

# **Introdução à Dinâmica de Aplicações do Tipo Twist**



# Publicações Matemáticas

## Introdução à Dinâmica de Aplicações do Tipo Twist

Clodoaldo G. Ragazzo  
USP

Mário J. Dias Carneiro  
UFMG

Salvador Addas-Zanata  
USP



25<sup>o</sup> Colóquio Brasileiro de Matemática

Copyright © 2005 by Clodoaldo G. Ragazzo, Mário J. Dias Carneiro e Salvador Addas-Zanata

Direitos reservados, 2005 pela Associação Instituto

Nacional de Matemática Pura e Aplicada - IMPA

Estrada Dona Castorina, 110

22460-320 Rio de Janeiro, RJ

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger / Sérgio R. Vaz

## 25º Colóquio Brasileiro de Matemática

- A Short Introduction to Numerical Analysis of Stochastic Differential Equations - Luis José Roman
- An Introduction to Gauge Theory and its Applications - Marcos Jardim
- Aplicações da Análise Combinatória à Mecânica Estatística - Domingos H. U. Marchetti
- Dynamics of Infinite-dimensional Groups and Ramsey-type Phenomena - Vladimir Pestov
- Elementos de Estatística Computacional usando Plataformas de Software Livre/Gratuito - Alejandro C. Frery e Francisco Cribari-Neto
- Espaços de Hardy no Disco Unitário - Gustavo Hoepfner e Jorge Hounie
- Fotografia 3D - Paulo Cezar Carvalho, Luiz Velho, Anselmo Antunes Montenegro, Adelailson Peixoto, Asla Sá e Esdras Soares
- Introdução à Teoria da Escolha - Luciano I. de Castro e José Heleno Faro
- **Introdução à Dinâmica de Aplicações do Tipo Twist - Clodoaldo G. Ragazzo, Mário J. Dias Carneiro e Salvador Addas-Zanata**
- Schubert Calculus: an Algebraic Introduction - Letterio Gatto
- Surface Subgroups and Subgroup Separability in 3-manifold Topology - Darren Long and Alan W. Reid
- Tópicos em Processos Estocásticos: Eventos Raros, Tempos Exponenciais e Metaestabilidade - Adilson Simonis e Cláudia Peixoto
- Topics in Inverse Problems - Johann Baumeister and Antonio Leitão
- Um Primeiro Curso sobre Teoria Ergódica com Aplicações - Krerley Oliveira
- Uma Introdução à Simetriação em Análise e Geometria - Renato H. L. Pedrosa

### Distribuição:

IMPA

Estrada Dona Castorina, 110

22460-320 Rio de Janeiro, RJ

E-mail: [ddic@impa.br](mailto:ddic@impa.br) - <http://www.impa.br>

ISBN: 85-244-0228 -8

*às nossas famílias*



## Apresentação:

O estudo qualitativo de algumas equações diferenciais ordinárias muitas vezes se reduz ao estudo de aplicações induzidas em seções transversais (aplicações de Poincaré). Este é o caso, por exemplo da dinâmica em uma vizinhança de uma órbita periódica.

Em seu trabalho pioneiro [16], sobre transformações em superfícies e aplicações à dinâmica, G.D. Birkhoff ressalta a importância da condição de "vorticidade" ("vortical condition") no estudo das transformações que preservam área e que possuem um ponto fixo elíptico.

Uma condição geométrica semelhante aparece também no chamado último teorema geométrico de Poincaré, provado no caso geral somente por Birkhoff. Este teorema afirma que um homeomorfismo que preserve área que possui um anel invariante limitado por duas curvas homotópicas e que restrito às quais suas trajetórias movem-se em direções opostas (torcem), possui infinitos pontos periódicos.

A condição de torção que aqui será chamada de twist (não necessariamente em direções opostas, mas com velocidades diferentes) aparece com frequência em várias situações aparentemente não relacionadas. Por exemplo: o modelo de Frenkel-Kontorova, as geodésicas no toro, as perturbações periódicas de hamiltonianos em dimensão dois, as aplicações do tipo bilhar em curvas convexas, os hamiltonianos com dois ou três graus de liberdade, em particular no problema restrito dos três corpos, a dinâmica numa vizinhança de um ponto fixo elíptico.

Portanto, a teoria das aplicações do tipo twist pode ser considerada como um modelo teórico unificador para vários fenômenos.

O objetivo dessas notas introdutórias é apresentar alguns resultados e exemplos básicos sobre a dinâmica de aplicações do tipo twist, ilustrando o alcance da teoria e tentando motivar o leitor para as questões e técnicas tratadas.

A partir dos resultados fundamentais de Birkhoff, obtidos no início do século XX e após o grande progresso alcançado nas décadas de 60 e de 80, é possível, hoje, enfrentar o desafio de apresentar a teoria de modo introdutório e acessível (assim esperamos) aos estudantes brasileiros do final de graduação.

Estas notas estão organizadas em quatro capítulos. No primeiro,

são apresentados alguns exemplos e são desenvolvidas as primeiras ferramentas, em especial, a função geratriz. Estes instrumentos são aplicados no Capítulo 2 que trata da existência de órbitas periódicas (conforme Birkhoff), de órbitas heteroclínicas e da teoria de Aubry-Mather. Nesta teoria, são obtidos conjuntos invariantes por uma aplicação do tipo twist que são constituídos de órbitas não periódicas, mas que satisfazem propriedades semelhantes às órbitas de Birkhoff, isto é, são "ordenados". Também é dada uma idéia de métodos puramente topológicos para a obtenção de tais conjuntos invariantes.

O Capítulo 3 trata das curvas rotacionais invariantes, que são curvas homotopicamente não triviais invariantes pela aplicação. Um exemplo de existência de tais curvas ocorre nos bilhares que possuem cústicas. São provados dois resultados: o Teorema da Curva Invariante de Birkhoff e o Teorema do Twist de Moser (para o caso analítico).

Finalmente, o Capítulo 4 trata das regiões (homeomorfas a um anel) limitadas por duas curvas rotacionais invariantes e que não contém outra curva rotacional invariante - a região de instabilidade. Prova-se a existência de órbitas que se acumulam (no passado e no futuro) nas duas curvas do bordo do anel. São apresentados também algumas extensões dos resultados para aplicações do tipo twist definidas no toro.



# Sumário

<b>1</b>	<b>Definições e Exemplos</b>	<b>7</b>
1.1	A dinâmica da aplicação “standard” . . . . .	7
1.2	O modelo de Frenkel-Kontorova . . . . .	16
1.3	Aplicações de tipo “twist” e princípios variacionais . . . . .	20
1.4	Billhares . . . . .	30
1.5	Dinâmica local em torno de uma órbita periódica . . . . .	35
1.6	Aplicação de primeiro retorno do laço sela-centro . . . . .	37
1.7	A forma normal de Birkhoff . . . . .	40
1.8	Hamiltonianos periódicos . . . . .	47
1.9	Passando do cilindro para o anel . . . . .	54
<b>2</b>	<b>Órbitas periódicas e conjuntos de Aubry-Mather.</b>	<b>56</b>
2.1	Órbitas periódicas minimizantes . . . . .	56
2.2	Monotonicidade e órbita minimizante . . . . .	65
2.3	Homeomorfismos do círculo e conjuntos de Aubry-Mather . . . . .	75
2.4	Órbitas minimais heteroclínicas . . . . .	79
2.5	A Fórmula de MacKay-Meiss . . . . .	88
2.6	Métodos topológicos . . . . .	91
2.6.1	Órbitas periódicas . . . . .	91
2.6.2	Teorema de Poincaré-Birkhoff . . . . .	95
<b>3</b>	<b>Curvas Invariantes</b>	<b>99</b>
3.1	O Teorema da Curva Invariante de Birkhoff . . . . .	100
3.2	Propriedade variacional das curvas invariantes . . . . .	106
3.3	Inexistência de Curvas Invariantes . . . . .	109
3.4	Cáusticas e curvas invariantes em billhares convexos . . . . .	112

3.5	Bilhares sem curvas invariantes . . . . .	116
3.6	Existência de Curvas Invariantes . . . . .	119
<b>4</b>	<b>Região de instabilidade</b>	<b>143</b>
4.1	Caminhos positivos e negativos . . . . .	145
4.2	Teorema fundamental . . . . .	146
4.3	Um pouco mais sobre a região de instabilidade . . . . .	154
4.4	Extensões para o cilindro infinito . . . . .	154
4.4.1	Aplicações do tipo twist no toro . . . . .	155

# Capítulo 1

## Definições e Exemplos

### 1.1 A dinâmica da aplicação “standard”

Nesta seção mostraremos de uma maneira informal diversos tipos de dinâmica que ocorrem em uma família a um parâmetro de aplicações de tipo “twist”, que será tomada como modelo (neste trabalho não traduziremos a palavra inglesa “twist”, que significa torção, pois torção é uma palavra que, em matemática, já possui outros significados). O modelo escolhido é o chamado “Standard Map”, ou aplicação “Standard”, ou aplicação de Chirikov, ou ainda aplicação standard de Chirikov, devido aos trabalhos pioneiros do físico Chirikov com respeito a mesma (veja [22] e referências ali citadas). A aplicação standard está definida sob o cilindro  $\{(\theta, y)\}$  e é dada por:

$$\begin{aligned}\theta' &= \theta + y', \\ y' &= y - \frac{\alpha}{2\pi} \text{sen}(2\pi\theta)\end{aligned}\tag{1.1}$$

onde:  $\theta$  é o ângulo longitudinal no cilindro, com período igual a um,  $y \in \mathbb{R}$  é a altura no cilindro e  $\alpha$  é um parâmetro real. A aplicação standard está associada de maneira imediata a uma aplicação do plano no plano, que também será chamada aplicação standard, através da identificação  $\theta = x \pmod{1}$ , onde  $x \in \mathbb{R}$ . Tal operação que associa a reta  $\{x\}$  ao círculo  $\{\theta\}$ , através da função  $x \rightarrow x \pmod{1} = \theta$  (denominada de aplicação de recobrimento), é chamada de levantamento

do círculo na reta (diz-se também que a reta recobre o círculo, ver [50]). Note que a aplicação  $x \rightarrow x(\text{mod } 1) = \theta$  é um difeomorfismo local não injetor. Esta é uma das propriedades das aplicações de recobrimento. A aplicação  $\{(x, y)\} \rightarrow \{(\theta, y)\}$  define um levantamento do cilindro para o plano. A aplicação standard levantada ao plano  $\{(x, y)\}$  é dada por

$$\begin{aligned} x' &= x + y', \\ y' &= y - \frac{\alpha}{2\pi} \text{sen}(2\pi x) \end{aligned} \quad (1.2)$$

À primeira vista pode parecer preciosismo a distinção entre  $\theta$  e  $x$  apresentada acima. No fim, a fórmula para a aplicação standard em termos de  $\theta$  e  $x$  é a mesma! Ao longo da leitura o leitor deverá estar atento ao número de vezes que transitaremos entre as duas “interpretações” da aplicação standard, ora usando propriedades topológicas do plano, ora do cilindro. Isto será feito tantas vezes que em muitas não será nem mencionado nem, provavelmente, percebido. Esperamos que a ênfase dada neste início à diferença entre  $\theta$  e  $x$ , aguace a atenção do leitor neste sentido. A importância da duplicidade de interpretações ficará evidente.

A aplicação standard aparece em diversos contextos distintos em física (ver [56]). Em particular, no contexto da física do estado sólido, a aplicação standard está associada ao chamado “Modelo de Frenkel-Kontorova”, que será apresentado e discutido na próxima seção. No contexto da matemática a aplicação standard tornou-se o paradigma de aplicação do tipo twist que preserva área. Espera-se que ela represente para a dinâmica conservativa bi-dimensional o que a aplicação quadrática representa para a dinâmica de aplicações unimodais do intervalo, no sentido de exibir um espectro de comportamentos dinâmicos que seja típico na classe das aplicações em que se insere.

Fora a simplicidade, a aplicação standard possui certas propriedades notáveis. Uma delas, a de twist, será tratada na seção 1.3. Uma outra, muito interessante, é a periodicidade por translações verticais, quer dizer, na direção  $y$ . De fato, da equação (1.1) segue que as imagens dos pontos  $(\theta, y)$  e  $(\theta, y + 1)$  são  $(\theta', y')$  e  $(\theta', y' + 1)$ , respectivamente. Então concluímos que a aplicação standard fica completamente caracterizada por seus valores no cilindro finito  $\{(\theta, y)\}$  com  $y \in [0, 1)$ . Mais ainda, definindo a aplicação de recobrimento

to  $(\theta, y) \rightarrow (\theta, y \bmod 1)$ , do cilindro no toro, segue que a aplicação standard induz um difeomorfismo no toro  $\{(\theta, y \bmod 1)\}$ :

$$\begin{aligned}\theta' &= \theta + y', \\ y' &= y - \frac{\alpha}{2\pi} \text{sen}(2\pi\theta) \pmod{1}\end{aligned}$$

Portanto, sem perda de generalidade, pode-se restringir o estudo da dinâmica da aplicação standard a  $(x, y) \in [0, 1] \times [0, 1]$ , o que é feito nas simulações numéricas mostradas abaixo.

A dinâmica da aplicação standard depende fortemente do parâmetro  $\alpha$ . Veremos que conforme  $\alpha$  cresce de zero até infinito a dinâmica complica-se gradativamente. Primeiramente, para  $\alpha = 0$  temos uma situação muito simples onde a aplicação standard reduz-se a:

$$\begin{aligned}x' &= x + y \\ y' &= y\end{aligned}$$

Isso implica que todos os círculos paralelos  $y = \text{constante}$  são invariantes pela aplicação, ou seja, o cilindro é folheado por círculos invariantes. Círculos invariantes que dão a volta no cilindro (homotopicamente não triviais), são chamados de círculos rotacionais invariantes. Uma aplicação do tipo twist para a qual toda órbita está contida em um círculo rotacional invariante é dita integrável. A Figura 1.1 ( $\alpha = 0$ ) ilustra esta situação. Note que a dinâmica da aplicação standard restrita ao círculo  $y = \bar{y} = \text{constante}$  pode assumir duas naturezas muito distintas. Se  $\bar{y}$  é um número racional  $m/n$ , então após  $n$  iteradas o ponto  $(x, \bar{y})$  será transformado em  $(x + m, \bar{y})$ , que quando projetado no cilindro é igual ao ponto de partida. Neste caso o ponto rodou  $m$  vezes em volta do cilindro após  $n$  iteradas, dizemos que seu número de rotação é  $m/n = \bar{y}$ . Note que, se  $\bar{y}$  é racional então a órbita de todo ponto do círculo  $y = \bar{y}$  possui o mesmo número de rotação e é periódica com o mesmo período (observe que, para contar o número de voltas que o ponto dá no cilindro  $\{\theta, y\}$  usamos o levantamento da aplicação standard ao plano  $\{x, y\}$ , enquanto que a órbita do mesmo ponto só é periódica quando projetada no cilindro). Agora, se  $\bar{y}$  é um número irracional, então a órbita de nenhum ponto do círculo  $y = \bar{y}$  é periódica (se alguma o fosse, chegaríamos a uma contradição). De fato, neste caso a dinâmica da aplicação standard

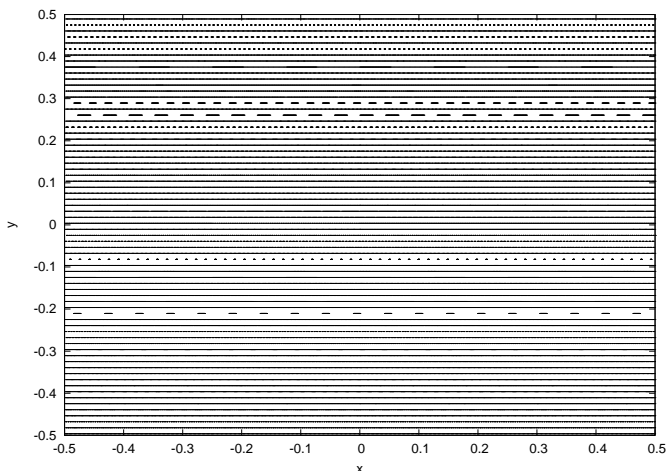


Figura 1.1: Órbitas da aplicação standard para  $\alpha = 0$ , caso integrável.

restrita ao círculo  $y = \bar{y}$  é a de uma rotação rígida irracional, e a órbita de qualquer ponto é densa no círculo. Dizemos que tal círculo invariante possui número de rotação irracional.

O próximo grau de complexidade ocorre quando  $\alpha$  é pequeno mas não é nulo. Neste caso um teorema importante devido a Kolmogorov, Arnold e Moser (KAM) garante que os círculos rotacionais invariantes do caso integrável  $\alpha = 0$ , associados à “maioria” dos números de rotação irracionais, são perturbados em sua forma, mas continuam invariantes e com a dinâmica de uma rotação rígida associada ao mesmo número de rotação. O enunciado preciso e a demonstração de uma das versões do teorema KAM, serão dados no capítulo 3 destas notas. A dinâmica de uma aplicação do tipo twist nestas condições é dita quase-integrável. A figura 1.2 ( $\alpha = 0.1$ ) ilustra tal situação. Note que quase toda órbita mostrada nesta figura parece estar contida em uma curva invariante.

O teorema KAM não diz nada quanto ao que ocorre com a dinâmi-

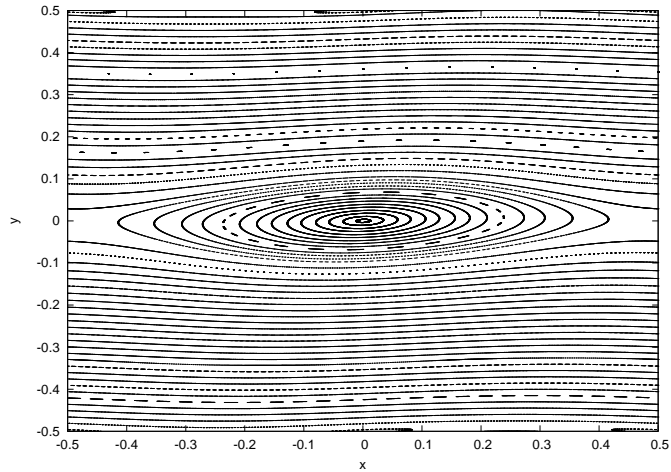


Figura 1.2: Órbitas da aplicação standard para  $\alpha = 0.1$ , caso “quase integrável”.

ca da aplicação standard próxima aos círculos rotacionais invariantes com número de rotação racional. Observa-se, e em grande parte pode se demonstrar, que tais círculos dão origem a um número finito de órbitas periódicas, todas com o mesmo número de rotação. A região entre dois círculos rotacionais invariantes, onde situam-se tais órbitas periódicas, é chamada de região de ressonância (o nome está ligado à racionalidade do número de rotação). As regiões de ressonância para quase todos números de rotação racionais são muito pequenas quando  $\alpha$  é pequeno. Por exemplo, para  $\alpha = 0.1$ , figura 1.2, a única região observável na escala da figura é aquela associada ao número de rotação zero. O círculo invariante com número de rotação nulo, que existia para  $\alpha = 0$ , foi substituído por dois pontos fixos. O primeiro, no centro da figura, é de tipo elíptico, o que significa que a derivada de  $F$  no ponto fixo está associada a autovalores complexos  $\lambda$  e  $\lambda^{-1}$ , com  $|\lambda| = 1$ . Tal ponto fixo é circundado por círculos invariantes que não enlaçam o cilindro (são homotopicamente triviais). O ponto

elíptico mais seus associados círculos invariantes são chamados de “ilha elíptica”. O segundo ponto fixo, localizado próximo a  $(x, y) = (0.5, 0)$ , é instável e de tipo hiperbólico, ou seja, possui autovalores reais  $\lambda$  e  $\lambda^{-1}$  com  $\lambda \neq 1$ .

As zonas de ressonância crescem com o aumento de  $\alpha$ . Na Figura 1.3 ( $\alpha = 0.8$ ) são mostradas algumas regiões de ressonância com suas respectivas ilhas elípticas. Note que, agora, a grande re-

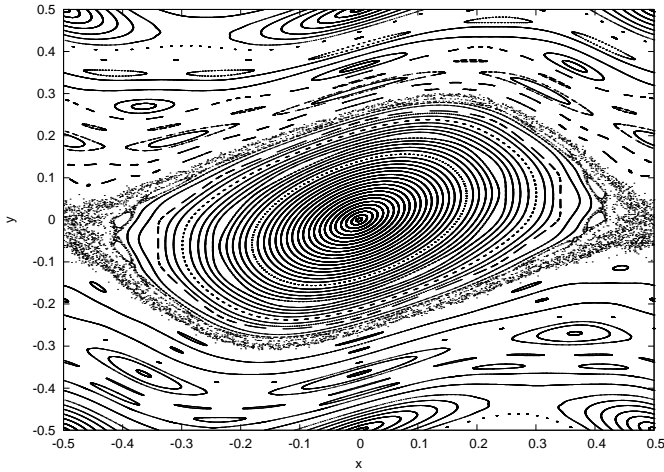


Figura 1.3: Órbitas da aplicação standard para  $\alpha = 0.8$ .

gião de ressonância do centro da figura, apresenta uma ilha elíptica circundada por uma coroa espessa de iteradas da aplicação standard. De fato, tais iteradas podem corresponder a uma única órbita. A dinâmica nesta coroa, que contém o ponto fixo hiperbólico, próximo a  $(x, y) = (0.5, 0)$ , é bastante complexa. Assim como ocorre em uma vizinhança do ponto fixo hiperbólico, pontos inicialmente próximos nesta coroa “tipicamente” (de acordo com observações numéricas) se afastam exponencialmente rápido. Por tal razão esta parte da região de ressonância é chamada de “zona estocástica”. Note que na figura 1.2 a zona estocástica da região de ressonância central é inobservável



(apesar de existir). Nesta figura, a ilha elíptica praticamente coincide com a região de ressonância. Já na figura 1.3 tanto a ilha elíptica quanto a zona estocástica dividem o espaço da região de ressonância.

Ao se ultrapassar o valor crítico de  $\alpha = \bar{\alpha} = 0.9716\dots$ , obtido numericamente [34], observa-se que os círculos rotacionais invariantes deixam de existir. De fato, como veremos no capítulo 4, após o rompimento do último círculo invariante aparecem órbitas que sobem arbitrariamente alto no cilindro. Para  $\alpha < \bar{\alpha}$  a existência de tais órbitas era impedida pela presença dos círculos rotacionais invariantes. A figura 1.4 ilustra como são as típicas órbitas da aplicação standard para  $\alpha = 0.971$ . Note que com o desaparecimento das curvas

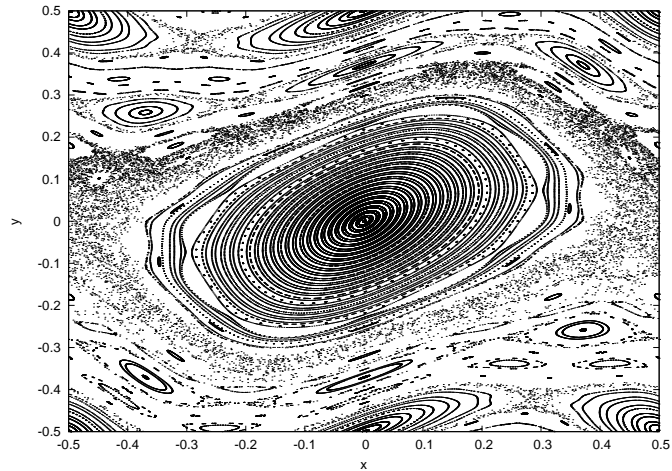


Figura 1.4: Órbitas da aplicação standard para  $\alpha = 0.971$ , próximo à quebra do último círculo rotacional invariante.

rotacionais invariantes ocorre uma fusão das regiões de ressonância. O elemento separador desapareceu. Na figura 1.4 o espaço ocupado pelas zonas estocásticas é maior do que na figura 1.3, mas ainda se observa aquela robusta ilha elíptica da zona de ressonância do centro da figura.

Conforme aumentamos  $\alpha$  a dinâmica da aplicação standard torna-se mais e mais complexa. O que se nota numericamente é que as ilhas elípticas perdem espaço para as zonas de ressonância, ao ponto de se tornarem praticamente inobserváveis. Por exemplo, na figura 1.5 ( $\alpha = 10$ ) já não se pode observar ilhas elípticas. Nesta figura estão

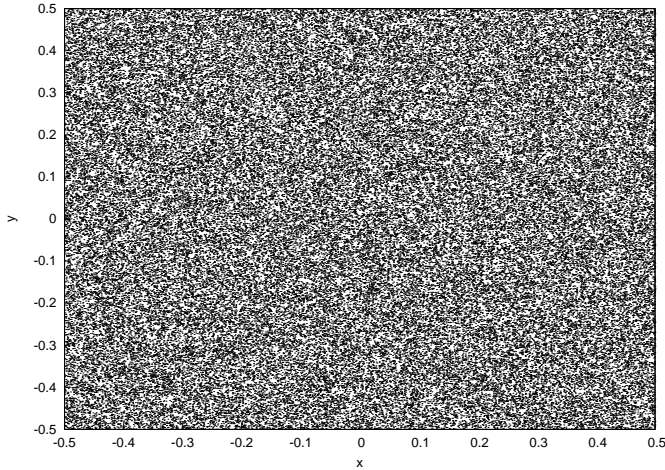


Figura 1.5: Uma órbita da aplicação standard para  $\alpha = 10$  e condição inicial  $(x_0, y_0) = (0, 0)$ , “caso aparentemente estocástico”.

mostrados os  $10^5$  primeiros pontos de uma única órbita com início em  $(x, y) = (0, 0)$ !

Resumidamente, podemos dizer que  $\alpha$  parametriza, na dinâmica da aplicação standard, uma disputa entre as curvas invariantes e as regiões estocásticas. Em  $\alpha = 0$  temos o triunfo absoluto das curvas rotacionais invariantes. Conforme  $\alpha$  cresce as curvas rotacionais invariantes vão sendo substituídas por outras curvas invariantes, as das ilhas elípticas, que são entremeadas por zonas estocásticas cada vez mais expressivas. Para  $\alpha > \bar{\alpha}$  as curvas rotacionais invariantes deixam de existir, e para  $\alpha$ 's ainda maiores a linhagem das curvas invariantes, incluindo aquelas das ilhas elípticas, parece totalmente

extinta. Até que ponto tal cenário obtido numericamente é verdadeiro? A resposta a essa pergunta ainda não foi dada. A primeira idéia, que brota da análise acima, de que existe um certo valor de  $\alpha$  a partir do qual não existem ilhas elípticas, como ocorre com os círculos rotacionais invariantes, é falsa. Foi provado por Duarte [28] que, existe  $\alpha_0$  tal que, para  $\alpha > \alpha_0$ , o conjunto de parâmetros  $\alpha$  para os quais a aplicação standard possui pontos periódicos elípticos é denso em  $[\alpha_0, \infty)$ ! Em que sentido, então, pode-se afirmar que a zona estocástica é dominante para  $\alpha$ 's grandes?

No cerne desta questão está a definição matemática do que se chama zona estocástica. A definição mais aceita atualmente é que a zona estocástica é o conjunto de pontos que possuem “expoente de Liapunov” positivo. O expoente de Liapunov  $\lambda(z)$  de um ponto  $z = (x, y)$  é igual ao seguinte limite, se ele existir,

$$\lim_{n \rightarrow \infty} \frac{\ln \|DF(z_n) \dots DF(z_1)DF(z_0)\|}{n}$$

onde:  $DF(z)$  é a derivada da aplicação em  $z$ ,  $z_n$  é a  $n$ -ésima iterada de  $z_0 = z$ , e  $\|DF(z_n) \dots DF(z_1)DF(z_0)\|$  é o módulo do maior autovalor do produto de matrizes  $DF(z_n) \dots DF(z_1)DF(z_0)$ . O expoente de Liapunov  $\lambda(z)$  é positivo se a dinâmica linearizada da aplicação próxima à órbita de  $z$  é exponencialmente instável. Note que se  $\lambda(z) = 0$  não podemos dizer que esta dinâmica linearizada é estável. Um ponto fixo hiperbólico tem expoente de Liapunov positivo e um ponto fixo elíptico tem expoente de Liapunov zero (mostre isto). Se a definição acima de zona estocástica é boa, deve se esperar que a maioria dos pontos (no sentido de “área”, ou melhor dizendo, da medida de Lebesgue) tenha expoente de Liapunov positivo para  $\alpha$  grande. Apesar de ter sido provado que, para todo  $\alpha > 0$ , a aplicação standard possui um conjunto não enumerável de pontos com expoente de Liapunov positivo, ainda é um problema em aberto provar que tal conjunto possui medida de Lebesgue positiva para algum valor de  $\alpha$ ! De fato, pode-se dizer que este é um dos (talvez o) problemas em aberto mais famosos da dinâmica conservativa.

## 1.2 O modelo de Frenkel-Kontorova

Como dissemos, aplicações do tipo twist aparecem em diversos contextos, entre eles na física do estado sólido. Em 1938, Y. Frenkel e T. Kontorova, no intuito de entender certos fenômenos associados a redes cristalinas, introduziram o seguinte modelo. Considere uma cadeia unidimensional de partículas, cada uma das quais ligada a suas vizinhas por uma mola linear com constante elástica igual a um e comprimento de repouso  $\ell$ . Se  $x_k$  é a posição da  $k$ -ésima partícula da cadeia, então a energia de interação entre partículas vizinhas é:

$$\frac{1}{2}(x_{k+1} - x_k - \ell)^2 + \frac{1}{2}(x_k - x_{k-1} - \ell)^2$$

Suponha agora que a cadeia encontra-se imersa em um cristal unidimensional e que cada partícula da cadeia interage com cada átomo do cristal (figura 1.6). Isto implica que a partícula  $k$  da cadeia está su-

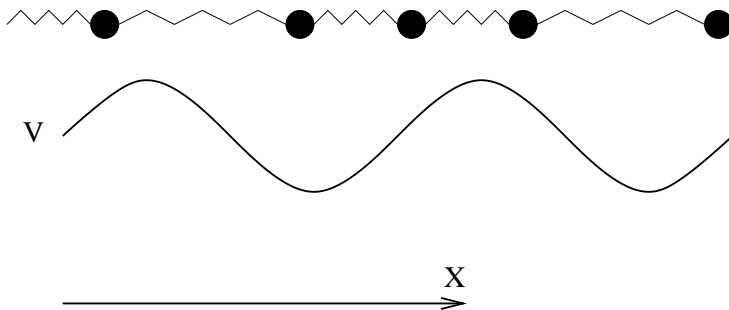


Figura 1.6: Cadeia de partículas do modelo de Frenkel-Kontorova.

jeita a um potencial periódico, devido a periodicidade do cristal, e se desprezarmos todos os harmônicos da série de Fourier deste potencial, exceto o primeiro, podemos escrevê-lo como:

$$V(x_k) = \frac{\alpha}{4\pi^2} \cos(2\pi x_k)$$

Sendo assim a energia da cadeia como um todo é dada por:

$$\begin{aligned}
 W &= \sum_k \left[ \frac{1}{2} (x_{k+1} - x_k - \ell)^2 + \frac{\alpha}{4\pi^2} \cos(2\pi x_k) \right] \\
 &\stackrel{\text{def}}{=} \sum_k h(x_k, x_{k+1})
 \end{aligned} \tag{1.3}$$

Cada sequência  $\dots, x_{-1}, x_0, x_1, \dots$ , que descreve as posições das infinitas partículas da cadeia, é chamada de configuração do sistema. Um problema interessante no estudo da física deste sistema é o de encontrar as configurações  $\dots, x_{-1}, x_0, x_1, \dots$  que “minimizam”  $W$ . Claramente tal questão precisa ser melhor definida, uma vez que para quase todas as configurações a série (1.3) que define  $W$  deve ter valor infinito. Isto motiva as seguintes definições:

**Definição 1 (Segmento Minimal).** *Um segmento  $\{x_a, \dots, x_b\}$  finito de uma configuração é dito minimal, se para qualquer outro segmento  $\{x_a, \tilde{x}_{a+1}, \dots, \tilde{x}_{b-1}, x_b\}$ , com os mesmos extremos, vale a desigualdade:*

$$W\{x_a, \dots, x_b\} \stackrel{\text{def}}{=} \sum_{k=a}^{b-1} h(x_k, x_{k+1}) \leq W\{x_a, \tilde{x}_{a+1}, \dots, \tilde{x}_{b-1}, x_b\}$$

**Definição 2 (Configuração Minimal).** *Uma configuração é dita minimal, se qualquer um de seus segmentos finitos é minimal.*

Note que, se uma configuração é minimal então cada um de seus segmentos  $\{x_a, \dots, x_b\}$  minimiza a função:

$$(\tilde{x}_{a+1}, \dots, \tilde{x}_{b-1}) \longrightarrow h(x_a, \tilde{x}_{a+1}) + \dots + h(\tilde{x}_{b-1}, x_b) = W\{x_a, \dots, x_b\}$$

Como pontos de mínimo de uma função diferenciável no  $\mathbb{R}^n$  são necessariamente pontos críticos da mesma, então  $\{x_a, x_{a+1}, \dots, x_{b-1}, x_b\}$ ,

deve satisfazer o seguinte conjunto de equações:

$$\begin{aligned} \frac{\partial W}{\partial x_{a+1}} \{x_a, \dots, x_b\} = 0 &\implies \\ \frac{\partial h}{\partial x_{a+1}}(x_a, x_{a+1}) + \frac{\partial h}{\partial x_{a+1}}(x_{a+1}, x_{a+2}) &= 0, \\ \frac{\partial W}{\partial x_{a+2}} \{x_a, \dots, x_b\} = 0 &\implies \\ \frac{\partial h}{\partial x_{a+2}}(x_{a+1}, x_{a+2}) + \frac{\partial h}{\partial x_{a+2}}(x_{a+2}, y_{a+3}) &= 0, \dots \end{aligned}$$

Ou seja, de uma maneira mais sucinta, os  $x_k$  de uma configuração minimal necessariamente satisfazem as equações:

$$\begin{aligned} \partial_2 h(x_{k-1}, x_k) + \partial_1 h(x_k, x_{k+1}) &= -x_{k+1} + 2x_k - x_{k-1} - \frac{\alpha}{2\pi} \sin(2\pi x) \\ &= 0 \quad k = 1, 2, 3, \dots \end{aligned} \quad (1.4)$$

onde  $\partial_1 h$  significa a derivada parcial de  $h$  com relação ao seu primeiro argumento e  $\partial_2 h$  o equivalente para o segundo argumento. Note que o parâmetro  $\ell$  não aparece na equação (1.4). Como estaremos primariamente interessados em configurações que satisfazem (1.4) suporemos a partir de agora que  $\ell = 0$ .

**Definição 3 (Configuração estacionária).** *Dizemos que uma configuração  $\dots, x_{-1}, x_0, x_1, \dots$  é estacionária, no sentido de seus segmentos serem pontos críticos de  $W$ , se seus elementos satisfazem a equação (1.4).*

Note que configurações minimais são necessariamente estacionárias, mas configurações estacionárias podem não ser configurações minimais. É interessante que dado qualquer par  $(x_0, x_1)$ , é possível gerar uma configuração estacionária que contenha  $(x_0, x_1)$ . Para ver isto é conveniente reescrever (1.4) de uma outra maneira. Primeiramente definimos, a partir de uma configuração qualquer  $\dots, x_{-1}, x_0, x_1, \dots$ , a sequência  $\dots, y_{-1}, y_0, y_1, \dots$  através de:

$$y_k = -\partial_1 h(x_k, x_{k+1}) = +x_{k+1} - x_k + \frac{\alpha}{2\pi} \sin(2\pi x_k)$$

onde usamos a definição de  $h$  dada em (1.3). Da equação acima segue que

$$x_{k+1} = y_k + x_k - \frac{\alpha}{2\pi} \text{sen}(2\pi x_k).$$

Agora, se  $\dots, x_{-1}, x_0, x_1, \dots$  é estacionária, então vale (1.4) e da definição de  $y_{k+1}$  obtemos:

$$y_{k+1} = -\partial_1 h(x_{k+1}, x_{k+2}) = \partial_2 h(x_k, x_{k+1}) = x_{k+1} - x_k$$

Em suma, para uma configuração estacionária, vale:

$$\begin{aligned} x_{k+1} &= x_k + y_{k+1} \\ y_{k+1} &= y_k - \frac{\alpha}{2\pi} \text{sen}(2\pi x_k), \end{aligned}$$

que é exatamente a recursão que define a dinâmica da aplicação standard, equação (1.2). Quer dizer, as configurações estacionárias do modelo de Frenkel-Kontorova são exatamente as projeções das órbitas da aplicação standard no eixo  $x$ ! Tal caracterização das órbitas da aplicação standard é chamada de “princípio variacional” da aplicação standard.

Do que foi dito na seção 1.1 pode-se imediatamente concluir que para certos valores de  $\alpha$ , como por exemplo  $\alpha = 10$  (ver figura 1.5), as configurações estacionárias do modelo de Frenkel-Kontorova podem apresentar um padrão bastante irregular. Tal correspondência entre as configurações estacionárias do modelo de Frenkel-Kontorova e as órbitas da aplicação standard sugerem a seguinte definição.

**Definição 4 (Órbita minimizante).** *Dizemos que uma órbita é minimizante se a sua correspondente configuração é minimal.*

Veremos a seguir que tais órbitas minimizantes possuem diversas propriedades notáveis. Na física do modelo de Frenkel-Kontorova elas são fundamentais, pois em certo sentido minimizam a energia. No entendimento da dinâmica da aplicação standard elas não são menos importantes.

### 1.3 Aplicações de tipo “twist” e princípios variacionais

Antes de prosseguir com a análise das órbitas minimizantes, é importante mostrar que os resultados da seção 1.2 não estão restritos à aplicação standard e ao modelo de Frenkel-Kontorova. De fato todas aplicações de tipo twist possuem princípios variacionais. Mais precisamente, seja  $C = \mathbb{R}/\mathbb{Z} \times \mathbb{R}$  o cilindro recoberto pelo plano  $\{(x, y) \in \mathbb{R}^2\}$  através de  $(x, y) \rightarrow (x \pmod{1} = \theta, y)$ . Considere agora um difeomorfismo  $f$  de  $C$ , cujo levantamento ao plano, denotado por  $F$ , possui as seguintes componentes

$$\begin{aligned} x' &= F_1(x, y) \\ y' &= F_2(x, y) \end{aligned} \tag{1.5}$$

Suponha que  $F$  e  $f$  possuam as seguintes propriedades:

- a)  $F_1(x + 1, y) = 1 + F_1(x, y)$  e  $f$  preserva os fins do cilindro (portanto  $f$  é homotópico a identidade e portanto preserva a orientação);
- b)  $(\partial_1 F_1)(\partial_2 F_2) - (\partial_2 F_1)(\partial_1 F_2) = 1$ , ou seja,  $f$  preserva área;
- c) existe uma constante  $c > 0$  tal que

$$0 < c \leq \partial_2 F_1(x, y) \leq \frac{1}{c}$$

ou seja,  $F$  satisfaz uma condição de twist uniforme.

Note que a aplicação standard é um típico exemplo de aplicação que satisfaz as condições acima. Para entender geometricamente a condição de twist consideremos uma reta vertical  $\gamma$  no plano  $\{x, y\}$ , dada por  $s \rightarrow (x, y) = (\bar{x}, s)$ , onde  $\bar{x}$  está fixo e  $s \in \mathbb{R}$ , ver figura 1.7. A imagem de tal reta pela aplicação, denotada por  $\gamma'$ , é  $s \rightarrow [x', y'] = [F_1(\bar{x}, s), F_2(\bar{x}, s)]$ . A condição  $\partial_2 F_1(x, y) > c > 0$  implica que o vetor tangente a  $\gamma'$  tem sempre uma componente  $x$  positiva maior que  $c$ , ou seja, a curva  $\gamma'$  projeta-se difeomorficamente sobre o eixo real  $x$ , ver figura 1.7. Note que, somente com a hipótese  $\partial_2 F_1 > 0$  não é possível garantir que tal projeção é sobrejetiva.



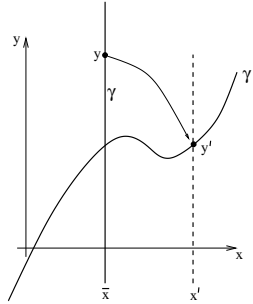


Figura 1.7: Método geométrico para resolver a equação  $x' = F_1(x, y)$  para  $y$ .

Da hipótese de twist uniforme segue que a equação  $x' = F_1(x, y)$  sempre pode ser resolvida para  $y$ , como mostrado na figura 1.7. Seja  $y = u(x, x')$  tal solução. A propriedade  $F_1(x + 1, y) = 1 + F_1(x, y)$ , implica que  $y = u(x + 1, x' + 1) = u(x, x')$ , e isso implica que a função  $u$ , a princípio definida em  $\mathbb{R}^2$ , também define uma função no cilindro  $\overline{C}$  que é obtido do plano  $\{x, x'\}$  através de identificação de pontos  $(x, x')$  e  $(\tilde{x}, \tilde{x}')$  que satisfaçam  $x = \tilde{x} + k$  e  $x' = \tilde{x}' + k$ , para  $k \in \mathbb{Z}$  (ver figura 1.8).

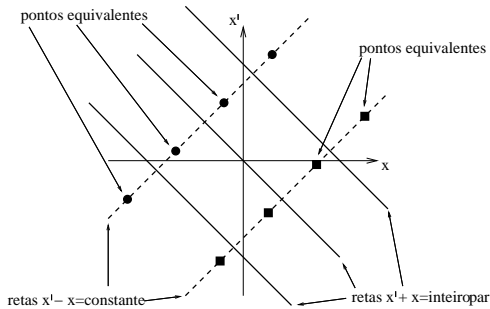


Figura 1.8: Periodicidade de  $h(x, x')$ .

Considere agora o campo de vetores no plano  $\{x, x'\}$  dado por

$$(x, x') \longrightarrow (-y, y') = (-u(x, x'), F_2(x, u(x, x'))) \stackrel{\text{def}}{=} (-u, v)$$

ou sua forma diferencial associada  $-udx + vdx'$ . A preservação de área,  $(\partial_1 F_1)(\partial_2 F_2) - (\partial_2 F_1)(\partial_1 F_2) = 1$ , implica que tal campo possui rotacional nulo, ou seja,  $\partial_{x'}u + \partial_x v = 0$  ou, equivalentemente, que a forma diferencial  $-udx + vdx'$  é fechada. Para verificar isso usamos que  $x' = F_1(x, u(x, x'))$  implica

$$\partial_{x'}u(x, x') = \frac{1}{\partial_2 F_1(x, u(x, x'))} \quad \text{e} \quad \partial_x u(x, x') = -\frac{\partial_1 F_1(x, u(x, x'))}{\partial_2 F_1(x, u(x, x'))}$$

e então:

$$\begin{aligned} \partial_{x'}u + \partial_x v &= \partial_{x'}u + \partial_1 F_2 + (\partial_2 F_2)\partial_x u \\ &= \frac{1}{\partial_2 F_1} + \partial_1 F_2 - (\partial_2 F_2)\frac{\partial_1 F_1}{\partial_2 F_1} \\ &= \frac{1 + (\partial_1 F_1)(\partial_2 F_2) - (\partial_2 F_1)(\partial_1 F_2)}{\partial_2 F_1} = 0. \end{aligned}$$

Sabemos do cálculo diferencial que um campo com rotacional nulo no plano (ou uma forma diferencial fechada) é o gradiente (ou o diferencial) de uma função no plano. Portanto existe uma função  $h(x, x')$  tal que:

$$\begin{aligned} -y &= -u(x, x') = \partial_1 h(x, x') \\ y' &= v(x, x') = F_2(x, u(x, x')) = \partial_2 h(x, x') \end{aligned} \quad (1.6)$$

Note que tal função  $h$ , a princípio definida no plano  $\{x, x'\}$ , não necessariamente satisfaz a condição  $h(x+1, x'+1) = h(x, x')$ , satisfeita pela função  $u$ , e portanto  $h$  não necessariamente define uma função no cilindro  $\overline{C}$ . De fato, para que valha tal propriedade é necessário e suficiente que a integral de linha do campo  $(-u, v)$  sobre qualquer curva fechada  $\gamma$ , que dá a volta no cilindro  $\overline{C}$  (ou qualquer curva no plano  $\{x, x'\}$  que comece em  $(x, x')$  e termine em  $(x+1, x'+1)$ ), seja igual a zero (isso é uma consequência do teorema de Stokes). Seja portanto  $\gamma$  a curva  $s \rightarrow (x = s, x' = F_1(s, 0))$ , com  $s \in [0, 1]$ . Note que  $\gamma$  começa em  $(0, F_1(0, 0))$  e termina em  $(1, F_1(1, 0) = F_1(0, 0) + 1)$ ,

devido à propriedade a). A definição de  $u$  implica que  $u(s, F_1(s, 0)) = 0$ , ou seja, sobre  $\gamma$  vale  $u = 0$ . Portanto a integral de linha do campo  $(-u, v)$  sobre  $\gamma$  reduz-se a

$$\int_0^1 v(s, F_1(s, 0)) ds = \int_0^1 F_2(s, u(F_1(s, 0))) ds = \int_0^1 F_2(s, 0) ds. \quad (1.7)$$

A interpretação geométrica desta integral é imediata. Se  $\beta$  é a curva  $y = 0$  que corresponde ao equador do cilindro  $C$  e  $\beta'$  é sua imagem pela aplicação  $F$ , então a integral acima representa a diferença entre a área abaixo de  $\beta'$  e acima de  $\beta$  (área hachurada na figura 1.9) e a área acima de  $\beta'$  e abaixo de  $\beta$  (área pontilhada na figura 1.9).

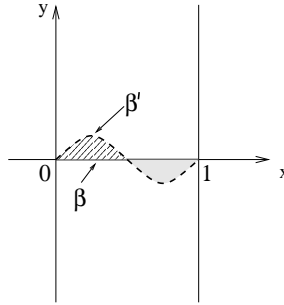


Figura 1.9: Condição para  $F$  ser exata.

A seguir, suporemos que  $f$ , ou seu levantamento  $F$ , além de possuir as propriedades a), b), c) acima, também satisfaz:

- d)  $F$  é tal que a integral (1.7) é nula. Uma  $F$  (ou sua projeção  $f$  no cilindro) que possui tal propriedade é dita exata.

Portanto a hipótese de  $F$  ser exata implica que

$$h(x + 1, x' + 1) = h(x, x')$$

ou seja,  $h$  também define uma função no cilindro  $\overline{C}$ . O que foi dito acima está sumarizado no seguinte lema.

**Lema 1 (Função Geratriz).** *Seja  $F$ , dada por (1.5), o levantamento de um difeomorfismo  $f$  do cilindro  $C$ , que satisfaz as propriedades a), b), c), d). Então existe uma função duas vezes continuamente diferenciável  $h(x, x')$ , chamada função geratriz de  $F$ , tal que:*

$$h(x + 1, x' + 1) = h(x, x'), \quad (1.8)$$

$$-\frac{1}{c} \leq \partial_1 h(x, x') \leq -c < 0, \quad (1.9)$$

(esta última desigualdade é consequência da propriedade (c) de twist uniforme) e

$$\begin{aligned} y &= -\partial_1 h(x, x') \\ y' &= \partial_2 h(x, x') \end{aligned} \quad (1.10)$$

onde  $x, x', y, y'$  satisfazem a equação (1.5).

Para dar um significado mais geométrico à função  $h$ , é conveniente calcular a integral de caminho que a define usando as variáveis  $(x, y)$ . Considere uma curva contínua  $\hat{\gamma} : s \rightarrow (x, x')$ , com  $s \in [0, 1]$ ,  $\hat{\gamma}(0) = (x_0, x'_0)$  e  $\hat{\gamma}(1) = (x_1, x'_1)$ . Suponha que, exceto por um número finito de pontos,  $\hat{\gamma}$  é continuamente diferenciável. Temos então que a função  $h$  satisfaz:

$$h(x_1, x'_1) - h(x_0, x'_0) = \int_{\hat{\gamma}} -udx + vdx'$$

Para calcular esta integral voltemos às variáveis  $(x, y) = (x, u(x, x'))$ . Denotemos por  $\gamma$  a representação de  $\hat{\gamma}$  nas novas variáveis e por  $\gamma(0) = (x_0, y_0)$ ,  $\gamma(1) = (x_1, y_1)$ , o valor de  $\gamma$  nos seus extremos. Nestas novas variáveis a integral de linha

$$\int_{\gamma} udx = \int_{\gamma} ydx$$

coincide com a “área embaixo da curva”  $\gamma$  como aprendemos no cálculo. Note que esta integral de linha faz sentido mesmo quando  $\gamma$  é vertical, caso em que a integral vale zero. A fim de dar uma interpretação geométrica para a integral

$$\int_{\hat{\gamma}} vdx' = \int_{\gamma} F_2(\partial_x F_1 dx + \partial_y F_1 dy) = \int_{\gamma} F_2 dF_1$$

é conveniente fazer uma nova mudança de variáveis:  $(x, y) \rightarrow (x', y') = F(x, y)$ . Nas novas variáveis temos  $F_2 dF_1 = y' dx'$  e  $\gamma(s) \rightarrow F \circ \gamma(s)$ . Portanto

$$\int_{\hat{\gamma}} v dx' = \int_{F \circ \gamma} y' dx'$$

e agora tal integral pode ser interpretada como a área embaixo da curva  $F \circ \gamma$ . Com isso podemos dar a seguinte caracterização geométrica para  $h$ . Seja  $\gamma$  uma curva qualquer no plano  $(x, y)$  com as propriedades de continuidade e diferenciabilidade acima, e tal que

$$\gamma(0) = (x_0, y_0), \quad \gamma(1) = (x_1, y_1).$$

Seja  $x'_0 = F_1(x_0, y_0)$  e  $x'_1 = F_1(x_1, y_1)$ . Então

$$h(x_1, x'_1) - h(x_0, x'_0) = \int_{F \circ \gamma} y' dx' - \int_{\gamma} y dx = A' - A \quad (1.11)$$

onde  $A$  e  $A'$  correspondem às áreas hachuradas na figura 1.10.

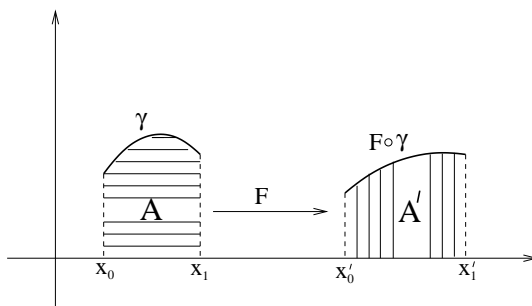


Figura 1.10: Interpretação geométrica de  $h$ , a saber:  $h(x_1, x'_1) - h(x_0, x'_0) = A' - A$ .

Note que para determinar a função  $h$  a partir da integral (1.11) é necessário impor o seu valor em algum ponto do plano  $\{x, x'\}$ . De fato, é sempre possível fazer uma mudança de coordenadas na aplicação  $F$ , sem alterar as propriedades a, b, c, d, de tal modo que  $h$  possua uma determinação geométrica razoavelmente simples. Faremos isso a seguir.

Considere a função  $x \rightarrow y = w(x) \stackrel{\text{def}}{=} u(x, x)$ . Tal função associa a cada  $x$  um  $y = w(x)$  tal que a imagem do ponto  $(x, w(x))$  por  $F$  permanece na mesma vertical, ou seja,  $F_1(x, w(x)) = x$ . Considere agora a mudança de variáveis que preserva área:

$$\begin{aligned}\tilde{x} &= x \\ \tilde{y} &= y - w(x)\end{aligned}$$

A aplicação  $F$  nas novas coordenadas satisfaz as mesmas propriedades a, b, c, d, que satisfazia nas antigas. Denotando as novas variáveis também por  $(x, y)$ , para não carregar a notação, temos que  $F$  nas, novas variáveis, passa a possuir seguinte propriedade adicional:

$$e) \quad x = F_1(x, 0)$$

Geometricamente isto significa que todo ponto da curva  $y = 0$  do plano  $(x, y)$  move-se verticalmente pela aplicação  $F$  (ver figura 1.11). Mais ainda, os pontos da curva  $y = 0$  que intersectam sua imagem

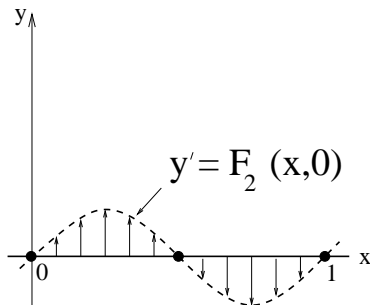


Figura 1.11: Curva de velocidade angular zero.

por  $F$  são pontos fixos, uma vez que nenhuma de suas coordenadas varia pela ação de  $F$ . Ou seja, esta simples construção mostra que  $f$ , a projeção de  $F$  no cilindro, tem pelo menos dois pontos fixos, como ilustrado na figura 1.11 (lembre-se que no cilindro os pontos fixos correspondentes a  $x = 0$  e  $x = 1$  são idênticos). Note que, com uma mudança de variáveis dada por uma simples translação de  $x$ , é possível colocar um ponto fixo qualquer de  $F$  em  $x = 0$ , como na

figura 1.11. Suponha, portanto, que  $(x, y) = (0, 0)$  é um ponto fixo de  $F$  e que  $h(0, 0) = 0$ . Dado  $x$ , seja  $\gamma_0$  um segmento horizontal no plano  $(x, y)$ , com  $y = 0$ , que liga a origem a  $x$ . Aplicando a identidade (1.11) a  $\gamma_0$  obtém-se que

$$h(x, x) = h(x, x) - h(0, 0) = \int_{F \circ \gamma_0} y' dx'$$

que é a integral ilustrada na figura 1.12. Note que os pontos críticos

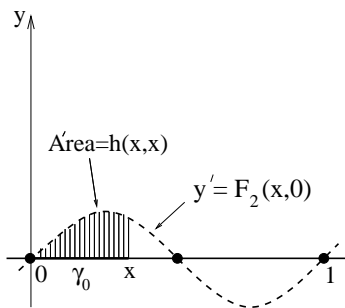


Figura 1.12: Interpretação geométrica de  $h(x, x)$ .

de  $h(x, x)$  correspondem aos pontos fixos de  $F$ , uma vez que pelo teorema fundamental do cálculo a função  $y' = F_2(x, 0)$  é a derivada de  $h(x, x)$ , e as soluções de  $F_2(x, 0) = 0$  correspondem aos pontos fixos de  $F$  (isso também pode ser deduzido das equações (1.10)). No caso ilustrado na figura 1.12, o ponto fixo em  $x = 0$  é aquele que corresponde ao mínimo da função  $h(x, x)$ . É sempre possível, e isto será suposto, que o ponto fixo colocado em  $x = 0$  através da translação de  $x$ , é aquele que corresponde a um mínimo da função  $h(x, x)$ . Com isso, assegura-se que  $h(x, x) \geq 0$  para qualquer  $x$ . Seja agora  $(x, x')$  um ponto qualquer, e considere o segmento vertical  $\gamma_1$  no plano  $(x, y)$  com extremos  $(x, 0)$  e  $(x, y)$ , onde  $x' = f(x, y)$ , ver figura 1.13. Seja  $\gamma$  a curva no plano  $(x, y)$  obtida da concatenação de  $\gamma_0$  e  $\gamma_1$ . Aplicando a identidade (1.11) a  $\gamma$  obtém-se

$$h(x, x') = h(x, x') - h(0, 0) = \int_{F \circ \gamma} y' dx'$$

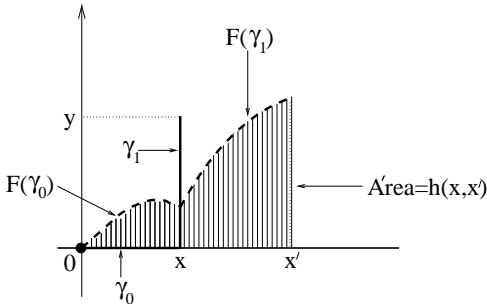


Figura 1.13: Interpretação geométrica de  $h(x, x')$ .

que é a integral ilustrada na figura 1.13. É fácil ver, a partir desta interpretação geométrica e da propriedade de twist, que  $h(x, x') \geq 0$  para qualquer par  $(x, x')$  (note que um segmento vertical com  $y$  negativo e orientado para baixo, sempre entorta para a esquerda, o que garante a positividade da integral acima quando  $x' < x$ ). Para resumir o que foi dito acima enunciamos a seguinte proposição.

**Proposição 1.** *Existem coordenadas  $(x, y)$  onde a aplicação  $F$ , além de possuir as propriedades a, b, c, d, satisfaz  $x = F_1(x, 0)$  e  $0 = F_2(0, 0)$ , o que implica que  $(x, y) = (0, 0)$  é um ponto fixo de  $F$ . Além disso  $F$  está associada a uma função  $h$  com as seguintes propriedades:  $h(0, 0) = 0$  e  $h(x, x') \geq 0$ .*

Há duas propriedades da função  $h$  que serão muito importantes na próxima seção. Elas são dadas pelas seguintes proposições.

