PART I: FORCED GENERATION AND STABILITY OF NONLINEAR WAVES
PART II: CHAOTIC ADVECTION IN A RAYLEIGH-BÉNARD FLOW

Thesis by

Roberto Camassa

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

June, 1990

(submitted December 18, 1989)

Acknowledgements

I would like to thank Professor Ted Wu for his guidance through the first part of this thesis and for his continued encouragement and understanding. The many discussions on this part of the thesis with Dr. George Yates are also thankfully acknowledged. Furthermore, I am grateful to Professor Steve Wiggins for his help and advice throughout the second part. I would also like to thank Professor Jim Knowles for his support during the difficult moments of my "career" as a graduate student. Completing the list of people I was fortunate enough to meet at Caltech and I am indebted with could cause some problem of space. However, to fellow students Kostas Papadimitriou, Luca Cortellezzi and Tasso Kaper in particular, I express my gratitude for helping out in various ways, and even more important, for their friendship. I am indebted to Ms. Helen Burrus in many ways: for her help, warm support and an occasional "boot in the pants" that kept me rolling when enthusiasm flagged. Also, I cannot miss the opportunity of thanking my parents for all they did for me. Last, but not least, I wish to thank my wife, Grazia. Unlike my enthusiasm, her love and constant care never flagged.

Abstract

Part I

The forced Korteweg-de Vries model has been found satisfactory in predicting the periodic generation of upstream-advancing solitary waves by a bottom topography moving in a layer of shallow water with a steady transcritical velocity. It is also known that with certain characteristic forcing distributions, there exist waves, according to the fKdV model, which can remain steady in accompanying the characteristic forcing, provided such a wave exists initially, whereas for a different initial condition the phenomenon of periodic generation can still manifest itself. The stability of two such transcritically forced steady solitary waves is investigated, with their bifurcation diagrams determined with respect to the velocity and the amplitude of the forcing as parameters. The linear stability analysis is first carried out; it involves solving a singular, non-self-adjoint eigenvalue problem, which is examined by applying techniques of matched asymptotic expansions with suitable multiscales for singular perturbations, about the isolated bifurcation points of the parameters. The eigenvalues and eigenfunctions for the full range of the parameters are then obtained by numerically summing a power series expansion for the solution. The numerical results, which accurately match with the local analysis, show that the eigenvalues have only four branches $\sigma = \pm\sigma_r \pm i\sigma_i$. The real part σ_r is nonvanishing for the velocity less than a certain supercritical value and for the amplitude greater than a certain marginal bound except at a single point in the parametric plane at which the external forcings vanish, reducing the forced waves to the classical free solitary wave. Within this parametric range, the real part of the four eigenvalues is algebraically two to five orders smaller than the imaginary part σ_i, wherever σ_i exists; such a small σ_r indicates physically a weak exponential growth rate of perturbed solutions and mathematically the need of a very accurate numerical method for its determination. Beyond this parametric range, linear stability theory appears to fail because no eigenvalues can there be found to exist. In this latter case a non linear analysis based on the functional Hamiltonian formulation is found to prevail, and our analysis predicts stability. Finally, extensive numerical simulations using various finite difference schemes are pursued, with results providing full confirmation to the predictions made in various regimes by the analysis.

We consider the Korteweg-de Vries equation in the semi-infinite real line with a boundary

condition at the origin. The numerical investigations of Chu et al.[2], are revisited and different new forms for the boundary forcing are assumed. In order to provide some qualitative description for the numerical simulations we develop a simple model based on the IST formalism. It is found that the model is also able to provide some quantitative predictions in agreement with the numerical results.

Part II

There has been considerable interest recently on chaotic advection, for the first time explored in the context of Rayleigh-Bénard roll (2D) convection by the experimental work of J. Gollub and collaborators. When the Rayleigh number increases across a (supercritical) value, depending on the wavelength of the rolls, an oscillatory instability sets in. The flow near the onset of the instability can still be modelled by a stream function, which can be split into a time independent part plus a small time dependent perturbation. The motion of fluid particles can therefore be regarded as the flow for a near integrable, "one-and a half" degree of freedom Hamiltonian vector field, with the phase space corresponding to the physical domain. In absence of molecular diffusivity, the evolution of a certain region of phase space can thus be viewed as the motion of a dyed part of fluid, when the tracer is perfectly passive. The most important objects for a theory of transport are the invariant manifolds for the Poincaré map of the flow homoclinic to fixed points, which physically correspond to the stagnation points. As fluid particles cannot cross invariant lines, these curves constitute a sort of "template" for their motion. For the time independent flow, the invariant manifolds connect the stagnation points and define the roll boundaries. Thus, no transport from roll to roll can occur in this case. Switching the perturbation on, these connections are broken and the manifolds are free to wander along the array of rolls. We use segments of stable and unstable manifold to define the time dependent analogue of the roll boundaries. Transport of fluid across a boundary can then be attributed to the way a region bounded by segments of stable and unstable manifold, or "lobe," is evolving under map iterations. This allows us to write explicit formulae for describing the fluid transport in terms of a few of these lobes, for a general cross section defining the Poincaré map. Using the symmetries of special cross sections, we are able to further reduce the number of necessary lobes to just one. Furthermore, these symmetries allow us to derive analytically a lower and upper bound for the first time tracer invades a roll, and a lower bound on the stretching of the interface

between dyed and clear fluid . These results are independent of the fact that the perturbation is small. When this is the case however, the analytical tools of the Melnikov and subharmonic Melnikov functions are available, so that an approximation to the lobe areas and location and size of the island bands can be determined analytically. It turns out that in our case these approximations are quite good, even for relatively large perturbations. The results we have produced regarding the strong dependence of transport on the period of the oscillation suggest an effect for which no experimental verification is currently available. The presence of molecular diffusivity introduces a (long) time scale into the problem. We discuss the applicability of the theory in this situation, by introducing a simple rule for determining when the effects of diffusivity are negligible, and perform numerical simulations of the flow in this case to provide an example.

Contents

Acknowledgements ii

Abstract iii

Table of Contents vi

List of Figures ix

List of Tables xii

I FORCED GENERATION AND STABILITY OF NONLINEAR WAVES 1

1 The stability of forced solitary waves 2

1.1 Introduction 2

1.2 Forced solitary waves 6

1.3 The stability of forced solitary waves 9

1.4 Linear stability analysis 11

1.4.1 A perturbation expansion for the eigenvalues 14

1.4.2 The inner problem for the case of $\mathbf{m = 2}$, $\mu_{\mathbf{m}} = \mathbf{m}^2 = \mathbf{4}$ . . . 16

1.4.3 The outer problem for the case of $\mu = \mathbf{4}$ 18

1.4.4 Global spectral behaviour 22
1.5 Nonlinear Stability 30
1.6 Existence of multiple stationary solutions 37
1.7 Numerical simulations 41
1.8 Conclusions 47
Appendix A 50
References 53

2 The KdV Model With Boundary Forcing 72
2.1 Introduction 72
2.2 Numerical Results 73
2.3 Approximate solutions using IST 76
References 82

II CHAOTIC ADVECTION IN A RAYLEIGH-BÉNARD FLOW 95

3 The Spreading of Passive Tracer in Chaotic Rayleigh-Bénard Flows 96
3.1 Introduction 96
3.2 The mathematical model and transport theory 101
3.2.1 The basic structures governing roll to roll transport 103
3.2.2 The spreading of tracer initially contained in one roll 107
3.2.3 The first visit time and the stretching of the interface 110
3.2.4 Chaotic fluid particle motion 115
3.2.5 The Melnikov method and analytical estimates of the lobe areas 117

3.2.6 The structures and transport within a roll 121

3.2.7 The relative time scales of chaotic advection and molecular diffusion 127

3.3 Numerical simulations for three "canonical" cases 130

3.3.1 Roll concentration of tracer and comparison with a Markov chain model 132

3.3.2 The case of compact phase space 136

3.3.3 The effects of molecular diffusivity 140

Appendix B 142

Appendix C 147

References 152

List of Figures

Figure 1.1 56
Figure 1.2 57
Figure 1.3 58
Figure 1.4 59
Figure 1.5 60
Figure 1.6 61
Figure 1.7 62
Figure 1.8 63
Figure 1.9 64
Figure 1.10 65
Figure 1.11 66
Figure 1.12 67
Figure 1.13 68
Figure 1.14 69
Figure 1.15 70
Figure 1.16 71

Figure 2.1 83
Figure 2.2 84
Figure 2.3 85

Figure 2.4 86
Figure 2.5 87
Figure 2.6 88
Figure 2.7 89
Figure 2.8 90
Figure 2.9 91
Figure 2.10 92
Figure 2.11 93
Figure 2.12 94

Figure 3.1 155
Figure 3.2 156
Figure 3.3 157
Figure 3.4 158
Figure 3.5 159
Figure 3.6 160
Figure 3.7 161
Figure 3.8 162
Figure 3.9 163
Figure 3.10 164
Figure 3.11 165
Figure 3.12 166
Figure 3.13 167
Figure 3.14 168
Figure 3.15 169

Figure 3.16 170
Figure 3.17 171
Figure 3.18 172
Figure 3.19 173
Figure 3.20 174
Figure 3.21 175
Figure 3.22 176
Figure 3.23 177
Figure 3.24 178
Figure 3.25 179
Figure 3.26 180
Figure 3.27 181
Figure 3.28 182
Figure 3.29 183
Figure 3.30 184
Figure 3.31 185
Figure 3.32 186
Figure 3.33 187

List of Tables

3.1 Comparison between the lobe area $\mu(L_{1,0})$ estimated by Melnikov function and numerically, with $A = 0.1$ 119

3.2 Comparison between the lobe area $\mu(L_{1,0})$ estimated by Melnikov function for the case of slip and no-slip boundary conditions, with $A = 0.1$ 121

3.3 Comparison between the island area $\mu(I_3)$ estimated by averaging and numerically for decreasing ϵ, $\omega = 0.6$, with $A = 0.1$. 125

3.4 The numerical estimates for r_T, α, with $A = 0.1$ 134

3.5 Comparison between the numerical estimate of r_T, and the one provided by the series, with $A = 0.1$ 139

Part I

FORCED GENERATION AND STABILITY OF NONLINEAR WAVES

Chapter 1

The stability of forced solitary waves

1.1 Introduction

Considerable attention has been drawn by recent studies to the phenomenon of nonlinear, dispersive waves generated in a soliton-bearing physical system by a moving forcing disturbance sustained at resonance. A remarkable feature of the phenomenon is that to a steadily moving disturbance imposed on such a physical system, the response may not be steady asymptotically with respect to the moving disturbance as would be expected for linear systems, but may result in periodic production of upstream-advancing solitary waves, in a process that may continue indefinitely. An example is given by the effects of a submerged topography or a surface pressure moving with a constant transcritical velocity over the top free surface of a water layer of uniform depth. Analogous phenomena can be expected to occur in various other soliton-supporting systems. A review of history and literature can be found in Wu (1987) and Lee et al. (1988).

A few theoretical models have been proposed for the description of this general class of motion. One of them is the generalized Boussinesq (gB) equation introduced by Wu (1979), which is applicable to wave generation and propagation by three-

dimensional forcing distribution moving in arbitrary manner through a medium which may vary gradually and slowly in two horizontal dimensions. Another approach is based on the director-sheet model of Green and Naghdi, as adopted by Ertekin et al. (1984, 1986). The most appealing theoretical model is perhaps the forced Korteweg de Vries (fKdV) equation which characterizes unidirectional, weakly nonlinear and weakly dispersive long waves being weakly forced at resonance. Because of its simplicity in structure, this model has been employed by Akylas (1984) and Cole (1985) for the pointed forcing and by Lee et al. (1988) and by Wu & Wu (1988) for distributed forcings. Comparisons between theory and experiment have been carefully examined by Lee at al. (1988), with results showing a broad agreement between experiment and various physical models in spite of some refined differences between these models. In addition, the roles played by the nonlinear and the dispersive effects during the periodic generation of upstream waves can be more directly evaluated by integration with mass and energy considerations based on the fKdV model, as illustrated by Grimshaw & Smyth (1986), Wu (1987) and by Lee et al. (1988). However, the basic mechanism underlying the manifestation of the periodically produced, upstream-moving waves remains so far unexplained.

This work is a study intended to explore the basic mechanism in question. It is carried out by first considering the stability and rate of growth of the waves that are furnished energy to grow at the expense of the continuing rate of working by an accompanying forcing agency. The nonlinear stability study is facilitated by employing two relatively simple solutions of the forced stationary solitary wave family found by Wu (1987) for the fKdV model, one of which was indeed proposed by Patoine & Warn (1982) to simulate a meteorological phenomenon.

In Section 1.3 a stability theory is formulated for the two forced steady solitary waves, one being transcritical, with a velocity range spanning across the critical speed, while the other pertaining only to a supercritical forcing, and both reducing to the classical free solitary wave of the KdV family as the forcing vanishes in the limit. Stability of the primary motion, designated in the wave frame by $\zeta_s(x)$ as a function of the one-dimensional space of $x \in \mathbf{R}$, independent of the time t, is examined with an arbitrary perturbation, $\eta(x,t)$, superimposed so that the resultant motion, $\zeta(x,t) = \zeta_s(x) + \eta(x,t)$, satisfies the fKdV equation while the forcing is kept fixed to that corresponding to ζ_s. As shown in Section 1.4, the linear stability analysis for η yields an eigenvalue problem with eigenvalue σ, whose real part gives the rate of growth of the imposed perturbation η. The difficulty of this eigenvalue problem is partly due to the feature that the governing ordinary differential equation, being third in order, is not self-adjoint and must be considered together with its adjoint counterpart.

In terms of certain appropriate similarity variables, the eigenvalue problem for forcings of type (i) forms a family of one parameter in μ, which is a speed detuning parameter, the forcing being supercritical or subcritical according as $\mu > 0$ or < 0. The single parameter for forcings of type (ii) is designated α, which characterizes the forcing amplitude such that the forcing reaches its maximum at $\alpha = 6$ and below this maximum the regimes of $\alpha > 6$ and $\alpha < 6$ constitute the two branches of solutions. For $\mu = 4$ and $\alpha = 12$, the forcings vanish and the two steady motions reduce to a free solitary wave. In addition to these single parameters for the two types of forcing, we can define another parameter, A, which is a measure of the amplitude of the perturbation η and hence characterizes the nonlinear effects.

For small perturbations, the analysis on linear theory for $\eta(x,t)$ given in Section 1.3 shows that the fixed-point solutions of η (corresponding to zero eigenvalue $\sigma = 0$) are found to exist at $\mu = 1, 4$, and 9 with $\alpha = 12$ for forcing (i) and at $\alpha = \nu(\nu+1), \nu = 2, 3, ...,$ with $\mu = 4$ for forcing (ii). Further, these stationary solutions are shown by perturbation techniques to be unstable, with an exponential rate of growth in a vicinity of these particular values of μ and α in the respective parametric plane provided that the characteristic forcing does not vanish (it vanishes at only one point $\mu = 4$ and $\alpha = 12$) and that $\mu < 9$ for case (i) and $\alpha > 6$ for case (ii). This analytical result is fully confirmed by the numerical results obtained in Section 1.4.4 for the global spectral behavior of the eigenvalue σ , which indeed has a nonvanishing real part over the span of (i) $\mu < 9$ (excluding $\mu = 4$) with $\alpha = 12$ and (ii) $\alpha > 6$ (excluding $\alpha = 12$) with $\mu = 4$, respectively. The only exception in this parametric domain is the single point at $\mu = 4$ and $\alpha = 12$ characterizing the free solitary wave, for which case the nonlinear stability has been concluded by Benjamin (1972). Outside this domain, there are two regions: (i) $\mu > 9, \alpha = 12$ and (ii) $\mu = 4, \alpha < 6$, in which both our perturbation and numerical techniques fail to find any nonvanishing eigenvalues. In these cases, however, we are able to utilize the Hamiltonian property of the system and show in Section 1.5 that the two primary motions ζ_s both satisfy a sufficiency criterion to imply stability on nonlinear theory.

Further study along this direction leads to the determination of bifurcation points of the primary motions for forcings of type (i) to other stationary solutions; this is first shown in Section 1.6 for the neighborhoods of $\mu = 1, 4$ and 9 , where approximations based on regular perturbation expansions can be provided for new stationary solutions. In particular, the solution curve for one of these, ζ_{ss}, in the parametric

$\mu - A$ plane, with A being a measure of the amplitude of the perturbation η, appears as a "valley bottom" line in the basin of attraction, to which all perturbed solutions will asymptotically evolve, no longer exchanging energy with the steady forcing. Detailed features of the transient solutions through bifurcation are illustrated in Section 1.7 for various cases with the numerical results so obtained. These results show that the evolution of nonlinear waves generated periodically in unstable regimes is at the rate in agreement with the prediction previously given by Wu (1987). In the stable regime, numerical solutions for initially perturbed motions are found to settle to the new, unique, and nontrivial stationary solution of η at a rate as predicted here by expansion methods. This new development, however, will require further exploration of numerical techniques in order to be continued outside some isolated small regions of the bifurcation diagram.

1.2 Forced solitary waves

In the context of shallow water waves, the fKdV equation is particularly suitable for describing weakly nonlinear, weakly dispersive and weakly forced waves and can be written as

$$\zeta_t + (F-1)\zeta_x - \frac{3}{2}\zeta\zeta_x - \frac{1}{6}\zeta_{xxx} = P_x, \tag{1.1}$$

where ζ is the free surface elevation of the water layer and the external forcing $P(x)$ is given by the sum of the applied surface pressure distribution and bottom topography (see Wu, 1987; Lee, 1985). Here $x \in \mathbb{R}$, $t \in (0, +\infty)$ and subscripts denote partial differentiation. Equation (1.1) is in nondimensional form in which the length, time and pressure have been scaled by h_0, $\sqrt{h_0/g}$, $\rho g h_0$ respectively, with h_0, g, ρ being the undisturbed water depth, gravity acceleration and uniform fluid density, respectively.

The reference frame used for (1.1) is fixed with the steadily moving disturbance, in which the fluid is moving to the right with uniform constant velocity U at $x = -\infty$, corresponding to the Froude number $F = U/\sqrt{gh_0}$, and the forcing distribution is fixed. The range of the physical parameters in which the model is derived (see Lee 1985; Wu 1987) is fixed by the following estimates

$$\bar{\epsilon} \equiv \left(\frac{h_0}{\lambda}\right)^2 \ll 1 \ , \quad \frac{a}{h_0} = O(\bar{\epsilon}) \ , \quad |P| = O(\bar{\epsilon}^2), \quad |F-1| = O(\bar{\epsilon}), \tag{1.2}$$

where λ is a typical wavelength, and a is a typical wave amplitude. The second equation in (1.2) represents the condition that the dispersive and nonlinear effects are so properly balanced that the phenomena can be described by the present theory. The particular numerical values of these parameters, which can be changed through the well known scaling properties of the KdV equation, are chosen in this way to facilitate comparisons with the notation used in the previous works.

The stationary solutions of (1.1) that vanish at infinity satisfy the ordinary differential equation

$$(F-1)\zeta - \frac{3}{4}\zeta^2 - \frac{1}{6}\zeta_{xx} = P(x), \tag{1.3}$$

which is obtained from (1.1) by dropping the time derivative and integrating once in x. Among the possible solutions of (1.3) for various choices of the forcing P, soliton-like solutions ζ_s of the form

$$\zeta(x,t) = \zeta_s(x) = a \ \mathrm{sech}^2(kx)$$

are of particular interest here, and they occur whenever the forcing term is chosen to be a $\mathrm{sech}^2(kx)$ or a $\mathrm{sech}^4(kx)$ distribution, provided their amplitudes and Froude number are related as follows (Wu, 1987):

(i)

$$\begin{aligned} \zeta_s(x) &= \frac{4}{3}k^2 \operatorname{sech}^2(kx) \\ P(x) &= \frac{4}{3}k^2\left(F-1-\frac{2}{3}k^2\right)\operatorname{sech}^2(kx), \end{aligned} \tag{1.4}$$

(ii)

$$\begin{aligned} \zeta_s(x) &= a \operatorname{sech}^2(kx) \\ P(x) &= a\left(k^2-\frac{3}{4}a\right)\operatorname{sech}^4(kx), \quad F-1 \equiv \frac{2}{3}k^2. \end{aligned} \tag{1.5}$$

In these equations, the Froude number F in (1.4) and amplitude a in (1.5), respectively for case (i) and (ii), can be regarded as playing the role of a free parameter, in addition to the parameter k which fixes the length scale of the disturbance. We note that for case (i) the flow can be either supercritical ($F > 1$) or subcritical ($F < 1$); the forcing amplitude, b_1 say, is always negative for $F < 1$, but for supercritical forcings ($F > 1$), $b_1 < 0$ or > 0 according as $(F-1) <$ or $> 2k^2/3$, whilst the resulting $\operatorname{sech}^2(x)$-wave remains positive in polarity for all these cases. For case (ii), the flow can only be supercritical; the forcing amplitude reaches its maximum of $k^4/3$ at the wave amplitude of $a = 2k^2/3$, and for forcing amplitudes below this maximum, there exist two branches of wave amplitudes a for given forcing. Nevertheless, the solution (i), or (ii), is a unique function of F in case (i), or of a in case (ii), and this will be so regarded here. Finally, we note that both solutions reduce, with vanishing forcing, to the free soliton solution

$$\zeta_s = \frac{4}{3}k^2 \operatorname{sech}^2(kx), \quad F-1 = \frac{2}{3}k^2. \tag{1.6}$$

Consistently with the estimates (1.2), k should be a small quantity of order $O(\bar{\epsilon}^{\frac{1}{2}})$. Solutions (1.4) and (1.5) are among the family of forced steady solitary waves

found by Wu (1987), while solution (1.5) was originally reported by Patoine and Warn (1982).

According to the well known uniqueness property for the initial value problem of the KdV equation on the real line, these steady solitary waves will be solutions of the fKdV equation (1.1), each of which may remain permanent in shape provided

$$\zeta(x,0) = \zeta_s(x). \tag{1.7}$$

The question of whether or not these solutions will also have a physical significance when being sufficiently perturbed is closely related to their stability properties, and this is the foremost problem we are going to investigate presently.

1.3 The stability of forced solitary waves

Equation (1.1) can be cast in homogeneous form using either one of the stationary solutions ζ_s corresponding to the $P(x)$ of (1.4) or (1.5) above, by setting

$$\zeta(x,t) = \zeta_s(x) + \eta(x,t), \tag{1.8}$$

so that η satisfies the nonlinear evolution equation,

$$\eta_t + \frac{\partial}{\partial x}\left[(F-1)\eta - \frac{3}{4}\eta^2 - \frac{1}{6}\eta_{xx} - \frac{3}{2}\zeta_s\eta\right] = 0. \tag{1.9}$$

Using the similarity transformation,

$$x' \equiv kx, \quad t' \equiv \frac{1}{6}k^3 t, \quad \eta' \equiv \frac{1}{k^2}\eta, \quad F' - 1 \equiv \frac{1}{k^2}(F-1), \tag{1.10}$$

we can eliminate the parameter k altogether from equation (1.9), and we can rewrite the resulting equation, after omitting the primes for η and the independent variables, as

$$\eta_t + \frac{\partial}{\partial x}\left[\left(\mu - \alpha\ \mathrm{sech}^2(x)\right)\eta - \frac{9}{2}\eta^2 - \eta_{xx}\right] = 0, \tag{1.11}$$

where

$$\mu \equiv 6(F' - 1) \quad \text{and} \quad \alpha = 12 \quad \text{for forcing (i)} \tag{1.12}$$

and

$$\mu = 4 \quad \text{and} \quad \alpha \equiv 9a/k^2 \quad \text{for forcing (ii)}\ . \tag{1.13}$$

Since k is of $O(\epsilon^{\frac{1}{2}})$, the free parameters μ and α are of $O(1)$ and can take on values in a relatively broad range even for F and a being held under the constraints (1.2).

The perturbed motion governed by (1.11) has two leading-order conservation laws. The first is the invariance of the excess mass m,

$$\frac{dm}{dt} = 0\ , \quad m\ \equiv\ \int_{-\infty}^{+\infty} \eta(x,t)dx\ ,$$

which is the first integral of (1.11) with respect to x under the regularity conditions at infinity. Therefore,

$$m\ =\ \text{const.} = m_0, \quad m_0 = \int_{-\infty}^{+\infty} \eta(x,0)dx\ \ , \tag{1.14}$$

m_0 being the initial excess mass. The second relation is for the energy,

$$\frac{dE}{dt} = \alpha \int_{-\infty}^{+\infty} \eta^2 \frac{d}{dx}\ \mathrm{sech}^2(x)dx, \quad E \equiv \int_{-\infty}^{+\infty} \eta^2 dx\ \ , \tag{1.15}$$

which is the first integral of the product of (1.11) with η. Expressing $\eta = \eta_e + \eta_o$, $\eta_e(x,t)$ and $\eta_o(x,t)$ being even and odd functions of x, we have

$$\frac{d}{dt}\int_{-\infty}^{+\infty} (\eta_e^2 + \eta_o^2)dx = 2\alpha \int_{-\infty}^{+\infty} \eta_e \eta_o \frac{d}{dx}\ \mathrm{sech}^2(x)dx \equiv \dot{W}\ \ , \tag{1.16}$$

which signifies that the total energy of the motion increases at the rate $\dot{W}$ equal to the rate of working by the external forcing on the system. A sufficient condition for

the total energy E to be invariant, or $\dot{W} = 0$, is that η is purely even or purely odd. In other words, E may vary only when η has both even and odd components. In such cases, the maximum rate of energy growth is readily found from (1.15) as

$$\frac{dE}{dt} \leq \beta E \ , \quad \text{or} \quad E(t) \leq E(0)e^{\beta t} \ , \quad \beta = \frac{4\alpha}{3\sqrt{3}} \ ,$$

β being the maximum rate of growth for $E(t)$, equal to the maximum of $\alpha \frac{d}{dx} \operatorname{sech}^2(x)$ over the real x. Thus, the total energy $E(t)$ can never grow at a rate faster than β.

1.4 Linear stability analysis

Linear stability analysis of the nonlinear system (1.11)-(1.13) is of significance since a state of growth or decay of $\eta(x,t)$ with a nonzero rate evaluated on linear analysis cannot be altered even with additional nonlinear effects taken into account. The stability of the forced solitary waves ζ_s versus small perturbation can be determined by the linearized form of equation (1.11)

$$\eta_t + \frac{\partial}{\partial x}\left[\mu\eta - \eta_{xx} - \alpha\eta \operatorname{sech}^2(x)\right] = 0. \tag{1.17}$$

where the initial data for η are supposed to be small in maximum norm. Introducing the usual separation of variables

$$\eta(x,t) = e^{\sigma t} f(x) \tag{1.18}$$

where σ and f are in general complex and the real part is understood for physical η's, we obtain the following equation for $f(x)$

$$\frac{d}{dx}\left[\frac{d^2 f}{dx^2} + \left(\alpha \operatorname{sech}^2(x) - \mu\right) f\right] = \sigma f, \tag{1.19}$$

or, in operator form

$$\mathcal{L}_{\mu,\alpha}(x)f = \sigma f \ , \tag{1.20}$$

where

$$\mathcal{L}_{\mu,\alpha}(x) \equiv \frac{d}{dx}\left[\frac{d^2}{dx^2} + \left(\alpha \ \mathrm{sech}^2(x) - \mu\right)\right], \tag{1.21}$$

and the boundary conditions for f will be taken to be regular at infinity, i.e.,

$$f^{(n)}(x) \underset{|x|\to\infty}{\longrightarrow} 0, \quad \text{exponentially fast, for at least } \ n = 0, 1, 2 \ , \tag{1.22}$$

where $f^{(n)}(x)$ denotes the nth derivative of $f(x)$. We note that $\mathcal{L}_{\mu,\alpha}$ is a non-self-adjoint operator. Equation (1.21), together with the regularity conditions at infinity for f, constitutes an eigenvalue problem in σ, and the stability is determined by the signature of the real part of the eigenvalues, Re $\sigma(\mu, \alpha) = 0$ giving the boundary of neutral stability in the (μ, α) space. Furthermore, by defining A as a measure of the perturbation amplitude, such as the initial excess mass pertaining to $\eta(x, 0)$, one can study the boundary of neutral stability in the (μ, A)-space for forcings of type (i), or Re $\sigma(\alpha, A) = 0$ in the (α, A)-space for forcing (ii). It is only in the state of neutral stability, with Re $\sigma = 0$ obtained by linear analysis that no statement can be made to imply nonlinear stability of the perturbed waves.

This eigenvalue problem has several features of basic interest. First, the symmetries possessed by the differential operator in (1.21), namely,

$$\mathcal{L}_{\mu,\alpha}(-x) = -\mathcal{L}_{\mu,\alpha}(x), \quad [\mathcal{L}_{\mu,\alpha}(x)]^* = \mathcal{L}_{\mu,\alpha}(x) \ , \tag{1.23}$$

where $(\cdot)^*$ denotes the complex conjugation, imply that if an eigenvalue σ and its eigenfunction $f(x)$ exist, then the same is true for the eigenvalues $-\sigma$, $\pm\sigma^*$ and their corresponding eigenfunctions $f(-x)$, $f^*(\pm x)$, respectively. Therefore we have instability whenever an eigenvalue whose real part is different from zero can be detected.

Second, the regularity conditions specified in (1.22) imply, upon integration of (1.19), that

$$\int_{-\infty}^{+\infty} f(x)dx = 0, \tag{1.24}$$

except of course when $\sigma = 0$; that is, for nonzero eigenvalues, the mass of the forced solitary waves cannot be changed by perturbations of the form (1.18). This integral condition also shows that the proper eigenfunctions of (1.20) cannot constitute a basis in the linear space of C^3 functions which satisfy the regularity conditions (1.22) at infinity. Finally, we further remark that nothing can be said about the nonlinear stability of these waves in case the real part of an eigenvalue should vanish, since we are only considering the linearized version of equation (1.11).

We define the inner product (f, g) between two well behaved functions $f(x)$ and $g(x)$ by

$$(f, g) = \int_{-\infty}^{+\infty} f(x)\ g^*(x)\ dx \quad . \tag{1.25}$$

Two functions f and g are said to be orthogonal if $(f, g) = 0$. By integration by parts of the inner product $((\mathcal{L}_{\mu,\alpha} - \sigma)f, g^*)$, with g required to be bounded at infinity, we find this product equal to $(f, (\mathcal{L}^\dagger_{\mu,\alpha} - \sigma^*)g^*)$, and thus obtain the adjoint equation as

$$\mathcal{L}^\dagger_{\mu,\alpha}(x)g = \sigma g, \tag{1.26}$$

where

$$\mathcal{L}^\dagger_{\mu,\alpha}(x) \equiv -\left[\frac{d^2}{dx^2} + \left(\alpha\ \mathrm{sech}^2(x) - \mu\right)\right]\frac{d}{dx}\ . \tag{1.27}$$

We note that with $D \equiv d/dx$, $D\mathcal{L}^\dagger_{\mu,\alpha} = -\mathcal{L}_{\mu,\alpha}D$. Hence if $f(x)$ is an eigenfunction of (1.20) with eigenvalue σ, then $g(x)$, called the adjoint eigenfunction of f and defined as the antiderivative of $f(-x)$,

$$g(x) \ = \ \int f(-x)dx\ , \tag{1.28}$$

is a solution of the adjoint equation (1.26) with σ its corresponding eigenvalue. Further, it can be shown that f and its adjoint function g are in general not orthogonal to each other, i.e., $(f,g) \neq 0$. In fact, we have

$$(f,g) \;=\; 2\int_{-\infty}^{+\infty} g_e(x)\frac{d}{dx}g_o(x)dx \tag{1.29}$$

where $g_e(x)$ and $g_o(x)$ are the even and odd components of $g(x)$, and hence the statement. A sufficient condition for $f(x)$ and its adjoint function $g(x)$ to be orthogonal is when f is either even or odd in x.

1.4.1 A perturbation expansion for the eigenvalues

The eigenfunctions corresponding to zero eigenvalues play a significant role in the present stability analysis; they can be determined explicitly since equation (1.20) can be integrated once when $\sigma = 0$, and the problem reduces to that of a Schrödinger equation with $\mathrm{sech}^2(x)$-potential:

$$-(K_\alpha + \mu)f \equiv \frac{d^2f}{dx^2} + \left[\alpha\,\mathrm{sech}^2(x) - \mu\right] f(x) = 0. \tag{1.30}$$

Such eigenvalue problems are, of course, quite familiar in quantum mechanics (see, e.g., Landau and Lifshitz 1958, §21).

