-6, -4] \times [-1, 1]$, distante da imagem do conjunto crítico — mas só sabemos disso porque calculamos o conjunto crítico da função. O programa então calculou todas as cinco pré-imagens de cada ponto, a partir das cinco pré-imagens (previamente conhecidas) do ponto $(-5, 0)$, usando um método de continuação muito simples que em geral era equivalente ao método de Newton habitual. Esta parte do programa levou 2.46 s para cálculos, e 1.64 s para entrada e saída. Consideramos a razão de 6 para 1 entre tempos bastante satisfatória, pelas seguintes razões. O protótipo testado está longe de ser ótimo e a tarefa que ele desempenha é substancialmente mais difícil: só sabemos que o método de continuação habitual encontrou todas as pré-imagens por estarmos limitados a uma região onde já sabíamos que a função era bem comportada. Além disso, os segmentos aparecendo nos métodos de continuação feitos na primeira parte dos testes estavam muito mais perto do conjunto crítico, forçando o processo de inversão a ser mais cuidadoso, e portanto mais lento, mesmo quando os segmentos não interceptam a imagem do conjunto crítico. Obviamente, um número maior de pré-imagens é responsável por mais tempo adicional na primeira parte do teste. Observe que o programa inverteu com sucesso todos os 199 pontos, apesar de sua proximidade provável à imagem do conjunto crítico. O tempo total, incluindo cálculos do conjunto crítico e a inversão dos 2×199 pontos foi de 28.70 segundos, dos quais 16.94 s foram gastos em quase 35000 avaliações da função e sua jacobiana.

No experimento seguinte, usamos uma grade mais modesta de $9 \times 9 = 81$ pontos para busca inicial de curvas críticas da mesma função F definida acima. O programa começou calculando corretamente o grau topológico da função (cinco) e, correndo pela grade, encontrou uma curva crítica que foi traçada com 158 pontos: nela foram encontradas quatro cúspides e a giração calculada de sua imagem foi 1. Verificando as

restrições de contagem, o programa concluiu corretamente que havia curvas críticas faltando na componente ilimitada do complemento da curva crítica calculada. Localizada pelo método de busca aleatória, a segunda curva foi traçada com 57 pontos, e mostrou-se que ela tinha 3 cúspides e giração igual a 1. O programa novamente concluiu pelas restrições de contagem que ainda havia curvas críticas faltando, e encontrou a última curva crítica, traçada com 53 pontos, tendo três cúspides e giração igual a 1. O programa verificou que este conjunto de curvas críticas passava pelo primeiro e segundo testes, e começou a calcular interseções. Duas foram encontradas entre as imagens das curvas 0 e 1, e outras duas entre as imagens das curvas 0 e 2. Isto permitiu ao programa verificar que este conjunto de curvas também satisfazia o terceiro teste (das palavras de Blank).

Agora mostraremos dois exemplos de uso do programa para visualizar o comportamento global de uma função através do cálculo de sua flor $\mathcal{F} = F^{-1}(F(C))$. Considere a família de funções

$$F_t(x, y) = (t(x^4 - 6x^2y^2 + y^4) + x^3 + xy^2 - 2x, t(4x^3y - 4xy^3) - x^2y - y^3 - 2y).$$

As flores mostradas nas Figuras 5.16(a) e 5.16(b) correspondem aos valores $t = 0.15$ e $t = 0.3$ respectivamente. Observe as pequenas pré-imagens dos triângulos na imagem. Observe ainda a presença de um recobrimento duplo na restrição da primeira função à região P .

Figura 5.16 (a,b)(duas flores com triângulo e quadrado críticos)

Em ainda outro experimento, o programa inverteu alguns pontos muito perto da imagem do conjunto crítico da função. Mais precisamente, considere a função $F(x, y) = (x^2 - y^2 + x, 2xy - y)$, com conjunto crítico dado por um círculo de raio $\frac{1}{2}$ centrado na origem (veja a Figura 1.4). A interseção da imagem do conjunto crítico com o semi-eixo positivo y é dada por $(0, \frac{(3-\sqrt{3})\sqrt[4]{12}}{8}) \equiv (0, 0.294989919892746)$, que é a imagem de um ponto de dobra. Para pontos $q_a = (0, 0.2949899198927)$ e $q_b = (0, 0.2949899198929)$, a inversão deu as respostas corretas: o primeiro ponto tem quatro pré-imagens $p_1 = (0.643433123248, 1.028318679856)$, $p_2 = (-1.009458527033, -0.097713820754)$, $p_3 = (0.183018329423, -0.465310690167)$ e $p_4 = (0.183007074333, -0.465294168935)$, e o segundo ponto tem apenas duas pré-imagens, que coincidem com os pontos p_1 e p_2 em todas as doze casas decimais dadas acima. Além disso, aplicando F aos pontos p_1 e p_2 , obtivemos os pontos q_a e q_b com um erro de aproximadamente 10^{-13} : nesta escala, q_a e q_b são indistinguíveis. Além disso, p_3 e p_4 estão de lados opostos da curva crítica: sua distância a ela é da ordem da raiz quadrada da distância do ponto q_a à imagem do conjunto crítico, que é o comportamento esperado de uma função perto de uma dobra. Calculando F em p_3 e p_4 , entretanto, desviamos do ponto q_a por um erro da ordem de 10^{-11} , indicando que a análise numérica não é perfeita mas é satisfatória. Invertamos então alguns pontos muito próximos de $(0.75, 0)$, que é a imagem da cúspide $(0.5, 0)$. O processo de inversão funcionou bem para os pontos $(0.75, 0.00000001)$, $(0.7500000001, 0)$ e $(0.7499999999, 0)$.