**Proposição 2.** *Se  $F$  satisfaz as propriedades a, b, c, d então para qualquer  $h$  associada a  $F$  vale a seguinte desigualdade*

$$h(x_b, x'_a) + h(x_a, x'_b) - h(x_b, x'_b) - h(x_a, x'_a) \geq c(x_b - x_a)(x'_b - x'_a)$$

onde  $c > 0$  é a constante de twist dada em c).

*Demonstração.* Para provar a proposição basta integrar a função  $-\partial_{12}h$  em um quadrado e usar que a desigualdade (1.9) implica:

$$\int_{x'_a}^{x'_b} \int_{x_a}^{x_b} -\partial_{12}h dx dx' \geq c \int_{x'_a}^{x'_b} \int_{x_a}^{x_b} dx dx'$$



□

**Proposição 3.** *Seja  $F$  e  $h$  como na proposição (1). Então vale a seguinte desigualdade*

$$h(x, x') \geq \frac{c}{2}(x' - x)^2$$

onde  $c > 0$  é a constante de twist dada em  $c$ ).

*Demonstração.* Para provar a proposição basta integrar a função  $-\partial_{12}h$  em um triângulo, usar que a desigualdade (1.9) implica:

$$\begin{aligned} h(x, x') &= h(x, x) + \int_x^{x'} \int_s^{x'} -\partial_{12}h(s, t) dt ds \\ &\geq h(x, x) + c \int_x^{x'} \int_s^{x'} dt ds \int_{x'_a}^{x'_b} \int_{x_a}^{x_b} dx dx' \end{aligned}$$

e finalmente usar que, da proposição 1, vale  $h(x, x) \geq 0$ . □

Vimos na seção 1.2 que a aplicação standard está associada ao modelo de Frenkel-Kontorova, que por sua vez é caracterizado por sua função energia  $W$ . De maneira análoga, a aplicação  $F$  está associada a uma função  $h$  a partir da qual definimos a função “energia”  $W$  ou, como é mais comumente chamada, a função “ação”  $W$ , como

$$W = \sum_k h(x_k, x_{k+1})$$

Da mesma forma que feito na seção 1.2, definimos para o par  $(F, W)$ : segmento minimal, configuração minimal e configuração estacionária. As relações (1.6) implicam que novamente vale: se  $(x_k, y_k)$ ,  $k \in \mathbb{Z}$ , é uma órbita de  $F$ , então  $x_k$ ,  $k \in \mathbb{Z}$ , é uma configuração estacionária de  $W$ , ou seja, valem as relações:

$$\partial_2 h(x_{k-1}, x_k) + \partial_1 h(x_k, x_{k+1}) = 0 \quad k \in \mathbb{Z}. \quad (1.12)$$

Portanto, neste caso vale a mesma definição de órbita minimizante que demos para a aplicação standard. Antes de usarmos o princípio variacional aqui obtido, para estudar o problema de existência de órbitas minimizantes para a aplicação  $F$ , veremos uma série de exemplos e circunstâncias onde aparecem aplicações do tipo twist.

## 1.4 Bilhares

Considere uma região  $U$  no plano  $\mathbb{R}^2$  limitada por uma curva  $\gamma$  simples (sem auto-interseção), regular, de classe  $C^k$ , fechada.

Associada a esta região temos uma aplicação bi-dimensional definida do seguinte modo:

Para cada ponto  $p \in \gamma$  tomemos a reta  $r(p, \phi)$  que passa por  $p$  e faz ângulo  $\phi$  com o vetor tangente a  $\gamma$  em  $p$ . Ao par  $(p, \phi)$  associamos o par  $(p_1, \phi_1)$  formado pelo ponto  $p_1 \in \gamma$  de interseção da reta  $r(p, \phi)$  com  $\gamma$  e o ângulo  $\phi_1$  entre  $r(p, \phi)$  e o vetor tangente a  $\gamma$  em  $p_1$ , de acordo com a regra de reflexão em um espelho: o ângulo de incidência é igual ao ângulo de reflexão.

Mais precisamente, se  $\gamma(s)$  é uma parametrização por comprimento de arco de  $\gamma$ , orientada positivamente, uma aplicação do tipo bilhar associado à  $\gamma$  é uma aplicação  $T(s, \phi) = (s_1, \phi_1)$  tal que se  $\phi$  é o ângulo entre  $\gamma'(s)$  e  $\gamma(s_1) - \gamma(s)$ , então  $\phi_1$  é o ângulo entre  $\gamma(s_1) - \gamma(s)$  e  $\gamma'(s_1)$ , nesta ordem.

Observe que  $0 \leq \phi \leq \pi$  e que  $\phi = 0$  e  $\phi = \pi$  são invariantes. Podemos, portanto, considerar a aplicação definida no cilindro aberto  $\mathcal{A} = [0, C] \times (0, \pi)$ , onde  $C$  é o comprimento de  $\gamma$  ( $0$  e  $C$  estão identificados).

Esta aplicação possui uma função geratriz:

$h(t, s) = \|\gamma(t) - \gamma(s)\|$  onde  $\|\cdot\|$  denota a distância euclidiana isto é,

$$h(t, s)^2 = \langle \gamma(s) - \gamma(t), \gamma(s) - \gamma(t) \rangle$$

Cabe-nos ressaltar que, usualmente a função geratriz é definida como sendo igual a  $-h$ , mas isto não afeta em nada a análise futura.

Usando a regra da cadeia,

$$2h\partial_1 h = -2 \langle \gamma'(t), \gamma(s) - \gamma(t) \rangle$$

e

$$2h\partial_2 h = 2 \langle \gamma'(s), \gamma(s) - \gamma(t) \rangle$$

Logo,

$$\partial_1 h = -\frac{1}{h} \langle \gamma'(t), \gamma(s) - \gamma(t) \rangle = -\cos(\phi)$$

e

$$\partial_2 h = \frac{1}{h} \langle \gamma'(s), \gamma(s) - \gamma(t) \rangle = \cos(\phi_1)$$

Portanto,  $T(t, \phi) = (s, \phi_1)$  se e somente se  $\cos(\phi) = -\partial_1 h$  e  $\cos(\phi_1) = \partial_2 h$ .

Observe que se  $\gamma$  for estritamente convexa (curvatura positiva,) então fixado  $t$ , a aplicação  $\phi \rightarrow s(t, \phi)$  é estritamente crescente, o que equivale à condição de twist:  $\frac{\partial s}{\partial \phi} > 0$

Provemos que  $\partial_{12} h > 0$  :

$$\begin{aligned} \partial_{12} h(t, s) &= \langle \gamma'(s), \frac{d}{dt} \left( \frac{\gamma(s) - \gamma(t)}{\sqrt{(\gamma(s) - \gamma(t)) \cdot (\gamma(s) - \gamma(t))}} \right) \rangle = \\ &= \frac{-\langle \gamma'(s), \gamma'(t) \rangle}{h(t, s)} - \frac{\partial_1 h}{h^2} \langle \gamma'(s), \gamma(s) - \gamma(t) \rangle \\ &= \frac{-\cos(\hat{\text{ângulo entre }} \gamma'(t) \text{ e } \gamma'(s)) + \cos(\phi) \cos(\phi_1)}{h} = \\ &= \frac{-\cos(\phi + \phi_1) + \cos(\phi) \cos(\phi_1)}{\sqrt{(\gamma(s) - \gamma(t)) \cdot (\gamma(s) - \gamma(t))}} = \frac{\text{sen}(\phi) \text{sen}(\phi_1)}{h} > 0, \end{aligned}$$

pois tanto  $\phi$  como  $\phi_1$  estão entre 0 e  $\pi$ .

Além disso, se usarmos as coordenadas  $(s, y)$  com  $y = \cos(\phi)$ , vemos que o elemento de área  $dy \wedge ds$  é preservado ou seja

$$d(y_1 ds_1 - y ds) = 0$$

De fato  $\cos(\phi_1) ds_1 - \cos(\phi) ds = \partial_2 h ds_1 + \partial_1 h ds = dh(s, s_1)$ . Portanto, bilhares associados a curvas regulares estritamente convexas são exemplos de aplicações do tipo twist que preservam área. A função geratriz neste caso é o comprimento da corda  $\gamma(s_1) - \gamma(s)$  e órbitas periódicas estão associadas a poligonais fechadas na região  $U$ .

Por exemplo, um diâmetro da curva corresponde a uma órbita de período dois.

Consideremos o caso simples do bilhar no círculo, onde podemos usar geometria elementar. Tomemos o ângulo  $\theta = \frac{\pi}{2} - \phi$  com a normal (raio do círculo). Logo,  $F(s, \theta) = (s + \pi - 2\theta, \theta)$ .

Como a segunda coordenada é preservada, vemos que o espaço de fase  $(s, \theta)$  fica totalmente decomposto (folheado) por curvas invariantes do tipo  $\theta = \theta_0$ .

Restrito a cada uma das folhas a aplicação é uma translação (rotação no cilindro) de ângulo  $\theta_0$ .

É possível observar também que para cada valor de um ângulo da saída  $\theta_0$  corresponde um círculo centrado na origem que é tangente a todas as cordas que fazem ângulo igual a  $\theta_0$  com a normal. Estes círculos são chamados *cáusticas do bilhar* e correspondem exatamente às curvas invariantes.

Outra observação de fácil verificação é que um ponto  $(s, \theta)$  é um ponto periódico, de período  $q$  se e somente se existe um número inteiro  $k$  tal que

$$\theta = \left(\frac{q - 2k}{2q}\right) \frac{\pi}{2}.$$

Um outro exemplo, mais interessante, com dinâmica simples é o do bilhar numa elipse, cujo espaço de fase é representado na seguinte figura.

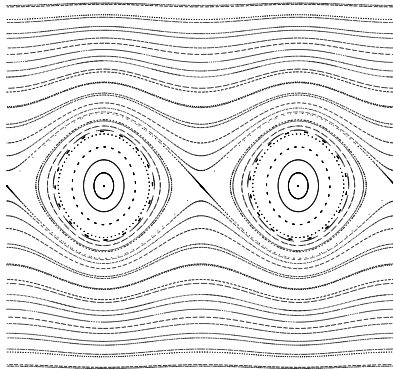


Figura 1.14: aplicação twist associada ao bilhar na elipse

É bastante fácil encontrar na Internet páginas onde podemos visualizar as propriedades do bilhar na elipse, inclusive com animação.

Destacamos algumas propriedades conhecidas do bilhar elíptico:

1) Existem dois pontos de período dois, que correspondem aos eixos da elipse. Apesar de ambos serem pontos periódicos de mesmo período observa-se a grande diferença na dinâmica próxima a estes pontos.

Para o diâmetro maior o ponto periódico associado é instável (hiperbólico do tipo sela) existem condições iniciais bem próximas ao diâmetro, após um certo tempo se afastam do diâmetro. Além disso, existe uma curva de condições iniciais cujas órbitas tendem assintoticamente à órbita periódica.

Ao passo que o diâmetro menor corresponde a um ponto periódico estável, ou seja, pontos suficientemente próximos do diâmetro menor e com ângulo de saída próximo a  $\frac{\pi}{2}$  permanecerão sempre numa vizinhança deste diâmetro.

2) Raios que passam por um dos focos da elipse quando refletidos passarão pelo outro foco e com o passar do tempo as cordas correspondentes terão inclinações cada vez menores tendendo (mas nunca atingindo) ao eixo maior da elipse.

Isto segue da propriedade da elipse: os raios que partem de um dos focos refletem em raios que passam pelo outro foco.

Esta propriedade geométrica implica as seguintes propriedades cujas demonstrações são exercícios de geometria elementar:

3) Cordas que cruzam o eixo maior da elipse entre os focos quando refletidas, geram cordas que cruzam o eixo maior entre os focos. As cordas correspondentes evoluem hipérbolas confocais com a elipse do bilhar.

4) Cordas que cruzam o eixo maior em pontos fora do segmento entre os focos quando refletidas voltarão a cruzar o eixo maior fora dos focos. As cordas correspondentes evoluem elipses (cáuticas) confocais com a elipse do bilhar.

As órbitas correspondentes a tais cordas estão contidas em curvas homotopicamente não triviais no cilindro. O bilhar na elipse guarda alguma semelhança com o bilhar no círculo: existe um sub-anel que está totalmente folheado por curvas invariantes restrita às quais a aplicação de bilhar é uma rotação.

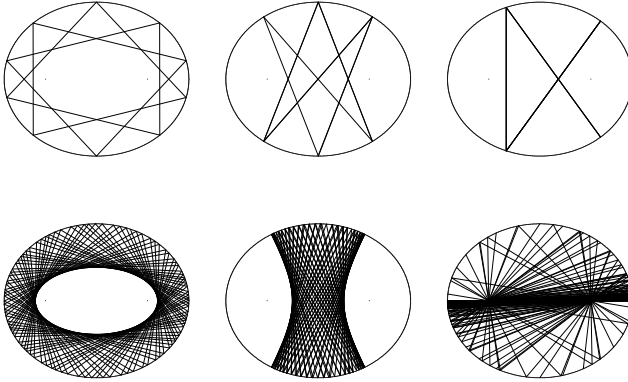


Figura 1.15: Bilhar na elipse

Uma figura totalmente distinta das que acabamos de ver é obtida considerando-se o bilhar na imagem da seguinte curva:

$$\alpha(t) = \left( \cos(t), \frac{3}{2 - \sin(t)} \right).$$

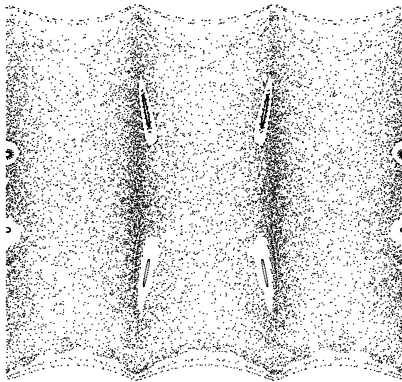


Figura 1.16: aplicação do twist associada a  $\alpha(t)$

Estes exemplos ilustram algumas das propriedades que buscamos descrever quando estudamos a dinâmica de aplicações do tipo twist,

algumas das quais estudaremos nos próximos capítulos:

- a) Existência de pontos periódicos, órbitas de Birkhoff
- b) Classificação dos pontos periódicos
- c) Existência de curvas rotacionais invariantes - Teoria de Kolmogorov-Arnold-Moser (KAM)
- d) Existência de outros conjuntos invariantes que são gráficos sobre subconjuntos compactos do círculo - Conjuntos de Aubry-Mather.
- e) Existência de órbitas que se aproximam, no futuro ou no passado, de outras órbitas - conjunto estável e instável de órbitas periódicas
- f) Ergodicidade de bilhares.

## 1.5 Dinâmica local em torno de uma órbita periódica

Considere uma função  $H : \mathbb{R}^4 \rightarrow \mathbb{R}$  de classe  $C^k$ ,  $k \geq 2$ ,  $H(x, y)$  com  $x = (x_1, x_2)$  e  $y = (y_1, y_2)$ . O campo Hamiltoniano associado a  $H$  é, o campo definido por  $X_H(x, y) = J \text{grad } H(x, y)$  onde  $\text{grad } H$  denota o campo gradiente de  $H$  e  $J(x, y) = (y, -x)$

$$\text{Em coordenadas, } X_H = \frac{\partial H}{\partial y_1} \frac{\partial}{\partial x_1} + \frac{\partial H}{\partial y_2} \frac{\partial}{\partial x_2} - \frac{\partial H}{\partial x_1} \frac{\partial}{\partial y_1} - \frac{\partial H}{\partial x_2} \frac{\partial}{\partial y_2}$$

É claro que  $H$  é constante ao longo das órbitas de  $X_H$ , isto é, se  $\phi_t$  é o fluxo gerado por  $X_H$ , então  $H \circ \phi_t$  não depende de  $t$ .

Seja  $(x_0, y_0)$  um ponto periódico de  $\phi_t$  de modo que a curva  $\gamma =: \phi_t(x_0, y_0)$  é fechada.

Como se sabe, as propriedades dinâmicas do fluxo  $\phi_t$ , em uma vizinhança de  $\gamma$  podem ser obtidas tomando-se uma seção transversal local  $\Sigma$  passando por  $(x_0, y_0)$  e analisando-se a aplicação de primeiro retorno (aplicação de Poincaré).

De fato, pela conservação de energia, basta observar a aplicação de primeiro retorno definida numa seção bi-dimensional  $\Sigma_e$  difeomorfa a um disco, e contida no nível de energia  $H^{-1}(e)$  que contém  $\gamma$ .

Esta aplicação é um difeomorfismo local que preserva área (ou uma forma de grau dois não degenerada, múltipla do elemento de área) e que possui um ponto fixo em  $(x_0, y_0)$ .

Vejamos porque:

Observe que  $(x_0, y_0)$  é um ponto regular do campo  $X_H$ , em particular,  $\text{grad } H(x_0, y_0) \neq 0$  e o nível de energia  $H^{-1}(e)$  é regular numa

vizinhança da curva  $\gamma$ .

Suponhamos que  $\frac{\partial H}{\partial y_1}(x_0, y_0) \neq 0$

Então o mapa  $F(x_1, x_2, y_1, y_2) = (x_1, x_2, H(x, y) - e, y_2)$  é um difeomorfismo local que leva o nível  $H^{-1}(e)$  no hiperplano  $y_1 = 0$ .

Além disso, usando o fato que  $\frac{dx_1}{dt}(x_0, y_0) \neq 0$ , podemos reparametrizar o tempo de modo que a obter um novo campo que se escreve da seguinte forma:

$$\begin{aligned}\frac{dx_2}{dx_1} &= \frac{\partial_{y_2} H}{\partial_{y_1} H} \\ \frac{dy_1}{dx_1} &= -\frac{\partial_{x_1} H}{\partial_{y_1} H} \\ \frac{dy_2}{dx_1} &= -\frac{\partial_{x_2} H}{\partial_{y_1} H}\end{aligned}$$

Definindo  $K(x_1, x_2, e, y_2)$  por  $F^{-1}(x_1, x_2, e, y_2) = (x_1, x_2, -K, y_2)$  de modo que  $H(x_1, x_2, -K, y_2) = e$  e usando a derivação implícita, podemos escrever o sistema acima na forma:

$$\begin{aligned}\frac{dx_2}{dx_1} &= \partial_{y_2} K \\ \frac{dy_2}{dx_1} &= -\partial_{x_2} K \\ \frac{dE}{dx_1} &= 0\end{aligned}$$

Ou seja, localmente o fluxo hamiltoniano restrito a uma superfície de nível regular é o fluxo  $\psi$  gerado por um novo Hamiltoniano, não autônomo (dependente do tempo), associado à função  $K$ .

Segue-se que este novo fluxo preserva a área de seções transversais pois  $|\det d\psi_{x_1}(x_2, y_2)| = 1$ .

Em particular, tomando-se a transformação de Poincaré ao longo da órbita  $\gamma$ , obtemos uma aplicação definida em uma vizinhança de  $(x_0, y_0)$  na seção  $\Sigma_e$ , que preserva área e que possui um ponto fixo em  $(x_0, y_0)$ .

Se este ponto fixo for elíptico, podemos usar a Forma Normal de Birkhoff e obter condições (genéricas) para que esta aplicação seja



do tipo twist. Finalmente, usando o Teorema do Twist de Moser, pode-se obter uma família de toros invariantes em torno da órbita periódica. Para uma discussão mais profunda sobre a dinâmica em torno de uma órbita periódica, consulte [1].

## 1.6 Aplicação de primeiro retorno do laço sela-centro

Nesta parte, apresentaremos um exemplo de aplicação do tipo twist que provém de um problema mecânico, do estudo de sistemas Hamiltonianos com 2 graus de liberdade com singularidades do tipo sela-centro e órbitas homoclínicas a estas. Apesar desta análise ser mais geral, para facilitar o entendimento o estudo será feito a partir de um exemplo. Seja a seguinte família a um parâmetro de sistemas Hamiltonianos em  $\mathbb{R}^4$ , com a forma simplética usual ( $\omega = dx \wedge dp_x + dy \wedge dp_y$ ):

$$H = \frac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2) + bx^2y - \frac{y^3}{3} \quad (1.13)$$

Então a dinâmica deste sistema é regida pelas seguintes equações:

$$\begin{cases} \dot{p}_x = -x - 2bxy \\ \dot{x} = p_x \\ \dot{p}_y = -y - bx^2 + y^2 \\ \dot{y} = p_y \end{cases} \quad (1.14)$$

Para todo valor de  $b \geq 0$ , este sistema tem o seguinte ponto de equilíbrio:

$\eta = (x, y, p_x, p_y) = (0, 1, 0, 0)$ . É fácil ver que  $\eta$  é uma singularidade do tipo sela-centro, com um par de autovalores reais  $\pm\nu = \pm 1$  e um par de imaginários  $\pm\omega = \pm\sqrt{1+2bi}$ .

Por outro lado, o plano ( $x = 0, y, p_x = 0, p_y$ ) é claramente invariante e nele a dinâmica é regida pela seguinte equação:

$$\ddot{y} = -y + y^2$$

Assim nesse plano temos um contínuo de órbitas periódicas com energia menor que  $\frac{1}{6}$  que se acumulam em uma órbita homoclínica à singularidade  $\eta$ , digamos  $\Gamma$ , com  $H(\Gamma) = \frac{1}{6}$ .

Seja agora  $\Sigma = \{(x, y = 0, p_x, p_y > 0)\}$  uma seção de Poincaré para  $\Gamma$ . Neste caso, como  $\Gamma$  não é órbita periódica, é claro que a aplicação de Poincaré  $f : \Sigma \rightarrow \Sigma$  nem sempre está definida, pois uma trajetória que sai de um ponto de  $\Sigma$ , próximo a  $\Gamma$  pode, ao se aproximar de  $\eta$ , acompanhar o ramo não homoclínico à  $\eta$  de  $W^U(\eta)$ , não retornando mais a  $\Sigma$  (ver [2] e [36]).

Da conservação da energia, é natural definir uma família a um parâmetro ( $E$ ) de aplicações de Poincaré  $f_E : \Sigma_E \rightarrow \Sigma_E$ , onde  $\Sigma_E$  é a restrição de  $\Sigma$  à superfície de energia  $H^{-1}(E)$  (para cada valor de  $E$ , tomamos  $p_y = \sqrt{2E - p_x^2 - x^2}$ ). Assim,  $(x, p_x)$  é um sistema de coordenadas para  $\Sigma_E$ , desde que  $2E - p_x^2 - x^2 > 0$ .

Como  $\Sigma$  é seção transversal a  $\Gamma$  em  $M$ ,  $\Sigma$  é 3-dimensional, assim  $\Sigma_E$  é 2-dimensional. Pode-se mostrar que para uma classe grande e interessante de sistemas Hamiltonianos (classe que inclui o exemplo acima), a aplicação  $f_{1/6} : \Sigma_{1/6} \rightarrow \Sigma_{1/6}$  está sempre bem definida (ver [36] e [57]).

Com as escolhas acima, para  $0 \leq E \leq \frac{1}{6}$ , o ponto  $(x, p_x) = (0, 0) \stackrel{def}{=} \mathbf{0}$  é fixo para  $f_E$ . Se  $E < \frac{1}{6}$ , esse ponto representa a órbita periódica do plano  $(x = 0, y, p_x = 0, p_y)$ , de energia  $E$ . A aplicação  $f_E$  é claramente diferenciável no  $\mathbf{0}$ . Se  $E = \frac{1}{6}$ , o ponto  $\mathbf{0}$  representa a órbita homoclínica  $\Gamma$  e  $f_E$  não é mais diferenciável nesse ponto.

Foi mostrado por Lerman [48] e Mielke et al. [57] (veja também [35] e [36]), que num sistema de coordenadas conveniente, para

$|E|, \|z\|$  ( $z = (x, p_x)$ ) suficientemente pequenos,  $f_E$  é dada aproximadamente pela seguinte família de aplicações (ver [35]):

$$F_E(z) = AR[c - \gamma \log \frac{1}{2}\omega \|z\|^2 - (E - \frac{1}{6})]z, \quad (1.15)$$

onde  $z = (x, p_x) \in \mathbb{R}^2$ ,  $\|z\| > \sqrt{\frac{2}{\omega}(E - \frac{1}{6})}$  para  $E > \frac{1}{6}$ ,  $F_{\frac{1}{6}}(\mathbf{0}) \stackrel{def}{=} \mathbf{0}$ ,  $c$  é uma constante e  $\gamma = \sqrt{1 + 2b}$ , com

$$R(\theta) \stackrel{def}{=} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad A \stackrel{def}{=} \begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix}.$$

Para o exemplo em questão, pode ser mostrado que  $\alpha = B + \sqrt{B^2 + 1}$ , onde  $B$  é dado por:

$$\left| \frac{\cos(\frac{\pi}{2}\sqrt{1-48b})}{\sinh(2\pi\sqrt{1+2b})} \right| \quad \text{ou} \quad \left| \frac{\cosh(\frac{\pi}{2}\sqrt{48b-1})}{\sinh(2\pi\sqrt{1+2b})} \right|,$$

para  $b \leq \frac{1}{48}$  e  $b \geq \frac{1}{48}$  respectivamente.

Mais geralmente, o número  $\alpha \geq 1$  é obtido a partir de um problema de espalhamento (ver referências [35] e [36]), sendo uma espécie de "coeficiente de Floquet" para a órbita homoclínica e  $\gamma = \left\lfloor \frac{\omega}{\nu} \right\rfloor$ . A constante  $c$ , no caso em que  $E = 1/6$ ,  $\sum_{1/6} \cap \Gamma \neq \emptyset$ , não afeta a dinâmica.

Para que

$$F_{1/6}(z) = AR[-\gamma \log \left( \frac{1}{2} \omega \|z\|^2 \right)]z$$

possa ser escrita de uma maneira mais familiar, vamos aplicar a seguinte mudança de coordenadas do tipo polar que remove a singularidade logaritmica que existe na origem (vamos supor que  $z = (z_1, z_2)$ ):

$$\begin{cases} z_1 = \sqrt{2}e^{\tilde{I}/(2\gamma)} \cos(\tilde{I} - \tilde{\phi}) \\ z_2 = \sqrt{2}e^{\tilde{I}/(2\gamma)} \sin(\tilde{I} - \tilde{\phi}). \end{cases}$$

Nessas novas coordenadas a aplicação se escreve como:

$$\tilde{f} : \begin{cases} \tilde{\phi}' = \mu(\tilde{\phi}) + \tilde{I}' \\ \tilde{I}' = \gamma \log J(\tilde{\phi}) + \tilde{I} \end{cases}, \text{ onde} \quad (1.16)$$

$$\begin{aligned} J(\tilde{\phi}) &= \alpha^2 \cos^2(\tilde{\phi}) + \alpha^{-2} \sin^2(\tilde{\phi}) \\ \mu(\tilde{\phi}) &= \arctan \left( \frac{\tan(\tilde{\phi})}{\alpha^2} \right), \text{ com } \mu(0) = 0 \end{aligned}$$

Nessas coordenadas, fica fácil identificar o plano  $z$  menos a origem, com o cilindro  $(\phi, I) \in S^1 \times \mathbb{R}$  onde  $S^1 = \mathbb{R}/\pi\mathbb{Z}$ , ou  $\phi = \phi \bmod \pi$ . Seja  $f$  a aplicação induzida por  $\tilde{f}$  (1.16) nesse cilindro. Claramente  $f$  é um difeomorfismo do tipo twist e do fato de  $F_{1/6}$  preservar área no plano, obtemos que  $f$  preserva a seguinte medida no cilindro:

$$\mu(A) = \int_A e^{I/\gamma} d\phi \wedge dI \quad (1.17)$$

A aplicação  $f$  tem mais uma propriedade notável,  $f(\phi, I + \pi) = f(\phi, I) + (0, \pi)$ . Assim  $f$  induz um difeomorfismo no toro, cujo estudo elucidava diversas propriedades do sistema hamiltoniano original, ver por exemplo [5] e [7].

## 1.7 A forma normal de Birkhoff

Um dos exemplos mais importantes de aplicações do tipo twist ocorre numa vizinhança de um ponto fixo elíptico de uma aplicação que preserva área. Neste caso, os autovalores da derivada estão no círculo unitário e são não reais e, diferentemente do caso hiperbólico, a derivada é insuficiente para descrever a dinâmica local.

Usando coordenadas polares, obtemos uma aplicação no cilindro aberto que preserva área e a propriedade de twist depende das derivadas de ordem superior, como pode ser visto a partir do seguinte resultado:

### Teorema 1. (Forma Normal de Birkhoff)

Seja  $f : (\mathbb{R}^2, 0) \rightarrow (\mathbb{R}^2, 0)$  um germe de difeomorfismo tal que

(a)  $f$  preserva área ( $f^*(dx \wedge dy) - dx \wedge dy = 0$ )

(b)  $Df(0)$  possui autovalores complexos  $\lambda, \bar{\lambda}$  no círculo unitário,  $\lambda = e^{2\pi i \alpha}$  tal que  $\lambda^n \neq 1$  para  $n \in \{1, \dots, q\}$ .

Então existe um germe de difeomorfismo  $h \in C^\infty$ , que preserva área, tal que se  $z = x + iy$ ,  $\bar{z} = x - iy$  são coordenadas complexas em uma vizinhança  $U$  de 0 em  $\mathbb{R}^2$  então

$h \circ f \circ h^{-1}(z) = \lambda e^{2\pi i P(|z|^2)} z + o(|z|^{q-1})$  onde  $P(X) = a_1 + \dots + a_m X^m$  é um polinômio real de grau  $m$  tal que  $2m + 1 < q$ .

**Corolário 1.** Segue da Forma Normal de Birkhoff que se  $P \not\equiv 0$ , ou seja se algum coeficiente  $a_j \neq 0$  então  $f$  satisfaz localmente a condição de twist.

*Demonstração do Corolário.* Basta escrever  $z = re^{i\theta}$ , de modo que

$$\begin{aligned} g(\theta, r) &= h \circ f \circ h^{-1}(re^{i\theta}) = \lambda r e^{i\theta} \cdot e^{2\pi i P(r^2)} + o(|r|^{q-1}) \\ &= (r e^{i(\theta + 2\pi[\alpha + P(r^2)])}) + o(|r|^{q-1}) = \\ &= (\theta + 2\pi\alpha + 2\pi P(r^2) + \theta(|r|^{q-1}), r + \theta(|r|^{q-1})) = (\theta_1, r_1). \end{aligned}$$

$$\frac{\partial \theta_1}{\partial r} = 2\pi P'(r) \cdot 2r + o(|r|^{q-2}) \neq 0 \quad \text{se } r \neq 0, |r| \text{ pequeno.}$$

□

Estamos interessados nas propriedades dinâmicas de  $f$  numa vizinhança arbitrariamente pequena de 0. Por isso, utilizamos o conceito de germe de uma aplicação:

**Definição 5.** Dizemos que duas aplicações  $f$  e  $f_1$  são equivalentes se existe uma vizinhança  $V$  de 0 em  $\mathbb{R}^2$  tal que  $f|_V = f_1|_V$ . Um germe de uma aplicação em 0 é uma classe de equivalência de aplicações em 0. Portanto um representante de um germe em 0 é uma aplicação definida em uma vizinhança de 0.

**Definição 6.** Dizemos que duas aplicações  $f$  e  $g$  possuem contato de ordem maior que  $k$  se  $d^{(j)}f(0) = d^{(j)}g(0)$  para  $0 \leq j \leq k$ , ou seja se os respectivos  $k$ -jatos em 0 (isto é, os polinômios de Taylor de grau  $k$ ) coincidem:  $j^k f(0) = j^k g(0)$ .

Neste caso usamos também a notação  $f \equiv g \pmod{(z^r)}$  para indicar que a diferença  $f - g$  é uma aplicação que possui todas as suas derivadas até ordem  $k$  nulas.

**Observação 1:** Se  $f$  e  $g$  são germes de difeomorfismo em 0, então  $j^k f(0) = j^k g(0)$  se e somente se  $j^k(f \circ g^{-1})(0) = id$  ou seja  $f \equiv g \pmod{(x^{k+1})}$  se e somente se  $f \circ g^{-1} \equiv id \pmod{(x^{k+1})}$  onde  $id$  é o germe da identidade.

*Demonstração.* Se  $j^n f(0) = j^n g(0)$  então  $f(x) = P_n(x) + R(x)$  e  $g(x) = P_n(x) + H(x)$  onde  $P_n$  é o polinômio de Taylor de  $f$  e  $g$  no ponto 0 e  $H$  e  $R$  são funções tais que  $d^k H(0) = d^k R(0) = 0$  para  $0 \leq k \leq n$ .

Logo  $g \circ g^{-1} = P_n(g^{-1}(x)) + H(g^{-1}(x)) = x$  e  $f \circ g^{-1}(x) = P_n(g^{-1}(x)) + R(g^{-1}(x)) = x - H(g^{-1}(x)) + R(g^{-1}(x))$ . Portanto,

$$j^n(f \circ g^{-1})(0) = j^n id(0) \quad \text{ou seja } f \circ g^{-1} = id \pmod{(x^n)}.$$

A recíproca é obtida escrevendo  $f(x) = P_n(x) + H(x)$  com  $d^k H(0) = 0$  para  $0 \leq k \leq n$ . Por hipótese,  $f \circ g^{-1}(x) = P_n(g^{-1}(x)) +$

$H(g^{-1}(x)) = x \pmod{x^n}$ . Portanto,  $P_n(g^{-1}(x)) = x \pmod{x^n}$ . Fazendo  $y = g^{-1}(x)$ , obtemos  $P_n(y) = g(y) \pmod{y^n}$  ou seja  $f(y) \equiv g(y) \pmod{y^n}$ .  $\square$

**Observação 2:** Vejamos como caracterizar a conservação de área nas coordenadas  $(z, \bar{z})$ . Se  $z = x + iy$  então  $\bar{z} = x - iy$  e  $dz \wedge d\bar{z} = -2i dx \wedge dy$ . Logo, se  $f(x, y) = (X, Y)$  e  $Z = X + iY$ , então  $dx \wedge dy - dX \wedge dY = 0$  se e somente se  $dZ \wedge d\bar{Z} - dz \wedge d\bar{z} = 0$ .

Seja  $F(z, \bar{z}) = Z(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i})$ . Suponha que  $F(z, \bar{z}) \equiv \lambda z + R_q(z, \bar{z}) + 0(|z|^{q+1})$  onde  $R_q(z, \bar{z}) = \sum_{k+l=q} a_{k,l} z^k \bar{z}^l$  é um polinômio homogêneo de grau  $q$ .

Então se  $Z = F(z, \bar{z})$ ,  $\bar{Z} = \overline{F(z, \bar{z})}$  então  $dZ \wedge d\bar{Z} =$

$$\begin{aligned} &([\lambda + \sum a_{k,l} k z^{k-1} \bar{z}^l] dz + [\sum a_{k,l} l z^k \bar{z}^{l-1}] d\bar{z}) \wedge \\ &\wedge ([\sum \bar{a}_{k,l} l \bar{z}^k z^{l-1}] d\bar{z} + [\bar{\lambda} + \sum \bar{a}_{k,l} k \bar{z}^{k-1} z^l] dz) = dz \wedge d\bar{z} \pmod{z^{q+1}} \end{aligned}$$

se e somente se

$$\begin{aligned} &(\lambda + \sum a_{k,l} k z^{k-1} \bar{z}^l)(\bar{\lambda} + \sum \bar{a}_{k,l} k \bar{z}^{k-1} z^l) - \\ &-(\sum a_{k,l} l z^k \bar{z}^{l-1})(\sum \bar{a}_{k,l} l \bar{z}^k z^{l-1}) = 1 \pmod{z^{q+1}}. \end{aligned}$$

o que implica

$$\begin{aligned} &\lambda \sum \bar{a}_{k,l} k \bar{z}^{k-1} z^l + \bar{\lambda} \sum a_{k,l} k z^{k-1} \bar{z}^l = 0 \\ &\text{ou seja } \lambda \sum \bar{a}_{k+1,l} (k+1) z^l \bar{z}^k + \bar{\lambda} \sum a_{k+1,l} (k+1) z^k \bar{z}^l = 0 \end{aligned}$$

Trocando os índices  $k \rightarrow l$  no primeiro somatório obtemos:

$$\lambda \sum \bar{a}_{l+1,k} (l+1) z^k \bar{z}^l + \bar{\lambda} \sum a_{k+1,l} (k+1) z^k \bar{z}^l = 0$$

Fixando o par  $(k, l) \in \mathbb{N} \times \mathbb{N}$ , tal que  $k + l = q$  obtemos as condições:

$$\lambda(l+1)\bar{a}_{l+1,k} + \bar{\lambda}(k+1)a_{k+1,l} = 0.$$

*Demonstração.* (do teorema 1) A prova do Teorema, que está baseada em [46], é feita por indução no grau  $q$ . Para  $q = 2$ ,  $f(z, \bar{z}) = \lambda z + O(|z|)$ , a fórmula é válida.

Para o passo de indução escreva

$$f(z, \bar{z}) = \lambda f_0(z, \bar{z}) + 0(|z|^{q-2}) \quad \text{onde } f_0(z, \bar{z}) = z e^{2\pi i Q(|z|^2)}$$

com  $Q$  um polinômio tal que  $2.\text{grau}(Q) + 1 < q - 1$  e  $f_0(z, \bar{z})$  preserva área.

**Lema 2.** Se  $f \circ f_0^{-1}(z, \bar{z}) = \lambda z + R(z, \bar{z}) \bmod(z^q)$  onde  $R(z, \bar{z}) = \sum_{k+j=q-1} a_{kj} z^k \bar{z}^j$  é um polinômio homogêneo de grau  $(q-1)$ , então existe um polinômio homogêneo  $S(z, \bar{z})$  de grau  $(q-1)$  que define um germe de difeomorfismo  $h(z, \bar{z}) = z + S(z, \bar{z})$  com

$$h \circ f \circ f_0^{-1} \circ h^{-1}(z, \bar{z}) = \begin{cases} \lambda z \bmod(z^q) & \text{se } q \text{ for ímpar} \\ \lambda z + Az^{(n+1)}\bar{z}^n \bmod(z^q) & \text{se } q = 2n + 2 \end{cases}$$

*Demonstração.* Observe que se  $h(z, \bar{z}) = z + S(z, \bar{z}) \bmod(z^q)$  então  $h^{-1}(z, \bar{z}) = z - S(z, \bar{z}) \bmod(z^q)$ . De fato

$$\begin{aligned} h(z - S(z, \bar{z}), \bar{z} - \bar{S}(z, \bar{z})) &= z - S(z, \bar{z}) + S(z - S(z, \bar{z}), \bar{z} - \bar{S}(z, \bar{z})) \\ &= z - S(z, \bar{z}) + S(z, \bar{z}) \bmod(z^q) = z \bmod(z^q). \end{aligned}$$

Como estamos resolvendo formalmente a equação, basta escrever  $S(z, \bar{z}) = \sum_{k+j=q-1} b_{kj} z^k \bar{z}^j$  e resolver a equação

$$\begin{aligned} f \circ f_0^{-1} \circ h^{-1}(z, \bar{z}) &= \lambda(z - S(z, \bar{z})) + R(z, \bar{z}) \bmod(z^q), \text{ de modo que} \\ h \circ f \circ f_0^{-1} \circ h^{-1}(z, \bar{z}) &= \lambda z - \lambda S(z, \bar{z}) + R(z, \bar{z}) + S(\lambda z, \bar{\lambda} \bar{z}) \bmod(z^q) \\ &= \lambda z - \sum \lambda b_{kj} z^k \bar{z}^j + \sum a_{kj} z^k \bar{z}^j + \sum b_{kj} \lambda^k \bar{\lambda}^j z^k \bar{z}^j = \\ &= \lambda z + \sum (-\lambda b_{kj} + a_{kj} + \lambda^{k-j} b_{kj}) z^k \bar{z}^j \bmod(z^q). \end{aligned}$$

Basta portanto, resolver  $a_{kj} + \lambda(-1 + \lambda^{k-j-1})b_{kj} = 0$ . Se  $q$  for ímpar, isso é sempre possível pois  $k + j = q - 1$  e  $\lambda^{k-j-1} = 1$  somente se  $k - j = 1$  ou seja  $q = 2k$ .

Portanto definimos  $b_{kj} = \frac{a_{kj}}{\lambda(-1 + \lambda^{k-j-1})}$  para obter a forma desejada.

Se  $q$  for par, a fórmula acima define  $b_{kj}$  para os índices tais que  $q \neq 2k$ . O único termo que não pode ser eliminado é o termo  $a_{\frac{q}{2}, \frac{q}{2}-1}$ . Portanto,

$$h \circ f \circ f_0^{-1} \circ h^{-1}(z, \bar{z}) = \lambda z + Az^{n+1}\bar{z}^n \quad \text{com } n + 1 = \frac{q}{2}.$$

□

**Lema 3.** Para  $h$  encontrado no Lema 2 temos:

$h \circ f \circ h^{-1}(z, \bar{z}) = \lambda z e^{2\pi i P(|z|^2)} \bmod(z^q)$  onde  $P(x) = Q(x)$ , se  $q$  for ímpar e  $P(x) = Q(x) + \frac{A}{2\pi i \lambda} x^n$ ,  $n = \frac{q}{2} - 1$  para  $q$  par.

*Demonstração.* De acordo com o Lema 2, se  $q$  for ímpar, então  $h \circ f \circ f_0^{-1} \circ h^{-1}(z, \bar{z}) = \lambda z \pmod{(z^q)}$  com  $h(z, \bar{z}) = z + S(z, \bar{z})$ .

Mas

$$f_0 \circ h(z, \bar{z}) = [z + S(z, \bar{z})]e^{2\pi i Q(|z+S(z, \bar{z})|^2)} = z e^{2\pi i Q(|z|^2)} + S(z, \bar{z}) \pmod{(z^q)} \text{ ou seja, } f_0 \circ h = h \circ f_0 \pmod{(z^q)}.$$

Logo  $(f_0 \circ h)^{-1} = h^{-1} \circ f_0^{-1} = f_0^{-1} \circ h^{-1} \pmod{(z^q)}$ , o que implica  $h \circ f \circ f_0^{-1} \circ h^{-1} = h \circ f \circ h^{-1} \circ f_0^{-1}(z, \bar{z}) = f_0(z, \bar{z}) = \lambda z e^{2\pi i P(|z|^2)} \pmod{(z^q)}$  no caso em que  $q$  é ímpar.

Para  $q$  par,

$$h \circ f \circ h^{-1} \circ f_0^{-1}(z, \bar{z}) = \lambda z + Az^{n+1}z^n \pmod{(z^q)}$$

ou  $h \circ f \circ h^{-1} \circ f_0^{-1}(z, \bar{z}) = \lambda z [1 + \frac{A}{\lambda} |z|^{2n}] \pmod{(z^q)}$ .

Logo,  $h \circ f \circ h^{-1}(z, \bar{z}) = \lambda z e^{2\pi i P(|z|^2)} (1 + \frac{A}{\lambda} |z|^{2n}) \pmod{(z^q)}$  ou  $h \circ f \circ h^{-1}(z, \bar{z}) = \lambda z e^{2\pi i [P(|z|^2) + A'|z|^{2n}]} \pmod{(z^q)}$ , onde  $A' = \frac{A}{\lambda 2\pi i}$ . Como queríamos demonstrar. □

**Lema 4.** *Se  $q$  é ímpar, existe  $H$ , um germe de difeomorfismo em  $0$ , que preserva área, tal que  $H = h \pmod{(z^q)}$ . Se  $q$  for par então  $\frac{A}{2\pi i \lambda}$  é real e podemos escolher o coeficiente  $b_{n+1, n}$  imaginário puro para obter  $H$  um germe de difeomorfismo no  $0$  que preserva área, tal que  $H = h \pmod{(z^q)}$ .*

*Demonstração.* Estamos em condições de aplicar a Observação 2 pois

$$f \circ f_0^{-1}(z, \bar{z}) = \lambda z + R(z, \bar{z}) \pmod{(z^q)}, \text{ com } R(z, \bar{z}) = \sum_{k+j=q-1} a_{k,j} z^k \bar{z}^j$$

Como  $f \circ f_0^{-1}(z, \bar{z})$  preserva área, vale

$$\lambda(j+1)\bar{a}_{j+1,k} + \bar{\lambda}(k+1)a_{k+1,j} = 0$$

Mas pela definição de  $h$ ,  $b_{k,j} = \frac{a_{k,j}}{\lambda(-1+\lambda^{k-j-1})}$ . Substituindo na equação acima temos

$$\begin{aligned} \lambda(j+1)\bar{b}_{j+1,k}\bar{\lambda}[-1+\bar{\lambda}^{j-k}] + \bar{\lambda}(k+1)b_{k+1,j}\lambda(-1+\lambda^{k-j}) &= 0 \\ \text{ou } (j+1)\bar{b}_{j+1,k}[-1+\bar{\lambda}^{j-k}] + (k+1)b_{k+1,j}[-1+\lambda^{k-j}] &= 0. \end{aligned}$$



Mas  $\bar{\lambda} = \lambda^{-1}$ , portanto,  $\bar{\lambda}^{j-k} = \lambda^{k-j}$  e  $(j+1)\bar{b}_{j+1,k} + (k+1)b_{k+1,j} = 0$  o que é exatamente a condição de  $h$  preservar área no grau  $q-1$ .

No caso em que  $q$  é par, obtemos as mesmas condições exceto para  $j = n = k$ . Como  $f \circ f_0^{-1}$  preserva área, temos  $\lambda(n+1)\bar{a}_{n+1,n} + \bar{\lambda}(n+1)a_{n+1,n} = 0$ . Logo,  $\lambda\bar{a}_{n+1,n}$  é imaginário puro, isto é,  $A' = \frac{a_{n+1,n}}{\lambda 2\pi i}$  é real e podemos tomar  $b_{n+1,n}$  um número imaginário puro qualquer.

Em ambos os casos, obtivemos que  $h(z, \bar{z})$  preserva área até ordem  $q$ , ou seja se  $h(z, \bar{z}) = (Z, \bar{Z})$  então  $dz \wedge d\bar{z} - dZ \wedge d\bar{Z} = 0 \pmod{(z^q)}$ .

Antes de obter  $H$  preservando área tal que  $H = h \pmod{(z^q)}$ , provemos um fato geral sobre aplicações que preservam área.

**Observação 3:** Se  $f : \mathbb{R}^2, 0 \rightarrow \mathbb{R}, 0$  é uma função de classe  $C^r$ ,  $r \geq 2$  tal que  $df(0, 0) = (0, 0)$ ,

$$\text{então o sistema de equações(E)} \quad \begin{cases} X = x + \partial_2 f(x, Y) \\ y = Y + \partial_1 f(x, Y) \end{cases} \text{ define}$$

implicitamente uma aplicação  $F(x, y) = (X, Y)$  em uma vizinhança de  $(0, 0) \in \mathbb{R}^2$ , que preserva área.

*Demonstração.* A existência da aplicação segue do Teorema da Função Implícita. Para checar que  $dX \wedge dY = dx \wedge dy$  basta calcular

$$dX \wedge dY - dx \wedge dy = [1 + \partial_{12} f(x, y)] dx \wedge dY - dx \wedge (1 + \partial_{21} f(x, Y)) dY = 0.$$

Podemos fazer um pouco mais: suponha que  $U(x, y)$  e  $V(x, y)$  sejam funções que se anulam em  $(0, 0)$  assim como os respectivos gradientes  $\nabla U(0, 0) = \nabla V(0, 0) = (0, 0)$ . Além disso, suponha que  $\partial_1 U + \partial_2 V = 0$ .

$$\text{Então as equações} \quad \begin{cases} X = x + U(x, Y) \\ y = Y - V(x, Y) \end{cases} \text{ definem localmente}$$

uma aplicação  $F(x, y) = (X, Y)$  que preserva área.

De fato, a forma  $-V(x, y)dx + U(x, y)dy$  é uma diferencial exata, logo, existe uma função  $f(x, y)$  definida numa vizinhança de  $(0, 0)$  em  $\mathbb{R}^2$  tal que  $\partial_1 f(x, y) = -V(x, y)$  e  $\partial_2 f(x, y) = U(x, y)$ . Portanto o sistema de equações acima, localmente se escreve como o sistema (E).

□

Voltando à conclusão da prova da Forma Normal de Birkhoff, vimos que o germe de difeomorfismo  $h(z, \bar{z}) = z + S(z, \bar{z})$  que obtivemos no Lema 2 preserva área até ordem  $q$ , isto é,  $S(z, \bar{z})$  é um polinômio homogêneo de grau  $(q-1)$  tal que se  $(Z, \bar{Z}) = h(z, \bar{z})$  então  $dZ \wedge d\bar{Z} = dz \wedge d\bar{z} \pmod{(z^q)}$ . Voltando às coordenadas  $(x, y)$ , isto significa que  $h(x, y) = (x + U(x, y), y + V(x, y))$  com  $U$  e  $V$  polinômios homogêneos de grau  $(q-1)$  tais que

$$\det \begin{bmatrix} 1 + \partial_1 U(x, y) & \partial_2 U(x, y) \\ \partial_1 V(x, y) & 1 + \partial_2 V(x, y) \end{bmatrix} = 1$$

ou

$$\partial_1 U + \partial_2 V + \partial_1 U \partial_2 V - \partial_1 V \partial_2 U = 0 \pmod{((x, y)^q)}.$$

O que implica  $\partial_1 U + \partial_2 V = 0$ .

Estamos então em condições de usar o argumento da observação 2, aplicando o Teorema da Função Implícita para as equações

$$\begin{cases} X = x + U(x, Y) \\ y = Y - V(x, Y) \end{cases} \text{ e obtendo uma aplicação } H(x, y) = (X, Y),$$

que preserva área tal que  $H = h \pmod{((x, y)^q)}$ .

Com isso concluímos a prova do Lema 4 e da Forma Normal de Birkhoff.

□

Na solução formal da equação do Lema 2, obtivemos coeficientes do tipo  $\frac{a_{ij}}{\lambda(-1+\lambda^{k-j-1})}$ . Portanto, se  $\lambda^n \neq 1$  para todo natural  $n$ , sempre será possível resolver a equação dos coeficientes da função  $S(z, \bar{z})$  do Lema 1 e portanto obter que existe uma mudança de coordenadas formal tal que  $f(z, \bar{z}) = \lambda z e^{2\pi P(|z|^2)}$ , agora com  $P(x)$  uma série de potências.

Em outras palavras, usando coordenadas polares, vemos que  $f$  é formalmente conjugada a uma aplicação que preserva todos os círculos  $r = r_0$  centrados na origem ("círculos invariantes") e que portanto possui uma dinâmica bastante simples.

Entretanto, a série de potências que define  $h$  em geral é divergente devido à presença de pequenos denominadores onde  $\lambda^n \neq 1$  mas  $\lambda^n - 1$  pode ser arbitrariamente pequeno. A obtenção de condições suficientes para a convergência de  $h$  (Teoria KAM) é um dos mais belos capítulos da história da Matemática do século XX, sobre o qual falaremos um pouco no Capítulo III, sobre Curvas Invariantes.

## 1.8 Hamiltonianos periódicos

As aplicações do tipo twist ajudam a entender a dinâmica do fluxo gerado por um campo hamiltoniano periódico.

**Exemplo 1.** *Seja  $H : T^*S^1 \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $H(\theta, p, t)$  uma função de classe  $C^3$  periódica, de período 1 na variável  $t$ , tal que  $\frac{\partial^2 H}{\partial p^2}(\theta, p, t) \geq \delta > 0$  para todo  $(\theta, p, t) \in T^*S^1 \times \mathbb{R}$  e  $\sup \|d^2 H(\theta, p, t)\| < M$ .*

Denotemos por  $J$  a matriz  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

Então o tempo 1 do fluxo hamiltoniano  $\varphi_t$  gerado pelo campo  $X^H(\theta, p, t) = -J \text{grad } H = \frac{\partial H}{\partial p} \frac{\partial}{\partial \theta} - \frac{\partial H}{\partial \theta} \frac{\partial}{\partial p}$  é a composta de aplicações do tipo twist que preservam área.

Ou seja existe uma partição  $0 = t_0 < t_1 < t_2 < \dots < t_n = 1$  tal que  $\varphi_1 = \varphi_{\delta_n} \varphi_{\delta_{n-1}} \dots \varphi_{\delta_1}$  onde cada  $\varphi_{\delta_i}$  é uma aplicação do tipo twist que preserva a área ( $\delta_i = t_i - t_{i-1}$ ).

*Demonstração.* Observe que o elemento área  $d\theta \wedge dp$  se escreve

$$d\theta \wedge d_p((u_1, u_2), (v_1, v_2)) = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} = u_1 v_2 - u_2 v_1 = \langle Ju, v \rangle .$$

Portanto, para provar que  $\varphi_t$  preserva área basta provar

$$\begin{aligned} (d\varphi_t)^T J d\varphi_t &= J \text{ pois } (\varphi_t)^*(d\theta \wedge dp)(u, v) = \langle Jd\varphi_t u, d\varphi_t v \rangle \\ &= \langle (d\varphi_t)^T J d\varphi_t u, v \rangle = \langle Ju, v \rangle = d\theta \wedge dp(u, v). \end{aligned}$$

Logo, basta provar que  $\frac{d}{dt}((d\varphi_t)^T J d\varphi_t^H) = 0 \forall t$ .

Lembrando que se  $\frac{\partial \varphi_t}{\partial t}(\theta, p, t) = X^H(\varphi_t(\theta, p, t))$  então sabemos que a matriz  $d\varphi_t(\theta, p, t)$  satisfaz à equação  $\frac{\partial}{\partial t}[d\varphi_t(\theta, p, t)] = dX^H(\varphi_t(\theta, p, t))d\varphi_t(\theta, p, t)$ .

Como  $X^H = -J \text{grad } H$  então  $dX^H = -Jd^2 H(\theta, p, t)$ .

Logo  $\frac{\partial}{\partial t}[d\varphi_t(\theta, p, t)] = -Jd^2 H(\varphi_t(\theta, p, t)) d\varphi_t(\theta, p, t)$ .

Usando isso no cálculo da derivada de  $d\varphi_t^T J d\varphi_t$  obtemos:

$$d\varphi_t^T d^2 H(-J)^T J d\varphi_t + d\varphi_t^T J(-J)d^2 H d\varphi_t = 0$$

$$\text{pois } -J^T = J \text{ e } J^2 = -I.$$

Como  $d\varphi_0 = Id$  então provamos o que queríamos,  $d\varphi_t^T J d\varphi_t = J$ .

Passemos agora à condição de twist. Observe que o que está sendo afirmado é que  $\varphi_1$  é **a composta** de aplicações do tipo twist, podendo ela mesma não ser uma aplicação twist. Precisamos observar o sinal do elemento  $[d\varphi_t]_{12}$  da matriz jacobiana do fluxo.

Tomemos uma trajetória  $\phi_t(\theta_0, p_0, t_0)$ , com  $0 < t_0 < 1$ , que satisfaz à condição inicial  $\phi_{t_0}(\theta_0, p_0, t_0) = (\theta_0, p_0, t_0)$ . Logo  $d\phi_{t_0}(\theta_0, p_0, t_0) = Id$ .

Queremos mostrar que existe  $t_1 > t_0$  tal que se  $\phi_{t_1} = (\Theta_{t_1}, P_{t_1})$  são as componentes espaciais, então  $\partial_p \Theta_{t_1} > 0$ .

Mas  $\frac{\partial}{\partial t} \Theta_t = \partial_p H(\theta_t, P_t, t)$ . Portanto derivando em relação a  $p$  obtemos:

$$\frac{\partial}{\partial t} \partial_p \Theta_t = \partial_\Theta \partial_p H(\phi_t, t) \partial_p \Theta_t + \partial_{pp} H \partial_p P_t$$

Logo, a expansão em série de Taylor de  $\partial_p \Theta_t$ , em torno do ponto  $t_0$ , se escreve:

$$\partial_p \Theta_t(\theta, p, t_0) = \partial_p \Theta_{t_0}(\theta, p, t_0) + (t - t_0) \frac{\partial}{\partial t} \partial_p \Theta_{t_0}(\theta, p, t_0) + O(t - t_0)^2$$

Ou seja, usando que  $d\phi_{t_0}(\theta_0, p_0, t_0) = Id$ :

$$\partial_p \Theta_t(\theta, p, t_0) = 0 + (t - t_0) \partial_{pp} H(\theta, p, t_0) + O(t - t_0)^2$$

Finalmente, usando a hipótese de convexidade estrita  $\partial_{pp} H(\theta, p, t_0) \geq \delta > 0$ , obtemos, para  $t - t_0 > 0$  suficientemente pequeno:

$$\partial_p \Theta_t(\theta, p, t_0) = (t - t_0) \delta + O(t - t_0)^2 > 0$$

Como queríamos demonstrar.  $\square$

Se o Hamiltoniano for da forma  $H(\theta, p) = \frac{1}{2}p^2 + V(\theta)$ , então pela conservação da energia:

$$H(\Theta_t(\theta, p), P_t(\theta, p)) = \frac{1}{2}P_t(\theta, p)^2 + V(\Theta_t(\theta, p)) = \frac{1}{2}p^2 + V(\theta) = E$$

Derivando esta expressão em relação a  $p$ , obtemos:

$$P_t(\theta, p) \partial_p P_t + V'(\Theta_t) \partial_p \Theta_t = p$$

Mas  $V'(\Theta_t) = -\frac{\partial}{\partial t}P_t$  e  $P_t(\theta, p) = \frac{\partial}{\partial t}\Theta_t$

Portanto, substituindo temos:

$$P_t(\theta, p) \frac{\partial}{\partial t} [\partial_p \Theta_t] - \frac{\partial}{\partial t} P_t [\partial_p \Theta_t] = p$$

Ou seja, como vimos acima anteriormente,  $\partial_p \Theta_t$  satisfaz uma equação diferencial linear.