For definiteness, let us consider the forcing (i) case first, i.e., set $\alpha = 12$ (see (1.12)). Making in (1.30) the substitution $x = \tanh(z)$ and $\mu = m^2$, we have

$$\frac{d}{dz}(1-z^2)\frac{df}{dz} + \left[12 - \frac{m^2}{1-z^2}\right] f = 0\,. \tag{1.31}$$

The solution of this equation has the form of the associated Legendre function, $f = P_3^m(z)$. Under boundary conditions that f vanishes at $z = \pm 1$, we have eigensolutions

only for the discrete spectrum of $m = 1,2,3$, or $\mu = m^2 = 1,4$ and 9, so that the entire set of eigenfunctions is

$$\mu_1 = 1, \quad f_0(x;\mu_1) = B \operatorname{sech}(x)\left(5\tanh^2(x) - 1\right), \tag{1.32}$$

$$\mu_2 = 4, \quad f_0(x;\mu_2) = B \operatorname{sech}^2(x)\tanh(x), \tag{1.33}$$

$$\mu_3 = 9, \quad f_0(x;\mu_3) = B \operatorname{sech}^3(x), \tag{1.34}$$

where B is an arbitrary constant. Since the eigenvalue σ is zero, none of these eigenfunctions has to satisfy the integral condition (1.24). Assuming that the spectrum has continuous dependence on the parameters, we can search for a perturbation expansion of the eigenvalues and eigenfunctions when μ is in a neighbourhood of the values given above. Defining

$$\mu = \mu_m + s\epsilon, \quad 0 < \epsilon \ll 1, \quad s = \pm 1, \quad m = 1,2 \text{ and } 3\ , \tag{1.35}$$

we can rewrite equation (1.20) as

$$\mathcal{L}_{\mu_m} f - s\epsilon \frac{d}{dx} f = \sigma(\epsilon) f, \tag{1.36}$$

where we have dropped the subscript α since it is fixed and equal to 12. Assume

$$\sigma(\mu) = \phi_1(\epsilon)\sigma_1 + \phi_2(\epsilon)\sigma_2 + \ldots, \tag{1.37}$$

$$f(x;\mu) = f_0(x) + \psi_1(\epsilon) f_1(x) + \psi_2(\epsilon) f_2(x) + \ldots, \tag{1.38}$$

where $f_0(x)$ represents the set of $f_0(x,\mu_m)$ in (1.32)-(1.34), and $\phi_i(\epsilon), \psi_i(\epsilon) \to 0$ as $\epsilon \to 0$, $\phi_{i+1} = o(\phi_i)$ etc., $i = 1,2,\ldots$. The terms of the product σf then show that whenever $\sigma(\mu) \neq 0$ the integral condition (1.24) cannot be satisfied for f_0 given in (1.32) and (1.34) above (except f_0 in (1.33)), and therefore it can already be seen that the perturbation expansion is a singular one. In fact for the "even" function cases

(1.32) and (1.34), the eigenfunction corresponding to zero eigenvalue spans an area of "mass" (which is the integral in (1.24)) different from zero whereas the perturbed one will be required to satisfy (1.24) as soon as the perturbation is switched on. For this reason and after some exploration we take for the "inner" problem

$$\phi_i(\epsilon) = \psi_i(\epsilon) = \epsilon^i, \quad \mu = 1,\ 9\ , \tag{1.39}$$

$$\phi_i(\epsilon) = \psi_i(\epsilon) = \epsilon^{\frac{i}{2}}, \quad \mu = 4\ . \tag{1.40}$$

1.4.2 The inner problem for the case of m = 2, $\mu_{\mathrm{m}} = \mathrm{m}^2 = 4$

For $\mu = 4 + \epsilon s$ we have

$O(1)$

$$\mathcal{L}_4 f_0 = 0, \tag{1.41}$$

$O(\epsilon^{\frac{1}{2}})$

$$\mathcal{L}_4 f_1 = \sigma_1 f_0, \tag{1.42}$$

$O(\epsilon)$

$$\mathcal{L}_4 f_2 = s\frac{d}{dx} f_0 + \sigma_2 f_0 + \sigma_1 f_1, \tag{1.43}$$

$O(\epsilon^{\frac{3}{2}})$

$$\mathcal{L}_4 f_3 = s\frac{d}{dx} f_1 + \sigma_3 f_0 + \sigma_2 f_1 + \sigma_1 f_2 \quad , \tag{1.44}$$

etc., with the boundary conditions

$$f_i(x) \underset{x\to-\infty}{\longrightarrow} 0, \quad i = 0, 1, 2, \ldots . \tag{1.45}$$

The solvability condition for the above equations requires that the right-hand side be orthogonal to $g_0 \equiv B\ \mathrm{sech}^2(x)$, (which is proportional to the integral of $f_0(x, \mu_2)$

given by (1.33) and hence by the Ansatz (1.28), is a solution to the appropriate adjoint problem of (1.27)), namely,

$$(g_0, f_R) = 0 \ , \tag{1.46}$$

where (f, g) denotes the inner product defined by (1.25), and f_R stands for the right-hand side members in (1.42)-(1.44). At order $O(\epsilon^{\frac{1}{2}})$ the solvability condition (1.46) is evidently satisfied, due to the integrand being a total differential, and f_1 can be determined by standard methods (such as by variation of parameters)

$$f_1(x) = \frac{\sigma_1}{4} \operatorname{sech}^2(x) \left[1 - x \tanh(x)\right], \tag{1.47}$$

where we have set the arbitrary constant in (1.33) $B = -2$. The homogenous solution of (1.42), i.e., f_0, need not be included in the expression for f_1, since it can always be absorbed in $\sigma_2 f_0$ at the next order. At order $O(\epsilon)$ the solvability condition determines σ_1 as

$$\sigma_1 = s\frac{||f_0||^2}{(g_0, f_1)} = \sqrt{s}\frac{8}{\sqrt{15}} \ . \tag{1.48}$$

Here $\|\cdot\|$ is the norm induced by the inner product, $\|\cdot\|^2 \equiv (\cdot, \cdot)$, and the factor $\sqrt{s}$ in (1.48) arises from (g_0, f_1) being proportional to σ_1. Thus we see that σ_1 is real for $\mu > 4$, $(s = +1)$, but is purely imaginary for $\mu < 4$, i.e., $s = -1$. We further note that f_1 does not satisfy the integral constraint (1.24), and this is reflected at the next order by the fact that for f_2 we can match the regularity condition at only one of the boundaries at $\pm\infty$, say at $x = -\infty$. This can be seen by integrating (1.43) once and taking the limit $x \to +\infty$,

$$f_2(x) \underset{x \to +\infty}{\longrightarrow} -\frac{\sigma_1}{4}\int_{-\infty}^{+\infty} f_1(x)dx = -\frac{\sigma_1^2}{16} \ . \tag{1.49}$$

The solvability condition for the $O(\epsilon^{\frac{3}{2}})$ problem determines the value of σ_2 as

$$\sigma_2 = -\sigma_1\frac{(g_0, f_2)}{(g_0, f_1)}. \tag{1.50}$$

To evaluate the numerator we do not need the explicit form of f_2 if we make use of the adjoint relationships, and the result is

$$\sigma_2 = -s\frac{4}{15}. \tag{1.51}$$

The eigenvalue thus has in the vicinity of $\mu = 4$ the inner solution

$$\sigma = \epsilon^{\frac{1}{2}}\frac{8}{\sqrt{15s}} - \epsilon\frac{4}{15s} + O(\epsilon^{\frac{3}{2}}) \tag{1.52}$$

which is purely real for $\mu > 4$ ($s > 0$), of order $O(\epsilon^{\frac{1}{2}})$ in magnitude, and is complex for $\mu < 4$ ($s < 0$), with its real part being of $O(\epsilon)$ and its imaginary part of $O(\epsilon^{\frac{1}{2}})$. As shown below, this inner expansion of σ can be matched well with the outer expansion so that the compound eigenfunction satisfies the regularity condition at infinity. The nonvanishing real part of σ will imply instability of the perturbed motion for μ in a neighbourhood of 4, as will be discussed below.

1.4.3 The outer problem for the case of $\mu = 4$

Since the expansion cannot satisfy the boundary conditions (1.22) at all orders, we look for an "outer" problem by defining the multiscale outer variables as

$$x^+ \equiv x\left(\epsilon^{\frac{1}{2}} + O(\epsilon^{\frac{3}{2}})\right), \quad \tilde{x} \equiv x\epsilon, \quad w(x^+, \tilde{x}) \equiv f(x), \tag{1.53}$$

so that

$$\frac{d}{dx} = \epsilon^{\frac{1}{2}}\frac{\partial}{\partial x^+} + \epsilon\frac{\partial}{\partial \tilde{x}} + O(\epsilon^{\frac{3}{2}})\,. \tag{1.54}$$

We take

$$w(x^+, \tilde{x}) = \epsilon w_2(x^+, \tilde{x}) + \epsilon^{\frac{3}{2}} w_3(x^+, \tilde{x}) + \ldots, \tag{1.55}$$

and by substituting these expressions in (1.36), we obtain for the first two terms

$O(\epsilon^{\frac{3}{2}})$

$$-4\frac{\partial}{\partial x^+}w_2 = \sigma_1 w_2, \tag{1.56}$$

$O(\epsilon^2)$

$$-4\frac{\partial}{\partial x^+}w_3 = \sigma_1 w_3 + 4\frac{\partial}{\partial \tilde{x}}w_2 + \sigma_2 w_2, \tag{1.57}$$

where we have neglected the term $\operatorname{sech}^2\left(\epsilon^{-\frac{1}{2}}x^+\right)$, since for x^+ fixed it is exponentially small as $\epsilon \to 0$. Suppressing the secularity from this equation, the first term in the outer expansion can be taken to be

$$w_2(x^+, \tilde{x}) = C \exp\left(-\frac{\sigma_1}{4}x^+ - \frac{\sigma_2}{4}\tilde{x}\right) , \tag{1.58}$$

which satisfies the boundary condition $w_2 \to 0$ for $x^+, \tilde{x} \to +\infty$ for both cases of $s = \pm 1$ by virtue of the expression (1.48) for σ_1, (1.51) for σ_2 and the relation $\tilde{x} = \epsilon^{\frac{1}{2}}x^+$. The constant C is determined by matching with the inner solution. By observing (1.49), we obtain a uniformly valid expansion up to order $O(\epsilon)$ (and up to a multiplicative constant) in the form

$$f(x;\mu) = f_0(x) + \epsilon^{\frac{1}{2}}f_1(x) + \epsilon\left\{f_2(x) - \frac{\sigma_1^2}{16}H(x)\left[\exp\left(-\frac{\sigma_1}{4}\epsilon^{\frac{1}{2}}x - \frac{\sigma_2}{4}\epsilon x\right) - 1\right]\right\} + O(\epsilon^{\frac{3}{2}}), \tag{1.59}$$

where $H(x)$ is the Heaviside step function. The eigenvalue corresponding to this eigenfunction is given by (1.52). The rate of growth of the perturbation (1.18) in the form of $\eta = f(x,\mu)\exp(\sigma t)$ with f given by (1.59) for $\mu \simeq 4$, is however quite different on the two sides of $\mu = 4$, being of $O(\epsilon^{-1})$ for $\mu < 4$ and of $O(\epsilon^{-\frac{1}{2}})$ for $\mu > 4$.

For the other two cases (1.32) and (1.34), the violation of the integral and boundary condition now occurs in both cases at order $O(1)$, $O(\epsilon)$, respectively, of the inner problem. The appropriate solution of the adjoint problem can be taken to be

$$g(x;\mu_m) = \int_{-\infty}^{x} f_0(s;\mu_m)ds, \quad m = 1, 3 , \tag{1.60}$$

and so it is in fact a distribution, being bounded but different from zero as $x \to +\infty$. The analysis relative to these cases is completely analogous to the one for $\mu = 4$, and we only quote the result of the uniform expansion for $\mu = \mu_m + s\epsilon$, to order $O(\epsilon)$:

$$f(x;\mu) = f_0(x;\mu_m) + \epsilon\left\{f_1(x;\mu_m) + \frac{\chi_m\sigma_1}{\mu_m}H(x)\left[1 - \exp\left(\frac{-\sigma_1}{\mu_m}\epsilon x\right)\right]\right\} + O(\epsilon^2), \tag{1.61}$$

where $\chi_m = \int_{-\infty}^{\infty} f_0(x;\mu_m)\,dx$. It should be noticed that this expression not only satisfies the boundary conditions (1.22) at order $O(\epsilon)$ but is also consistent with the integral constraint (1.24) at the same order. The explicit form of f_1 is not important at this order, only the limits for $x \to \pm\infty$ are, and from (1.36) it can be shown that, for $m = 1, 3$

$$f_1(x;\mu_m) \underset{x\to-\infty}{\longrightarrow} 0, \qquad f_1(x;\mu_m) \underset{x\to+\infty}{\longrightarrow} \frac{-\sigma_1}{\mu_m}\,. \tag{1.62}$$

We find for the first term in the eigenvalue expansion

$$\sigma_1(\mu_m) = -\frac{2s}{\chi_m^2}(f_0, f_0)\,, \tag{1.63}$$

so that

$$\begin{aligned}
\sigma(\mu) &= \epsilon\,\sigma_1(\mu_m) + O(\epsilon^2) \\
&= \epsilon\,\frac{128}{27\pi^2} + O(\epsilon^2)\,, \quad \text{for } \mu = 1 - \epsilon \\
&= \epsilon\,\frac{128}{15\pi^2} + O(\epsilon^2)\,, \quad \text{for } \mu = 9 - \epsilon\,.
\end{aligned} \tag{1.64}$$

We notice that in (1.61) it is necessary to have $\sigma_1 > 0$ to meet the boundary condition at $+\infty$ and therefore by (1.63) we get eigenvalues of the form (1.64) for $\mu < \mu_m$, i.e., $s = -1$, only. Of course, this does not exclude the existence of other eigenvalues outside a neighbourhood of $\sigma = 0$ when $\mu > \mu_m$, but at this stage nothing can be said about the stability of the forced waves in this range.

We further notice that equation (1.30), for the forcing under consideration, i.e., type (i), for $\mu = 0$ admits the solution

$$f_0(x;0) = B\,\tanh(x)\left[5\,\tanh^2(x) - 3\right]\ , \tag{1.65}$$

which violates the regularity conditions (1.22), and therefore is not an eigenfunction as already noted, corresponding to zero eigenvalue. However, for μ in a neighbourhood of 0, it might be possible to use this solution as an "inner" one and correct the behaviour at infinity by means of an "outer" solution, which in this case should occur already at order $O(1)$. Such an analysis is not pursued here, but will be supplemented by numerical calculations given in Section 1.4.4.

If we now fix $\mu = 4$ and let α vary, for the case (ii) forcing, we have solutions $\sigma = 0$ for $\alpha = \nu(\nu+1)$, $\nu = 2, 3, \ldots,$. The first two are

$$\alpha_0 = 6, \quad f_0(x;\alpha_0) \;=\; B\,\mathrm{sech}^2(x)\,, \tag{1.66}$$

$$\alpha_1 = 12, \quad f_0(x;\alpha_1) \;=\; B\,\mathrm{sech}^2(x)\tanh(x), \tag{1.67}$$

The above procedure of determining the eigenvalues and eigenfunctions now with a perturbation of the parameter $\alpha = \alpha_\nu + s\epsilon$ can be applied here without change, the presence of the first order term $-s\frac{d}{dx}\left[\,\mathrm{sech}^2(x) f_0(x;\alpha_n)\right]$ rather than $s\frac{d}{dx} f_0(x;\mu_m)$ (see eq.(1.43)) merely changes the numerical expressions for σ_1 and σ_2 (and the role played by the signature s). We have

$$\sigma(\alpha_0 + s\epsilon) = s\epsilon\,\frac{16}{5} + O(\epsilon^2)\ , \tag{1.68}$$

$$\sigma(\alpha_1 + s\epsilon) = \epsilon^{\frac{1}{2}}\frac{16i}{\sqrt{105s}} + \epsilon\,\frac{16}{105s} + O(\epsilon^{\frac{3}{2}})\ , \tag{1.69}$$

and the corresponding asymptotic expansion for the eigenfunctions shows that for $\nu = 0$ (or more general for ν even) there is no solution for $s = -1$, i.e., $\alpha < \alpha_0$.

Since by (1.68), σ has its leading term real, the forced solitary wave (1.5) is unstable for $\alpha > \alpha_0$, but the linear mode analysis fails to provide information for $\alpha < \alpha_0$. However, a similar analysis for a neighbourhood of α_1 must take into account the slow time scale provided by the second term in (1.69), i.e., ϵt, and the linear theory shows that the stationary wave is unstable on this time scale, at least for $\alpha < \alpha_1$.

1.4.4 Global spectral behaviour

The relative simplicity of the operator $\mathcal{L}_{\alpha,\mu}$, (1.21), makes it possible to consider values of the parameters α, μ away from the special values listed above, and to find exact solutions to the eigenvalue problem of (1.19) for the whole range of the parameters. Although a closed form solution seems out of reach, we shall see that the problem can be cast in a form which allows some numerical computation with a high degree of accuracy. Keeping the notation compact, we consider type (i) forcing first. By setting $\alpha = 12$, and by transforming the independent variable through

$$z = \frac{1}{2}\left(1 - \tanh(x)\right) \ , \quad \hat{f}(z) \equiv f(x), \tag{1.70}$$

equation (1.19) becomes (dropping the 'hats' from now on)

$$\frac{d^3 f}{dz^3} + \frac{3(1-2z)}{z(1-z)}\frac{d^2 f}{dz^2} + \frac{1}{4}\left[\frac{10 - \mu - 6(1-2z)^2}{z^2(1-z)^2}\right]\frac{df}{dz} + \frac{12z(1-z)(1-2z) + \frac{\sigma}{8}}{z^3(1-z)^3} f = 0 \ . \tag{1.71}$$

This is a third order ordinary differential equation with three regular singularities at $z = 0,\ 1,\ +\infty$, and the boundary conditions (1.22) are now

$$f(0) = f(1) = 0 \ . \tag{1.72}$$

The indicial equation at $z = 0$ is

$$\kappa^3 - \frac{\mu}{4}\kappa + \frac{\sigma}{8} = 0 \ , \tag{1.73}$$

and in order to satisfy (1.72) we must retain only the roots whose real part is greater then zero. If no roots are coincident or differ by an integer number, we only need to consider the case when one of them, say κ_1, has real part greater then zero. In fact, due to the symmetry (1.23), the indicial equation at $z = 1$ is obtained from (1.70) by changing σ into $-\sigma$, so that $\kappa_i \to -\kappa_i$, where κ_i, $i = 1, 2, 3$ are the solutions of (1.73). A function f which satisfies (1.71) is therefore an eigenfunction for (1.19), if we can find coefficients of linear combination c_2, c_3 in such a way that it has the following behaviour in a neighbourhood of the singular points,

$$f(z) \sim \begin{cases} z^{\kappa_1} g_1(z) & z \sim 0 \\ c_2(1-z)^{-\kappa_2} h_2(1-z) + c_3(1-z)^{-\kappa_3} h_3(1-z) & z \sim 1 \end{cases} , \tag{1.74}$$

where g_1 and h_2, h_3 are analytic functions in a neighbourhood of, respectively, $z = 0$ and $z = 1$. Here $\{\kappa_i\}_{i=1,2,3}$ are three solutions of (1.73) with $\mathrm{Re}\kappa_1 > 0$, $\mathrm{Re}\kappa_3 < \mathrm{Re}\kappa_2 < 0$. The algebra is simplified if we remove the singular behaviour at $z = 0$, by setting

$$f(z) = \frac{z^{\kappa_1}}{(1-z)^{\kappa_1}} p(z), \tag{1.75}$$

so that equation (1.71) reduces to

$$\frac{d^3p}{dz^3} + \frac{Az+B}{z(1-z)}\frac{d^2p}{dz^2} + \frac{Cz^2+Dz+E}{z^2(1-z)^2}\frac{dp}{dz} + \frac{Fz+G}{z^2(1-z)^2}p = 0 , \tag{1.76}$$

where

$$\begin{aligned} A &= -6 , \quad B = 3(\kappa_1+1) , \\ C &= -6 , \quad D = 6(1-\kappa_1) , \quad E = 3\kappa_1(\kappa_1+1) + 1 - \frac{\mu}{4} , \\ F &= -24 , \quad G = 12(\kappa_1+1) . \end{aligned} \tag{1.77}$$

The power series

$$p(z) = \sum_{n=0}^{\infty} a_n z^n \tag{1.78}$$

is a solution of (1.76) if the coefficients $\{a_n\}$ satisfy the second order difference equation

$$Q(n)a_{n+1} + P(n)a_n + R(n)a_{n-1} = 0 \ , \tag{1.79}$$

where

$$\begin{aligned} Q(n) &\equiv (n+1)\left[(n-1)n + 3(\kappa_1+1)n + 3\kappa_1(\kappa_1+1) + 1 - \frac{\mu}{4}\right] \ , \\ P(n) &\equiv n\left[-2(n-1)(n-2) - 3(3+\kappa_1)(n-1) + 6(1-\kappa_1)\right] + 12(\kappa_1+1) \ , \\ R(n) &\equiv (n-1)(n-2)(n-3) + 6(n-1)(n-2) - 6(n-1) - 24 \ . \end{aligned} \tag{1.80}$$

Let us first check if certain values of the parameters μ, σ exist such that the series terminates. In order for this to be the case, we must have

$$a_{N-1} \neq 0, \quad a_N = a_{N+1} = 0 \tag{1.81}$$

for some integer $N > 0$. Substituting this condition into (1.79), we obtain

$$R(N)a_{N-1} = 0 \ , \tag{1.82}$$

and since $R(n)$ is independent of κ_1, this has only one solution

$$N = 4 \ . \tag{1.83}$$

Therefore $a_4 = 0$ and so

$$\begin{aligned} 0 &= \frac{P(3)}{Q(3)}a_3 + \frac{R(3)}{Q(3)}a_2 \\ &= \left[-a_2\frac{P(2)}{Q(2)} - a_1\frac{R(2)}{Q(2)}\right]\frac{P(3)}{Q(3)} + a_2\frac{R(3)}{Q(3)} \\ &= \left[-a_1\frac{P(1)}{Q(1)} - a_0\frac{R(1)}{Q(1)}\right]\left[\frac{R(3)}{Q(3)} - \frac{P(3)}{Q(3)}\frac{P(2)}{Q(2)}\right] - a_1\frac{R(2)}{Q(2)}\frac{P(3)}{Q(3)} \ . \end{aligned} \tag{1.84}$$

Since

$$a_1 = -\frac{P(0)}{Q(0)} a_0 \tag{1.85}$$

and $a_0 \neq 0$, the final equation for μ, σ becomes

$$\left[\frac{P(0)}{Q(0)}\frac{P(1)}{Q(1)} - \frac{R(1)}{Q(1)}\right]\left[\frac{R(3)}{Q(3)} - \frac{P(3)}{Q(3)}\frac{P(2)}{Q(2)}\right] + \frac{P(3)}{Q(3)}\frac{R(2)}{Q(2)}\frac{P(0)}{Q(0)} = 0 \, . \tag{1.86}$$

After some lengthy calculations, this reduces to

$$(4\kappa_1^2 - \mu)(\mu - 4) = 0 \, , \tag{1.87}$$

hence

$$\mu = 4 \, , \quad \text{any} \;\; \kappa_1 \; ; \quad \text{or} \;\; \kappa_1 = \frac{\sqrt{\mu}}{2} \, , \quad \text{any} \;\; \mu \, . \tag{1.88}$$

In the first case, solving for a_n, $n = 1,2,3$ in (1.79), we retrieve the solution found by Jeffrey and Kakutani (1972), i.e., a third order polynomial in z times the factor $z^{\kappa_1}(1-z)^{-\kappa_1}$, which becomes, in the original independent variable x,

$$f(x) = e^{2\kappa_i x}\left[\kappa_i(\kappa_i - 1)^2 + e^{-x}\left\{\kappa_i(\kappa_i - 1) - (2\kappa_i - 1)\tanh x + \tanh^2 x\right\}\right] \, , \tag{1.89}$$

for $i = 1,2,3$. The other case in (1.88) has already been considered in Section 1.4.1 for the perturbation expansions, since by (1.73) it implies $\sigma = 0$. With the exception of these two particular cases, the difference equation (1.79) does not seem to have a closed form solution in general, but it does show that the series (1.78) has radius of convergence equal to 1. Therefore the eigenvalue σ can be calculated to any desired order of accuracy, limited only by the round off error, by summing the two series in (1.74) numerically, one about $z = 0$ and the other about $z = 1$, and matching the two expansions at some intermediate point. By requiring continuity of the functions and the derivatives up to second order, which can of course be determined analytically by

differentiating the power series, the eigenvalues σ are therefore the zeros of

$$\det \begin{vmatrix} -p(z) & [(1-z)^{\kappa_1-\kappa_2} q_2(1-z)] & [(1-z)^{\kappa_1-\kappa_3} q_3(1-z)] \\ -p'(z) & [(1-z)^{\kappa_1-\kappa_2} q_2(1-z)]' & [(1-z)^{\kappa_1-\kappa_3} q_3(1-z)]' \\ -p''(z) & [(1-z)^{\kappa_1-\kappa_2} q_2(1-z)]'' & [(1-z)^{\kappa_1-\kappa_3} q_3(1-z)]'' \end{vmatrix} = 0 \ , \quad (1.90)$$

where $(\cdot)'$ denotes derivative with respect to z, and the determinant depends on σ through the indices κ_i. Here q_2 and q_3 are the power series obtained by h_2,h_3 in (1.74) via multiplication by the (analytic at $z = 1$) factor $z^{-\kappa_1}$. Their coefficients can be easily determined by the difference equation (1.79), using $-\kappa_1$ rather then κ_1 in the definitions (1.77) and setting $n \rightarrow n + \kappa_1 - \kappa_i$, $i = 2,3$. Choosing a point z sufficiently far removed from 0, 1, the power series have a fast convergence rate and the roots of equation (1.90) can be found by a Newton-Raphson scheme. This requires a good initial guess, and in order to gain some insight a simple Galerkin code was developed, using sines and Legendre polynomials as basis. The results pertaining to the eigenvalues so determined are shown in Figure 1.1, while Figure 1.2 depicts a typical eigenfunction for $1 < \mu < 4$. In the common domain of existence, the two expressions in (1.74) coincide up to six significant digits, when the eigenvalues are evaluated with a series truncated after 200 terms. Figure 1.1 shows that the branch originating at $\mu = 4$ has a small real part, as expected by perturbation analysis, which stays small throughout the range $-1 < \mu < 4$, including the subcritical range $(-1,0)$ but is always different from zero. Being at its maximum still two orders of magnitude smaller then the imaginary part, the real part can only be found by high precision numerical schemes, and in fact the Galerkin expansions we tried were unable to provide an estimate for it. For the branch between $\mu = 4$ and $\mu = 9$ however, the eigenvalue is found to be purely real as predicted by the perturbation analysis for μ near 4 and 9.

The calculations relative to the forcing (ii) are completely analogous and results for the eigenvalues are depicted in Figure 1.3.

The following tables provide a comparison between some of the eigenvalues found by the present procedure and the corresponding ones evaluated by the perturbation techniques of the preceding paragraph:

forcing (i) $\mu = 4 + s\epsilon$

s	ϵ	σ (by series)	σ (by perturbation theory)
-1	0.1	$0.10953i + 4.164 \cdot 10^{-3}$	$0.1088i + 4.\overline{44} \cdot 10^{-3}$
-1	0.01	$0.03444i + 4.41 \cdot 10^{-4}$	$0.034426i + 4.\overline{44} \cdot 10^{-4}$
$+1$	0.01	0.033967	0.033982

forcing (i) $\mu = 9 - \epsilon$

ϵ	σ_{series}	$\sigma_{\text{pert.}}$
0.1	0.01426	0.01441

forcing (ii) $\alpha = 12 + s\epsilon$

s	ϵ	σ_{series}	$\sigma_{\text{pert.}}$
$+1$	0.1	$0.083184i + 2.4861 \cdot 10^{-3}$	$0.0823i + 2.54 \cdot 10^{-3}$
$+1$	0.01	$0.026052i + 2.534 \cdot 10^{-4}$	$0.026024i + 2.54 \cdot 10^{-4}$

Numerical experiments for the full evolution equation (1.1) have completely established the behaviour predicted by the previous spectral analysis. The inherent error of the numerical scheme should in fact play the role of a perturbation of the exact solutions (1.4) and (1.5). For (i) type forcings, the stationary solution (1.4) for $4 < \mu < 9$ is indeed unstable and decays to a lower amplitude stationary wave as shown in Figure 1.10. The instability for $\mu < 4$ is driven by the real part of the eigenvalues according to (1.18), and the previous results show that for subcritical and

critical cases, $\mu \leq 0$, it is too small to be detected numerically, since one should then use maximum times (in dimensionless form) of order $O(10^5)$. However, the imaginary part should give rise to periodic oscillations in time, and this is indeed shown by the computed wave resistance coefficient with the oscillation period agreeing well with the one predicted by the imaginary part of the eigenvalue, as described in detail in Section 1.7. The special case $\mu = 4$, $\alpha = 12$ where no eigenvalue different from zero could be detected corresponds to the free solitary wave, and the linearized analysis reduces to the one of Jeffrey and Kakutani (1971). Therefore, in the special case of $\mu = 4$, $\alpha = 12$ for the free solitary waves, no definite conclusions can be drawn from linear theory and the stability problem has to be resolved by nonlinear analysis, as shown by Benjamin (1972) or by other means, including the inverse scattering formalism (Newell 1985, §3f).

Combining the information provided by the perturbation approach, the power series solution and some implication of the nonlinear stability analysis of the next Section, we are actually able to predict some structure for the spectrum of the operator $\mathcal{L}_{\alpha,\mu}$ as α and μ may both vary in general. In fact, the results obtained around the special points $\{\alpha_n\}$, $\{\mu_n\}$, corresponding to zero eigenvalues, depend only on the symmetries of the appropriate eigenfunctions, which in turn are the solutions of the Schrödinger equation (1.30), satisfying the regularity conditions imposed at infinity. The structure of the solutions to this problem, as well as for any symmetric "potential" replacing the $\mathrm{sech}^2(x)$-like with a minimum $-\alpha$ at $x = 0$, and decaying fast enough at $x = \infty$, is well known. Recalling that $-\mu$ is playing the role of an eigenvalue in (1.30), we know that the first eigenfunction, corresponding to the lowest "eigenvalue" $-\mu(\alpha)$, is always symmetric, the second always antisymmetric and so on in alternating

fashion. The total number of "eigenvalues" μ will depend on α and on the rate of decay of the potential at ∞. For exponential decay, as in $-\alpha\ \mathrm{sech}^2(x)$, the number of eigenvalues is finite, and these will constitute the special points corresponding to the $\{\mu_n\}$ on which the analysis of Section 1.4.1 is based. Thus, in general, we can say that for fixed α, a μ corresponding to an even eigenfunction, μ_e say, is the starting point of a branch of real eigenvalues of $\mathcal{L}_{\alpha,\mu}$ for $\mu < \mu_e$. Similarly, from a $\mu \equiv \mu_o$ corresponding to an odd eigenfunction, a branch of complex eigenvalue emanates for $\mu < \mu_o$, and a branch of real eigenvalues issues forth for $\mu > \mu_o$, the scaling of the real and imaginary parts with respect to the (small) distance $|\mu - \mu_o|$ being similar to the case of μ_2 considered above. Of course this only holds in a neighbourhood of the special points, and the power series approach is needed to follow the branches as μ varies further on. In the next Section we will show that when $\mu > \bar{\mu}_e$, $\bar{\mu}_e \equiv \max\{\mu_e\}$, or $\alpha < 0$, the positivity of the operator $K_\alpha + \mu$ implies that no eigenvalue of $\mathcal{L}_{\alpha,\mu}$ with real part different from zero can exist. This finding agrees with the perturbation analysis which shows that no eigenvalue is found to exist in a neighbourhood of this point when $\mu > \bar{\mu}_e$. Hence the curve $\mu = \bar{\mu}_e(\alpha)$ provides a limit in the μ, α plane for the existence of eigenvalues with real part different from zero. We illustrate the above results using a $\mathrm{sech}^2(x)$ potential. It is convenient to introduce the parameter $\nu = \nu(\alpha)$ by

$$\alpha = \nu(\nu + 1)\ ,$$

so that, using the same transformation as for (1.31), equation (1.30) has the form of a Legendre operator of order $\sqrt{\mu}$ and degree ν (Landau and Lifshitz, *ibid.*). It can then be shown that the "eigenvalues" μ are given by

$$\nu(\alpha) - \sqrt{\mu} \ = \ n \qquad n = 0, 1, 2 \cdots$$

their number being determined by the obvious requirement $\nu(\alpha) > n$. We note that for $\nu = n$, the solution of (1.30) is not an eigenfunction since it does not vanish at $x = \pm\infty$. The curve (a parabola in the (μ, ν) as well as the $(\mu, \alpha$ plane))

$$\nu(\alpha) = \sqrt{\mu} \tag{1.91}$$

for $\mu > 0$, and the line

$$\alpha = \nu(\alpha) = 0$$

for $\mu < 0$ thus divide the (μ, ν) (or (μ, α)) plane in two parts, the one contained in the upper half plane (above the line (4.75) for $\mu > 0$) being the region that allows existence of eigenvalues with nonzero real part. The other special points at which the eigenvalue branches described above are generated lie on the parabolae given by (1.91) with $n > 0$, see Figure 1.4.

By using the power series approach, it can be seen that, in general, the complex eigenvalue branch originating from the "odd" parabolae continue into the quarter plane $\mu < 0$, while the real branches sprouting from the"even" parabolae merge with the ones coming from the "odd" ones in the direction of increasing μ (or decreasing α). The overall features of the dependence of the spectrum on the parameters are then similar to the cases $\alpha = 12$, μ varying, or $\mu = 4$, α varying, considered above, the only difference being the position (and the number) of the points in the parametric (μ, α) space corresponding to eigenvalues equal to zero.

1.5 Nonlinear Stability

In the preceding Section, the linear instability for the forced solitary waves in the range $0 < \mu < 9$, $\alpha > 6$ was established for forcing (i) and (ii) respectively, but we

were unable to ascertain any eigenvalues when $\mu > 9$ or $\alpha < 6$, and so the question of stability of the forced solitary waves in this range of parameters is still to be resolved. A principal difficulty which is characteristic of the present problem lies in the rather unique feature of the linearized stability equation that an eigenvalue σ with a negative real part cannot be used to indicate a decay of small perturbations $f(x)$ because $-\sigma$ is then the eigenvalue for the perturbation $f(-x)$. However, a sufficient condition for linear stability, i.e., non-existence of eigenvalues with real part greater then zero, can actually be provided.

We first notice that the fKdV equation (1.1) can be interpreted, as well known for its homogeneous counterpart, as a Hamiltonian system. Introducing the Hamiltonian functional

$$\mathcal{H}(\zeta) = \frac{1}{2}\int_{-\infty}^{+\infty} \left[\zeta_x^2 + \mu\zeta^2 - 3\zeta^3 - 6P\zeta\right] dx, \tag{1.92}$$

and the Poisson bracket

$$\{\mathcal{H}, \mathcal{G}\} \equiv \int_{-\infty}^{+\infty} \frac{\delta\mathcal{H}}{\delta\zeta}\frac{\partial}{\partial x}\frac{\delta\mathcal{G}}{\delta\zeta} dx \ , \tag{1.93}$$

where $\mathcal{G}$ is any functional of ζ and $\frac{\delta}{\delta\zeta}$ denotes the functional derivative, we find that

$$\zeta_t = -\frac{\partial}{\partial x}\frac{\delta\mathcal{H}}{\delta\zeta} \ , \tag{1.94}$$

as can be readily verified by direct computation. The Hamiltonian (1.92) is one of the two "obvious " conserved quantities for the fKdV equation with the forcings under consideration, the other being the total excess mass (the equation itself being a conservation law),

$$\mathcal{M} = \int_{-\infty}^{+\infty} \zeta \ dx \ . \tag{1.95}$$

(These are obvious in the sense that they are the ones obtained by Noether's first theorem, due to invariance of the Lagrangian for equation (1.1) with respect to time

translations and addition of an arbitrary constant to the potential for the dependent variable, see appendix). The presence of a non-constant forcing term has the effect of destroying invariance with respect to spatial translations, and so the excess energy integral

$$\mathcal{E} = \frac{1}{2}\int_{-\infty}^{+\infty} \zeta^2 \, dx \ , \tag{1.96}$$

is in general no longer a constant of motion.