Nosso último exemplo é o estudo da função $F(x, y) = (x^2 - y^2 + 20\text{sen}x, 2xy + 20\text{cos}y)$. Como deve ser C ? O conjunto crítico de $(x, y) \mapsto (\text{sen}x, \text{cos}y)$ é uma grade de retas da forma $x = k\pi + \frac{\pi}{2}$ e $y = k\pi$, onde k é um inteiro. Devido ao termo quadrático, F se comporta no infinito como $z \mapsto z^2$, em notação complexa. Em particular, o conjunto crítico de F é limitado e seu grau topológico é igual a 2. Os conjuntos C e $F(C)$ são mostrados na Figura 5.17: observe o grande número de curvas críticas no domínio (17) e o intrincado padrão de interseções na imagem. Nos cálculos para este exemplo, primeiro estimamos uma cota para o conjunto crítico e fizemos uso de uma grade inicial muito fina. De fato, só desta maneira poderíamos esperar encontrar as duas pequenas curvas críticas nos cantos superior direito e esquerdo da figura do domínio: elas são indetetáveis por qualquer teste (lembre-se dos exemplos na seção 4.5). A combinatória relacionada ao teste da palavra é intratável pelo menos para nosso algoritmo, devido ao grande número de palavras não triviais associadas às várias curvas críticas. Nesta execução, inibimos o teste da palavra. Mesmo assim, o programa não teve dificuldade em inverter os pontos $(-32, 32)$, $(8, 32)$, $(-8, 16)$, $(8, 16)$ e $(-2, 10)$, que têm respectivamente, 2, 2, 8, 10 e 10 pré-imagens. Os últimos três pontos foram tomados em uma região da imagem particularmente superpovoada com imagens de curvas críticas.

Figura 5.17 (logotipo PUC)

Bibliografia

- [AG] Allgower, E.L., Georg, K., *Numerical continuation methods: an introduction*, Springer-Verlag, New York, 1991.
- [B] Blank, S.J., *Extending immersions and regular homotopies in codimension one*, Tese de Doutorado, Brandeis University, 1967.
- [BS] Bulirsch, R., Stoer, J., *Introduction to numerical analysis*, Springer-Verlag, New York, 1979.
- [F] Fletcher, R., *Practical methods of optimization*, Wiley-Interscience, Chichester, 1981.
- [L1] Lima, E.L., *Curso de análise, vol. II*, IMPA, Projeto Euclides, 1981.
- [L2] Lima, E.L., *Grupo fundamental e espaços de recobrimento*, 11^o Colóquio Brasileiro de Matemática, IMPA, Poços de Caldas, 1977.
- [Ma] Massey, W.M., *Algebraic topology: an introduction*, Springer-Verlag, New York, 1977.
- [Mi] Milnor, J.W., *Topology from the differentiable viewpoint*, University Press of Virginia, Charlottesville, 1969.
- [Mu] Munkres, J. R., *Elements of algebraic topology*, Addison-Wesley, Menlo Park, 1984.
- [MST] Malta, I.P., Saldanha, N.C., Tomei, C., *The numerical inversion of functions from the plane to the plane*, Preprint Mat-006/92, Departamento de Matemática, PUC-Rio.
- [MT] Malta, I.P., Tomei C., *Singularities of vector fields arising from one-dimensional Riemann problems*. J. Diff. Eqs 94 (1991), 165-190.
- [N] Nehari, Z., *Conformal mapping*, Dover, New York, 1952.
- [P] Poénaru, V., *Extension des immersions en codimension 1*, Séminaire Bourbaki 1967/1968, Exposé 342.
- [S] Sotomayor, J., *Singularidades de aplicações diferenciáveis*, III ELAM, IMPA, 1976.
- [SSMP] Schaeffer, D. G., Shearer, M., *The classification of 2×2 systems of non-strictly hyperbolic conservation laws, with applications to oil recovery*, apêndice por Marchesin, D., Paes-Leme, P. J., Comm. Pure Appl. Math. XL (1987), 141-178.
- [T] Troyer, S. F., *Extending a boundary immersion to the disk with n holes*, Tese de Doutorado, Northeastern University, 1973.
- [W] Whitney, H., *On singularities of mappings of Euclidean spaces, I: mappings of the plane into the plane*, Ann. of Math. 62 (1955), 374-410.