Suponhamos que  $E > \max V(\theta)$ . Escolhendo a raiz quadrada positiva,  $p = \sqrt{2(E - V(\theta))}$ , vemos que  $p > 0$  de modo que  $P_t(\theta, p) > 0, \forall t$ .

E, usando o método da variação dos parâmetros, obtemos a solução explícita

$$\partial_p \Theta_t = P_t(\theta, p) \int \frac{p}{P_s(\theta, p)^2} ds$$

que é estritamente positiva.

Suponhamos que  $V(\theta)$  seja uma função de Morse, isto é, todos os seus pontos críticos são não degenerados com valores críticos distintos.

Nestas condições, é fácil fazer o retrato de fase do fluxo gerado pelo campo hamiltoniano. Aos pontos de máximo de  $V$  correspondem selas hiperbólicas.

A condição  $E > \max V(\theta)$  no nível de energia significa que estamos considerando a região complementar à região do cilindro cujo bordo é formado pelos ramos das selas. Esta região está totalmente folheada por níveis de energia que são curvas homotopicamente não triviais no cilindro.

Um dos exemplos da situação que estamos tratando é quando consideramos uma perturbação periódica do caso autônomo, isto é, para Hamiltonianos do tipo  $H(\theta, p, t) = \frac{p^2}{2} + V(\theta) - g(t)$  com  $|g(t)|$  suficientemente pequeno. Neste caso, a condição de twist segue imediatamente por tratar-se de uma condição aberta no espaço de aplicações que preservam área.

De certa forma, um difeomorfismo definido pelo tempo 1 de um fluxo Hamiltoniano em  $T^*S^1 \times \mathbb{R}$  periódico no tempo é o exemplo mais geral de aplicação do tipo twist exata. Isto é o que nos diz o teorema de suspensão de Moser.

**Teorema 2. (Moser, 1986)** *Seja  $f : S^1 \times \mathbb{R} \rightarrow S^1 \times \mathbb{R}$  uma aplicação do tipo twist exata de classe  $C^r$ ,  $r \geq 3$ .*

*Então existe uma função  $H : T^*S^1 \times \mathbb{R} \rightarrow \mathbb{R}$  de período 1 em  $t$ , tal que  $f = \varphi_1$  tempo 1 do fluxo gerado por  $H$ .*

*Além disso, existe uma constante  $d > 0$  tal que  $d^{-1} < \frac{\partial^2 H}{\partial p^2}(\theta, p, t) < d$ , para todo  $(\theta, p, t) \in T^*S^1 \times \mathbb{R}$ .*

A prova deste teorema que é baseada em [52], está descrita em uma série de lemas, o primeiro deles estabelece condições suficientes para que um campo periódico seja Hamiltoniano.

**Lema 5.** *Suponha que  $X(\theta, p, t)$  seja um campo vetorial em  $T^*S^1$  periódico em  $t$ , de período 1 que gera uma família de difeomorfismos  $\varphi_t : T^*S^1 \rightarrow T^*S^1$  tal que*

$$\begin{aligned} (1) \varphi_{t+1}(\theta, p) &= \varphi_t \circ \varphi_1(\theta, p) \\ (2) \varphi_t &\text{ é exata isto é } \varphi_t^*(\alpha) - \alpha = du_t \end{aligned}$$

onde  $\alpha$  é a forma de Liouville,  $\alpha = pd\theta$ .

*Então existe uma função  $H : T^*S^1 \times \mathbb{R} \rightarrow \mathbb{R}$  periódica em  $t$  tal que  $\varphi_t = \varphi_t^H =$  fluxo do campo Hamiltoniano  $X_H$ .*

*Em outras palavras,  $\omega(X_t, Y) = dH_t \cdot Y$  para todo campo  $Y$  definido em  $T^*S^1$  e  $i_{X_t}\omega(X_t, Y) = dH_t$ , onde  $\omega = -d\alpha$  é a forma simplética.*

*Demonstração.* Primeiramente prova-se que a 1-forma  $Y \rightarrow \omega(X_t, Y)$  é fechada.

Observe que  $\varphi_t^*\alpha = du_t$  implica  $\varphi_t^*(\omega) = -\varphi_t^*(d\alpha) = -d(\varphi_t^*\alpha) = -d\alpha = \omega$ . Logo  $\lim_{s \rightarrow 0} \frac{\varphi_{t+s}^*\omega - \varphi_t^*\omega}{s} = 0$  ou seja  $L_{X_t}\omega = 0$ .

Mas  $L_{X_t}(\omega) = d(i_{X_t}\omega) + e_{X_t}(d\omega) = 0$  (veja [50]). Como  $d\omega = 0$ , temos  $i_{X_t}(\omega)$  é uma forma fechada.

Para provar que  $i_{X_t}(\omega)$  é exata, basta provar que para qualquer curva fechada  $\gamma$  em  $T^*S^1$  temos  $\int_\gamma i_{X_t}(\omega) = 0$  ou  $\int_\gamma i_{X_t}(d\alpha) = \int_\gamma d(i_{X_t}\alpha) = 0$ .

Usando novamente o fato de que  $\varphi_t$  é exata

$$L_{X_t}(\alpha) = \lim_{s \rightarrow 0} \frac{\varphi_{t+s}^*\alpha - \varphi_t^*\alpha}{s} = \lim_{s \rightarrow 0} \frac{du_{t+s} - du_t}{s}.$$

Logo  $\int_\gamma L_{X_t}(\alpha) = \lim_{s \rightarrow 0} \frac{1}{s} \int_\gamma du_{t+s} - du_t = 0$ .

Usando novamente a fórmula

$$L_{X_t}(\alpha) = d(i_{X_t}\alpha) + i_{X_t}(d\alpha) \text{ obtem-se } \int_{\gamma} d(i_{X_t}\alpha) + \int_{\gamma} i_{X_t}(d\alpha) = 0.$$

Como a primeira parcela é nula, conclui-se que  $\int_{\gamma} L_{X_t}(\omega) = 0$ , como queríamos. Logo existe  $H(\theta, p, t)$  tal que  $i_{X_t}(\omega) = dH_t$  ( $H_t(\theta, p) = H(\theta, p, t)$ ) o que implica  $X_t = X_t^H = -J \text{ grad } H_t$  com  $H$  periódica em  $t$ .  $\square$

**Lema 6.** *Seja  $f : T^*S^1 \rightarrow T^*S^1$  um difeomorfismo de classe  $C^\infty$ . Então existem um difeomorfismo vertical  $S(\theta, p) = (\theta, \Upsilon(\theta, p))$  e uma translação  $T_c(\theta, p) = (\theta, p + c)$  tais que  $T_c \circ S \circ f$  é exata.*

*Demonstração.* Primeiramente definimos  $S$  de modo que  $S \circ f$  preserve área.  $d(S \circ f) = dS \circ f \cdot df$  e  $dS = \begin{bmatrix} 1 & 0 \\ \frac{\partial \Upsilon}{\partial \theta} & \frac{\partial \Upsilon}{\partial p} \end{bmatrix}$  logo  $\det d(S \circ f) = \frac{\partial \Upsilon}{\partial p} \circ f \cdot \det df$ .

Ou seja, queremos  $\frac{\partial \Upsilon}{\partial p} \circ f = \frac{1}{\det df}$ . Basta então definir  $\Upsilon(\theta, p) = \int \frac{dp}{\det df \circ f^{-1}}$ .

É claro que para qualquer translação  $T_c$  a composta  $T_c \circ S \circ f$  preserva área. Logo, basta encontrar um valor  $c$  para o qual  $\int_{\gamma} (T_c \circ S \circ f)^* \alpha = \int_{\gamma} \alpha$  para uma curva  $\gamma$  homotopicamente não trivial.

Mas  $T_c^*(\alpha) = (p + c)d\theta$ , portanto

$$\int_{\gamma} (T_c \circ S \circ f)^* \alpha = \int_{\gamma} (S \circ f)^*(pd\theta) + c \int_{\gamma} (S \circ f)^* d\theta.$$

Como  $S$  é um difeomorfismo vertical,  $S^*(d\theta) = d\theta$  e então

$$\int_{\gamma} (T_c \circ S \circ f)^* \alpha = \int_{\gamma} (S \circ f)^* \alpha + c \int_{\gamma} f^* d\theta.$$

Basta então definir  $c = \frac{\int_{\gamma} (S \circ f)^*(\alpha) - \int_{\gamma} \alpha}{\int_{\gamma} f^*(d\theta)}$  para obter a condição

$$\int_{\gamma} (T_c \circ S \circ f)^* \alpha = \int_{\gamma} \alpha.$$

$\square$

*Demonstração.* (do teorema de Moser) Tomando um levantamento  $\tilde{f} = (\tilde{f}_1, \tilde{f}_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  de  $f$  provamos inicialmente que existem difeomorfismos verticais  $\tilde{R}$  e  $\tilde{Q}$  tais que  $\tilde{f} = \tilde{Q} \circ \tilde{F} \circ \tilde{R}$  onde  $\tilde{F}(x, y) = (x + y, y)$ . De fato, se  $\tilde{R}(x, y) = (x, \tilde{f}_1(x, y) - x)$ , como  $\frac{\partial \tilde{f}_1}{\partial y} > 0$  então  $\tilde{R}$  é um difeomorfismo. Além disso,

$$\tilde{F}(\tilde{R}(x, y)) = \tilde{F}(x, \tilde{f}_1(x, y) - x) = (\tilde{f}_1(x, y), \tilde{f}_1(x, y) - x).$$

Queremos encontrar  $\tilde{Q}(x, y) = (x, \tilde{Q}_2(x, y))$  tal que  $\tilde{Q} \circ \tilde{F} \circ \tilde{R} = \tilde{f}$ , isto é,  $\tilde{Q}_2(\tilde{f}_1(x, y), \tilde{f}_1(x, y) - x) = \tilde{f}_2(x, y)$ . É suficiente definir  $\tilde{Q}_2 = \tilde{f}_2 \circ \tilde{R}^{-1} \circ \tilde{F}^{-1}$ .

Observemos finalmente que

$$\tilde{R}(x + 1, y) = (x + 1, \tilde{f}_1(x + 1, y) - x - 1) = (x, \tilde{f}_1(x, y) - x) + (1, 0)$$

$$\tilde{F}(x + 1, y) = \tilde{F}(x, y) + (1, 0)$$

$$\tilde{Q}(x + 1, y) = (x + 1, \tilde{Q}_2(x + 1, y)) = (x, \tilde{Q}_2(x, y)) + (1, 0).$$

Logo  $\tilde{Q}, \tilde{F}, \tilde{R}$  projetam-se em aplicações  $Q, F$  e  $R$  tais que  $f = Q \circ F \circ R$ .

Consideremos famílias parametrizadas de difeomorfismos  $Q_t, F_t$  e  $R_t$ , dependendo  $C^\infty$  de  $t$ ,  $t \in [0, 1]$  tais que  $Q_1 = Q$ ,  $R_1 = R$ ,  $Q_0 = Id$ ,  $R_0 = Id$  e  $F_t(\theta, p) = (\theta + tp, p)$ .

Para  $\epsilon > 0$  suficientemente pequeno, fazendo uma reparametrização linear adequada em  $t$ , obtemos uma família  $H_t$  de difeomorfismos dependendo  $C^\infty$  de  $t$ , tal que  $H_{1-\epsilon} = Q \circ F \circ R = f$  e  $H_\epsilon = F_{2\epsilon}$ .

Para  $0 \leq t \leq 1$  definimos  $\Psi_t = F_t$  para  $0 \leq t \leq \epsilon$

$\Psi_t = F_{-\epsilon} \circ H_t$  para  $\epsilon \leq t \leq 1 - \epsilon$  e

$\Psi_t = F_{t-1} \circ f$  para  $1 - \epsilon \leq t \leq 1$

E estendemos para  $t \in \mathbb{R}$  pela fórmula  $\Psi_t = \Psi_{t-[t]} \circ \Psi_1^{[t]}$  onde  $[t]$  = parte inteira de  $t$  e  $\Psi_1^{[t]} = \Psi_1 \circ \dots \circ \Psi_1$  composição  $[t]$  vezes. Verifica-se então que

$$\Psi_{t+1} = \Psi_{t+1-[t+1]} \circ \Psi_1^{[t+1]} = \Psi_{t-[t]} \circ \Psi_1^{[t]+1} \text{ ou seja } \Psi_{t+1} = \Psi_t \circ \Psi_1.$$

Observe que  $X_t = \frac{\partial \Psi_t}{\partial t} \circ \Psi_t^{-1}$  é um campo possivelmente descontínuo para os valores  $t = \epsilon$  e  $t = 1 - \epsilon$ . Usando-se uma função



auxiliar, podemos modificar  $\Psi_t$  em uma vizinhança desses valores sem afetar a parte importante da interpolação, isto é, próximo aos valores inteiros de  $t$ .

Este é um detalhe importante que pode ser visto no artigo de Moser ([60]), onde se consegue diretamente obter a condição de convexidade.

Para podermos aplicar o Lema 5, é preciso obter uma família de difeomorfismos exatos. Para isso, usando o Lema 6, definimos uma família de difeomorfismos verticais  $S_t$  e de translações  $T_t$ , dependendo  $C^\infty$  de  $t$  tais que  $\varphi_t = T_t \circ S_t \circ \Psi_t$  é exato, com  $S_1 = Id$  e  $T_1 = Id$  pois  $\Psi_1 = f$  é exato, por hipótese.

Logo  $\varphi_{t+1} = T_{t+1} \circ S_{t+1} \circ \Psi_{t+1}$ .

Note que, de acordo com a definição do difeomorfismo vertical

$$\begin{aligned} S_{t+1}(\theta, p) &= (\theta, V_{t+1}(\theta, p)), \text{ com } V_{t+1} = \int \frac{dp}{\det d\Psi_{t+1} \circ \Psi_{t+1}^{-1}} = \\ &= \int \frac{dp}{\det d\Psi_t \circ \Psi_t^{-1} \det \Psi_1 \circ \Psi_1^{-1}} \\ &= \int \frac{dp}{\det d\Psi_t \circ \Psi_t^{-1}} \end{aligned}$$

pois  $\Psi_{t+1} = \Psi_t \circ \Psi_1$  e  $\det d\Psi_1 = \det f = 1$ . Portanto  $V_{t+1} = V_t$  e  $S_{t+1} = S_t \circ S_1$ ,  $S_1 = Id$ . Analogamente obtemos  $T_1 = Id$  e  $T_{t+1} = T_t = T_t \circ T_1$ , usando a definição da constante de translação  $c$ .

Portanto  $\varphi_{t+1} = T_t \circ S_t \circ \Psi_t \circ \Psi_1 = \varphi_t \circ T_t \circ S_1 \circ \Psi_1 = \varphi_t \circ \varphi_1$ .

Concluimos que o campo  $X_t = \frac{\partial \varphi_t}{\partial t} \circ \varphi_t^{-1}$  que gera a família  $\varphi_t$  é Hamiltoniano.

Resta provar que o Hamiltoniano  $H(\theta, p, t)$  satisfaz à condição de convexidade estrita nas fibras ou seja  $\frac{\partial^2 H}{\partial p^2}(\theta, p, t) > d$ .

Escrevendo  $\Psi_{t+\epsilon} \circ \Psi_t^{-1} = Q_{t+\epsilon} \circ F_{t+\epsilon} \circ R_{t+\epsilon} \circ R_t^{-1} \circ F_t^{-1} \circ Q_t^{-1}$ .

Como  $F_{t+\epsilon}(\theta, p) = (\theta + (t + \epsilon)p, p)$ ,  $Q_{t+\epsilon}$ ,  $Q_t$  e  $R_{t+\epsilon} \circ R_t^{-1}$  são difeomorfismos verticais, é fácil ver que, calculando a matriz jacobiana da composta, se escrevemos  $\Psi_{t+\epsilon} \circ \Psi_t^{-1} = (A_{t+\epsilon}, B_{t+\epsilon})$  então

existe uma constante  $C_1$  tal que  $C_1^{-1}\epsilon < \frac{\partial A_{t+\epsilon}}{\partial p} < C_1\epsilon$  (twist uniforme). Da mesma forma,

$$\varphi_{t+\epsilon} \circ \varphi_t^{-1} = T_{t+\epsilon} \circ S_{t+\epsilon} \circ \Psi_{t+\epsilon} \circ \Psi_t^{-1} \circ S_t^{-1} \circ T_t^{-1}, \quad T_{t+\epsilon} \text{ e } T_t$$

translações e  $S_{t+\epsilon}$  e  $S_t$ , verticais. Obtemos que existe uma constante  $d > 0$  tal que  $\epsilon d^{-1} < \frac{\partial G_{t+\epsilon}}{\partial p} < \epsilon d$  onde  $G_{t+\epsilon} = \pi_1 \circ \varphi_{t+\epsilon} \circ \varphi_t^{-1}$ , primeira coordenada de  $\varphi_{t+\epsilon} \circ \varphi_t^{-1}$ .

Como  $\frac{\partial G_t}{\partial p} = 0$  então  $d^{-1} < \frac{1}{\epsilon} \left[ \frac{\partial G_{t+\epsilon}}{\partial p} - \frac{\partial G_t}{\partial p} \right] < d$ , o que implica, tomando-se  $\epsilon \rightarrow 0$ , que  $d^{-1} < \frac{\partial}{\partial t} \left[ \frac{\partial G_t}{\partial p} \right] < d$  ou  $d^{-1} < \frac{\partial}{\partial p} \left[ \frac{\partial G_t}{\partial t} \right] < d$ .

Isto é, a primeira coordenada  $X_t^1$  do campo  $X_t = \frac{\partial \varphi_t}{\partial t} \circ \varphi_t^{-1}$  satisfaz  $d^{-1} < \frac{\partial X_t^1}{\partial p} < d$ . Mas o campo  $X_t$  é Hamiltoniano, logo  $X_t^1 = \frac{\partial H}{\partial p}(\theta, p, t)$ . Concluimos então que  $d^{-1} < \frac{\partial^2 H}{\partial p^2}(\theta, p, t) < d$ .  $\square$

A convexidade estrita nas fibras da função Hamiltoniana é uma condição bastante útil e comum, que permite-nos utilizar métodos variacionais para encontrar órbitas periódicas ou subconjuntos invariantes mais gerais. Esta é uma condição importante para a generalização desses métodos em dimensões mais altas.

## 1.9 Passando do cilindro para o anel

Os resultados sobre aplicações do tipo twist são válidos para um anel invariante limitado pelo gráfico de duas funções. Isto é interessante ao estudarmos, por exemplo, as aplicações do tipo bilhar.

Esta propriedade de extensão decorre do seguinte lema cuja prova encontra-se em [55]

### Lema 7. (Lema de Extensão)

Sejam  $\beta_-, \beta_+ : \mathbb{R} \rightarrow \mathbb{R}$  difeomorfismos de classe  $C^{r-1}$ ,  $r \geq 1$  que satisfazem  $\beta_- < \beta_+$  e  $\beta_-(x+1) = \beta_-(x) + 1$ .

Seja  $W = \{(x, x') \mid \beta_-(x) \leq x' \leq \beta_+(x)\}$ .

Se  $h : W \rightarrow \mathbb{R}$  é uma função de classe  $C^{r+1}$ , tal que  $h(x+1, x'+1) = h(x, x')$  e  $\partial_{12}h < 0$  (ou  $\partial_{12}h > 0$ ), então  $h$  possui uma extensão  $\hat{h}$  de classe  $C^{r+1}$  em  $\mathbb{R}^2$ , que satisfaz  $\hat{h}(x+1, x'+1) = \hat{h}(x, x')$ ,  $\partial_{12}\hat{h} < 0$  (ou  $\partial_{12}\hat{h} > 0$ )

*Demonstração.* A idéia é estender a função  $\partial_{12}h$  a todo o plano e integrar duas vezes a extensão, levando em conta as condições de fronteira.

Observe inicialmente que o conjunto  $W$  é invariante pela translação  $T(x, x') = (x + 1, x' + 1)$ , portanto o cilindro quociente  $\mathcal{C} = \mathbb{R}^2 \text{ mod}(T)$  contém a superfície com fronteira  $\hat{W} = W \text{ mod}(T)$  difeomorfa a um anel fechado.

Logo, existe um número  $\delta > 0$  tal que  $\partial_{12}h < -\delta$  em  $\hat{W}$ .

Seja  $\rho(x, x')$  uma extensão de classe  $C^{r-1}$  de  $\partial_{12}h$  ao  $\mathbb{R}^2$  satisfazendo  $\rho(x + 1, x' + 1) = \rho(x, x')$ ,  $\rho(x, x') = -\delta$  no subconjunto  $U = \{(x, x') \mid \beta_-(x) - 1 \leq x' \leq \beta_+(x) + 1\}$ .

Definimos  $\partial_{12}\hat{h}(u, u') = \rho(u, u')$  de modo que

$$\partial_1\hat{h}(u, x') = \int_0^{x'} \rho(u, u') du' + \tau(u).$$

A condição

$$\partial_1\hat{h}(u, \beta_-(u)) = \int_0^{\beta_-(u)} \rho(u, u') du' + \tau(u) = \partial_1h(u, \beta_-(u))$$

define a função  $\tau$  de modo que

$$\partial_1\hat{h}(u, x') = \int_{\beta_-(x)}^{x'} \rho(u, u') du' + \partial_1h(u, \beta_-(u)).$$

Observe que  $\rho|_W = \partial_{12}h$  implica que  $\partial_1\hat{h}(u, \beta_+(u)) = \partial_1h(u, \beta_+(u))$ .

Integrando em relação a  $u$  :

$$\hat{h}(x, x') = \int_0^x \left[ \int_{\beta_-(u)}^{x'} \rho(u, u') du' + \partial_1h(u, \beta_-(u)) \right] du + \gamma(x')$$

e usando novamente a condição de fronteira, juntamente com o fato de que  $\beta_-$  é um difeomorfismo, obtemos a função  $\gamma(x')$  e

$$h(x, x') = \int_{(\beta_-)^{-1}(x')}^x [\partial_1\hat{h}(u, x')] du + h((\beta_-)^{-1}(x'), x')$$

A condição  $\hat{h}(x, \beta_+(x)) = h(x, \beta_+(x))$  está também satisfeita.  $\square$

Desta forma, usando  $\hat{h}(x, x')$  como função geratriz, estendemos qualquer aplicação do tipo twist que preserva área no anel  $\mathcal{A}$  para todo o cilindro  $\mathcal{C}$ .

## Capítulo 2

# Órbitas periódicas e conjuntos de Aubry-Mather.

### 2.1 Órbitas periódicas minimizantes

Nesta seção seguiremos a notação e suporemos as hipóteses introduzidas na seção 1.3. Há três referências, entre muitas, na quais o leitor pode encontrar os resultados provados a seguir: [13] (Bangert), [55] (Forni-Mather) e [56] (Meiss). A última é a mais elementar e mais indicada para uma primeira leitura. A referência [13] (Bangert) é, do ponto de vista matemático, a padrão no assunto. É importante mencionar que a exposição abaixo é fortemente influenciada por [13] (principalmente) e [56], que por sua vez foram fortemente influenciadas pela exposição original de Aubry [12]. Os resultados principais desta seção são devidos a Aubry [12] e Mather [53].

Qualquer órbita periódica  $(x_k, y_k)_{k \in \mathbb{Z}}$  de uma aplicação do tipo twist  $F$  está associada a um par de inteiros  $m$  e  $n$ , tais que  $x_{k+n} = x_k + m$  para qualquer  $k \in \mathbb{Z}$ . Chamaremos tal órbita de órbita periódica tipo  $(m, n)$ . O inteiro positivo  $n$  é o período da órbita e  $m$  é o número de voltas, positivo ou negativo, que a órbita dá no

cilindro  $C$ . O número de rotação de uma órbita periódica é definido como a razão  $m/n$ . Mais geralmente, vale a seguinte:

**Definição 7.** *O número de rotação de uma órbita  $(x_k, y_k)_{k \in \mathbb{Z}}$  qualquer de  $F$  é definido, caso os limites abaixo existam e coincidam, como:*

$$\rho = \lim_{k \rightarrow \infty} \frac{x_k - x_0}{k} = \lim_{k \rightarrow -\infty} \frac{x_k - x_0}{k}$$

Caso o número de rotação de uma órbita exista, ele mede a velocidade angular média (ângulo/iteração) com que a órbita dá voltas no cilindro. Para uma configuração  $(x_k)_{k \in \mathbb{Z}}$ , vale uma definição análoga.

O objetivo principal desta seção é provar o seguinte teorema.

**Teorema 3.** *A aplicação  $F$  admite órbitas periódicas minimizantes com todos os possíveis números de rotação  $m/n$ .*

A prova deste teorema será consequência de uma série de resultados interessantes, que evidenciam diversas propriedades das órbitas minimizantes.

Seja  $W_{m,n} : \mathbb{R}^n \rightarrow \mathbb{R}$  a função  $W$  restrita a configurações finitas do seguinte modo:

$$W_{m,n}(x_0, \dots, x_{n-1}) = h(x_0, x_1) + h(x_1, x_2) + \dots + h(x_{n-1}, x_0 + m)$$

Note que a seqüência das componentes  $(x_k)_{k \in \mathbb{Z}}$  de uma órbita periódica tipo  $(m, n)$  é uma configuração estacionária de  $W$ , ou seja  $(x_k)_{k \in \mathbb{Z}}$  satisfaz as equações (1.12). Por sua vez a periodicidade da seqüência  $(x_k)_{k \in \mathbb{Z}}$  implica que esta é uma configuração estacionária de  $W$  se, e só se,  $(x_0, x_1, \dots, x_{n-1})$  é ponto crítico de  $W_{m,n}$ . Portanto, procurar órbitas periódicas tipo  $(m, n)$  de  $F$  é equivalente a procurar pontos críticos de  $W_{m,n}$ . Isto motiva o estudo da função  $W_{m,n}$ . Para entender alguns aspectos de  $W_{m,n}$  é conveniente fazer a seguinte mudança de variáveis  $(x_0, \dots, x_{n-1}) \rightarrow (s, \eta_1, \dots, \eta_{n-1})$ :

$$\begin{aligned} s &= x_0 \\ \eta_1 &= x_1 - x_0 \\ \eta_2 &= x_2 - x_1, \dots \\ \eta_{n-1} &= x_{n-1} - x_{n-2} \end{aligned}$$

ou

$$\begin{aligned} x_0 &= s \\ x_1 &= \eta_1 + s \\ x_2 &= \eta_1 + \eta_2 + s, \dots \\ x_{n-1} &= \eta_1 + \dots + \eta_{n-2} + \eta_{n-1} + s \end{aligned}$$

Nas novas variáveis vale:

$$\begin{aligned} W_{m,n}(s, \eta_1, \dots, \eta_{n-1}) &= h(s, \eta_1 + s) + h(\eta_1 + s, \eta_1 + \eta_2 + s) + \\ &h(\eta_1 + \eta_2 + s, \eta_1 + \eta_2 + \eta_3 + s) \dots \\ &+ h(\eta_1 \dots + \eta_{n-1} + s, m + s) \end{aligned}$$

Desta expressão e da propriedade,  $h(x+1, x'+1) = h(x, x')$ , fica evidente que  $W_{m,n}(s, \eta)$  é 1-periódica na variável  $s$ , ou seja,  $W_{m,n}(s+1, \eta) = W_{m,n}(s, \eta)$ . Portanto, se  $s \rightarrow \theta = s \pmod{1}$ ,  $s \in \mathbb{R}$ ,  $\theta \in S^1$ , é uma aplicação de recobrimento do círculo, então a função  $\hat{W}_{mn}(\theta, \eta) \stackrel{\text{def}}{=} W_{mn}(s, \eta)$  está definida sobre o cilindro

$$C_{mn} \stackrel{\text{def}}{=} \{(\theta, \eta) : \theta \in S^1, \eta \in \mathbb{R}^{n-1}\}$$

Agora, aplicando a proposição 3, na seção 1.3, a cada parcela da função  $W_{mn}(s, \eta)$  obtém-se a seguinte desigualdade:

$$\hat{W}_{mn}(\theta, \eta) = W_{m,n}(s, \eta) \geq \frac{c}{2} [\eta_1^2 + \eta_2^2 + \dots + \eta_{n-1}^2 + (\eta_1 + \eta_2 + \dots + \eta_{n-1})^2] \quad (2.1)$$

Isto implica que sobre cada fibra  $\theta = \text{constante}$ , o gráfico da função  $\hat{W}_{mn}(\theta, \cdot)$  está acima do parabolóide definido no lado direito da equação (2.1). Isso e a continuidade de  $\hat{W}_{mn}$  implicam o seguinte lema (faça a prova).

**Lema 8.** *A função  $\hat{W}_{mn}$ , ou equivalentemente  $W_{mn}$ , possui um ponto de mínimo global, que será denotado por  $p$ .*

A diferenciabilidade de  $W_{mn}$  implica que  $p$  é um ponto crítico desta função. Portanto, as componentes  $(x_0, \dots, x_{n-1})$  de  $p$  correspondem a uma órbita periódica tipo  $(m, n)$  de  $F$ . É possível encontrar um segundo ponto crítico de  $\hat{W}_{mn}$ , ou de  $W_{mn}$ , usando um argumento “mini-max”, da seguinte forma.

**Lema 9 (Mini-Max).** *A função  $\hat{W}_{mn}$ , ou equivalentemente  $W_{mn}$ , possui um segundo ponto crítico distinto do ponto  $p$  do lema 8.*

*Demonstração.* Seja  $A$  o conjunto de todas as curvas  $\gamma : S^1 \rightarrow C_{mn}$ : diferenciáveis, homotópicas à curva  $\psi \rightarrow (\theta = \psi, \eta = 0)$ ,  $\psi \in S^1$ , e que passam pelo ponto de mínimo  $p$ . Seja  $\alpha$  o valor mini-max

$$\alpha = \inf_{\gamma \in A} \max_{\psi \in S^1} \hat{W}_{mn} \circ \gamma(\psi)$$

Primeiro consideremos o caso  $\alpha = \hat{W}_{mn}(p)$ . Seja  $\bar{\theta}$  um valor qualquer de  $\theta$ , distinto daquele que corresponde à coordenada  $\theta$  de  $p$ . Note que qualquer curva  $\gamma$  de  $A$  é homotopicamente não trivial, portanto possui um ponto cuja coordenada  $\theta$  vale  $\bar{\theta}$ . Então, da definição de  $\alpha$ , existe uma sequência de curvas  $\gamma^k$ ,  $k = 1, 2, \dots$ , e uma correspondente sequência de pontos  $(\bar{\theta}, \eta^k) \in \gamma^k$  tal que  $\lim_{k \rightarrow \infty} \hat{W}_{mn}(\bar{\theta}, \eta^k) = \alpha$ . Devido à desigualdade (2.1) a sequência  $(\bar{\theta}, \eta^k)_{k \in \mathbb{N}}$  é limitada e portanto possui uma subsequência convergente a um ponto  $(\bar{\theta}, \bar{\eta})$ . Da continuidade de  $\hat{W}_{mn}$  concluí-se que  $\hat{W}_{mn}(\bar{\theta}, \bar{\eta}) = \alpha = \hat{W}_{mn}(p)$ , ou seja, existe um segundo ponto distinto de  $p$  que minimiza  $\hat{W}_{mn}$  e, portanto, um segundo ponto crítico. Como a escolha de  $\bar{\theta}$  é arbitrária, concluí-se que para todo  $\theta \in S^1$ , a função  $\hat{W}_{mn}$  possui um ponto crítico  $(\theta, \bar{\eta}(\theta))$ . Ou seja, no caso particular em que  $\alpha$  coincide com o valor mínimo da função  $\hat{W}_{mn}$ , a aplicação  $F$  possui um contínuo de órbitas periódicas do tipo  $(m, n)$ .

Agora consideremos o caso  $\alpha > \hat{W}_{mn}(p)$ . Seja  $\epsilon > 0$  suficientemente pequeno tal que  $\alpha - \epsilon > \hat{W}_{mn}(p)$ , e considere o conjunto  $B \stackrel{\text{def}}{=} \{(\theta, \eta) \in C_{mn} : \alpha - \epsilon \leq \hat{W}_{mn}(\theta, \eta) \leq \alpha + \epsilon\}$ . Mostraremos que, se  $\hat{W}_{mn}$  não possui nenhum ponto crítico em  $B$ , então  $\alpha$  não pode ser o valor mini-max. Isso prova o lema. Portanto, suponha que  $\hat{W}_{mn}$  não possui ponto crítico em  $B$ . Se  $z = (\theta, \eta)$  denota um ponto de  $C_{mn}$ , considere o campo de vetores gradiente  $\dot{z} = -\nabla \hat{W}_{mn}(z)$ , onde  $\nabla = (\partial_\theta, \partial_{\eta_1}, \dots, \partial_{\eta_{n-1}})$  é o gradiente usual do  $R^n$ . Seja  $\phi_t(z)$ ,  $t \geq 0$ , a curva integral deste campo de vetores que, para  $t = 0$ , passa por  $z$ . Sobre esta curva vale

$$\frac{d}{dt} \hat{W}_{mn} \circ \phi_t(z) = -\|\nabla \hat{W}_{mn}(\phi_t(z))\|^2$$

Devido à desigualdade (2.1) o conjunto  $B$  é limitado e, uma vez que é fechado, é compacto. Como  $\hat{W}_{mn}$  não possui pontos críticos em  $B$ ,

então existe  $k > 0$  tal que  $k = \min_{z \in B} \|\nabla \hat{W}_{mn}(z)\|^2$ . Agora, seja  $z$  um ponto qualquer em  $B$  e  $T(z) > 0$  o valor máximo de  $t$  tal que  $\phi_t(z) \in B$  para  $0 \leq t \leq T(z)$  (a princípio  $T(z)$  pode ser infinito). Para  $0 \leq t \leq T(z)$  vale

$$\begin{aligned} \frac{d}{dt} \hat{W}_{mn} \circ \phi_t(z) &= -\|\nabla \hat{W}_{mn}(\phi_t(z))\|^2 \leq -k \implies \\ \hat{W}_{mn}(z) - \hat{W}_{mn} \circ \phi_t(z) &\geq kt \end{aligned}$$

Como  $\hat{W}_{mn}(z) - \hat{W}_{mn} \circ \phi_t(z) \leq 2\epsilon$  para  $z$  e  $\phi_t(z)$  em  $B$ , tem-se que  $2\epsilon \geq kt$  o que implica  $T(z) \leq 2\epsilon/k$ . Ou seja, um ponto qualquer em  $B$ , quando carregado pelo fluxo  $\phi_t$ , permanece em  $B$  no máximo pelo tempo  $2\epsilon/k$ . A definição de  $\alpha$  implica a existência de uma curva  $\gamma \in A$  tal que  $\hat{W}_{mn}|_\gamma < \alpha + \epsilon$ . Considere a família a um parâmetro de curvas  $\gamma_t = \phi_t \circ \gamma$ ,  $t \in [0, 3\epsilon/k]$ . Qualquer curva nesta família pertence a  $A$  (lembre-se que  $p$  é um ponto crítico de  $\hat{W}_{mn}$  e portanto não se move sobre a ação de  $\phi_t$ ). Mas  $\hat{W}_{mn}|_{\gamma_t} < \alpha - \epsilon$  se  $t > 2\epsilon/k$ , o que contradiz a definição de  $\alpha$ . Logo  $B$  tem que conter um ponto crítico de  $\hat{W}_{mn}$ . Como o valor de  $\epsilon > 0$  utilizado na construção acima pode ser arbitrariamente pequeno, conclui-se que o valor de  $\hat{W}_{mn}$  no novo ponto crítico é  $\alpha$ . □

Nosso próximo lema (“lema de cruzamento de Aubry”) é um resultado fundamental no contexto das órbitas minimizantes (ver definições 1 e 2 na seção 1.2). Ele afirma que duas configurações minimais podem se “cruzar” no máximo uma vez. Antes de apresentá-lo, é preciso definir o termo cruzar. Sejam  $a$  e  $b$ ,  $a < b$ , dois inteiros, e seja  $I = \{a, a+1, \dots, b\}$ . A partir de um segmento qualquer  $(x_k)_{k \in I}$  define-se uma função  $t \rightarrow x_t$ , com  $t \in [a, b]$ , tal que  $x_t = x_k$  para  $t = k$ , e  $x_t$  é linear para  $k \leq t \leq k+1$ ,  $k \in I$ . O gráfico da função  $t \rightarrow x_t$  será chamado de gráfico do segmento  $(x_k)_{k \in I}$ . Substituindo-se  $I$  por  $\mathbb{Z}$  define-se o gráfico de uma configuração  $(x_k)_{k \in \mathbb{Z}}$  do mesmo modo. Dizemos que dois segmentos  $(x_k)_{k \in I}$  e  $(z_k)_{k \in I}$  se cruzam se seus gráficos se cruzam, ou seja, se existem  $t_1, t_2, t_3$  e  $t_4$ , com  $a \leq t_1 < t_2 \leq t_3 < t_4 \leq b$ , tal que  $x_t < z_t$  para  $t \in [t_1, t_2)$ ,  $x_t = z_t$  para  $t \in [t_2, t_3]$ , e  $x_t > z_t$  para  $t \in (t_3, t_4]$ . Dizemos que duas configurações se cruzam se elas contém segmentos que se cruzam. No lado esquerdo das figuras 2.1, 2.2 e 2.3 são mostrados gráficos de configurações que se cruzam.



**Lema 10.** *Dois configurações (ou dois segmentos) minimais (ver definições 1 e 2 na seção 1.2) se cruzam no máximo uma vez.*

*Demonstração.* Para provar o lema basta considerar o caso de cruzamento de dois segmentos. Sejam  $(x_k)_{k \in I}$  e  $(\xi_k)_{k \in I}$  dois segmentos minimais,  $I = \{a, \dots, b\}$ . Primeiramente mostraremos que, ou estes segmentos são iguais, ou eles não tem dois pontos consecutivos iguais. De fato, se, por exemplo,  $x_a = \xi_a$  e  $x_{a+1} = \xi_{a+1}$ , então da minimalidade dos segmentos decorre que as órbitas de  $F$  associadas a estes segmentos possuem a mesma coordenada  $y$  em  $k = a$  (ver figura 1.7), e portanto coincidem para todo  $k \in I$ . Disso, segue que os cruzamentos dos segmentos podem ser de dois tipos, como ilustrados no lado esquerdo das figuras 2.1 e 2.2, respectivamente. Suponha que os dois

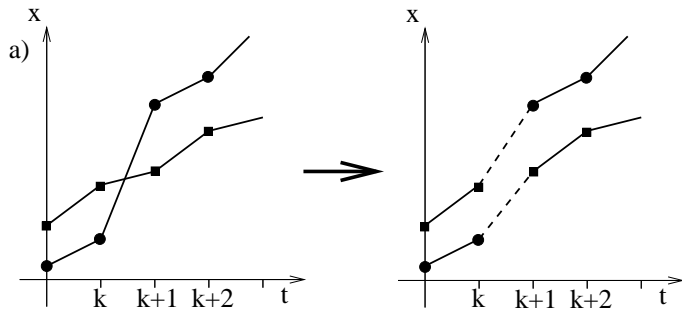


Figura 2.1: Eliminação de cruzamento.

segmentos minimais se cruzam duas vezes, como no lado esquerdo da figura 2.3. Nesta figura  $(x_k)_{k \in I}$  é representada por círculos e  $(\xi_k)_{k \in I}$  por quadrados. Suponha que pelo menos um dos cruzamentos seja do tipo da figura 2.1, por exemplo o que ocorre entre  $k = i$  e  $k = i + 1$ , e que o outro, em  $k = j$ , seja do tipo da figura 2.2. Os casos em que ambos os cruzamentos são do tipo da figura 2.1, ou em que o primeiro é do tipo da figura 2.2 e o segundo do tipo da figura 2.1 podem ser tratados da mesma maneira. O caso em que os dois cruzamentos são do tipo da figura 2.2 será tratado posteriormente. Então, alterando os subsegmentos de  $(x_k)_{k \in I}$  e  $(\xi_k)_{k \in I}$  entre  $k = i$  e  $k = j$ , como ilustrado no lado direito da figura 2.3 (ou no lado direito das figuras 2.1 e

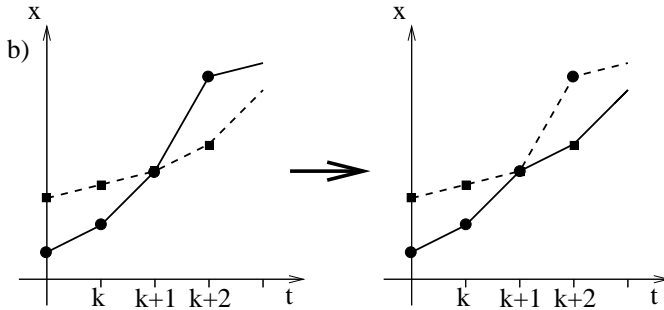


Figura 2.2: Eliminação de cruzamento.

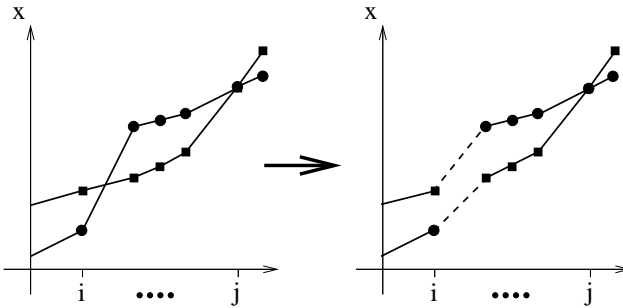


Figura 2.3: Eliminação de cruzamentos.

2.2), obtém-se dois novos segmentos, denotados respectivamente por  $(w_k)_{k \in I}$  e  $(\eta_k)_{k \in I}$ , dados por:

$$\begin{aligned} w_k &= x_k, & \eta_k &= \xi_k & \text{se} & & k \leq i \\ w_k &= \xi_k, & \eta_k &= x_k & \text{se} & & i < k < j \\ w_k &= x_k, & \eta_k &= \xi_k & \text{se} & & k \geq j \end{aligned}$$

Denotando a restrição de  $W$  ao segmento  $\{x_i, \dots, x_j\}$  por  $W_{ij}^-(x) \stackrel{\text{def}}{=} W(x_i, \dots, x_j)$ , obtém-se, da definição de configuração minimal, as seguintes desigualdades:

$$W_{ij}^-(x) \leq W_{ij}^-(w) \quad \text{e} \quad W_{ij}^-(\xi) \leq W_{ij}^-(\eta). \quad (2.2)$$

Usando a desigualdade da proposição 2, seção 1.3, e que  $W = \sum_k h(x_k, x_{k+1})$ , obtém-se:

$$\begin{aligned} W_{\overline{ij}}(x) + W_{\overline{ij}}(\xi) - [W_{\overline{ij}}(w) + W_{\overline{ij}}(\eta)] &= h(x_i, x_{i+1}) + h(\xi_i, \xi_{i+1}) \\ &\quad - h(x_i, \xi_{i+1}) - h(\xi_i, x_{i+1}) \geq c(x_i - \xi_i)(\xi_{i+1} - x_{i+1}) \end{aligned}$$

Mas o fato dos gráficos se cruzarem entre  $i$  e  $i + 1$  implica que  $(x_i - \xi_i)(\xi_{i+1} - x_{i+1}) > 0$ , o que contradiz as desigualdades (2.2) e portanto a minimalidade dos segmentos  $(x_k)_{k \in I}$  e  $(\xi_k)_{k \in I}$ . O ponto crucial neste argumento foi, que ao mudarmos os segmentos do lado esquerdo da figura 2.1, para as do lado direito, decrescemos o valor da soma das suas ações  $W$ . Ou seja, eliminar cruzamentos diminui a soma das ações  $W$ .

Para concluir a prova do lema é preciso considerar o caso em que os dois cruzamentos são do tipo da figura 2.2, ou seja  $x_{i+1} = \xi_{i+1}$  e  $x_{j-1} = \xi_{j-1}$ . Neste caso, constrói-se segmentos minimais  $(w_k)_{k \in I}$  e  $(\eta_k)_{k \in I}$  de modo semelhante, mas ocorre que  $W_{\overline{ij}}(x) = W_{\overline{ij}}(w)$  e  $W_{\overline{ij}}(\xi) = W_{\overline{ij}}(\eta)$ , e o argumento acima não se aplica. No entanto, neste caso, tanto  $(x_k)_{k \in [i, j]}$  quanto  $(w_k)_{k \in [i, j]}$  são segmentos minimais que possuem dois elementos consecutivos em comum, a saber  $x_i = w_i$  e  $x_{i+1} = w_{i+1}$ . Mas, como visto no começo da prova, isto implica que  $x_k = w_k = \xi_k$  para todo  $k$ . Isto viola a hipótese de cruzamento, que requer que  $(x_k)_{k \in I}$  e  $(\xi_k)_{k \in I}$  sejam distintas.  $\square$

O lema 10 possui uma variante que será necessária futuramente. Uma vez que sua prova é praticamente idêntica à do lema 10 ela será enunciada aqui.

**Lema 11.** *Sejam  $(x_k)_{k \in I}$  e  $(\xi_k)_{k \in I}$ ,  $I = \{a, \dots, b\}$ , duas configurações minimais tais que  $x_a \neq \xi_a$  e  $x_b = \xi_b$  (ou  $x_a = \xi_a$  e  $x_b \neq \xi_b$ ). Então  $(x_k)_{k \in [a, b-1]}$  e  $(\xi_k)_{k \in [a, b-1]}$  (ou  $(x_k)_{k \in [a+1, b]}$  e  $(\xi_k)_{k \in [a+1, b]}$ ) não se cruzam.*

*Demonstração.* Considere o caso  $x_a \neq \xi_a$  e  $x_b = \xi_b$ , o outro é análogo. Suponha que ocorra um cruzamento entre  $(x_k)_{k \in [a, b-1]}$  e  $(\xi_k)_{k \in [a, b-1]}$  como no lado esquerdo da figura 2.1. Então o mesmo argumento usado na prova do lema 10, e ilustrado no lado direito da figura 2.1, mostra que tal cruzamento não pode ocorrer. Agora, suponha que o cruzamento é como no lado esquerdo da figura 2.2, ou seja, existe  $i$ ,  $a < i < b - 1$ , tal que  $x_i = \xi_i$ . Então, os segmentos

$(x_k)_{k \in [i, b]}$  e  $(\xi_k)_{k \in [i, b]}$  possuem as mesmas ações  $W$ . Isso implica que tanto  $(x_k)_{k \in [i-1, b]}$ , quanto  $\{x_{i-1}, x_i, \xi_{i+1}, \dots, \xi_b\}$  são segmentos minimais, distintos e que possuem dois pontos consecutivos iguais. Mas isto é impossível.  $\square$

O lema 10 possui inúmeras consequências importantes. Uma delas é o seguinte.

**Corolário 2.** *Duas configurações minimais periódicas de tipo  $(m, n)$  nunca se cruzam.*

De fato se elas se cruzassem uma vez, então, devido a periodicidade, elas se cruzariam infinitas vezes, o que violaria o lema 10. Um outro corolário, que de fato implica o corolário acima, é o seguinte.

**Corolário 3.** *Duas configurações periódicas tipo  $(m, n)$ , que correspondem a mínimos globais da função  $W_{mn}$  (ver lema 8), nunca se cruzam.*

De fato, se elas se cruzassem uma vez, por exemplo entre  $k = 0$  e  $k = 1$ , elas teriam que possuir um segundo cruzamento no intervalo  $[1, n]$ . Mas isto violaria o lema 10, pois restrito a estes intervalos as configurações que minimizam  $W_{mn}$  são segmentos minimais. O corolário 3 possui a seguinte consequência.

**Lema 12.** *Considere a função  $W_{m'n'}$ , onde  $m' = im$ ,  $n' = in$ , com  $i$  inteiro positivo e  $(m, n)$  primos entre si. Seja  $p$  um ponto de mínimo global de  $W_{m'n'}$ , que existe pelo lema (8). Então,  $p$  corresponde a uma configuração periódica tipo  $(m, n)$  que minimiza  $W_{mn}$ .*

*Demonstração.* Seja  $(x_k)_{k \in \mathbb{Z}}$  uma configuração que corresponde a um mínimo global de  $W_{m'n'}$ . Então  $(\xi_k)_{k \in \mathbb{Z}}$ , onde  $\xi_k = x_{k+n} - m$ , também corresponde a um mínimo global de  $W_{m'n'}$ . Se  $(x_k)_{k \in \mathbb{Z}}$  é igual a  $(\xi_k)_{k \in \mathbb{Z}}$ , então  $(x_k)_{k \in \mathbb{Z}}$  corresponde a uma órbita periódica tipo  $(m, n)$  e a prova está concluída. Se  $(x_k)_{k \in \mathbb{Z}}$  é diferente de  $(\xi_k)_{k \in \mathbb{Z}}$  então considere a configuração  $w_k = \xi_k - x_k$ . Note que,  $w_{k+n'} = w_k$  e que a média de  $w_k$  é zero, pois

$$\sum_{k=1}^{n'} w_k = \sum_{k=1}^{in} (x_{k+n} - m) - \sum_{k=1}^{in} x_k = -inm + \sum_{k=1}^n (x_{k+in} - x_k) = 0.$$

Portanto,  $w_k$  muda de sinal pelo menos duas vezes, o que implica que  $(x_k)_{k \in \mathbb{Z}}$  e  $(\xi_k)_{k \in \mathbb{Z}}$  se cruzam, o que viola o corolário 3. Ou seja,  $(x_k)_{k \in \mathbb{Z}}$  é do tipo  $(m, n)$ . Agora, se  $(\eta_k)_{k \in \mathbb{Z}}$  é uma configuração qualquer do tipo  $(m, n)$ , então  $W_{m'n'}(\eta) = iW_{mn}(\eta)$ . Portanto, se  $(x_k)_{k \in \mathbb{Z}}$  minimiza  $W_{m'n'}$  então ela minimiza  $W_{mn}$ .  $\square$

Finalmente, usando os lemas 8 e 12 mostraremos o teorema 3. Do lema 8 segue a existência de uma órbita periódica associada a uma configuração  $(x_k)_{k \in \mathbb{Z}}$  que minimiza  $W_{mn}$ , com  $m$  e  $n$  primos entre si. Vamos mostrar que esta órbita é de fato uma órbita minimizante. Suponha o contrário. Então existem inteiros  $a$  e  $b$ ,  $b > a$ , e um segmento  $\xi_a, \xi_{a+1}, \dots, \xi_b$ , com  $x_a = \xi_a$ ,  $x_b = \xi_b$ , tal que  $W_{ab}^-(x) > W_{ab}^-(\xi)$ . Seja  $i$  o menor inteiro tal que  $i \cdot n > b - a$ , e considere a configuração periódica  $(w_k)_{k \in \mathbb{Z}}$  que satisfaz  $w_k = \xi_k$  se  $a \leq k \leq b$ , e  $w_k = x_k$  se  $b < k \leq a + in$ . Note que  $w_{a+in} = x_{a+in} = x_a + im = w_a + im$ , ou seja  $(w_k)_{k \in \mathbb{Z}}$  é uma configuração tipo  $(im, in)$ . Finalmente, a desigualdade  $W_{ab}^-(x) > W_{ab}^-(\xi)$  implica que  $W_{im,in}^-(x) > W_{im,in}^-(w)$ . Isso, no entanto, é impossível, pois o lema 12 implica que as configurações minimais de  $W_{im,in}$  são precisamente as configurações minimais de  $W_{mn}$ , portanto  $(x_k)_{k \in \mathbb{Z}}$  é configuração minimal de  $W_{(im,in)}$ . Isso conclui a prova do teorema 3.

## 2.2 Monotonicidade e órbita minimizante

Nesta seção veremos a força do corolário 2, que essencialmente implica na existência de certas órbitas notáveis das aplicações do tipo twist. Antes de iniciar tal discussão precisamos apresentar a seguinte definição.

**Definição 8 (Configuração (Órbita) monótona).** *Uma configuração  $(x_k)_{k \in \mathbb{Z}}$  é dita monótona se  $x_i < x_{i+j+k}$  implica  $x_l < x_{l+j+k}$ , para quaisquer inteiros  $i, j, k, l$ . Outra maneira de expressar isto é dizer que, duas “translações” inteiras do gráfico de uma configuração monótona ou nunca se intersectam ou coincidem. Uma órbita é dita monótona se sua configuração associada é monótona.*

Abaixo, usaremos a seguinte caracterização alternativa de monotonicidade. Uma configuração  $(x_k)_{k \in \mathbb{Z}}$  é monótona se, e só se, para

quaisquer inteiros  $i, j$  e  $k$  vale:

$$x_i < x_j + k \implies x_{i+1} < x_{j+1} + k \quad (2.3)$$

Uma generalização desta definição é a seguinte.

**Definição 9 (Conjunto invariante monótono).** *Seja  $F$  uma aplicação de tipo twist como na seção 1.3. Um conjunto  $S$ , invariante por  $F$ , é dito monótono, quando ele possui a seguinte propriedade: se  $(x, y)$  e  $(\bar{x}, \bar{y})$  são dois pontos de  $S$  tal que  $x < \bar{x}$  então  $F_1(x, y) < F_1(\bar{x}, \bar{y})$ . Ou seja, a ordem da coordenada  $x$  dos pontos de  $S$  é mantida pela ação de  $F$ .*

Note que se  $(x_k)_{k \in \mathbb{Z}}$  é uma configuração minimal periódica tipo  $(m, n)$ , então também o é  $(\xi_k)_{k \in \mathbb{Z}}$ , onde  $\xi_k = x_{k+a} + b$ , com  $a$  e  $b$  inteiros quaisquer. Logo os gráficos de  $(x_k)_{k \in \mathbb{Z}}$  e  $(\xi_k)_{k \in \mathbb{Z}}$  nunca se cruzam, pelo corolário 2. Ou seja, vale a seguinte proposição.

**Proposição 4.** *Uma configuração (ou órbita) minimal periódica tipo  $(m, n)$  é monótona.*

Uma consequência fundamental do corolário 2 é o seguinte lema, que diz que, uma órbita minimizante periódica tipo  $(m, n)$  possui seu gráfico contido em uma estreita faixa como ilustrado na figura 2.4.

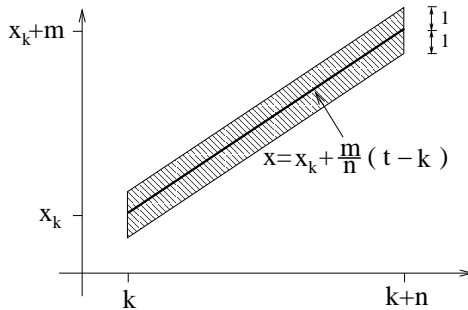


Figura 2.4: Região de confinamento de uma configuração periódica minimal.

**Lema 13.** *Se  $(x_k)_{k \in \mathbb{Z}}$  é uma configuração minimal periódica tipo  $(m, n)$ , então para todo  $k \in \mathbb{Z}$  vale*

$$\left| x_k - x_0 - k \frac{m}{n} \right| < 1$$

*Demonstração.* Primeiramente provaremos que, se  $x_k - x_0 = \alpha$  então para todo  $i \in \mathbb{Z}$  vale

$$\alpha - 1 < x_{ik} - x_{(i-1)k} < \alpha + 1 \quad (2.4)$$

Dado  $i$ , existe um inteiro  $b$  tal que a configuração  $(\xi_j)_{j \in \mathbb{Z}}$ , dada por  $\xi_j = x_{j+(i-1)k} - b$ , é minimal e satisfaz  $x_0 < \xi_0 < x_0 + 1$  ou, equivalentemente,

$$-1 < x_0 - \xi_0 < 0 \quad (2.5)$$

Logo, pelo corolário 2, o gráfico de  $(\xi_j)_{j \in \mathbb{Z}}$  fica entre o gráfico de  $(x_j)_{j \in \mathbb{Z}}$  e  $(x_j + 1)_{j \in \mathbb{Z}}$ , ver figura 2.5. Mas isso implica que  $x_k <$

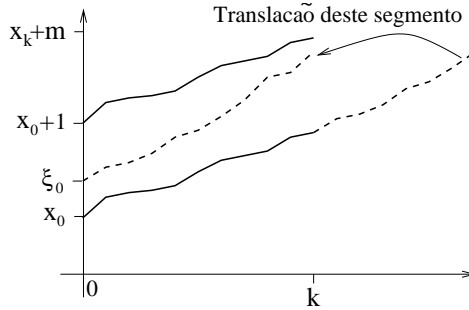


Figura 2.5: Construção usada na prova do lema 13.

$\xi_k < x_k + 1$ , o que implica  $x_k - x_0 < \xi_k - x_0 < x_k + 1 - x_0$ , ou seja,  $\alpha < \xi_k - x_0 < \alpha + 1$ . Somando esta desigualdade à desigualdade (2.5), obtém-se  $\alpha - 1 < \xi_k - \xi_0 < \alpha + 1$ , que é equivalente a (2.4).