Equation (1.94) shows that the stationary solutions $\zeta_s(x)$ of (1.1) can now be seen as extrema of the Hamiltonian (1.92), i.e., solutions of

$$\frac{\delta \mathcal{H}}{\delta \zeta} = 0 \ . \tag{1.97}$$

The second variation of the Hamiltonian at ζ_s is determined as

$$\delta^2 \mathcal{H} = \frac{1}{2}\int_{-\infty}^{+\infty} \left[\eta_x^2 - 9\zeta_s(x)\eta^2 + \mu\eta^2\right] dx \ , \tag{1.98}$$

where η is, same as in (1.8), $\eta(x,t) = \zeta(x,t) - \zeta_s(x)$, and is assumed to be a square integrable function together with its first derivative, i.e., it is an element of the Sobolev space H^1. Taking the time t fixed for the moment, this equation can be rewritten, with respect to the inner product (1.25) in L^2, the usual Hilbert space of square integrable functions, as

$$\delta^2 \mathcal{H} = (\eta, K_\alpha \eta) + \mu \, (\eta, \eta) \ , \tag{1.99}$$

where K_α is the Schrödinger operator introduced in (1.30), i.e.,

$$K_\alpha \equiv -\frac{d^2}{dx^2} - \alpha \ \mathrm{sech}^2(x) \ . \tag{1.100}$$

From elementary quantum mechanics, it is well known that

$$(\eta, K_\alpha \eta) \geq \lambda_0 \, (\eta, \eta) \, , \tag{1.101}$$

where $\lambda_0 = \lambda_0(\alpha) < 0$ is the eigenvalue of K_α corresponding to the ground state, if the $\mathrm{sech}^2(x)$-potential in it is able to support one. In our case $\lambda_0(12) = -9$ for forcing (i), while the minimum value of α for having a bound state is $\alpha = 6$ for forcing (ii), corresponding to $\lambda_0(6) = -4$. Thus

$$\begin{aligned} \delta^2\mathcal{H} &\geq [\mu + \lambda_0(\alpha)]\,\|\eta\|^2 , \\ &\geq (\mu - 9)\|\eta\|^2 \qquad \text{for forcing (i)} , \qquad (1.102) \\ &\geq [4 + \lambda_0(\alpha)]\,\|\eta\|^2 \qquad \text{for forcing (ii)} , \qquad (1.103) \end{aligned}$$

and for $\mu > 9$ or $\alpha < 6$, for forcing (i) or (ii) respectively, the second variation $\delta^2\mathcal{H}$ is strictly positive, a result commonly referred to as formal stability, which is a sufficient condition for excluding linear instability (Holm et al., 1985). Formal stability is a step towards establishing nonlinear stability and, owing critically to the presence of the parameters μ and α, it is indeed possible to show that these forced solitary waves are (nonlinearly) stable in the range of parameters under consideration. Nonlinear stability is here intended in the usual Lyapunov sense, i.e., it is possible to find a norm $d(\cdot,\cdot)$ in the appropriate functional space on which the evolution equation (1.1) is defined, so that for any $\epsilon > 0$ one can determine a $\delta > 0$ such that if $d(\zeta,\zeta_s) < \delta$ at $t = 0$, then $d(\zeta,\zeta_s) < \epsilon$ at any time $t > 0$. Although a second nontrivial conserved quantity appears to be missing for the forced case, the original argument of Benjamin (*ibid.*) can be adapted to this case, and convexity estimates can be provided for the Hamiltonian functional itself.

The fact that the excess energy of (1.1), $\mathcal{E}$, is no longer conserved actually simplifies the analysis, since this is a consequence of the absence of translation invariance, which eliminates the need to consider quotient spaces with respect to the translations.

Specifically, the total variation $\Delta\mathcal{H}$ of the functional $\mathcal{H}$ at ζ_s is

$$\Delta\mathcal{H} = \frac{1}{2}\int_{-\infty}^{+\infty}\left[\eta_x^2 - 9\zeta_s(x)\eta^2 + \mu\eta^2\right]dx - \frac{3}{2}\int_{-\infty}^{+\infty}\eta^3\,dx\,, \tag{1.104}$$

and the last term can be estimated by using the Sobolev inequality

$$\frac{3}{2}\int_{-\infty}^{+\infty}\eta^3\,dx \;\leq\; \frac{3}{2}|\eta|_\infty\int_{-\infty}^{+\infty}\eta^2\,dx\,, \tag{1.105}$$

$$\leq\; \frac{3}{2\sqrt{2}}\|\eta\|_1\,\|\eta\|^2\,, \tag{1.106}$$

where $|\cdot|_\infty$ is the supremum norm, $\|\cdot\|_1$ is the Sobolev norm $\|\eta\|_1^2 \equiv \|\eta\|^2 + \|\eta_x\|^2$ and η is assumed to be in the corresponding space at times $t > 0$. Using (1.106) it then follows that

$$\Delta\mathcal{H} \geq \frac{1}{2}\left[\mu + \lambda_0(\alpha)\right]\|\eta\|^2 - \frac{3}{2\sqrt{2}}\|\eta\|_1\|\eta\|^2\,. \tag{1.107}$$

If we take the norm d to be $\|\cdot\|_1$, and impose

$$\|\eta(\cdot,0)\|_1 \leq \delta\,, \tag{1.108}$$

at time $t = 0$, it is easy to provide an upper bound on $\Delta\mathcal{H}$, again using Sobolev inequality (1.106),

$$\begin{aligned}\Delta\mathcal{H} \;&\leq\; M\|\eta\|_1^2 + \frac{3}{2\sqrt{2}}\|\eta\|_1\|\eta\|^2\,,\\ &\leq\; M\|\eta\|_1^2 + \frac{3}{2\sqrt{2}}\|\eta\|_1^3\,,\\ &\leq\; \gamma(\delta)\quad\text{say}\,,\end{aligned} \tag{1.109}$$

where $M = \frac{1}{2}\max\left(1, |\mu + \lambda_0(\alpha)|\right)$. Because of the invariance of $\mathcal{H}$ with time, this bound holds for all time $t > 0$ for which $\eta(x,t)$ exists in H^1.

Following Bona (1975), introducing for sake of tidiness the notation

$$A(t) \equiv \|\eta\| \ , \quad B(t) \equiv \|\eta_x\| \ , \tag{1.110}$$

we can further provide a bound on $B(t)$ for any time $t \geq 0$ from the previous inequality, isolating the η_x term in it,

$$B^2 \leq 2\gamma + \frac{3}{\sqrt{2}}(A+B)A^2 + |\mu + \lambda_0(\alpha)|A^2 \ . \tag{1.111}$$

Once this inequality is solved for B, we can estimate B in terms of A,

$$B \leq F(A) \ , \tag{1.112}$$

where

$$F(A) = \frac{1}{2}\left\{3A^2 + \left[9A^4 + 8\gamma + 4\left(3A^3 + |\mu + \lambda_0(\alpha)|A^2\right)\right]^{\frac{1}{2}}\right\} \ . \tag{1.113}$$

Since (1.107) implies

$$\gamma \geq \Delta\mathcal{H} \geq \frac{1}{2}A^2\left[(\mu + \lambda_0(\alpha)) - \frac{3}{\sqrt{2}}(A+B)\right] \ , \tag{1.114}$$

using (1.112) we find

$$\gamma \geq \frac{1}{2}A^2\left[(\mu + \lambda_0(\alpha)) - \frac{3}{\sqrt{2}}F(A)\right] - \frac{3}{2\sqrt{2}}A^3 \ . \tag{1.115}$$

Now, since $F(0) = \sqrt{2\gamma}$, if

$$\frac{1}{3}\left[\mu + \lambda_0(\alpha)\right] > \sqrt{\gamma} \ , \tag{1.116}$$

there is certainly some positive interval $I_\gamma \equiv [0, A_\gamma]$ such that, if A belongs to I_γ, the right hand side of (1.115) is positive, and monotonically increasing with A, as can be realized by studying it as a function of A. Therefore, with γ (and hence δ) held fixed sufficiently small and taking the corresponding A, A_γ say, we have a bound on A in

terms of γ, $A \leq A_\gamma$ and in turn, using (1.112), we also have a bound on B in terms of γ only, $B \leq B_\gamma$, for any $t \geq 0$. Hence, by choosing δ such that both (1.116) and

$$A^2 + B^2 \leq A^2_{\gamma(\delta)} + B^2_{\gamma(\delta)} \leq \epsilon^2 \tag{1.117}$$

hold, the Lyapunov condition is established with $d = \| \ \|_1$. We again see that the condition expressed by (1.116) is crucial, and this of course is attained for $\mu > 9$ and $\alpha < 6$ respectively for forcing (i) and (ii). Underlying these considerations is the assumption that for sufficiently smooth initial data $\zeta(x,0)$ (and forcing P in general) the solution $\zeta(x,t)$ exists, continuous in time, and belongs to at least H^1 for all time $t > 0$ (see Bona and Smith, 1975).

It is interesting to notice that an almost identical argument can be provided for the nonlinear stability of the stationary solutions to the regularized fKdV equation, which is the one to be used in most of the numerical simulations of Section 1.7. In the coordinate frame and notation of (1.1), this equation can be written as

$$\zeta_t + 6(F-1)\zeta_x - 9\zeta\zeta_x - \zeta_{xxt} - F\zeta_{xxx} = 6P(x) \ . \tag{1.118}$$

The analogue of (1.4) and (1.5), i.e., the stationary solutions $\zeta_{s_r}(kx)$, in this case are

(i_r)

$$\begin{aligned} \zeta_{s_r}(x) &= \frac{4}{3}k^2 F \ \mathrm{sech}^2(kx) \\ P_r(x) &= \frac{4}{3}k^2\left(F - 1 - \frac{2}{3}k^2F\right) \ \mathrm{sech}^2(kx), \end{aligned} \tag{1.119}$$

(ii_r)

$$\begin{aligned} \zeta_{s_r}(x) &= a \ \mathrm{sech}^2(kx) \\ P_r(x) &= a\left(k^2F - \frac{3}{4}a\right) \ \mathrm{sech}^2(kx), \quad F - 1 \equiv \frac{2}{3}k^2 \ , \end{aligned} \tag{1.120}$$

and the Hamiltonian for the this regularized fKdV equation only requires a change in the coefficient of the derivative term in (1.92) (see appendix),

$$\mathcal{H}_r \equiv \frac{1}{2}\int_{-\infty}^{+\infty} dx \left[F\zeta_x^2 + 6(F-1)\zeta^2 - 9\zeta^3 - 6P(x)\zeta\right] . \tag{1.121}$$

Thus, introducing the notation

$$x' = kx, \quad \mu_r \equiv 6\frac{(F-1)}{k^2F}, \quad \alpha_r \equiv \frac{9a}{k^2F} , \tag{1.122}$$

in analogy with (1.12) and (1.13) we can write the total variation of the Hamiltonian based on $\zeta_{s_r}(x)$ as

$$\frac{1}{kF}\Delta\mathcal{H}_r = \frac{1}{2}\int_{-\infty}^{+\infty}\left[\eta_x^2 - \alpha_r\,\mathrm{sech}^2(x')\eta^2 + \mu\eta^2\right]dx' - \frac{3}{2k^2F}\int_{-\infty}^{+\infty}\eta^3\,dx' , \tag{1.123}$$

and from here on the previous arguments apply virtually unchanged, with the exception of the coefficient in front of the cubic term in (1.104).

1.6 Existence of multiple stationary solutions

The spectrum of the operator K_α in (1.30) is once again useful since it determines bifurcation points of the solutions (1.4) and (1.5) to other stationary solutions. Specifically, consider the time independent counterpart of equation (1.11) in a neighbourhood of the parameter values $\mu = \mu_m$ (or $\alpha = \alpha_n$), which can be written, after integrating in x and using the regularity condition at infinity for η, as

$$-\eta_{xx} - \alpha\eta\,\mathrm{sech}^2(x) + \mu\eta - \frac{9}{2}\eta^2 = 0 . \tag{1.124}$$

We note that (1.124) has a solution of the form $\eta_s = b\,\mathrm{sech}^2(x)$ if $\mu = 4$ and $b = 2(6-\alpha)/9$, so that for forcing (i), $\mu = 4$ ($\alpha = 12$), we have $b = -4/3$, making the resultant motion trivial (i.e., $\zeta = \zeta_s + \eta_s = 0$) with $P = 0$, and for forcing (ii), $b = 0$,

$P = 0$, giving a free soliton for the resultant ζ. We proceed below to determine nontrivial solutions of (1.124) for other values of μ and α.

For definiteness, let us first restrict ourselves to forcing (i), with $\alpha = 12$ so that dropping the subscript of the operator K_α does not cause ambiguity, and define μ as in (1.35), i.e.,

$$\mu = \mu_m + s\epsilon, \quad 0 < \epsilon \ll 1, \quad s = \pm 1, \quad m = 1, 2, 3 \; . \tag{1.125}$$

We look for solutions of (1.124) by regular perturbation expansion,

$$\eta(x;\mu) = \psi_1(\epsilon)\eta_1(x) + \psi_2(\epsilon)\eta_2(x) + \psi_3(\epsilon)\eta_3(x) + \ldots, \tag{1.126}$$

where the $\{\psi_i\}$ are taken as in (1.39) and (1.40), i.e., $\psi_i(\epsilon) = \epsilon^i$ for $\mu = 1,\ 9$ and $\psi_i(\epsilon) = \epsilon^{\frac{i}{2}}$, for $\mu = 4$. Substituting these expressions in (1.124) we have for $m = 1, 3$

$$(K + \mu_m)\eta_1 = 0, \quad (K + \mu_m)\eta_n = s\eta_{n-1} + \frac{9}{2}\sum_{k=1}^{n-1}\eta_{n-k}\eta_k, \quad (n > 1) \; , \tag{1.127}$$

and for $m = 2$

$$(K + 4)\eta_1 = 0, \quad (K + 4)\eta_2 = \frac{9}{2}\eta_1^2, \tag{1.128}$$

$$(K + 4)\eta_n = s\eta_{n-2} + \frac{9}{2}\sum_{k=1}^{n-1}\eta_{n-k}\eta_k, \quad (n > 2) \; , \tag{1.129}$$

where K is the Schrödinger operator defined in (1.30). Thus, for the "even" case $m = 2$, $\mu_m = 4$, we find for the first order (in the notation of (1.32-1.34))

$$\eta_1(x) = c_1 f_0(x;4) = 2c_1 \tanh(x)\ \mathrm{sech}^2(x) \; . \tag{1.130}$$

In analogy with (1.42), the inhomogeneous equation in (1.129) is solvable, due to the symmetries of the functions involved, and the particular solution for η_2 can be determined as

$$\eta_2(x) = -\frac{3}{4}c_1^2\ \mathrm{sech}^2(x)\left[3\ \mathrm{sech}^2(x) - 2\ \tanh(x) - 2\right] \; . \tag{1.131}$$

The constant c_1 is determined by the solvability condition at the next order, $n = 3$,

$$\frac{s}{9}\|f_0\|^2 - (f_0^2, \eta_2) = 0$$

to be

$$c_1 = \pm \frac{\sqrt{-11s}}{3} , \tag{1.132}$$

the square root arising from the proportionality $\eta_2 \propto c_1^2$. Therefore, in order to have η_2 real we must take $s = -1$, i.e., the bifurcating solutions only exists for $\mu < 4$, and these small amplitude deviations from the exact solution ζ_s (1.4) for $\mu < 4$ have the form, to the first order,

$$\zeta(x) = \frac{4}{3}\,\mathrm{sech}^2(x) \pm \frac{2(11\epsilon)^{\frac{1}{2}}}{3}\,\mathrm{sech}^2(x)\tanh(x) + O(\epsilon) \quad \text{at} \quad \mu = 4 - \epsilon\,. \tag{1.133}$$

By similar calculations, the solutions bifurcating at the other two eigenvalues (corresponding to $\mu = 1$ and $\mu = 9$) of the operator K are

$$\zeta(x) = \frac{4}{3}\,\mathrm{sech}^2(x) + s\epsilon c_1(\mu_m) f_0(x;\mu_m) + O\left(\epsilon^2\right) , \quad m = 1, 3 \tag{1.134}$$

where

$$c_1(\mu_m) = \frac{2}{9}\,\frac{\int f_0^2(x;\mu_m)\,dx}{\int f_0^3(x;\mu_m)\,dx} , \tag{1.135}$$

with the integrals ranging from $-\infty$ to $+\infty$. Thus, the amplitude constant $c_1(\mu_3)$ is

$$c_1(\mu_3) = \frac{2^{12}}{3^3 \cdot 5^2 \cdot 7 \cdot \pi} \simeq 0.276\,. \tag{1.136}$$

Furthermore, the forcing function P in (1.4) becomes of order ϵ for $\mu = 4 + s\epsilon$, as the time independent equivalent of equation (1.1) shows upon using the similarity transformation (1.10),

$$-\zeta_{xx} + 4\zeta - \frac{9}{2}\zeta^2 + s\epsilon\left[\zeta - \frac{4}{3}\,\mathrm{sech}^2(x)\right] = 0\,, \tag{1.137}$$

and so we can look for a perturbation expansion of the solution about $\zeta_s = 0$, $\mu = 4$, that is, there exists a branch of solutions $\zeta(x;\epsilon)$ bifurcating from the zero solution at $\mu = 4$

$$\zeta(x;\epsilon) = \epsilon\zeta_1(x) + \epsilon^2\zeta_2(x) + \dots \ . \tag{1.138}$$

We find, for the first order

$$\zeta_{1_{xx}} - 4\zeta_1 + \frac{4s}{3}\,\mathrm{sech}^2(x) = 0, \tag{1.139}$$

and so

$$\zeta(x;\epsilon) \ = \ \frac{s\epsilon}{3}\int_{-\infty}^{+\infty} d\xi e^{-2|x-\xi|}\,\mathrm{sech}^2(\xi) \ + O(\epsilon^2) \ . \tag{1.140}$$

By numerically integrating (1.137) regarding ζ as even, i.e., $\zeta_x(0) = 0$, it is possible to see that this new branch of solutions, from now on denoted by ζ_{ss}, joins the one originated at $\mu = 9$ with wave amplitude $a = \frac{4}{3}$ (see (1.134)), as represented by figure 1.8, where the amplitudes pertaining to $\eta_{ss} = \zeta_{ss} - \zeta_s$ are plotted versus the parameter μ.

The stability character of the small amplitude solutions found by the above regular perturbation methods, as well as the full bifurcation curve away from the special parameter values $\{\mu_m\}$ is an issue that needs to be addressed. According to the nonlinear stability analysis of Section 1.5., the first question can be answered once the ground state eigenvalue of the operator K with "potential" $\zeta_s(x)$ given by one of these stationary solutions has been evaluated. Once again, in a neighbourhood of μ_m, this is possible by using perturbation techniques. For instance, for the case $m = 3$ we have

$$K_\epsilon \equiv -\frac{d^2}{dx^2} - 12\,\mathrm{sech}^2(x) - s\epsilon\left[9c_1(\mu_3)\,\mathrm{sech}^3(x)\right] \ , \tag{1.141}$$

and by an elementary perturbation calculation the corresponding ground state eigenvalue is

$$\lambda_0(12;\epsilon) = \lambda_0(12) - s\epsilon \left[9c_1 \frac{\int f_0^3(x;\mu_m)\, dx}{\int f_0^2(x;\mu_m)\, dx}\right] + O(\epsilon^2) = -9 - 2s\epsilon + O(\epsilon^2)\ . \quad (1.142)$$

Thus, for $s = -1$, i.e., $\mu = 9 - \epsilon$, we have

$$\mu + \lambda_0(12;\epsilon) = \epsilon > 0\ , \quad (1.143)$$

and so the positive lower bound estimate for the corresponding $\Delta\mathcal{H}$ can be established. Under the forcing (i), the bifurcation of stationary solutions to (1.1) taking place at $\mu = 9$ is therefore of the transcritical type.

1.7 Numerical simulations

The salient features of the forced solitary wave solutions delineated by the preceding discussion can constitute an ideal trial ground for any numerical code devised for solving the fKdV (1.1). In fact, if an initial condition for $\zeta(x,t)$ corresponding to one of the two types of forcing (i) or (ii) is chosen, an accurate code must be able to show evidence of such stability features as determined analytically in the foregoing. Since the numerical results are by nature approximations, with errors usually estimable, we would expect that for the cases where the exact solution is found to be unstable, there should arise, in due time, in the corresponding numerical results an automatic onset of the instability, which can be attributed to the ever present numerical errors, with no need to superimpose a perturbation to the exact solution. By decreasing the truncation error of the numerical code, assuming for instance a finer grid, the manifestation of any instability should accordingly be delayed in time. Conversely, by

perturbing an initial condition which leads to a stable stationary wave, the numerical results should show evidence of the tendency of the system to recover the stationary state which is stable, provided the perturbation is not too large.

We shall focus our numerical simulations for forcing (i) on the three intervals of the parameter μ pertaining to the velocity of forcing (see (1.12)), as identified by the previous analysis, namely μ around zero (critical), $4 < \mu < 9$ and $\mu \geq 9$ (supercritical). We further notice that at $\mu = 0$ the amplitude by the forcing given by (1.4) is negative. For forcing (ii), since it provides an analytic expression for solitary waves of negative amplitude, we shall concentrate our numerical study on values of $\alpha < 0$. One of the codes is based on the one developed by Wu & Wu (1982) and uses the modified Euler's predictor-corrector algorithm in advancing time, with the space derivatives approximated by central differences. The forward difference computation for $\zeta(x,t)$ is implicit, in order to achieve the desired numerical stability and accuracy with a relatively large time step (up to 0.2 in dimensionless form), and no iteration is required on the corrector stage. Furthermore, open boundary conditions have been devised at the ends of the computation domain, requiring the waves adjacent to each boundary to leave the computation domain at the rate of the linear wave velocity $c_0 = \sqrt{gh}$. In order to avoid propagation of short wavelength disturbances that can be generated by numerical errors, equation (1.118), the so called "regularized" fKdV, has been adopted rather then the fKdV equation itself, the two models being equivalent in the limit of applicability, i.e., for long waves (small k, see (1.6)), and small amplitude. The differences between the results obtained from the two models is expected to be $O(k^2)$ (Benjamin et al. 1972, Whitham 1974, Wu 1987) and, as already mentioned, the previous stability analysis applies to this equation as well, with

only minor modifications. We notice that according to the similarity transformation (1.10), the choice of taking k small has to be counterbalanced by the drawback that a smaller k implies a larger number of time steps. In all the cases in which this code was used, the parameter k is kept fixed at a compromised value equal to 0.3. Throughout the simulations the wave resistance coefficient $C_R(t)$, defined as

$$C_R(t) = -\int_{-\infty}^{+\infty} P(x)\, \zeta_x(x,t)\, dx \tag{1.144}$$

has been evaluated; this quantity has been found very useful to better emphasize some major features such as small amplitude oscillations, the time period of variations, etc., which otherwise would become obscure to detect. Physically, its significance is that of a nondimensional measure of the power being supplied by the forcing, as shown by the energy balance equation,

$$\frac{d}{dt}\int_{-\infty}^{+\infty} \zeta^2(x,t)\, dx = -\int_{-\infty}^{+\infty} P(x)\, \zeta_x(x,t)\, dx\ , \tag{1.145}$$

which is obtained by multiplying (1.1) by ζ and integrating in x. In order to fully appreciate the previous remarks, we also have developed a simple finite difference code to numerically solve the fKdV equation. It is based on the explicit leap-frog scheme introduced by Zabuski and Kruskal (1964), with three point average on the nonlinear term to comply with the energy balance equation (1.145). Being explicit, the code is much less efficient then its counterpart for the regularized fKdV model, and for numerical stability the typical time step Δt has to be $\frac{1}{50}$ the spatial step Δx, for $k = 0.5$, as compared to the typical case of $\Delta x = 0.1$, $\Delta t = \frac{1}{2}\Delta x$ for the first code. Therefore, the usage of the latter code has been limited to the case when spot comparisons with the results of the regularized fKdV equation were desired. All the plots shown here are referred to the coordinate frame in which the forcing is viewed

as stationary, i.e., the water is entering the computational domain from the left side, with nondimensional velocity F.

The results pertaining to the critical case $\mu \approx 0$ for forcing (i) can be summarized as follows:

a) When the initial condition of the rest state $\zeta = 0$ is chosen, the phenomenon of periodic generation of the so called "runaway" solitons as reported by Wu & Wu (1982) is invariably observed. This can be seen in Figure 1.5 together with the wave resistance coefficient $C_R(t)$. The nondimensional period $T_s' = T_s k^3$ of the wave resistance coefficient, which in this case coincides with the period of generation of runaway solitons, is $T_s' \simeq 7.9$, in agreement with the value determined by Wu (1987) numerically as well as by applying a mass-energy theorem.

b) If the initial condition is chosen to be exactly $\zeta_s(x)$, the resulting wave is close to the stationary solution within a 2% error when the (nondimensional) time reaches $t = 800$, as illustrated by Figure 1.6a. This is in agreement with the fact that at $\mu = 0$ the real part of the eigenvalue is 6.84×10^{-4} and so, with the time scaled according to (1.10), it would require a time of order $O(10^5)$ at $k = 0.3$ to show an appreciable effect of $O(1)$. However, the wave resistance coefficient in Figure 1.6b, as calculated by the first code, exhibits a regular oscillation of period

$$T' = k^3 T_{num.} \simeq 5.75, \tag{1.146}$$

though with a very small amplitude ($\simeq 5 \times 10^{-5}$), and, interestingly, this is in good agreement with the period predicted by the imaginary part of the eigenvalue ($\omega_{e.v.} \equiv$ Im $\frac{\sigma}{6} = 1.144$ at $\mu = 0$, see Figure 1.1b)

$$T'_{e.v.} = \frac{2\pi}{\omega_{e.v.}} = 5.494 \tag{1.147}$$

This result can be further improved by using a smaller k, or by using the original fKdV model, as Figure 1.6c indeed shows, along with the persistence of the short wavelength oscillation mentioned above. Here for numerical convenience, the equation solved is in the form

$$\bar{\zeta}_{\bar{t}} + (F-1)\bar{\zeta}_{\bar{x}} - \frac{3}{2}\bar{\zeta}\bar{\zeta}_{\bar{x}} - \frac{k^2}{6}\bar{\zeta}_{\bar{x}\bar{x}\bar{x}} = \bar{P}_{\bar{x}} \ , \tag{1.148}$$

where

$$\bar{x} = kx \ , \quad \bar{t} = kt \ , \quad \bar{\zeta} = \zeta \ , \quad \bar{P} = P \ ,$$

and $k = 0.5$ for the result of Figure 1.6c. Therefore, the period given by the similarity relation (1.10) is

$$T' = k^3 \, \frac{\bar{T}_{num.}}{k} = 0.25 \, \bar{T}_{num} \simeq 5.54 \ .$$

For μ in the range $4 < \mu < 9$, the initial state of rest is found to gradually evolve under the forcing (i) into a stationary wave as shown in Figure 1.7, where $\mu = 6.\bar{6}$ corresponding to $F = 1.1$ at $k = 0.3$. This wave is not the forced solitary wave ζ_s of (1.4) which, as we have seen, is indeed unstable in this range. Instead, the terminal value of amplitude $\simeq 0.66\alpha$ is seen to be an element of the family of solutions which bifurcates from $\mu = 9$, as it can be shown by following the bifurcation curve (Figure(1.8)) and comparing the amplitudes. The instability associated with the real eigenvalue $\sigma = 0.23$ at $\mu = 6.\bar{6}$ manifests itself around the nondimensional time $t = 600$, a little insufficient to have the resulting wave fully developed, over a time interval of about 800 that we have kept as reasonable size for the computation. To pursue this point, rather than extending the time interval, we perturb the initial condition, by increasing the initial wave amplitude by 10%, and the result is shown in Figure 1.9. We also have doubled the size of the grid with similar results. There are two remarkable features in the result: a) the perturbed stationary wave first exhibits

instability through emitting a single upstream advancing solitary wave, and b) the remaining wave in the vicinity of the origin decays to the previously identified new stationary solution by emitting some "radiation," i.e., dispersive waves moving with non zero (subcritical) speed to the right. As a final proof of the point, we found that a decrease rather than an increase in the initial wave amplitude introduced as a weaker perturbation, makes the local solitary wave decay to the new stable state by emitting only slight radiation, as can be seen in Figure 1.10. In the soliton emission case, the wave resistance coefficient exhibits in Figure 1.9b a positive hump in the course of time when the solitary wave is being generated, signifying that the system is absorbing energy from the forcing. Conversely, the case of gradual decay to the stable state with no soliton emission correspond to a release of energy from the system, and $C_R(t)$ remains negative. The fact that C_R seems to settle to a negative, albeit small, asymptotic negative value is seen to be of numerical origin, since it is halved as the spatial grid is halved, and is partly due to the fact that the finite difference scheme for the regularized fKdV model is not exactly complying with the discrete counterpart of the energy balance equation (1.145), or its equivalent for the regularized KdV (see appendix).

For the supercritical case of $\mu > 9$, our numerical results show evidence of their strong stability through the fact that even when perturbed from the initial rest state of ζ, with $\eta(x,0) = -\zeta_s(x)$, the forced steady solitary wave shape is invariably regained after emitting some radiation downstream. Results pertaining to the case of $\mu = 18$ are graphically represented in Figure 1.11 and 1.12, where in order to get a larger computation time interval the grid size of $\Delta x = 0.2$, $\Delta t = 0.1$ was taken.

As mentioned above, we also have examined the case of forcing (ii), for a negative

$\zeta_s(x)$, with $\alpha = -10$. According to the previous stability analysis, when interpreted as a potential for the operator K_α, such a ζ_s cannot support any bound state, the spectrum being only continuous and positive. Therefore, as a forced solitary wave this ζ_s should be stable and this can indeed be confirmed by the simulations. Figure 1.13 presents the result for a positive perturbation of 10% of the steady wave amplitude, and shows that the solitary wave solution is regained in finite time with emission of some radiation. The numerical results seem to indicate that the "basin of attraction" of this solution with respect to perturbations having changes only in the amplitude extends to a point between 0.3 and 0.5 of the unperturbed amplitude, as illustrated by Figure 1.14 and 1.15. Beyond this level of variation of the amplitude, the solution evolves like the one which is strongly perturbed from the rest state, so that the regime of periodic generation of solitons again seems to occur (Figure 1.16). The evolution of the wave resistance coefficient as shown by Figure 1.14 and 1.15 clearly indicates this trend. Naturally, the system cannot possess a true attractor since it is Hamiltonian, but the dispersive wave trains, carrying energy away from a neighbourhood of the forced solitary wave and distributing over ever increasing spatial intervals towards $+\infty$, play here the role of dissipation.

1.8 Conclusions

We now summarize our results on the stability of stationary solutions of the fKdV equation. We have considered two basic types of forcing in the class given by Wu (1987), which have the common feature of supporting symmetric forced solitary waves similar in analytic expression to a free solitary wave. The linearized stability analysis about these waves allow us to identify three basic transcritical regimes. The first

occurs at "high" speed ($\mu > 9$), and "low" amplitude ($\alpha < 6$) for forcing (i) and (ii), respectively, the meaning of high and low being viewed in the sense of (1.2) and (1.10). For this regime, the solitary wave is shown to be stable in Lyapunov sense for the full nonlinear problem. For lower speeds, $4 < \mu < 9$ for forcing (i) or higher amplitudes, $6 < \alpha < 12$ for forcing (ii), there exists a second regime, in which the eigenvalues, determined by numerical methods of high accuracy, are purely real and therefore the corresponding stationary state becomes unstable. Accordingly, the evolution of the system from these states does not involve any time periodic phenomenon, as indeed the numerical simulations show. For still lower speeds and higher amplitudes, $\mu < 4$ and $\alpha > 12$ for forcing (i) and (ii) respectively, a third regime is identified, as the eigenvalues pass through zero and become complex, however with the real part being many orders smaller than the imaginary part. In this regime the initial growth of the instability is oscillatory, and the numerical simulations carried out with the full fKdV equation show evidence that this period is moderately amplitude dependent. These salient features actually conform with the contention made by Chen & Chang (1987). The simulations also provide some indication as to where the system may evolve in departing from the unstable stationary states. In the second regime, there exists for both forcings a second stable stationary wave of lower amplitude (but still positive), to which the system relaxes, and this is accomplished in a manner which strongly depends on the initial conditions. For forcing (i) in particular, this second solution, $\zeta_{ss}(x)$, becomes of higher amplitude than $\zeta_s(x)$ in (1.4) for $\mu > 9$, and the system undergoes a transcritical bifurcation in a neighbourhood of this parameter value. For the third regime, the second (stable) stationary solution exists with a negative amplitude. The weak instability of the positive amplitude solutions seems to evolve

into the same regime as obtained by taking the rest initial condition. This feature in particular seems amenable for further analytical investigations.

From the present study we may take note of the richness of new physical and mathematical contents of the general subject of nonlinear, dispersive systems sustaining forcings at resonance. The time-periodic births of upstream radiating solitary waves in response to a steady transcritical forcing can now be attributed to the instability of the primary wave over a fairly broad range of transcritical velocity and finite amplitude of forcing disturbances. For the velocity parameter above a certain threshold ($\mu > 9$ for forcing (i)) and for the amplitude parameter below a certain margin ($\alpha < 6$ for forcing (ii)), the phenomenon of periodic generation of solitary waves ceases to manifest because the forcing is moving too fast to be out raced by any free wave in case (i) or because the resulting wave is too weak to bring the nonlinearity to effect. No special attention is given here separately to the subcritical forcings since the upstream radiation becomes relatively weak in this regime, though the phenomenon can persist to velocities as low as the Froude number $F = 0.2$ (Lee 1985; Lee et al. 1989).

In the mathematical context, the theory of eigenvalue problems seems to be still under-developed for ordinary differential equations of the third order as in the present case. The different physical features exhibited of the phenomenon in different regimes of the parametric space are found to correlate closely with the eigenvalues being either non- existing, or purely real, or being complex conjugate. For the last case, we found that numerical methods of high accuracy are indispensable. In fact, the previous attempts by Wu (1988), using the Galerkian modal expansion method up to four terms retaining the nonlinearity, and by Camassa (1986) for the linearized

problem with up to 400 terms and using various basis functions, failed to reach a definitive determination of the real component of the eigenvalue because of its minuteness compared with the imaginary component. Such a broad disparity between the real and imaginary parts of the eigenvalue seems to be a hallmark characteristic of this class of problems which requires further investigation. From the theoretical standpoint, the variational approach expounded by Whitham (1967, 1974) and by Lighthill (1967) may be valuable in problems involving nonlinear dispersive waves governed by equations too complicated for analytical or numerical treatment, such as those considered here.

Finally, we note that type of bifurcation of solution from the primary wave seems to be new in nature. The significance and possible impact of the issues pointed out here deserves a continued attention and investigation which is underway.

Appendix A

We briefly report here some of the results mentioned in the text regarding the application of Noether's first theorem (Bogoliubov and Shirkov, 1980, §2) to the evolution equations of interest.

Introducing the potential

$$\zeta(x,t) \equiv \phi_x(x,t) \ , \tag{A.1}$$

a Lagrangian for the fKdV equation (1.1) can be written as (see Whitham, 1974, §16.14)

$$\mathcal{L}(\phi) = \frac{1}{2}\int_{-\infty}^{+\infty} \left[\phi_t\phi_x + \mu\phi_x^2 - 3\phi_x^3 + \phi_{xx}^2 - 6P(x)\phi_x\right]\ dx \ , \tag{A.2}$$

as one can check by the Euler-Lagrange equations,

$$\frac{d}{dt}\frac{\delta \mathcal{L}}{\delta \phi_t} - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \,, \tag{A.3}$$

which reproduce equation (1.1), once the similarity transformation (1.10) is applied. We note that (A.2) has a structure similar to a Lagrangian describing the scattering of a field by an external one, whose evolution is not affected by the interaction (see Bogoliubov and Shirkov, *ibid.* §24). By Noether's first theorem, if

$$\phi'(x',t') = \phi(x,t) + \delta\phi(x,t) \tag{A.4}$$

is a one parameter transformation, ξ say, of the field ϕ and the coordinates x, t, for which

$$\mathcal{L}'(\phi') = \mathcal{L}(\phi) + \frac{d}{dt}\mathcal{F}(\phi;\xi) \,, \tag{A.5}$$

then the quantity

$$Q \equiv \left(\int_{-\infty}^{+\infty} dx \frac{\delta \mathcal{L}}{\delta \phi_t} \frac{\partial \phi'}{\partial \xi}|_{\xi=0} \right) - \frac{\partial \mathcal{F}}{\partial \xi}|_{\xi=0} \tag{A.6}$$

is a constant of motion. Now, from the structure of the Lagrangian (A.2), it is obvious that we have invariance under the transformation

$$\phi'(x',t') = \phi(x,t) + \xi, \quad \mathcal{F} = 0 \,, \tag{A.7}$$

and so, according to (A.6),

$$\mathcal{M} \equiv \frac{1}{2}\int_{-\infty}^{+\infty} \phi_x \, dx = \frac{1}{2}\int_{-\infty}^{+\infty} \zeta \, dx \tag{A.8}$$

is conserved. The spatial translation

$$\phi'(x',t') = \phi(x+\xi,t) \tag{A.9}$$

does not yield the form (A.5) for the transformed Lagrangian, unless $P(x) = \text{const.}$, in which case is easy to verify that $\mathcal{F} = 0$ and the associated conserved quantity is

$$\mathcal{E} \equiv \frac{1}{2}\int_{-\infty}^{+\infty} \phi_x^2 \, dx = \frac{1}{2}\int_{-\infty}^{+\infty} \zeta^2 \, dx \, . \tag{A.10}$$

If the forcing P is independent of time, then translation with respect to time

$$\phi'(x',t') = \phi(x, t+\xi) \tag{A.11}$$

when ξ is infinitesimal, leads to a transformed Lagrangian

$$\mathcal{L}'(\phi') = \mathcal{L}(\phi) + \xi \frac{d}{dt}\mathcal{L}(\phi) + O(\xi^2) \, , \tag{A.12}$$

and so (A.5) holds with $\mathcal{F} \equiv \mathcal{L}$, and the associated conserved quantity is

$$\begin{aligned} \mathcal{H} = \left(\int_{-\infty}^{+\infty} dx \frac{\delta \mathcal{L}}{\delta \phi_t}\frac{\partial \phi}{\partial t}\right) - \mathcal{L} &= \frac{1}{2}\int_{-\infty}^{+\infty} \left[\mu\phi_x^2 - 3\phi_x^3 + \phi_{xx}^2 - 6P(x)\phi_x\right] \, dx \, , \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} \left[\mu\zeta^2 - 3\zeta^3 + \zeta_x^2 - 6P(x)\zeta\right] \, dx \, , \end{aligned} \tag{A.13}$$

which is the Hamiltonian considered in the text. There are no other obvious invariances possessed by the Lagrangian, at least for a general forcing $P(x)$. For the regularized fKdV equation (1.118), the Lagrangian (A.2) has to be modified into

$$\mathcal{L}_r = \frac{1}{2}\int_{-\infty}^{+\infty} \left[\phi_t\left(\phi_x - \phi_{xxx}\right) + 6(F-1)\phi_x^2 - 3\phi_x^3 + F\phi_{xx}^2 - 6P(x)\phi_x\right] \, dx \, , \tag{A.14}$$

and the invariants corresponding to the one parameters transformations considered above are

$$\mathcal{E}_r \equiv \frac{1}{2}\int_{-\infty}^{+\infty} \left[\zeta^2 + \zeta_x^2\right] \, dx \, , \tag{A.15}$$

and

$$\mathcal{H}_r \equiv \frac{1}{2}\int_{-\infty}^{+\infty} \left[6(F-1)\zeta^2 - 9\zeta^3 + F\zeta_x^2 - 6P(x)\zeta\right] \, dx \, , \tag{A.16}$$

respectively for (A.9), when P is a constant, and (A.11).

References

[1] Ablowitz, M.J., Kaup, D.J., Newell, A.C. & Segur, H. 1974 The Inverse Scattering Transform-Fourier analysis for nonlinear problems. *Studies in Appl. Math.* **53**, 249-315.

[2] Benjamin, T.B. 1971 The stability of solitary waves. *Proc. R. Soc. Lond.* **328**, 153-183.

[3] Benjamin, T.B., Bona, J.L. & Mahony, J.J. 1972 Model equation for long waves in nonlinear dispersive systems. *Phil. Trans. R. Soc. Lond.* **A272**, 47-78.

[4] Bogoliubov, N.N. & Shirkov, D.V. 1980 *Introduction to the Theory of Quantized Fields.* Wiley Interscience, New York.

[5] Bona, J. 1975 On the stability of solitary waves. *Proc. R. Soc. Lond.* **344**, 363-374.

[6] Bona, J.L. & Smith, R. 1975 The initial value problem for the Korteweg-De Vries equation. *Phil. Trans. R. Soc. Lond.* **A278**, 556-601.

[7] Boussinesq, J. 1872 Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. *J. Math. Pures Appl.* **17(2)**, 55-108.

[8] Camassa, R. 1986 Graduate Research Report, California Institute of Technology.

[9] Cheng, S.L. 1987 Generation of solitons in shallow waters. M.Sc. thesis, National Taiwan University.

[10] Cole, S.L. 1985 Transient waves produced by flow past a bump. *Wave Motion* **7**, 579-587.

[11] Gardner, C.S., Greene, J.M., Kruskal, M.D. & Miura, R.M. 1967 Method for solving the KdV equation. *Phys. Rev. Lett.* **19**, 1095-1097.

[12] Grimshaw, R.H.J. & Smyth, N. 1986 Resonant flow of a stratified fluid over a topography. *J. Fluid. Mech.* **169**, 429-464.

[13] Holm, D.D., Marsden, J.E., Ratiu, T. & Weinstein, A. 1985 Nonlinear stability of fluid and plasma equilibria. *Phys. Rep.* **123**, 1-116.

[14] Jeffrey, A. & Kakutani, T. 1972 Weak nonlinear dispersive waves: a discussion centered around the Korteweg-de Vries equation. *SIAM Rev.* **14**, 582-643.

[15] Kevorkian, J. & Cole, J.D. 1981 *Perturbation Methods in Applied Mathematics.* Springer-Verlag, New York.

[16] Korteweg, D.J. & de Vries, G. 1895 On the change of form of long waves advancing in a rectangular channel, and on a new type of long stationary waves. *Phil. Mag.* **39**, 422-443.

[17] Landau, L. & Lifshitz, E. 1955 *Quantum Mechanics, Nonrelativistic Theory.* Pergamon Press, London.

[18] Lee, S.J. 1985 Generation of long waterwaves by moving disturbances. Ph.D. thesis, California Institute of Technology.

[19] Lighthill, M.J. 1967 Some special cases treated by the Whitham theory. *Proc. R. Soc. Lond. A* **299**, 28-53.

[20] Malanotte-Rizzoli, P.M. 1984 Boundary-forced nonlinear planetary radiation. *J. of Phys. Oceanogr.* **14**, 1032-1047.

[21] Newell, A.C. 1985 *Solitons In Mathematics and Physics.* SIAM, Philadelphia.

[22] Patoine, A. & Warn, T. 1982 The interaction of long, quasi-stationary baroclinic waves with topography. *J. Atmos. Sci.* **39**, 1019-1025.

[23] Tanaka, M. 1985 The stability of solitary waves. *Phys. Fluids* **29**, 650-655.

[24] Whitham, G.B. 1974 *Linear and Nonlinear Waves.* Wiley Interscience, New York.

[25] Whitham, G.B. 1967 Variational methods and applications to water waves. *Proc. R. Soc. Lond. A* **299**, 6-25.

[26] Wu, T.Y. 1981 Long waves in ocean and coastal waters. *J. Eng. Mech. Div., Proc. ASCE* **107**, 502-522.

[27] Wu, T.Y. 1987 On generation of solitary waves by moving disturbances. *J. Fluid Mech.* **184**, 75-99.

[28] Wu, T.Y. 1988 Forced generation of solitary waves. In *Applied Mathematics, Astrophysics, A Symposium to Honor C.C. Lin. 22-24 June 1987, M.I.T. Cambridge, MA.* (D.J. Benney, F.H. Shu & C. Yuan eds.), 198-212, World Scientific, Singapore.

[29] Wu, T.Y. & Wu, D.M. 1982 Three dimensional nonlinear long waves due to moving surface pressure. *Proc. 14th Symp. on Naval Hydrodynamics, Washington D.C.*, 103-125.

[30] Zabusky, N.J. & Kruskal, M.D. 1965 Interaction of solitons in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**, 241-243.

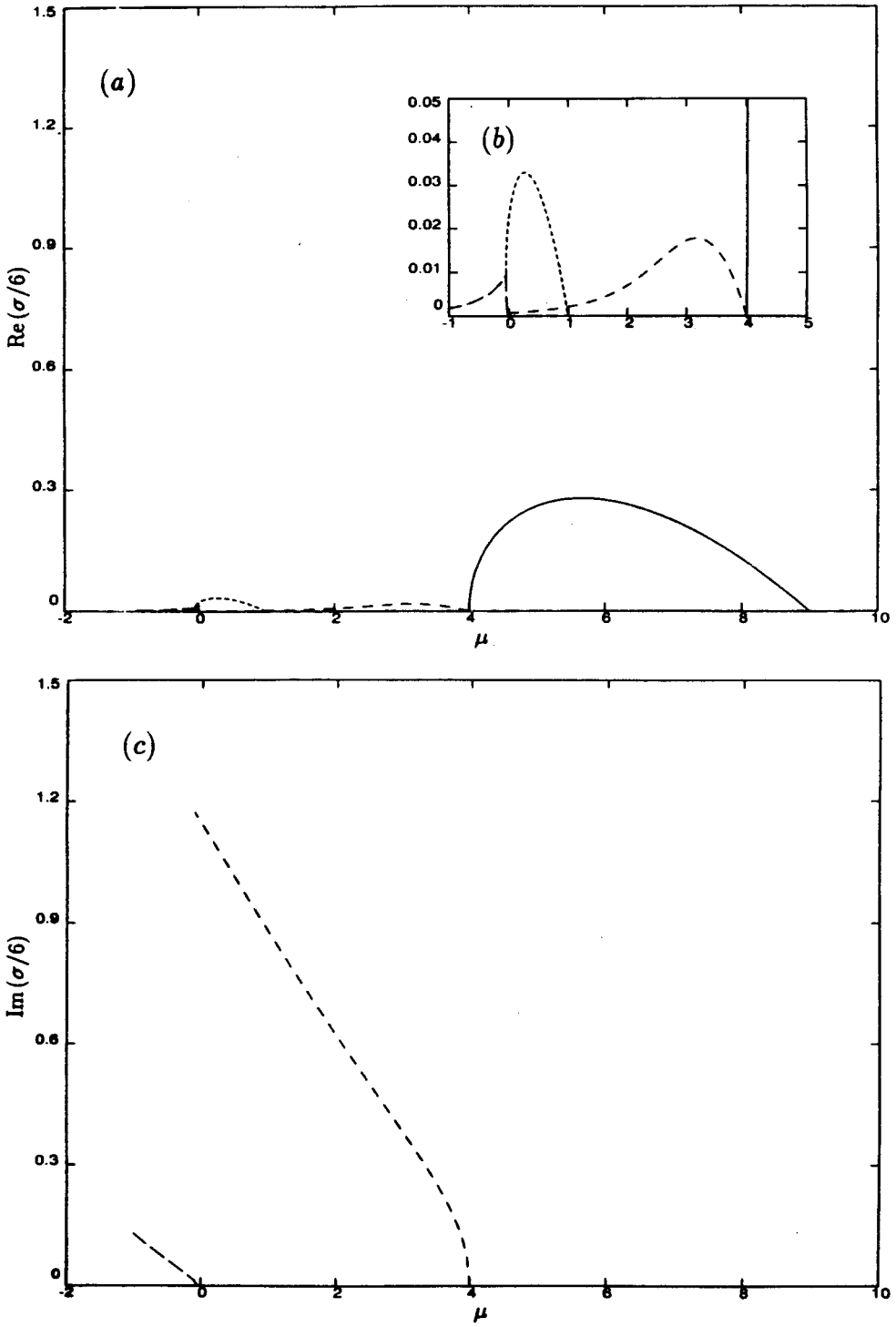

Figure 1.1. a) The real part of $\sigma/6$ vs. μ for forcing (i), as determined by the power series method of Section 1.4.4. b) A blow-up of the region close to the axis Re $\sigma = 0$ for $-1 < \mu < 5$. c) The imaginary part of $\sigma/6$ vs. μ.

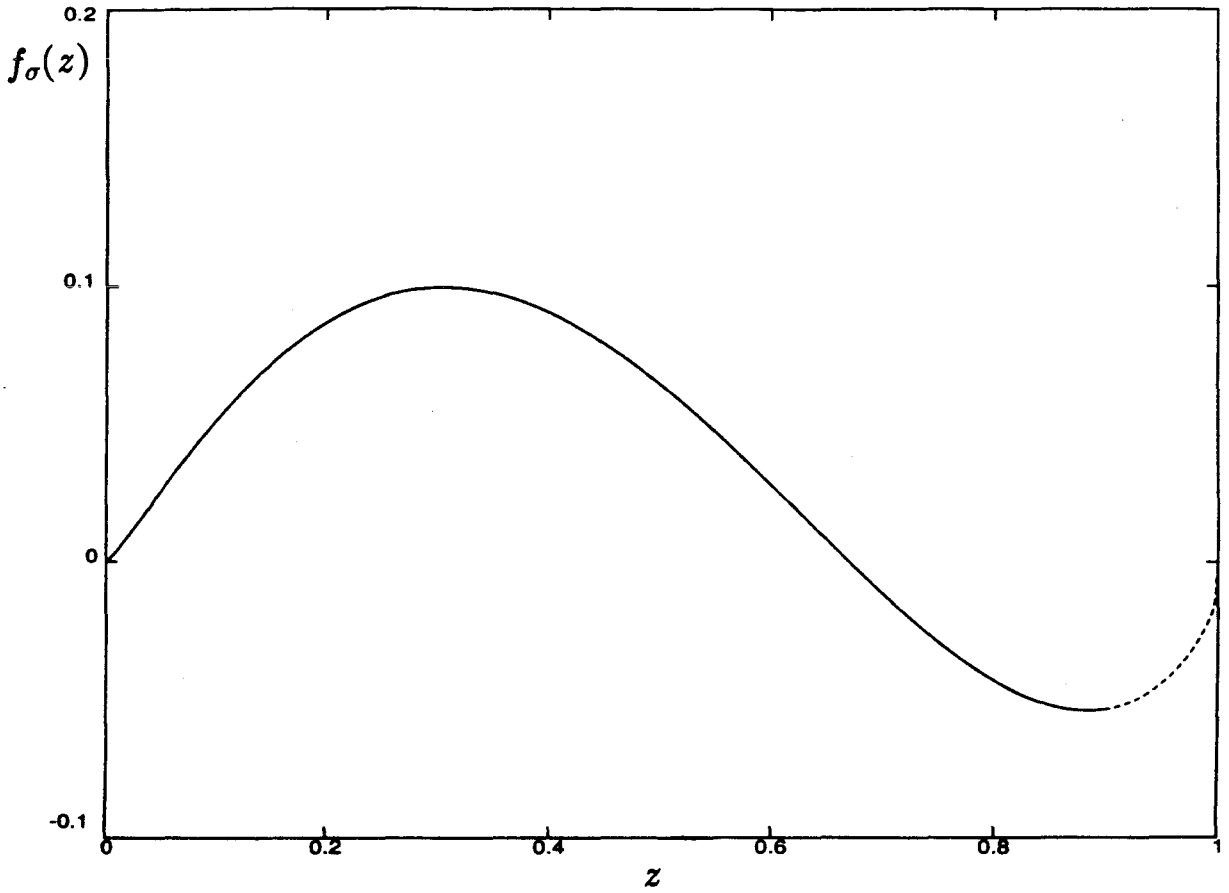

Figure 1.2. Eigenfunction corresponding to the eigenvalue $\frac{\sigma}{6} = 0.2605$ at $\mu = 5$ (forcing (i)), with respect to the independent variable z (equation (1.70)). The solid line is for $0 \leq z \leq 0.9$, the dashed for $0.1 \leq z \leq 1.0$.

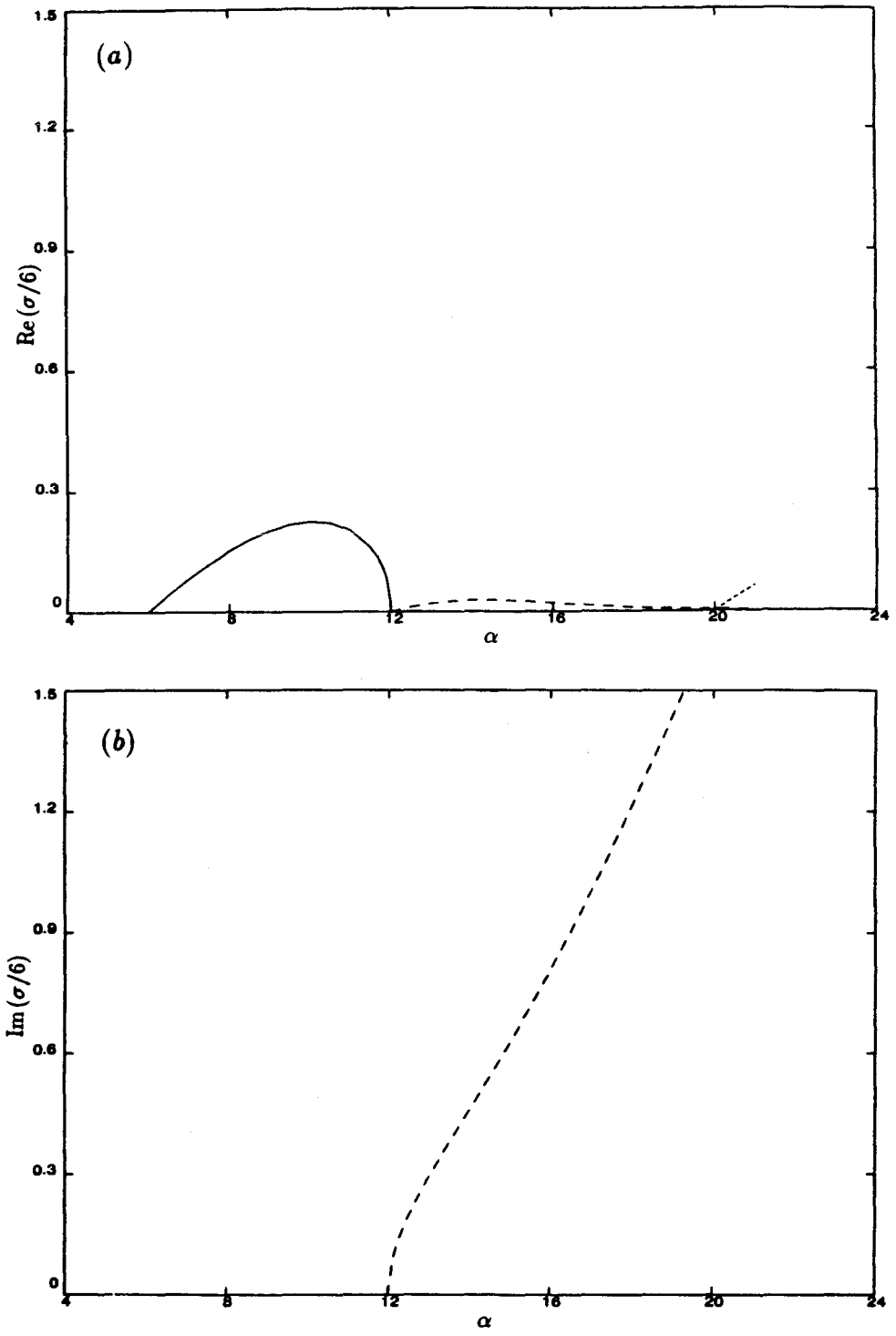

Figure 1.3. a) The real part of $\sigma/6$ vs. α for forcing (ii), as determined by the power series method of Section 1.4.4. b) Imaginary part of $\sigma/6$ vs. α.

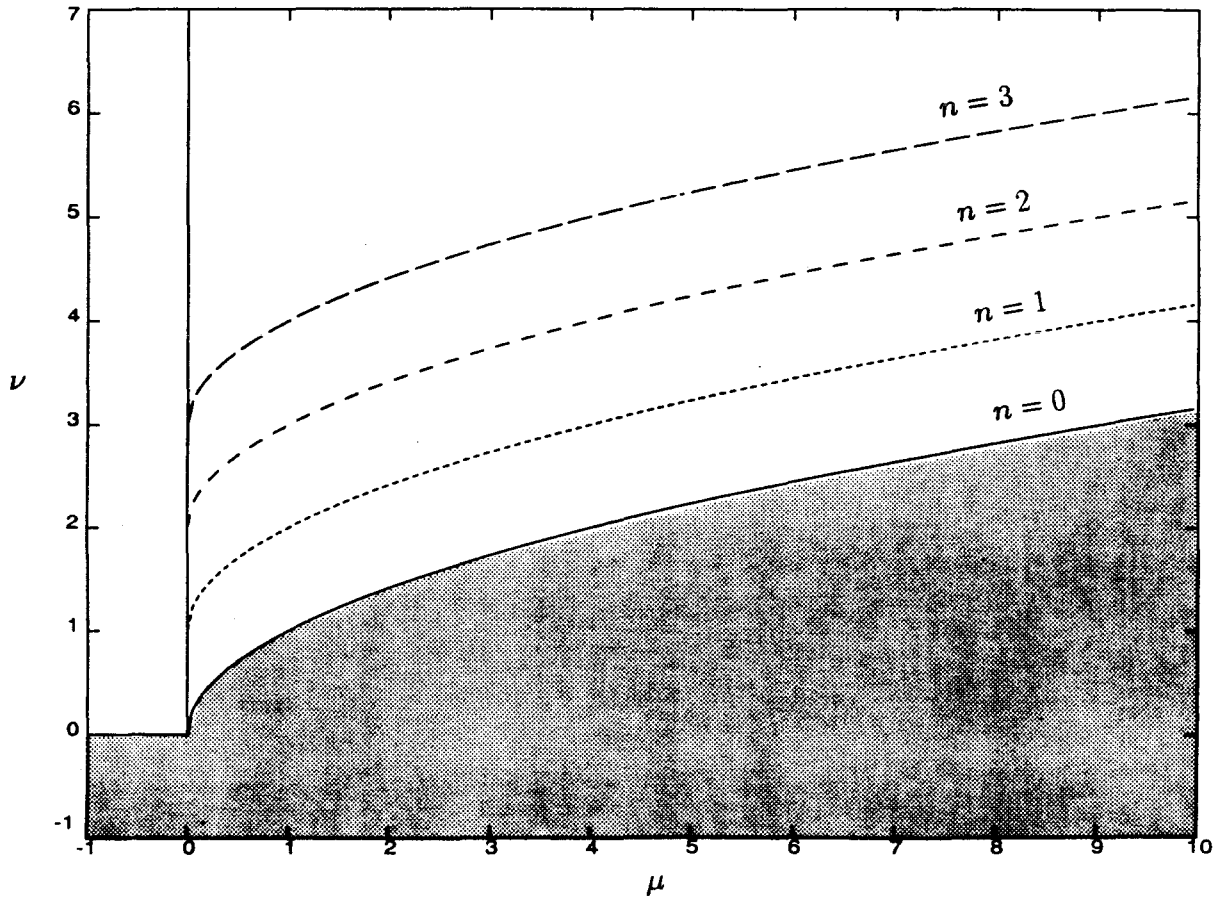

Figure 1.4. Subdivision of the parametric $(\mu, \nu(\alpha))$ plane for the existence of eigenvalues of the operator $\mathcal{L}_{\mu,\alpha}$. In the shaded region no eigenvalue with real part different from zero can exist.

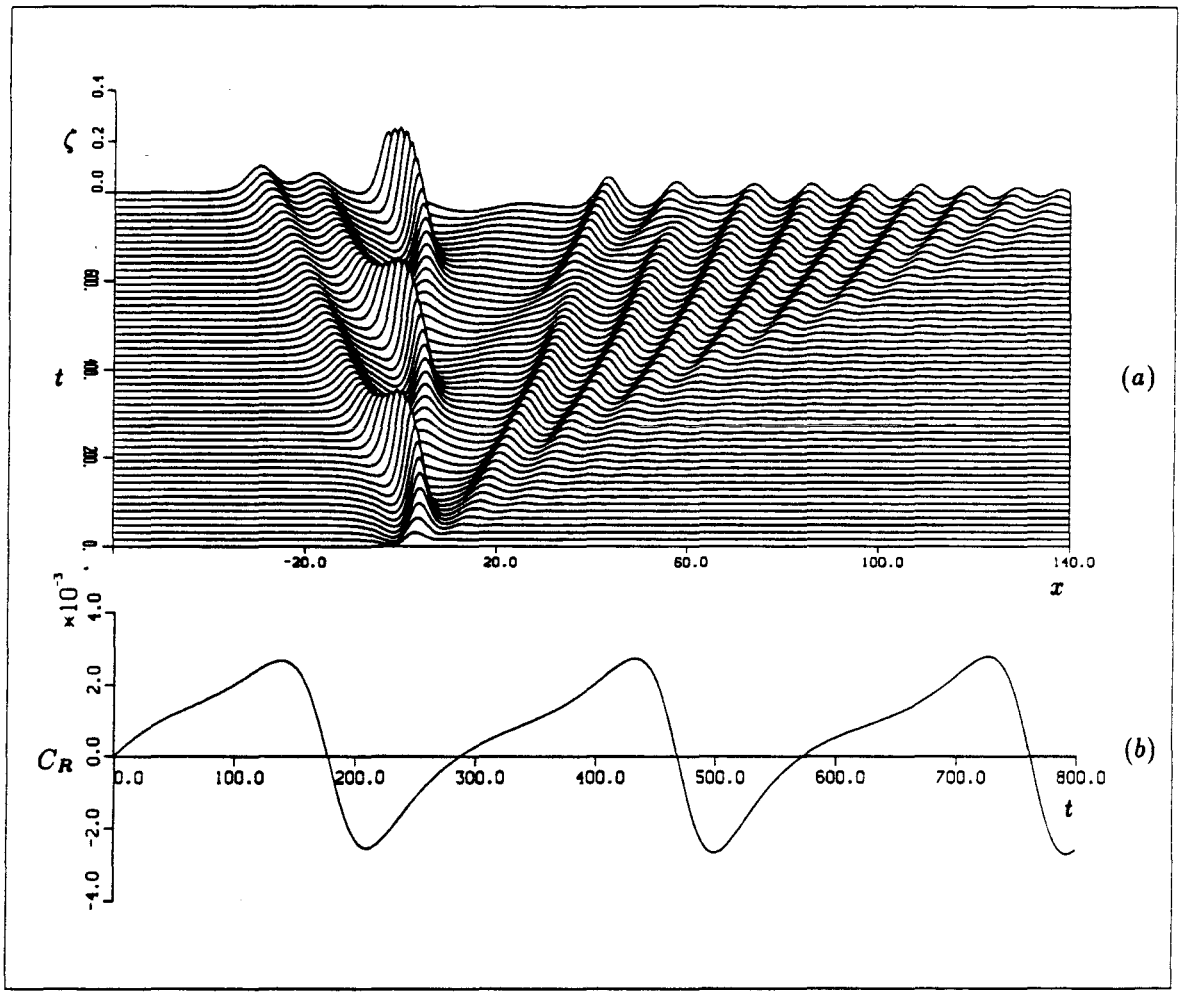

Figure 1.5. a) The phenomenon of generation of upstream running waves for forcing (i), with the initial condition $\zeta(x,0) = 0$ ($\eta(x,0) = -\zeta_s(x)$). The parameters for this simulation are $\mu = 0$ ($F = 1$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. Its nondimensional period of oscillation T' is $\simeq 7.9$ and coincides with the period of birth of an upstream running wave.

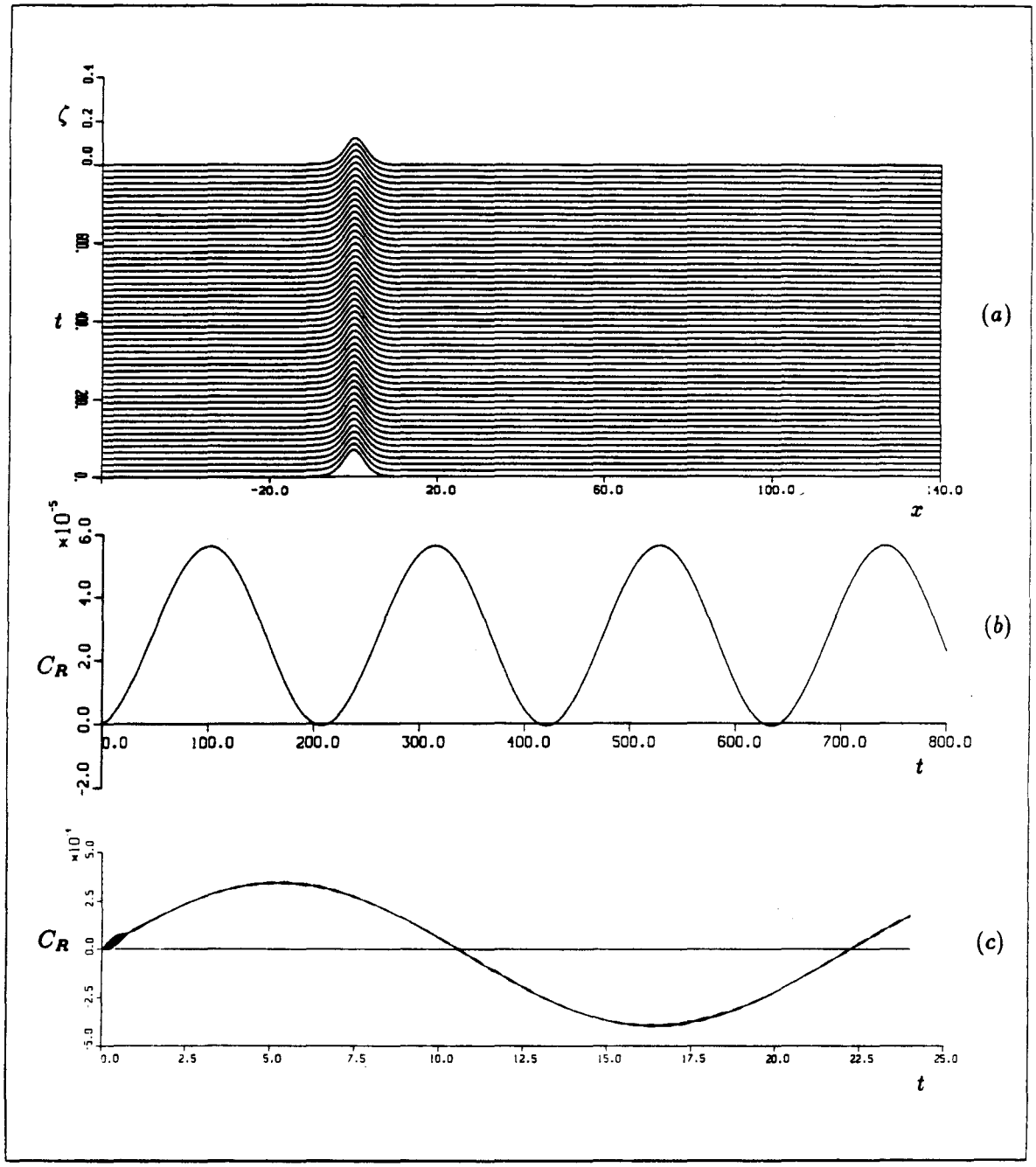

Figure 1.6. a) Evolution of the stationary solution $\zeta_s(x)$ for forcing (i) with the initial condition $\zeta(x,0) = \zeta_s(x)$. The parameters for this simulation are $\mu = 0$ ($F = 1$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The nondimensional period of oscillation $T' \simeq 5.75$. c) Same as b) but using the explicit code, with $k = 0.5$ and $\Delta x = 0.1$, $\Delta t = 0.002$. In this case $T' \simeq 5.54$.