Agora, contrariando a tese do lema, suponha que  $x_k - x_0 > k \frac{m}{n} + 1$  (o caso  $x_k - x_0 < k \frac{m}{n} - 1$  é análogo). Então da desigualdade (2.4) segue que para qualquer  $i$  vale  $km/n < x_{ik} - x_{(i-1)k}$ . Portanto,

$$\begin{aligned} km = x_{kn} - x_0 &= (x_{kn} - x_{(n-1)k}) + (x_{(n-1)k} - x_{(n-2)k}) \cdots \\ &\quad + (x_k - x_0) > n \frac{km}{n} = km \end{aligned}$$

o que é uma contradição.  $\square$

Agora, enunciaremos o principal resultado desta seção.

**Teorema 4.** *Dado qualquer número de rotação irracional  $\rho$ , existe uma configuração  $(\bar{x}_k)_{k \in \mathbb{Z}}$  que possui este número de rotação, que é minimal e monótona.*

*Demonstração.* Primeiramente, dado  $\epsilon_1 > 0$  qualquer, seja  $(\rho_i)_{i>0}$  uma seqüência de números racionais que converge para  $\rho$ , com  $|\rho - \rho_i| < \epsilon_1$ , para todo  $i > 0$ . Para cada  $\rho_i$  seja  $(x_k^i)_{k \in \mathbb{Z}}$  uma configuração minimal periódica com número de rotação  $\rho_i$  e tal que  $x_0^i \in [0, 1]$ . Tal configuração existe pelo teorema 3. Da compacidade de  $[0, 1]$  segue que existe uma subsequência  $(x_0^{i_p})_{p>0}$  de  $(x_0^i)_{i>0}$ , tal que  $x_0^{i_p} \rightarrow \bar{x}_0$  quando  $p \rightarrow \infty$ . Portanto, dado  $\epsilon_2 > 0$  existe um inteiro  $J$  tal que  $|x_0^{i_p} - \bar{x}_0| < \epsilon_2$  se  $p > J$ . Para um  $j > 0$  qualquer seja  $p = J + j$  e seja  $(x_k^{0,j})_{k \in \mathbb{Z}}$  uma nova notação para a seqüência  $(x_k^{i_p})_{k \in \mathbb{Z}}$ . Note que  $x_0^{0,j} \rightarrow \bar{x}_0 \in [0, 1]$  quando  $j \rightarrow \infty$  e  $|x_0^{0,j} - \bar{x}_0| < \epsilon_2$  para todo  $j > 0$ . Do lema 13 segue que para qualquer  $j > 0$  a seqüência  $(x_k^{0,j})_{k \in \mathbb{Z}}$  satisfaz

$$\begin{aligned}
 |x_k^{0,j} - \bar{x}_0 - k\rho| &= |x_k^{0,j} - x_0^{0,j} - k\rho + x_0^{0,j} - \bar{x}_0| \\
 &\leq |x_k^{0,j} - x_0^{0,j} - k\rho| + |x_0^{0,j} - \bar{x}_0| \\
 &< |x_k^{0,j} - x_0^{0,j} - k\rho_{i_p}| + |k\rho_{i_p} - k\rho| + \epsilon_2 \\
 &< 1 + |k|\epsilon_1 + \epsilon_2
 \end{aligned} \tag{2.6}$$

Ou seja, todos os pontos das seqüências  $(x_k^{0,j})_{k \in \mathbb{Z}}$ , para  $j > 0$ , estão contidos na região ilustrada na figura 2.6. Portanto, para qualquer  $k$  fixo, a seqüência  $(x_k^{0,j})_{j>0}$  está contida em um intervalo compacto. Em particular, isto vale para a seqüência  $(x_1^{0,j})_{j>0}$ , o que implica a existência de uma subsequência  $(x_1^{0,j_i})_{i>0}$  que converge para um  $\bar{x}_1$ . Denotemos a seqüência  $(x_k^{0,j_i})_{k \in \mathbb{Z}}$  por  $(x_k^{1,i})_{k \in \mathbb{Z}}$ . Aplicando o mesmo raciocínio, a partir de  $(x_{-1}^{1,i})_{i>0}$  obtém-se uma subsequência  $(x_{-1}^{1,i_j})_{j>0}$  que converge para  $\bar{x}_{-1}$ . Denotemos a seqüência  $(x_k^{1,i_j})_{k \in \mathbb{Z}}$  por  $(x_k^{-1,j})_{k \in \mathbb{Z}}$ . Repetindo tal operação constrói-se sucessivamente as seqüências de configurações  $(x_k^{2,i})_{k \in \mathbb{Z}}$ ,  $(x_k^{-2,i})_{k \in \mathbb{Z}}$ ,  $(x_k^{3,i})_{k \in \mathbb{Z}}$ , etc... Finalmente, a seqüência “diagonal” de configurações minimais

$$(x_k^{0,0})_{k \in \mathbb{Z}}, (x_k^{1,1})_{k \in \mathbb{Z}}, (x_k^{-1,2})_{k \in \mathbb{Z}}, (x_k^{2,3})_{k \in \mathbb{Z}}, (x_k^{-2,4})_{k \in \mathbb{Z}}, \dots$$



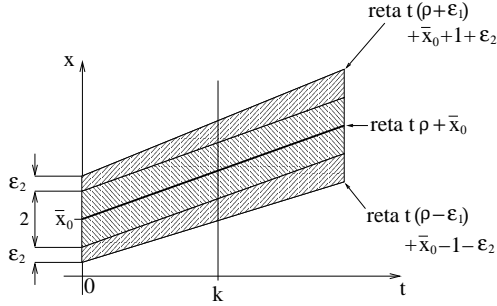


Figura 2.6: Construção usada na prova do teorema 4.

é tal que para qualquer  $k$  dado vale:

$$x_k^{0,0}, x_k^{1,1}, x_k^{-1,2}, x_k^{2,3}, x_k^{-2,4}, \dots \longrightarrow \bar{x}_k$$

A configuração limite  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é aquela que estávamos procurando. O método usado para a obtenção da mesma é o chamado “método diagonal” de Cantor, comumente usado em análise, como por exemplo na prova do teorema de Arzelá-Ascoli (ver [14]) ou do teorema de Tikhonov (ver [41]). Um subproduto imediato desta construção é que, para qualquer  $k$ , vale a desigualdade (2.6) para o valor limite  $\bar{x}_k$ , ou seja,

$$|\bar{x}_k - \bar{x}_0 - k\rho| \leq |k|\epsilon_1 + 1 + \epsilon_2$$

Como  $\epsilon_1$  e  $\epsilon_2$  podem ser escolhidos tão pequenos quanto o desejado, segue que, para qualquer  $k$ , vale:

$$|\bar{x}_k - \bar{x}_0 - k\rho| \leq 1$$

Isso implica que o número de rotação de  $(\bar{x}_k)_{k \in \mathbb{Z}}$  (ver definição 7) é  $\rho$ .

A prova que  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é minimal será por absurdo. Suponha que  $(\bar{x}_k)_{k \in \mathbb{Z}}$  não é minimal. Então existem inteiros  $a$  e  $b$  e um segmento  $(\xi_k)_{k \in [a,b]}$ , com  $\bar{x}_a = \xi_a$  e  $\bar{x}_b = \xi_b$ , tal que  $W_{ab}(\bar{x}) = W_{ab}(\xi) + \epsilon_1$ , para algum  $\epsilon_1 > 0$ . Como  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é o limite de configurações minimais e  $h$  é contínua, para qualquer  $\epsilon_2 > 0$  dado, existe uma configuração

minimal  $(x_k^i)_{k \in \mathbb{Z}}$  tal que:

$$\begin{aligned} |W_{ab}^-(\bar{x}) - W_{ab}^-(x^i)| &< \epsilon_2, \\ |h(\bar{x}_a, \xi_{a+1}) - h(x_a^i, \xi_{a+1})| &< \epsilon_2, \\ |h(\xi_{b-1}, \bar{x}_b) - h(\xi_{b-1}, x_b^i)| &< \epsilon_2 \end{aligned} \quad (2.7)$$

Considere agora o segmento  $(\eta_k)_{k \in [a, b]}$ , que é uma pequena variação de  $(\xi_k)_{k \in [a, b]}$ , dado por:  $\eta_k = \xi_k$  se  $a+1 \leq k \leq b-1$ ,  $\eta_a = x_a^i$ , e  $\eta_b = x_b^i$ . Então, escolhendo  $\epsilon_2 < \epsilon_1/3$  vale:

$$\begin{aligned} W_{ab}^-(x^i) - W_{ab}^-(\eta) &= W_{ab}^-(x^i) - W_{ab}^-(\bar{x}) + W_{ab}^-(\bar{x}) - W_{ab}^-(\xi) \\ &\quad - h(\eta_a, \xi_{a+1}) + h(\xi_a, \xi_{a+1}) \\ &\quad - h(\xi_{b-1}, \eta_b) + h(\xi_{b-1}, \xi_b) \\ &> \epsilon_1 - |W_{ab}^-(x^i) - W_{ab}^-(\bar{x})| \\ &\quad - |h(x_a^i, \xi_{a+1}) - h(\bar{x}_a, \xi_{a+1})| \\ &\quad - |h(\xi_{b-1}, x_b^i) - h(\xi_{b-1}, \bar{x}_b)| \\ &> \epsilon_1 - 3\epsilon_2 > 0, \end{aligned}$$

o que viola a minimalidade de  $(x_k^i)_{k \in \mathbb{Z}}$ .

Finalmente, provemos a monotonicidade de  $(\bar{x}_k)_{k \in \mathbb{Z}}$ . Dados  $i$  e  $k$  existe uma sequência de configurações minimais  $(x_p^j)_{p \in \mathbb{Z}}$ ,  $j = 0, 1, \dots$ , tal que  $x_p^j \rightarrow \bar{x}_p$ , quando  $j \rightarrow \infty$ , para  $p = k, k+1, k+2, i, i+1, i+2$ . Logo da monotonicidade de  $(x_p^j)_{p \in \mathbb{Z}}$ , ver equação (2.3), segue  $x_k^j < x_k^{j+1} < x_{k+1}^{j+1} < x_{k+1}^j + a$  e, portanto, vale  $\bar{x}_k < \bar{x}_i + a \Rightarrow \bar{x}_{k+1} \leq \bar{x}_{i+1} + a$ . Queremos mostrar que na última desigualdade não pode valer  $\bar{x}_{k+1} = \bar{x}_{i+1} + a$ . De fato, se esta igualdade vale então é possível mostrar que  $\bar{x}_{k+2} > \bar{x}_{i+2} + a$ . Mas isso implica que para valores de  $j$  suficientemente grandes vale  $x_{k+2}^j > x_{i+2}^j + a$ , o que viola a monotonicidade de  $(x_p^j)_{p \in \mathbb{Z}}$ . Portanto para concluir a prova da monotonicidade de  $(\bar{x}_k)_{k \in \mathbb{Z}}$ , basta mostrar que, se  $\bar{x}_k < \bar{x}_i + a$  e  $\bar{x}_{k+1} = \bar{x}_{i+1} + a$  então  $\bar{x}_{k+2} > \bar{x}_{i+2} + a$ . Provemos esta afirmação de uma forma geométrica. A configuração  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é minimal, e portanto está associada a uma órbita de  $F$ . Então, o que queremos provar decorre facilmente da propriedade de twist como mostrado na figura 2.7 (se voce não se convenceu escreva a prova).  $\square$

A configuração minimal  $(\bar{x}_k)_{k \in \mathbb{Z}}$ , garantida pelo teorema 4, está associada a uma órbita  $(\bar{\theta}_k, \bar{y}_k)_{k \in \mathbb{Z}}$  de  $f$  no cilindro  $C$  (lembre-se que

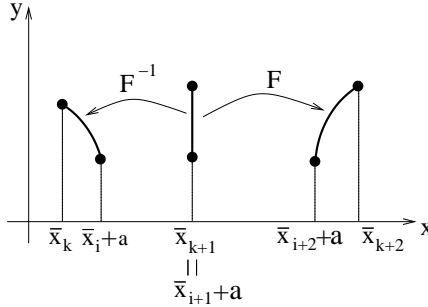


Figura 2.7: Argumento geométrico usado na prova da monotonicidade de  $(\bar{x}_k)_{k \in \mathbb{Z}}$ .

$\theta = x \pmod{1}$ ). O fecho de tal órbita, que será denotado por  $A_C$ , é formado pela união da órbita mais seus conjuntos  $\alpha$  e  $\omega$ -limite. Logo  $A_C$  é invariante por  $f$ . Seja  $A$  o levantamento de  $A_C$  para o plano. Denotaremos por  $A_1$  a projeção de  $A$  na reta  $\{x\}$ . O conjunto  $A_1$  também pode ser caracterizado como o fecho de todas as translações inteiras de  $(\bar{x}_k)_{k \in \mathbb{Z}}$ , ou seja:

$$A = \text{fecho} \{x_k^{pq} \stackrel{\text{def}}{=} \bar{x}_{k+p} + q : k \in \mathbb{Z}, p \in \mathbb{Z}, q \in \mathbb{Z}\}$$

**Lema 14.** *O conjunto  $A$  é monótono.*

*Demonstração.* Sejam  $x_0$  e  $\xi_0$  as componentes  $x$  de dois pontos quaisquer de  $A$  com  $x_0 < \xi_0$  e,  $(x_{k_i}^{p_i q_i})_{i > 0}$  e  $(x_{j_i}^{r_i s_i})_{i > 0}$ , duas seqüências de aproximantes de  $x_0$  e  $\xi_0$ , respectivamente, tais que  $x_{k_i}^{p_i q_i} < x_{j_i}^{r_i s_i}$ , para todo  $i > 0$ . Uma vez que  $(x_k^{pq})_{k \in \mathbb{Z}}$  é monótona para todo par  $(p, q)$ , então usando as desigualdades na definição 8 obtém-se  $x_{1+k_i}^{p_i q_i} < x_{1+j_i}^{r_i s_i}$ , o que implica

$$\lim_{i \rightarrow \infty} x_{1+k_i}^{p_i q_i} = x_1 \leq \lim_{i \rightarrow \infty} x_{1+j_i}^{r_i s_i} = \xi_{i+1},$$

onde  $x_1$  e  $\xi_1$  denotam as componentes  $x$  da imagem por  $F$  dos pontos associados a  $x_0$  e  $\xi_0$ . Novamente, para mostrar a monotonicidade de  $A$  basta mostrar que não ocorre a igualdade na desigualdade acima. De fato, se ocorrer a igualdade então, assim como ilustrado na figura

2.7 e na prova do teorema 4, vale que  $x_2 > \xi_2$ . Novamente, isto contradiz o fato que  $x_2 = \lim_{i \rightarrow \infty} x_{2+k_i}^{P_i Q_i}$ ,  $\xi_2 = \lim_{i \rightarrow \infty} x_{2+j_i}^{r_i s_i}$ , e que  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é monótona.  $\square$

Os conjuntos monótonos invariantes por  $F$  possuem a seguinte interessante propriedade geométrica.

**Teorema 5.** *Seja  $S$  um conjunto invariante e monótono de  $F$ . Suponha que  $S$  está contido entre os paralelos  $y = a$  e  $y = b$ . Então não existem dois pontos de  $S$  sobre a mesma vertical (ou seja, a projeção de  $S$  no eixo  $x$  é injetora) e se  $(\tilde{x}, \tilde{y})$ ,  $(\bar{x}, \bar{y})$  são dois pontos distintos quaisquer de  $S$  então vale:*

$$\frac{|\bar{y} - \tilde{y}|}{|\bar{x} - \tilde{x}|} < L,$$

onde  $L$  é uma constante que só depende dos valores de  $a$  e  $b$  e derivadas da aplicação  $F$ . Uma maneira mais sintética de expressar a tese, é dizer que:  $S$  é o gráfico de uma função Lipshitz  $x \rightarrow y$ , cujo domínio é a projeção de  $S$  no eixo  $x$ .

*Demonstração.* Se  $S$  é monótono então não pode conter dois pontos  $(x_1, y_1)$  e  $(\bar{x}_1, \bar{y}_1)$ , na mesma vertical com  $x_1 = \bar{x}_1$  e  $y_1 < \bar{y}_1$ , pois, como ilustrado na figura 2.7, isto implicaria  $\bar{x}_0 < x_0$ , o que contradiz a propriedade de monotonicidade de  $S$ .

A idéia para provar a desigualdade da tese do teorema é simples. Sem perda de generalidade, suponha que  $\tilde{x} < \bar{x}$ . Então como mostra a figura 2.8, se  $\tilde{y}$  for muito maior que  $\bar{y}$ , então a propriedade de twist implica que  $\tilde{x}_1 = F_1(\tilde{x}, \tilde{y}) > F_1(\bar{x}, \bar{y}) = \bar{x}_1$  o que viola a propriedade de monotonicidade. Para provar a desigualdade acima suporemos, primeiramente, que  $\tilde{y} > \bar{y}$ . Note que para todo  $y \in [a, b]$  e  $x \in \mathbb{R}$  existe uma constante  $\alpha$  tal que  $|\partial_1 F_1(x, y)| < \alpha$ , pois  $F_1$  é continuamente diferenciável e  $F_1(x + 1, y) = F_1(x, y) + 1$ . Isso implica que

$$|F_1(\bar{x}, \bar{y}) - F_1(\tilde{x}, \bar{y})| < \alpha |\bar{x} - \tilde{x}|$$

Lembrando que  $\bar{y} - \tilde{y} < 0$  e usando a hipótese  $c \leq \partial_2 F_1(x, y)$  de twist uniforme obtém-se:

$$F_1(\tilde{x}, \bar{y}) - F_1(\tilde{x}, \tilde{y}) = \int_{\tilde{y}}^{\bar{y}} \partial_2 F_1(\tilde{x}, s) ds < c(\bar{y} - \tilde{y})$$

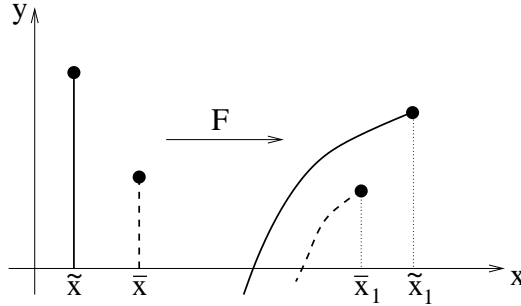


Figura 2.8: Argumento geométrico para mostrar o teorema 5.

Portanto, usando que  $\bar{x} > \tilde{x} \Rightarrow F_1(\bar{x}, \bar{y}) - F_1(\tilde{x}, \tilde{y}) > 0$ , obtém-se das desigualdades acima

$$\begin{aligned} 0 &< F_1(\bar{x}, \bar{y}) - F_1(\tilde{x}, \tilde{y}) = F_1(\bar{x}, \bar{y}) - F_1(\tilde{x}, \bar{y}) + F_1(\tilde{x}, \bar{y}) - F_1(\tilde{x}, \tilde{y}) \\ &< |F_1(\bar{x}, \bar{y}) - F_1(\tilde{x}, \bar{y})| + c(\bar{y} - \tilde{y}) < \alpha|\bar{x} - \tilde{x}| + c(\bar{y} - \tilde{y}) \end{aligned}$$

que é a desigualdade desejada no caso  $\tilde{y} > \bar{y}$ . O caso  $\tilde{y} < \bar{y}$  é provado de maneira análoga, trocando  $F$  por  $F^{-1}$  e usando desigualdades como as acima. Lembre que  $F^{-1}$  também é uma aplicação de tipo twist que desvia a vertical para a esquerda. A constante  $L$  do enunciado é a maior das constantes que aparece nos casos  $\tilde{y} > \bar{y}$  e  $\tilde{y} < \bar{y}$  (o caso  $\bar{y} = \tilde{y}$  é trivial).  $\square$

Aplicando o teorema 5 ao conjunto  $A$  definido acima, e usando que  $A$  é fechado, obtém-se que sua projeção no eixo  $x$ , denotada por  $A_1$ , também é um conjunto fechado. Logo o complementar de  $A_1$  é uma união de intervalos abertos. A fim de entender a topologia de  $A$  e a dinâmica de  $F$  restrito a  $A$ , faremos a construção auxiliar de um homeomorfismo do círculo que, “essencialmente”, possui a mesma dinâmica que  $F$  quando restrito a  $A$ . De acordo com o teorema 5, tal construção começa com a definição de uma função  $Y : A_1 \rightarrow \mathbb{R}$ , dada por

$$x \rightarrow y = Y(x) \quad \text{tal que} \quad (x, Y(x)) \in A. \quad (2.8)$$

A partir desta função e da invariância de  $A$  por  $F$ , constrói-se uma outra função  $v : A_1 \rightarrow A_1$  dada por  $x' = v(x) = F_1(x, Y(x))$ . O

teorema 5 garante que  $v$  também é uma função Lipschitz, a monotonicidade de  $A$  garante que  $v$  é monótona e a periodicidade de  $A_1$  e  $F_1$  garantem que  $v(x+1) = v(x) + 1$ . Agora a função  $v$  é estendida a toda a reta da seguinte forma: se  $x$  não pertence a  $A_1$  então  $v(x)$  possui o valor linearmente interpolado entre aqueles que  $v$  possui nos pontos de  $A_1$  mais próximos a  $x$ , ver figura 2.9. Tal extensão de  $v$ ,

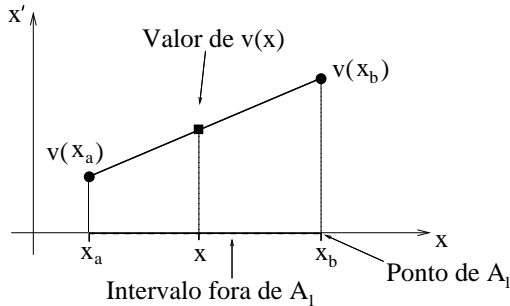


Figura 2.9: Extensão da função  $v$ .

que continuará a ser denotada por  $v$ , é um homeomorfismo Lipschitz da reta e  $v(x+1) = v(x) + 1$ , o que implica que  $v$  induz um homeomorfismo do círculo dado por  $\theta = x \pmod{1}$ ,  $\theta \in S^1$ . Além disso, por construção,  $v$  deixa invariante o conjunto  $A_1$ . Note que o gráfico da função Lipschitz

$$x \rightarrow y = u(x, v(x)) \stackrel{\text{def}}{=} \overline{Y}(x),$$

onde  $u$  é definida por  $x' = F_1(x, u(x, x'))$ , define uma curva que passa pelos pontos de  $A$ . O homeomorfismo  $v$  desempenhará um papel fundamental no entendimento da dinâmica de  $F$  restrito a  $A$ , o que será feito na próxima seção. Concluimos esta seção com o seguinte lema.

**Lema 15.** *Toda órbita em  $A$  é minimizante.*

*Demonstração.* O argumento usado nesta prova é muito semelhante àquele usado na parte da minimalidade na prova do teorema 4. Seja  $(x_0, y_0)$  um ponto qualquer de  $A$  e  $(x_k)_{k \in \mathbb{Z}}$  a configuração associada à sua órbita. Queremos mostrar que  $(x_k)_{k \in \mathbb{Z}}$  é uma configuração

minimal. Se  $x_0$  é um ponto de  $(\tilde{x}_k)_{k \in \mathbb{Z}}$ , onde  $(\tilde{x}_k)_{k \in \mathbb{Z}}$  é a configuração do teorema 4, então não há o que provar. Suponha que isso não ocorre ou seja, contrariamente ao que queremos provar,  $(x_k)_{k \in \mathbb{Z}}$  não é minimal. Então existem inteiros  $a$  e  $b$ , e um segmento  $(\xi_k)_{k \in [a,b]}$ , com  $\xi_a = x_a$  e  $\xi_b = x_b$ , tal que  $W_{ab}^-(x) - W_{ab}^-(\xi) > \epsilon_2$ , para algum  $\epsilon_2 > 0$ . Da continuidade de  $h$  e  $F$ , e do fato que  $A$  é o levantamento de  $A_C$  (que é o fecho da órbita  $(\tilde{x}_k, \tilde{y}_k)_{k \in \mathbb{Z}}$  projetada no cilindro  $C$ ), segue que dado qualquer  $\epsilon_1 > 0$  existem inteiros  $p$  e  $q$  tais que, a configuração minimal  $(\tilde{x}_k)_{k \in \mathbb{Z}}$  dada por  $\tilde{x}_k = \tilde{x}_{k+p} + q$ , satisfaz:

$$\begin{aligned} |W_{ab}^-(\tilde{x}) - W_{ab}^-(x)| &< \epsilon_1, \\ |h(\tilde{x}_a, \xi_{a+1}) - h(x_a, \xi_{a+1})| &< \epsilon_1, \\ |h(\xi_{b-1}, \tilde{x}_b) - h(\xi_{b-1}, x_b)| &< \epsilon_1 \end{aligned}$$

Considere agora o segmento  $(\eta_k)_{k \in [a,b]}$ , que é uma pequena variação de  $(\xi_k)_{k \in [a,b]}$ , dado por:  $\eta_k = \xi_k$  se  $a+1 \leq k \leq b-1$ ,  $\eta_a = \tilde{x}_a$ , e  $\eta_b = \tilde{x}_b$ . Então, escolhendo  $\epsilon_1 < \epsilon_2/3$  vale:

$$\begin{aligned} W_{ab}^-(\tilde{x}) - W_{ab}^-(\eta) &= W_{ab}^-(\tilde{x}) - W_{ab}^-(x) + W_{ab}^-(x) - W_{ab}^-(\xi) \\ &\quad - h(\eta_a, \xi_{a+1}) + h(\xi_a, \xi_{a+1}) \\ &\quad - h(\xi_{b-1}, \eta_b) + h(\xi_{b-1}, \xi_b) \\ &> \epsilon_2 - |W_{ab}^-(\tilde{x}) - W_{ab}^-(x)| \\ &\quad - |h(\tilde{x}_a, \xi_{a+1}) - h(x_a, \xi_{a+1})| \\ &\quad - |h(\xi_{b-1}, \tilde{x}_b) - h(\xi_{b-1}, x_b)| \\ &> \epsilon_2 - 3\epsilon_1 > 0, \end{aligned}$$

o que viola a minimalidade de  $(\tilde{x}_k)_{k \in \mathbb{Z}}$ . □

## 2.3 Homeomorfismos do círculo e conjuntos de Aubry-Mather

Antes de prosseguir com o estudo da dinâmica de  $F$  restrita ao conjunto  $A$ , é necessário apresentar alguns resultados sobre homeomorfismos do círculo (ver [11], [38]). Seja  $v$  como na seção anterior, um homeomorfismo  $v : \mathbb{R} \rightarrow \mathbb{R}$  com a propriedade  $v(x+1) = v(x) + 1$ ,

que induz um homeomorfismo  $\hat{v} : S^1 \rightarrow S^1$  do círculo  $\theta \rightarrow \hat{v}(\theta)$  através de  $\theta = x \pmod{1}$ .

Se  $x_1, x_2, \dots$  denotam as iteradas de um ponto  $x_0$  então o número de rotação  $\rho(x_0)$  de  $x_0$ , é definido exatamente como na definição 7. Como  $v$  induz um homeomorfismo do círculo, também o faz sua  $k$ -ésima iterada  $v^k$ . De fato,  $v^k(x+1) = v^k(x) + 1$ . Isso implica que se  $\tilde{x}$  e  $\bar{x}$  são dois pontos quaisquer de  $\mathbb{R}$  com  $|\tilde{x} - \bar{x}| < 1$ , então  $|\tilde{x}_k - \bar{x}_k| < 1$ . Isso por sua vez implica:

**Proposição 5.** *O número de rotação  $\rho(x)$  de  $v$  no ponto  $x$  independe do valor de  $x$ . Ou seja,  $\rho$  depende apenas do homeomorfismo  $v$ .*

*Demonstração.* De fato, se  $\tilde{x}$  e  $\bar{x}$  são dois pontos quaisquer de  $\mathbb{R}$  com  $|\tilde{x} - \bar{x}| < 1$ , então  $|\tilde{x}_k - \bar{x}_k| < 1$ . Isso por sua vez implica:

$$\begin{aligned} \rho(\tilde{x}) - \rho(\bar{x}) &= \lim_{k \rightarrow \infty} \frac{\tilde{x}_k - \bar{x}_k - (\tilde{x}_0 - \bar{x}_0)}{k} \\ &\leq \lim_{k \rightarrow \infty} \frac{|\tilde{x}_k - \bar{x}_k| + |\tilde{x}_0 - \bar{x}_0|}{k} \leq \lim_{k \rightarrow \infty} \frac{2}{k} = 0 \end{aligned}$$

Ou seja,  $\rho$  é constante em intervalos de comprimento um, e portanto constante em  $\mathbb{R}$ .  $\square$

Da proposição 5, segue imediatamente que se  $v$  possui uma órbita periódica tipo  $(m, n)$ , ou seja  $v^n(x_0) = x_0 + m$ , então o número de rotação de  $v$  é  $\rho = m/n$ . A recíproca de tal afirmação também vale.

**Proposição 6.** *O homeomorfismo  $v$  possui número de rotação racional  $\rho = m/n$ , com  $m$  e  $n$  primos entre si, se, e só se,  $v$  possui uma órbita periódica tipo  $(m, n)$ .*

*Demonstração.* Se  $v$  possui uma órbita periódica tipo  $(m, n)$ , então tal proposição é imediata. Resta provar que, se  $\rho = m/n$  então  $v$  possui uma órbita periódica de tipo  $(m, n)$ , ou equivalentemente, se  $v$  não possui órbita periódica de tipo  $(m, n)$  então  $\rho \neq m/n$ . Portanto, supondo que  $v$  não possui órbita periódica de tipo  $(m, n)$ , para todo  $x \in [0, 1]$  vale,  $v^n(x) > x + m$  ou  $v^n(x) < x + m$ . Considere o caso  $v^n(x) > x + m$ , o outro é análogo. Neste caso existe  $\epsilon > 0$  tal que, devido a periodicidade de  $v$ , para todo  $x \in \mathbb{R}$  vale,  $v^n(x) - x > +m + \epsilon$ . Isso implica que para qualquer  $i > 0$  inteiro vale:

$$v^{in}(x) - x = v^{in}(x) - v^{(i-1)n}(x) + v^{(i-1)n}(x) - \dots - v(x) - x > im + i\epsilon$$



Portanto

$$\rho = \lim_{i \rightarrow \infty} \frac{v^{in}(x) - x}{in} > \frac{m}{n} + \frac{\epsilon}{n},$$

ou seja,  $\rho \neq m/n$ . □

Provemos agora o resultado acerca de homeomorfismos do círculo que mais nos interessa. Lembramos que um conjunto de Cantor em  $\mathbb{R}$  pode ser caracterizado como: um conjunto fechado, perfeito (ou seja todo ponto do conjunto está arbitrariamente próximo de outros pontos do conjunto) e que não contém nenhum intervalo.

**Teorema 6.** *Suponha que o número de rotação de  $v$  é irracional. Então o conjunto  $\omega$ -limite  $E$  de qualquer ponto  $\theta \in S^1$ ,  $E \stackrel{\text{def}}{=} \omega(\theta)$ , é o mesmo para qualquer  $\theta$ , e  $E$ , ou é todo o círculo, ou é um conjunto de Cantor.*

*Demonstração.* Começaremos provando o seguinte resultado auxiliar. Seja  $\bar{\theta}$  um ponto qualquer de  $S^1$ , e  $a$  e  $b$  dois inteiros positivos quaisquer. Então a semi-órbita positiva de qualquer ponto  $\theta \in S^1$ , passa dentro do intervalo fechado com extremos  $\hat{v}^a(\bar{\theta})$  e  $\hat{v}^b(\bar{\theta})$  ( de fato, há dois destes intervalos em  $S^1$  e o resultado vale para ambos). Para mostrar isso basta provar que as pré-imagens de  $I = [\hat{v}^a(\bar{\theta}), \hat{v}^b(\bar{\theta})]$  pela aplicação  $\hat{v}$  cobrem  $S^1$ . A fim de simplificar a prova é conveniente analisar não as pré-imagens deste intervalo por  $\hat{v}$ , mas sim por  $\bar{v} = \hat{v}^i$  onde  $i = b - a$ . Se as pré-imagens de  $I$  por  $\bar{v}$  cobrem  $S^1$ , então o mesmo vale para as pré-imagens por  $\hat{v}$ . Note primeiramente que  $I, \bar{v}^{-1}(I), \bar{v}^{-2}(I), \dots$  são intervalos adjacentes, pois  $\bar{v}^{-1}(\hat{v}^b(\bar{\theta})) = \hat{v}^a(\bar{\theta})$ , etc. Portanto se estes intervalos não preenchem o círculo então existe um ponto  $\bar{\theta}_{-\infty}$  tal que  $\lim_{k \rightarrow \infty} \bar{v}^{-k}(\bar{\theta}) \rightarrow \bar{\theta}_{-\infty}$ . Da continuidade de  $\hat{v}$  segue que  $\bar{v}(\bar{\theta}_{-\infty}) = \bar{\theta}_{-\infty}$ , o que por sua vez implica que o número de rotação de  $\bar{v}$  é inteiro, o que viola a hipótese de irracionalidade de  $\rho(v)$ .

Seja  $\theta$  um ponto qualquer de  $S^1$  e  $p$  um ponto qualquer no conjunto  $\omega(\theta)$ . Então existe uma sequência de inteiros  $(k_i)_{i>0}$  tal que  $\hat{v}^{k_i}(\theta) = \theta_{k_i} \rightarrow p$ , quando  $i \rightarrow \infty$ . Seja  $z$  um outro ponto qualquer de  $S^1$ . Pelo resultado acima, dado qualquer  $i > 0$ , existe um elemento da semi-órbita positiva de  $z$  que entra no intervalo  $[\theta_{k_i}, \theta_{k_{i+1}}]$ . Logo o ponto  $p$  também está em  $\omega(z)$ . Isto mostra que  $E = \omega(\theta) = \omega(z)$ , ou seja o conjunto  $\omega$ -limite de qualquer ponto de  $S^1$  é  $E$ .

Note que  $E$  é o único conjunto fechado, não vazio, invariante por  $\hat{v}$  e minimal (aqui, por minimal entendemos que,  $E$  não contém estritamente nenhum outro conjunto fechado não vazio invariante por  $\hat{v}$ , ou equivalentemente, a órbita de todo ponto de  $E$  é densa em  $E$ ). De fato, seja  $B$  um conjunto não vazio, fechado e invariante por  $\hat{v}$ . Se  $\theta \in B$  então  $E = \omega(\theta) \subset B$ , pois  $B$  é fechado e invariante. Agora, como a fronteira  $\partial E$  do conjunto  $E$  também é um conjunto fechado invariante, concluímos que  $\partial E$  ou é vazio ou é igual a  $E$ . No primeiro caso temos que  $E = S^1$ . No segundo caso temos que  $E$  não contém nenhum intervalo, pois caso isso acontecesse então  $\partial E$  seria diferente de  $E$ . Como a órbita de qualquer ponto  $\theta$  em  $E$  se acumula em todos os pontos de  $E$ , pois  $\omega(\theta) = E$ , e  $\hat{v}^k(\theta) \neq \theta$  para todo  $k$ , pois de outro modo  $\rho$  seria, pela proposição 6, racional, segue que  $E$  é perfeito. Ou seja, se  $\partial E = E$  então  $E$  é um conjunto de Cantor.  $\square$

Voltemos agora a analisar o particular homeomorfismo  $v$  construído no fim da seção anterior, com o objetivo de estudar a dinâmica de  $F$  restrita a  $A$ . A notação segue aquela anterior ao lema 15. A configuração minimal  $(\bar{x}_k)_{k \in \mathbb{Z}}$ , obtida no teorema 4, está contida em  $A_1$  e possui número de rotação irracional  $\rho$ . Pela proposição 6, como  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é uma órbita de  $v$ , então o número de rotação de todas as órbitas de  $v$  é  $\rho$ . Seja  $\hat{v}$  a aplicação em  $S^1$  induzida por  $v$ . Se  $\hat{E}$  denota o conjunto  $\omega$ -limite de qualquer órbita de  $\hat{v}$ , como no teorema 6, seja  $E_1$  o levantamento de  $\hat{E}$  para a reta  $\{x\}$ . Pelo teorema 6 o conjunto  $\hat{E}$  é o conjunto  $\omega$ -limite da órbita  $(\bar{\theta}_k)_{k \in \mathbb{Z}}$ . Portanto, o conjunto  $A_1$  é a união de  $E_1$  mais o levantamento da órbita  $(\bar{\theta}_k)_{k \in \mathbb{Z}}$  para a reta, ou seja,  $(\bar{x}_{k+i} + j)_{k \in \mathbb{Z}, i \in \mathbb{Z}, j \in \mathbb{Z}}$ . Note que, não necessariamente  $A_1$  e  $E_1$  coincidem, pois a órbita  $(\bar{\theta}_k)_{k \in \mathbb{Z}}$  pode ser errante, ou seja, nenhum ponto da órbita é ponto de acumulação da própria. Seja  $E \subset A$  o levantamento de  $E_1$  para o plano  $\{x, y\}$  através de  $E = \{(x, y) : x \in E_1, y = Y(x)\}$ , onde  $Y$  é definida na equação (2.8). Como o número de rotação de qualquer ponto em  $E_1$  é  $\rho$ , o mesmo vale para  $E$  (e também para  $A$ ). O conjunto  $E \subset A$  é invariante por  $F$ , monótono, pelo lema 14 e minimizante, pelo lema 15. Agora, do teorema 6 segue que  $E$  pode assumir duas naturezas distintas. Primeiramente se  $\hat{E} = S^1$ , o que implica  $A_1 = \mathbb{R}$ , então  $E$  é o gráfico da função Lipschitz  $x \rightarrow Y(x)$ ,  $x \in \mathbb{R}$ , dada na equação (2.8). Ou seja, neste caso  $f$  possui um círculo invariante que é o gráfico de uma função Lipschitz  $\theta \rightarrow y = Y(\theta)$ , e toda órbita neste

círculo é minimizante e possui o mesmo número de rotação irracional.

Finalmente, se  $\hat{E}$  é um conjunto de Cantor,  $E$  também o é. Neste caso  $E$  está estritamente contido no gráfico de uma função Lipschitz  $x \rightarrow Y(x)$ ,  $x \in E_1$ . Ou seja,  $F$  (e conseqüentemente  $f$ ) possui um conjunto de Cantor invariante, chamado conjunto de “Aubry-Mather”, tal que toda órbita neste conjunto é minimizante e possui o mesmo número de rotação irracional. Em suma, vale o seguinte notável teorema.

**Teorema 7 (Aubry-Mather).** *Seja  $\rho$  um número irracional qualquer e  $f$  uma aplicação de tipo twist que satisfaz as hipóteses  $a$ ,  $b$ ,  $c$  da seção 1.3. Então  $f$  possui um conjunto invariante  $E$  no qual toda órbita é minimizante e possui número de rotação  $\rho$ . Tal conjunto  $E$  possui uma das duas seguintes caracterizações:*

*ou  $E$  é uma curva rotacional invariante dada pelo gráfico de uma função Lipschitz  $\theta \rightarrow y = Y(\theta)$ ;*

*ou  $E$  é um conjunto de Cantor invariante que está contido no gráfico de uma função Lipschitz  $\theta \rightarrow y = Y(\theta)$ . Neste caso  $E$  é chamado um conjunto de Aubry-Mather.*

## 2.4 Órbitas minimais heteroclínicas

Nesta seção voltaremos a estudar órbitas minimizantes com número de rotação racional. Na seção 2.1 foi provada a existência de órbitas minimizantes periódicas com número de rotação racional  $\rho = m/n$ . Considere o conjunto de todas essas órbitas, para um dado par de inteiros  $m$  e  $n > 0$  primos entre si, e seja  $P_{mn}$  o conjunto de todas as suas configurações minimais associadas. Pelo corolário 2 duas configurações de  $P_{mn}$  nunca se cruzam. Isso implica que, se  $w$  e  $z$  são duas configurações em  $P_{mn}$  e  $w_0 < z_0$ , então  $w_k < z_k$  para todo  $k \in \mathbb{Z}$ . Neste caso diz-se que  $w$  é menor que  $z$ ,  $w < z$ .

**Definição 10 (Configurações, ou órbitas, vizinhas).** *Duas configurações  $w$  e  $z$  de  $P_{mn}$  são ditas vizinhas se não existe nenhuma outra  $x \in P_{mn}$  tal que  $w < x < z$ . Duas órbitas minimizantes de tipo  $(m, n)$  de  $F$  são ditas vizinhas se suas respectivas configurações são vizinhas.*

Note que, se os gráficos de todas as configurações de  $P_{mn}$  são representados simultaneamente, como ilustrado na figura 2.10, então  $w$  e  $z$  são vizinhas se entre os gráficos de  $w$  e  $z$  não passar o gráfico de nenhuma outra configuração de  $P_{mn}$ . A princípio, pode ocorrer

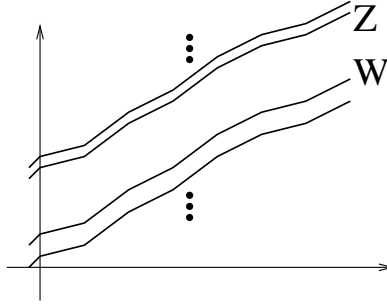


Figura 2.10: Duas configurações vizinhas.

que haja um contínuo de órbitas minimizantes (ou configurações minimais) de tipo  $(m,n)$  de modo que os gráficos das configurações em  $P_{mn}$  preencham o plano. Neste caso, que ocorre para a aplicação standard com  $\alpha = 0$  (ver figura 1.1), não existem configurações vizinhas em  $P_{mn}$ .

Dizemos que uma configuração  $x$  é heteroclínica entre duas configurações  $w$  e  $z$  se  $\lim_{k \rightarrow -\infty} |x_k - w_k| \rightarrow 0$  e  $\lim_{k \rightarrow \infty} |x_k - z_k| \rightarrow 0$ . Note que, se  $w$ ,  $x$  e  $z$  são todas configurações minimais e  $x$  é heteroclínica entre  $w$  e  $z$ , então a órbita associada a  $x$  é heteroclínica entre as órbitas associadas a  $w$  e  $z$ , no sentido de ser assintótica à órbita de  $w$  para  $k \rightarrow -\infty$  e assintótica à órbita de  $z$  para  $k \rightarrow \infty$ . Nosso objetivo nesta seção é provar o seguinte teorema.

**Teorema 8.** *Se  $W$  e  $Z$  são duas órbitas periódicas minimizantes de tipo  $(m,n)$  de  $F$ , então existe uma órbita heteroclínica minimizante entre  $W$  e  $Z$  e uma outra entre  $Z$  e  $W$ .*

*Demonstração.* Sejam  $(w_k)_{k \in \mathbf{Z}}$  e  $(z_k)_{k \in \mathbf{Z}}$  as configurações minimais associadas a  $W$  e  $Z$ , respectivamente. Iremos mostrar que existe uma configuração minimal,  $(x_k)_{k \in \mathbf{Z}}$ , que é heteroclínica entre  $(w_k)_{k \in \mathbf{Z}}$  e

$(z_k)_{k \in \mathbb{Z}}$ . A prova da existência de uma outra, heteroclínica entre  $(z_k)_{k \in \mathbb{Z}}$  e  $(w_k)_{k \in \mathbb{Z}}$ , é essencialmente a mesma. A prova será feita através da combinação de diversos resultados já apresentados e alguns outros a serem apresentados nesta seção. O primeiro destes resultados é análogo ao teorema de Tonelli do cálculo das variações.

**Proposição 7 (Existência de segmentos minimais).** *Dados um par de inteiros  $a, b$ , e um par de números reais  $x_a, x_b$ , existe um segmento minimal  $\{x_a, x_{a+1}, \dots, x_{b-1}, x_b\}$  que “conecta”  $x_a$  a  $x_b$ .*

*Demonstração.* Da definição 1 de segmento minimal segue que, para provar a proposição, basta mostrar que existe  $(x_{a+1}, \dots, x_{b-1}) \in \mathbb{R}^p$ ,  $p = b - a - 1$ , que minimiza a função

$$(r_1, \dots, r_p) \rightarrow h(x_a, r_1) + h(r_1, r_2) \dots + h(r_p, x_b) \stackrel{\text{def}}{=} W_{ab}^-(r)$$

Da proposição 3 decorre a desigualdade

$$W_{ab}^-(r) > \frac{c}{2} [(r_1 - x_a)^2 + (r_2 - r_1)^2 + \dots + (x_b - r_p)^2]$$

Se  $W_{ab}^-(0) = A$  então existe  $R > 0$ , suficientemente grande, tal que o lado direito da desigualdade acima é maior do que  $A$  se  $\|r\| \geq R$ . Logo a função  $W_{ab}^-$  possui um ponto de mínimo global na bola  $\|r\| < R$ .  $\square$

Sejam  $a_1$  e  $b_1$  um par de inteiros positivos. A proposição 7 implica que existe um segmento minimal  $\xi^1$  que conecta os pontos  $w_{-na_1}$  e  $z_{nb_1}$ . Dado outros dois inteiros  $a_2 \geq a_1$  e  $b_2 \geq b_1$ , existe um outro segmento minimal  $\xi^2$  que conecta  $w_{-na_2}$  a  $z_{nb_2}$ , e assim sucessivamente. Sejam  $x^j$ ,  $j = 1, 2, \dots$ , as seguintes configurações:  $x_k^j = \xi_k^j$  se  $na_j \leq k \leq nb_j$ ,  $x_k^j = w_k$  se  $k < -na_j$ , e  $x_k^j = z_k$  se  $k > nb_j$  (ver figura 2.11). A idéia central na prova do teorema 8, é mostrar que a sequência de configurações  $(x^j)_{j>0}$  possui uma subsequência, que converge para a desejada configuração heteroclínica. Esse é o conteúdo do seguinte lema.

**Lema 16.** *Existem duas sequências de inteiros  $(a_j)_{j>0}$ , e  $(b_j)_{j>0}$ , tais que a sequência de configurações  $(x^j)_{j>0}$  possui uma subsequência que converge para uma configuração minimal  $(x_k)_{k \in \mathbb{Z}}$ . A configuração  $(x_k)_{k \in \mathbb{Z}}$  : é distinta de  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$ ; seu gráfico está*

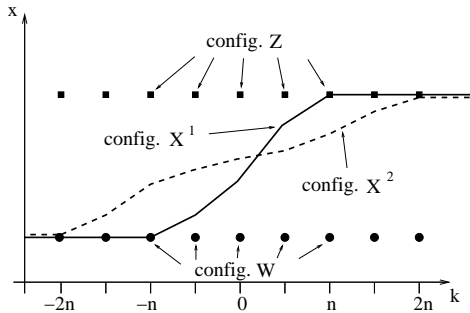


Figura 2.11: Idéia da prova de existência de uma configuração heteroclínica.

*contido entre os gráficos de  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$ ; satisfaz a propriedade de  $x_{k-n} + m \leq x_k$ , para qualquer  $k \in \mathbb{Z}$ ; e possui número de rotação  $m/n$ .*

*Demonstração.* Suponha que as sequências  $(a_j)_{j>0}$  e  $(b_j)_{j>0}$  sejam escolhidas, a priori, como  $a_j = b_j = j$ ,  $j = 1, 2, \dots$ . O lema 11, seção 2.1, implica que o gráfico do segmento minimal  $\xi^j$  não intesec-ta os gráficos de  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$ , para  $|k| < j - 1$  e  $j = 1, 2, \dots$ . Portanto, o gráfico de todas as configurações  $x^j$ ,  $j = 1, 2, \dots$ , está contido entre os gráficos de  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$ , e do mesmo argu-mento de compacidade usado na prova do teorema 4, envolvendo o método diagonal de Cantor, segue que existe uma subsequência de configurações  $(x^{j_i})_{i>0}$  que converge a uma configuração  $(\bar{x}_k)_{k \in \mathbb{Z}}$ , para  $i \rightarrow \infty$ . Tal configuração possui, por construção, número de rotação  $m/n$ . A configuração  $(\bar{x}_k)_{k \in \mathbb{Z}}$  é minimal. O argumento para provar isto é o mesmo usado na prova do teorema 4. Ou seja, se  $(\bar{x}_k)_{k \in \mathbb{Z}}$  não é minimal, então ela possui um segmento não minimal. Mas tal segmento pode ser arbitrariamente bem aproximado por segmentos minimais, e isso leva a uma contradição (ver o argumento envolvendo a equação (2.7) na prova do teorema 4). O problema com esta natural escolha de sequência de configurações  $(x^j)_{j>0}$  é provar que a a confi-guração limite obtida,  $(\bar{x}_k)_{k \in \mathbb{Z}}$ , não coincide nem com  $(w_k)_{k \in \mathbb{Z}}$  nem  $(z_k)_{k \in \mathbb{Z}}$ . Infelizmente, para demonstrar este fato precisamos modi-

ficar (e complicar) a escolha da sequência de configurações  $(x^j)_{j>0}$ . Neste ponto, deve-se enfatizar que a idéia central na prova do teorema 8 está contida neste parágrafo e na figura 2.11. O leitor não deve esquecer disso durante, e após, a leitura do particular e intrincado raciocínio que segue abaixo.

Seja  $\xi^1$  o segmento minimal que liga  $w_{-a_1n}$  a  $z_{b_1n}$ , onde  $a_1 = b_1 = 1$ . Se  $\xi_0^1 \geq (z_0 + w_0)/2$  então defina os segmentos  $A^1$  e  $B^1$  como  $A_k^1 = \xi_k^1$ ,  $k \in [-a_1n, b_1n]$  e  $B_k^1 = \xi_{k-n}^1 + m$ ,  $k \in [-(a_1 - 1)n, (b_1 + 1)n]$ . Neste caso defina  $a_2 = a_1$  e  $b_2 = b_1 + 1$ , ver figura 2.12 (para simplificar, nas próximas figuras  $m$  é sempre igual a zero). Por outro lado, se  $\xi_0 < (z_0 + w_0)/2$  então defina os segmentos  $A^1$

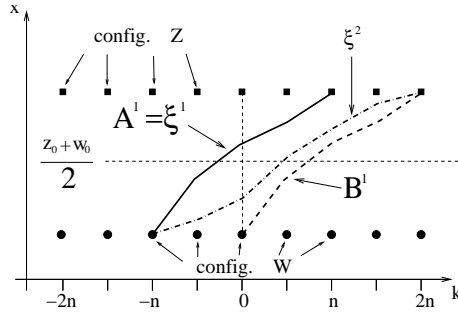


Figura 2.12: Algoritmo para a construção dos segmentos  $\xi^j$ .

e  $B^1$  como  $A_k^1 = \xi_{k+n}^1 - m$ ,  $k \in [-(a_1 + 1)n, (b_1 - 1)n]$  e  $B_k^1 = \xi_k^1$ ,  $k \in [-a_1n, b_1n]$ . Neste caso  $a_2 = a_1 + 1$  e  $b_2 = b_1$ , ver figura 2.13. Agora, seja  $\xi^2$  o segmento minimal que liga  $w_{-a_2n}$  a  $z_{b_2n}$ . Note que:  $A_{-n}^1 < (z_0 + w_0)/2 - m$ ,  $B_n^1 > (z_0 + w_0)/2 + m$  e  $A^1$  e  $B^1$  são minimais e não se cruzam (consequência do lema 10). O lema 10 também implica que o gráfico do segmento  $\xi^2$  esta abaixo do gráfico de  $A^1$  ( $\xi_k^2 \leq A_k^1$ ) e acima do gráfico de  $B^1$  ( $\xi_k^2 \geq B_k^1$ ), ver figuras 2.12 e 2.13. Se isso não ocorresse estes segmentos teriam que se cruzar duas vezes. Disso decorre que  $\xi_{-n}^2 < (z_0 + w_0)/2 - m$  e  $\xi_n^2 > (z_0 + w_0)/2 + m$ . Agora, repitamos o procedimento feito com  $\xi^1$  para obter  $\xi^2$ . Se  $\xi_0^2 \geq (z_0 + w_0)/2$  então defina os segmentos  $A^2$  e  $B^2$  como  $A_k^2 = \xi_k^2$ ,  $k \in [-a_2n, b_2n]$  e  $B_k^2 = \xi_{k-n}^2 + m$ ,  $k \in [-(a_2 - 1)n, (b_2 + 1)n]$ . Neste caso defina  $a_3 = a_2$  e  $b_3 = b_2 + 1$ . Note que  $A_{-n}^2 < (z_0 + w_0)/2 - m$

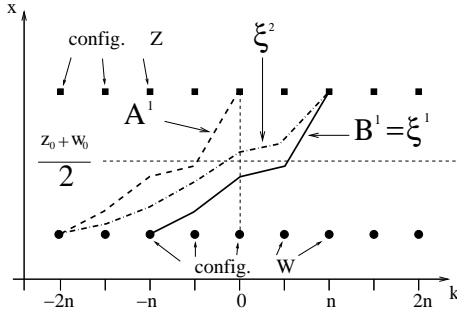


Figura 2.13: Algoritmo para a construção dos segmentos  $\xi^j$ .

e  $B_n^2 > D1(z_0 + w_0)/2 + m$ . Se  $\xi_0^2 < (z_0 + w_0)/2$  então defina os segmentos  $A^2$  e  $B^2$  como  $A_k^2 = \xi_{k+n}^1 - m$ ,  $k \in [-(a_1 + 1)n, (b_1 - 1)n]$  e  $B_k^2 = \xi_k^1$ ,  $k \in [-a_1n, b_1n]$ . Neste caso  $a_3 = a_2 + 1$  e  $b_3 = b_2$ . Agora, seja  $\xi^3$  o segmento minimal que liga  $w_{-a_3n}$  a  $z_{b_3n}$ . Pode-se repetir a construção acima e obter  $a_4$ ,  $b_4$  e  $\xi^4$ , etc. Ao final obtém-se uma seqüência infinita de segmentos minimais  $(\xi^j)_{[a_j, b_j]}$ ,  $j = 1, 2, \dots$ , tais que:  $\xi_{-n}^j < (z_0 + w_0)/2 - m$ ,  $\xi_n^j > (z_0 + w_0)/2 + m$ ,  $\xi_{-na_j}^j = w_{-na_j}$  e  $\xi_{nb_j}^j = z_{nb_j}$ , para todo  $j$ . Além disso,  $a_j \rightarrow \infty$  e  $b_j \rightarrow \infty$ , para  $j \rightarrow \infty$ . Destas propriedades as únicas que não são facilmente obtidas da construção acima são  $a_j \rightarrow \infty$  e  $b_j \rightarrow \infty$ , para  $j \rightarrow \infty$ . Elas serão provadas abaixo. Seja agora  $(x^j)_{j>0}$  a seqüência de configurações definida por:  $x_k^j = \xi_k^j$  se  $a_j \leq k \leq b_j$ ,  $x_k^j = w_k$  se  $k < a_j$ , e  $x_k^j = z_k$  se  $k > b_j$ . A minimalidade de dos segmentos  $\xi^j$  e o lema 10 implicam que o gráfico de  $(\xi_k^j)_{k \in [a_j, b_j]}$  está acima do gráfico do seu transladado  $(\xi_{k-n}^j + m)_{k \in [a_j, b_j]}$ , pois estes segmentos não se cruzam. Logo vale  $x_k^j \geq x_{k-n}^j + m$ . Para a seqüência de configurações  $(x^j)_{j>0}$  pode-se repetir o raciocínio feito no primeiro parágrafo desta prova, e obter uma configuração minimal  $(x_k)_{k \in \mathbb{Z}}$ , com número de rotação  $m/n$  e tal que  $w_k \leq x_k \leq z_k$ . Além disso, por construção  $x_{-n} < (z_0 + w_0)/2 - m$  e  $x_n > (z_0 + w_0)/2 + m$  ou seja  $(x_k)_{k \in \mathbb{Z}}$  não coincide, nem com  $(w_k)_{k \in \mathbb{Z}}$ , nem com  $(z_k)_{k \in \mathbb{Z}}$ . Finalmente, a propriedade  $x_k^j \geq x_{k-n}^j + m$  implica que a configuração  $(x_k)_{k \in \mathbb{Z}}$  satisfaz  $x_k \geq x_{k-n} + m$ ,  $k \in \mathbb{Z}$ .