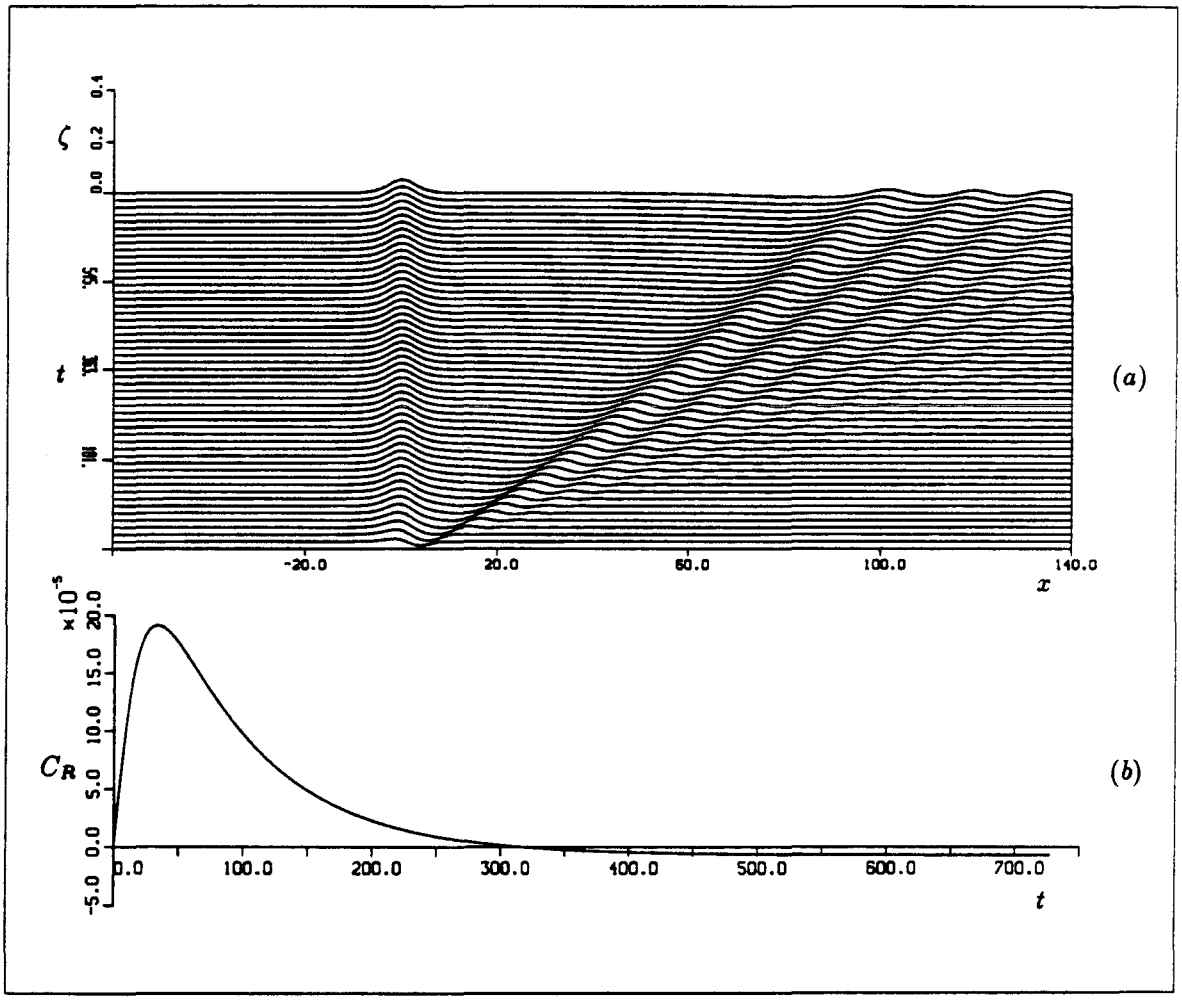

Figure 1.7. a) Generation of the second (stable) stationary wave ζ_{s_2} for forcing (i), when the initial condition is $\zeta(x,0) = 0$ ($\eta(x,0) = -\zeta_s(x)$). The parameters for this simulation are $\mu = 6.\bar{6}$ ($F = 1.1$), $k = 0.3$ computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time exhibits no periodic behaviour. The negative small asymptote is seen to be of numerical origin, as it is halved by halving the grid size.

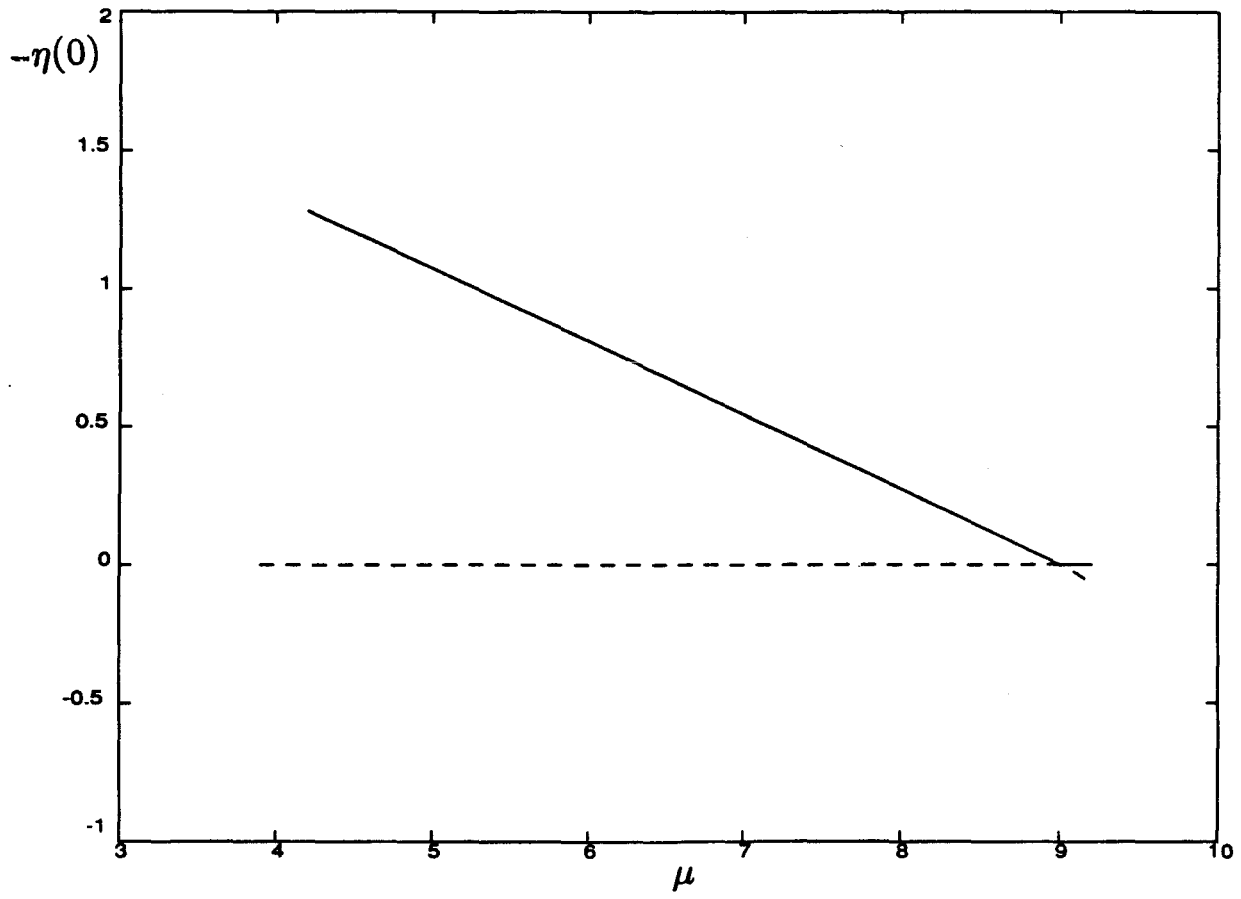

Figure 1.8. Bifurcation diagram pertaining to the second stationary solution for forcing (i), ζ_{ss}, as obtained by numerically solving (1.3), using a 4th order Runge-Kutta integrator and bisection method for matching the zero boundary condition at large x . Plotted here is the stationary point of $-\eta_{ss}(x) = \zeta_s(x) - \zeta_{ss}(x)$ vs. μ, the dashed line referring to instability, the solid to stability.

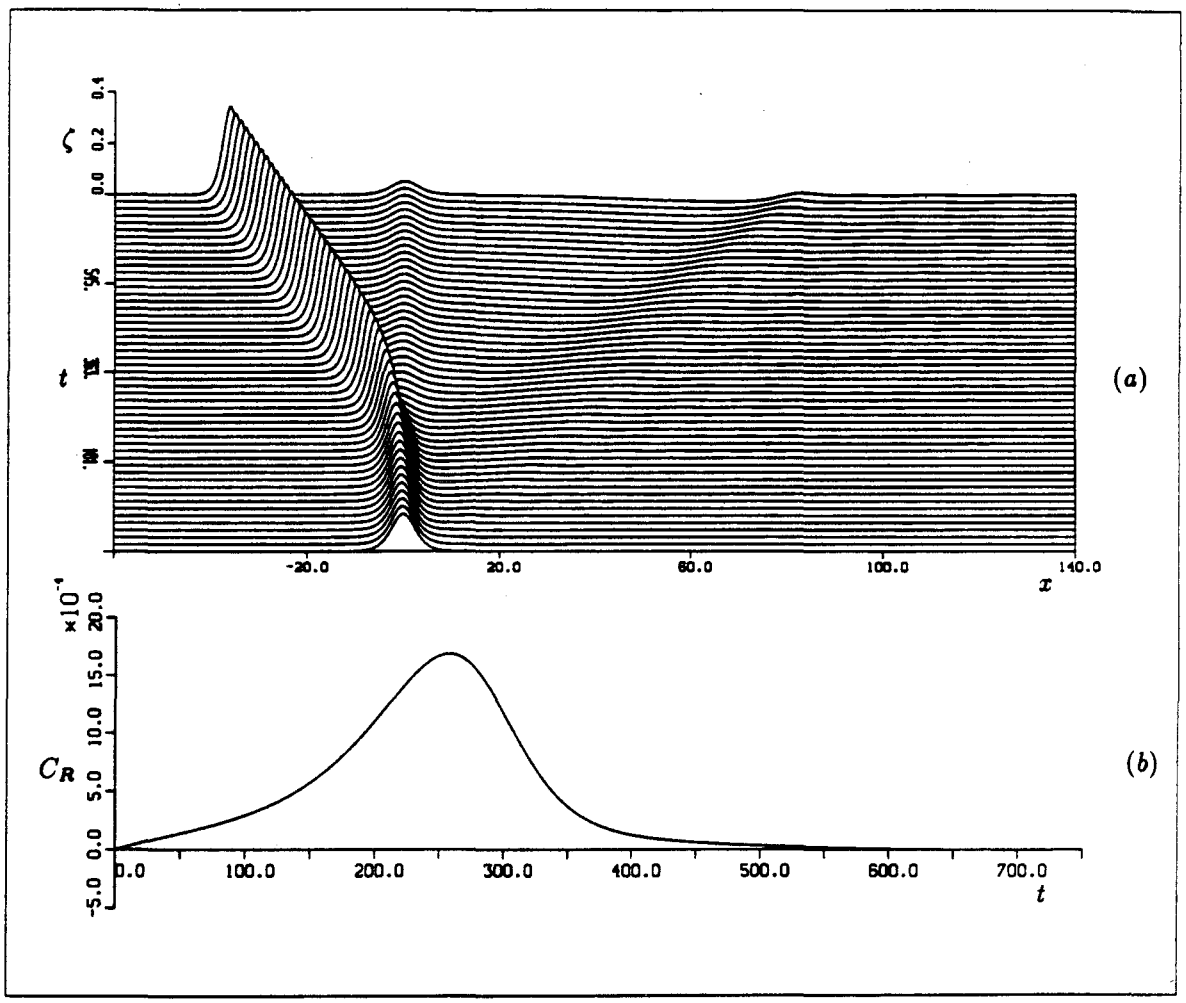

Figure 1.9. a) Growth of the instability for forcing (i), with the perturbed initial condition $\zeta(x,0) = \zeta_s(x) + \frac{1}{10}\zeta_s(x)$ ($\eta(x,0) = \frac{1}{10}\zeta_s(x)$). The system sheds off the extra excess mass with respect to the new stable state by sending a solitary wave upstream and a dispersive wave train downstream. The parameters for this simulation are $\mu = 6.\bar{6}$ ($F = 1.1$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time.

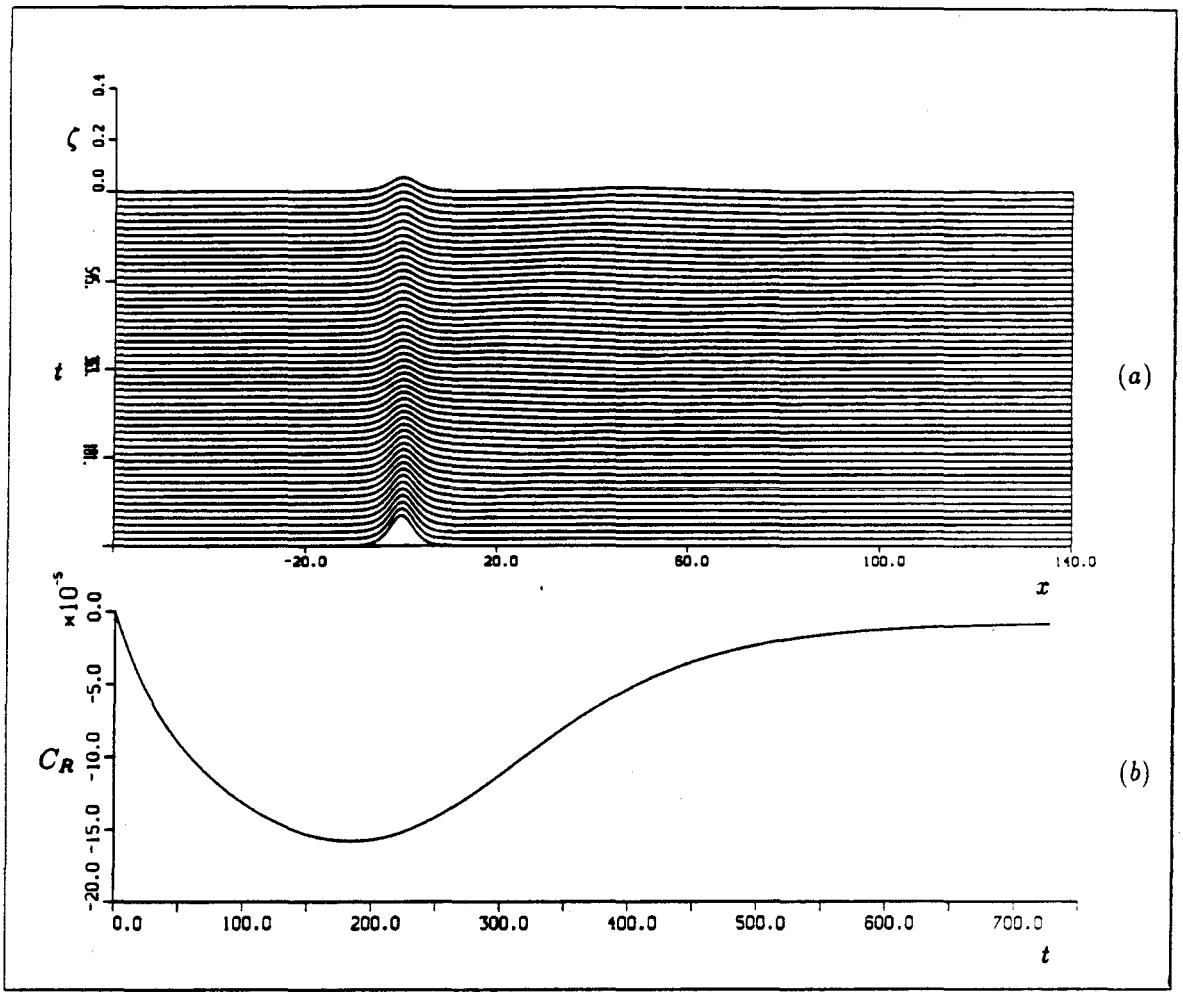

Figure 1.10. a) Growth of the instability for forcing (i), for the perturbed initial condition $\zeta(x,0) = \zeta_s(x) - \frac{1}{10}\zeta_s(x)$ ($\eta(x,0) = -\frac{1}{10}\zeta_s(x)$). The system gets rid of the extra mass with respect to the new stable state by sending only a weak dispersive wave train downstream. The parameters for this simulation are $\mu = 6.\bar{6}$ ($F = 1.1$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The system is releasing a small amount of energy to the forcing in connection with emission of downstream running waves.

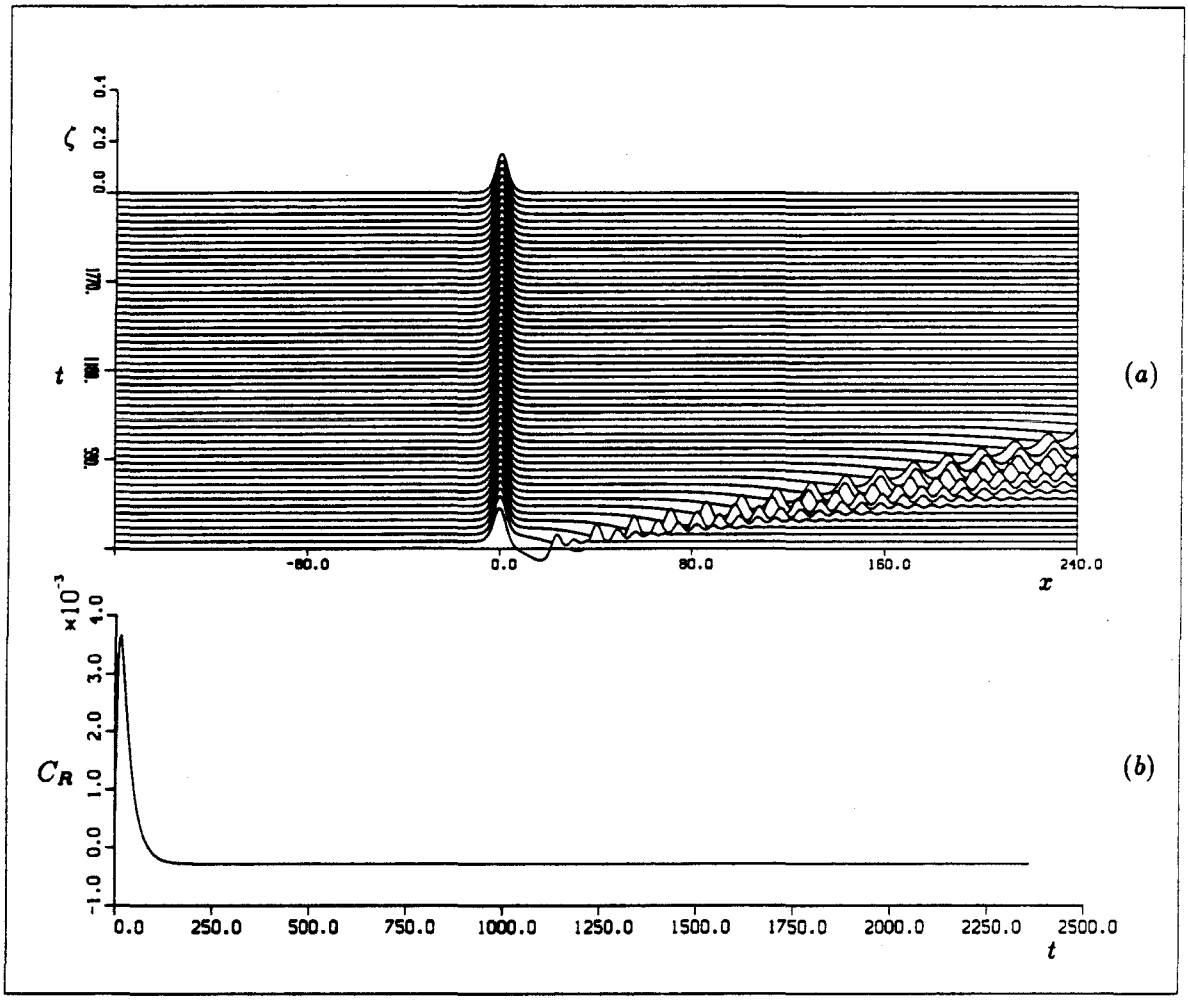

Figure 1.11. a) The generation of the stationary solution $\zeta_s(x)$ by forcing (i), with the initial condition $\zeta(x,0) = 0$ ($\eta(x,0) = -\zeta_s(x)$). The system goes to the stationary solution $\zeta_s(x)$, emitting a train of dispersive waves downstream. The parameters for this simulation are $\mu = 18$ ($F = 1.271$), $k = 0.3$, computed with $\Delta x = 0.2$, $\Delta t = 0.1$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The system absorbs energy initially in generating a stationary wave.

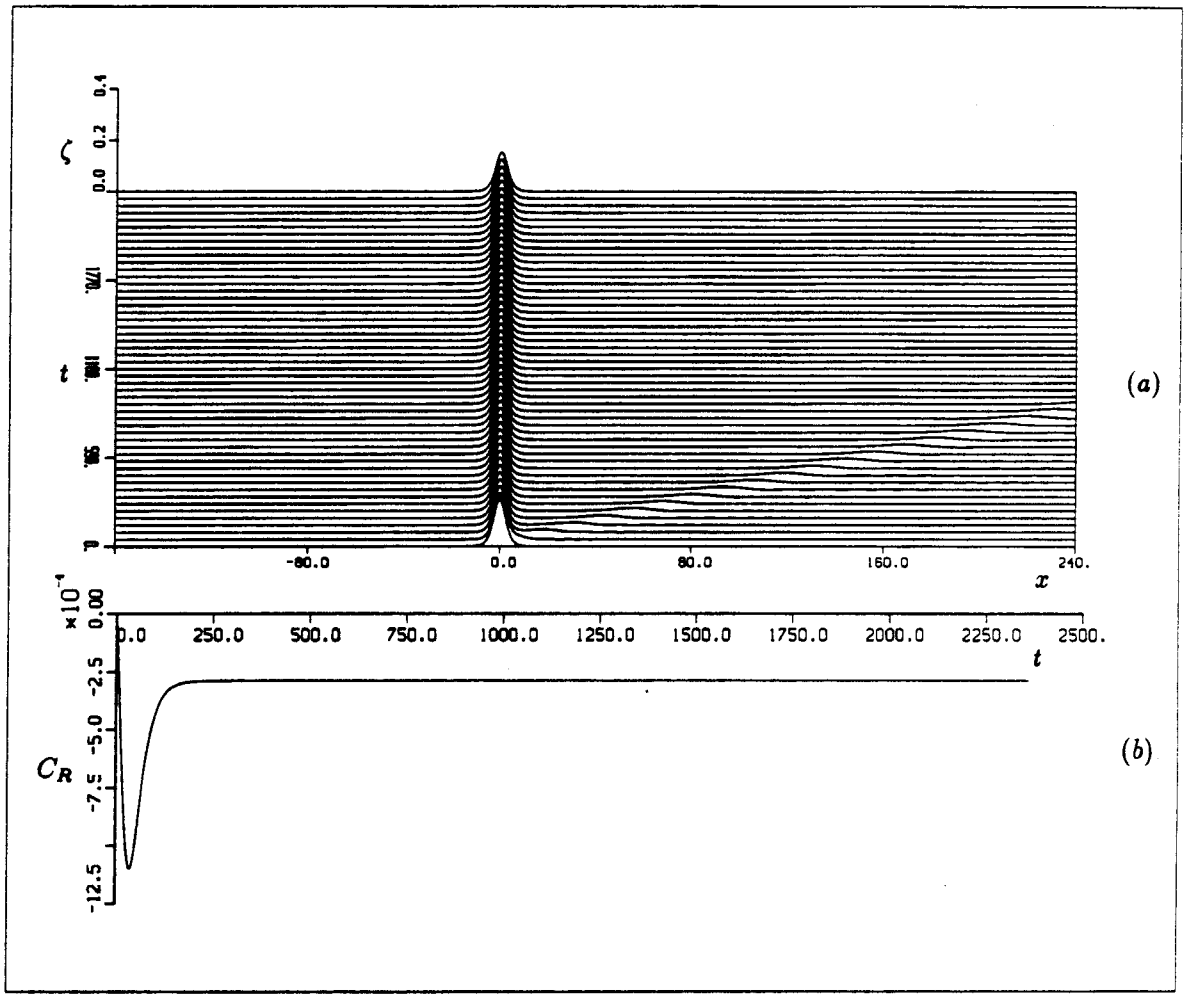

Figure 1.12. a) Stability of the stationary solution $\zeta_s(x)$ for forcing (i), with the initial condition $\zeta(x,0) = \zeta_s(x) + \frac{1}{5}\zeta_s(x)$ ($\eta(x,0) = \frac{1}{5}\zeta_s(x)$). The system goes to the stationary solution $\zeta_s(x)$, emitting a train of dispersive waves downstream. The parameters for this simulation are $\mu = 18$ ($F = 1.271$), $k = 0.3$, computed with $\Delta x = 0.2$, $\Delta t = 0.1$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The system initially releases energy through emission of a dispersive wave train downstream.

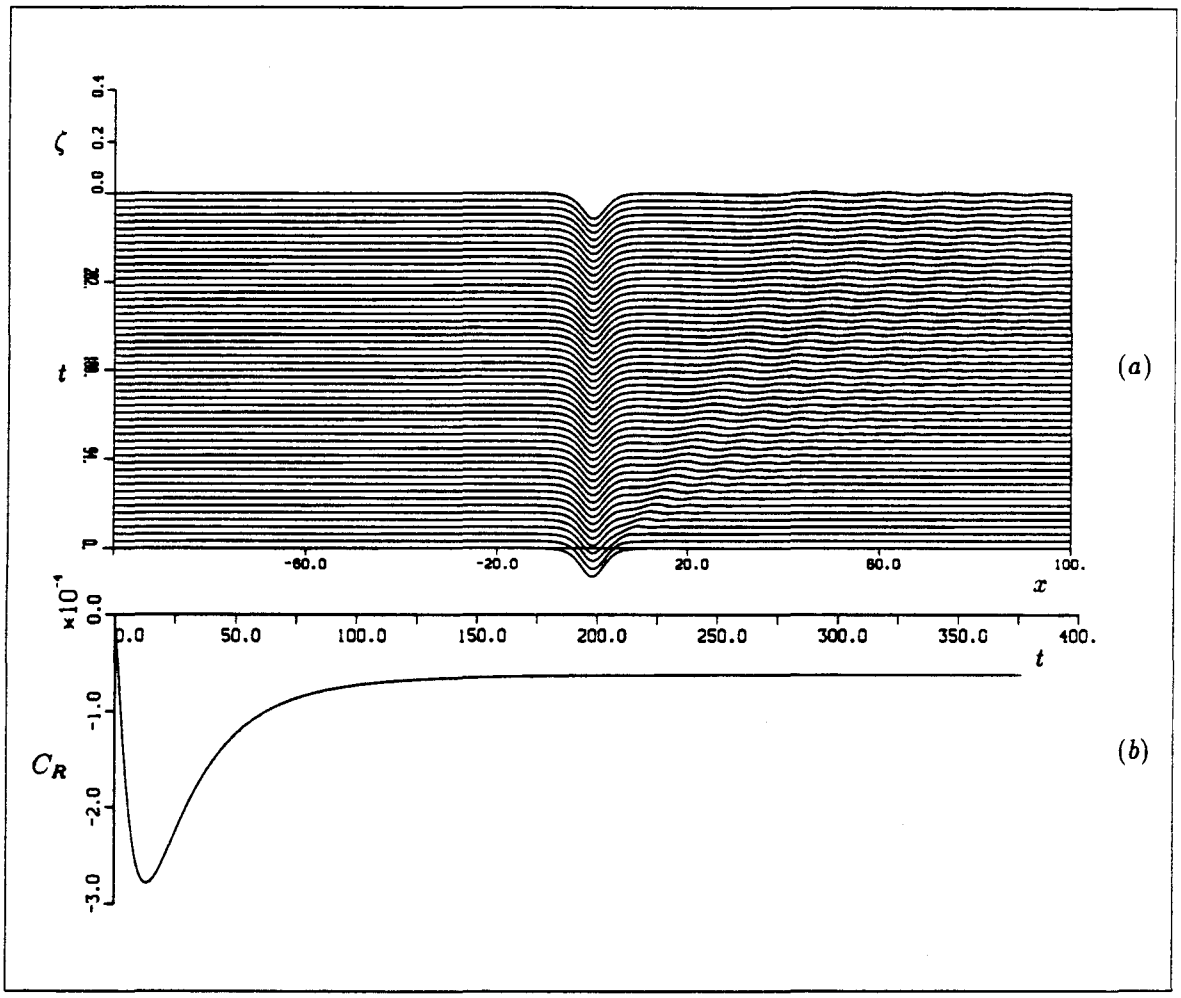

Figure 1.13. a) Stability of the smaller of the stationary solutions $\zeta_s(x)$ for forcing (ii), with the initial condition $\zeta(x,0) = \zeta_s(x) + \frac{1}{10}\zeta_s(x)$ ($\eta(x,0) = \frac{1}{10}\zeta_s(x)$). The system goes to the stationary solution $\zeta_s(x)$, emitting a train of dispersive waves downstream. The parameters for this simulation are $\alpha = -10$ ($F = 1.064$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The system initially releases energy through emission of a dispersive wave train downstream.

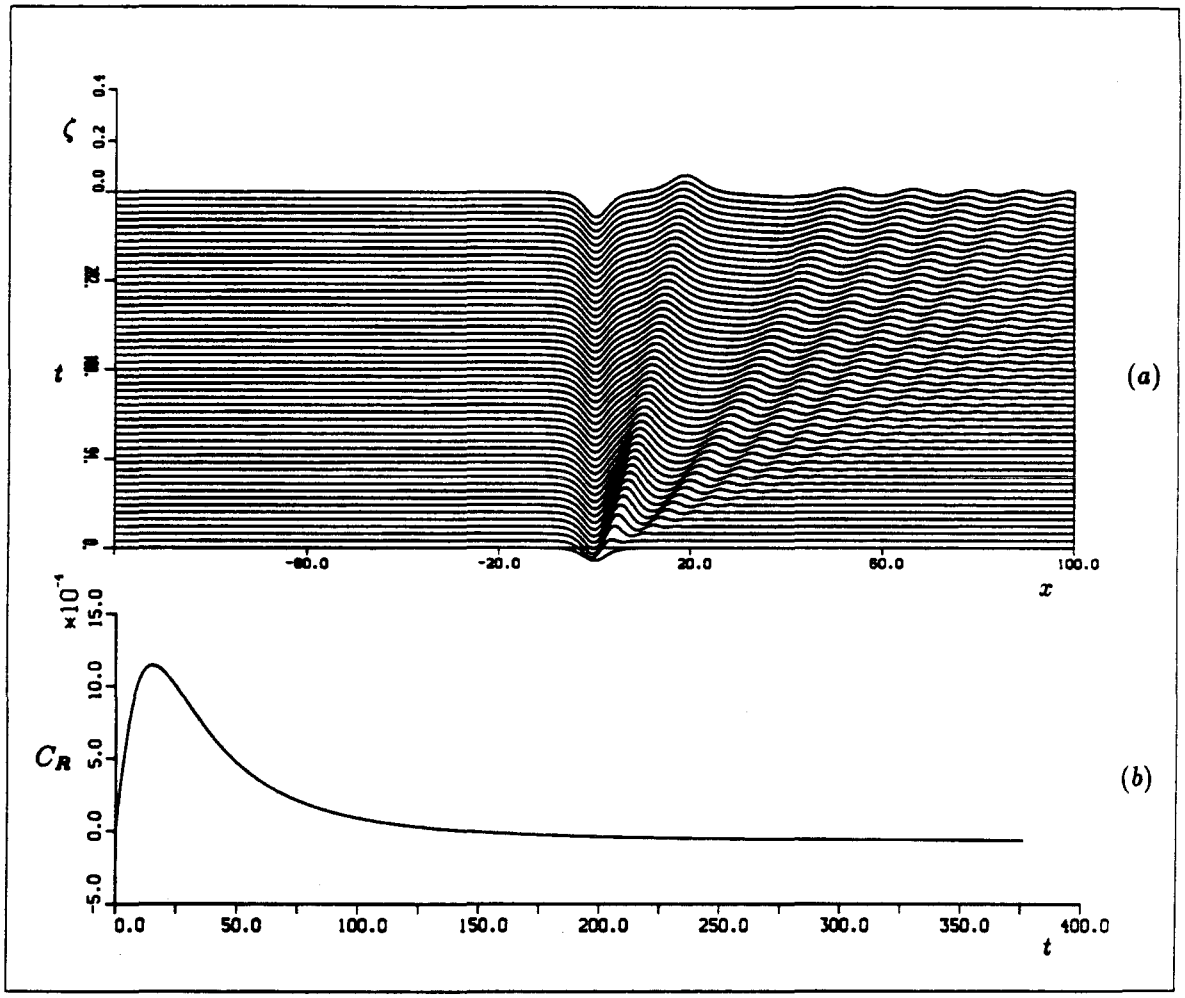

Figure 1.14. a) Evolution of the system for forcing (ii), with the initial condition perturbed around the smaller of the stationary solutions $\zeta_s(x)$ with $\zeta(x,0) = \zeta_s(x) - \frac{1}{2}\zeta_s(x)$ ($\eta(x,0) = -\frac{1}{2}\zeta_s(x)$). The system still goes to the stationary solution $\zeta_s(x)$, emitting a train of dispersive waves downstream. The parameters for this simulation are $\alpha = -10$ ($F = 1.064$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The system initially absorbs energy together with emission of a dispersive wave train downstream.

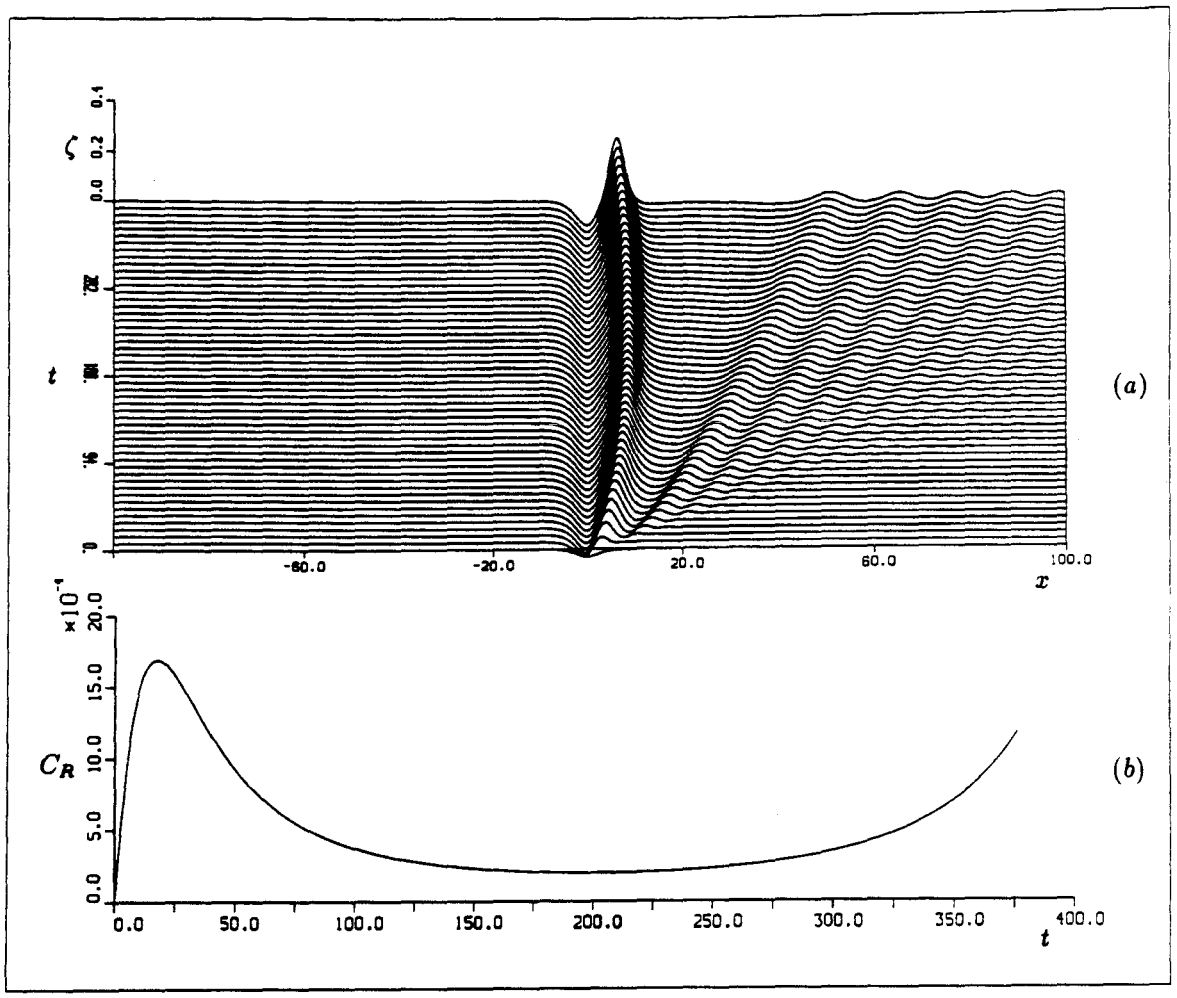

Figure 1.15. a) Evolution of the system for forcing (ii), with the initial condition perturbed around the smaller of the stationary solutions $\zeta_s(x)$ with $\zeta(x,0) = \zeta_s(x) - \frac{7}{10}\zeta_s(x)$ $(\eta(x,0) = -\frac{7}{10}\zeta_s(x))$. An upstream running wave is being created in the region following the forcing. The parameters for this simulation are $\alpha = -10$ $(F = 1.064)$, $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The energy absorption of the system grows monotonically after reaching a positive minimum.

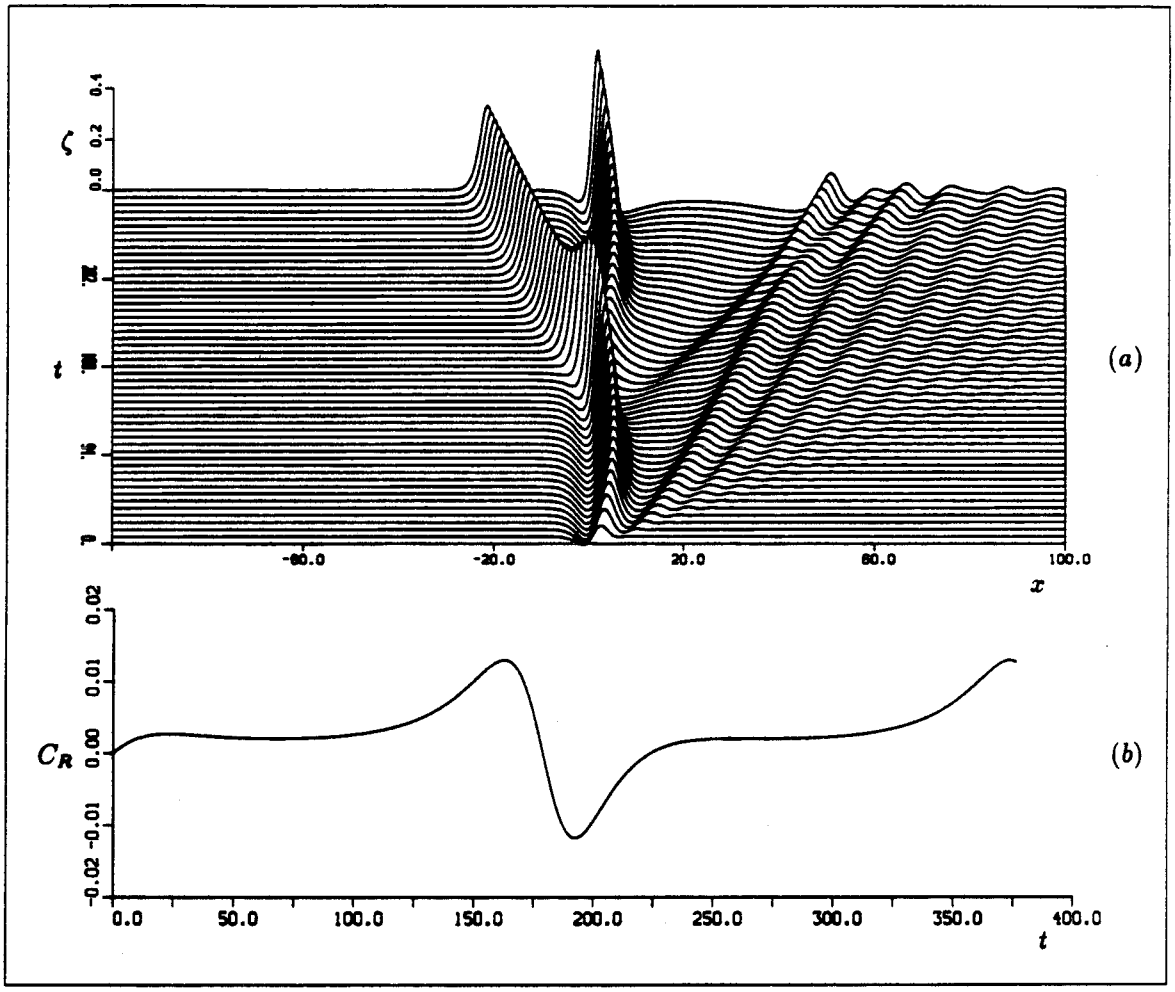

Figure 1.16. a) The evolution of the system for forcing (ii), with the initial condition at rest $\zeta(x,0) = 0$ ($\eta(x,0) = -\zeta_s(x)$). The phenomenon of upstream running wave is regained. The parameters for this simulation are $\alpha = -10$ ($F = 1.064$), $k = 0.3$, computed with $\Delta x = 0.1$, $\Delta t = 0.05$, using the implicit code. b) The corresponding wave resistance coefficient vs. time. The energy absorption of the system from the forcing is periodic, after an initial transient similar to the evolution depicted in Figure 15.

Chapter 2

The KdV Model With Boundary Forcing

2.1 Introduction

The equation of Korteweg and de Vries (KdV; 1895)[8] belongs to a class of nonlinear evolution equations which provide good models for predicting a variety of physical phenomena, and have the further advantage of being solvable by analytical means. However, the rather ingenious techniques developed so far for these models are primarily meant for the unbounded space, and cannot be applied in the presence of boundaries occurring at finite distances, with arbitrary boundary conditions or when external forcings exist, a situation which certainly is of great practical importance. For instance, it is known that this equation offers a satisfactory description of the evolution of long surface waves initially prescribed for shallow water in a channel, when the end boundaries are sufficiently far removed to avoid reflections. The integrability of the KdV equation would afford determination of the asymptotic behaviour of the solution evolved from some smooth initial shape of the free surface. In a more general situation of wave generation however, the free surface may be initially undisturbed and waves are created by the introduction of fluid mass or by the action of some mechanical device like a wave maker at one end, or by a moving bottom topography.

This generalization would lead to consideration of boundary conditions and forcing functions in supplement to the appropriately chosen evolution equation. In order to provide some insight for developing analytical methods for the general case, we investigate the KdV equation in presence of a boundary forcing, both numerically and expositively by introducing an approximate method based on the classical inverse scattering transformation. Section 2 describes numerical results with respect to two types of forcing functions, i.e., a trapezoidal and a Gaussian shape with various parameters for the time variation of the dependent variable at the boundary, which is taken to be the origin. In the same spirit as in Kaup et al. [5], we present an approximate model for this problem based on the inverse scattering formalism. This approach is carried out in section 2.3, where the model is described in detail and some comparison with the numerical experiments is presented. Rather surprisingly, it is found that this new model is able to provide good qualitative agreement and to exhibit some quantitative predictions.

2.2 Numerical Results

The KdV equation,

$$u_t + uu_x + \epsilon^2 u_{xxx} = 0, \quad \epsilon = 0.022, \tag{2.1}$$

is studied in the semi-infinite domain $x \geq 0$, for $t \geq 0$. For given initial condition $u(x,0)$ and boundary conditions $u(0,t) = f(t)$ and $u(x,t)$, $u_x(x,t) \to 0$ as $x \to +\infty$, the solution is uniquely specified. This problem was examined by Chu et al.[2] and the equation is expressed in this particular form for ease of comparison with their results. For application to water waves this equation can be regarded in non-dimensional form with length and time scaled to water depth h and $\sqrt{h/g}$, g being the gravitational

constant, respectively. The numerical scheme adopted here is the same as in Zabuski and Kruskal [1], i.e. central difference in time and three-point average for suppressing nonlinear instability. We assume

$$u(x,0) = 0, \qquad x \geq 0, \tag{2.2}$$

and we first reproduce Chu's results, by taking for the only remaining boundary condition,

$$u(0,t) = \begin{cases} \frac{U_0}{\tau} t & t \in [0, \tau) \\ U_0 & t \in [\tau, t_0 - \tau) \\ \frac{U_0}{\tau} (t_0 - t) & t \in [t_0 - \tau, t_0) \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

A typical solution is shown in figure 2.1a, where snapshots of the field $u(x,t)$ were taken at two different times, namely at $t = 0.38$ and 0.75, for $U_0 = 2$, $\tau = 0.1$, $t_0 = 0.6$. In this case, positive waves are periodically generated at the boundary at $x = 0$, then propagating away from it and growing to an amplitude which is about twice the forcing U_0, the process continuing indefinetly in time, as long as the forcing is kept being applied. When the forcing is turned off at $t_0(= 0.6)$, the growth of the wave adjacent to the boundary stops, as can be clearly seen from the result for $t = 0.75$ as shown in figure 2.1a. With no further forcing after generation, these positive peaks would be expected to evolve like solitons for a free (unforced) KdV equation, and the wave amplitude would return to effectively zero at $x = 0$ and remain zero there as these waves propagate to the right, after having left the region of influence of the boundary condition. This feature of final recovery at the boundary

is illustrated by figure 2.2, where $t_0 = 0.25$ was chosen together with $U_0 = 2$ and $\tau = 0.1$ to produce only one soliton. Subsequently, as soon as the time reaches $t = 1$, this newly generated soliton has already evolved so perfectly that it becomes nearly identical with a free soliton of the same amplitude as shown in figure 2.2.

Taking the instant of first occurrence of each relative maximum that arises at the origin as the generation time, and following the propagation of these maxima with time, the curves of figure 2.1b,2.3b are obtained, from which the periodicity of soliton birth can be clearly seen.

The initial slope of the boundary forcing from 0 to U_0 (see (2.3) for $t \in [0, \tau)$) can be made steeper without introducing appreciable numerical errors, up to a value of τ of about 0.01, for $U_0 = 2$. The results for this case, $\tau \equiv 0.01$, $U_0 = 2$, $t_0 = \infty$, are shown in figure 2.3. As expected, the generation of the first soliton occurs sooner and the period of generation diminishes slightly, as compared with the case of $\tau = 0.1$ presented in figure 2.1. As shown in figure 2.3c, the soliton speed closely approaches the value $\frac{2}{3}U_0$, which is the speed of a free solitary wave of the same amplitude.

In order to have a smoother start from zero, we also examine the case with a gaussian boundary function,

$$u(0,t) = U_0 \exp\left[-\tau (t - t_0)^2\right] \tag{2.4}$$

where τ and t_0 are now chosen in order to have $u(0,0)$ of the same order of the time step used in the numerical calculations (typically $O(10^{-4})$). With $t_0 = 0.4$, $\tau = 60$ so fixed, we vary the amplitude U_0, over the range from 1.5 to 4. The results for these cases are shown with $U_0 = 2.5$ in figure 2.4 and with $U_0 = 4$ in figure 2.5. The amplitudes eventually reached by these solitons long after their birth are different, but the soliton generation still occurs almost periodically in time. After the forcing is shut

off, or practically so for case (2.4), the peak amplitude reaches a steady value, as it should for the KdV solitons; a collection of data for the asymptotic soliton amplitudes is shown in figure 2.6 and figure 2.7 for forcing (2.3) and (2.4), respectively.

The numerical experiments can also be utilized to provide information about $u_x(0,t)$ and $u_{xx}(0,t)$, two quantities of crucial importance for the Inverse Scattering Transform (IST) of the KdV equation [4]. Figures 2.8 and 2.9 show that they oscillate about zero, periodically in the case of steady forcing after the transients, with an average very nearly zero.

2.3 Approximate solutions using IST

For the purely initial value problem in $(-\infty, +\infty)$, the KdV equation is known to be solvable in terms of the IST formalism [3],[4]. Specifically, the equation

$$q_t + 6qq_x + q_{xxx} = 0, \tag{2.5}$$

can be transformed into (2.1) by using

$$u \to \beta q, \quad x \to \gamma x; \qquad \gamma = \epsilon^{\frac{2}{3}}, \quad \beta = 6\gamma, \tag{2.6}$$

and (2.5) is the solvability condition for the following two systems of equations

$$\begin{aligned} v_{1x} + i\zeta v_1 &= q v_2 \\ v_{2x} - i\zeta v_2 &= -v_1 \end{aligned} \tag{2.7}$$

and

$$\begin{aligned} v_{1t} &= A v_1 + B v_2 \\ v_{2t} &= C v_1 - A v_2, \end{aligned} \tag{2.8}$$

where

$$\begin{aligned} A &= -4i\zeta^3 + 2i\zeta q - q_x = i\zeta C - q_x \\ B &= 4q\zeta^2 + 2i\zeta q_x - 2q^2 - q_{xx} = -qC + A_x \\ C &= -4\zeta^2 + 2q\,. \end{aligned} \tag{2.9}$$

Equation (2.7) can be considered as an eigenvalue problem in ζ for the eigenfunction $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, and (2.8) gives the time evolution for the eigenfunctions and scattering data that can be determined with given initial value of $q(x,0)$. As can be realized by solving for v_2 in (2.7), the operator for the KdV equation is actually self-adjoint, being equivalent to the Schroedinger operator with potential $q(x,t)$.

When the domain is the whole real line of x, one can exactly solve for the scattering data and determine how they evolve in time by examining the asymptotic behaviour of the eigenfunctions, which depends critically on the regularity conditions at infinity on q and its derivatives, q_x, q_{xx}. However, for the present quarter-plane problem, $0 \leq x < \infty$, $t \geq 0$, only one set of such rest conditions can be assumed for the boundary at $x = +\infty$, the behaviour of q_x, q_{xx} at $x = 0$ being a part of the solution to the problem. To see how the evolution of the scattering data would depend on these two quantities, let us define the Jost function [4] ψ for the problem (2.7) as the solution with the following asymptotic behaviour

$$\psi(x,t;\zeta) \underset{x\to+\infty}{\longrightarrow} \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i\zeta x}. \tag{2.10}$$

Assuming $q = 0$ for $x < 0$, the scattering coefficients $a(\zeta,t)$ and $\overline{b}(\zeta,t)$[3] are then given by

$$\psi(0,t;\zeta) = \begin{pmatrix} -2i\zeta\overline{b} \\ a - \overline{b} \end{pmatrix}, \tag{2.11}$$

and their time evolution is therefore known once the one for ψ is. Since the behaviour of ψ in time is determined by the equation

$$\psi_t = \left[N - (\lim_{x \to +\infty} A)\, I \right] \psi, \tag{2.12}$$

where

$$N \equiv \begin{pmatrix} A & B \\ C & -A \end{pmatrix} \tag{2.13}$$

and I is the identity matrix, one can see that the quantities $q(0,t)$, $q_x(0,t)$, $q_{xx}(0,t)$ in A, B, C are needed to find $\psi(x=0)$ at any time $t \geq 0$. As already pointed out, only q is given by the boundary condition at $x = 0$, so that the solution to the problem is not possible until one devises a way of extracting from it this necessary information regarding the values of $q_x(0,t)$ and $q_{xx}(0,t)$. The question of whether or not this can be done does not seem to have an answer yet [5], but one can hope to find a reasonably good approximation by assuming some form for these derivatives. We will take

$$\begin{aligned} q_x(0,t) &= 0 \\ q_{xx}(0,t) &= 0, \end{aligned} \tag{2.14}$$

as the first approximation which is motivated in part by the numerical simulations showing that these quantities have almost zero average. Even with such a crude approximation, we have not succeded in obtaining the solution, in closed form, of equation (2.12) for the forcing (2.4), so we further assume that $q(0,t)$ be given by

$$q(0,t) = \begin{cases} Q & t \in [0,1) \\ 0 & \text{otherwise} \end{cases} \tag{2.15}$$

for some real constant $Q > 0$.

For seeking the zeros of $a(\zeta,t)$ in the upper half ζ plane, i.e., discrete spectrum eigenvalues of (2.7), since these will correspond to solitons [3],[4], and due to the self-adjointness of (2.7), one can set

$$\zeta = ik, \quad k > 0. \tag{2.16}$$

Equations (2.11-2.15) then lead to

$$\begin{aligned} a(ik,t) &= \psi_2(0,t;ik) + \frac{1}{2k}\psi_1(0,t;ik) \\ &= e^{-4k^3t}\left[\cos\Omega t + \frac{4k^4 - Q^2}{k\Omega}\sin\Omega t\right], \end{aligned} \tag{2.17}$$

where

$$\Omega^2 = 4\left(2k^2+Q\right)^2\left(Q-k^2\right). \tag{2.18}$$

In order to have zeros for $a(ik,t)$ it is necessary to require that $k^2 < Q$. Since $a(ik,t) = a(ik,1)$ for $t > 1$, the final number of solitons can be found by solving the equation $a(ik,1) = 0$, which can be determined graphically, e.g., by rewriting it in the form

$$\begin{aligned} y &= \cot\left[2\left(2k^2+Q\right)\sqrt{Q-k^2}\right] \\ y &= \frac{Q-2k^2}{2k\sqrt{Q-k^2}.} \end{aligned} \tag{2.19}$$

The first curve of (2.19) has asymptotes at values of k^2 such that

$$\left(2k^2+Q\right)\sqrt{Q-k^2} = \frac{n\pi}{2}, \quad n = 0,1,\ldots N, \tag{2.20}$$

where, as shown by figure 2.10a, the finite number N of solutions to this equation depends on Q, and some of the $k_n^2 \to Q-$ as Q increases. The asymptotes of the first

curve will 'force' intersections with the second curve in (2.19) among them, as seen in figure 2.10b. Each eigenvalue k_n corresponds to a solitary wave of amplitude $2k_n^2$ [3],[6],

$$q(x,t) = 2k_n^2 \operatorname{sech}^2\left(k_n x - k_n x_0 + 4k_n^3 t\right), \tag{2.21}$$

so that a train of solitons is generated, their amplitudes progressively increasing to the upper limit of $2Q$. This is in good agreement with the numerical calculations for the boundary forcing (2.3) when $t_0 \to \infty$.

In order to compare the results of the present approximate theory with the numerical ones, an appropriate scale for the model has to be found. By requiring that the integral

$$\int_0^\infty q(0,t)\, dt\ , \tag{2.22}$$

be the same for the approximate model and the numerical experiments, one has

$$\begin{aligned} Q &= (\frac{1}{\beta}U_0)\ [t_0 - \tau] \\ &= 2.122 \cdot U_0\ [t_0 - \tau]\ , \end{aligned} \tag{2.23}$$

for the boundary condition (2.3), and

$$Q = 2.122 \cdot U_0 \sqrt{\frac{\pi}{\tau}}\ . \tag{2.24}$$

for (2.4). In order to take into account the different shapes of the forcing functions (2.3) and (2.4) with respect to (2.15), the time can be rescaled according to some typical time length, say the widths of the forcing functions at half height. This implies a rescaling of the velocities, which are proportional to k^2 for the model. Thus, one should have

$$k^2 \longrightarrow \frac{k^2}{t_0 - \tau} \tag{2.25}$$

for case (2.3) and

$$k^2 \longrightarrow \frac{k^2}{2\sqrt{\log 2/\tau}} \tag{2.26}$$

for case (2.4). A direct comparison between the approximate IST results and the numerical ones is provided by figure 2.11 and figure 2.12 for case (2.3) and (2.4) respectively, for the first three solitons. As it can be seen, the IST model approximates the numerical results within 10 - 20%, which is surprisingly good, as the drastic assumptions made might be expected to limit the model to being able to produce only qualitative agreement, at most. However, it should be noticed that the model seems unable to provide a good description for small amplitude solitons. From (2.19) and figure 2.10b one can see that as Q increases, there are intersections between the two curves for which the corresponding k^2 moves to zero, but there is no evidence from the numerical experiment of the corresponding solitons. Solitary waves of amplitude Q, i.e. $k^2 = \frac{Q}{2}$, appear whenever

$$Q = Q_n \equiv \left[\left(n + \frac{1}{2}\right)\frac{\pi}{2}\right]^{\frac{2}{3}}, \quad n = 0, 1, \ldots . \tag{2.27}$$

and these are taken as the threshold of soliton creation for the model in figure 2.11 and 2.12. Further investigations are being conducted to determine whether the model can be improved on this.

References

[1] Zabusky, N.J. & Kruskal, M.D. 1965 Interaction of 'solitons' in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**, 240-243.

[2] Chu, C.K., Xiang, L.W. & Baransky, Y. 1983 Solitary waves induced by boundary motion. *Comm. Pure and Appl. Math.* **36**, 495-504.

[3] Gardner, C.S., Greene, J.M., Kruskal, M.D. & Miura, R.M. 1967 Method for solving the KdV equation. *Phys. Rev. Letters* **19**, 1095-1097.

[4] Ablowitz, M.J., Kaup, D.J., Newell, A.C. & Segur, H. 1974 The inverse scattering transform-Fourier analysis for nonlinear problems. *Studies in Appl. Math.* **53**, 249-315.

[5] Kaup, D.J. & Hansen, P.J. 1986 The forced nonlinear Schroedinger equation. *Physica* **18D**, 77-84.

[6] Whitham, G.B. 1974 *Linear and Nonlinear Waves.* Wiley Interscience, New York.

[7] Wu, T.Y. & Wu, D.M. 1982 Three-dimensional nonlinear long waves due to moving surface pressure. *Proc. 14th Symp. on Naval Hydrodynamics, Washington D.C.*, 103-125.

[8] Korteweg, D.J. & de Vries, G. 1895 On the change of form of long waves advancing in a rectangular channel, and on a new type of long stationary waves. *Phil. Mag.* **39**, 422-443.

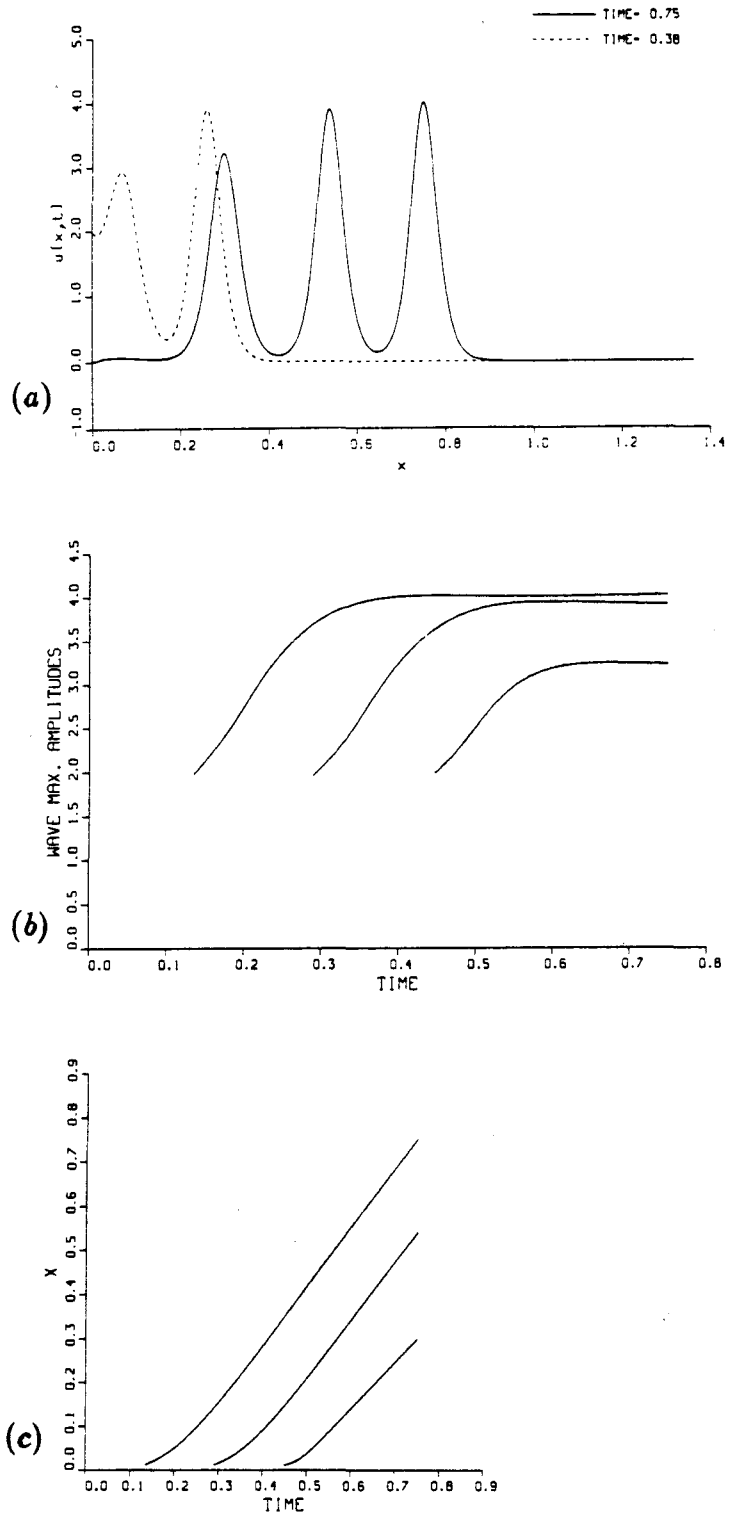

Figure 2.1. a) Solution for $u(x,t)$ at time $t = 0.38$ (dashed line) and $t = 0.75$ (solid), taking $U_0 = 2$, $\tau = 0.1$, $t_0 = 0.6$ in (3). Three solitons are generated, two reaching amplitude $\simeq 2U_0$ before turning the forcing off. b) Evolution of the relative maxima of u with time. First soliton birth occurs at $t \simeq 0.132$, period of generation (average) $T \simeq 0.156$. An interaction between the first and second soliton can be noticed around time 0.55. c) Trajectory of the maxima of u vs. time.

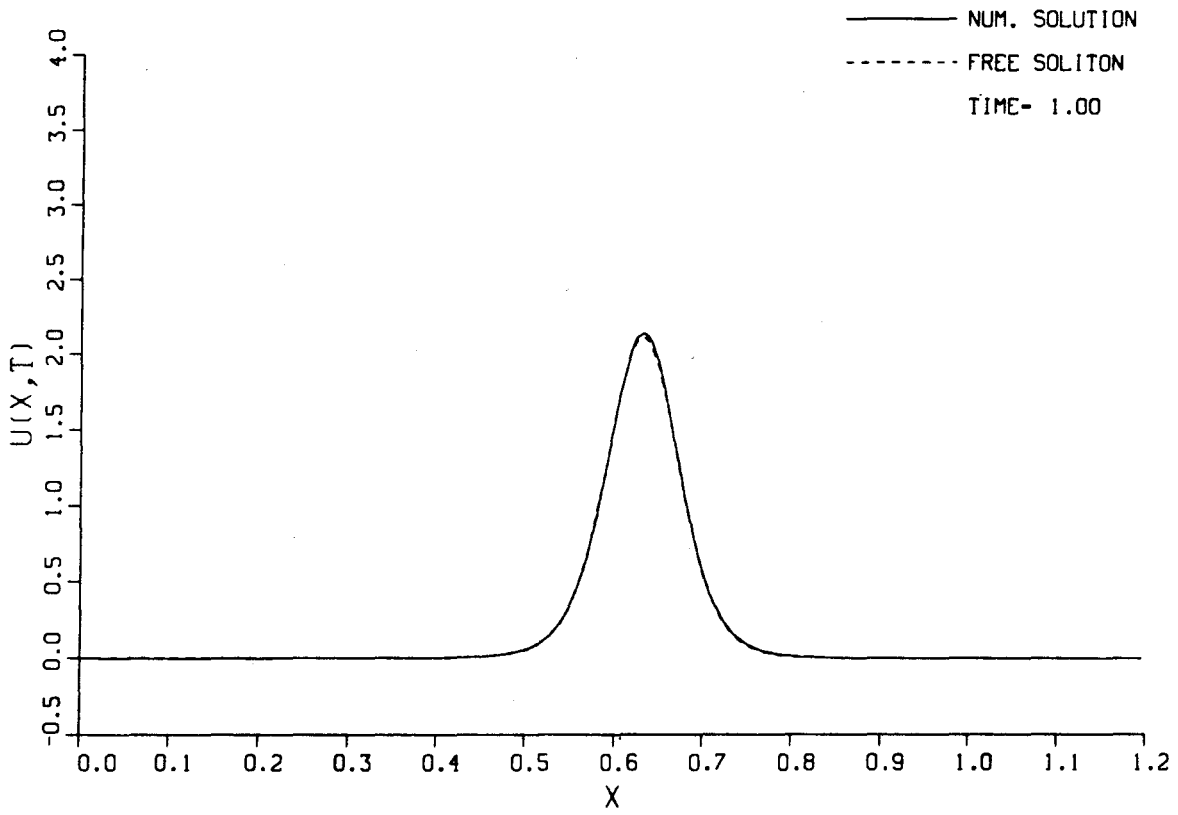

Figure 2.2. Comparison between the numerical solution at $t = 1$, $U_0 = 2$, $\tau = 0.1$, $t_0 = 0.25$ and the solitary wave solution $12\alpha^2\epsilon^2 \mathrm{sech}^2[\alpha(x - x_0)]$, $\alpha^2\epsilon^2 = \frac{1}{12} \times$ (amplitude), $x_0 \simeq 0.63$. The soliton amplitude is 2.18.

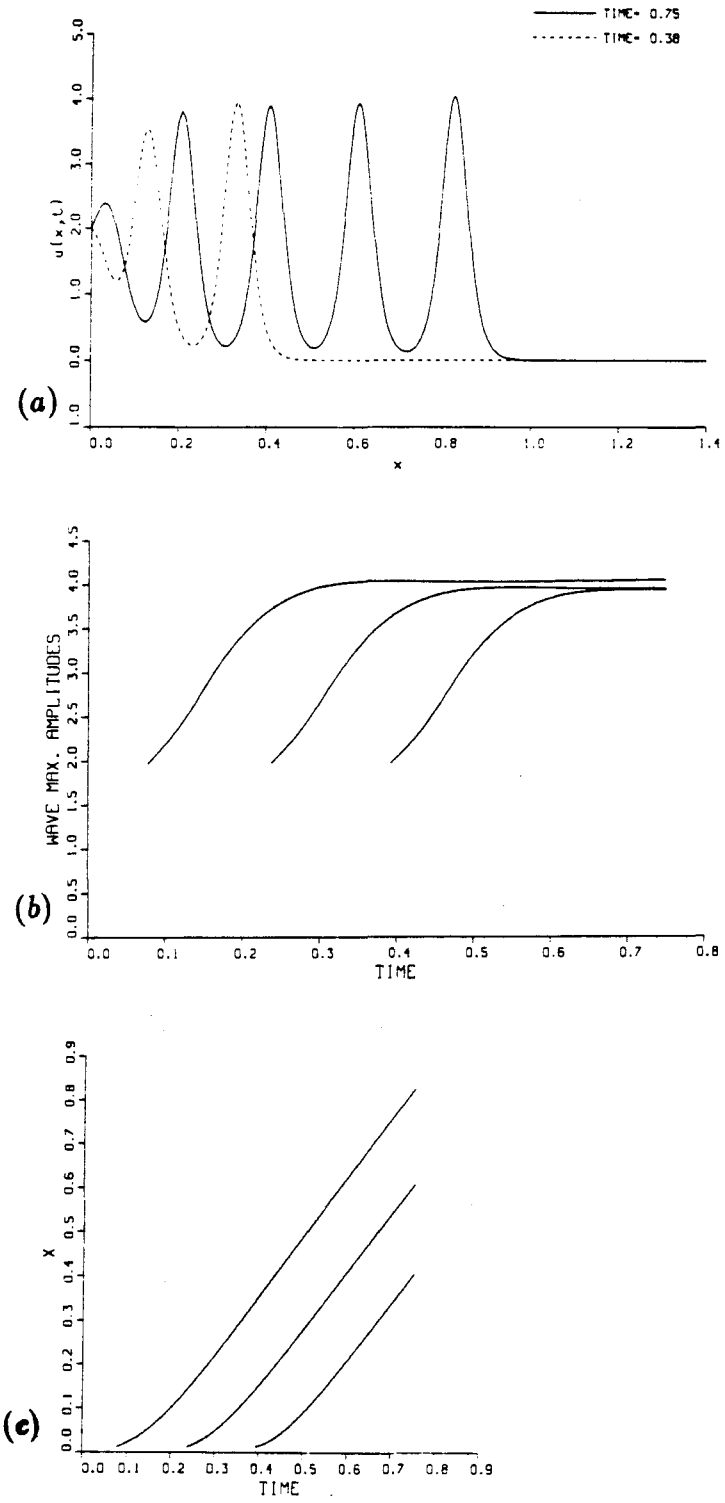

Figure 2.3. a) Solution for $u(x,t)$ at time $t = 0.38$ (dashed line) and $t = 0.75$ (solid), with $U_0 = 2$, $\tau = 0.01$, $t_0 = \infty$. b) Evolution of the first three relative maxima of u with time. The first soliton birth occurs at $t \simeq 0.075$, the period of generation (average) being $T \simeq 0.15$. c) Position of the maxima of u vs. time. The slope of these curves after the initial transient is 1.33 ± 0.05, with reference to the speed of a free solitary wave solution to (2.1), i.e. $\frac{4}{3}\times$ (amplitude) .