Para concluir a prova do lema é necessário provar que vale  $a_j \rightarrow \infty$  e  $b_j \rightarrow \infty$ , para  $j \rightarrow \infty$ . Da construção dessas seqüências segue que  $a_n \leq a_{n+1}$  e  $b_n \leq b_{n+1}$ . Contrariando a tese, suponha que uma das seqüências, por exemplo  $(b_j)_{j>0}$ , seja limitada. Neste caso existe um valor  $J$  de  $j$ , tal que  $b_j = b_J$  e  $a_{j+1} = a_j + 1$ , para todo  $j > J$ . Mais ainda, do lema 11 segue que, para todo  $j > J$ , o gráfico de  $\xi^{j+1}$  está acima do gráfico de  $\xi^j$ , uma vez que  $\xi_{b_J}^j = \xi_{b_J}^{j+1}$ , ver figura 2.14. Isso implica que a seqüência das configurações  $(x_k^j)_{[-\infty, b_J]}$ ,

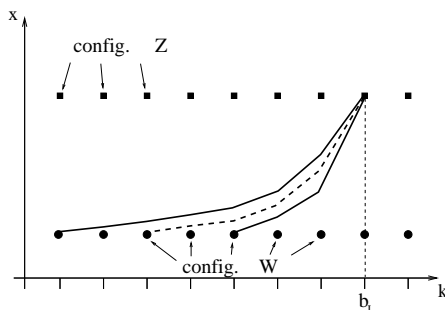


Figura 2.14: Seqüência de configurações convergindo para  $(\bar{x}_k)_{[-\infty, b_J]}$ .

covergirá a uma configuração minimal  $(\bar{x}_k)_{[-\infty, b_J]}$ . Novamente, o lema 10 implica que o gráfico de  $(\xi_k^j)_{k \in [a_j, b_J]}$  está acima do gráfico do seu transladado  $(\xi_{k-n}^j + m)_{k \in [a_j, b_J]}$ , pois estes segmentos não se cruzam. Disto decorre que a configuração limite satisfaz  $\bar{x}_k \geq \bar{x}_{k-n} + m$ ,  $k \in [-\infty, b_J]$ . Para terminar a prova do lema precisamos da seguinte proposição, onde finalmente será usada a hipótese que  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$  são configurações vizinhas. Tal proposição será também usada na finalização da prova do teorema 8.

**Proposição 8.** *As configurações  $(x_k)_{k \in \mathbb{Z}}$  e  $(\bar{x}_k)_{[-\infty, b_J]}$ , dadas na prova do lema 16, se existirem, satisfazem as seguintes desigualdades:  $x_k < x_{k-n} + m$ ,  $k \in \mathbb{Z}$  e  $\bar{x}_k < \bar{x}_{k-n} + m$ ,  $k \in [-\infty, b_J]$ . Além disso, existem inteiros  $a$ ,  $b$  e  $c$  tais que valem os seguintes limites:  $\lim_{k \rightarrow -\infty} x_k - w_{k+a} = 0$ ,  $\lim_{k \rightarrow \infty} x_k - z_{k+b} = 0$  e  $\lim_{k \rightarrow -\infty} \bar{x}_k - w_{k+c} = 0$ .*

*Demonstração.* Vamos provar apenas que  $x_k < x_{k-n} + m$ ,  $k \in \mathbb{Z}$ , e  $\lim_{k \rightarrow -\infty} x_k - w_{k+a} = 0$ . As outras afirmações são provadas usando-se os mesmos argumentos. Da prova do lema 16 segue que  $x_k \geq x_{k-n} + m$ . Suponha que exista  $k$  tal que  $x_k = x_{k-n} + m$ . Uma vez que  $(x_k)_{k \in \mathbb{Z}}$  é uma configuração minimal, isso implica que  $(x_k)_{k \in \mathbb{Z}}$  é uma configuração tipo  $(m, n)$ . Como o gráfico de  $(x_k)_{k \in \mathbb{Z}}$  está entre os gráficos de  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$ , e não coincide com nenhum desses, obtém-se que  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$  não são configurações vizinhas, o que é impossível. Logo vale  $x_k < x_{k-n} + m$ ,  $k \in \mathbb{Z}$ . Considere agora a sequência de configurações  $\tilde{x}_k^j = x_{k-jn} + jm$ ,  $k \in \mathbb{Z}$  e  $j = 1, 2, \dots$ . Note que, para todo  $j > 0$ , o gráfico de  $\tilde{x}^j$  está acima do gráfico de  $\tilde{x}^{j+1}$ , que está acima do gráfico de  $(w_k)_{k \in \mathbb{Z}}$ , e vale  $\tilde{x}_0^j > \tilde{x}_0^{j+1} = \tilde{x}_{-n}^j + m > w_0$ . Isso, como antes, implica que a sequência de configurações minimais  $(\tilde{x}^j)_{j>0}$  converge para uma configuração minimal  $(\hat{x}_k)_{k \in \mathbb{Z}}$ , que satisfaz  $\hat{x}_{-n} + m = \hat{x}_0 \geq w_0$ . Ou seja,  $(\hat{x}_k)_{k \in \mathbb{Z}}$  é uma configuração minimal periódica do tipo  $(m, n)$ , cujo gráfico está entre os gráficos de  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$ , podendo coincidir apenas com  $(w_k)_{k \in \mathbb{Z}}$ , pois a sequência  $(\tilde{x}^j)_{j>0}$  é decrescente. Como  $(w_k)_{k \in \mathbb{Z}}$  e  $(z_k)_{k \in \mathbb{Z}}$  são configurações vizinhas,  $(\hat{x}_k)_{k \in \mathbb{Z}}$  tem que coincidir com  $(w_k)_{k \in \mathbb{Z}}$ . Isso implica o limite enunciado na proposição.  $\square$

Voltemos a prova do lema 16. Seja  $\tilde{x}_k = \bar{x}_{k-n} + m$ ,  $k \in [-\infty, b_J + n]$ . A proposição 8 implica que as configurações  $(\bar{x}_k)_{[-\infty, b_J]}$  e  $(\tilde{x}_k)_{[-\infty, b_J + n]}$  satisfazem:  $\bar{x}_k > \bar{x}_{k-n} + m = \tilde{x}_k$ ,  $k \in [-\infty, b_J]$ ,  $\lim_{k \rightarrow -\infty} \bar{x}_k - w_{k+c} = 0$  e  $\lim_{k \rightarrow -\infty} \tilde{x}_k - w_{k+c} = 0$ . Considere agora a configuração  $(\hat{x}_k)_{[-\infty, b_J + n]}$ , dada pela colagem de  $(\bar{x}_k)_{[-\infty, b_J]}$  com um período da configuração  $(z_k)_{k \in \mathbb{Z}}$ , ou seja,  $\hat{x}_k = \bar{x}_k$ , para  $k \leq b_J$  e  $\hat{x}_k = z_k$  para  $b_J < k \leq b_J + n$ , ver figura 2.15. Abaixo será mostrado que a configuração  $(\hat{x}_k)_{[-\infty, b_J + n]}$  é minimal. Mas isto é impossível, pois  $(\hat{x}_k)_{[-\infty, b_J + n]}$  possui um segmento que coincide com  $(z_k)_{k \in \mathbb{Z}}$ , e como ambas são minimais deveriam coincidir, pois geram a mesma órbita da aplicação  $F$ . Isso por sua vez mostra que  $(b_j)_{j>0}$  não pode ficar limitada quando  $j \rightarrow \infty$ .

Portanto, para concluir a prova do lema basta mostrar que se a configuração  $(\hat{x}_k)_{[-\infty, b_J + n]}$  existe ela é minimal. Para simplificar a notação, façamos uma translação nas configurações envolvidas de modo que  $k = b_J$  seja levado em  $k = 0$ . Seja  $a < 0$  um inteiro qualquer e considere as funções  $(\eta \rightarrow \tilde{W}_a(\eta))$  e  $(\eta \rightarrow \tilde{W}_a(\eta))$ ,  $\eta =$

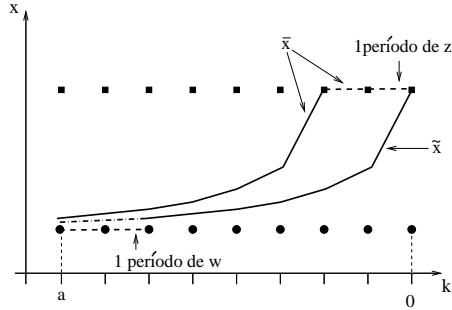


Figura 2.15: Construção usada para mostrar que  $(\bar{x}_k)_{-\infty,0}$  não pode existir.

$\eta_{a+1}, \dots, \eta_{-1}$ ), dadas por:

$$\begin{aligned} \hat{W}_a(\eta) &= h(\hat{x}_a, \eta_{a+1}) + h(\eta_{a+1}, \eta_{a+2}) + \dots + h(\eta_{-1}, z_0) \\ \tilde{W}_a(\eta) &= h(\tilde{x}_a, \eta_{a+1}) + h(\eta_{a+1}, \eta_{a+2}) + \dots + h(\eta_{-1}, z_0) \end{aligned}$$

Note que para qualquer  $\eta$  dado:

$$|\hat{W}_a(\eta) - \tilde{W}_a(\eta)| = |h(\hat{x}_a, \eta_{a+1}) - h(\tilde{x}_a, \eta_{a+1})| \rightarrow 0, \quad \text{para } a \rightarrow -\infty, \quad (2.9)$$

uma vez que  $\lim_{a \rightarrow -\infty} |\hat{x}_a - \tilde{x}_a| = 0$ , ver figura 2.15. Note também que (ver figura 2.15):

$$\begin{aligned} |\hat{W}_a(\hat{x}) - \tilde{W}_a(\tilde{x})| &= |h(z_{-n}, z_{-n+1}) + \dots + h(z_{-1}, z_0) \\ &\quad - h(\tilde{x}_a, \tilde{x}_{a+1}) + \dots - h(\tilde{x}_{a+n-1}, \tilde{x}_{a+n})| \\ &\rightarrow 0 \quad \text{para } a \rightarrow -\infty, \end{aligned} \quad (2.10)$$

pois,  $\lim_{k \rightarrow -\infty} \tilde{x}_k - w_{k+c} = 0$  implica que, para  $a \rightarrow \infty$ , vale

$$\begin{aligned} h(\tilde{x}_a, \tilde{x}_{a+1}) + \dots + h(\tilde{x}_{a+n-1}, \tilde{x}_{a+n}) &\rightarrow \\ h(w_{a+c}, w_{a+c+1}) + \dots + h(w_{a+c+n-1}, w_{a+c+n}) &\end{aligned}$$

e

$$\begin{aligned} h(w_{a+c}, w_{a+c+1}) + \dots + h(w_{a+c+n-1}, w_{a+c+n}) &= \\ h(z_{-n}, z_{-n+1}) + \dots + h(z_{-1}, z_0) &\end{aligned}$$

pois ambas configurações periódicas minimizam globalmente a mesma função  $W_{mn}$ , definida na seção 2.1. Agora suponha, ao contrário do que queremos provar, que  $(\hat{x}_k)_{-\infty,0}$  não é minimal. Então existem inteiros negativos  $p$  e  $q$ , um  $\epsilon > 0$ , e uma variação  $\eta$  de  $(\hat{x}_k)_{[-\infty,0]}$ , dada por,  $\{\dots\hat{x}_p, \eta_{p+1}, \dots, \eta_{q-1}, \hat{x}_q, \dots, \hat{x}_0\}$ , tal que para todo  $a < q$  vale  $\tilde{W}(\eta) < \hat{W}_a(\hat{x}) - \epsilon$ , onde  $\epsilon$  independe de  $a$ . Dessa desigualdade e da minimalidade de  $\tilde{x}$  segue que (na desigualdade abaixo “ $\eta$ ” denota as devidas restrições da configuração  $\eta$ ):

$$\tilde{W}_a(\eta) - \tilde{W}_a(\eta) < \hat{W}_a(\hat{x}) - \epsilon - \tilde{W}_a(\eta) < \hat{W}_a(\hat{x}) - \tilde{W}_a(\tilde{x}) - \epsilon$$

Tomando o limite  $a \rightarrow -\infty$  dos dois lados desta desigualdade, e usando os limites (2.9) e (2.10), obtém-se a contradição desejada. Ou seja, se a configuração  $(\hat{x}_k)_{[-\infty,0]}$  existe ela é minimal.  $\square$

Finalmente, o teorema 8 é uma consequência imediata do lema 16 e da proposição 8.  $\square$

## 2.5 A Fórmula de MacKay-Meiss

A classificação dos pontos periódicos (hiperbólico, elíptico ou parabólico) de uma aplicação que preserva área depende apenas do traço da derivada. Isto ocorre porque o determinante da derivada é igual a 1.

Nesta seção, apresentamos uma fórmula, devida a Mackay e Meiss, que relaciona o traço da derivada da aplicação de twist num ponto periódico com o determinante hessiano da função geratriz ao longo da configuração associada à órbita.

Suponha que uma aplicação do tipo twist  $F$  que preserva área tenha função geratriz  $h(x, x')$ .

Sabemos que uma sequência  $(x_i)_{i \in \mathbb{Z}}$  está associada a uma órbita  $F(x_i, y_i) = (x_{i+1}, y_{i+1})$  se e somente se  $y_i = -\partial_1 h(x_i, x_{i+1}) = \partial_2 h(x_{i-1}, x_i)$ .

Ou seja, a sequência  $(x_i)_{i \in \mathbb{Z}}$  é um ponto crítico do funcional  $W(x) = \sum_{i \in \mathbb{Z}} h(x_i, x_{i+1})$  : ou  $\partial_1 h(x_i, x_{i+1}) + \partial_2 h(x_{i-1}, x_i) = 0$ .

Escrevendo  $z_i = (x_i, y_i)$  e  $\begin{pmatrix} u_i \\ v_i \end{pmatrix} \in T_{z_i}(\mathcal{C})$ , espaço tangente de  $C$  em  $z_i$ ,

$$dF(z_i) \begin{pmatrix} u_i \\ v_i \end{pmatrix} = \begin{pmatrix} u_{i+1} \\ v_{i+1} \end{pmatrix} \quad \text{então}$$

$$v_i = -\partial_{11}h(x_i, x_{i+1})u_i - \partial_{12}h(x_i, x_{i+1})u_{i+1} \quad \text{e}$$

$$v_i = \partial_{21}h(x_{i-1}, x_i)u_{i-1} + \partial_{22}h(x_{i-1}, x_i)u_i$$

Isto significa que a sequência de vetores  $\begin{pmatrix} u_i \\ v_i \end{pmatrix}$  é um campo de vetores ao longo de uma órbita de  $F$  então satisfaz à equação de Jacobi:

$$\partial_{22}h(x_{i-1}, x_i) + \partial_{11}h(x_i, x_{i+1})]u_i + \partial_{21}h(x_{i-1}, x_i)u_{i-1} + \partial_{22}h(x_i, x_{i+1})u_{i+1} = 0$$

Em particular se  $(z_i)$  é uma órbita periódica, de período  $q$  então  $F^q(z_i) = z_i + (p, 0)$ , ou seja, existe um inteiro  $p$  tal que  $x_{i+q} = x_i + p$ ,  $\forall_i \in \mathbb{Z}$ , e a equação de Jacobi acima nos permite relacionar a natureza de ponto crítico  $(x)$  com o tipo de órbita periódica.

**Teorema 9. (MacKay-Meiss )**

$$\text{tr } dF^q(z_0) - 2 = \frac{(-1)^q \text{Hess } W(x)}{\prod \partial_{12}h(x_i, x_{i+1})}.$$

Como  $\det dF^q(z_0) = 1$  então a classificação das órbitas periódicas (hiperbólica, elíptica ou parabólica) depende apenas do sinal de  $|\text{tr } dF^q(z_0)| - 2$ , pois a condição de twist nos garante que  $\partial_{12}h(x_i, x_{i+1}) < 0$  e o sinal do produto no denominador é igual a  $(-1)^q$ .

**Corolário 4.** *Se  $\text{Hess } W(x) > 0$  então  $(z_0, \dots, z_{q-1})$  é uma órbita periódica elíptica ou hiperbólica .*

*Se  $\text{Hess } W(x) < 0$  então  $(z_0, \dots, z_{q-1})$  é hiperbólica*

*Se  $\text{Hess } W(x) = 0$  então  $(z_0, \dots, z_{q-1})$  é parabólica*

*Demonstração.* Se  $\lambda$  é um autovalor de  $dF^q(z_0)$  e  $\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$  é o autovetor associado então  $dF^q(z_0) \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \lambda \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$ .

Usando a regra da cadeia e a periodicidade da órbita  $(z_i, \dots, z_q)$  temos  $dF^q(z_i) \begin{pmatrix} u_i \\ v_i \end{pmatrix} = \lambda \begin{pmatrix} u_i \\ v_i \end{pmatrix}$  onde  $\begin{pmatrix} u_{i+1} \\ v_{i+1} \end{pmatrix} = dF(z_i) \begin{pmatrix} u_i \\ v_i \end{pmatrix}$ . Isto significa que  $\lambda u_i = u_{i+q}$ .

Substituindo esta igualdade na equação de Jacobi para  $i = 1$  e  $i = q$  respectivamente, obtemos:

$$\partial_{21}h(x_0, x_1)\lambda^{-1}u_q + [\partial_{11}h(x_1, x_2) + \partial_{22}h(x_0, x_1)]u_1 + \partial_{12}h(x_1, x_2)u_2 = 0$$

$$\partial_{21}h(x_q, x_{q+1})\lambda u_1 + [\partial_{11}h(x_q, x_{q+1}) + \partial_{22}h(x_{q-1}, x_q)]u_q + \partial_{12}h(x_{q-1}, x_q)u_{q-1} = 0$$

Escrevendo na forma matricial  $M(\lambda).u = 0$  com  $\begin{pmatrix} u_i \\ \vdots \\ u_q \end{pmatrix}$  e a matriz

$M(\lambda)$  tem as seguintes entradas  $M_{ij}$  :

$$M_{11} = \partial_{11}h(x_1, x_2) + \partial_{22}h(x_0, x_1),$$

$$M_{12} = \partial_{12}h(x_1, x_2) \text{ e } M_{1q} = \partial_{21}h(x_0, x_1)\lambda^{-1}$$

Para  $i = 2, \dots, q-1$ :

$$M_{i \ i-1} = \partial_{21}h(x_{i-1}, x_i) \text{ e } M_{i \ i+1} = \partial_{12}h(x_i, x_{i+1})$$

$$M_{ii} = \partial_{11}h(x_i, x_{i+1}) + \partial_{22}h(x_{i-1}, x_i)$$

$$M_{q1} = \lambda \partial_{12}h(x_q, x_{q+1}) \text{ e } M_{q \ q-1} = \partial_{12}h(x_{q-1}, x_q)$$

$M_{qq}\partial_{11}h(x_q, x_{q+1}) + \partial_{22}h(x_{q-1}, x_q)$  e as demais posições são todas nulas.

Portanto,  $\lambda$  é um autovalor de  $dF^q(z_1)$  se e somente se  $\det M(\lambda) = 0$ .

Se  $D(\lambda) = \det M(\lambda)$ , então

$$D(\lambda) = D(1) + (\lambda - 1)(-1)^{q-1} \prod_{i=1}^q \partial_{12}h(x_i, x_{i+1}) + (\lambda^{-1} - 1)(-1)^{q-1} \prod_{i=1}^q \partial_{12}h(x_i, x_{i-1})$$

Portanto  $D(\lambda) = 0$  se e somente se

$$D(1) = (-1)^q (\lambda + \lambda^{-1} - 2) \prod_{i=1}^q \partial_{12}h(x_i, x_{i+1}).$$

Lembrando que  $\lambda^{-1}$  também autovalor de  $dF^q(Z_1)$ , se  $\lambda$  o for, podemos escrever

$$tr(dF^q(z_1)) - 2 = \frac{(-1)^q D(1)}{\prod_{i=1}^q \partial_{12}h(x_i, x_{i+1})}$$

Pela condição de twist o denominador não se anula, em particular com as nossas hipóteses temos  $\partial_{12}h < 0$ . Obviamente  $z_i$  é ponto

periódico simples ( $\lambda \neq 1$ ) se e somente se  $D(1) \neq 0$ . Escrevendo  $W(x_1, \dots, x_q) = h(x_1, x_2) + \dots + h(x_q, x_{1+p})$  (lembre-se que  $x_{q+1} = x_1 + p$ ) de modo que

$$\text{grad } W(x_1, \dots, x_q) =$$

$$(\partial_1 h(x_1, x_2) + \partial_2 h(x_q, x_1 + p), \dots, \partial_2 h(x_{q-1}, x_q) + \partial_1 h(x_q, x_1 + p))$$

e calculando a derivada segunda conclui-se que  $d^2W(x_1, x_2, \dots, x_q) = M(1)$ .

Portanto  $\det M(1) = \det d^2W(x_1, \dots, x_q) = \text{Hess } W(x_1, \dots, x_q)$  para uma sequência que é o ponto crítico de  $W$  e, que satisfaz ( $x_{i+q} = x_i + p$ ).

Com isto fica provada a fórmula.  $\square$

De acordo com o corolário 4, a fórmula de Mackay e Meiss nos ajuda a classificar os pontos periódicos de aplicações do tipo twist que preservam área usando a função geratriz.

Por exemplo, se a configuração é um ponto de mínimo não degenerado então a órbita periódica  $O(p)$  correspondente é *hiperbólica* e o Teorema de Hartman e Grobman [24] nos garante que existe uma vizinhança  $U$  do ponto  $p$  tal que  $f^n|U$  é topologicamente conjugada à sua parte linear  $Df^n(p)$ .

Por outro lado, se a órbita periódica for elíptica, ou seja se os autovalores  $\lambda$  e  $\lambda^{-1}$  são complexos com  $|\lambda| = 1$ , a derivada, cuja matriz é semelhante a uma rotação, é insuficiente para dar informações sobre a dinâmica local em torno do ponto periódico.

Para tratar desta questão, o primeiro passo é dado pela Forma Normal de Birkhoff, cuja consequência fundamental é dar condições suficientes para que a aplicação restrita a uma vizinhança de um ponto fixo elíptico seja do tipo twist. Este é um dos exemplos mais importantes de aplicações do tipo twist que preservam área.

## 2.6 Métodos topológicos

### 2.6.1 Órbitas periódicas

Nesta subseção demonstraremos a existência de órbitas periódicas para as aplicações do tipo twist de uma maneira construtiva, através

da análise da intersecção de certos conjuntos e de suas propriedades. Suporemos apenas que  $f : A \rightarrow A$  ou  $f : C \rightarrow C$  satisfaz, além da propriedade de twist, a seguinte propriedade de intersecção de curvas:

**PIC** : Dada qualquer curva fechada simples  $\gamma \subset \text{int}(A)$  ou  $\gamma \subset C$ , cujo complementar não contenha um disco (intuitivamente,  $\gamma$  circunda o anel ou cilindro), então  $f(\gamma) \cap \gamma \neq \emptyset$ . É claro que se  $f$  preserva área no anel ou é exata no cilindro, então  $f$  satisfaz PIC.

No que se segue, denotaremos o anel  $A$  ou o cilindro  $C$  por  $M$  e um levantamento  $F$  de  $f$  está fixado. No caso do anel, como  $f$  é um homeomorfismo que fixa as componentes do bordo de  $A$ , em cada componente do bordo  $f$  age como um homeomorfismo do círculo. Assim podemos associar um número de rotação a cada uma destas componentes. Denotando as componentes do bordo de  $A$  por  $\partial_a$  e  $\partial_b$ , vamos supor que

$$\rho_a = \rho(f \text{ restrita a } \partial_a) < \rho(f \text{ restrita a } \partial_b) = \rho_b.$$

Se  $f$  preservar área, então é possível mostrar que a desigualdade abaixo é consequência disso mais a propriedade de twist, ver cap. 3, corolário 6.

Vamos agora considerar o seguinte conjunto

$$K(p, q) = \{(x, y) \in M : (F^q)_1(\tilde{x}, \tilde{y}) = \tilde{x} + p\}, \quad (2.11)$$

onde  $(\tilde{x}, \tilde{y}) \in \mathbb{R}^2$  é um levantamento qualquer de  $(x, y)$ . Então vale a seguinte:

**Proposição 9.** *Para todo racional  $\rho_a < p/q < \rho_b$  (se  $M$  for o cilindro, vale para todo racional),  $K(p, q)$  contém um compacto conexo  $C(p, q)$ , cujo complementar possui pelo menos 2 componentes conexas, uma contendo o fim inferior de  $M$  e outra contendo o fim superior.*

*Demonstração.* Para  $q = 1$ , a prova é imediata a partir da condição de twist, pois para cada  $x \in S^1$ , existe um único ponto de  $K$  ( $K$  é conexo e é gráfico sobre  $S^1$ ).



Se  $q > 1$ ,  $K$  pode não ser mais um conjunto conexo, assim é preciso procurar um compacto conexo, contido em  $K$ , que separe  $M$  como no enunciado da proposição. Primeiramente, definamos o seguinte conjunto aberto:

$$O_1 = \{(x, y) \in M : (F^q)_1(\tilde{x}, \tilde{y}) < \tilde{x} + p\},$$

onde  $(\tilde{x}, \tilde{y}) \in \mathbb{R}^2$  é um levantamento qualquer de  $(x, y)$ .

Se  $M$  for o cilindro, vamos observar que para qualquer  $\tilde{x} \in \mathbb{R}$  fixado, temos o seguinte:

$$\lim_{t \rightarrow \pm\infty} (F^q)_2(\tilde{x}, t) = \lim_{t \rightarrow \pm\infty} (F^q)_1(\tilde{x}, t) = \pm\infty \quad (2.12)$$

Onde (2.12) é consequência da condição de twist e do fato de que as nossas hipóteses implicam

$$\lim_{t \rightarrow \pm\infty} F_2(\tilde{x}, t) = \pm\infty.$$

Assim, existe  $Const > 0$  suficientemente grande, tal que :

$$O_1 \supset S^1 \times ] - \infty, -Const[ \text{ e } O_1 \cap S^1 \times ]Const, +\infty[ = \emptyset$$

No caso de  $M$  ser o anel é fácil ver que  $O_1$  contém uma vizinhança do  $\partial_a$  e  $\overline{O_1}$  não intersecta  $\partial_b$ .

Agora, seja  $O_2$  a componente conexa de  $O_1$ , que contém  $S^1 \times ] - \infty, -Const[$ , ou no caso do anel, a que contém  $\partial_a$ . E, seja  $O_3$  a componente conexa de  $M \setminus \overline{O_2}$ , que contém o fim superior de  $M$ . Então é claro que  $\partial(O_3)$  é um compacto conexo  $C$ , como no enunciado da proposição. Além disso, da definição de  $O_1$  temos que para todo  $z$  pertencente à fronteira de  $O_3$ :

$$(F^q)_1(\tilde{z}) = (\tilde{z})_1 + p \Rightarrow C \subset K(p, q).$$

□

Vamos agora lembrar a seguinte:

**Proposição 10.** *Se  $z \in f(C(p, q)) \cap C(p, q) \Rightarrow z$  é ponto  $(p, q)$  periódico para  $f$*

*Demonstração.* Para cada  $\tilde{z} \in \mathbb{R}^2$ , seja  $V_{\tilde{z}}$  a reta vertical por  $\tilde{z}$ . Como  $F$  é do tipo twist,  $F(V_{F^{-1}(\tilde{z})}) \cap V_{\tilde{z}} = \tilde{z}$ . Agora suponha que  $z \in f(C(p, q)) \cap C(p, q)$  :

$$\begin{aligned} F^q(\tilde{z}) &= F(V_{F^{q-1}(\tilde{z})}) \cap V_{F^q(\tilde{z})} = \\ &= F(V_{F^{-1}(\tilde{z})+(p,0)}) \cap V_{\tilde{z}+(p,0)} = F(V_{F^{-1}(\tilde{z}+(p,0))}) \cap V_{\tilde{z}+(p,0)} = \\ &= \tilde{z} + (p, 0) \end{aligned}$$

Assim,  $z$  é ponto  $(p, q)$  periódico para  $f$ . □

Finalmente estamos prontos para mostrar a existência de órbitas periódicas com todos os números de rotação possíveis:

**Teorema 10.** *Dado um racional  $\rho_a < p/q < \rho_b$  (se  $M$  for o cilindro,  $p/q$  é qualquer),  $f$  possui ao menos uma órbita  $q$ -periódica com esse número de rotação.*

*Demonstração.* Pela proposição 10, é suficiente mostrar que

$$f(C(p, q)) \cap C(p, q) \neq \emptyset.$$

Assim, por absurdo, vamos supor que tal intersecção é vazia. Neste caso, como  $C(p, q)$  é o bordo de um aberto homeomorfo a  $M$ , existe uma curva fechada simples  $\gamma \subset M$ , que circunda  $M$ , tal que  $C(p, q)$  pertence a uma componente conexa de  $\gamma^c$  e  $f(C(p, q))$  pertence a outra. Mas isso implica que  $f(\gamma) \cap \gamma = \emptyset$ , o que contradiz a hipótese de intersecção de curvas. □

Tudo o que foi feito anteriormente pode ser encontrado em [43] e [47].

O teorema acima não nos dá certas propriedades importantes que as órbitas periódicas das aplicações do tipo twist muitas vezes tem. Lembremos a definição de órbita bem ordenada (ver sec. 2.2):

**Definição:** Dizemos que uma órbita  $q$ -periódica  $\{z, f(z), \dots, f^{q-1}(z)\}$  de uma aplicação do tipo twist  $f$  é do tipo Birkhoff, ou bem ordenada se:

- para  $\tilde{z} = (\tilde{x}, \tilde{y}), \tilde{z}' = (\tilde{x}', \tilde{y}') \in \pi^{-1}(\{z, f(z), \dots, f^{q-1}(z)\})$  com  $\tilde{x} < \tilde{x}'$ , então  $F_1(\tilde{z}) < F_1(\tilde{z}')$ , onde  $\pi : \tilde{M} \rightarrow M$  é a aplicação de recobrimento e  $\tilde{M}$  é o recobrimento universal de  $M$ .

Finalmente citaremos, sem demonstração, o seguinte teorema devido a G. R. Hall [42]. Este teorema diz que se uma aplicação do tipo twist tem uma órbita periódica com um determinado número de rotação, então ela tem uma órbita bem ordenada com esse número de rotação.

**Teorema 11.** (Hall) *Dada uma aplicação do tipo twist  $f : M \rightarrow M$  e um racional  $p/q$ , se  $f$  possui ao menos uma órbita  $q$ -periódica com número de rotação  $p/q$ , então  $f$  possui uma órbita periódica do tipo Birkhoff com esse número de rotação.*

Para uma demonstração quase elementar desse resultado, ver [9]. Finalmente, se  $f$  satisfaz a propriedade de intersecção PIC acima citada, ela tem, pelos teoremas 10 e 11, órbitas periódicas bem ordenadas com todos os números de rotação possíveis. Agora aplicando os métodos de aproximação vistos anteriormente, sec. 2.2, obtemos a existência dos conjuntos de Aubry-Mather com números de rotação irracionais.

### 2.6.2 Teorema de Poincaré-Birkhoff

O teorema de Poincaré-Birkhoff, ou "último teorema geométrico de Poincaré" é um resultado muito importante, não só porque foi motivado por aplicações mecânicas (problema restrito dos 3 corpos), mas também porque deu início a uma importante área de pesquisa dentro dos sistemas dinâmicos. Basicamente, ele consiste numa generalização do teorema 10 sem a hipótese de twist. Vamos então a um enunciado.

Suponha que  $h : S^1 \times [0, 1] \rightarrow S^1 \times [0, 1]$  seja um homeomorfismo que preserva área, homotópico a identidade, isto é, se  $\tilde{h} : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R} \times [0, 1]$  é um levantamento de  $h$ , então  $\tilde{h}(\tilde{z} + (1, 0)) = \tilde{h}(\tilde{z}) + (1, 0)$  para todo  $\tilde{z} \in \mathbb{R} \times [0, 1]$  e  $h(S^1 \times \{i\}) = S^1 \times \{i\}$ , para  $i = 0, 1$ . Do fato de  $h$  ser um homeomorfismo, segue que  $h|_{S^1 \times \{i\}}$  é um homeomorfismo do círculo, para  $i = 0, 1$ . No que se segue vamos supor que os números de rotação desses homeos do círculo satisfazem a seguinte condição:

$$\rho(h|_{S^1 \times \{0\}}) < 0 < \rho(h|_{S^1 \times \{1\}})$$

O teorema que será demonstrado nessa seção nos diz que, dadas as hipóteses acima, então  $h$  tem pelo menos 1 ponto fixo. Na verdade é possível provar a existência de pelo menos 2 pontos fixos, mas isso não será feito aqui. Existem diversas versões desse teorema, desde "correções" das primeiras demonstrações feitas por Poincaré, até provas modernas, que dão mais informação sobre as órbitas.

Uma tendência muito importante que se iniciou com o trabalho de Kerekjarto foi a ligação entre essa teoria e o estudo dos chamados homeomorfismos de Brouwer, que são homeomorfismos do plano que preservam orientação, sem pontos fixos, ver [40], [37], [65] e [32] por exemplo.

Podemos citar como referências básicas as seguintes: Poincaré [62] e os trabalhos de Birkhoff [15] e [17], no primeiro provando a existência de apenas um ponto fixo e no segundo com uma prova mais topológica, garantindo os 2.

Para mais algumas versões interessantes do teorema de Poincaré-Birkhoff, ver por exemplo [20], [21], [8], [31] e [43].

A demonstração apresentada aqui é a que aparece em [15].

*Demonstração.* A prova será feita por absurdo. Assim, vamos supor que existe  $h$  satisfazendo as hipóteses acima, sem pontos fixos. Nesse caso, como  $A = S^1 \times [0, 1]$  é compacto, existe  $\delta > 0$  tal que

$$\text{distância}(h(z), z) > \delta, \text{ para todo } z \in A.$$

Seja  $\tilde{h} : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R} \times [0, 1]$  um levantamento de  $h$ . É óbvio que podemos estender  $\tilde{h}$  ao plano todo da seguinte maneira:

$$\text{dado } \tilde{z} = (\tilde{x}, \tilde{y}) \notin \mathbb{R} \times [0, 1], \text{ definimos}$$

$$\tilde{h}_{ext}(\tilde{z}) = \begin{cases} \tilde{h}(\tilde{x}, 1) + (0, \tilde{y} - 1), & \text{se } \tilde{y} > 1 \\ e \\ \tilde{h}(\tilde{x}, 0) + (0, \tilde{y}), & \text{se } \tilde{y} < 0 \end{cases}$$

É claro que  $\tilde{h}_{ext}$  é um homeomorfismo do plano, sem pontos fixos tal que  $\text{distância}(\tilde{h}_{ext}(\tilde{z}), \tilde{z}) > \delta$  para todo  $\tilde{z} \in \mathbb{R}^2$ . Vamos agora chamar  $\tilde{h}_{ext}$  simplesmente de  $\tilde{h}$  e consideremos a seguinte perturbação de  $\tilde{h}$ : Para  $0 < \epsilon < \delta$ , seja  $\tilde{h}_\epsilon(\tilde{z}) = \tilde{h}(\tilde{z}) + (0, \epsilon)$ .

Da escolha de  $\delta > 0$ , é claro que  $\tilde{h}_\epsilon$  não tem pontos fixos. Mais ainda, como  $\tilde{h}(\mathbb{R} \times \{0\}) = \mathbb{R} \times \{0\}$ , temos que  $\tilde{h}_\epsilon(\mathbb{R} \times \{0\}) = \mathbb{R} \times \{\epsilon\}$ . Como  $\tilde{h}_\epsilon$  preserva área na faixa  $\mathbb{R} \times [0, 1]$  e  $\tilde{h}_\epsilon(\mathbb{R} \times [0, \epsilon]) \cap \mathbb{R} \times [0, \epsilon] = \emptyset$ , temos que para algum inteiro  $n > 0$ ,  $\tilde{h}_\epsilon^n(\mathbb{R} \times [0, \epsilon]) \cap \mathbb{R} \times \{1\} \neq \emptyset$ . Assim, existe  $k > 1$  tal que  $\tilde{h}_\epsilon^k(\tilde{z}_0) = \tilde{z}_k$ , com  $\tilde{z}_0 \in \mathbb{R} \times \{0\}$  e  $\tilde{z}_k \in \mathbb{R} \times \{1\}$ .

Vamos agora considerar o segmento de reta  $r_0$  de extremos  $\tilde{z}_0$  e  $\tilde{z}_1 = \tilde{h}_\epsilon(\tilde{z}_0)$ . É claro que  $r_0 \subset \mathbb{R} \times [0, \epsilon]$  e  $r_0 \cup \tilde{h}_\epsilon(r_0) \cup \dots \cup \tilde{h}_\epsilon^{k-1}(r_0)$  é um arco simples contínuo  $\Gamma_\epsilon$  que começa em  $\tilde{z}_0 \in \mathbb{R} \times \{0\}$  e termina em  $\tilde{z}_k \in \mathbb{R} \times \{1\}$ .

Agora é necessário apresentar a seguinte definição:

**Definição (Índice de uma curva):** Dada uma curva simples  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  definimos o seu índice com relação a um homeomorfismo  $\tilde{h}$  da seguinte forma; seja  $H : [0, 1] \rightarrow S^1$  dada por

$$H(t) = \frac{\tilde{h}(\gamma(t)) - \gamma(t)}{\|\tilde{h}(\gamma(t)) - \gamma(t)\|}, \text{ para } t \in [0, 1].$$

Vamos agora tomar um levantamento de  $H$ , denominado  $\tilde{H} : [0, 1] \rightarrow \mathbb{R}$ . Então o índice é definido como  $\text{ind}(\gamma |_{\tilde{h}}) = \tilde{H}(1) - \tilde{H}(0)$ . É claro que o índice não depende da particular escolha do levantamento, pois 2 levantamentos distintos diferem por um inteiro.

É fácil ver que  $\tilde{h}_\epsilon(\Gamma_\epsilon) \subset \Gamma_\epsilon \cup \tilde{h}_\epsilon^k(r_0)$ . Assim, para  $\epsilon \rightarrow 0$ , o índice da correspondente  $\Gamma_\epsilon$  orientada com início em  $\tilde{z}_0$  e fim em  $\tilde{z}_k$  satisfaz  $\text{ind}(\Gamma_\epsilon |_{\tilde{h}_\epsilon}) \rightarrow -1/2$ .

Seja agora  $\beta$  um arco simples, tal que  $\beta \subset \mathbb{R} \times ]0, 1[$  e um extremo de  $\beta$  pertence a  $\mathbb{R} \times \{0\}$  e o outro pertence a  $\mathbb{R} \times \{1\}$ . Tomando um transladado de  $\Gamma_\epsilon$  podemos supor que  $\Gamma_\epsilon \cap \beta = \emptyset$ . Seja agora  $\Omega$  a curva fechada simples formada por  $\beta$ ,  $\Gamma_\epsilon$  e dois segmentos horizontais,  $\eta_0 \subset \mathbb{R} \times \{0\}$  e  $\eta_1 \subset \mathbb{R} \times \{1\}$ . Como  $\tilde{h}_\epsilon$  não possui pontos fixos,  $\text{ind}(\Omega |_{\tilde{h}_\epsilon}) = 0$ . Como  $\tilde{h}_\epsilon(\tilde{z} + (1, 0)) = \tilde{h}_\epsilon(\tilde{z}) + (1, 0)$  para todo  $\tilde{z} \in \mathbb{R}^2$ , obtemos que o  $\text{ind}(\beta |_{\tilde{h}_\epsilon}) \rightarrow -1/2$ , quando  $\epsilon \rightarrow 0$  e  $\beta$  é percorrida de baixo para cima, isto é, de  $\mathbb{R} \times \{0\}$  até  $\mathbb{R} \times \{1\}$ .

Assim,  $\text{ind}(\beta |_{\tilde{h}_\epsilon}) \rightarrow -1/2$ . Seja agora  $\alpha = \tilde{h}(\beta)$ . Fazendo a análise anterior utilizando  $\tilde{h}^{-1}$  ao invés de  $\tilde{h}$  e  $\alpha$  ao invés de  $\beta$  obtemos que

$ind(\alpha_{\uparrow} |_{\tilde{h}_{-1}}) = 1/2$ . Mas isto contradiz a seguinte propriedade sobre índices, de fácil verificação:

$$ind(\beta_{\uparrow} |_{\tilde{h}}) = ind(\tilde{h}(\beta)_{\uparrow} |_{\tilde{h}_{-1}})$$

O que garante a existência de pelo menos um ponto fixo para  $\tilde{h}$ .

□

## Capítulo 3

# Curvas Invariantes

Este Capítulo trata da noção de curva rotacional invariante por uma aplicação do tipo twist e de suas propriedades.

**Definição 11.** *Dizemos que  $\Gamma$  é uma curva rotacional invariante por uma aplicação do tipo twist  $f$  se  $\Gamma$  é a imagem de uma curva parametrizada contínua, fechada e simples (sem auto-interseção), homotopicamente não trivial, tal que  $f(\Gamma) = \Gamma$ .*

Por exemplo, o cilindro está completamente folheado por curvas rotacionais invariantes pela aplicação que no recobrimento, se escreve:  $F(x, y) = (x+y, y)$  e restrita a cada uma das curvas ( $y = \text{const.}$ ) temos uma rotação.

Vale portanto, perguntar se aplicações suficientemente próximas a  $f$  possuem curvas rotacionais invariantes. Birkhoff [18] provou é possível destruir as curvas invariantes rotacionais com número de rotação racional.

A existência de curvas rotacionais invariantes com número de rotação irracional é parte dos resultados obtidos por Kolmogorov, Ar-

nold e Moser, num contexto  $m$ -dimensional, que fornecem condições suficientes para a existência de toros invariantes ( Teoria KAM).

O caso bidimensional é particularmente interessante, pois o complementar de uma curva rotacional invariante tem duas componentes conexas homeomorfas ao cilindro que são invariantes por  $f$  e, portanto, as órbitas ficam confinadas à essas regiões.

Em particular, se existem curvas invariantes contidas em todas as vizinhanças de um ponto fixo elíptico, então este ponto é *estável*.

De fato, este exemplo é uma das motivações históricas para a investigação da existência de curvas invariantes, conforme pode ser visto no livro de Moser, *Stable and Random Motions in Dynamical Systems* [60].

De acordo com Moser, esta motivação vem da Mecânica Celeste e trata do problema da estabilidade do sistema solar. Este problema fundamental dos sistemas dinâmicos envolveu matemáticos como Laplace, Lagrange, Poisson, Weierstrass, Poincaré, Birkhoff, entre outros e tema de intensa pesquisa até hoje.

### 3.1 O Teorema da Curva Invariante de Birkhoff

Iniciaremos com a prova do Teorema da Curva Invariante de Birkhoff, cuja consequência mais importante é que toda curva rotacional invariante projeta-se injetivamente sobre  $S^1$  e portanto, é um gráfico.

As referências que usamos aqui são [29] e [38]

**Teorema 12.** *Seja  $f : S^1 \times \mathbb{R} \rightarrow S^1 \times \mathbb{R}$  uma aplicação do tipo twist que preserva área. Suponha que  $U \subset S^1 \times \mathbb{R}$  seja um subconjunto aberto tal que*

(i)  *$U$  é homeomorfo ao cilindro  $S^1 \times \mathbb{R}$*

(ii)  *$U$  é invariante por  $f$*

(iii) *existem números  $a < b$  tais que*

*$S^1 \times (-\infty, a] \subset U \subset S^1 \times [b, \infty)$*

(iv)  *$\text{int}(\bar{U}) = U$*

*Então  $\partial U$ , a fronteira de  $U$ , é o gráfico de uma função Lipschitz  $g : S^1 \rightarrow \mathbb{R}$*



**Corolário 5.** *Se  $\gamma$  é uma curva invariante rotacional, isto é, não homotopicamente trivial, por uma aplicação do tipo twist que preserva área então  $\gamma$  é o gráfico de uma função de Lipschitz.*

*Demonstração.* (do teorema 12) Observemos inicialmente que:

(a)  $U$  e  $\bar{U}$  são conexos

(b) O complementar  $U^c = S^1 \times \mathbb{R} - U$  é conexo

(c)  $\partial U$  é conexo (para a prova disso veja Newman [61]) e se uma curva simples fechada (de Jordan) contida em  $U$  limita uma região simplesmente conexa, então não há pontos de  $\partial U$  no interior da curva.

Provaremos inicialmente que a restrição da projeção canônica  $\pi_1 : S^1 \times \mathbb{R} \rightarrow S^1$  ao subconjunto  $\partial U$  é injetiva ( pois já é sobrejetiva).

Em seguida provamos que a função  $g(\theta) = \sup\{y \in \mathbb{R} \mid \{\theta\} \times (-\infty, y] \subset U\}$  é uma função de Lipschitz.

Vamos trabalhar no recobrimento universal  $\pi : \mathbb{R}^2 \rightarrow S^1 \times \mathbb{R}$  do cilindro e fixar um levantamento  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  de  $f$ .

Considere o subconjunto  $U_1 = \pi^{-1}(U)$ , aberto conexo e simplesmente conexo, invariante por  $F$  e invariante por translações horizontais ( $T(x, y) = (x, y) + (1, 0)$ ).

Seja  $V = \{(x, y) \mid (x, y_1) \in U_1, \forall y_1 \in [a, \leq y]\}$  subconjunto de pontos de  $U_1$  acessíveis por semi-retas verticais.

Então  $V$  é um aberto em  $\mathbb{R}^2$  conexo por caminhos. Provaremos que  $U_1 = V$  usando uma sequencia de lemas.

**Lema 17.** *Sejam  $I = [x_1, x_2]$  um intervalo e  $y \in \mathbb{R}$  tais que  $(x_1, y)$  e  $(x_2, y)$  pertencem a  $V$  e  $I \times \{y\} \subset U_1$ , então  $I \times [a, y] \subset U_1$ .*

*Demonstração.* Se  $D^i(x, y)$  denota a semi-reta vertical  $\{x\} \times (-\infty, y)$ . Considere a curva  $\Delta \subset U_1$  simples e fechada, formada por  $I \times \{a\} \cup I \times \{y\}$  e por segmentos verticais contidos em  $\cup D^i(x_1, y) \cup D^i(x_2, y)$  e acima de  $y = a$ . Então  $\Delta$  separa o aberto  $U_1$ .

Se  $\Delta$  contém um ponto da fronteira de  $U_1$  no seu interior, então, como  $\partial U_1$  é conexa, toda a fronteira de  $U_1$  está contida no interior de  $\Delta$ , o que é um absurdo.

Mas se interior de  $\Delta$  contém um ponto do complementar a  $U_1$  no então contém também um ponto da fronteira, o que, como acabamos de ver, não ocorre.

Logo, todo ponto do interior de  $\Delta$  está em  $U_1$ . Isto significa que  $I \times [a, y] \subset U_1$ .  $\square$

Denotemos por  $\partial_{U_1} V$  a fronteira de  $V$  em  $U_1$  ou seja, o subconjunto de pontos  $p$  de  $U_1$  tais que todo aberto que contém  $p$ , intersecta  $V$  e o seu complementar em  $U_1$ ,  $U_1 - V$ .

**Lema 18.**  $\partial_{U_1} V$  é a união disjunta (possivelmente enumerável) de segmentos verticais  $S_i$  com extremidades em  $\partial U$ .

Além disso, se  $p \in S_i$  então toda bola centrada em  $p$  é subdividida por  $S_i$  em dois abertos  $U_i^+$  e  $U_i^-$  um dos quais é disjuncto de  $V$ .

*Demonstração.* Para cada  $x \in \mathbb{R}$ , considere a interseção  $\{x\} \times [a, \infty) \cap U_1$  da semi-reta vertical, denotada por  $D^s(x)$ , que passa por  $x$  com o aberto  $U_1$ .

Seja  $S$  uma componente de  $D^s(x)$  que contenha um ponto  $(x, y) \in \partial_{U_1} V$ . Podemos escrever  $S = \{x\} \times (y_1, y_2)$  com  $(x, y_1)$  e  $(x, y_2)$  pertencentes a  $\partial U_1$  e  $V \cap S = \emptyset$ . Note que qualquer ponto de  $\partial_{U_1} V$  está contido em um intervalo  $S$

Provemos que  $S \subset \partial_{U_1} V$ . Para isso, basta provar que  $S \cap \partial_{U_1} V$  é aberto e fechado em  $S$

Tomemos uma vizinhança aberta de  $(x, y)$  em  $U_1$  da forma  $W = I \times (y - \epsilon, y + \epsilon)$  onde  $I$  é um intervalo aberto que contém  $x$ .

Observe que se  $(x', y') \in W \cap V$  então  $\{x'\} \times (y - \epsilon, y + \epsilon) \subset V$ .

Logo  $V \cup W = \cup \{x'\} \times (y - \epsilon, y + \epsilon)$  é a união de segmentos verticais e  $\{x\} \times (y - \epsilon, y + \epsilon) \subset \partial_{U_1}(V \cap W)$ . Portanto  $S \cap \partial_{U_1} V$  é aberto em  $S$ . Como é também fechado em  $S$ , concluímos que  $S \cap \partial_{U_1} V = S$ , ou seja  $S \subset \partial_{U_1} V$ , como queríamos demonstrar.

Observe agora que se  $(u, y)$  e  $(v, y)$  pertencem a  $W \cap V$ , com  $u < v$ , então pelo Lema anterior  $[u, v] \times (y - \epsilon, y + \epsilon) \subset V$ . Isto significa que se  $W$  contém pontos de  $V$  em ambos os lados de  $W - S$  então toda a faixa (incluindo os pontos de  $S$ ) estará contida em  $V$ , o que é um absurdo.

Portanto, existe um número  $\delta > 0$  pequeno tal que  $(x - \delta, x + \delta) \times (y - \epsilon, y + \epsilon) - [\{x\} \times (y - \epsilon, y + \epsilon)]$  possui duas componentes, uma contida em  $V$ , outra em  $U_1 - V$ .

Como existem no máximo uma quantidade enumerável de intervalos do tipo  $S \subset \partial_{U_1} V$  acima, concluímos a prova do Lema 18.  $\square$

Denotemos por  $S_i$  os "portões verticais": sub-intervalos verticais que formam o conjunto  $\partial_{U_1} V$ .

**Lema 19.** *Para todo  $i$ ,  $U_1 - S_i$  é a união de duas componentes conexas, uma das quais é disjunta de  $V$  e tem fronteira disjunta de  $S_j$ , para  $j \neq i$ .*

*Segue então que a componente de  $U_1 - S_i$ , disjunta de  $V$ , tem fronteira formada por  $S_i$  e um subconjunto da fronteira de  $U_1$  (uma "bolsa" inacessível por segmentos verticais contidos em  $U_1$ )*

*Demonstração.* Vimos, no Lema 18, que todo ponto  $p$  de  $S_i$  contém uma vizinhança  $Z_p$  tal que  $Z_p - S_i = V' \cup U'$  com  $V' \subset V$  e  $U' \subset U_1 - V$ .

Seja  $A_i$  a componente contida em  $U_1 - V$ .  $A_i$  é uma "bolsa" inacessível por verticais que situa-se à esquerda ou à direita de  $S_i$ .

Suponha, por absurdo, que exista um caminho em  $U_1$  iniciando em algum ponto de  $A_i$  e finalizando em um ponto no interior do segmento  $S_j$ , para  $j \neq i$  sem passar por nenhum portão  $S_k$ . Prolongando este caminho até o interior de  $S_i$ , construímos um caminho  $\gamma$  em  $U_1 - \cup_k S_k$  ligando  $S_i$  com  $S_j$ .

Prolongando um pouco mais  $\gamma$  para a esquerda ou para a direita, de modo a entrar no subconjunto  $V$  e finalmente, usando duas verticais, obtemos uma curva fechada  $\Gamma$  que liga dois pontos de  $S^1 \times \{a\}$ , passa pelos portões verticais  $S_i$  e  $S_j$  e é disjunta de  $\partial U_1$ .

Pelo Teorema de Jordan, esta curva separa o plano em duas componentes conexas.

Mas as extremidades dos segmentos  $S_i$  e  $S_j$ , que ficam em componentes distintas de  $\mathbb{R}^2 - \Gamma$  pertencem à fronteira  $\partial U_1$ , que é conexa. Obtemos assim uma contradição.

Com isso concluímos que a fronteira de cada bolsa  $A_i$  é a união de  $S_i$  com um pedaço da fronteira de  $U_1$ . □

Segue do Lema 18 que  $cl_{U_1} A_i = A_i \cup S_i$  e  $V \cap cl_{U_1} A_i = \emptyset$ .

Designemos por  $R_i$ ,  $i \in J$ , todas as componente conexas de  $U_1 - cl_{U_1} V$  que situam-se à direita de  $S_i$  e por  $E_i$  aquelas que se encontram à esquerda de  $S_i$ .

Observe que se  $R = \cup R_i$  e  $E = \cup E_i$  então,  $U_1 = cl_{U_1} R \cup cl_{U_1} E \cup V$ .

Vamos agora observar como se comportam os subconjuntos  $V, R$  e  $E$  em relação à aplicação  $F$ , que estamos supondo é do tipo twist para a direita.

**Lema 20.** *Nas hipóteses do Teorema de Birkhoff (Teorema 12), e usando a notação definida nos Lemas 18 e 19 temos:*

- (1)  $F(V) \cap cl_{U_1} E_k = \emptyset$
- (2)  $F^{-1}(V) \cap cl_{U_1} R_k = \emptyset$
- (3)  $F^{-1}(cl_{U_1} E_k) \cap V = \emptyset$
- (4)  $F^{-1}(cl_{U_1} E) \cap cl_{U_1} R = \emptyset$
- (5)  $F(cl_{U_1} R) \cap cl_{U_1} E = \emptyset$

*Demonstração.* (1) Se  $(x, y) \in V$  então, por definição  $(x, sy) \in U_1$ , para  $s \leq 1$ , e segue da invariância deste conjunto, que  $F(x, sy) \in U_1$ . Como  $F$  é do tipo twist para a direita, a curva  $F(x, sy)$  intersecta cada reta vertical no sentido da esquerda para a direita.

Mas, por definição as bolsas  $E_i$  estão sempre à esquerda de  $S_i$  e portanto, se houvesse um ponto de  $F(x, sy)$  em  $E_i$  então, usando uma construção semelhante à feita no Lema 4, isto é, conectando um ponto de  $E_i$  a um ponto de  $V$ , obteríamos uma curva que separa pontos da fronteira de  $U_1$ , o que é um absurdo.

O mesmo é válido para a intersecção da imagem de  $V$  com  $cl_{U_1} E_k$

A prova de (2) segue de modo análogo aplicando a  $F^{-1}$ , que é do tipo twist para a esquerda.

A prova de (3) é imediata a partir de (1), uma vez que se  $(x, y) \in F^{-1}(cl_{U_1} E_k) \cap V$  então  $F(x, y) \in F(V) \cap cl_{U_1} E_k = \emptyset$ .

(4) Se  $F^{-1}(cl_{U_1} E) \cap cl_{U_1} R \neq \emptyset$  então  $F^{-1}(cl_{U_1} E_k) \cap cl_{U_1} R_i \neq \emptyset$  para algum par  $i, k$ .

Como as bolsas  $cl_{U_1} A_i$  são disjuntas duas a duas e conexas, em vista do item (3) do Lema acima, devemos ter  $F^{-1}(cl_{U_1} E_k) \subset cl_{U_1} R_i$ .

Porém,  $F^{-1}(S_k)$  intersecta  $S_i$  transversalmente, no sentido da direita para a esquerda. Portanto,  $F^{-1}(S_k) \cap R_i \neq \emptyset$  e  $F^{-1}(S_k) \cap V \neq \emptyset$ .

Pela monotonicidade de  $F^{-1}$  restrita a cada segmento vertical, vemos isso implica que  $F^{-1}(V) \cap R_j \neq \emptyset$ . O que é um absurdo.

Logo  $F^{-1}(cl_{U_1} E) \cap cl_{U_1} R = \emptyset$  isto é,  $F^{-1}(cl_{U_1} E) \subset E$

Analogamente obtem-se que  $F((cl_{U_1} R) \cap cl_{U_1} E) = \emptyset$ , concluindo a prova do Lema 20.  $\square$

Usando a invariância de  $U_1$  e a decomposição  $U_1 = cl_{U_1} R \cup cl_{U_1} E \cup V$ , obtemos, em particular que

$$F^{-1}(cl_{U_1} E) \subset E \text{ e } F(cl_{U_1} R) \subset R$$

Mas se isso ocorresse, então o aberto  $F^{-1}(R - F(cl_{U_1}R))$  seria um subconjunto errante, contradizendo o Teorema de Recorrência de Poincaré (veja [10]).

De fato, se  $p \in F^{-1}(R - F(cl_{U_1}R))$  então, como  $F$  preserva área, o Teorema de Poincaré nos diz que para toda bola aberta  $B_p$  contendo  $p$ , existe um inteiro  $n$  tal que  $F^n(B_p) \cap B_p \neq \emptyset$ .

Tomemos  $B_p \subset F^{-1}(R - F(cl_{U_1}R))$  de modo que  $F(B_p) \subset (R - F(cl_{U_1}R))$ ,  $F^2(B_p) \subset (F(cl_{U_1}R) - F^2(cl_{U_1}R))$ , e assim por diante, até  $n$  para chegar à conclusão de que  $F^n(B_p) \subset F^n(cl_{U_1}R) \subset F(cl_{U_1}R) \subset R$

Logo se  $q \in F^n[B_p] \cap B_p$ , então,  $F(q) \in F(cl_{U_1}R) \subset R$  ao mesmo tempo que  $F(q) \in R - F(cl_{U_1}R)$ . Uma contradição.

Podemos desta forma concluir que  $R = \emptyset$  e analogamente  $E = \emptyset$ .

Segue-se que  $U = V$ , ou seja, todo ponto de  $U$  é acessível por retas verticais.

Para concluir a prova do Teorema de da Curva Invariante de Birkhoff, basta provar que  $\partial U_1$  não contém segmento vertical.

Mas isto segue da condição de twist pois se  $S \subset \partial U_1$  é um segmento vertical, então pontos de  $U_1$  à direita (ou à esquerda) de  $U_1$  seriam levados por  $F$  em pontos de  $U$  não acessíveis por verticais. Um absurdo.