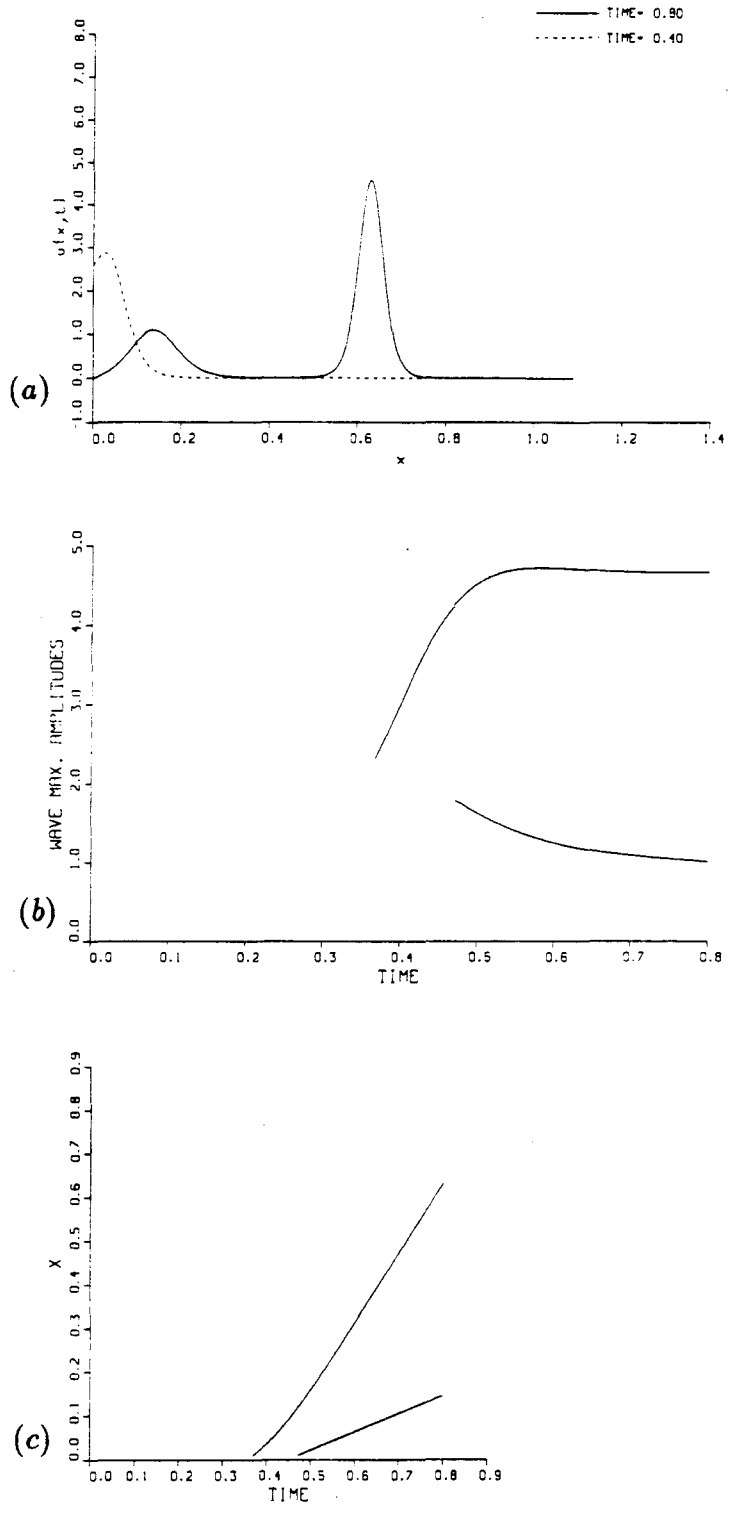

Figure 2.4. a) Solution for $u(x,t)$ at time $t = 0.4$ (dashed line) and $t = 0.8$ (solid) for the forcing (2.4), with $U_0 = 2.5$, $\tau = 60$, $t_0 = 0.4$. b) Evolution of the first two relative maxima of u with time. c) Trajectory of the maxima of u vs. time.

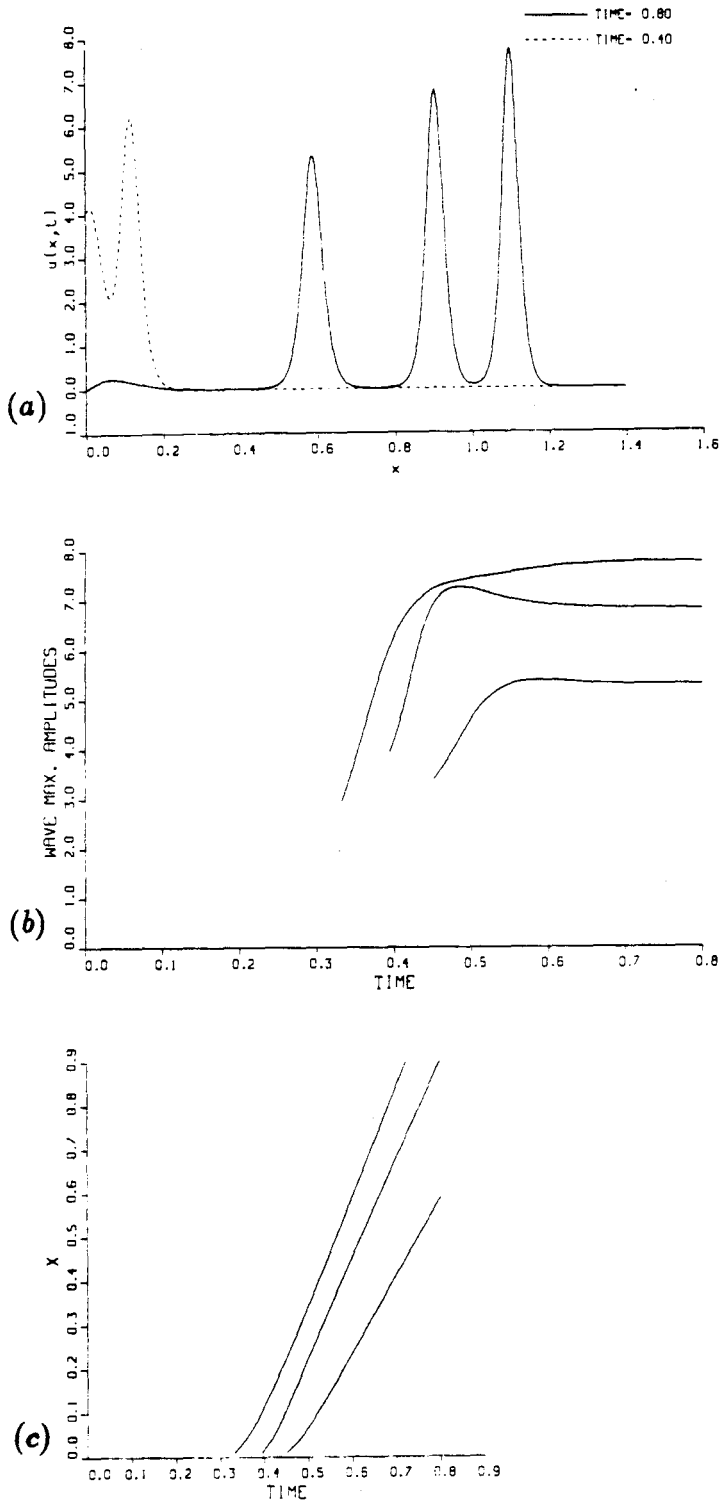

Figure 2.5. a) Solution for $u(x,t)$ at time $t = 0.4$ (dashed line) and $t = 0.8$ (solid) for the forcing (2.4), with $U_0 = 4$, $\tau = 60$, $t_0 = 0.4$. b) Evolution of the first three relative maxima of u with time. The interaction between the first two solitons at time $\simeq 0.5$ can be clearly seen. c) Trajectory of the maxima of u vs. time.

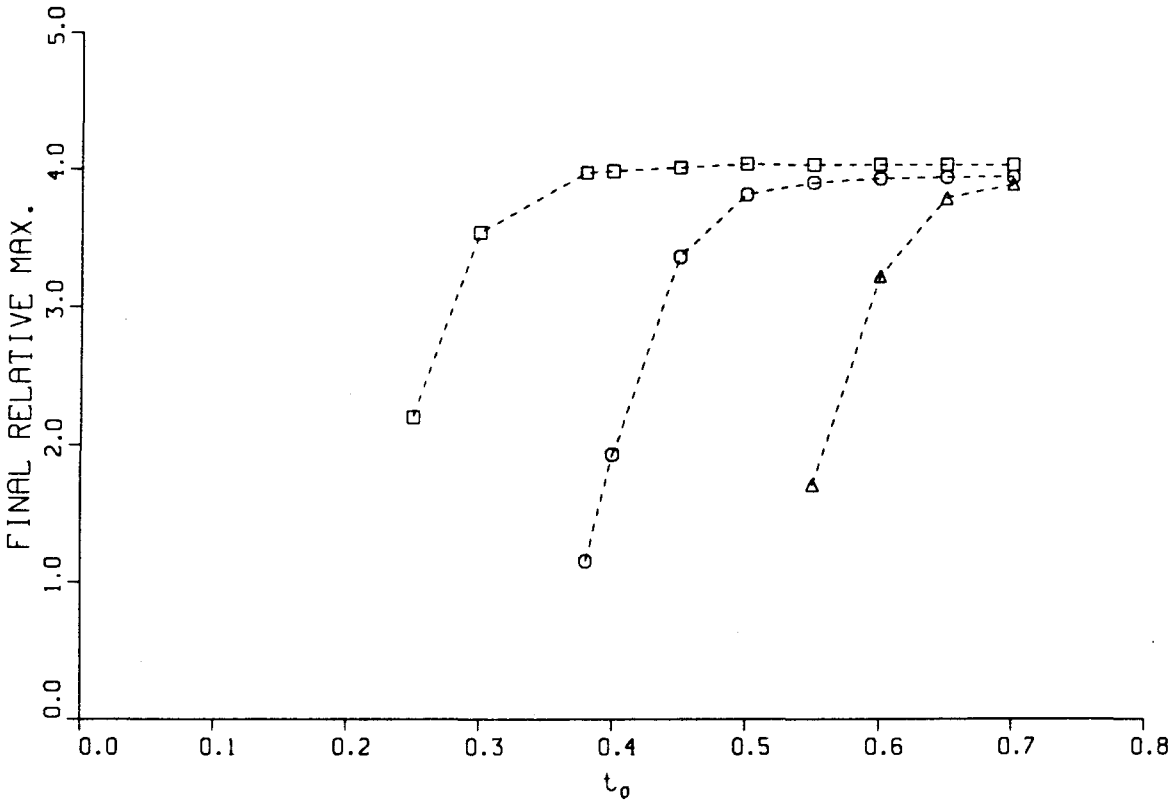

Figure 2.6. Final amplitudes for the first three solitons generated by the box-shaped forcing (2.3), with $U_0 = 2$, $\tau = 0.1$ and t_0 varying from 0.25 to 0.7 .

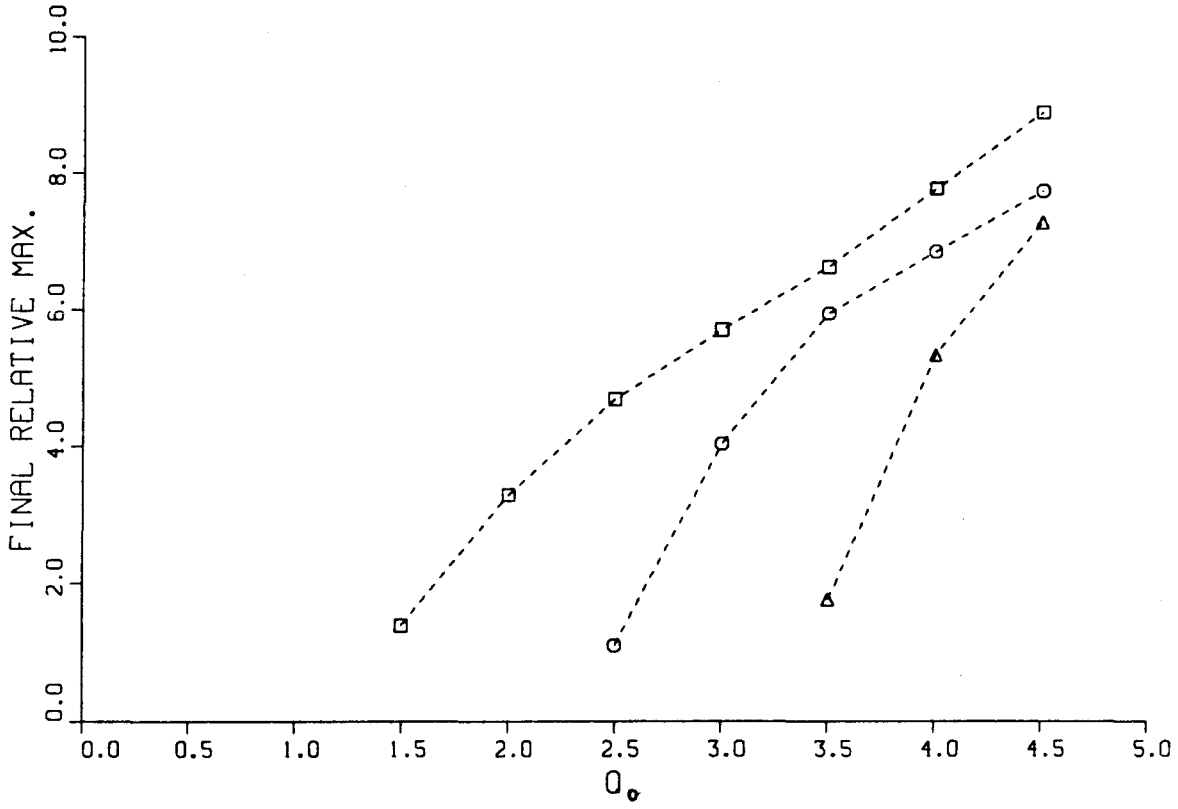

Figure 2.7. Final amplitudes for the first three solitons generated by the gaussian boundary forcing (2.4), with $t_0 = 0.4$, $\tau = 60$ and U_0 ($U_0 \equiv Q_0$) varying from 1.5 to 4 .

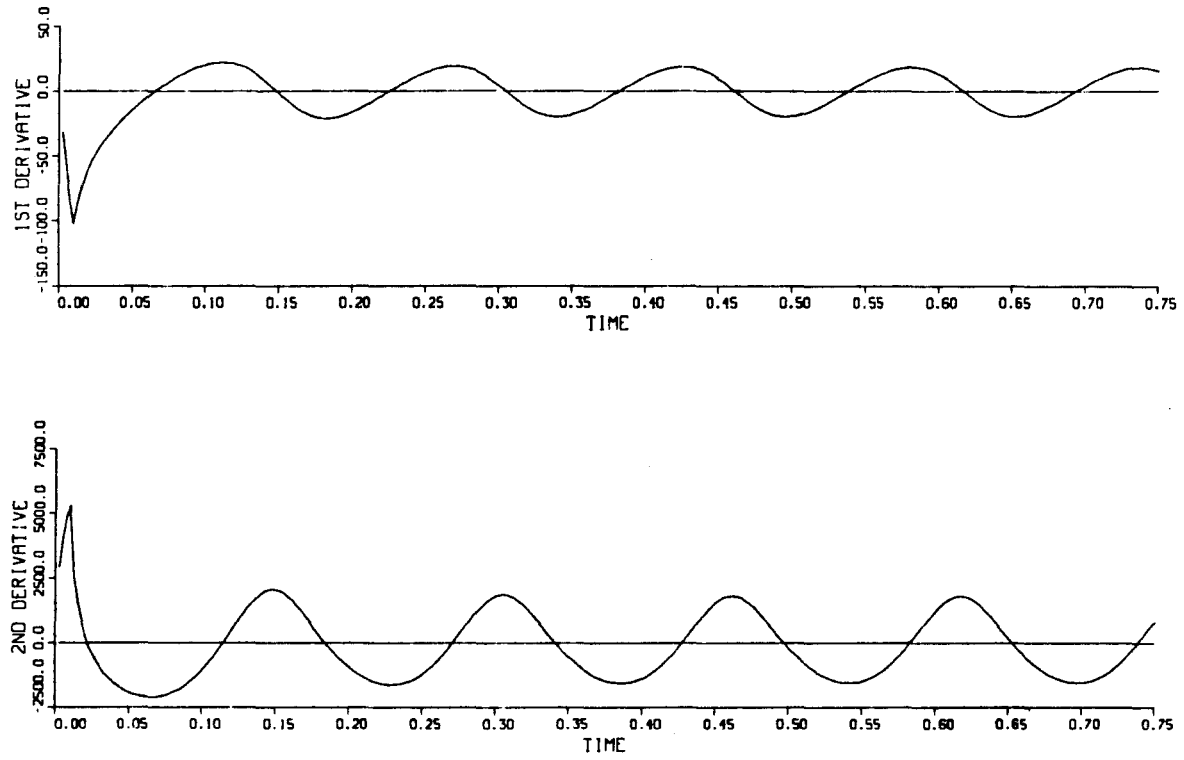

Figure 2.8. Evolution of $u_x(0,t)$ and $u_{xx}(0,t)$ in time for the forcing (2.3), with $U_0 = 2$, $\tau = 0.01$, $t_0 = \infty$. After an initial transient, they are almost perfectly periodic, with an average period of 0.15.

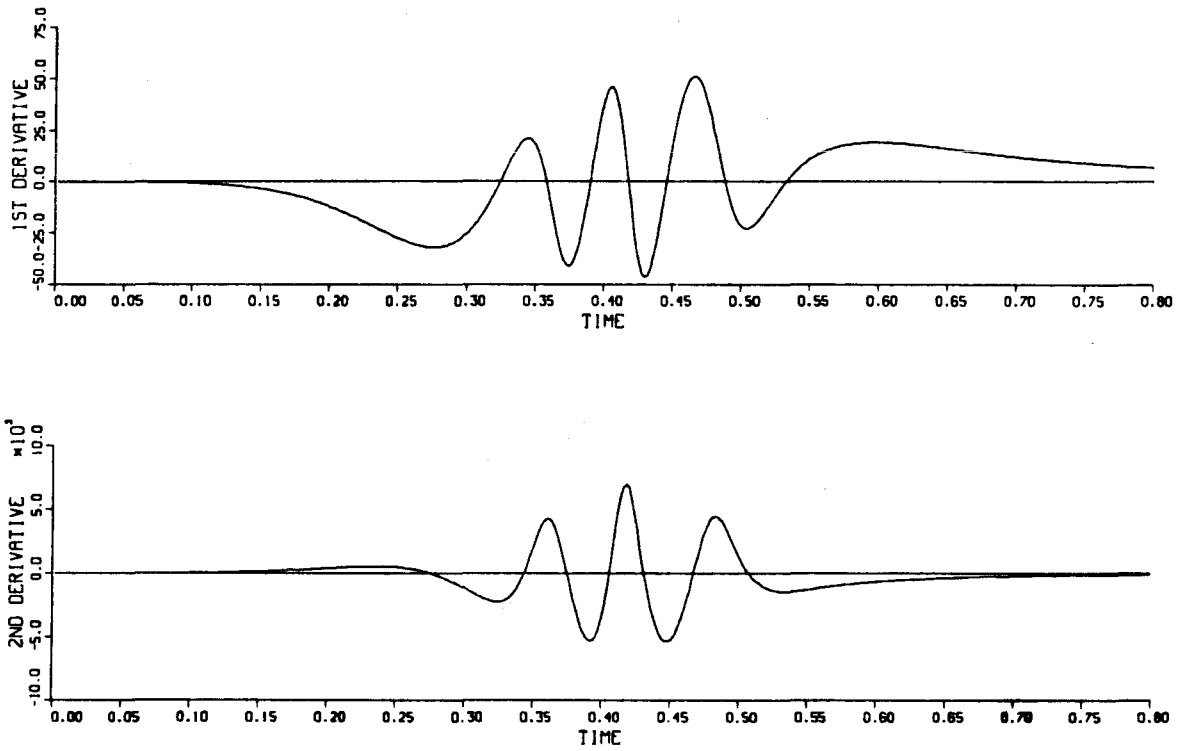

Figure 2.9. Evolution of $u_x(0,t)$ and $u_{xx}(0,t)$ in time for the forcing (2.4), with $U_0 = 4$, $\tau = 60$, $t_0 = 0.4$. The first curve is very closely skew-symmetric in t, and both have average close to zero.

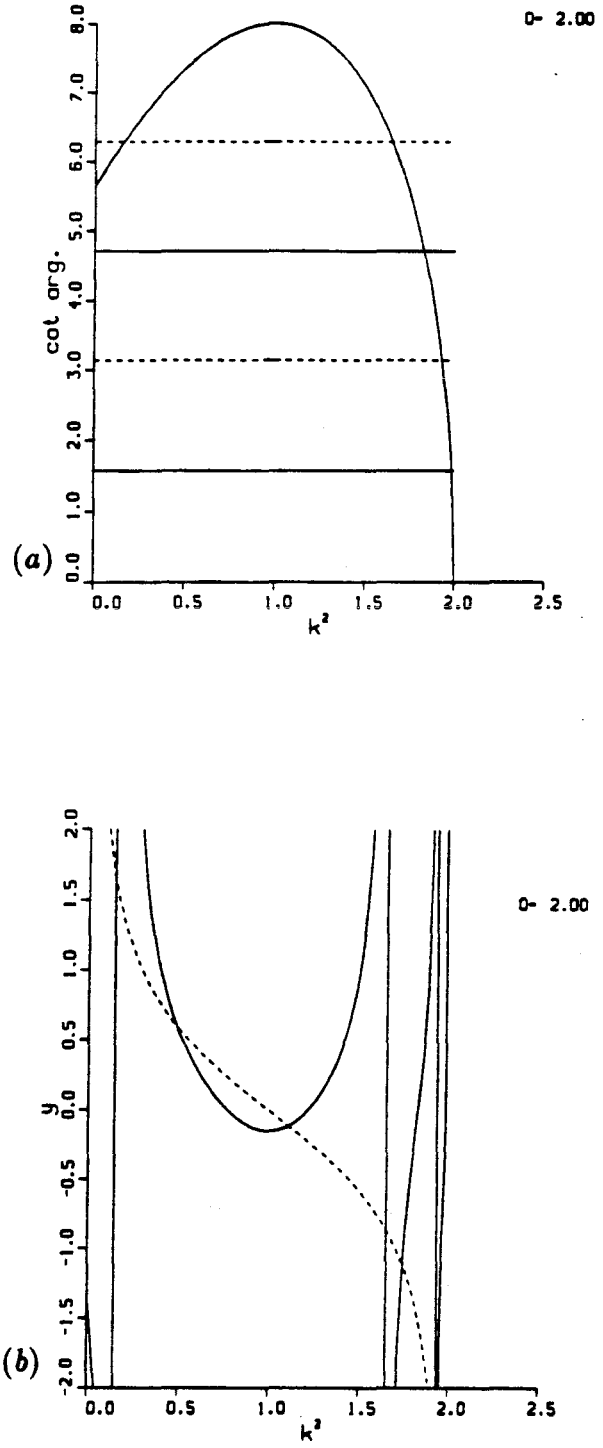

Figure 2.10. a) Argument of the first function in (2.19) vs. k^2. The horizontal lines correspond to the first two asymptotes (dashed) and zeros (solid) of the cotangent. The zeros of $a(ik, 1)$ lie between the intersections of these lines with the curve. b) Graphic solution of system (2.19), for $Q = 2$. The zero at $k^2 \cong 0.5$ moves to the origin as Q increases, and the relative minimum at $\cong 1.2$ plunges to $-\infty$ to create a new asymptote at $Q/2$.

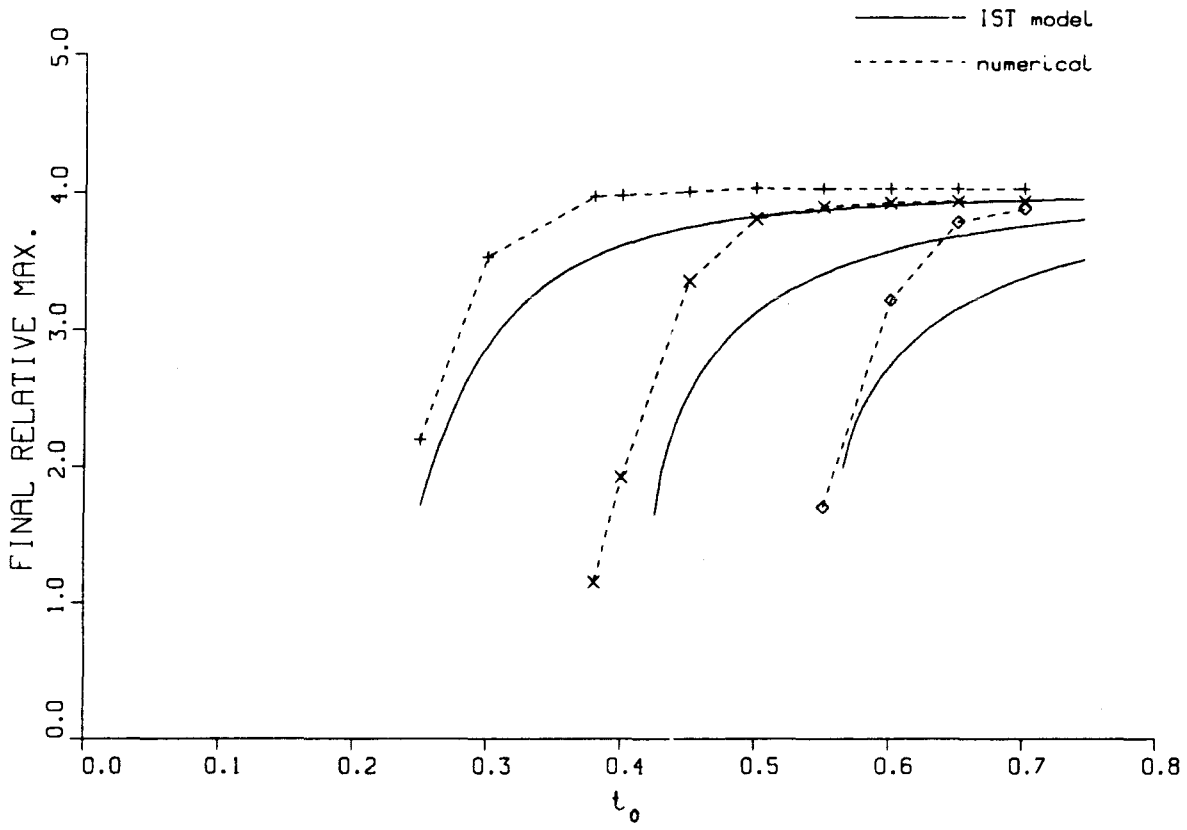

Figure 2.11. Direct comparison between the final amplitudes for the first three solitons obtained numerically and the ones predicted by the approximate IST model, for the box-shaped forcing function (2.3).

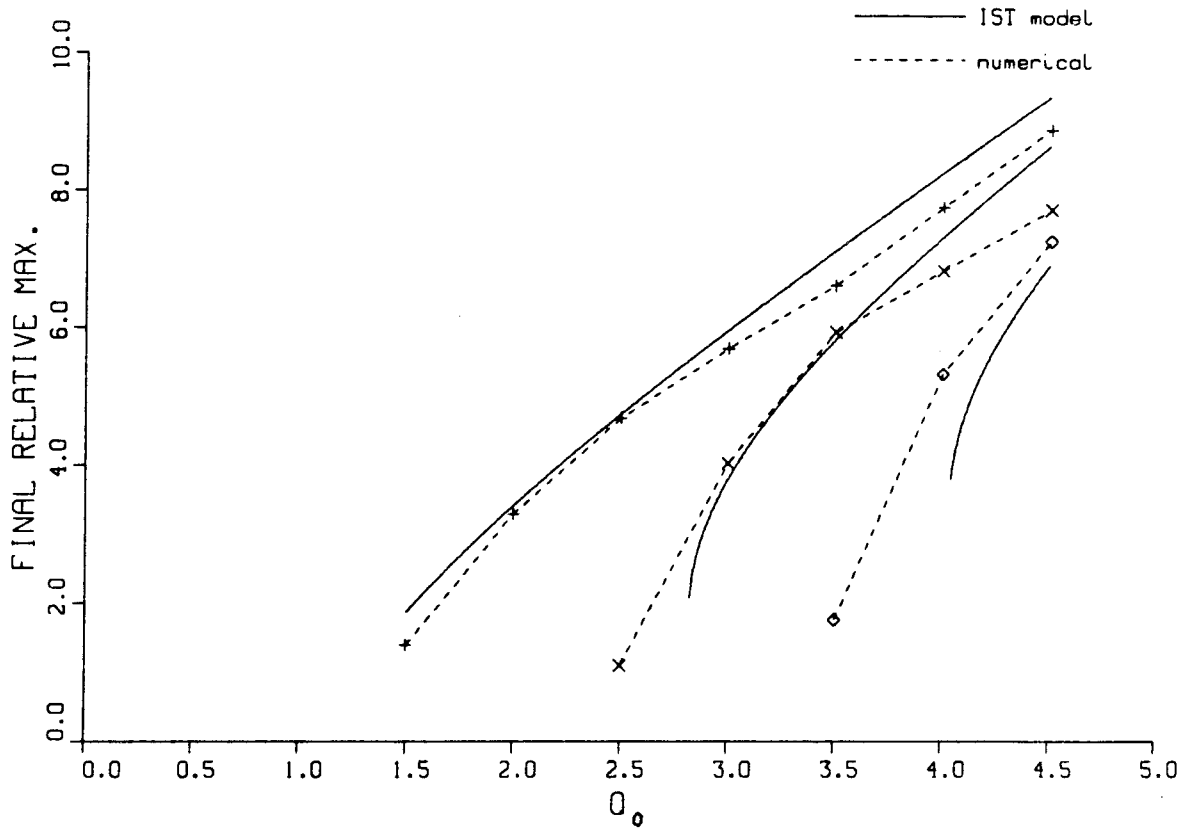

Figure 2.12. Direct comparison between the final amplitudes for the first three solitons obtained numerically and the ones predicted by the approximate IST model, for the gaussian forcing function (2.4) ($Q_0 \equiv U_0$).

Part II

CHAOTIC ADVECTION IN A RAYLEIGH-BÉNARD FLOW

Chapter 3

The Spreading of Passive Tracer in Chaotic Rayleigh-Bénard Flows

3.1 Introduction

In many fluid flows, knowledge of the velocity field can not be regarded as the solution. In physical applications, one is often interested in how quantities like mass or heat are transported in the flow. With good approximation, these quantities can often be described as moving with the fluid, i.e., their transport is essentially governed by convective processes. Therefore, an understanding of the motion of fluid particles becomes of fundamental importance. Unfortunately, this would imply going one extra step in integrating the Navier-Stokes equations, the well known difficulties usually involved with the first step of finding the velocity field notwithstanding. In fact, it is well known that even when the velocity field is explicitly determined and has a very simple form, the individual fluid elements can have an extremely rich and complicated dynamics[1].

It is in this context that the recent progress in dynamical system theory provides for the first time some analytical tool that can offer a description of this motion. By assuming the Lagrangian point of view, these techniques fit naturally with the exper-

imental ones, where usually the method of choice for studying the particle motion is to follow the evolution of a passive tracer.

The present study is meant to be an example in this direction. However, we try not to limit ourselves to the level of qualitative description, such as merely proving that the motion of a class of fluid particles can be chaotic. We strive throughout this study to compute quantities of direct physical interest. Our motivation comes from the recent series of experimental investigations carried out on transport in the context of the Rayleigh-Bénard convection[4],[5],[7],[8]. By exhibiting a wide variety of behaviours, ranging from stationary to travelling waves to fully turbulent flow, this system seems to offer an ideal trial ground of increasing (and controllable) complexity for experimental observation and theoretical testing. The wealth of information in the literature about this problem indeed seems to show this. However, for the particular problem of mass transport in this system very little has been done from the theoretical point of view, and what has been done is mainly limited to stationary flows. In this case the fluid particle trajectories coincide with the streamlines and an Eulerian approach is possible. In particular, this approach shows that if the flow is cellular, advection can only be responsible for the intra-cellular transport of a tracer, while molecular diffusivity is the only agent governing transport across the cellular boundaries [13],[14]. This fact is reflected in the large time asymptotics of the tracer concentration, which is basically given by a diffusion equation with an enhanced diffusion coefficient.

The experimental work of Solomon and Gollub[4] focuses on the the transition from stationary to time dependent flow, while still remaining in cellular regime[2] where the streamlines are closed and hence no net mass transport is possible. The

enhancement on the spreading of tracer they observe in the experiment is therefore entirely due to the radically different behaviour exhibited by the particle paths in the time dependent case. No approach based solely on the knowledge of the velocity field would be able to extract information about transport in this case.