Concluimos assim que  $\partial U_1$  é gráfico de uma função contínua  $\phi_1(x) = \sup\{y | (-\infty, y) \subset U_1\}$ , tal que  $\phi_1(x+1) = \phi_1(x)$ .  $\square$

**Lema 21.**  $\phi_1(x)$  é uma função de Lipschitz.

*Demonstração.* Suponhamos que  $F$  seja do tipo twist para a direita, de modo que existe um número  $\delta$  tal que  $\partial_2 F_1 > \delta$  em uma vizinhança de  $\partial U_1 = \text{graf}(\phi_1)$ .

Tomemos um par de pontos  $(x_0, \phi_1(x_0))$  e  $(x'_0, \phi_1(x'_0))$  com  $x_0 < x'_0$ .

Se  $\phi_1(x_0) > \phi_1(x'_0)$ , então

$$F_1(x'_0, \phi_1(x_0)) - F_1(x'_0, \phi_1(x'_0)) > \delta[\phi_1(x_0) - \phi_1(x'_0)].$$

Mas pela continuidade da derivada e da função  $\phi_1$ , existe uma constante positiva  $L$  tal que

$$F_1(x'_0, \phi_1(x_0)) - F_1(x_0, \phi_1(x_0)) < L[x'_0 - x_0]$$

Usando a monotonicidade da função  $x \rightarrow F_1(x, \phi_1(x))$  temos  $F_1(x'_0, \phi_1(x'_0)) > F_1(x_0, \phi_1(x_0))$ .

Juntando todas essas desigualdades obtemos:

$$\begin{aligned} F_1(x'_0, \phi_1(x_0)) - L[x'_0 - x_0] &< F_1(x_0, \phi_1(x_0)) < F_1(x'_0, \phi_1(x'_0)) \\ &< F_1(x'_0, \phi_1(x_0)) - \delta[\phi_1(x_0) - \phi_1(x'_0)]. \end{aligned}$$

Isto é:

$$[\phi_1(x_0) - \phi_1(x'_0)] < L\delta^{-1}[x'_0 - x_0]$$

Se  $\phi_1(x_0) < \phi_1(x'_0)$ , procedemos analogamente usando  $F^{-1}$  para obter

$$[\phi_1(x'_0) - \phi_1(x_0)] < L'\delta'^{-1}[x'_0 - x_0],$$

para outras constantes  $L'$  e  $\delta'$ , concluindo assim a prova do Lema.  $\square$

Usando a periodicidade de  $\phi_1$  obtemos que  $\partial U$  também coincide com o gráfico de uma função de Lipschitz,  $\phi$ , definida em  $S^1$ .

Com isso, fica concluída a prova do teorema 12  $\square$

## 3.2 Propriedade variacional das curvas invariantes

No Capítulo 2, utilizamos a função geratriz (ou princípio variacional) para obter órbitas periódicas e os conjuntos de Aubry-Mather.

Provaremos nesta seção que as órbitas contidas em curvas rotacionais invariantes estão associadas a configurações minimizantes.

**Proposição 11.** *Seja  $\Gamma$  uma curva invariante rotacional por uma aplicação do tipo twist que preserva área  $f : S^1 \times \mathbb{R} \rightarrow S^1 \times \mathbb{R}$ .*

*Então toda órbita em  $\Gamma$  é minimizante global, isto é, se  $(x_i, y_i) = f(x_{i-1}, y_{i-1})$ ,  $i \in \mathbb{Z}$  é uma órbita contida em  $\Gamma$  e  $h$  é a função geratriz de  $f$ , então dados quaisquer dois números inteiros  $n < m$*

e uma sequência qualquer  $z_n, z_{n+1}, \dots, z_m$  com  $z_n = x_n$  e  $z_m = x_m$  vale:

$$W_{nm}(z_n, \dots, z_m) = \sum_{i=n}^{m-1} h(z_i, z_{i+1}) \geq \sum_{i=n}^{m-1} h(x_i, x_{i+1}) = W_{nm}(x_n, \dots, x_m)$$

*Demonstração.* Pelo Teorema da Curva Invariante de Birkhoff, existe uma função de Lipschitz  $\phi : S^1 \rightarrow \mathbb{R}$  tal que  $\Gamma = \text{graf}(\phi) = \{(x, \phi(x))\}$

Seja  $h^*(x, x') = h(x, x') - \int_x^{x'} \phi(t)dt$ , onde  $h$  é a função geratriz da aplicação  $f$ .

Observe que pela invariância de  $\Gamma$  obtemos uma função  $g(x) = \pi_1 \circ f(x, \phi(x))$  ( $\pi_1$  é a projeção no primeiro fator) que é um homeomorfismo de Lipschitz.

Portanto  $f(x, \phi(x)) = (g(x), \phi(g(x)))$ .

Logo,  $h^*(x, g(x)) = h(x, g(x)) - \int_x^{g(x)} \phi(t)dt$ . Mas o Teorema de Rademacher nos diz que  $\phi$ , e portanto  $g$  é diferenciável exceto num conjunto de medida de Lebesgue zero.

Portanto, tomando a derivada total na expressão acima temos:

$$\frac{d}{dx} h^*(x, g(x)) = \partial_1 h(x, g(x)) + \partial_2 h(x, g(x))g'(x) - \phi(g(x))g'(x) + \phi(x) \quad (11)$$

Mas pela definição de função geratriz,  $f(x, \phi(x)) = (g(x), \phi(g(x)))$  se e somente se  $\phi(g(x)) = \partial_2 h(x, g(x))$  e  $\phi(x) = -\partial_1 h(x, g(x))$ .

Susbtituindo estas expressões em (11) temos

$$\frac{d}{dx} h^*(x, g(x)) = -\phi(x) + \phi(g(x))g'(x) - \phi(g(x))g'(x) + \phi(x) = 0$$

Ou seja  $h^*(x, g(x))$  é uma função de Lipschitz com derivada nula (onde existir).

Logo,  $h^*(x, g(x)) = \text{constante} = c$ .

Usemos agora a condição de twist que, nas nossas hipóteses, implica  $\partial_{12}h < 0$ , para obter que fixado  $x \in \mathbb{R}$  a aplicação  $x' \rightarrow h^*(x, x')$  tem valor mínimo igual a  $h^*(x, g(x))$ .

De fato,  $g(x)$  é um ponto crítico pois  $\partial_2 h^*(x, g(x)) = \partial_2 h(x, g(x)) - \phi(g(x)) = 0$ .

Além disso, como  $\partial_{12}h^*(x, x') = \partial_{12}h(x, x') < 0$ , a função  $x' \rightarrow \partial_1 h^*(x, x')$  é decrescente.

Portanto,  $x' > g(x)$  implica  $\partial_1 h^*(x, x') < \partial_1 h^*(x, g(x)) = 0$  e  $x' < g(x)$  implica  $\partial_1 h^*(x, x') > \partial_1 h^*(x, g(x)) = 0$ .

Ou seja, para  $(x, x')$  acima (abaixo) do gráfico de  $g$ ,  $\partial_1 h^*(x, x') < 0$  (respectivamente  $\partial_1 h^*(x, x') > 0$ ). Isso claramente implica que o mínimo de  $x \rightarrow h^*(x, x')$  ocorre exatamente quando  $x = g^{-1}(x')$ . Analogamente, a função  $x \rightarrow \partial_2 h^*(x, x')$  é decrescente. ou seja,  $x' > g(x)$  implica  $\partial_2 h^*(x, x') < \partial_2 h^*(x, g(x))$  e  $x' < g(x)$  implica  $\partial_2 h^*(x, x') > \partial_2 h^*(x, g(x))$ .

Geometricamente isto significa que o vetor gradiente  $\nabla h^*(x, x')$  aponta para fora do gráfico de  $g$ .

Portanto  $c = h^*(x, g(x)) \leq h^*(x, x')$  para todo par  $(x, x')$ .

Dados dois números inteiros  $n < m$  considere uma seqüência qualquer  $z_n, z_{n+1}, \dots, z_m$  com  $z_n = x_n$  e  $z_m = x_m$ , onde, como no enunciado da proposição, a seqüência  $(x_i, \phi(x_i))$  corresponde a uma órbita em  $\Gamma$ .

Então

$$\begin{aligned} W_{nm}(z_n, \dots, z_m) &= \sum_n^{m-1} h(z_i, z_{i+1}) = \\ &= \sum_{i=n}^{m-1} h^*(z_i, z_{i+1}) + \sum_{i=n}^{m-1} \left[ \int_{z_i}^{z_{i+1}} \phi(t) dt \right] = \\ &= \sum_{i=n}^{m-1} h^*(z_i, z_{i+1}) + \int_{x_n}^{x_m} \phi(t) dt \end{aligned}$$

A segunda parcela só depende das extremidades.

Mas  $h^*(x, x') \geq h^*(x, g(x)) = c = h^*(x_i, x_{i+1})$ , donde concluímos que

$$W_{nm}(z_n, z_{n+1}, \dots, z_m) \geq (m - n)c + \int_{x_n}^{x_m} \phi(t) dt =$$

$$\sum_{i=n}^{m-1} h^*(x_i, x_{i+1}) + \int_{x_n}^{x_m} \phi(t) dt = W_{nm}(x_n, x_{n+1}, \dots, x_m)$$

como queríamos demonstrar.  $\square$



### 3.3 Inexistência de Curvas Invariantes

Vejam algumas aplicações do Teorema da Curva Invariante de Birkhoff. A referência para esta seção é [39].

**Proposição 12.** *Suponha que  $F$  seja uma aplicação do tipo twist e  $z$  um ponto tal que*

$$\limsup \frac{\pi_1(F^k(z))}{k} = b$$

$$\liminf \frac{\pi_1(F^k(z))}{k} = a \text{ com } a < b$$

*Então não existe curva invariante  $\Gamma$  cujo número de rotação  $\rho(\Gamma)$  satisfaz*

$$a < \rho(\Gamma) < b$$

*Demonstração.* Escreva  $z = (x_0, y_0)$  e seja  $\Gamma$  uma curva invariante que é o gráfico de uma função periódica de Lipschitz,  $\phi$ .

Suponha que  $z$  esteja abaixo da curva  $\Gamma$  isto é,  $y_0 < \phi(x_0)$ . Seja  $C_0 = (x_0, \phi(x_0))$ .

Usando a condição de twist  $\delta$ -uniforme em um compacto ( $\partial_2(\pi_1 \circ F) > \delta$ ) e o Teorema Fundamental do Cálculo aplicado à função  $\pi_1(F(x_0, (1-\lambda)y_0 + \lambda\phi(x_0)))$  obtemos

$$\pi_1 \circ F(C_0) - \pi_1 \circ F(z) = \int_0^1 \partial_2 \pi_1(F(x_0, (1-\lambda)y_0 + \lambda\phi(x_0))) (\phi_0(x_0) - y_0) d\lambda$$

$$\pi_1 \circ F(C_0) - \pi_1 \circ F(z) > \delta(\phi(x_0) - y_0)$$

Defina  $x_1 = \pi_1 \circ F(z)$ ,  $C_1 = (x_1, \phi(x_1))$  e, por indução  $x_k = \pi_1 \circ F^k(z)$ ,  $C_k = (x_k, \phi(x_k))$  de modo que os pontos  $C_k$  e  $F^k(z)$  estão sobre a mesma reta vertical.

Aplicando o mesmo raciocínio acima obtemos

$$\pi_1 \circ F(C_k) - \pi_1 \circ F^k(z) > 0$$

Mas o homeomorfismo  $g(x) = \pi_1 \circ F(x, \phi(x))$  definido pela restrição de  $F$  à curva  $\Gamma$ , é crescente. Logo,  $g(x_0) > x_1$  implica  $g^2(x_0) > g(x_1)$ .

E  $x_2 < g(x_1) < g^2(x_0)$  e assim por diante, provamos por indução que  $x_k < g^k(x_0)$ :

$$x_{k+1} = \pi_1 \circ F^{k+1}(z) < \pi_1 \circ F(C_{k+1}) = g(x_k) < g^{k+1}(x_0)$$

Concluimos assim que

$$\frac{x_k}{k} < \frac{g^k(x_0)}{k}$$

O que implica

$$\limsup \frac{\pi_1(F^k(z))}{k} \leq \rho(\Gamma)$$

Analogamente se o ponto  $z$  estiver acima da curva invariante  $\Gamma$ , obteremos

$$\rho(\Gamma) \leq \liminf \frac{\pi_1(F^k(z))}{k}$$

□

**Corolário 6.** *Suponha que  $\phi_1 < \phi_2$  sejam duas funções definidas no círculo cujos gráficos  $\Gamma_1$  e  $\Gamma_2$  sejam curvas invariantes rotacionais por um aplicação  $F$  do tipo twist que preserva área.*

*Se  $\rho(\Gamma_1)$  e  $\rho(\Gamma_2)$  denotam os respectivos números de rotação, então  $\rho(\Gamma_1) < \rho(\Gamma_2)$*

*Demonstração.* Segue imediatamente da Proposição 12 que  $\rho(\Gamma_1) \leq \rho(\Gamma_2)$ . Vejamos como usar a condição de preservar área para obter a desigualdade estrita.

*Caso A: Número de rotação irracional*

Se  $\rho(\Gamma_1)$  ou  $\rho(\Gamma_2)$  é irracional, então  $\rho(\Gamma_1) < \rho(\Gamma_2)$ .

Suponha que  $\rho(\Gamma_1)$  seja irracional. Sejam  $g_1$  e  $g_2$  os homeomorfismos crescentes definidos pela restrição de  $F$  a  $\Gamma_1$  e  $\Gamma_2$  respectivamente. É claro que  $\rho(g_i) = \rho(\Gamma_i)$ , para  $i = 1, 2$ . Além disso, como estamos supondo que  $F$  é do tipo twist, a hipótese  $\phi_1(x) < \phi_2(x)$  implica a desigualdade  $g_1(x) < g_2(x)$ .

Seja  $m > 0$  tal que para todo  $x \in \mathbb{R}$ ,  $g_2(x) \geq g_1(x) + m$ . Tomemos um convergente ímpar por frações contínuas de  $\rho(g_1)$ ,  $p/q \in \mathbb{Q}$  tal que  $1/q < m$ . Então,  $\frac{p}{q} - \frac{m}{q} < \rho(g_1) < \frac{p}{q}$ .

Observe que existe um ponto  $x_0$  tal que  $g_1^q(x_0) - x_0 > p - m$ . Caso contrário, para todo  $x \in \mathbb{R}$  teríamos  $g_1^q(x) - x \leq p - m$ , o que implicaria, usando indução e uma soma telescópica,

$$g_1^{nq}(x) - x \leq np - nm$$

Dividindo tudo por  $n$  e tomando o limite quando  $n \rightarrow \infty$  teríamos uma contradição

$$\rho(g_1) \leq \frac{p}{q} - \frac{m}{q}$$

Portanto, segue por indução e da definição do número  $m$ , que  $g_2^q(x_0) > g_1^q(x_0) + m > p + m$  pois, para todo  $k$  temos que  $g_2^k(x) > g_1^k(x)$  implica  $g_2^{k+1}(x) > g_1(g_2^k(x)) + m > g_1^{k+1}(x) + m$ .

Logo,  $g_2^q(x_0) > g_1(g_2^{q-1}(x_0)) + m > g_1^q(x_0) + m > p + m$

Usando o mesmo argumento anterior, isso implica que  $\rho(g_2) \geq p/q > \rho(g_1)$  como queríamos demonstrar. Se  $\rho(g_2)$  for irracional, procedemos analogamente. Observe que até agora apenas usamos a condição de twist.

Resta provar o caso onde os dois números de rotação são racionais. Aqui, usa-se que  $F$  preserva área.

*Caso B: Número de rotação racional*

Suponha que  $\rho(\Gamma_1) = \rho(\Gamma_2) = p/q$  e que  $\Gamma_1 \neq \Gamma_2$ . A idéia é contruir um quadrilátero  $Q$  de área não nula, limitado por curvas contínuas, que contém estritamente a sua imagem por uma aplicação que preserva área. Isso é obviamente impossível.

Escreva, como acima,  $\Gamma_i = \text{graf}(\phi_i)$ , com  $\phi_1 < \phi_2$ . Nesse caso, existe um ponto  $(x_1, \phi_1(x_1))$  tal que  $F^q(x_1, \phi_1(x_1)) = (x_1, \phi_1(x_1)) + (p, 0)$ .

Ou melhor,  $z_1 = (x_1, \phi_1(x_1))$  é um ponto fixo da aplicação  $T_{-p} \circ F^q$ , onde  $T_{-p}(x, y) = (x, y) - (p, 0)$  é a translação.

Seja  $z_2 = (x_2, \phi_1(x_2))$  o primeiro ponto fixo de  $T_{-p} \circ F^q$  situado em  $\Gamma_2$  à direita de  $z_1$ . Não está descartada à priori a coincidência  $x_1 = x_2$ .

O quadrilátero  $Q$  é limitado pelas seguintes curvas:

os segmentos verticais  $V_1 =: [(x_1, \phi_1(x_1)); (x_1, \phi_2(x_1))]$  e

$V_2 =: [(x_2, \phi_1(x_2)); (x_2, \phi_2(x_2))]$

o arco em  $\Gamma_1$  entre  $(x_1, \phi_1(x_1))$  e  $(x_2, \phi_1(x_2))$

o arco em  $\Gamma_2$  entre  $(x_1, \phi_2(x_1))$  e  $(x_2, \phi_2(x_2))$ .

Provemos então que a imagem de  $Q$  por  $T_{-p} \circ F^q$  está contida em  $Q$ . Para ver isso, basta verificar as imagens dos segmentos verticais  $V_i$  são curvas inteiramente contidas no interior de  $Q$ . De fato, a imagem

de  $V_1$  é um *caminho negativo* (veja Cap. 4), portanto, está sempre à direita de  $V_1$ , mais precisamente, se  $\gamma(t)$  é uma parametrização de  $T_{-p} \circ F^q(V_1)$  com  $\gamma(0) = z_1$  então  $\pi_1(\gamma(t)) > x_1$ , para todo  $t \neq 0$ .

Além disso,  $x_1 < \pi_1 \circ T_{-p} \circ F^q(x_1, \phi_2(x_1)) < x_2$ .

Analogamente, a imagem de  $V_2$  é também um *caminho negativo* que termina no ponto fixo  $z_2$  e que está inteiramente à esquerda de  $V_2$  e inicia em  $T_{-p} \circ F^q(x_2, \phi_1(x_2))$ .

Pela definição de número de rotação, e pela invariância por  $F$ , conclui-se que os arcos em  $\Gamma_i$ ,  $i = 1, 2$  são invariantes por  $T_{-p} \circ F^q$ . Dessa forma, conclui-se que imagem de  $Q$  está inteiramente contida em  $Q$ , absurdo. No caso em que  $x_1 = x_2$ , também chegamos a um absurdo, pois a imagem de uma vertical por  $T_{-p} \circ F^q$  é um *caminho negativo*, portanto um arco vertical não pode ser invariante.  $\square$

### 3.4 Cáusticas e curvas invariantes em bilhares convexos

No Capítulo 1, vimos o exemplo do bilhar elíptico, cujo espaço de fase é totalmente folheado por curvas invariantes, algumas rotacionais outras homotopicamente triviais. Como vimos, as curvas invariantes rotacionais estão associadas aos raios que intersectam o eixo maior da elipse em pontos fora do segmento entre os focos.

As curvas homotopicamente triviais estão associadas aos raios que intersectam o segmento entre os focos.

Separando estes dois subconjuntos de curvas, temos os raios que passam pelos focos, e que correspondem a duas curvas invariantes rotacionais que se intersectam na órbita periódica de período dois correspondente ao diâmetro (eixo maior).

As curvas invariantes rotacionais definem uma família de raios que tangenciam uma elipse confocal com a elipse bordo do bilhar. Esta bela propriedade, é válida em geral para os bilhares convexos. Em outras palavras, cada curva rotacional invariante define uma família de raios que tangenciam uma curva, denominada *cáustica*, no interior do bilhar. Esta seção, que é baseada em [27] e [64], descreve algumas propriedades das cáusticas nos bilhares.

**Definição 12.** *Uma cáustica no bilhar definido por uma curva  $\Gamma$  é*

uma curva  $\gamma$  situada no interior de  $\Gamma$  tal que todo raio tangente a  $\gamma$  se reflete em  $\Gamma$  em um raio tangente a  $\gamma$ .

**Proposição 13.** *A cada curva rotacional invariante e diferenciável, contida no anel  $S^1 \times (0, \frac{\pi}{2})$  ou em  $S^1 \times (\frac{\pi}{2}, \pi)$ , está associada uma cáustica no interior do bilhar.*

*Demonstração.* Seja  $\Lambda$  uma curva rotacional invariante que é o gráfico de uma função diferenciável  $\varphi$  definida no bordo do bilhar  $\Gamma$ . Se  $\alpha(s)$  é a parametrização de  $\Gamma$  por comprimento de arco, então  $\Lambda = (s, \varphi(s))$ . Vamos supor  $0 < \varphi(s) < \frac{\pi}{2}$  de classe  $C^1$ .

Seja  $v_\varphi(s) = \cos(\varphi(s))\alpha'(s) + \text{sen}(\varphi(s))\eta(s)$ , onde  $\eta(s)$  é o vetor normal a  $\Gamma$  no ponto  $\alpha(s)$ .  $v_\varphi(s)$  é um campo de vetores unitário que faz ângulo  $\varphi(s)$  com a tangente a  $\Gamma$ .

Considere a família de raios definida por  $r(s, \lambda) = \alpha(s) + \lambda v_\varphi(s)$ . Queremos calcular a envoltória desta família ou seja queremos encontrar uma função  $\lambda(s)$  tal que a curva  $\beta(s) = \alpha(s) + \lambda(s)v_\varphi(s)$  é tangente ao raio  $r(s, \lambda)$  no ponto  $\beta(s)$ .

Mas, usando as fórmulas de Frenet,  $\alpha'(s) = k(s)\eta(s)$  e  $\eta'(s) = -k(s)\alpha'(s)$  temos:

$$\beta'(s) = \alpha'(s) + \lambda'(s)\eta(s) + \lambda(s)[k(s) + \varphi'(s)]v_\varphi^\perp(s),$$

onde  $v_\varphi^\perp(s) = -\text{sen}(\varphi(s))\alpha'(s) + \cos(\varphi(s))\eta(s)$ .

A condição de tangência implica

$$\langle \beta'(s), v_\varphi^\perp(s) \rangle = 0$$

ou seja:

$$-\text{sen}(\varphi(s)) + \lambda(s)[k(s) + \varphi'(s)] = 0$$

Esta é a chamada condição de focalização de família de raios.

Se  $[k(s) + \varphi'(s)] \neq 0$ , então obtemos uma cáustica associada à curva invariante  $\beta(s) = \alpha(s) + \frac{\text{sen}(\varphi(s))}{[k(s) + \varphi'(s)]}v_\varphi(s)$ .

Se  $[k(s_0) + \varphi'(s_0)] = 0$ , então a condição de focalização implica que  $\text{sen}(\varphi(s_0)) = 0$ , ou seja  $(s_0, \varphi(s_0))$  é um ponto da fronteira do anel,  $\varphi(s_0) = 0$  ou  $\pi$ , contradizendo a hipótese.  $\square$

Assim como as evolutas das curvas planas, em geral as cáusticas de bilhares podem não ser curvas regulares. Entretanto, é fácil descrever condições para que se a curva invariante estiver suficientemente próxima do bordo do anel então a cáustica é uma curva regular.

Observe também que a aplicação do bilhar é invariante pela reflexão em torno da reta  $\theta = \frac{\pi}{2}$ . Isto implica que se uma curva invariante cruza esta reta então a cáustica pode não ser regular.

Vejam novamente o caso do caso do bilhar na elipse: seja  $O$  a órbita periódica, de período dois, correspondente às extremidades do eixo maior. Esta órbita é hiperbólica e as variedades estável e instável,  $W^s(O)$  e  $W^u(O)$ , respectivamente, são duas curvas invariantes que contém  $O$ . A cáustica associada a essas curvas coincide com os dois focos da elipse.

A próxima proposição trata de uma das propriedades interessantes das cáusticas. Ela nos ensina a desenhar a fronteira  $\Gamma$  de um bilhar a partir de uma curva convexa  $C$  dada, de modo que  $C$  seja uma cáustica. Para isso, faça um laço, com folga, em torno de  $C$ , estique-o com um lápis e faça-o deslizar, bem esticado, em torno da curva. A curva traçada  $\Gamma$  é o bordo de um bilhar com cáustica  $C$ .

**Proposição 14. (O parâmetro de Lazutkin)** *Seja  $C$  uma cáustica de um bilhar convexo com fronteira  $\Gamma$  ambas orientadas positivamente (sentido anti-horário).*

*Dado um ponto  $y \in C$ , considere o raio positivo  $r_y$  tangente a  $C$  em  $y$ . Sejam  $x \in r_y \cap \Gamma$  e  $z \in C$  o ponto da cáustica tangente ao raio refletido em  $x$ .*

*Se  $m(y, z) =$  comprimento de arco em  $C$  entre  $y$  e  $z$ ,  $|y - x| = \text{dist}(y, x)$ ,  $|x - z| = \text{dist}(x, z)$  então:*

$$|y - x| + |x - z| - m(y, z) = \text{constante}$$

*Chama-se parâmetro de Lazutkin a constante associada à cáustica.*

*Demonstração.* Utilizando a notação da proposição anterior, seja  $\alpha(s)$  a parametrização do bordo  $\Gamma$ . Uma parametrização de  $C$  é dada por  $\gamma_+(s) = \alpha(s) + \lambda(s)v(s)$ , onde  $v(s)$  é um vetor unitário.

Portanto, como  $\gamma_+'(s)$  é paralelo ao vetor unitário  $v(s)$  e se  $s_1 < s_2$ , temos

$$m(\gamma_+(s_1), \gamma_+(s_2)) = \int_{s_1}^{s_2} |\gamma'_+(s)| ds =$$

$$\lambda(s_2) - \lambda(s_1) + \int_{s_1}^{s_2} \langle \alpha'(s), v(s) \rangle ds$$

Por outro lado, a cáustica pode ser parametrizada no sentido contrário,  $\gamma_-(s) = \alpha(s) + \delta(s)\hat{v}(s)$ , usando-se o raio refletido  $\hat{v}(s)$  tal que  $\langle \alpha'(s), v(s) \rangle = -\langle \alpha'(s), \hat{v}(s) \rangle$  e  $\langle \eta(s), v(s) \rangle = \langle \eta(s), \hat{v}(s) \rangle$ .

Portanto,

$$m(\gamma_-(s_1), \gamma_+(s_1)) - m(\gamma_-(s_2), \gamma_+(s_2)) =$$

$$m(\gamma_-(s_1), \gamma_-(s_2)) + m(\gamma_-(s_2), \gamma_+(s_1))$$

$$- m(\gamma_-(s_2), \gamma_+(s_1)) - m(\gamma_+(s_1), \gamma_+(s_2))$$

Mas

$$m(\gamma_+(s_1), \gamma_+(s_2)) = \lambda(s_2) - \lambda(s_1) + \int_{s_1}^{s_2} \langle \alpha'(s), v(s) \rangle ds$$

$$m(\gamma_-(s_2), \gamma_-(s_1)) = \delta(s_1) - \delta(s_2) + \int_{s_2}^{s_1} \langle \alpha'(s), \hat{v}(s) \rangle ds$$

E como,  $m(\gamma_-(s_2), \gamma_-(s_1)) = m(\gamma_-(s_1), \gamma_-(s_2))$  e  $\langle \alpha'(s), v(s) \rangle = -\langle \alpha'(s), \hat{v}(s) \rangle$ , obtemos:

$$m(\gamma_-(s_1), \gamma_+(s_1)) - m(\gamma_-(s_2), \gamma_+(s_2)) =$$

$$\lambda(s_1) - \lambda(s_2) + \delta(s_1) - \delta(s_2)$$

Isto é:

$$\lambda(s_2) + \delta(s_2) - m(\gamma_-(s_2), \gamma_+(s_2)) = \lambda(s_1) + \delta(s_1) - m(\gamma_-(s_1), \gamma_+(s_1))$$

Para concluir a prova da proposição, basta usar o fato de que  $\lambda(s_i) = |\gamma_+(s_i) - \alpha(s_i)|$  e  $\delta(s_i) = |\gamma_-(s_i) - \alpha(s_i)|$ .  $\square$

O próximo teorema é conseqüência do Teorema da Curva Invariante de Moser. De acordo com R. Douady, este teorema foi provado

inicialmente exigindo-se um alto grau de diferenciabilidade para a curva do bordo do bilhar, esta versão encontra-se em [27]. Na seção 3.6, provaremos uma versão mais simplificada do teorema de Lazutkin.

**Teorema 13. (Lazutkin)** *Se a curva do bilhar for estritamente convexa (curvatura positiva) e suficientemente diferenciável (de classe  $C^k, k \geq 6$ ), então existe uma família de cáusticas numa vizinhança do bordo do bilhar, cuja união tem área positiva.*

### 3.5 Bilhares sem curvas invariantes

Nesta seção, mostraremos que para um bilhar em uma região convexa limitada do plano, cujo bordo é uma curva  $C^2$  com pelo menos um ponto de curvatura zero, não existem cáusticas. Ou seja, o twist mapping associado não possui curvas rotacionais invariantes. Essa análise está baseada no trabalho [54].

Considerando agora esse resultado juntamente com o teorema 19, obtemos:

**Teorema 14.** *Seja  $\Gamma$  uma curva  $C^2$  com pelo menos um ponto de curvatura zero, tal que a componente conexa limitada do seu complementar é convexa. Então o bilhar a ela associado possui a seguinte propriedade:*

*Existe  $(p_0, \phi_0) \in \Gamma \times ]0, \pi[$  tal que se  $(p_n, \phi_n)$  denota a órbita de  $(p_0, \phi_0)$ , então  $\phi_n \xrightarrow{n \rightarrow \infty} 0$  e  $\phi_n \xrightarrow{n \rightarrow -\infty} \pi$ .*

Como já foi dito, é natural definir a aplicação do bilhar no anel  $(s, r = \cos(\phi)) \in S^1 \times [-1, 1]$ , onde  $p = \Gamma(s)$ . O que o teorema diz, é que existe uma órbita cujo  $\omega$ -limite está contido no bordo superior e o  $\alpha$ -limite está contido no bordo inferior. Geometricamente, do ponto de vista da bolinha quicando na curva, esta condição é bastante surpreendente. Convidamos o leitor a fazer alguns desenhos para entendê-la corretamente.

A prova do teorema consiste em verificar que a existência de pelo menos um ponto de curvatura zero implica na não existência de curvas rotacionais invariantes. Isto será feito exibindo uma condição geral sobre as funções geratrizes de aplicações do tipo twist.



**Lema 22.** *Seja  $f : S^1 \times [-1, 1] \rightarrow S^1 \times [-1, 1]$  um difeomorfismo  $C^1$  do tipo twist, com função geratriz  $h$ . Se para algum  $x$  real,  $\partial_{22}h(x^*, x) + \partial_{11}h(x, x^\#) > 0$  para quaisquer  $x^*, x^\#$ , então  $f$  não possui curvas rotacionais invariantes.*

*Demonstração.* Para que esse resultado possa ser mais facilmente aplicado ao problema do bilhar, faremos as seguintes convenções:

- $\partial_{12}h > 0$  em todo ponto
- se  $(x_1, y_1) = f(x, y)$ , então  $y = -\partial_1h(x, x_1)$  e  $y_1 = \partial_2h(x, x_1)$

Por absurdo, vamos supor que existe uma curva rotacional invariante contida em  $S^1 \times (-1, 1)$ . Então, pelo teorema da curva invariante de Birkhoff, temos que ela coincide com o gráfico de uma certa função Lipschitz  $\gamma : S^1 \rightarrow (-1, 1)$ . Assim,  $f(x, \gamma(x)) = (s(x), \gamma(s(x)))$ , onde  $s(x)$  é um homeomorfismo do círculo. Como  $f$  é pelo menos  $C^1$ , preserva orientação (consequência trivial de  $\det[Df] \equiv 1 > 0$ ) e  $\gamma$  é Lipschitz, segue que  $s$  também é Lipschitz e  $s(x+1) = s(x) + 1$ . Fixado  $x$ , vamos definir  $x_1 = s(x)$  e  $x_{-1} = s^{-1}(x)$ . Do fato do gráfico de  $\gamma$  ser invariante, segue que

$$\partial_2h(x_{-1}, x) + \partial_1h(x, x_1) = 0, \text{ para todo } x \in S^1.$$

Vamos agora lembrar que apesar de uma função Lipschitz não ser necessariamente diferenciável em todo ponto, ela é diferenciável em quase todo ponto (q.t.p.), com respeito a medida de Lebesgue. Assim, podemos escrever, para quase todo  $x$ :

$$\partial_{12}h(x_{-1}, x) \frac{ds^{-1}(x)}{dx} + \partial_{22}h(x_{-1}, x) + \partial_{12}h(x, x_1) \frac{ds(x)}{dx} + \partial_{11}h(x, x_1) = 0,$$

o que nos dá

$$\partial_{12}h(x_{-1}, x) \frac{ds^{-1}(x)}{dx} + \partial_{12}h(x, x_1) \frac{ds(x)}{dx} = -(\partial_{22}h(x_{-1}, x) + \partial_{11}h(x, x_1)).$$

Agora note que:

- para todo ponto  $(x, x')$ ,  $\partial_{12}h(x, x') > 0$

- as funções  $s(x)$  e  $s^{-1}(x)$  são ambas estritamente crescentes (levantamentos de homeomorfismos do círculo) e Lipschitz, assim como elas são deriváveis q.t.p., segue que  $\frac{ds^{-1}(x)}{dx} > 0$  e  $\frac{ds(x)}{dx} > 0$  para quase todo  $x$

Assim,  $\partial_{22}h(x_{-1}, x) + \partial_{11}h(x, x_1) < 0$  para quase todo  $x$ . O que implica que  $\partial_{22}h(x_{-1}, x) + \partial_{11}h(x, x_1) \leq 0$  para todo  $x$ . Mas isto contradiz a hipótese do lema e portanto termina a prova.  $\square$

Vamos agora voltar ao problema do bilhar. Para que possamos aplicar o lema 22, codificaremos o problema da seguinte maneira, já explicada anteriormente: Seja  $\Gamma(t)$  uma parametrização do bordo por comprimento de arco. Dado um ponto  $W$  pertencente a curva  $\Gamma$  e um ângulo de saída  $\phi$ , seja  $W_1$  o ponto de encontro da semi-reta que tem  $W$  como origem e forma ângulo  $\phi$  com a tangente a  $\Gamma$  em  $W$ . Seja  $\phi_1$  o ângulo que a semi-reta anterior forma com a tangente a  $\Gamma$  em  $W_1$ . Então a nossa aplicação do bilhar fica:

$$(t, u) \rightarrow f(t, u) = (s, w), \quad (3.1)$$

onde  $\Gamma(t) = W$ ,  $u = \cos(\phi)$ ,  $\Gamma(s) = W_1$  e  $w = \cos(\phi_1)$ . Já foi mostrado que  $f$  tem a seguinte função geratriz:

$$h(t, s) = \sqrt{(\Gamma(s) - \Gamma(t)) \cdot (\Gamma(s) - \Gamma(t))}$$

Isto decorre de

$$\partial_1 h(t, s) = \frac{-\Gamma'(t) \cdot (\Gamma(s) - \Gamma(t))}{\sqrt{(\Gamma(s) - \Gamma(t)) \cdot (\Gamma(s) - \Gamma(t))}} = -\cos(\phi) = -u \quad (3.2)$$

$$\partial_2 h(t, s) = \frac{\Gamma'(s) \cdot (\Gamma(s) - \Gamma(t))}{\sqrt{(\Gamma(s) - \Gamma(t)) \cdot (\Gamma(s) - \Gamma(t))}} = \cos(\phi_1) = w \quad (3.3)$$

Assim, como  $\Gamma$  é suposta  $C^2$ ,  $f$  é  $C^1$ . Também já vimos que  $\partial_{12}h > 0$ .

Portanto, para concluir a prova, vamos mostrar que se  $\Gamma(s)$  é um ponto de curvatura zero, então podemos aplicar o lema 22. Como estamos tomando uma parametrização da curva por comprimento de

arco, a hipótese de  $\Gamma(s)$  ser um ponto de curvatura zero se reduz a  $\Gamma''(s) = (0, 0)$ . Portanto, derivando as expressões (3.2) e (3.3), obtemos:

$$\partial_{11}h(s, w) = \frac{\|\Gamma'(s)\|^2 \|\Gamma(s) - \Gamma(w)\|^2 - \|\Gamma'(s) \cdot (\Gamma(s) - \Gamma(w))\|^2}{\|\Gamma(s) - \Gamma(w)\|^3}$$

$$\partial_{22}h(t, s) = \frac{\|\Gamma'(s)\|^2 \|\Gamma(s) - \Gamma(t)\|^2 - \|\Gamma'(s) \cdot (\Gamma(s) - \Gamma(t))\|^2}{\|\Gamma(s) - \Gamma(t)\|^3}$$

Como as expressões acima são iguais, nos concentraremos em apenas uma delas. Pela desigualdade de Cauchy-Schwarz,  $\partial_{22}h(t, s) \geq 0$ . A única maneira de  $\partial_{22}h(t, s) = 0$  é se  $\Gamma'(s)$  e  $\Gamma(s) - \Gamma(t)$  forem vetores paralelos ( $\phi_1 = 0$  ou  $\pi$ ). Mas como estamos considerando um bilhar convexo, isso só ocorre se o ângulo  $\phi$  entre  $\Gamma'(t)$  e  $\Gamma(s) - \Gamma(t)$  for 0 ou  $\pi$ . Assim, o ponto de saída  $(t, \cos(\phi)) \in S^1 \times \{-1 \text{ ou } 1\}$ . Como estamos buscando curvas invariantes em  $S^1 \times (-1, 1)$ , pois os bordos do anel já são invariantes, este caso não interessa.

Caso a aplicação do bilhar  $f$  (3.1) possua uma curva rotacional invariante em  $S^1 \times (-1, 1)$ , que pelo teorema da curva invariante de Birkhoff coincide com o gráfico de uma função Lipschitz  $\gamma : S^1 \rightarrow (-1, 1)$ , temos como antes, que  $f(s, \gamma(s)) = (g(s), \gamma(g(s)))$ , onde  $g(s)$  é um homeomorfismo do círculo que preserva orientação. Como o gráfico de  $\gamma$  está contido em  $S^1 \times (-1, 1)$ , finalmente obtemos que

$$\partial_{22}h(g^{-1}(s_0), s_0) + \partial_{11}h(s_0, g(s_0)) > 0,$$

onde  $s_0$  é o parâmetro onde a curvatura vale 0. Assim o lema 22 se aplica e a demonstração está completa.

### 3.6 Existência de Curvas Invariantes

Apresentamos, nesta seção, uma versão do Teorema da Curva Invariante de Moser também conhecido como o Teorema do Twist [59], que segue de um resultado sobre perturbações de aplicações integráveis. A versão aqui exposta vale para o caso analítico e está baseada na referência [49] de Levi e Moser.

O objetivo é provar a existência de curvas rotacionais invariantes numa vizinhança de um ponto fixo elíptico genérico, isto é, que possui algum invariante de Birkhoff não nulo.

**Teorema 15.** *Suponha que  $f : U \subset (\mathbb{R}^2, 0) \rightarrow (\mathbb{R}^2, 0)$ ,  $U$  aberto de  $\mathbb{R}^2$  seja uma aplicação analítica real que preserva área com ponto fixo elíptico na origem. Se a Forma Normal de Birkhoff de  $f$ , numa vizinhança de 0, se escreve em coordenadas complexas  $(z, \bar{z})$ ,  $\varphi(z) = e^{iP(|z|^2)}.z + O(|z|^q)$  onde  $P$  é um polinômio real de grau  $\leq \frac{q}{2} - 1$  e não constante, então 0 é um ponto fixo estável.*

*Mais precisamente, para todo  $\epsilon > 0$  suficientemente pequeno existe uma curva (analítica)  $\gamma$ , contida num anel em torno de 0, homotopicamente não trivial e invariante por  $f$ ,  $f(\gamma) = \gamma$ .*

*Demonstração.* Este teorema é consequência de um resultado de existência de curvas invariantes para pequenas perturbações de aplicações do tipo twist suficientemente próximas de aplicações integráveis. A formulação do problema dessa maneira é conseguida após uma mudança de coordenadas numa vizinhança da origem.

Escrevendo  $P(x) = \alpha + \beta x^m + P_1(x)$  com grau  $P_1 > m$ , onde,  $\beta$  é o primeiro invariante de Birkhoff não nulo, obtemos

$\varphi(z) = e^{i(\alpha + \beta|z|^{2m})}.z + \varphi_1(z, \bar{z})$  onde  $\varphi_1(z, \bar{z}) \in O(|z|^{2m+2})$ , isto é  $|\varphi_1(z, \bar{z})| \leq |z|^{-(2m+2)}$  é limitado.

Observe que  $\varphi^{-1}(w) = e^{i(-\alpha + \beta|w|^{2m})}w + O(|w|^{2m+2})$  de modo que trocando  $\varphi$  por  $\varphi^{-1}$ , se necessário, podemos supor  $\beta > 0$ .

Provaremos a existência de curva invariante contida no anel

$$\Gamma_\epsilon =: \epsilon^2(1 - \epsilon^\nu) < |z|^2 < \epsilon^2(1 + \epsilon^\nu)$$

para  $\epsilon > 0$  suficientemente pequeno e  $\nu = \frac{1}{3}$  (por exemplo).

Escreva as "coordenadas polares" com mudança de escala no anel  $\Gamma_\epsilon$ ,  $z = \epsilon\sqrt{1 + \epsilon^\nu y} e^{2\pi i x}$  de modo que  $|z|^2 = \epsilon^2(1 + \epsilon^\nu y)$ , assim o anel  $\Gamma_\epsilon$  transforma-se no anel  $|y| < 1$ .

É fácil verificar, usando a equação de conjugação, que a aplicação  $(x, y) \mapsto \epsilon\sqrt{1 + \epsilon^\nu y}(\cos(2\pi x), \sin(2\pi x))$  tem determinante jacobiano constante, portanto a aplicação  $\tilde{\varphi}$  igual a  $\varphi$  restrita ao anel  $\Gamma_\epsilon$

escrita nas coordenadas  $(x, y)$  é exata, isto é, existe uma função  $S$  periódica na variável  $x$ , tal que  $\tilde{\varphi}^*(ydx) - ydx = dS(x, y)$ .

**Afirmção:** Se  $\gamma = \frac{\beta m}{2\pi} \epsilon^{2+\nu} > 0$ ,  $a(\gamma) = \frac{1}{2\pi}(\alpha + \epsilon^{2m})$  então podemos escrever:

$$\tilde{\varphi}(x_1, y_1) = (x_1 + a(\gamma) + \gamma y_1 + f(x_1, y_1, \gamma), y_1 + g(x_1, y_1, \gamma))$$

com  $f$  e  $g$  analíticas em  $|y_1| < 1$  e  $|Im(x_1)| < r$ ,  $f, g \in O(\epsilon^{2(m+\nu)})$ , onde  $\nu = \frac{1}{3}$ ,  $r$  é uma constante positiva e  $|Im(x_1)|$  significa o módulo do argumento do complexificado da coordenada  $x_1$ .

Isto é, a restrição de  $\varphi$  ao anel  $\Gamma_\epsilon$  é uma aplicação do tipo twist.

$$\tilde{\varphi}(x_1, y_1) = (x_1 + a + \gamma y_1 + 0(\epsilon^{2(m+\nu)}), y_1 + 0(\epsilon^{2(m+\nu)})).$$

*Demonstração.* Escreva  $\tilde{\varphi}(x_1, y_1) = (x_2, y_2)$  de modo que

$$\epsilon \sqrt{1 + \epsilon^\nu y_2} e^{2\pi i x_2} = e^{i(\alpha + \beta |z|^{2m})} z + O(|z|^{2m+2})$$

tomando o quadrado do módulo desta expressão obtemos:

$$\epsilon^2(1 + \epsilon^\nu y_2) = \epsilon^2(1 + \epsilon^\nu y_1) + 0(|z|^{2m+3}) \text{ o que implica}$$

$$y_2 = y_1 + 0(\epsilon^{2m+1-\nu}).$$

Por outro lado,

$$2\pi x_2 = \alpha + 2\pi x_1 + \beta \epsilon^{2m} (1 + \epsilon^\nu y_1)^m + 0(|z|^{2m+2})$$

$$\text{ou } x_2 = x_1 + \frac{1}{2\pi}(\alpha + \beta \epsilon^{2m}) + \frac{\beta}{2\pi} m \epsilon^{2m+\nu} y_1 + O(\epsilon^{2m+2\nu}) +$$

$$+ O(\epsilon^{2m+1-\nu}). \text{ Como } \nu = \frac{1}{3} \text{ temos } 1 - \nu = 2\nu \text{ isto é}$$

$$x_2 = x_1 + \frac{1}{2\pi}(\alpha + \beta \epsilon^{2m}) + \frac{\beta}{2\pi} m \epsilon^{2m+\nu} y_1 + 0(\epsilon^{2m+1-\nu}).$$

Fazendo  $a = \frac{1}{2\pi}(\alpha + \beta \epsilon^{2m})$  e  $\gamma = \frac{\beta m}{2\pi} \epsilon^{2m+\nu} > 0$ , obtemos a expressão de  $\tilde{\varphi}(x_1, y_1)$ .  $\square$

Portanto o Teorema 15 da Curva Invariante de Moser para o caso analítico segue do seguinte:

**Teorema 16.** *Seja  $F(x, y) = (x + a(\gamma) + \gamma y + f(x, y, \gamma), y + g(x, y, \gamma))$  onde  $f, g$  são funções analíticas em  $|y| < 1, |Im(x)| \leq r$ , tal que  $F$  preserva área e  $0 < \gamma < 1$ .*

*Existe uma constante  $\delta > 0$ , independente de  $\gamma$ , tal que se  $|f|_{r,1} + |g|_{r,1} < \gamma\delta$  então  $F$  possui uma curva invariante analítica  $x = u(\theta), y = v(\theta)$  com  $|v| < 1$  e  $u'(\theta) > 0, \hat{u}(\theta) = u(\theta) - \theta$  e  $v(\theta)$  periódicas de período 1.*

*Aqui  $|f|_{r,s} = \sup |f|$  para  $|y| < s, |Im(x_1)| < r$ . (Veja o 2º passo na demonstração abaixo).*

A prova deste teorema será apresentada no final desta seção. Iremos mostrar a existência de curvas invariantes dentro do contexto Lagrangiano, tal como apresentado no artigo de Levi e Moser [49].

Apesar das hipóteses aqui utilizadas serem bem mais fortes do que as usuais da Teoria KAM, vários aspectos interessantes desta teoria podem ser ressaltados.

O primeiro deles, é o uso de um "método variacional" para achar a curva invariante, isto é, usando a função geratriz da aplicação do tipo twist, o problema de achar uma curva invariante traduz-se em encontrar uma solução para uma equação de diferenças (equação homológica).

O segundo aspecto é a modificação do Método de Newton para encontrar aproximações para a solução. Isto significa que em cada etapa do processo iterativo ao invés de resolver a equação linearizada, como no método de Newton usual, resolve-se a mesma a menos de um erro quadrático.

Finalmente, mostra-se como associar a um problema de existência de curva invariante numa vizinhança de um ponto fixo elíptico um problema perturbativo.

Estes três aspectos por si só justificam, se não por razões estéticas, a escolha da exposição que fizemos.

Para outras abordagens e para uma ampla discussão sobre a teoria KAM referimos ao artigo [25] de R. de La LLave.

**Passos preparatórios para a prova do teorema 16:** Após reformular a questão da existência de curvas invariantes em termos da função geratriz, a idéia é usar um método de aproximações sucessivas para resolver uma equação de diferenças não linear. Para

isso, usa-se um método de Newton modificado em que ao invés de resolver a equação linearizada usual, esta é resolvida a menos de um erro quadrático. Além do controle deste erro em cada etapa do processo, há necessidade de acompanhar uma diminuição do domínio de analiticidade da solução em relação aos dados iniciais. Este é o conteúdo do Lema de Aproximação (Lemma 23). Finalmente prova-se o resultado principal, Teorema 17, que prova a existência de solução única, desde que iniciemos com uma "quase solução".

**1º Passo:** Reformulação do problema da existência da curva invariante usando a função geratriz  $h$  (princípio variacional).

Seja  $h(x_1, x_2)$  uma função geratriz da aplicação, isto é,  $h$  é uma função analítica que satisfaz  $h(x_1+1, x_2+1) = h(x_1, x_2)$ ,  $\partial_{12}h(x_1, x_2) < 0$  e  $f(x_1, y_1) = (x_2, y_2)$  se somente se  $\partial_1h(x_1, x_2) = -y_1$  e  $\partial_2h(x_1, x_2) = y_2$ .

Suponha que  $\theta \mapsto (u(\theta), v(\theta))$  seja uma curva  $\delta$ , de classe  $C^1$ , tal que  $u$  é crescente e  $u(\theta + 1) = u(\theta) + 1$ .

Podemos usar a função geratriz  $h$  para descrever a condição de invariância de  $\delta$  de modo que  $f \mid \delta$  é conjugada a uma rotação de ângulo  $\omega$ , isto é

$$f(u(\theta), v(\theta)) = (u(\theta + \omega), v(\theta + \omega)).$$

Em termos da função  $h$  isto significa:

$$\begin{cases} v(\theta) = -\partial_1h(u(\theta), u(\theta + \omega)) \\ v(\theta + \omega) = \partial_2h(u(\theta), u(\theta + \omega)) \end{cases}$$

ou

$$\begin{cases} v(\theta) = -\partial_1h(u(\theta), u(\theta + \omega)) \\ v(\theta) = \partial_2h(u(\theta - \omega), u(\theta)) \end{cases}$$

Portanto, estamos procurando uma função  $\theta \mapsto u(\theta)$  tal que

$$\partial_1h(u(\theta), u(\theta + \omega)) + \partial_2h(u(\theta - \omega), u(\theta)) = 0$$

Se usarmos a seguinte notação  $u^+(\theta) = u(\theta + \omega)$ ,  $u^-(\theta) = u(\theta - \omega)$  podemos escrever a equação acima como  $E(u(\theta)) = 0$ , onde

$$E(u(\theta)) = \partial_1 h(u(\theta), u^+(\theta)) + \partial_2 h(u^-(\theta), u(\theta)).$$

**Observação:** a função  $u'(\theta)E(u(\theta))$  tem média zero, isto é,

$$\int_0^1 u'(\theta)E(u(\theta))d\theta = 0.$$

*Demonstração.* Se denotarmos  $\nabla G(\theta) = G(\theta + \omega) - G(\theta)$  e  $\nabla^* G(\theta) = G(\theta) - G(\theta - \omega)$  então

$$\begin{aligned} u'(\theta)E(u(\theta)) &= u'(\theta)\partial_1 h(u(\theta), u^+(\theta)) + u'(\theta)\partial_2 h(u^-(\theta), u(\theta)) \\ &= \frac{d}{d\theta}h(u(\theta), u^+(\theta)) - \nabla[u'(\theta)\partial_2 h(u^-(\theta), u(\theta))]. \end{aligned}$$

Portanto

$$\begin{aligned} &\int_0^1 u'(\theta)E(u(\theta))d\theta = \\ &\int_0^1 \frac{d}{d\theta}h(u(\theta), u^+(\theta))d\theta - \int_0^1 \nabla[u'(\theta)\partial_2 h(u^-(\theta), u(\theta))]d\theta = \\ &h(u(1), u^+(1)) - h(u(0), u^+(0)) - \int_0^1 \nabla[u'(\theta)\partial_2 h(u^-(\theta), u(\theta))]d\theta. \end{aligned}$$

Mas  $h(u(1), u^+(1)) = h(u(0) + 1, u^+(0) + 1) = h(u(0), u^+(0))$ , ou seja, o primeiro termo é nulo. Quanto ao segundo,

$$\int_0^1 u'(\theta + \omega)\partial_2 h(u(\theta), u(\theta + \omega))d\theta - \int_0^1 u'(\theta)\partial_2 h(u(\theta - \omega), u(\theta))d\theta$$

fazemos a mudança  $\theta + \omega = \varphi$  na primeira integral para obter:

$$\int_{\omega}^{\omega+1} u'(\varphi)\partial_2 h(u(\varphi - \omega), u(\varphi))d\varphi - \int_0^1 u'(\theta)\partial_2 h(u(\theta - \omega), u(\theta))d\theta$$

Como

$$\int_0^{\omega} u'(\theta)\partial_2 h(u^-(\theta), u(\theta))d\theta = \int_1^{1+\omega} u'(\theta)\partial_2 h(u^-(\theta), u(\theta))d\theta$$



Então reescrevemos

$$\int_{\omega}^{\omega+1} u'(\theta) \partial_2 h(u^-(\theta), u(\theta)) d\theta + \int_0^{\omega} u'(\theta) \partial_2 h(u^-(\theta), u(\theta)) d\theta - \int_0^1 u'(\theta) \partial_2 h(u^-(\theta), u(\theta)) d\theta - \int_1^{1+\omega} u'(\theta) \partial_2 h(u^-(\theta), u(\theta)) d\theta = 0$$

□

**2º Passo:** Definição de um espaço de funções onde procurar a solução de  $E(u) = 0$ .

Se  $g : \mathbb{R} \rightarrow \mathbb{R}$  é uma função periódica, de período 1 e analítica real então usando a expansão de Taylor, podemos estendê-la a uma vizinhança do eixo real em  $\mathcal{C}$ . Vale também a seguinte estimativa: se  $\sup |g^{(n)}(\theta)| = M_n$  então,  $\exists R > 0$  tal que  $M_n \leq n! R^n$

Além disso,  $g = \sum g_k e^{2\pi i k \theta}$  onde  $g_k = \int_0^1 g(\theta) e^{2\pi i k \theta} d\theta$  e existem constantes  $K > 0$  e  $a > 0$  tais que

$$|g_k| \leq K e^{-a|k|}$$

*Demonstração.* Se  $g(z)$  é analítica em  $|Im(z)| < \lambda$  então mudando o caminho de integração  $\int_0^1 g(\theta) e^{2\pi k i \theta} d\theta$  em  $\mathcal{C}$  para a curva

$$\psi = \{t\delta i\}_{0 \leq t \leq 1} \circ \{t + \delta i\}_{0 \leq t \leq 1} \circ \{1 + (1-t)\delta i\}_{0 \leq t \leq 1}$$

com  $|\delta| < \lambda$  de modo que

$$g_k = \int_0^1 g(t + \delta i) e^{2\pi k \delta} e^{-2\pi i k t} dt.$$

Logo

$$|g_k| = \int_0^1 |g(t + \delta i)| e^{2\pi k \delta} dt.$$

Seja  $0 < \sigma < \lambda$ . Se  $k > 0$  definimos  $\delta = \sigma - \lambda$  e se  $k < 0$  fazamos  $\delta = \lambda - \sigma$ , de modo que  $e^{2\pi k \delta} = e^{-2\pi |k|(\lambda - \sigma)}$  e  $|g_k| \leq |g|_{\lambda} e^{-2\pi |k|(\lambda - \sigma)}$ , onde  $|g|_{\lambda} = \sup |g(z)|$  em  $|Im(z)| < \lambda$ .

Fazendo  $\sigma \rightarrow 0$  obtemos o decaimento exponencial do coeficiente de Fourier de uma função analítica:

$|g_k| \leq |g|_\lambda e^{-2\pi|k|\lambda}$ . onde  $|g|_\lambda = \sup |g(z)|$  em  $|Im(z)| < \lambda$ .  $\square$

Observe que a função  $g(z) = \sum g_k e^{2k\pi iz}$  é analítica em um domínio  $|Im(z)| < \frac{\lambda}{2}$  pois,

$$|g_k e^{2\pi ikz}| = |g_k| e^{-2\pi ky} \leq |g|_\lambda e^{-2\pi(|k|\lambda + ky)}$$

Mas

$$|k|\lambda + ky > |k|\lambda - |k|\frac{\lambda}{2} = |k|\frac{\lambda}{2} \text{ para } |y| < \frac{\lambda}{2}.$$

De modo que  $|g_k e^{2\pi ikz}| \leq |g|_\lambda e^{-\pi|k|\lambda}$  e portanto a série acima converge uniformemente.

Obtemos então, que se  $W_r$  é o espaço das funções analíticas periódicas, de período 1, em  $|Imz| < r$  então para todo  $g \in W_r$  vale:  $|g_k| \leq |g|_r e^{-2\pi|k|r}$ .

O próximo resultado, Teorema 17, diz que se iniciamos com uma função  $u_0(\theta)$  tal que  $E(u_0)$  é suficientemente pequeno, então existe uma solução (única com média nula) para o problema  $E(u) = 0$  descrito no 1º Passo, para  $u$  próxima de  $u_0$ .

Suponhamos que a função geratriz  $h(x_1, x_2)$  seja analítica e que se estenda a  $h(z_1, z_2)$  em um domínio  $D \subset \mathcal{C}^2$ .

Fixemos um número  $R > 0$  tal que  $\emptyset \neq D_R \subset D$  está definido por

$$D_R = \{(z_1, z_2) / \text{dist}((z_1, z_2), D^c) > R\},$$

isto é,  $D_R$  é o maior subconjunto de  $D$  cuja  $R$ -vizinhança está contida em  $D$ .

**Teorema 17.** *Suponha que  $\min_D |\partial_{12}h| > N_1$  e que exista  $M > 0$  tal que  $|h|_{\mathcal{C}^3} < M$  em  $D$  (todas as derivadas até ordem três limitadas por  $M$ ).*

*Suponha que exista uma função  $u_0(\theta)$  tal que:  $u_0(\theta) - \theta \in W_r$  para algum  $0 < r < 1$  e que a curva  $(u_0(z), u_0^+(z))$  esteja inteiramente contida em  $D_R$  para  $|Im(z)| < r$ .*

*Além disso, suponha que exista  $N_o > 0$  tal que  $|u'_0|_r < N_o$ ,  $|(u'_0)^{-1}| < N_o$ .*

*Se  $\alpha$  é um número que satisfaz à Condição Diofantina (CD): existem  $K > 0, \sigma > 0$  tais que*

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{K}{q^{2+\sigma}} \quad \forall p, q > 0 \quad \text{inteiros.}$$

Então existe  $\delta = \delta(r, R, M, N_o, K, \sigma, N_1) > 0$  tal que se  $|E(u_0)|_r < \delta$ , existe uma única função  $u$  próxima de  $u_0$ , tal que  $E(u) = 0$  com  $u(\theta) - \theta \in W_{r/2}$  e  $\int_0^1 (u(\theta) - \theta) d\theta = 0$ .

**Exemplo:** Se  $(h(x_1, x_2) = \frac{1}{2}(x_2 - x_1)^2 + V(x_1))$  então  $E(u) = 2u(\theta) - u^+(\theta) - u^-(\theta) + V'(u)$  e a condição  $|E(u_0)|_r < \delta$  depende fortemente da norma  $|V'(x_1)|$ .

No caso da aplicação standard,  $V(x_1) = \frac{k}{4\pi^2} \cos(2\pi x_1)$ , assim começando com função  $u(\theta) = \theta$ , basta tomar  $k$  suficientemente pequeno.

Os próximos passos e lemas são dedicados à prova do Teorema 17.

**3º Passo:** Modificação do Método de Newton

Escrevendo  $E(u + v) = E(u) + E'(u).v + Q(v)$  com  $|Q(v)| \leq c|v|^2$  para  $|v|$  pequeno, o método de Newton consiste em resolver sucessivamente a equação linear tipo  $E(u) + E'(u).v = 0$ , para obter uma sequência convergente à solução de  $E(u) = 0$ .

Ao invés disso, Levi e Moser propõem uma outra equação linear (Equação Homológica) que se escreve na forma:

$$u'E(u) + u'E'(u).v - vE'(u).u' = 0$$

Lembrando que  $u^+(\theta) = u(\theta + \alpha)$  e  $u^-(\theta) = u(\theta - \alpha)$ , temos

$$\begin{aligned} & E'(u).v = \\ & = \partial_{11}h(u, u^+)v + \partial_{12}h(u, u^+)v^+ + \partial_{21}h(u^-, u)v^- + \partial_{22}h(u^-, u)v \end{aligned}$$

e

$$\begin{aligned} & u'.E'(u)v - v.E'(u)u' = \\ & u'\partial_{11}h(u, u^+).v + u'\partial_{12}h(u, u^+)v^+ + u'\partial_{21}h(u^-, u)v^- \\ & + u'\partial_{22}h(u^-, u)v - v\partial_{11}h(u, u^+)u' - v\partial_{12}h(u, u^+)(u')^+ \\ & - v\partial_{21}h(u^-, u)(u')^- - v\partial_{22}h(u^-, u)(u')^- = \\ & \partial_{12}h(u, u^+)(u'v^+ - v(u')^+) + \partial_{21}h(u^-, u)(u'v^- - v(u')^-) \end{aligned}$$

Fazendo  $v = u'w$ , de modo que  $v^+ = (u')^+w^+$  etc..., tem-se

$$\begin{aligned} & u'E'(u)v - vE'(u)u' = \\ & \partial_{12}h(u, u^+)(u'(u')^+w^+ - u'w(u')^+) + \partial_{21}h(u^-, u)(u'(u')^-w^- - u'w(u')^-) \\ & = \partial_{12}h(u, u^+)u'(u')^+.(w^+ - w) - \partial_{12}h(u^-, u)u'(u')^-(w - w^-) \end{aligned}$$

Lembrando que

$$\nabla G(\theta) = G(\theta + \alpha) - G(\theta) = (G^+ - G)(\theta)$$

$$\nabla^* G(\theta) = G(\theta) - G(\theta - \alpha) = (G - G^-)(\theta)$$

Escrevemos:

$$\partial_{12}h(u, u^+)u'(u')^+\nabla w - \partial_{12}h(u^-, u)u'(u')^-\nabla^* w$$

Mas

$$\begin{aligned} & \nabla^*(\partial_{12}h(u, u^+)u'(u')^+) = \\ & \partial_{12}h(u, u^+)u'(u')^+ - \partial_{12}h(u^-, u)u'(u')^- \\ & \text{e} \\ & \nabla^*(\partial_{12}h(u, u^+)u'(u')^+\nabla w) = \\ & \partial_{12}h(u, u^+)u'(u')^+\nabla w - \partial_{12}h(u^-, u)u'(u')^-\nabla^* w \end{aligned}$$

Isto significa que podemos reescrever a equação homológica na seguinte forma:

$$\nabla^*(\partial_{12}h(u, u^+)u'(u')^+\nabla w) = -u'E(u).$$

Esta equação pode ser então resolvida em duas etapas:

$$\begin{cases} \nabla^*(\psi) = -u'E(u) \\ \partial_{12}h(u, u^+)u'(u')^+\nabla w = \psi + \eta, \eta \text{ constante} \end{cases}$$

a ser determinada.

Como  $\partial_{12}h(u, u^+)u'(u')^+ < 0$ , a segunda equação se escreve na forma  $\nabla w = g$ .

O próximo lema afirma a existência da solução para o sistema acima, mas com uma perda no domínio da analiticidade.

**Lema 23. (de Aproximação)**

Suponha que  $\alpha$  satisfaça à condição diofantina (CD)

$$|\alpha q - p| \geq \frac{K}{q^{\sigma+1}}$$

e  $g \in W_r$  tenha média zero. Então, para todo  $0 < r' < r$ , a equação  $\nabla\varphi = g$  tem solução única, com  $\varphi \in W_{r'}$ , e  $\int_0^1 \varphi d\theta = 0$  (média zero) tal que existe uma constante  $C(K, \sigma)$

$$|\varphi|_{r'} \leq C(K, \sigma) \frac{|g|_r}{(r - r')^\tau}, \quad \tau = 2 + \sigma$$

*Demonstração.* Escreva as séries de Fourier  $\varphi(\theta) = \sum \varphi_n e^{2n\pi i\theta}$  e  $g(\theta) = \sum g_n e^{2n\pi i\theta}$

$$\nabla\varphi(\theta) = \varphi(\theta + \alpha) - \varphi(\theta) = \sum \varphi_n e^{2n\pi i\theta} (e^{2n\pi i\alpha} - 1).$$

Para que  $\nabla\varphi(\theta) = g(\theta)$  devemos ter  $\varphi_n = \frac{g_n}{e^{2n\pi i\alpha} - 1}$  e  $\varphi_0 = 0$ .

Esta última condição é necessária para que  $\varphi$  tenha média zero e é possível porque este coeficiente fica indeterminado na equação.

A condição Diofantina (CD) nos dá a existência de uma constante  $K$  tal que escolhendo  $m$  com  $|\alpha n - m| < \frac{1}{2}$ , temos

$$|e^{2n\pi i\alpha} - 1| = 2 |\operatorname{sen}(n\pi\alpha)| = 2 \operatorname{sen}(\pi |\alpha n - m|) \geq 4 |\alpha n - m| > \frac{4K}{|n|^{\sigma+1}}$$

Além disso,  $|g_n| \leq |g|_r e^{-2\pi|n|r}$ , pois  $g \in W_r$ .

Logo,  $|\varphi_n| \leq \frac{|g|_r e^{-2\pi|n|r} \cdot |n|^{1+\sigma}}{4K}$

Se  $0 < \frac{r}{2} < s < r$ ,

$$|\varphi_n| \leq |g|_r (4K)^{-1} \cdot e^{-2\pi s|n|} e^{-2\pi(r-s)|n|} \cdot |n|^{1+\sigma}$$

usando que  $x e^{-x} \leq e^{-1}$  e fazendo  $x = \frac{2\pi(r-s)|n|}{1+\sigma}$  obtemos:

$$\frac{2\pi(r-s)|n|}{1+\sigma} \cdot e^{\frac{-2\pi(r-s)|n|}{1+\sigma}} < e^{-1}$$

ou  $\left(\frac{2\pi(r-s)}{1+\sigma}\right)^{1+\sigma} |n|^{1+\sigma} e^{-2\pi(r-s)|n|} < e^{-(1+\sigma)}$

isto é,

$$|\varphi_n| \leq |g|_r (4K)^{-1} e^{-2\pi s|n|} \cdot \frac{(1+\sigma)^{1+\sigma}}{(2\pi)^{1+\sigma}} \frac{1}{(r-s)^{1+\sigma}} e^{1+\sigma}$$

ou seja:  $|\varphi_n| \leq |g|_r C(K, \sigma) \frac{e^{-2\pi|n|s}}{(r-s)^{1+\sigma}}$

Portanto, se  $r' > 0$  é tal que  $s = \frac{r+r'}{2}$  obtemos

$$|\varphi|_{r'} \leq \Sigma |\varphi_n| e^{2|n|\pi r'} \leq \Sigma \frac{|g|_r C(K, \sigma) e^{2\pi|n|(r'-s)}}{(r-s)^{1+\sigma}}$$

$$\text{ou } |\varphi|_{r'} \leq \frac{2|g|_r C(K, \sigma)}{(r-s)^{1+\sigma}} \cdot \Sigma e^{2\pi|n|(r'-s)}.$$

$$|\varphi|_{r'} \leq \frac{2|g|_r C(K, \sigma)}{(r-s)^{1+\sigma}} \cdot \frac{1}{1 - e^{2\pi(r'-s)}}.$$

Usando que  $\frac{1}{1-e^{-2\pi x}} < \frac{1}{x}$  para  $0 < x < \frac{1}{2}$  temos

$$|\varphi|_{r'} \leq \frac{2C(K, \sigma)}{(r-s)^{1+\sigma}} \cdot \frac{|g|_r}{(s-r')} =$$

$$\frac{2C(K, \sigma) |g|_r}{(s-r')^{2+\sigma}} = \frac{2^{(3+\sigma)} C(K, \sigma) |g|_r}{(r-r')^{2+\sigma}}$$

Concluimos assim a demonstração do Lema de Aproximação 23.  $\square$

Analogamente procedemos para a equação  $\nabla^* \psi = g$  onde  $g$  tem média zero e  $\nabla^* \psi(\theta) = \psi(\theta) - \psi(\theta - \alpha)$ .

$$\nabla^* \psi(\theta) = \Sigma \psi_n e^{2n\pi i \theta} [1 - e^{-2\pi i n \alpha}] \text{ e obviamente}$$

$$|1 - e^{-2\pi i n \alpha}| = |e^{2n\pi i \alpha} - 1|.$$

Voltemos nossa atenção agora para o sistema

$$\begin{cases} \nabla^* \psi = g(u) = -u' E(u) \\ p^{-1} \nabla w = \psi + \eta \end{cases}$$

Onde  $p^{-1} = \partial_{12}h(u(\theta), u^+(\theta))u'(\theta)(u^+)'(\theta)$  e o parâmetro  $\eta$  é determinado pela condição  $\int_0^{2\pi} p(\psi + \eta)d\theta = 0$

$$\text{isto é } \eta = -\frac{\int p\psi d\theta}{\int p d\theta}.$$

Primeiramente resolvemos  $\nabla^*\psi = -u'E(u)$  com  $u$  satisfazendo as hipóteses do Teorema 16:

$$u(\theta) - \theta \in W_r, | (u')^{-1} |_r < N, | u' |_r < N.$$

Pelo Lema 23 de Aproximação,

$$\begin{aligned} | \psi |_{r'} &\leq \frac{C(K, \sigma) | u' |_r | E(u) |_r}{| r - r' |^\tau} \\ &\leq C(K, \sigma, N) \frac{| E(u) |_r}{| r - r' |^\tau}. \end{aligned}$$

Para obtermos uma estimativa para a solução de  $\nabla(w) = p(\psi + \eta)$ , observemos que  $\nabla^*(\psi + \eta) = \nabla^*\psi = -u'E(u)$  e

$$\begin{aligned} | \eta | &\leq \frac{\int | p\psi |}{\left| \int p \right|} \leq | \psi |_{r'} \frac{\int | p |}{\left| \int p \right|} = \\ &| \psi |_{r'} \frac{\int \frac{1}{|\partial_{12}hu'(u^+)'|}}{\left| \int \frac{1}{\partial_{12}hu'(u^+)'^\tau} \right|}. \end{aligned}$$

Usando a desigualdade de Cauchy-Schwarz obtemos

$$\frac{1}{\left| \int \frac{1}{\partial_{12}hu'(u^+)'^\tau} \right|} \leq \int \partial_{12}hu'(u^+)' \leq N^2.$$

$$\text{Portanto } | \eta | \leq | \psi |_{r'} N^2 M \cdot N^2 N_1 = | \psi |_{r'} M N_1 N^4$$

$$\text{isto é: } | \eta | \leq C(k, \sigma, N, N_1, M) \frac{| E(u) |_r}{| r - r' |^\tau}$$

Usando o Lema 23, de Aproximação, concluímos que existe uma constante  $C'_1$  tal que a solução de  $\nabla w = p(\psi + \eta)$  satisfaz

$$| w |_{\rho} \leq C'_1(K, \sigma) \frac{| p(\psi + \eta) |_{r'}}{| r' - \rho |^\tau} \leq C_1 \cdot \frac{| E(u) |_r}{| r - r' |^\tau | r' - \rho |^\tau}.$$

para todo  $0 < \rho < r' < r$  com  $C_1 = C'_1(K, N_1, \sigma, N, M)$

Fazendo  $r' = \frac{\rho+r}{2}$ , obtemos  $|w|_\rho \leq C_1 \frac{|E(u)|_r}{|r-\rho|^{2\tau}}$

Finalmente, se  $v = u'w$  então obtemos uma solução para  $u'E(u) + u'E'(u)v - vE'(u)u' = 0$  analítica, que satisfaz:

$$|v|_\rho \leq C_2(K, \sigma, N_1, N, M) \frac{|E(u)|_r}{|r-\rho|^{2\tau}}$$

e usando a estimativa de Cauchy, provamos que:

$$|v'|_\rho \leq 2C_2 \frac{|E(u)|_r}{|(r-\rho)|^{2\tau+1}}$$

Com efeito, se  $v(z) = \frac{1}{2\pi i} \int_\gamma \frac{v(w)}{w-z} dw$ , então  $v'(z) = \frac{1}{2\pi i} \int_\gamma \frac{v(w)}{(w-z)^2} dw$

Para  $z$  tal que  $|Im(z)| < \rho$ , tomemos o círculo  $\gamma =: \{|w-z| = r-\rho\}$  para obter a seguinte estimativa:

$$|v'(z)| \leq \frac{1}{2\pi} \int_\gamma \frac{|v(w)|}{|w-z|^2} |dw|$$

$$\leq |v|_{|Im(z)|+r-\rho} \frac{1}{|r-\rho|} \leq C_2 \frac{|E(u)|_r}{|r-\rho| \cdot ||Im(z)|-\rho|^{2\tau}}$$

Isso significa, que para verificar a estimativa para a derivada de  $v$  é suficiente que se tenha  $|Im(z)| + r - \rho < \rho$ .

Concluimos assim que, dado  $\rho < r$  seja  $\tilde{\rho} = \frac{\rho+r}{2}$ . Então vale:

$$|v|_\rho \leq |v|_{\tilde{\rho}} \leq C_2 \frac{|E(u)|_r}{|r-\tilde{\rho}|^{2\tau}} = 2^{2\tau} C_2 \frac{|E(u)|_r}{|r-\rho|^{2\tau+1}}$$

e

$$|v'|_\rho \leq |v'|_{\tilde{\rho}} \leq \frac{1}{|\tilde{\rho}-\rho|} \leq 2^{2\tau+1} C_2 \frac{|E(u)|_r}{|r-\rho|^{2\tau+1}}.$$

Logo, fazendo  $C = 2^{2\tau} C_2$  valem as seguintes desigualdades para a solução de  $u'E(u) + u'E'(u)v - vE'(u)u' = 0$ :

$$|v|_\rho \leq C \frac{|E(u)|_r}{|(r-\rho)|^{2\tau}}$$



$$|v'|_{\rho} \leq 2C \frac{|E(u)|_r}{|(r-\rho)^{2\tau}}.$$

4º **Passo:** Estimativa do erro quadrático

**Lema 24.** *Suponha que  $v = u'w$  com  $w \in W_{\rho}$  ( $0 < \rho < r$ ) seja uma solução da equação  $u'E(u) + u'E'(u)v - vE'(u)u' = 0$ , tal que  $|w|_{\rho} \leq C_1 \frac{|E(u)|_r}{|r-\rho|^{2\tau}}$ .*

$$\text{Então } |E(u+v)|_{\rho} \leq C_6 \frac{|E(u)|_r^2}{|r-\rho|^{4\tau}}$$

para  $C_6 = C_6(k, \sigma, N_1, N, M)$

*Demonstração.* Vimos que a equação  $u'E(u) + u'E'(u).v = vE'(u)u'$  equivale a:

$$E(u) + E'(u).v = (u')^{-1}vE'(u)u' = w \frac{d}{d\theta}[E(u)].$$

Pela estimativa de Cauchy  $|\frac{d}{d\theta}[E(u)]|_{\rho} \leq D \frac{|E(u)|_r}{|r-\rho|}$  para alguma constante  $D$ .

$$\text{Logo, } |E(u) + E'(u).v|_{\rho} \leq |w|_{\rho} D \frac{|E(u)|_r}{|r-\rho|}.$$

Por hipótese,  $|w|_{\rho} \leq \frac{C_1|E(u)|_r}{|r-\rho|^{2\tau}}$ . Portanto, existe uma constante  $C_3 > 0$  tal que

$$|E(u) + E'(u).v|_{\rho} \leq C_3 \frac{|E(u)|_r^2}{|r-\rho|^{2\tau+1}}.$$

Pela Fórmula de Taylor,  $E(u+v) = E(u) + E'(u).v + Q(u, v)$  com  $|Q|_{\rho} \leq \tilde{C} |v|_{\rho}^2$ , e além disso  $|v|_{\rho} = |u'w|_{\rho} \leq N |w|_{\rho} \leq \frac{NC_1|E(u)|_r}{|r-\rho|^{2\tau}}$ .

Donde concluímos que

$$|E(u+v)|_{\rho} \leq \frac{C_3 |E(u)|_r^2}{|r-\rho|^{2\tau+1}} + \frac{\tilde{C}N^2C_1^2 |E(u)|_r^2}{|r-\rho|^{4\tau}}.$$

Lembrando que  $\tau = 2 + \sigma$ , existe uma constante  $C_6 > 0$  tal que,  
 $|E(u+v)|_\rho \leq C_6 \frac{|E(u)|_\rho^2}{|r-\rho|^{4\tau}}$ .  $\square$

**5º Passo:** Convergência do processo iterativo para uma solução de  $E(u) = 0$ .

Se  $r > 0$  é dado como nas hipóteses do teorema 17, fixado  $r_\infty < r$ , definimos a sequência  $r_n = \frac{r_{n-1} + r_\infty}{2}$ , com  $r_0 = r$ . É claro que em cada etapa a distância entre  $r_n$  e  $r_\infty$  é dividida por 2,  $r_n - r_\infty = \frac{r_{n-1} + r_\infty}{2} - r_\infty = \frac{r_{n-1} - r_\infty}{2}$  de modo que  $r_n \rightarrow r_\infty$  e  $r_n - r_{n+1} = \frac{(r_0 - r_\infty)}{2^{n+1}}$ .

Vamos supor que resolvemos sucessivamente a equação  $u'_n E(u_n) + u'_n E'(u_n) \cdot v_n = v_n E'(u_n) u'_n$  com  $u_0$  satisfazendo as hipóteses do Teorema 17, e  $u_{n+1} = u_n + v_n$ ,  $u_n \in W_n$  e  $v_n$  satisfazendo as seguintes estimativas:

$$|v_n|_{r_{n+1}} \leq C_1 \frac{|E(u_n)|_{r_n}}{|r_n - r_{n+1}|^{2\tau}} \text{ e}$$

$$|v'_n|_{r_{n+1}} \leq \frac{2C_1 |E(u_n)|_{r_n}}{|r_n - r_{n+1}|^{2\tau+1}}.$$

De modo que pela estimativa quadrática do erro (4º Passo) temos:

$$|E(u_{n+1})|_{r_{n+1}} \leq C_6 \frac{|E(u_n)|_{r_n}^2}{|r_n - r_{n+1}|^{4\tau}} = \frac{C_6 (2^{4\tau})^{n+1} |E(u_n)|_{r_n}^2}{|r_0 - r_\infty|^{4\tau}}$$

Sejam  $\epsilon_n = |E(u_n)|_{r_n}$  e  $a = 2^{4\tau}$ , temos então uma expressão tipo  $\epsilon_{n+1} \leq C_7 a^n \epsilon_n^2$  para alguma constante  $C_7$ .

É preciso encontrar um número  $\delta > 0$  tal que iniciando com uma condição  $|E(u_0)|_{r_0} = \epsilon_0 < \delta$ , como no enunciado do Teorema 17, temos  $\epsilon_n^2 \rightarrow 0$  suficientemente rápido de maneira que o produto  $a^n \epsilon_n^2$  convirja a zero.

Escrevendo  $\eta_n = C_7 a^{n+1} \epsilon_n$  temos

$$\eta_{n+1} = C_7 a^{n+2} \epsilon_{n+1} \leq C_7^2 a^{2n+2} \epsilon_n^2 = \eta_n^2.$$

De modo que se escolhermos  $\epsilon_0$  tal que  $C_7 a \epsilon_0 < 1$ , então  $\eta_n \leq \eta_0^{2^n}$ , que converge a 0 e obviamente  $\epsilon_n \rightarrow 0$ .

Mas

$$u_n = u_0 + \sum_{j=1}^n (u_j - u_{j-1}) = u_0 + \sum_{j=0}^{n-1} v_j$$

e

$$|v_n|_{r_\infty} \leq |v_n|_{r_{n+1}} \leq C_1 \frac{\epsilon_n}{|r_n - r_{n+1}|^{2\tau}} = \frac{C_1(2^{2\tau})^{n+1} \epsilon_n}{(r_0 - r_\infty)^{2\tau}}$$

Como acabamos de ver, o lado direito é o termo geral de uma série convergente.

Logo  $\sum v_n$  converge absolutamente em  $|Im(z)| < r_\infty$ .

O mesmo pode ser afirmado sobre  $u_n$  e além disso,  $E(\lim(u_n)) = \lim E(u_n) = 0$ . Isto é  $u_n$  converge uniformemente a uma solução de  $E(u) = 0$  no domínio  $|Im(z)| < r_\infty$ .

**6º Passo:** Resta apenas provar que em cada passo do processo iterativo obtemos uma função  $u_n$ , para a qual valem as hipóteses usadas no 3º Passo e no Lema 23 (de Aproximação).

Isto é, resta provar que  $u_n \in W_{r_n}$  com

$$(u_n, u_n^+) \subset D_R \text{ para } |Im(z)| < r_n \text{ e}$$

$$|u'_n|_{r_n} < N, \quad |(u'_n)^{-1}|_{r_n} < N.$$

**Lema 25.** *Suponha que  $u_0(\theta)$  seja uma função (condição inicial) tal que  $u_0(\theta) - \theta \in W_r$  para algum  $0 < r < 1$ . Suponha que  $(u_0, u_0^+) \in D_R$  para  $|Im(z)| < r$  e que  $|u'_0|_r < N$ ,  $|(u'_0)^{-1}|_r < N$  onde  $R$  e  $N$  são dados na hipóteses do Teorema 17*

*Se  $u_n = u_{n-1} + (u'_{n-1})w_{n-1}$  com  $w_{n-1} \in W_{r_n}$  a solução da equação homológica então*

$$|u_n - u_0|_{r_n} < \frac{R}{2} \quad \text{e} \quad |u'_n - u'_0|_{r_n} < \frac{2}{N} \quad \text{para todo } n = 1, 2, \dots$$

*Demonstração.* Usando a notação da prova do Lema 23,  $a = 2^{4\tau}$ ,  $\eta_n = C_7 a^{n+1} \epsilon_n$ ,  $r_n - r_{n-1} = 2^{-(n+1)}(r_0 - r_\infty)$ , se  $\lambda \leq 4\tau$  então

$$\begin{aligned} \Sigma \frac{\epsilon_n}{(r_n - r_{n-1})^\lambda} &= \Sigma \frac{\epsilon_n 2^{\lambda(n+1)}}{(r_0 - r_\infty)^\lambda} \leq \\ &\leq (r_0 - r_\infty)^{-\lambda} C_7^{-1} \Sigma C_7 \epsilon_n a^{n+1} = \\ &(r_0 - r_\infty)^{-\lambda} C_7^{-1} \Sigma \eta_n \leq \\ &\leq \frac{2\eta_0}{C_7 (r_0 - r_\infty)^\lambda} = \\ &\frac{2a\epsilon_0}{(r_0 - r_\infty)^\lambda} \text{ para } \eta_0 < \frac{1}{2}. \end{aligned}$$

Logo

$$\begin{aligned} |u_n - u_0|_{r_n} &\leq \Sigma_{k=0}^{n-1} |v_k|_{r_k} \leq \Sigma \frac{C_1 \epsilon_k}{|r_k - r_{k+1}|^{2\tau}} \\ &\leq \frac{2aC_1\epsilon_0}{|r_0 - r_\infty|^{2\tau}} \\ \text{e } |u'_n - u'_0|_{r_n} &\leq \Sigma_{k=0}^{n-1} |v'_k|_{r_k} \leq \Sigma |v'_k|_{r_k} \\ &\leq \Sigma \frac{C_1 \epsilon_k}{|r_k - r_{k+1}|^{2\tau+1}} \leq \frac{2C_1 \epsilon_0 a}{|r_0 - r_\infty|^{2\tau+1}}. \end{aligned}$$

Basta portanto, escolher  $\epsilon_0 = |u_0|_r$  suficientemente pequeno para obtermos as desigualdades enunciadas.  $\square$

**Lema 26.** *Nas mesmas condições do Lema 25, se  $|u_0| < N$  então*

$$\begin{aligned} (u_n(z), u_n^+(z)) &\in D_{\frac{R}{2}} \text{ em } |Im(z)| < r_n \\ |u'_n|_{r_n} &< 2N, \quad |(u'_n)^{-1}| < 2N, \quad \forall n = 1, 2, \dots \end{aligned}$$

*Demonstração.* Podemos supor que  $N > 2$ . Suponha que  $(z_1, z_2)$  sejam tais que

$\text{dist}((z_1, z_2), (u_n(\theta), u_n^+(\theta))) < \frac{R}{2}$  para todo  $\theta$  tal que  $|Im(\theta)| < r_n$ . Então

$$\text{dist}((z_1, z_2), (u_0(\theta), u_0^+(\theta))) \leq$$

$$\text{dist}((z_1, z_2), (u_n(\theta), u_n^+(\theta))) + \text{dist}((u_n(\theta), u_n^+(\theta)), (u_0(\theta), u_0^+(\theta))) < \frac{R}{2} + \frac{R}{2} = R, \text{ para todo } n \text{ suficientemente grande.}$$

Como  $(u_0, u_0^+) \in D_R$ , segue que  $(z_1, z_2) \in D$ . Isto quer dizer que se  $\text{dist}((z_1, z_2), (u_n(\theta), u_n^+(\theta))) < \frac{R}{2}$ , então  $(z_1, z_2) \in D$ . Logo, por definição do subconjunto  $D_{\frac{R}{2}}$ , segue que  $(u_n, u_n^+) \in D_{\frac{R}{2}}$ .

$$\begin{aligned} \text{Além disso, pelo lema 25, } |u'_n|_{r_n} &\leq |u'_n - u'_0|_{r_n} + |u'_0|_{r_n} \leq \\ &\leq \frac{2}{N} + N < 2N. \end{aligned}$$

E ainda,  $|(u'_0)^{-1}|_{r_n} \leq |(u'_0)^{-1}|_r < N$ , o que implica

$$|u'_0|_{r_n} > \frac{1}{N}.$$

Isso, juntamente com a desigualdade  $|u'_n - u'_0|_{r_n} < \frac{1}{2N}$ , nos fornece  $\frac{1}{2N} < |u'_n|_{r_n}$  ou  $|(u'_n)^{-1}| < 2N$ .

De modo, concluímos a prova do Lema 26 e do Teorema 17.  $\square$

Passemos à prova do Teorema 16 (Teorema da Perturbação).

Este resultado não é consequência direta do que acabamos de provar mas sim da sua demonstração. A partir da construção de um princípio variacional adequado prova-se que todos os passos descritos na prova do Teorema 17 podem ser dados.

Em particular, mostramos que é possível escolher um número  $\alpha$  que satisfaz à condição diofantina (CD), de modo que a restrição da aplicação  $F$  ao círculo invariante seja conjugada (no nosso caso analiticamente conjugada) à rotação de ângulo  $\frac{\alpha}{2\pi}$ .

Iniciamos escrevendo o princípio variacional em coordenadas convenientes:  $F(x_1, y_1) = (x_1 + a(\gamma) + \gamma y_1 + f(x_1, y_1, \gamma), y_1 + g(x_1, y_1, \gamma))$

Se  $p = \gamma^{-1}(x_2 - x_1 - a(\gamma))$ , então  $p = y_1 + \gamma^{-1}f(x_1, y_1, \gamma)$ . Vamos escrever a função geratriz nas coordenadas  $(x_1, p)$ . Por hipótese  $\gamma^{-1}|f|_{r,1} < \delta < 1$ , logo a aplicação  $(x_1, y_1) \mapsto (x_1, p)$  é um difeomorfismo global de modo que  $y_1 = p + \tilde{f}(x_1, p, \gamma)$  com  $\gamma^{-1}|\tilde{f}| < \delta$  em aberto  $|Im(x_1)| < r$  e  $|p| < \delta_1$ .

Isto segue da seguinte propriedade das funções analíticas reais, cuja prova pode ser vista, no Lema A.3 da página 730 de Poschel [63]:

**Lema 27.** *Seja  $G$  uma função analítica real que se estende a uma vizinhança  $\mathcal{U}_h$  da forma  $| \operatorname{Im}(y) | < h$  no plano complexo, tal que  $| G(y) - id | \leq \delta \leq \frac{h}{4}$ . Então  $G$  possui uma inversa analítica real  $\varphi$ , definida em uma vizinhança  $\mathcal{U}_{\frac{h}{4}}$  que satisfaz*

$$| \varphi - id | < \delta, \quad \frac{h}{4} | d\varphi - Id | < \delta.$$

*Demonstração.* Seja  $\eta = \frac{h}{4}$ . Tomemos dois pontos  $u, v \in U_{2\eta}$  tais que  $G(u) = G(v)$ . Então  $u - v = u - G(u) - v + G(v)$  e  $| u - v | \leq | G(u) - u | + | G(v) - v | \leq 2\delta < 2\eta$ . Portanto o segmento  $[u, v] = (1 - s)u + sv$ ,  $s \in [0, 1]$  está contido no aberto  $U_{3\eta}$ .

Pela estimativa de Cauchy para funções analíticas reais, ao longo de  $[u, v]$  temos

$$| dG - Id |_{3\eta} \leq \frac{4 | G - id |_h}{h} \leq \frac{\delta}{\eta} < 1.$$

Ou seja  $G$  é uma perturbação da identidade e pelo Teorema do Valor Médio,  $| u - v | \leq | dG - Id | \leq \frac{\delta}{\eta} | u - v |$ . O que implica  $u = v$ , isto é,  $G$  é injetiva em  $U_{2\eta}$ . Além disso,  $G(U_{2\eta})$  contém  $U_\eta$ , tendo em vista que, dado  $z \in U_\eta$  a equação  $G(y) = z$  se escreve  $y + (G(y) - y) = z$  com  $| G - id | \leq \delta$ .

Logo, se  $z \in U_\eta$  então a aplicação  $T_z(y) = y - G(y) + z$  é uma contração em uma vizinhança compacta contida em  $U_{2\eta}$ .

Encontramos assim  $\varphi : U_\eta \rightarrow U_{2\eta}$  tal que  $\varphi = G^{-1}$ .

Como  $| G - id | \leq \delta$  em  $U_{2\eta}$ , valem as estimativas:  $| \varphi - id | \leq \delta$  em  $U_\eta$  e

$$\begin{aligned} | d\varphi - Id |_\eta &\leq | (dG)^{-1} - Id |_{2\eta} = \\ &= | (dG)^{-1} |_{2\eta} | dG - Id |_{2\eta} \leq \frac{| dG - Id |_{2\eta}}{1 - | dG - Id |_{2\eta}} \leq \frac{\delta}{\eta}. \end{aligned}$$

□

Continuando, de modo análogo, para  $y_2 = y_1 + g(x_1, y, \gamma)$  com  $\gamma^{-1} |g|_{r_1} < \delta < 1$ , obtemos  $y_2 = p + \tilde{g}(x_1, p, \gamma)$  com  $\gamma^{-1} |\tilde{g}|_{r_1, \delta_1} < \delta < 1$ .

Seja  $h(x_1, x_2)$  a função geratriz da aplicação  $F$  ou seja, tal que  $F(x_1, y_1) = (x_2, y_2)$  se somente se  $dh = y_2 dx_2 - y_1 dx_1$ . Como  $dx_2 = \gamma dp + dx_1$  temos

$$\begin{aligned} \gamma^{-1}(y_2 dx_2 - y_1 dx_1) &= \gamma^{-1}(y_2 \gamma dp + y_2 dx_1 - y_1 dx_1) \quad \text{ou} \\ \gamma^{-1}(\partial_1 h + \partial_2 h) dx_1 + \partial_2 h dp &= y_2 dp + \gamma^{-1}(y_2 - y_1) dx_1. \end{aligned}$$

Portanto, se  $l(x_1, p) = \gamma^{-1} h(x_1, x_1 + \gamma p + a(\gamma))$  concluímos que  $f(x_1, y_1) = (x_2, y_2)$  se somente se  $y_2 = \partial_2 l(x_1, p)$ ,  $y_2 - y_1 = \gamma \partial_1 l(x_1, p)$ .

Finalmente, observe que  $l(x_1, p) = \frac{p^2}{2} + l^*(x_1, p, \gamma)$  com  $l^*$  analítica satisfazendo  $|l^*|_{r_1, \delta_1} < \delta$  em  $|p| < \delta_1$ .

Vejam agora como ficam os passos para a obtenção de uma curva invariante  $x_1 = u(\theta) = \theta + \hat{u}(\theta)$  e  $y_1 = v(\theta)$  com  $\hat{u}$  e  $v$  periódicas e com número de rotação  $\beta$  a ser determinado, tal que  $x_2 = u(\theta + \beta)$ .

Substituindo em  $p = \gamma^{-1}(x_2 - x_1 - a(\gamma))$  obtemos  $p(\theta) = \gamma^{-1}(u(\theta + \beta) - u(\theta) - a(\gamma))$ . Usando a notação da prova do Teorema 17,  $(\nabla g(\theta) = g(\theta + \beta) - g(\theta) = g^+(\theta) - g(\theta)$  etc...), podemos escrever:

$$p(\theta) = \gamma^{-1} \nabla \hat{u}(\theta) + \gamma^{-1}(\beta - a(\gamma)).$$

Definindo  $\alpha = \gamma^{-1}(\beta - a(\gamma))$  obtemos  $p(\theta) = \gamma^{-1} \nabla \hat{u}(\theta) + \alpha$ .

A equação  $E(u) = \partial_1 h(u, u^+) + \partial_2 h(u^-, u) = 0$ , usando o novo princípio variacional  $l(x_1, p) = \gamma^{-1} h(x_1, x_1 + \gamma p + a(\gamma))$ , se escreve:

$$\begin{aligned} \gamma \partial_1 l(u, u^+ - u - a(\gamma)) - \partial_2 l(u, \gamma^{-1}(u^+ - u - a(\gamma))) + \\ + \partial_2 l(u^-, \gamma^{-1}(u - u^- - a(\gamma))) = 0 \end{aligned}$$

Substituindo  $\gamma^{-1}(u^+ - u - a(\gamma)) = \gamma^{-1} \nabla \hat{u}(\theta) + \alpha = p(\theta)$  e  $\gamma^{-1}(u - u^- - a(\gamma)) = \gamma^{-1} \nabla \hat{u}(\theta - \beta) + \alpha = p(\theta - \beta)$  na expressão acima tem-se que:

$$\begin{aligned} \gamma \partial_1 l(u(\theta), p(\theta)) - \partial_2 l(u(\theta), p(\theta)) + \partial_2 l(u(\theta - \beta), p(\theta - \beta)) = 0 \\ \gamma^{-1} [\partial_2 l(u(\theta), p(\theta)) - \partial_2 l(u(\theta - \beta), p(\theta - \beta))] - \partial_1 l(u(\theta), p(\theta)) = 0 \end{aligned}$$

Como mudamos as coordenadas do princípio variacional, agora denotando  $\nabla^*g(\theta) = \gamma^{-1}[g(\theta) - g(\theta - \beta)]$ , podemos escrever

$$E(u) = \nabla^* \partial_2 l(u(\theta), p(\theta)) - \partial_1 l(u(\theta), p(\theta)).$$

Portanto, se  $p(u) = \gamma^{-1}(u^+ - u - a(\gamma))$  então  $p(u^-) = \gamma^{-1}(u - u^- - a(\gamma))$  e  $p'(u).v = \gamma^{-1}(v^+ - v)$ .

Derivando e fazendo as simplificações, chegamos à seguinte expressão para o método de Newton modificado:

$$\begin{aligned} u'.E'(u)v - v.E'(u).u' &= \gamma^{-2}l_{22}(u, p(u))[u'v^+ - v(u')^+] - \\ \gamma^{-1}l_{21}(u^-, p(u^-))[u'v^- + v(u')^-] &+ \gamma^{-2}l_{22}(u^-, p(u^-))[u'v^- - v(u')^-] - \\ -\gamma^{-1}l_{21}(u, p(u))[u'v^+ - v(u')^+] &, \text{ onde } l_{ij} = \partial_{ij}l \end{aligned}$$

As substituições  $v = u'w$ ,  $v^+ = (u')^+w^+$ , etc. na equação  $u'.E'(u)v - v.E'(u).u' = -u'E(u)$  resultam na seguinte equação:

$$\begin{aligned} u'.E'(u)v - v.E'(u).u' &= \\ [\gamma^{-2}l_{22}(u, p(u))u'(u')^+ - \gamma^{-1}l_{21}(u, p(u))u'(u')^+] &(w^+ - w) + \\ [\gamma^{-2}l_{22}(u^-, p(u^-))u'(u')^- - \gamma^{-1}l_{21}(u^-, p(u^-))u'(u')^-] &(w^- - w) = -u'E(u) \end{aligned}$$

Que é equivalente à equação:

$$\nabla^*[[l_{22}(u, p(u)) - \gamma l_{12}(u, p(u))]u'(u')^+ \nabla w] = -u'E(u)$$

e ao sistema

$$\nabla^*(\psi) = -u'E(u)$$

$$[l_{22}(u, p(u)) - \gamma l_{12}(u, p(u))]u'(u')^+ \nabla w = \psi + \eta$$

Antes de prosseguir imitando a prova do Teorema 17, precisamos escolher um número  $\beta$  tal que  $\alpha = \gamma^{-1}(\beta - a(\gamma))$  esteja no domínio de  $l$ , por exemplo  $|w - a(\gamma)| < \frac{\gamma}{2}$ , de modo que  $|\alpha| < \frac{1}{2}$ . Para o Lema de Aproximação, como as equações obtidas anteriormente dependem do fator de escala  $\gamma$ , a condição diofantina também sofre uma dilatação e supomos que  $|\beta - \frac{p}{q}| \geq \frac{\gamma k}{q^{\sigma+2}}$ .

É preciso, portanto, mostrar que podemos fazer uma escolha de  $\beta$  satisfaz a estas duas condições.

A possibilidade desta escolha segue do fato de que o conjunto dos números que satisfazem à condição diofantina tem medida total em  $[0, 1]$ , ver [58].



*Demonstração.* Provemos que o complementar no intervalo  $[0, 1]$  dos números que não satisfazem à condição diofantina (CD) tem medida nula.

De fato, se  $\alpha$  não satisfaz à condição diofantina, então  $\forall \epsilon, K > 0$  a inequação  $|n\alpha - m| < Kn^{-(2+\epsilon)}$  tem sempre uma solução com  $n > 0, m$  inteiros.

Considere o conjunto  $B(K, \epsilon) = \{\alpha \in [0, 1] \mid |n\alpha - m| < Kn^{-(2+\epsilon)}\}$ . O conjunto  $B$  dos números que não satisfazem à condição diofantina se escreve  $B = \bigcap_K B(K, \epsilon)$

Para  $n, m$  fixos, a inequação acima define um intervalo  $I(K, \epsilon, n)$  em torno do número  $\frac{m}{n}$  de tamanho  $2Kn^{-(3+\epsilon)}$ . Além disso, a inequação  $|n\alpha - m| < Kn^{-(2+\epsilon)}$  implica que  $-K < m < K + n$  e este intervalo contém, no máximo,  $2K + n + 1$  números inteiros.

Portanto, denotando por  $m(A)$  a medida de Lebesgue de um subconjunto  $A$ , obtemos

$$m(B(K, \epsilon)) \leq \sum_{n=1}^{\infty} m(I(K, \epsilon, n)) \leq \sum_{n=1}^{\infty} (2K + n + 1)2Kn^{-(3+\epsilon)}$$

Mas  $2K + n + 1 < 2(K + 1)n$ , portanto

$$m(B(K, \epsilon)) < 4K(K + 1) \sum_{n=1}^{\infty} n^{-(2+\epsilon)} = 4K(K + 1)S(\epsilon)$$

onde fizemos  $S(\epsilon) = \sum_{n=1}^{\infty} n^{-(2+\epsilon)}$ .

Portanto,  $m(B) \leq 4K(K + 1)S(\epsilon)$ ,  $\forall K > 0$  ou seja,  $m(B)$  tem medida nula. Como queríamos demonstrar.  $\square$

Escolhido o número  $\beta$ , resta observar que pelo fato de  $l(x_1, p) = \frac{p^2}{2} + 0(\delta)$ , podemos iniciar com a função  $u_0(\theta) = \theta$  de modo que  $\dot{u}_0(\theta) = 0$ ,  $p(u(\theta)) = \alpha$  e  $\partial_2 l(\theta, \alpha) = \alpha + 0(\delta)$ ,  $\partial_1 l(\theta, \alpha) = 0(\delta)$  e  $E(u_0) = 0(\delta)$ . Além disso,  $\partial_{22} l = 1 + 0(\delta)$  o que implica que o sistema associado ao método de Newton modificado pode ser resolvido de modo análogo ao descrito na prova do Teorema 17.

Com isso, concluímos a prova do Teorema 16 e do Teorema 15 da Curva Invariante de Moser.

**Exemplo:** Voltando às aplicações do tipo bilhar em uma curva convexa, sabemos que existem pelo menos dois pontos periódicos

de período dois, que correspondem à largura e ao diâmetro. Neste pontos, o ângulo de partida com a tangente é igual a  $\frac{\pi}{2}$ .

Nas coordenadas  $(s, p)$ , com  $p = \cos(\phi)$ , vemos que  $(s_0, 0)$  é um ponto periódico do tipo elíptico se e somente se

$$L_0 - R_0 - R_1 < 0 \text{ e } (L_0 - R_0)(L_0 - R_1) > 0,$$

onde  $L_0$  é o comprimento da corda,  $R_0$  e  $R_1$  são os raios de curvatura nas respectivas extremidades.

Para verificar se  $(s_0, 0)$  é um ponto fixo de  $f^2$  que possui o primeiro invariante de Birkhoff não nulo, e que portanto satisfaz às hipóteses do Teorema de Twist de Moser, é necessário calcular a derivada terceira de  $f^2$  em  $(s_0, 0)$ . Na referência [26], este cálculo é feito e é demonstrado que para curvas convexas genéricas com pontos periódicos de período dois elípticos, existem curvas invariantes que circundam estes pontos ("ilhas elípticas").

## Capítulo 4

# Região de instabilidade

Nesta parte, suporemos que  $f : S^1 \times [0, 1] \rightarrow S^1 \times [0, 1]$  é uma aplicação do tipo twist que preserva área, tal que para qualquer função contínua  $\psi : S^1 \rightarrow [0, 1]$  não identicamente nula nem identicamente igual a 1, o gráfico de  $\psi$  não é invariante por  $f$ .

Pelo teorema da curva invariante de Birkhoff, ver capítulo 3, os resultados que obtivermos aqui poderão ser aplicados para a região entre duas curvas rotacionais invariantes disjuntas de uma aplicação do tipo twist qualquer, que preserve área, desde que não existam outras curvas invariantes no interior da região.

Mas dada  $f : S^1 \times \mathbb{R} \rightarrow S^1 \times \mathbb{R}$ , o conjunto das curvas invariantes rotacionais para  $f$  é um conjunto fechado. Então, se  $f$  possuir pelo menos duas destas curvas e a região entre elas não for folheada por curvas rotacionais invariantes, poderemos encontrar um anel, cujo bordo é formado pela união de duas curvas rotacionais invariantes, em cujo interior não existe nenhuma outra.

Inicialmente, vamos introduzir um pouco de notação para este capítulo:

1.  $A = S^1 \times [0, 1]$ ,  $C^- = S^1 \times \{0\}$  and  $C^+ = S^1 \times \{1\}$
2.  $\mathcal{D} =$  funções contínuas  $\psi : S^1 \rightarrow [0, 1]$  que verificam  $0 < \psi < 1$
3. dado  $z = (x, y) \in A$ ,  $V^-(z) = \{x\} \times [0, y]$  e  $V^+(z) = \{x\} \times [y, 1]$

4. dada  $\psi \in \mathcal{D}$ , seja  $U_\psi^{-(+)} = \{(x, y) \in S^1 \times [0, 1] : 0 \leq y \leq \psi(x) \text{ (}\psi(x) \leq y \leq 1)\}$
5.  $\pi_1, \pi_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$  são as projeções canônicas, dadas por  $\pi_1(\tilde{x}, \tilde{y}) = \tilde{x}$  e  $\pi_2(\tilde{x}, \tilde{y}) = \tilde{y}$
6. os levantamentos para  $\mathbb{R}^2$  ou para  $\mathbb{R} \times [0, 1]$  de aplicações e conjuntos do cilindro ou anel serão denotados com um  $\sim$

Vamos agora enunciar um teorema que é apenas uma outra forma de escrever o teorema da curva invariante de Birkhoff.

**Teorema 18.** *Seja  $V \subset A \subset S^1 \times \mathbb{R}$  um conjunto fechado conexo, tal que  $(S^1 \times \mathbb{R}) \setminus V$  tem ao menos 2 componentes conexas, uma delas contendo o fim inferior do cilindro e a outra contendo o fim superior, e ainda  $f(V) = V$ . Então:*

- i)  $V = C^-$
- ii)  $V = C^+$
- iii)  $V \supset C^- \cup C^+$

O teorema acima implica o seguinte resultado:

**Lema 28.** *Dada  $\psi \in \mathcal{D}$ ,  $C^+ \subset \text{fecho}(\cup_{n=0}^{\infty} f^n(U_\psi^-))$ .*

*Demonstração.* Seja  $V = \text{fecho}(\cup_{n=-\infty}^{\infty} f^n(U_\psi^-))$ . Então  $V$  é fechado, conexo,  $f$ -invariante, contém  $C^-$  e não é igual a  $C^-$ , logo pelo teorema acima contém  $C^+$ . Vamos agora notar o seguinte:

$$\cup_{k=-\infty}^{\infty} f^k(U_\psi^-) = \cup_{m=0}^{\infty} f^{-m}(\cup_{n=0}^{\infty} f^n(U_\psi^-))$$

É imediato ver que  $f^{-(m+1)}(\cup_{n=0}^{\infty} f^n(U_\psi^-)) \supset f^{-m}(\cup_{n=0}^{\infty} f^n(U_\psi^-))$ , assim o conjunto  $\cup_{k=-\infty}^{\infty} f^k(U_\psi^-)$  pode ser escrito como uma reunião crescente de abertos, todos com a mesma área, pois  $f$  preserva área. Assim, como

$$\cup_{k=-\infty}^{\infty} f^k(U_\psi^-) \supset \cup_{n=0}^{\infty} f^n(U_\psi^-)$$

e ambos os conjuntos são abertos com a mesma área, os seus fechos coincidem e portanto o lema está provado.

□

Vamos agora introduzir um conceito bastante importante no estudo de aplicações do tipo twist.

## 4.1 Caminhos positivos e negativos

**Definição:** Um caminho positivo partindo de  $C^-$  (resp.  $C^+$ ) é um arco simples  $\gamma : [0, 1] \rightarrow A$  de classe  $C^1$  tal que:

- i)  $\gamma(0) \in C^-$  (resp.  $C^+$ )
- ii) o levantamento a  $\mathbb{R}$  do ângulo que a tangente em  $\gamma(t)$  faz com o vetor  $(0, 1)$  (resp.  $(0, -1)$ ) é sempre estritamente positivo, para  $t \in [0, 1]$

Observação: Para definir um caminho negativo, basta trocar estritamente positivo por estritamente negativo na definição anterior.

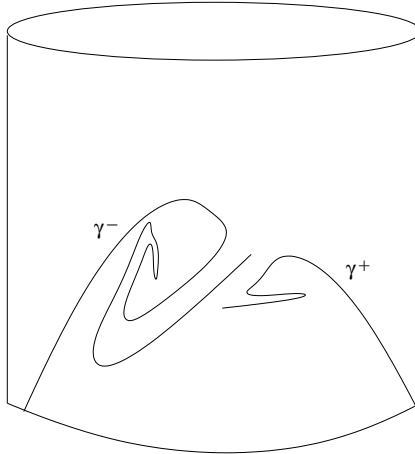


Figura 4.1: Caminhos Positivos e Negativos.

Vamos agora enunciar dois resultados, cujas demonstrações serão deixadas como exercício para o leitor (ambas podem ser encontradas em [19] e em [45]).

**Proposição 15.** *A imagem por  $f^{-1}$  de um caminho positivo é um caminho positivo e a imagem por  $f$  de um caminho negativo é um caminho negativo.*

Apenas como comentário, gostaríamos de dizer que o argumento que prova esse resultado é bastante semelhante ao da demonstração da proposição 19.

**Proposição 16.** *Seja  $U$  um conjunto fechado conexo cujo complementar possui pelo menos 2 componentes conexas, uma contendo  $C^-$  e disjunta de  $C^+$ . Suponha que existam caminhos  $\gamma_-$  e  $\gamma_+$  saindo de  $C^-$ , respectivamente negativo e positivo, que terminam num mesmo ponto  $z$  sem intersectar  $U$ . Nesse caso,  $V^-(z) \cap U = \emptyset$ .*

## 4.2 Teorema fundamental

Nesta seção iremos demonstrar o teorema fundamental sobre regiões de instabilidade. Existem alguns enunciados distintos, mas um dos mais interessantes é o seguinte:

**Teorema 19.** *Existe um ponto em  $A$  cujo conjunto  $\alpha$ -limite está contido em  $C^-$  e o  $\omega$ -limite está contido em  $C^+$ . Em outras palavras, existe um ponto de  $A$  cuja órbita positiva converge para  $C^+$  e a órbita negativa para  $C^-$ .*

Após introduzir uma série de lemas e proposições auxiliares que serão utilizados na demonstração deste teorema, citaremos outros resultados que podem ser provados com as mesmas idéias e técnicas.

Começemos com uma função  $\psi \in \mathcal{D}$  e vamos considerar os seguintes conjuntos fechados

$$\alpha_{\psi}^- \subset \left( \bigcap_{n=0}^{\infty} f^n(U_{\psi}^-) \right) \text{ e } \omega_{\psi}^+ \subset \left( \bigcap_{n=0}^{\infty} f^{-n}(U_{\psi}^+) \right)$$

que são definidos da seguinte forma:

- como  $C^- \subset \left( \bigcap_{n=0}^{\infty} f^n(U_{\psi}^-) \right)$  e  $C^+ \subset \left( \bigcap_{n=0}^{\infty} f^{-n}(U_{\psi}^+) \right)$  e ambas as intersecções anteriores são fechadas, vamos chamar de  $\alpha_{\psi}^-$  a componente conexa de  $\left( \bigcap_{n=0}^{\infty} f^n(U_{\psi}^-) \right)$  que contém  $C^-$  e de  $\omega_{\psi}^+$  a componente conexa de  $\left( \bigcap_{n=0}^{\infty} f^{-n}(U_{\psi}^+) \right)$  que contém  $C^+$ .

Temos então a seguinte proposição:

**Proposição 17.** *Para toda  $\psi \in \mathcal{D}$  valem as seguintes propriedades:*

- i)  $f^{-1}(\alpha_{\psi}^{-}) \subset \alpha_{\psi}^{-}$  e  $f(\omega_{\psi}^{+}) \subset \omega_{\psi}^{+}$
- ii)  $\lim_{n \rightarrow \infty} f^{-n}(\alpha_{\psi}^{-}) = C^{-}$  e  $\lim_{n \rightarrow \infty} f^n(\omega_{\psi}^{+}) = C^{+}$
- iii) o gráfico de  $\psi$  intersecta  $\alpha_{\psi}^{-}$  e  $\omega_{\psi}^{+}$

*Demonstração.* Como  $f^{-1}(\cap_{n=0}^{\infty} f^n(U_{\psi}^{-})) \subset (\cap_{n=0}^{\infty} f^n(U_{\psi}^{-}))$ ,  $C^{-} \subset \alpha_{\psi}^{-}$  e  $f(C^{-}) = C^{-} \Rightarrow f^{-1}(\alpha_{\psi}^{-}) \subset \alpha_{\psi}^{-}$ . Um argumento análogo serve para mostrar que  $f(\omega_{\psi}^{+}) \subset \omega_{\psi}^{+}$ . Assim i) está provado.

Para provar ii) vamos observar que  $f^{-(n+1)}(\alpha_{\psi}^{-}) \subset f^{-n}(\alpha_{\psi}^{-})$  para todo  $n > 0$  e portanto quando  $n \rightarrow \infty$ ,  $f^{-n}(\alpha_{\psi}^{-})$  converge a  $(\cap_{n=0}^{\infty} f^{-n}(\alpha_{\psi}^{-})) \subset U_{\psi}^{-}$ , que é fechado, conexo, contém  $C^{-}$  e é  $f$ -invariante. Como  $U_{\psi}^{-}$  é disjunto de  $C^{+}$ , pelo teorema 18,

$$(\cap_{n=0}^{\infty} f^{-n}(\alpha_{\psi}^{-})) = C^{-}.$$

O outro limite se prova de maneira análoga.

Para provar iii) notemos que se para  $0 < \epsilon < 1$  definimos  $\psi_{\epsilon}(x) = \epsilon$ , para todo  $x \in S^1$ , do lema 28 temos que

$$C^{+} \subset \text{fecho}(\cup_{n=0}^{\infty} f^n(U_{\psi_{\epsilon}}^{-}))$$

e portanto o conjunto  $\cup_{n=0}^{\infty} f^n(U_{\psi_{\epsilon}}^{-})$  encontra o gráfico de  $\psi$ . Suponhamos que  $\epsilon > 0$  seja suficientemente pequeno de maneira que  $\psi_{\epsilon} < \psi$ .

Seja então  $N_{\epsilon} > 0$  o primeiro inteiro tal que  $f^{N_{\epsilon}}(U_{\psi_{\epsilon}}^{-})$  encontra o gráfico de  $\psi$ . Como  $f^{N_{\epsilon}}(U_{\psi_{\epsilon}}^{-})$  é um aberto conexo, é conexo por caminhos, assim existe um arco  $\Gamma_{\epsilon} \subset f^{N_{\epsilon}}(U_{\psi_{\epsilon}}^{-}) \cap U_{\psi}^{-}$  cujos extremos estão, um em  $C^{-}$  e o outro no gráfico de  $\psi$ . Da escolha de  $N_{\epsilon}$ , temos que  $f^{N_{\epsilon}-k}(U_{\psi_{\epsilon}}^{-})$  não intersecta o gráfico de  $\psi$  para  $k = 1, 2, \dots, N_{\epsilon}$ . Assim,  $f^{-k}(\Gamma_{\epsilon}) \subset U_{\psi}^{-}$ , para  $k = 1, 2, \dots, N_{\epsilon}$ . Portanto

$$\Gamma_{\epsilon} \subset \cap_{k=0}^{N_{\epsilon}} f^k(U_{\psi}^{-}). \quad (4.1)$$

Agora vamos notar o seguinte:

1. quando  $\epsilon \rightarrow 0$ ,  $N_\epsilon \rightarrow \infty$
2. como o anel  $S^1 \times [0, 1]$  é compacto, o conjunto dos subconjuntos fechados do anel é compacto na topologia de Hausdorff (ver [23]). Assim, existe uma sequência  $\epsilon_n \rightarrow 0$  tal que  $\Gamma_{\epsilon_n} \rightarrow \Gamma$ , um fechado conexo que intersecta  $C^-$  e o gráfico de  $\psi$ . Da expressão (4.1),  $\Gamma \subset \bigcap_{k=0}^{\infty} f^k(U_\psi^-)$  e portanto  $\alpha_\psi^-$  intersecta o gráfico de  $\psi$ . A outra intersecção é demonstrada de maneira análoga.

□

Uma consequência simples das definições acima é que o conjunto  $\alpha$ -limite dos pontos de  $\alpha_\psi^-$  está contido em  $C^-$  e o conjunto  $\omega$ -limite dos pontos de  $\omega_\psi^+$  está contido em  $C^+$ . Vamos agora definir as seguintes funções de  $S^1$  em  $[0, 1]$ :

$$\begin{aligned} g_{\alpha_\psi^-}(x) &= \sup\{\pi_2(z) : z \in \alpha_\psi^- \text{ e } \pi_1(z) = x\} \\ g_{\omega_\psi^+}(x) &= \inf\{\pi_2(z) : z \in \omega_\psi^+ \text{ e } \pi_1(z) = x\} \end{aligned} \tag{4.2}$$

Uma das possíveis formas de concluirmos a demonstração do teorema seria mostrar que os gráficos das duas funções definidas acima se intersectam (verifique!). No entanto, por razões técnicas, faremos algo um pouco diferente.

**Lema 29.** *Para toda  $\psi \in \mathcal{D}$  valem as seguintes propriedades:*

- 1) os gráficos de  $g_{\alpha_\psi^-}$  e  $g_{\omega_\psi^+}$  intersectam o gráfico de  $\psi$  e  $g_{\alpha_\psi^-} \leq \psi \leq g_{\omega_\psi^+}$
- 2)  $g_{\alpha_\psi^-}$  é semi-contínua superiormente e  $g_{\omega_\psi^+}$  é semi-contínua inferiormente
- 3)  $g_{\alpha_\psi^-}$  e  $g_{\omega_\psi^+}$  são ambas contínuas pela esquerda
- 4) existem  $\varphi, \phi \in \mathcal{D}$  tais que  $g_{\alpha_\varphi^-} \geq \psi$  e  $g_{\omega_\phi^+} \leq \psi$

*Demonstração.* A propriedade 1) acima se verifica trivialmente a partir da definição de  $\alpha_\psi^-$  e  $\omega_\psi^+$  e da propriedade iii) da proposição 17.

Dizemos que uma função real  $g$  é semi-contínua superiormente (inferiormente) em  $x$ , se dado  $\epsilon > 0$ , existir  $\delta > 0$  tal que para  $|x - y| < \delta \Rightarrow g(y) < g(x) + \epsilon$  ( $g(y) > g(x) - \epsilon$ ).



Como os conjuntos  $\alpha_{\psi}^{-}$  e  $\omega_{\psi}^{+}$  são fechados, a definição das funções  $g_{\alpha_{\psi}^{-}}$  e  $g_{\omega_{\psi}^{+}}$  (4.2) implica na validade da propriedade 2). Vamos agora enunciar uma proposição que será usada na prova de 3).

**Proposição 18.** *Seja  $z = (x, y) \in A$  tal que  $g_{\alpha_{\psi}^{-}}(x) < y \leq \psi(x)$ . Então, existe  $z' = (x, y')$  verificando  $g_{\alpha_{\psi}^{-}}(x) < y' \leq y$  tal que  $f(V^{+}(f^{-1}(z'))) \cap \alpha_{\psi}^{-} = \emptyset$ .*

Observação: A hipótese desta proposição é bastante razoável porque caso  $g_{\alpha_{\psi}^{-}}(x) = \psi(x)$  para todo  $x \in S^1$ , então pela propriedade iii) da proposição 17 os gráficos de  $g_{\alpha_{\psi}^{-}}$  e  $g_{\omega_{\psi}^{+}}$  se intersectariam e o resultado fundamental, teorema 19, estaria provado.

*Demonstração.* Seja  $z = (x, y)$  como no enunciado da proposição e vamos considerar o conjunto  $\alpha_{\psi}^{-} \cup (\{x\} \times [g_{\alpha_{\psi}^{-}}(x), y])$ . Como ele é conexo e é distinto de  $\alpha_{\psi}^{-}$ , ele não está contido em  $\bigcap_{n=0}^{\infty} f^n(U_{\psi}^{-})$ . Portanto existe  $z' = (x, y')$  tal que  $g_{\alpha_{\psi}^{-}}(x) < y' \leq y \leq \psi(x)$  e um inteiro  $k \geq 0$  tal que  $f^{-k}(z') \notin U_{\psi}^{-}$ . Como  $y' \leq \psi(x)$ ,  $k$  é estritamente positivo. Deste modo,  $f^{-1}(V^{+}(z'))$  define um caminho positivo  $\gamma_1$  que começa em  $C^{+}$  e chega em  $f^{-1}(z')$ , que não intersecta  $f^{-1}(\alpha_{\psi}^{-})$ . Por outro lado, como  $f^{-k}(z') \notin U_{\psi}^{-}$ , temos que  $V^{+}(f^{-k}(z')) \cap \alpha_{\psi}^{-} = \emptyset$  e como  $f^{-k}(\alpha_{\psi}^{-}) \subset \alpha_{\psi}^{-}$ , finalmente  $V^{+}(f^{-k}(z')) \cap f^{-k}(\alpha_{\psi}^{-}) = \emptyset$ . Desta forma,  $f^{k-1}(V^{+}(f^{-k}(z'))) \cap f^{-1}(\alpha_{\psi}^{-}) = \emptyset$ . Agora temos 2 casos, se  $k > 1$ ,  $f^{k-1}(V^{+}(f^{-k}(z')))$  define um caminho negativo  $\gamma_2$  que também começa em  $C^{+}$  e termina em  $f^{-1}(z')$  e se  $k = 1$ ,  $f^{k-1}(V^{+}(f^{-k}(z'))) = V^{+}(f^{-1}(z'))$ . Em ambos os casos, graças a proposição 16 temos que  $V^{+}(f^{-1}(z')) \cap f^{-1}(\alpha_{\psi}^{-}) = \emptyset$ , assim  $f(V^{+}(f^{-1}(z'))) \cap \alpha_{\psi}^{-} = \emptyset$  o que prova a proposição.  $\square$

Vamos agora continuar a demonstração da propriedade 3) do lema 29. Fixemos  $x \in S^1$  e seja  $x_n$  uma sequência que converge para  $x$  pela esquerda tal que  $g_{\alpha_{\psi}^{-}}(x_n) \rightarrow r$ . Como  $g_{\alpha_{\psi}^{-}}$  é semi-contínua superiormente, então  $r \leq g_{\alpha_{\psi}^{-}}(x)$ . Vamos supor que  $r < g_{\alpha_{\psi}^{-}}(x)$ . Como chegaremos a um absurdo, concluiremos que  $r = g_{\alpha_{\psi}^{-}}(x)$  e portanto  $g_{\alpha_{\psi}^{-}}$  é contínua pela esquerda.

Como  $r < g_{\alpha_{\psi}^{-}}(x) \leq \psi(x)$  e  $\psi$  é função contínua, se  $n > 0$  for suficientemente grande,  $g_{\alpha_{\psi}^{-}}(x_n) < \psi(x_n)$ . Seja agora  $z = (x, r)$  e  $z_n = (x_n, r_n)$ , tal que para todo  $n \geq N$ , (para uma escolha conveniente de  $N$ ),  $g_{\alpha_{\psi}^{-}}(x_n) < r_n \leq \psi(x_n)$  e  $r_n \rightarrow r$ . Vamos agora aplicar a proposição 18 e notar que existe uma sequência  $z'_n = (x_n, r'_n)$  tal que  $g_{\alpha_{\psi}^{-}}(x_n) < r'_n \leq r_n$  e  $f(V^+(f^{-1}(z'_n))) \cap \alpha_{\psi}^{-} = \emptyset$ . Como  $g_{\alpha_{\psi}^{-}}(x_n) \rightarrow r$ ,  $r_n \rightarrow r$ , fica claro que  $z'_n \rightarrow z$ . Como estamos supondo que  $g_{\alpha_{\psi}^{-}}(x) > r$ , se  $n > 0$  for suficientemente grande, o ponto  $(x, g_{\alpha_{\psi}^{-}}(x))$  esta acima da curva  $f(V^+(f^{-1}(z'_n)))$ , o que contradiz a conexidade de  $\alpha_{\psi}^{-}$ , pois  $V^+(z'_n) \cup f(V^+(f^{-1}(z'_n)))$  não intersecta  $\alpha_{\psi}^{-}$ . A demonstração de que  $g_{\omega_{\psi}^{+}}$  é contínua pela direita é análoga.

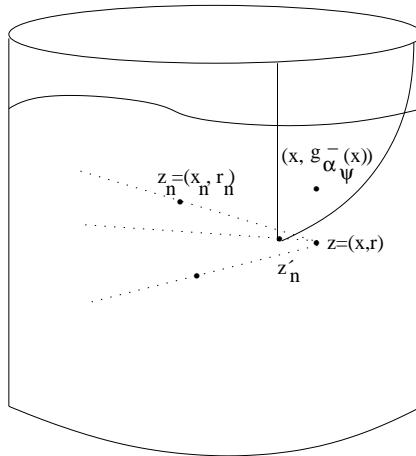


Figura 4.2: Contradiz a conexidade de  $\alpha_{\psi}^{-}$ .

Por fim, vamos provar a propriedade 4). Para cada  $x \in S^1$ , vamos definir os seguintes conjuntos:

$$\Gamma_x = f^{-1}(V^+(x, \psi(x))) \text{ e } H_x = \{z \in A : V^-(z) \cap \Gamma_x \neq \emptyset\}$$

A sequência de conjuntos  $f^n(\alpha_{\psi}^{-})$  é crescente, assim o

$fecho(\cup_{n=0}^{\infty} f^n(\alpha_{\psi}^-))$  é um fechado conexo que separa o anel, não reduzido a  $C^-$  e  $f$ -invariante. Portanto, pelo teorema 18 ele contém  $C^+$ .

Vamos agora tentar entender geometricamente o que é o conjunto  $H_x$ . Se pensarmos no levantamento de  $\Gamma_x$  para a faixa  $\mathbb{R} \times [0, 1]$ , como  $f$  tem a propriedade de twist,  $\tilde{\Gamma}_x$  se projeta injetivamente em  $\mathbb{R}$  e  $\tilde{H}_x$  é o conjunto dos pontos acima deste gráfico. Entendido isso, fica fácil de ver que, como  $fecho(\cup_{n=0}^{\infty} f^n(\alpha_{\psi}^-)) \supset C^+$  e a sequência  $f^n(\alpha_{\psi}^-)$  é crescente, para todo  $x_0 \in S^1$  existe  $n > 0$  tal que  $interior(H_{x_0}) \cap f^n(\alpha_{\psi}^-) \neq \emptyset$ . Mas então para todo  $x$  próximo de  $x_0$ ,  $interior(H_x)$  também intersecta  $f^n(\alpha_{\psi}^-)$ . Desta forma, da compacidade de  $S^1$ , segue que existe um natural  $N > 0$  (a compacidade garante que  $N$  não vai para infinito) tal que  $f^N(\alpha_{\psi}^-) \cap interior(H_x) \neq \emptyset$  para todo  $x \in S^1$ .

Vamos agora escolher  $\varphi' \in \mathcal{D}$  tal que  $f^N(\alpha_{\psi}^-)$  está abaixo do gráfico de  $\varphi'$ . Isto claramente é possível, pois  $f^N(\alpha_{\psi}^-)$  é um fechado que não intersecta  $C^+$ . Seja agora  $\alpha_{\varphi'}^-$  a componente conexa de  $\cap_{k=0}^{\infty} f^k(U_{\varphi'}^-)$  que contém  $C^-$ . Vamos mostrar que  $\alpha_{\varphi'}^-$  intersecta  $\Gamma_x$ , para todo  $x \in S^1$ . Inicialmente vamos notar que  $f^{-k+N}(\alpha_{\psi}^-) \subset f^N(\alpha_{\psi}^-) \subset U_{\varphi'}^-$ , para todo  $k \geq 0$ . Assim,  $f^N(\alpha_{\psi}^-) \subset f^k(U_{\varphi'}^-)$  e portanto  $f^N(\alpha_{\psi}^-) \subset \cap_{k=0}^{\infty} f^k(U_{\varphi'}^-)$ . Como  $f^N(\alpha_{\psi}^-) \supset C^-$ , então  $\alpha_{\varphi'}^- \supset f^N(\alpha_{\psi}^-)$ .

Seja agora o seguinte conjunto, para  $x \in S^1$  fixado:

$$B = \{z \in \Gamma_x : V^+(z) \cap \alpha_{\varphi'}^- = \emptyset\}$$

Como  $\alpha_{\varphi'}^-$  é fechado,  $B$  é um aberto de  $\Gamma_x$ . Mas  $B$  não pode ser todo  $\Gamma_x$ , pois  $\alpha_{\varphi'}^- \supset f^N(\alpha_{\psi}^-)$  e  $f^N(\alpha_{\psi}^-) \cap H_x \neq \emptyset$ . Assim, uma das componentes conexas de  $B$  se escreve como  $f^{-1}(\{x\} \times ]r, 1])$ , com  $\psi(x) < r < 1$ . Como  $g_{\alpha_{\varphi'}^-}$  é contínua pela esquerda, se  $g_{\alpha_{\varphi'}^-}(\pi_1 \circ f^{-1}(x, r)) > \pi_2 \circ f^{-1}(x, r)$ , então para algum  $r < r_1 < 1$ , teríamos  $(\pi_1 \circ f^{-1}(x, r_1), g_{\alpha_{\varphi'}^-}(\pi_1 \circ f^{-1}(x, r_1))) \in V^+(f^{-1}(x, r_1))$ , o que contradiz a escolha de  $r$ . Assim,  $g_{\alpha_{\varphi'}^-}(\pi_1 \circ f^{-1}(x, r)) = \pi_2 \circ f^{-1}(x, r)$  porque da escolha de  $r$ ,  $V^+(f^{-1}(x, r)) \cap \alpha_{\varphi'}^- \neq \emptyset$ . Desta forma,  $\Gamma_x \ni f^{-1}(x, r) \in \alpha_{\varphi'}^-$  o que nos diz que  $f(\Gamma_x) = V^+(x, \psi(x))$  intersecta

$f(\alpha_{\varphi'}^-)$  para todo  $x \in S^1$ . Seja agora  $\varphi \in \mathcal{D}$  tal que  $f(\alpha_{\varphi}^-)$  esteja abaixo do gráfico de  $\varphi$ . Como feito anteriormente, a definição de  $\alpha_{\varphi}^-$  como sendo a componente conexa de  $\bigcap_{k=0}^{\infty} f^k(U_{\varphi}^-)$  que contém  $C^-$  implica no seguinte, para todo  $k \geq 0$ :

$$f^{-k+1}(\alpha_{\varphi'}^-) \subset f(\alpha_{\varphi'}^-) \subset U_{\varphi}^- \Rightarrow f(\alpha_{\varphi'}^-) \subset f^k(U_{\varphi}^-),$$

assim como  $f(\alpha_{\varphi'}^-) \supset C^-$ , obtemos que  $\alpha_{\varphi}^- \supset f(\alpha_{\varphi'}^-)$ . Portanto  $\alpha_{\varphi}^-$  intersecta cada um dos conjuntos  $V^+(x, \psi(x))$ ,  $x \in S^1$  e  $g_{\alpha_{\varphi}^-} \geq \psi$ . A existência de  $\phi$  é provada de maneira análoga.  $\square$

Agora, finalmente estamos prontos para provar o teorema principal.

*Demonstração.* (do teorema 19)

Dada  $\psi \in \mathcal{D}$ , mostraremos que existe  $\varphi \in \mathcal{D}$  tal que  $\alpha_{\varphi}^-$  intersecta  $\omega_{\psi}^+$  em pelo menos um ponto. Como já foi visto, o  $\alpha$ -limite de um ponto em  $\alpha_{\varphi}^-$  está contido em  $C^-$  e o  $\omega$ -limite de um ponto de  $\omega_{\psi}^+$  está contido em  $C^+$ , assim isto concluirá a demonstração do teorema.

Vamos fixar  $\psi \in \mathcal{D}$  e supor que para toda  $\varphi \in \mathcal{D}$ ,  $\alpha_{\varphi}^- \cap \omega_{\psi}^+ = \emptyset$ . Assim, dada  $\varphi \in \mathcal{D}$  vamos considerar os seguintes conjuntos:

$$K_1 = \{x \in S^1 : g_{\alpha_{\varphi}^-}(x) < g_{\omega_{\psi}^+}(x)\}$$

$e$

$$K_2 = \{x \in S^1 : g_{\alpha_{\varphi}^-}(x) > g_{\omega_{\psi}^+}(x)\}$$

Do fato de estarmos supondo que  $\alpha_{\varphi}^- \cap \omega_{\psi}^+ = \emptyset$ , obtemos que  $S^1 = K_1 \cup K_2$ . No que se segue, mostraremos que obrigatoriamente um dos conjuntos acima é vazio.

Assim, por absurdo vamos supor ambos não-vazios. Como  $g_{\alpha_{\varphi}^-}$  é semi-contínua superiormente e  $g_{\omega_{\psi}^+}$  é semi-contínua inferiormente,  $K_1$  é aberto. Por outro lado, como  $g_{\alpha_{\varphi}^-}$  e  $g_{\omega_{\psi}^+}$  são ambas contínuas pela esquerda, obtemos que se  $x_2 \in K_2$  então todo ponto suficientemente próximo, a esquerda de  $x_2$ , também pertence a  $K_2$ . Vamos agora considerar o levantamento destes conjuntos para o plano, isto é, se definirmos a seguinte aplicação de recobrimento

$$p : \mathbb{R} \rightarrow S^1,$$

então  $\tilde{K}_1 = p^{-1}(K_1)$  e  $\tilde{K}_2 = p^{-1}(K_2)$  satisfazem  $\mathbb{R} = \tilde{K}_1 \cup \tilde{K}_2$ . Vamos agora considerar  $\tilde{x}_2 \in \tilde{K}_2$  e seja  $\tilde{x}_0 = \sup \tilde{K}_1 \cap ]-\infty, \tilde{x}_2]$ . Temos então duas possibilidades:

i)  $\tilde{x}_0 \in \tilde{K}_2$ . Neste caso, como já comentamos acima, todo ponto, suficientemente próximo, a esquerda de  $\tilde{x}_0$ , também pertence a  $\tilde{K}_2$ , o que é uma contradição, pois existem pontos arbitrariamente próximos de  $\tilde{x}_0$ , à sua esquerda, que pertencem a  $\tilde{K}_1$ .

ii)  $\tilde{x}_0 \in \tilde{K}_1$ . Como  $K_1$  é aberto, existe uma vizinhança de  $\tilde{x}_0$  inteiramente contida em  $\tilde{K}_1$ , mas isto contradiz a definição deste ponto.

Logo um dos elementos de  $\{K_1, K_2\}$  é vazio.

Portanto podemos particionar o conjunto  $\mathcal{D}$  da seguinte maneira:

$\mathcal{D} = \mathcal{D}_\gamma \cup \mathcal{D}_f$ , onde:

$$\mathcal{D}_\gamma = \{\varphi \in \mathcal{D} : g_{\alpha_\varphi^-} < g_{\omega_\psi^+}\} \text{ e } \mathcal{D}_f = \{\varphi \in \mathcal{D} : g_{\alpha_\varphi^-} > g_{\omega_\psi^+}\}$$

A demonstração será concluída ao mostrarmos que  $\mathcal{D}_\gamma$  e  $\mathcal{D}_f$  são ambos não-vazios e abertos de  $\mathcal{D}$  na topologia da convergência uniforme. Mas isto é uma contradição, pois  $\mathcal{D}$  é claramente um conjunto conexo na topologia da convergência uniforme. Assim, existe  $\varphi \in \mathcal{D}$  tal que  $\alpha_\varphi^- \cap \omega_\psi^+ \neq \emptyset$  e portanto o teorema está provado. Vamos então a prova de que  $\mathcal{D}_\gamma$  e  $\mathcal{D}_f$  são ambos não-vazios e abertos.

1. sobre  $\mathcal{D}_\gamma$ . As funções  $g_{\alpha_\varphi^-}$  e  $\varphi$  coincidem em pelo menos um ponto. Assim, se  $\varphi \in \mathcal{D}_\gamma$ , existe  $x \in S^1$  tal que  $g_{\omega_\psi^+}(x) > \varphi(x)$ . Claramente esta desigualdade vale para todas as funções  $\varphi'$  pertencentes a uma vizinhança suficientemente pequena de  $\varphi$ . Assim, como  $g_{\alpha_{\varphi'}^-} \leq \varphi'$  e  $\varphi' \in \mathcal{D}_\gamma \cup \mathcal{D}_f$ , obtemos que  $\varphi' \in \mathcal{D}_\gamma$  e portanto  $\mathcal{D}_\gamma$  é aberto. Claramente  $\mathcal{D}_\gamma$  é não vazio, pois da definição de  $\alpha_\varphi^-$  e de  $\omega_\psi^+$ ,  $g_{\alpha_\varphi^-} \leq \varphi$  e  $g_{\omega_\psi^+} \geq \psi$ . Assim se  $\varphi < \psi$ , então  $\varphi \in \mathcal{D}_\gamma$ .
2. sobre  $\mathcal{D}_f$ . Seja  $\varphi \in \mathcal{D}_f$ . Como  $\alpha_\varphi^-$  e  $\omega_\psi^+$  são ambos fechados que estamos supondo disjuntos, existe  $\eta > 0$  tal que  $g_{\omega_\psi^+} < g_{\alpha_\varphi^-} - \eta$ . Como  $g_{\alpha_\varphi^-} \leq \varphi$ , então  $g_{\omega_\psi^+} < \varphi - \eta$ . Seja agora uma vizinhança de  $\varphi$  em  $\mathcal{D}$  tal que para toda  $\varphi'$  nessa vizinhança vale  $g_{\omega_\psi^+} <$

$\varphi' - \eta/2$ . Como  $g_{\alpha_{\varphi'}}^-$  e  $\varphi'$  coincidem em pelo menos um ponto  $x' \in S^1$ , então  $g_{\omega_{\psi}}^+(x') < g_{\alpha_{\varphi'}}^-(x') = \varphi'(x)$ , o que implica que  $\varphi' \in \mathcal{D}_f$ . Assim  $\mathcal{D}_f$  é aberto. Como  $\psi$  e  $g_{\omega_{\psi}}^+$  coincidem em pelo menos um ponto, se  $\varphi \in \mathcal{D}$  for tal que  $g_{\alpha_{\varphi}}^- > \psi$  (tal  $\varphi$  existe pela propriedade 4 do lema 29), então como  $\varphi \in \mathcal{D}_\gamma \cup \mathcal{D}_f$ , obteremos que  $\varphi \in \mathcal{D}_f$  e portanto este conjunto também é não vazio.

Isto conclui a prova do teorema. □

### 4.3 Um pouco mais sobre a região de instabilidade

Através de argumentos semelhantes aos desenvolvidos acima é possível mostrar também o seguinte resultado:

**Teorema 20.** *Dada uma vizinhança de  $C^+$ , existe um ponto nessa vizinhança cujos conjuntos  $\alpha$  e  $\omega$ -limite estão contidos em  $C^-$  e vice-versa.*

No fundo, estes resultados mostram que existem órbitas passeando pelo anel todo, com os comportamentos mais diversos possíveis.

Tudo que foi feito acima pode ser encontrado em [44].

### 4.4 Extensões para o cilindro infinito

Inicialmente, vamos salientar que os resultados acima podem ser estendidos para difeomorfismos do tipo twist definidos no cilindro. Vamos enunciar uma destas extensões, cuja demonstração pouco difere da do teorema 19 e depois vamos analisar o caso de difeomorfismos do cilindro que induzem aplicações no toro com detalhe.

Seja  $f : C \rightarrow C$  um difeomorfismo do tipo twist no cilindro, que preserva área e é exato, isto é:

- para todo aberto  $A \subset C$ , homeomorfo a  $C$  e com bordo conexo e compacto, temos que

$$\text{Área}(f(A) \setminus A) = \text{Área}(A \setminus f(A))$$

Note que esta definição de exatidão coincide com a da sec. 1.3. Então vale o seguinte:

**Teorema 21.** *Se  $f$  não possui curvas rotacionais invariantes, existem  $z, w \in C$  tais que  $\pi_2 \circ f^n(z) \xrightarrow{n \rightarrow \pm\infty} \pm\infty$  e  $\pi_2 \circ f^n(w) \xrightarrow{n \rightarrow \pm\infty} \mp\infty$ .*

No que se segue, vamos considerar uma classe especial de difeomorfismos do toro, que num certo sentido generaliza a aplicação Standard.

#### 4.4.1 Aplicações do tipo twist no toro

Seja *Dehn\_twist* o conjunto dos difeomorfismos  $\tilde{f}$  do plano tais que:

$$\tilde{f}(\tilde{x}, \tilde{y}) \text{ tem a propriedade de twist, isto é: } \partial_{\tilde{y}} \pi_1 \circ \tilde{f}(\tilde{x}, \tilde{y}) > K > 0, \text{ e}$$

$$\begin{cases} \tilde{f}(\tilde{x} + 1, \tilde{y}) = \tilde{f}(\tilde{x}, \tilde{y}) + (1, 0) \\ \tilde{f}(\tilde{x}, \tilde{y} + 1) = \tilde{f}(\tilde{x}, \tilde{y}) + (1, 1) \end{cases}$$

É fácil ver que  $\tilde{f} \in \text{Dehn\_twist}$  induz difeomorfismos  $f$  e  $\bar{f}$  respectivamente no cilindro  $S^1 \times \mathbb{R} = (\mathbb{R}/\mathbb{Z}) \times \mathbb{R}$  e no toro  $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ .

O resultado que provaremos para essa classe de aplicações pode ser pensado como uma extensão do teorema 21. Ele tem algumas aplicações interessantes no estudo de certos problemas sobre aplicações do tipo twist no toro, como por exemplo o estudo da ruptura de curvas rotacionais invariantes para famílias e o problema da ergodicidade. Para mais detalhes, ver por exemplo [3], [4] e [6].

**Teorema 22.** *Dada  $\tilde{f} \in \text{Dehn\_twist}$  tal que a aplicação  $f : C \rightarrow C$  induzida por  $\tilde{f}$  é exata e não possui círculos rotacionais invariantes, então existem dois números,  $\rho^- < 0 < \rho^+$  tais que para todo racional  $p/q$ , com  $\rho^- < p/q < \rho^+$ , existe  $z \in C$  tal que  $f^q(z) = z + (0, p)$ .*

Observações:

1. É claro que à órbita de  $z$  corresponde uma órbita  $q$ -periódica para  $\bar{f}$
2. para um irracional  $\omega$  entre  $\rho^-$  e  $\rho^+$  também podemos associar um conjunto compacto invariante para  $\bar{f}$ , que quando levantado para o cilindro, possui a seguinte propriedade: a órbita de seus pontos percorrem  $C$  verticalmente com velocidade  $\omega$ , ver [3]

Para provar esse teorema, inicialmente vamos apresentar alguns resultados auxiliares.

Para cada  $q \geq 1$ ,  $\tilde{x} \in \mathbb{R}$ , vamos definir:

$$\mu_q(t) = \tilde{f}^q(\tilde{x}, t), \text{ para } t \in \mathbb{R} \quad (4.3)$$

Diremos, para  $\mu_q(t)$  definida como em (4.3), que o primeiro encontro de  $\mu_q$  com a vertical por  $\tilde{x}_0$ , se dá para

$$t_P \in \mathbb{R}, \text{ tal que:} \\ t_P = \min\{t \in \mathbb{R} : \pi_1 \circ \mu_q(t) = \pi_1 \circ \tilde{f}^q(\tilde{x}, t) = \tilde{x}_0\}$$

Analogamente, o último encontro de  $\mu_q$  com a vertical por  $\tilde{x}_0$ , se dá para

$$t_U \in \mathbb{R}, \text{ tal que:} \\ t_U = \max\{t \in \mathbb{R} : \pi_1 \circ \mu_q(t) = \pi_1 \circ \tilde{f}^q(\tilde{x}, t) = \tilde{x}_0\}$$

É claro que, para todo  $\tilde{x}, \tilde{x}_0 \in \mathbb{R}$ ,  $t_P \leq t_U$ .

Então, vale (ver [47]):

**Proposição 19.** *Para todo  $\tilde{x}_0, \tilde{x} \in \mathbb{R}$ , seja  $\mu_q(t) = \tilde{f}^q(\tilde{x}, t)$ , como em (4.3). Então valem as seguintes desigualdades:  $\pi_2 \circ \mu_q(t_U) \leq \pi_2 \circ \mu_q(t) \leq \pi_2 \circ \mu_q(t_P)$ , para todo  $t \in \mathbb{R}$  tal que  $\pi_1 \circ \mu_q(t) = \tilde{x}_0$ .*

*Demonstração.* Para  $q = 1$  é evidente, pois  $\mu_1(t)$  encontra cada vertical em um único ponto (da condição de twist).

Para  $q > 1$ , a prova será feita por indução, assim, suponha que o resultado seja válido para  $1, 2, \dots, q-1$ .

Seja  $\tilde{z}_0 = (\tilde{x}_0, \tilde{y}_0) = \mu_q(t_P)$ , o primeiro ponto de encontro de  $\mu_q$  com a vertical por  $\tilde{x}_0$ .

Seja  $\tilde{z}_1 = (\tilde{x}_1, \tilde{y}_1) = \mu_{q-1}(t_P) = \tilde{f}^{-1}(\tilde{z}_0)$ . Então, vale :

$$\mu_{q-1} (] - \infty, t_P]) \cap \tilde{f}^{-1} (\tilde{x}_0, [\tilde{y}_0, +\infty]) = \emptyset. \quad (4.4)$$

Vamos agora provar (4.4).

Suponha, por absurdo, que  $\tilde{z} \in \mu_{q-1} (] - \infty, t_P]) \cap \tilde{f}^{-1} (\tilde{x}_0, [\tilde{y}_0, +\infty])$ . Então, seja  $\tilde{w} = \tilde{f}^{-q+1}(\tilde{z}) \Rightarrow \tilde{w} = (\tilde{x}, T)$ , com  $T < t_P$ . Mas então,  $\mu_q(T) = \tilde{f}^q(\tilde{w}) = \tilde{f}(\tilde{z}) \in \text{vertical por } \tilde{x}_0$ , absurdo, pois  $T < t_P$ .



Suponha agora que existe  $\tilde{w} \in \mu_{q-1} (] - \infty, t_P[)$ , com  $\pi_1(\tilde{w}) > \tilde{x}_1$ . Como  $\lim_{t \rightarrow -\infty} \pi_1 \circ \mu_{q-1}(t) = -\infty$ , temos que existe  $\hat{t} < t_P$ , tal que  $\pi_1 \circ \mu_{q-1}(\hat{t}) = \tilde{x}_1$  e, da hipótese de indução,  $\pi_2 \circ \mu_{q-1}(\hat{t}) > \tilde{y}_1 = \pi_2 \circ \mu_{q-1}(t_P)$ .

Por outro lado, como  $\tilde{f} \in \text{Dehn\_twist}$ , temos que :

$$\begin{aligned} \pi_1 \left( \tilde{f}^{-1}(\tilde{x}_0, ]\tilde{y}_0, +\infty[) \right) &= ] - \infty, \tilde{x}_1[ \\ \pi_2 \left( \tilde{f}^{-1}(\tilde{x}_0, ]\tilde{y}_0, +\infty[) \right) &\supseteq ]\tilde{y}_1, \infty[ \end{aligned}$$

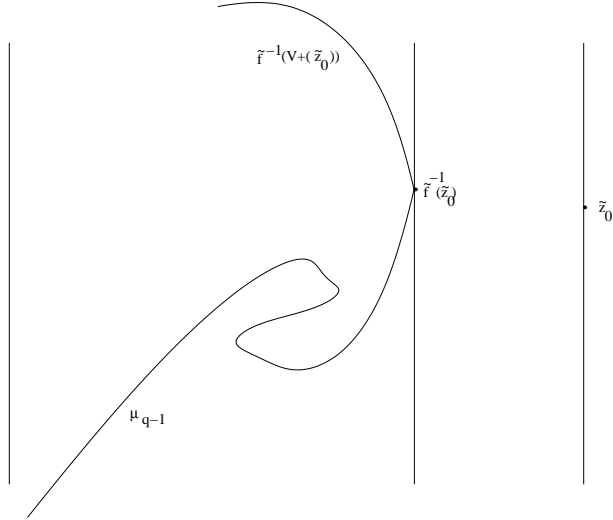


Figura 4.3: último ponto de encontro.

O que implica que, existe  $t_* < \hat{t} < t_P$ , tal que :

$$\mu_{q-1}(t_*) \in \tilde{f}^{-1}((\tilde{x}_0, ]\tilde{y}_0, +\infty[)),$$

absurdo, pois contradiz (4.4). Dessa forma,  $\mu_{q-1} (] - \infty, t_P[)$  está a esquerda da vertical por  $\tilde{x}_1$  e portanto  $\mu_{q-1}(t_P)$  é o primeiro ponto de encontro de  $\mu_{q-1}$  com a vertical por  $\tilde{x}_1$ , sendo assim o de altura máxima, pela hipótese de indução.

Suponha agora, que existe  $t_{\#} > t_P$ , tal que  $\pi_1 \circ \mu_q(t_{\#}) = \tilde{x}_0$  e,  $\pi_2 \circ \mu_q(t_{\#}) > \tilde{y}_0 = \pi_2 \circ \mu_q(t_P)$ .

Mas então,  $\mu_{q-1}(t_{\#}) \in \tilde{f}^{-1}((\tilde{x}_0, ]\tilde{y}_0, +\infty[)) \Rightarrow$  existe  $t_* > t_P$ , tal que  $\pi_1 \circ \mu_{q-1}(t_*) = \tilde{x}_1$  e,  $\pi_2 \circ \mu_{q-1}(t_*) > \pi_2 \circ \mu_{q-1}(t_P)$ , o que contradiz  $\mu_{q-1}(t_P)$  ser o ponto de altura máxima na vertical por  $\tilde{x}_1$ . Assim,  $\mu_q(t_P)$  é o ponto de máxima altura na vertical por  $\tilde{x}_0$ . A prova da proposição para o último ponto de encontro ( $t_U$ ) é análoga.  $\square$

Vamos agora relembrar a proposição (9). Dados  $s$  inteiro e  $q$  natural não-nulo,  $C(s, q) \subset K(s, q)$ , onde  $K(s, q)$  é como em (2.11) e  $C(s, q)$  é um compacto conexo que separa o cilindro. Definamos agora, para cada  $x \in S^1$ , as seguintes funções:

$$\begin{aligned} \mu^-(x) &= \min\{\pi_2(z) : z \in K(s, q) \text{ e } \pi_1(z) = x\} \\ &\quad \text{e} \\ \mu^+(x) &= \max\{\pi_2(z) : z \in K(s, q) \text{ e } \pi_1(z) = x\} \end{aligned} \quad (4.5)$$

Podemos associar funções análogas a  $f^q(K(s, q))$  :

$$\begin{aligned} \nu^-(x) &= \min\{\pi_2(z) : z \in f^q(K(s, q)) \text{ e } \pi_1(z) = x\} \\ &\quad \text{e} \\ \nu^+(x) &= \max\{\pi_2(z) : z \in f^q(K(s, q)) \text{ e } \pi_1(z) = x\} \end{aligned} \quad (4.6)$$

Um corolário trivial da proposição 19 é o seguinte:

**Corolário 7.**  $f^q(x, \mu^-(x)) = (x, \nu^+(x))$  e  $f^q(x, \mu^+(x)) = (x, \nu^-(x))$ .

O corolário acima, juntamente com o próximo resultado são um exemplo da enorme rigidez que a condição de twist acarreta.

**Lema 30.** Para todo  $x \in S^1$ ,  $(x, \mu^{\pm}(x)) \in C(s, q)$ .

*Demonstração.* É uma consequência trivial do corolário 7 juntamente com o fato de  $C(s, q)^c$  ter pelo menos duas componentes conexas, uma contendo o fim inferior do cilindro, denotada por  $U$ , e a outra contendo o fim superior,  $V$ . No caso de, por exemplo  $(x, \mu^-(x)) \notin C(s, q)$ , então  $(x, \mu^-(x)) \in U$ , o que implica que  $(x, \nu^+(x)) = f^q(x, \mu^-(x)) \in f^q(U)$ , absurdo. O outro caso é análogo.  $\square$

O lema que se segue é a base da prova do teorema 22.

**Lema 31.** *Dada  $\tilde{f} \in \text{Dehn\_twist}$ , tal que  $f$  é exata e não tem curvas rotacionais invariantes, existe  $\tilde{z} = (\tilde{x}, \tilde{y}) \in [0, 1]^2$ , tal que  $\tilde{f}^q(\tilde{z}) = \tilde{z} + (s, des)$ , para algum  $q \in \mathbb{N}^*$ ,  $s \in \mathbb{Z}$  e  $des > 2$ .*

*Demonstração.* Inicialmente, vamos tomar constantes  $K > 0$ ,  $a > 0$  e  $b > 0$  tais que para todo  $\tilde{z} \in \mathbb{R}^2$  valem as seguintes desigualdades:

- $\pi_2 \circ \tilde{f}(\tilde{z}) - \pi_2(\tilde{z}) > -a$
- $\partial_{\tilde{y}} \tilde{f}_x(\tilde{z}) > K$
- $\left| \partial_{\tilde{x}} \tilde{f}_x(\tilde{z}) \right| < b$

Com  $\tilde{f}$  não possui C.R.I.'s, por um argumento análogo ao da prova do lema 28, existe  $\tilde{w} = (\tilde{x}', \tilde{y}') \in [0, 1]^2$  tal que  $\pi_2 \circ f^N(\tilde{w}) > 4 + a + \frac{2+b}{K}$ , para algum  $N > 1$ .

Seja  $r = \{(\tilde{x}, \tilde{y}) \in [0, 1]^2 : \tilde{x} = \tilde{x}'\}$

Definindo agora:

$$C.H.Max(\tilde{f}^n(r)) = \sup_{a, b \in [0, 1]} \left| \pi_1 \circ \tilde{f}^n(\tilde{x}', a) - \pi_1 \circ \tilde{f}^n(\tilde{x}', b) \right| \quad (4.7)$$

Fica claro que

$$C.H.Max(\tilde{f}^n(r)) \geq \left| \pi_1 \circ \tilde{f}^n(\tilde{x}', 0) - \pi_1 \circ \tilde{f}^n(\tilde{x}', 1) \right| = n. \quad (4.8)$$

Logo para todo  $n > 1$ , existe pelo menos um inteiro  $s$  tal que  $\tilde{x}' + s \in \pi_1(\tilde{f}^n(r))$ .

A prova será feita agora, por contradição. Assim vamos supor que,  $\forall \tilde{z} \in r$ , tal que

$$\pi_1 \circ \tilde{f}^n(\tilde{z}) = \tilde{x}' \pmod{1}, \text{ para algum } n > 0, \quad (4.9)$$

então

$$\pi_2 \circ \tilde{f}^n(\tilde{z}) < 4 \quad (4.10)$$

Seja agora  $\tilde{w}_1 \in r$ , tal que:

$$\begin{aligned} \pi_2 \circ \tilde{f}^N(\tilde{w}_1) &= 4 + a \\ &e, \\ \forall \tilde{z} \in \{ \text{segmento de } r \text{ entre } \tilde{w} \text{ e } \tilde{w}_1 \}, \\ \pi_2 \circ \tilde{f}^N(\tilde{z}) &> 4 + a \end{aligned}$$

É claro que tal  $\tilde{w}_1$  existe, pois, como já foi dito, existe pelo menos um inteiro  $s$  tal que  $\tilde{x}' + s \in \pi_1(\tilde{f}^N(r))$ . E de (4.9) e (4.10),  $\tilde{f}^N(r)$  cruza a reta  $l$  dada por:

$$l = \{(\tilde{x}, 4 + a), \text{ com } \tilde{x} \in \mathbb{R}\}$$

Definindo agora  $I = \{\text{intervalo de extremos } \pi_2(\tilde{w}) \text{ e } \pi_2(\tilde{w}_1)\}$ , temos (também de (4.9) e (4.10)):

$$\sup_{\tilde{y}_1, \tilde{y}_2 \in I} \left| \pi_1 \circ \tilde{f}^N(\tilde{x}', \tilde{y}_1) - \pi_1 \circ \tilde{f}^N(\tilde{x}', \tilde{y}_2) \right| < 1$$

Seja agora  $\gamma_N : I \rightarrow \mathbb{R}^2$  a seguinte curva :

$$\gamma_N(t) = \tilde{f}^N(\tilde{x}', t), \text{ com } t \in I$$

Então temos:

$$\pi_2 \circ \gamma_N(\pi_2(\tilde{w})) - \pi_2 \circ \gamma_N(\pi_2(\tilde{w}_1)) > \frac{(2 + b)}{K}$$

Agora vamos observar que:

**Proposição 20.** *Seguindo a notação anterior, dada uma curva  $\gamma : J = [a, b] \rightarrow \mathbb{R}^2$ , tal que :*

$$\sup_{t, s \in J} |\pi_1 \circ \gamma(t) - \pi_1 \circ \gamma(s)| < 1 \quad (4.11)$$

$$|\pi_2 \circ \gamma(b) - \pi_2 \circ \gamma(a)| > \frac{(2 + b)}{K} \quad (4.12)$$

Então, existe  $s$  inteiro tal que  $\tilde{x}' + s \in \pi_1 \circ \tilde{f}(\gamma(J))$ .

*Demonstração.*

$$\begin{aligned} & \sup_{t,s \in J} \left| \pi_1 \circ \tilde{f}(\gamma(t)) - \pi_1 \circ \tilde{f}(\gamma(s)) \right| = \\ & = \sup_{t,s \in J} \left| \tilde{f}_x(\gamma(t)) - \tilde{f}_x(\gamma(s)) \right| \geq \left| \tilde{f}_x(\gamma(a)) - \tilde{f}_x(\gamma(b)) \right| \geq \\ & \geq -b + K \cdot \frac{(2+b)}{K} = 2 \end{aligned}$$

Assim, a proposição está provada.  $\square$

É claro que,  $\gamma_N(t)$  satisfaz as hipóteses da proposição (20), por construção. Assim, existe  $s$  inteiro tal que  $\tilde{x}' + s \in \pi_1 \circ \tilde{f}(\gamma_N(I)) = \pi_1 \circ \tilde{f}^{N+1}(\tilde{x}', I)$ .

Por outro lado, como

$$\inf_{t \in I} \pi_2 \circ \gamma_N(t) = \pi_2 \circ \gamma_N(\pi_2(\tilde{w}_1)) = 4 + a,$$

da escolha de  $a > 0$  temos:

$$\inf_{t \in I} \pi_2 \circ \tilde{f}(\gamma_N(t)) > 4$$

Assim, está provado que existe  $t^* \in I$  e  $\tilde{z}^* = (\tilde{x}', t^*) \in r$ , tal que:

$$\begin{aligned} \pi_1 \circ \tilde{f}^{N+1}(\tilde{z}^*) &= \tilde{x}' \pmod{1} \\ \pi_2 \circ \tilde{f}^{N+1}(\tilde{z}^*) &> 4 \end{aligned}$$

O que contradiz (4.9) e (4.10) e portanto conclui a prova.  $\square$

*Observação :*

• É claro que uma demonstração análoga nos dá que existe  $\tilde{w} = (\tilde{x}, \tilde{y}) \in [0, 1]^2$ , tal que  $\tilde{f}^q(\tilde{w}) = \tilde{w} + (s, des)$ , para algum  $q \in \mathbb{N}^*$ ,  $s \in \mathbb{Z}$  e  $des < -2$ .

Já estamos prontos para provar o teorema 22.

*Demonstração.* (do teorema 22)

A demonstração pode ser dividida em 2 casos:

i)  $0 < p/q < \rho^+$                       ii)  $\rho^- < p/q < 0$

Como os 2 casos são análogos, analisaremos apenas o i).

Já que  $f : C \rightarrow C$  é exata e não possui círculos rotacionais invariantes, pelo lema 31, existe  $\tilde{z} = (\tilde{x}, \tilde{y}) \in [0, 1]^2$ , tal que  $\tilde{f}^q(\tilde{z}) =$

$\tilde{z} + (s, des)$ , para algum  $q \in \mathbb{N}^*$ ,  $s \in \mathbb{Z}$  e  $des > 2$ . Assim, a  $\tilde{z}$  corresponde um ponto  $z = (x, y) \in K(s, q) \subset C$ . Pela definição das funções  $\mu^\pm$  e  $\nu^\pm$ , ver (4.5) e (4.6), temos que  $\nu^+(x) - \mu^-(x) \geq des > 2$ .

Por outro lado, como  $f$  é exata, ela satisfaz a propriedade PIC. Agora basta notar que:

- $(x, \mu^-(x)) \in C(s, q)$ , pelo lema 30
- pela prova do teorema 10, existe  $z_s \in C(s, q)$  tal que  $f^q(z_s) = z_s$
- $C(s, q)$  é conexo

Das 3 propriedades acima, segue que a função  $\pi_2 \circ f^q(\bullet) - \pi_2(\bullet) - 1$  se anula em pelo menos um ponto de  $C(s, q)$ , digamos  $z_1$ . Mas então  $f^q(z_1) = z_1 + (0, 1)$ .

Vamos agora provar a segunda parte do teorema. Suponhamos que existe  $w \in C$  tal que

$$\lim_{n \rightarrow \infty} \frac{\pi_2 \circ f^n(w) - \pi_2(w)}{n} = \omega > 0. \quad (4.13)$$

Seja  $0 < p/q < \omega$ . Suponha que  $H(\bullet) = \pi_2 \circ f^q(\bullet) - \pi_2(\bullet) - p \mid_{C(0, q)} < 0$ . Como acima, existe  $z_0 \in C(0, q)$  tal que  $f^q(z_0) = z_0$ . Assim caso  $H$  não se anule em  $C(0, q)$ , como  $H(z_0) = -p < 0$  e  $C(0, q)$  é conexo,  $H$  deve ser negativa em todo ponto de  $C(0, q)$ .

Assim, para todo  $x \in S^1$ ,  $\nu^+(x) < \mu^-(x) + p$ . Como  $C(0, q)$  é fechado, existe uma função contínua  $\theta : S^1 \rightarrow \mathbb{R}$  tal que o seu gráfico separa  $f^q(C(0, q))$  de  $C(0, q) + (0, p)$ , em outras palavras,  $f^q(C(0, q))$  está abaixo do seu gráfico e  $C(0, q) + (0, p)$  está acima.

Vamos chamar de  $U$  a componente conexa de  $(\text{gráfico}(\theta))^c$  que contém o fim inferior do cilindro. Por construção,  $(f^q(U) - (0, p)) \subset U$ . Mas isto implica que

$$\limsup_{n \rightarrow \infty} \frac{\pi_2 \circ f^n(r) - \pi_2(r)}{n} \leq p/q, \text{ para todo } r \in C,$$

o que contradiz (4.13). Assim  $H$  se anula e o teorema está provado.  $\square$

# Referências Bibliográficas

- [1] Abraham R. e Marsden J. (1978): *Foundations of mechanics*, Perseus Books, Cambridge, Mass.
- [2] Addas-Zanata e Grotta-Ragazzo C. (2001): Critical number in scattering and escaping problems in classical mechanics *Physical Review E* **64**
- [3] Addas-Zanata S. (2002): On the existence of a new type of periodic and quasi-preriodic orbits for twist maps of the torus. *Nonlinearity* **15**, 1399-1416
- [4] Addas-Zanata S. (2005): On properties of the vertical rotation interval of twist mappings. A aparecer em *Erg. Th. and Dyn. Sys*
- [5] Addas-Zanata S. e Grotta-Ragazzo C. (2002): On the stability of some periodic orbits of a new type for twist maps. *Nonlinearity* **15**, 1385-1397
- [6] Addas-Zanata S. (2004): Stability for the vertical rotation interval of twist mappings. *preprint*
- [7] Addas-Zanata S. e Grotta-Ragazzo C. (2004): Conservative dynamics: unstable sets for saddle-center loops. *J. Differential Equations* **197**, 118-146

- [8] Alpern S. and Prasad V. (1993): Combinatorial proofs of the Conley-Zehnder-Franks theorem on a fixed point for torus homeomorphisms. *Adv. Math.* **99**, 238-247
- [9] Angenent S. (1990): Monotone recurrence relations, their Birkhoff orbits and topological entropy. *Erg. Th. and Dyn. Sys.* **10**, 15-41
- [10] Arnold V. (1978): *Mathematical methods of classical mechanics*, Springer-Verlag, NY
- [11] Arnold V. I. (1988): *Geometrical methods in the theory of ordinary differential equations*, 2nd ed., Springer-Verlag, New-York.
- [12] Aubry S. and Le Daeron P. Y. (1983): The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states. *Phys. D* **8**, 381-422.
- [13] Bangert V. (1988): Mather sets for twist maps and geodesics on tori. *Dynamics Reported* **1**. Ed. by U. Kirchgraber and H. O. Walther. Wiley and Sons, New York.
- [14] Bartle R. G. (1976): *The elements of real analysis*. John Wiley and Sons, New York.
- [15] Birkhoff G. (1913): Proof of Poincaré's geometric theorem. *Trans. Amer. Math. Soc.* **14**, 14-22
- [16] Birkhoff G. (1920): Surface transformations and their dynamical applications. *Acta Math.* **43**, 1-119
- [17] Birkhoff G. (1925): An extension of Poincaré's last geometric theorem. *Acta Math.* **47**, 297-311
- [18] Birkhoff G. (1932): Sur quelques courbes fermées remarquables. *Bull. Soc. Math. de France* **60**, 1-26
- [19] Birkhoff G. *Collected Math. Papers, Vol. II*, 530-561
- [20] Brown M. and Neumann W. (1977): Proof of the Poincaré-Birkhoff fixed point theorem. *Michigan Math. J.* **24**, 21-31



- [21] Carter P. (1982): An improvement of the Poincaré-Birkhoff fixed point theorem. *Trans. Amer. Math. Soc.* **269**, 285-299
- [22] Chirikov B. V. (1979): A universal instability of many-dimensional oscillator systems. *Physics Reports* **52**, 263-379
- [23] Choquet G. (1969): *Lectures on Analysis*, W. A. Benjamin, Inc. New York **vol II**
- [24] de Melo W. e Palis J. (1982): *Geometric theory of dynamical systems*, Springer Verlag, NY
- [25] de la Llave R. (2001): A tutorial on KAM theory. "Smooth Ergodic Theory and Its Applications", *Proceedings of Symposia in Pure Mathematics*, **69**, American Math Soc., 175-292
- [26] Dias Carneiro M.J., Oliffson Kampfhurst S. e Pinto de Carvalho S. (2003): Elliptical Islands in strictly convex billiards. *Ergodic Theory and Dynamical Systems* **23**, 799-812
- [27] Douady R. (1982): Applications du Théorème des tores invariants. *Thèse 3ème cycle- Université Paris VII*
- [28] Duarte P. (1994): Plenty of elliptic islands for the standard family of area preserving maps. *Ann. Inst. H. Poincaré Anal. Non. Linéaire* **11**, 359-409.
- [29] Fathi A. (1983): Une interprétation plus topologique de la démonstration du théorème de Birkhoff, Appendice du Chap. I de [39]
- [30] Franks, J. (1987): A variation on the Poincaré-Birkhoff theorem. *Contemp. Math.* **81**, 111-117
- [31] Franks, J. (1988): Recurrence and fixed points of surface homeomorphisms. *Erg. Th. and Dyn. Sys.* **8\***, 99-107
- [32] Franks J. (1988): Generalizations of the Poincaré-Birkhoff theorem. *Ann. of Math.* **128**, 139-151
- [33] Golé, C. (2001): *Symplectic Twist Maps- Global Variational Techniques*. *Advanced Series in Nonlinear Dynamics* **18**, World Scientific

- [34] Greene J. M. (1979): A method for computing the stochastic transition. *J. Math. Phys.* **20**, 1183-1201
- [35] Grotta Ragazzo C. (1997): On the Stability of Double Homoclinic Loops. *Comm. Math. Phys.* **184**, 251-272
- [36] Grotta Ragazzo C. (1997): Irregular Dynamics and Homoclinic Orbits to Hamiltonian Saddle Centers. *Comm. Pure App. Math.* **L**, 105-147
- [37] Guillou L. (1994): Théorème de translation plane de Brouwer et généralisations du théorème de Poincaré-Birkhoff. *Topology* **33** , 331-351
- [38] Hassenblatt B. e Katok A. (1995): *Introduction to the modern theory of dynamical systems*, Cambridge University Press, Cambridge
- [39] Herman M.R. (1983): Sur les courbes invariantes par difféomorphismes de l'anneau, *Astérisque* **103-104**, Soc. Math. France.
- [40] Kerekjarto B. (1928-29): The plane translation theorem of Brouwer and the last geometric theorem of Poincaré. *Acta Sci. Math. Szeged.* **4** ,86-102
- [41] Kolmogorov A. N. e Fomin S. V. (1970): *Introductory real analysis*. Prentice-Hall, Englewood Cliffs, N.J.
- [42] Hall G. R. (1984): A topological version of a theorem of Mather on twist maps. *Erg. Th. and Dyn. Sys.* **4**, 585-603.
- [43] Le Calvez P. (1986): Existence d'orbites quasi-périodiques dans les attracteurs de Birkhoff. *Comm. Math. Phys.* **106**, 383-394
- [44] Le Calvez P. (1987): Propriétés Dynamique des Regions D'instabilité. *Ann. Scient. Éc. Norm. Sup.* **4**, **20**, 443-464
- [45] Le Calvez P. (1987): Propriétés des attracteurs de Birkhoff. *Erg. Th. and Dyn. Sys.* **8**, 241-310

- [46] Le Calvez P. (1990): *Etude topologique des applications déviant la vertical*. Ensaios Matemáticos, Soc. Brasileira de Matemática, vol.2
- [47] Le Calvez P. (1991): Propriétés Dynamiques des Difféomorphismes de L'Anneau et du Tore. *Astérisque* **204**
- [48] Lerman L. M. (1991): Hamiltonian systems with loops of a separatrix of a saddle-center. *Sel. Math. Sov.* **10**, 297-306.
- [49] Levi M. e Moser J. (2001): A Lagrangian proof of the invariant curve theorem for twist maps. "Smooth Ergodic Theory an Its Applications", *Proceedings of Symposia in Pure Mathematics* **69**, American Math Soc. 733-746
- [50] Lima E. L. (1993) *Grupo fundamental e espaços de recobrimento*. Projeto Euclides, Soc. Bras. Mat., Rio de Janeiro.
- [51] Mackay R. e Meiss J (1983): Linear satability of periodic orbits in Lagrangian systems. *Phys, Letters A* **98**,92-94
- [52] Mañé R. (1993): *Global variational methods in conservative dynamics*. XIX Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro
- [53] Mather J. N. (1982): Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology* **21**, 457-467.
- [54] Mather J. (1982): Glancing billiards. *Erg. Th. and Dyn. Sys.* **3-4**, 397-403.
- [55] J. N. Mather and G. Forni (1994): Action minimizing orbits in Hamiltonian systems. *Lect. Notes Math.* **1589**, Springer, Berlin, 92-186
- [56] Meiss J. D. (1992): Symplectic maps, variational principles, and transport. *Rev. Modern Phys.* **64**, 795-848.
- [57] Mielke, A., Holmes P., O'Reilly O. (1992): Cascades of homoclinic orbits to, and chaos near, a Hamiltonian saddle center. *J. Dyn. Diff. Eqns.* **4**, 95-126.

- [58] Moser J. e Siegel K. (1971): *Lectures on celestial mechanics*, Springer Verlag, NY
- [59] Moser J. (1973): *Stable and Random Motions in Dynamical Systems*, Princeton University Press, Princeton, NJ
- [60] Moser J. (1986): Monotone twist maps and the calculus of variations, *Ergodic Th. and Dyn. Systems* **6**, 401-413
- [61] Newman M.H.A (1939): *Elements of the topology of plane sets of points*, Cambridge University Press, Cambridge
- [62] Poincaré H. (1912): Sur un théorème de géométrie. *Rend. Circ. Mat. Palermo* **33**, 375-407
- [63] Poschel J. (2001): A Lecture in the Classical KAM Theorem, *Smooth Ergodic Theory and Its Applications*, Proceedings of Symposia in Pure Mathematics, **69**, 707-732
- [64] Tabachnikov S. (1995): *Billiards*. Panoramas et Synthèse, **1**, Soc. Math. France
- [65] Winkelkemper H. (1988): A generalisation of the Poincaré-Birkhoff theorem. *Proc. Amer. Math. Soc.* **102**, 1028-1030