To model the flow after the onset of the time dependent instability, we use the stream function introduced by Solomon and Gollub[4] and based on the analysis of Busse[2],[3]. The velocity field generated by this stream function basically describes two-dimensional convection rolls. The time dependence would correspond to the collective oscillation of the roll boundaries in the direction perpendicular to the roll axes, a phenomenon known as the "even" oscillatory instability. Whenever the velocity field is obtained from a stream function, the problem can be cast into the framework of the Hamiltonian formalism, with the Hamiltonian corresponding to the stream function. The physical space can then be interpreted as the phase space, and the transition from time independent to time dependent flow would correspond to the loss of integrability due to a non-autonomous component in the Hamiltonian. Achieving this point of view is more than just an intellectual exercise. Thanks to it, we are able to identify the relevant invariant structures present in the flow, and show how they effectively govern the motion of fluid particles. Exploiting this information, we derive several results not previously available.

By providing a precise definition of the time dependent equivalent of the rolls, we can identify the mechanism of inter-roll convective transport and derive formulae for explicitly computing the transport rates, in the spirit of Rom-Kedar et al.[11],[12] . Making use of the symmetries of the problem, we are able to predict that the lateral spreading of tracer along the rolls would follow a linear law in time, for as long as

the molecular diffusion effects can be neglected. As a by-product of this derivation, a lower bound on the rate of stretching of the interface between clear and dyed fluid can be obtained. In the limit of small lateral oscillation of the roll boundaries, we use the Melnikov technique to analytically estimate the amount of fluid exchanged between rolls during each period. This analysis confirms the findings of Solomon and Gollub of linear dependence on the strength of the perturbation and points out the strong (nonlinear) dependence on the ratio of the frequency of oscillation and the maximum speed of the fluid along the roll boundary. The latter observation needs to be investigated experimentally, since the existing data[4] do not seem to address this point.

We carry out detailed computations, based on the formalism we derive, for three examples, which are good representatives of the main effects induced by varying the parameters of the flow. By considering a particular way of obtaining a bounded phase space, and hence a bounded region of fluid flow, we also use these results to investigate the mixing properties for the portion of a roll that participates in the transport process. Here mixing is understood in the mathematical sense, where an explicit definition is possible[1]. In the more general context of transport in Hamiltonian systems, we show that the proposal[10] of modelling the mechanism of transport of phase points by a Markov chain is inappropriate, at least for these types of Rayleigh-Bénard flows.

Finally, we discuss the validity of a purely convective approach in the presence of molecular diffusivity, and introduce a simple criterion for the time scale in which the effects of diffusivity can be neglected. As an example, we carry out some numerical simulation of the time dependent flow with a term representing the Brownian motion

a tracer particle would undergo in the presence of molecular diffusivity.

The model for the stream function can satisfy both slip and non-slip conditions at the horizontal plates. Which type is actually enforced has no effect on the technical implementation of the numerical computations, and the results from global analysis are largely independent of the boundary conditions. However, the Melnikov analysis and the proof of existence of chaotic motion does require that some technical point be resolved in the presence of non-slip boundary conditions. We address this case with some detail in the appendices.

This work is organized as follows. In Section 3.2 we introduce the relevant definitions and concentrate on the analytical results. These include the formulae for computing the transport rates, the time scale of spreading of a tracer and interface stretching, the Melnikov estimates, the comparison between the time scales of molecular diffusivity and convective transport. In Section 3.3 we work out the specific examples, provide the comparisons with the results of the Markov model, introduce the compact phase space and check for its mixing properties. We conclude the section by reporting on the computation simulating molecular diffusivity. The two appendices discuss the use of symmetries in the computation of the tracer content in a given roll, and derive the Melnikov function for the case of non-slip boundary conditions.

3.2 The mathematical model and transport theory

As a model for the flow after the onset of the time dependent instability, we consider the following stream function[3],[4],

$$\begin{aligned} \psi(x,z,t) &= H_0(x,z) + \epsilon H_1(x,z,t) \\ &= \frac{A}{k}\sin(kx)\sin(\pi z) + A\epsilon f(t)\cos(kx)\sin(\pi z) \end{aligned} \tag{3.1}$$

which yields the velocity field

$$\begin{aligned} \dot{x} &= -\frac{\partial \psi}{\partial z} = -\frac{A\pi}{k}\cos(\pi z)\left[\sin(kx) + \epsilon k f(t)\cos(kx)\right] \\ \dot{z} &= \frac{\partial \psi}{\partial x} = A\sin(\pi z)\left[\cos(kx) - \epsilon k f(t)\sin(kx)\right] \end{aligned} \tag{3.2}$$

with $x \in \mathbb{R}$, $z \in [0,1]$ and $f(t)$ a function of time we will specify later. From the Lagrangian point of view, (3.2) describes the motion of "fluid particles" and hence, in particular, particles of a passive tracer in the fluid, corresponding to an assigned initial configuration. Here length measures have been nondimensionalized with respect to the distance between the top ($z = 1$) and bottom ($z = 0$) surfaces, k is the wave number $k = \frac{2\pi}{\lambda}$, λ being the period of the roll pattern, and A represents an estimate of the maximum vertical velocity. The amplitude of the perturbation ϵ is proportional to $(\mathcal{R} - \mathcal{R}_t)^{\frac{1}{2}}$ with $\mathcal{R}_t$ being the Rayleigh's number corresponding to the onset of the time dependent instability[6].

The unperturbed ($\epsilon = 0$) flow given by (3.2) corresponds to single mode, two dimensional convection with slip boundary conditions, and the perturbation introduced by the term in ϵ describes the (small) oscillation of the roll boundaries along the x (lateral) direction. Thus, the model refers to a simplified version of the "even" oscillatory instability[3],[2] by considering the motion of the roll boundaries as independent of the coordinate along the roll axes. This is a good approximation near the

onset of the time dependent instability[2], and is verified in practice by the experiment performed by Solomon and Gollub[4]. On the other hand, the assumption of stress free boundary conditions is almost never verified in practice, but it is certainly of theoretical interest since it allows closed form calculations (see § 3.2.5). Furthermore, the basic mechanism of fluid transport from roll to roll, in the time dependent case, relies on the existence of the structures we are going to describe in detail in the next sections. These structures are based on the invariance of certain curves (surfaces for the full 3D space) under the action of the flow, and these exist regardless of the type of boundary conditions, which have only mild quantitative effects. As fluid particles cannot cross these invariant curves, they act as a geometrical "template" through which their motion is forced. An understanding, from the global point of view, of the geometry of these curves is hence going to be the most important piece of information on which to build a theory of transport.

The bulk of this Section is organized as follow. We first identify the invariant structures (§3.2.1) we use in deriving formulae for computing the tracer content of each roll (§3.2.2). We derive the results about the time scale of tracer spreading and the lower bound on the interface stretching in §3.2.3. We briefly mention in which sense the orbits of fluid particles can be chaotic in §3.2.4. In §3.2.5 we use the Melnikov theory to evaluate the volume of fluid transported at each iteration across a roll boundary. We then look at the existence of invariant structures that could actually prevent mixing inside one roll in §3.2.6. In these regards, perhaps a more serious deficiency of the model is the fact that it does not describe a weak recirculating flow induced by the vertical boundaries, the existence of which can be shown experimentally in the time independent case[5]. However, as discussed in more

detail at the end of § 3.2.6, this 3D effect should only bear consequences for the transport inside a roll and does not affect the inter-roll transport mechanism. We conclude (§3.2.7) the Section with a discussion on how the molecular diffusivity could affect the results obtained, in anticipation of the numerical simulations reported in Section 3.3.

3.2.1 The basic structures governing roll to roll transport

First, we note that (3.2) and (3.2), as suggested by the notation used, can be interpreted as a Hamiltonian system, with an integrable component H_0 and a nonintegrable perturbation ϵH_1, formally equivalent to a two degree of freedom Hamiltonian system. The unperturbed (time independent) flow is characterized by the hyperbolic fixed points $\{p^+_{j,0}\}$, $\{p^-_{j,0}\}$ respectively along the top and bottom surfaces

$$p^{\pm}_{j,0} \equiv (x_{j,0}, z^{\pm}_j) \ , \quad \text{with} \quad x_{j,0} = \frac{j\pi}{k}, \quad z^-_j = 0 \text{ and } z^+_j = 1 \ , \quad j = 0, \pm 1, \ldots , \tag{3.3}$$

physically corresponding to the stagnation points in the flow, joined by the heteroclinic connections (stagnation streamlines) corresponding to the roll boundaries. Switching the perturbation on, this ordered structure is seemingly lost as the path lines of the fluid particles can now intersect themselves and no longer coincide with the (instantaneous) streamlines.

For periodic time-dependency, say $f(t) = \cos(\omega t)$, a well known and effective way of unveiling the structure still possessed by the flow is to study the Poincaré map of (3.2) $F : \mathbf{R} \times [0,1] \to \mathbf{R} \times [0,1]$

$$F_{t_0} \begin{pmatrix} x(t_0) \\ z(t_0) \end{pmatrix} \equiv \begin{pmatrix} x(t_0+T) \\ z(t_0+T) \end{pmatrix}, \quad t_0 \in (0,T], \quad T = \frac{2\pi}{\omega}. \tag{3.4}$$

This map corresponds to taking stroboscopic snapshot of the fluid, and in mathematical terms is an area and orientation preserving two dimensional diffeomorphism. For

the Poincaré map, the hyperbolic fixed points present in the time independent flow persist with the same stability type, i.e., they become periodic points for the time dependent flow. We will denote the map fixed points by $p^{\pm}_{j,\epsilon}(t_0)$. The notation used stresses the fact that their x coordinate depends on ϵ and on the section t_0, however, for sake of tidiness, we will drop the subscript ϵ and t_0 dependence from now on, simply referring to $p^{\pm}_{j,\epsilon}(t_0)$ as $p^{\pm}_j$ and restoring the full notation whenever ambiguities can arise.

The heteroclinic orbits at the top and bottom surfaces also persist for the map, since the lines $z = 0$, $z = 1$ are left invariant by the perturbation H_1. However, the stable and unstable manifolds, which in the unperturbed case coincide to form the vertical heteroclinic connections between $p^{+}_{j,0}(t_0)$ and $p^{-}_{j,0}(t_0)$, break apart under the action of the perturbation and intersect transversely. We will denote the perturbed manifolds by $W^s_\epsilon(p^{\pm}_{j,\epsilon}(t_0))$ and $W^u_\epsilon(p^{\pm}_{j,\epsilon}(t_0))$ for the stable and unstable one respectively (see figure 3.1 and §3.2.5). Due to the invariance of the manifolds under the map, a heteroclinic point, i.e., a point that belongs to both $W^s_\epsilon(p^{\pm}_j)$ and $W^u_\epsilon(p^{\mp}_j)$, must remain on both manifolds under the action of F_{t_0} and $F^{-1}_{t_0}$ (the inverse of F_{t_0}), resulting in the wild oscillations, or "tangle", of one manifold about the other when approaching a fixed point (see figure 3.2). Intersections are not limited to manifolds of the same j-couple, however only "hybrid," i.e., stable with unstable, manifold crossings are allowed.

In the following, we will only be concerned with a particular class of heteroclinic points, which we are now going to introduce. For definiteness, let us focus on the case where the unperturbed flow has a stagnation streamline oriented from $p^{-}_{j,0}$ to $p^{+}_{j,0}$, which occurs for instance at $x = 0$. A primary intersection point (pip) of $W^s_\epsilon(p^{+}_j)$ and

$W^u_\epsilon(p^-_j)$ is then defined as a heteroclinic point q such that the segments $S_j[q, p^+_j]$ and $U_j[q, p^-_j]$ of stable and unstable manifold connecting q with p^+_j and p^-_j, respectively, do not intersect at any other heteroclinic point other than q itself (see figure 3.3).

Using pieces of manifold and the pip's as building blocks, we can now define the time dependent analogue of the roll boundaries. Although a similar construction can be carried out for any Poincaré section, i.e., any choice of t_0 in (3.4), the symmetries possessed by the Poincaré map for a particular t_0 may suggest a natural definition. For instance, when $t_0 = 0$ and $f(t) = \cos(\omega t)$ it is easy to show that the flow given by (3.2) is invariant under

$$t \to -t, \quad x \to x, \quad z \to 1 - z \ , \tag{3.5}$$

and therefore $W^s_\epsilon(p^+_0)$ is mapped into $W^u_\epsilon(p^-_0)$ by $z \to 1 - z$. An immediate consequence of this symmetry for this Poincaré section is that the manifold will always have a pip, say c_j, along the line $z = \frac{1}{2}$. The periodicity of the vector field also implies, for any cross section t_0, the invariances

$$t \to -t, \quad x \to x + \frac{\pi}{k}, \quad z \to z \ , \tag{3.6}$$

and

$$t \to t, \quad x \to x + \frac{2\pi j}{k}, \quad z \to z \ , \qquad j = \pm 0, \pm 1, \dots \ , \tag{3.7}$$

Using (3.6) for the Poincaré section $t_0 = 0$, one can see that for the couple p^-_1 to p^+_1 the vertical heteroclinic connection breaks up into the same structure as the one at $x = 0$, the only difference being that the manifold stability is now reversed.

We therefore define R_1 as the region bounded by the heteroclinic orbits connecting p^-_1 to p^-_0 and p^+_0 to p^-_1, the segments $S_0[c_0, p^+_0]$, $S_1[c_1, p^-_1]$ and $U_0[p^-_0, c_0]$, $U_1[p^+_1, c_1]$ of, respectively, stable, $W^s_\epsilon(p^+_0)$, $W^s_\epsilon(p^-_1)$, and unstable, $W^u_\epsilon(p^+_0)$, $W^u_\epsilon(p^-_1)$ manifolds

(see figure 3.4). Using the periodicity in x (3.7) one can similarly define the regions R_j for all j, thus tessellating the whole strip $\mathbf{R} \times [0,1]$. We note that for $\epsilon = 0$ these regions go over to the ones representing time independent rolls.

In order to see how the map can describe the transport of fluid from one (time dependent) roll to another we need to introduce one more object, the turnstile lobe. If the segments $S[q_0, q_1]$ and $U[q_0, q_1]$ of stable and unstable manifold between two pip's q_0, q_1 do not contain another pip, we refer to the region enclosed by these segments as a lobe. Under the action of the map F_{t_0}, pip's are mapped into pip's[9]. Without loss of generality, we assume that between a pip q_0 and its image under the map, $F_{t_0}(q_0)$, only one pip can exist, one being the minimum required for preserving orientation. We will then have two lobes for each couple q_0, $F_{t_0}(q_0)$. In addition, we choose $t_0 = 0$, so that we can drop the subscript t_0 from now on, and $F(q_0) = c_j$ as in figure 3.5, where $j = 0$. We label the lobe lying in region R_1 as $L_{1,0}$ and the lobe lying in region R_0 as $L_{0,1}$, the meaning of this notation being clear when one considers that $L_{1,0}$ is mapped from R_1 to R_0 under F and vice versa for $L_{0,1}$, see figure 3.5. Thus, the fluid transported across the boundary between R_0 and R_1 in one period is precisely the one contained into $L_{1,0}$ or $L_{0,1}$. This pair of lobes has come to be known as "turnstile" in the literature[10]. We note that, because of the symmetries (3.5) and (3.6) and area preservation, all of the turnstile lobes will have equal areas. This is a consequence of the fact that the flow given by (3.2) does not have a preferred convective direction, or the average mass transport is zero. Under the action of the map F, the turnstile lobes are then stretched and folded following the tangle of the unstable manifold $W^u_\epsilon(p_j^\pm)$ with the stable one, $W^s_\epsilon(p_j^\mp)$ (see figure 3.2). Because of orientation preservation, the fluid that crosses the boundary between two regions at some iteration n must

be contained in the $(n-1)$-th pre-image of the turnstile lobes. This is the crucial observation for constructing a theory of transport[12] based on the dynamics of the lobes.

3.2.2 The spreading of tracer initially contained in one roll

We illustrate how the definitions given in the previous Section can be used by considering the following problem. Suppose we have some passive tracer in the fluid that initially is all contained in one roll, say the one corresponding to region R_1, with uniform concentration C. Of course, due to the symmetries (3.5-3.7), as long as the tracer is initially uniformly concentrated in one roll, we can always label the regions in order to reduce to this configuration. We would then like to know what is the average tracer concentration in any region R_j at any later time $t = n$ corresponding to the n-th cycle. Neglecting any molecular diffusivity for the moment, this is equivalent to asking how much fluid which is initially in R_1 is contained in R_j at time n. We denote this quantity by $T_j(n)$ and refer to the fluid in R_1 at $t = 0$ (or $n = 0$) as the R_1-species. For definiteness, let us assume that the lobes forming a turnstile are entirely contained in neighbouring regions, which for small enough ϵ and large enough ω can always be shown to be the case (see §3.2.5). Typically, we find that this is true for ω as small as 0.2 (at $A = 0.1$), and expect it to be true for ϵ up to $O(1)$. Thus, the only way fluid can enter region R_j is to be contained in R_{j-1} or R_{j+1} at the previous cycle. It should be noticed that in general the mass transport will depend more on frequency than amplitude, as low frequency means that the fluid particles have have longer times to wander over the phase space. Having stated the problem, we proceed to construct the formulae necessary for its solution. We will show that knowledge of

the evolution of one turnstile lobe only is all that is required.

From the preceding remarks, it should be clear that the R_1-species entering R_j, $j \neq 1$, at iteration n can only be contained in the intersections of $L_{j-1,j}$, $L_{j+1,j}$ with the lobes that have left R_1, i.e., $F^k L_{1,0}$ and $F^k L_{1,2}$, $k = 1, 2 \ldots, n-1$. However, not all of the iterates $F^k L_{1,0}$ and $F^k L_{1,2}$ will consist purely of R_1-species. As time increases, iterates of the lobes $L_{0,1}$ and $L_{2,1}$ will transport fluid coming from R_0 and R_2 into region R_1, from which it can eventually escape, the only way it is allowed to do so being through the lobes $L_{1,0}$, $L_{1,2}$. Hence, this fluid has to be subtracted from $F^k L_{1,0}$ and $F^k L_{1,2}$ in order to get the net content of R_1-species in these lobes. Denoting by $\mu(L^1_{j-1,j}(n))$ the amount of R_1-species in lobe $L_{j-1,j}$ entering R_j at cycle n, the above considerations can be summarized into the following formula[12]

$$\mu(L^1_{j-1,j}(n)) = \delta_{j,2}\mu\left(L_{1,2}\right) + \sum_{k=1}^{n-1} \sum_{s=0,2} \left[\mu\left(L_{j-1,j} \bigcap F^k L_{1,s}\right) - \mu\left(L_{j-1,j} \bigcap F^k L_{s,1}\right)\right] \tag{3.8}$$

Here $\mu(\mathcal{M})$ denotes the area measure of a subset $\mathcal{M} \subset \mathbb{R}^2$, and $\delta_{j,k}$ is the Kronecker delta. A similar relation for $\mu(L^1_{j+1,j}(n))$ can be obtained by replacing the first term in (3.8) with $\delta_{j,0}\mu\left(L_{1,0}\right)$ and $L_{j-1,j}$ with $L_{j+1,j}$.

The same arguments apply of course to R_1-species leaving R_j at iteration n, so that the variation $T_j(n) - T_j(n-1)$ of R_1-species in region R_j at cycle n can be written as

$$\begin{aligned} T_j(n) - T_j(n-1) = & \\ = & \sum_{r=0,2} \left[\mu\left(L^1_{j-1+r,j}(n)\right) - \mu\left(L^1_{j,j-1+r}(n)\right)\right] \\ = & \ \delta_{j,2}\mu\left(L_{1,2}\right) + \delta_{j,0}\mu\left(L_{1,0}\right) + \\ & \sum_{k=1}^{n-1} \Big\{ \sum_{r,s=0,2} \left[\mu\left(L_{j-1+r,j} \bigcap F^k L_{1,s}\right) - \mu\left(L_{j-1+r,j} \bigcap F^k L_{s,1}\right)\right] \end{aligned}$$

$$- \sum_{r,s=0,2} \left[\mu \left(L_{j,j-1+r} \bigcap F^k L_{1,s} \right) - \mu \left(L_{j,j-1+r} \bigcap F^k L_{s,1} \right) \right] \Big\} \qquad (3.9)$$

with $j \neq 1$. This equation can be "integrated" once to solve for $T_j(n)$,

$$T_j(n) = n \left[\delta_{j,2} \mu \left(L_{1,2} \right) + \delta_{j,0} \mu \left(L_{1,0} \right) \right] + \sum_{k=1}^{n} (n-k) a_{j,k}, \qquad (3.10)$$

where $a_{j,k}$ stands for the terms in curly brackets in (3.9), and use has been made of the initial condition $T_j(0) = 0$. The content of R_1-species in R_1 is best obtained by using the mass conservation property,

$$T_1(n) = \mu(R_1) - \sum_{j \neq 1} T_j(n) \qquad (3.11)$$

From these formulae, it can be seen that in order to evaluate the amount of R_1-species in any region R_j at time $t = n$, only the dynamics of the four lobes of the turnstiles for region R_1 is needed. In view of the symmetry property (3.6), this number can immediately be reduced to two lobes of just one turnstile. By careful examination of the symmetries possessed by the Poincaré map for the cross section $t_0 = 0$ and the ones corresponding to the sections $t_0 = \frac{T}{4}$ and $t_0 = \frac{T}{2}$ (see figure 3.6) one can show that the computation of (3.10) can be carried out by following the dynamics of just one lobe. The details of how this is done are reported in the Appendix.

Before concluding this Section, a few remarks are in order. Firstly, the formulae derived above refer to the purely convective case. The presence of molecular diffusivity adds an extra mechanism for the transport of tracer particles across the lobe boundaries and hence the equivalence between average concentration of tracer and average content of R_1-species would not hold in this case (see Section 3.2.7 for a discussion of the validity of the present theory). Secondly, the method exposed here relies on the natural structures arising in the Lagrangian description of the time dependent case, and this is the only way one can effectively describe the transport in

absence of molecular diffusivity. In particular, the methods used by Shraiman[13] and the analysis of Young et al.[14] for the time independent case would be inadequate to deal with the present case, at least sufficiently far away from the limit of zero lobe area, as their Eulerian description relies on the fact that all the fluid particle trajectories are closed and transport can only occur by molecular diffusion.

Until now we have focussed on heteroclinic points for manifolds of the same j-couple of fixed points. In the next Section we examine some of the consequences of intersections between manifolds from different tangles.

3.2.3 The first visit time and the stretching of the interface

Once inside a region, the fluid corresponding to a turnstile lobe is of course still subject to the action of the map, and in the following we would like to see if some more information can be extracted from the dynamics of F. One question we may ask is how long it takes (a portion of) the fluid just transported into a region to completely traverse it, that is, the speed at which the tracer invades unpolluted rolls.

From the lobe dynamics described above, it can be seen that the crucial quantity we have to determine in order to answer this question (and incidentally, also the question raised in §3.2.4) is the number of iterations necessary for the image of a turnstile lobe to first intersect an adjacent turnstile, e.g., the smallest integer, $\bar{m}$ say, such that $F^{\bar{m}}L_{1,0} \bigcap L_{0,-1} \neq \emptyset$. For definiteness, let us suppose that, for some choice of the parameters ϵ, ω, A, the turnstile intersection has the form depicted in figure 3.7, i.e., the segment of unstable manifold of lobe $F^{\bar{m}}L_{1,0}$ intersects the stable segment of $L_{0,-1}$ in two points, see also figure 3.10 for a blow-up. The next iteration will take the "tip" of $F^{\bar{m}}L_{1,0}$ (the shaded area in figure 3.10) into $FL_{0,-1}$ and hence the first

time fluid from R_1 will enter R_{-1} is $\bar{m}+1$. In view of the symmetry (3.6), $FL_{0,-1}$ is geometrically equal to (and will follow the same evolution as) $FL_{1,0}$, i.e., $F^{\bar{m}}L_{0,-1}$ will intersect $L_{-1,-2}$ in exactly the same way as $F^{\bar{m}}L_{1,0} \bigcap L_{0,-1}$. As a subset of $FL_{0,-1}$, $F^{\bar{m}+1}L_{1,0} \bigcap L_{0,-1}$ will undergo stretching and will eventually intersect $L_{-1,-2}$, but in general it will do so in more than $\bar{m}-1$ iterations. Let us denote by $\bar{m}'$ the number of iterations required for $F^{\bar{m}}L_{1,0}$ to first intersect $L_{-1,-2}$ in such a way that the unstable segment of $L_{1,0}$ completely "pushes through" it, i.e., intersects its boundary in four points, and hence two of these points are preceding and two are following the ones of $F^{\bar{m}}L_{0,-1}$, according to the arc length measure of $W^u_\epsilon(p^-_{-2})$, see figures 3.7 and 3.8. From now on, we will refer to the integers $\bar{m}$ and $\bar{m}'$ as the signatures of the map F corresponding to a particular choice of parameters ϵ, ω and A.

Since manifolds of the same stability type cannot have common points, a glance at figure 3.7 shows that $F^{\bar{m}+\bar{m}'}L_{1,0} \bigcap L_{-1,-2}$ is "trapped" between $W^u_\epsilon(p^-_{-1})$ and $W^u_\epsilon(p^+_{-2})$ and hence the first visit time $t_{f.v.}$ for roll R_{-3} will be bounded by

$$(3\bar{m}+1)T \leq t_{f.v.} \leq (2\bar{m}'+\bar{m}+1)T \ . \tag{3.12}$$

The symmetry property (3.6) assures that, at the next turnstile intersections, the role of trapping manifolds will played by the ones of the adjacent couples of hyperbolic fixed points, and therefore, for any roll R_{-j} $(j>0)$,

$$(j\bar{m}+1)T \leq t_{f.v.} \leq [(j-1)\bar{m}'+\bar{m}+1)]\,T \ , \tag{3.13}$$

a similar relation holding for rolls to the right of R_1. Of course, the amount of fluid actually making the first visit can be expected to rapidly go to zero in general, as the lobe $L_{1,0}$ gets stretched and thinner so that the area "clipped" by the intersections with the turnstile lobes goes to zero. An example illustrating the previous consid-

erations is offered by the case $\epsilon = 0.1$, $\omega = 0.6$, $A = 0.1$. For this choice of the parameters we have $\bar{m} = 3$, the unstable segment of $F^{\bar{m}}L_{1,0}$ has a four-point intersection with the boundary of $L_{0,-1}$, and $\bar{m}' = \bar{m} = 3$, with an eight-point intersection at the boundaries for $F^{\bar{m}+\bar{m}'}L_{1,0} \bigcap L_{0,-1}$. Therefore the lower and upper bound for the first visit time in (3.13) coincide and the tracer pollutes a new cell every three periods.

We remark here that these results only depend on the signatures $\bar{m}$ and $\bar{m}'$ and are valid for non-slip boundary conditions as well. Of course, the type of boundary conditions can have some influence on the actual value of these signatures for a given set of parameters ϵ, ω and A, however once $\bar{m}$ and $\bar{m}'$ are fixed the pace of roll invasion by the tracer is set and does not "feel" the boundary any more. Furthermore, when the lobe area is large enough, i.e., the volume of fluid transported across the roll boundary is large, $\bar{m}$ and $\bar{m}'$ are small and it is easy to see that the boundary conditions are not going to have an influence in this case, since the turnstile intersection can happen before the images of the turnstile lobe have entered the boundary layer. In fact, since one typical length scale for the turnstile lobe is fixed by the upper bound of half the distance between the top and bottom boundaries, a large area of the lobe would necessarily mean that the distance between the lobe segments of stable and unstable manifolds is large. Hence, since the map is area preserving, it will take fewer iterations for the lobe to stretch across the roll width, i.e., $\bar{m}$ and $\bar{m}'$ will be small.

The upper and lower bounds (3.13) also show that the "natural" time scale for the lateral spreading of the tracer is linear in t, at least initially, i.e., within the time scale of applicability of our theory (see § 3.2.7). This is quite different from the time independent case, where the number of invaded cells grows initially like $t^{\frac{1}{4}}$, $t^{\frac{1}{3}}$ for slip

and non-slip boundary conditions respectively[14], and once again this estimate relies entirely on the existence of molecular diffusivity. In particular, this implies that the spreading of tracer for the time independent case cannot be modelled by a Gaussian distribution, for as long as the molecular diffusivity effects can be neglected.

We will now use the signatures $\bar{m}$, $\bar{m}'$ for estimating a lower bound on the stretching of a turnstile lobe, which in turn will provide one for the unstable manifold. The region $F^{\bar{m}+\bar{m}'+1}L_{1,0} \bigcap L_{-1,-2}$ is stretched and folded by the next application of the map (see figure 3.9 for the example $\bar{m} = 1$, $\bar{m}' = 2$) and hence, at the next turnstile encounter, $F^{\bar{m}+2\bar{m}'+1}L_{1,0} \bigcap L_{-2,-3}$ will consist of at least four disjoint strips. Each of these strips will be trapped in exactly the same way as $F^{\bar{m}+\bar{m}'}L_{1,0} \bigcap L_{-1,-2}$ and, due to symmetry, will undergo a similar evolution. It is then possible to provide a lower bound for the stretching of lobes, or more physically of the interface between dyed and clear fluid. This bound is not sharp as it only takes into account the fate of particular segments of the unstable manifold, but it is rigorous and does provide an exponential estimate for the rate of stretching of the lobes. In a region $L_{U,S}$ bounded by a segment U of and a segment S between two intersection points of an unstable and stable manifolds respectively, we will define the distance of a point $p \in L_{U,S}$ from S as

$$l_p(L_{U,S}) = \inf_{q \in S} \inf_{c_{p,q}} l(c_{p,q}), \tag{3.14}$$

where $c_{p,q}$ is any continuous curve $\subset L_{U,S}$ connecting p and q, and $l(c_{p,q})$ is its length, see figure 3.11. Thus, $l(L_{U,S})$ is the distance of the point p from S *within* the region $L_{U,S}$. Furthermore, we will define the length of region $L_{U,S}$ by

$$l(L_{U,S}) = \sup_{p \in U} l_p(L_{U,S}), \tag{3.15}$$

and denote by $\bar{p}_{L_{U,S}}$ the point on U at which the sup is achieved. As time n increases,

if $F^n L_{U,S}$ gets stretched and thinner, this distance will approach the measure of half the length of the unstable segment.

Having introduced the necessary definitions, we now provide the estimate of the turnstile lobe length. Let us consider the region, L_f say, bounded by the unstable segment of $F^{\bar{m}} L_{1,0}$ and the stable segment of $L_{0,-1}$ between the first two points of intersection of boundaries of these lobes (according to the arc length of $W^u_\epsilon(p_0^-)$), i.e., the "tip" of $F^{\bar{m}} L_{1,0} \bigcap L_{0,-1}$ in the case of figure 3.7 and figure 3.10. Let us denote by $l_1, l_2 \ldots l_{\bar{m}'}$ the distances *in* $FL_{0,-1}, \ldots F^{\bar{m}'} L_{0,-1}$ of $\bar{p}_{FL_f}$, $\bar{p}_{F^2 L_f}$ etc., see figure 3.12. Thus, l_m is a shorthand notation for $l_{\bar{p}_{F^m L_f}}(F^m L_{0,-1})$, $m = 1, \ldots, \bar{m}'$. Now, consider the fate of $F^{\bar{m}+1} L_{1,0}$ under one iteration of the map. The portion $F^{\bar{m}+1} L_{1,0} \bigcap L_{0,-1}$ will have to "curl" around the region FL_f while still remaining in $FL_{0,-1}$. A lower bound for the length of $F^{\bar{m}+2} L_{1,0}$ can safely assumed to be twice the distance (in $FL_{0,-1}$) of $\bar{p}_{FL_f}$ from the stable segment of $FL_{0,-1}$, i.e., $2l_1$. We then have, carrying on with similar arguments,

$$l(F^{\bar{m}'+\bar{m}+1} L_{1,0}) \geq 2 \sum_{i=1}^{\bar{m}'} l_i \tag{3.16}$$

and since $F^{\bar{m}'+\bar{m}+1} L_{1,0} \bigcap L_{-1,-2}$ will consist of at least two strips playing the role of $F^{\bar{m}'+\bar{m}+1} L_{1,0} \bigcap L_{0,-1}$, (see figure 3.9),

$$l(F^{2\bar{m}'+\bar{m}+1} L_{1,0}) \geq (4+2) \sum_{i=1}^{\bar{m}'} l_i. \tag{3.17}$$

Hence, in general,

$$l(F^{n\bar{m}'+\bar{m}+1} L_{1,0}) \geq (2^{n+1}+1) \sum_{i=1}^{\bar{m}'} l_i \tag{3.18}$$

which shows explicitly the exponential character of the turnstile lobe and therefore of the interface stretching with time.

One can of course refine this estimate by further distinguishing between the types of turnstile intersections (whether 2 or 4 point), by taking into account the fate of manifold segments not trapped by the intersection, etc.. The information on the interface stretching can be of great practical value when, rather than a tracer, one considers the transport of species that can chemically react with each other, when usually the objective is to maximize the length of the interface between the species, which is the "core" of the layer where the reaction takes place[1].

3.2.4 Chaotic fluid particle motion

In this section we briefly examine an issue, the existence of chaotic particle motion, that has drawn considerable attention in the literature on dynamical systems[16], but whose usefulness for fluid dynamics is not completely clear, since no quantitative information can easily be extracted from it. In particular, for the theory of transport outlined above this issue is largely irrelevant.

The splitting of (some) heteroclinic orbits in a heteroclinic cycle, generally implies that a horseshoe construction can be carried out, and hence a zero-measure set of initial conditions can be found for which the motion is chaotic (see, e.g., Wiggins[15]). Being of measure zero, this set is of no physical interest in itself. However, it can induce some transient chaotic-like behaviour for orbits whose initial condition falls into a neighbourhood of this set, and hence it can be of interest in case individual trajectories of fluid particles, or very small dyed regions, were to be followed. Furthermore, one can heuristically expect that the presence of horseshoes in the flow would enhance mixing, although is not clear how to quantify this.

Various constructions of a horseshoe are possible in our case, but two in particular

are interesting, since they are responsible for two different types of chaotic motion, i.e., an inter-roll "transport" horseshoe, which leads to orbits that can move all over the array of rolls, and an "internal" one, with chaotic orbits confined inside one roll. Accordingly, the transport horseshoe can be expected to enhance inter-roll mixing whereas the internal one would have an influence on the mixing inside the roll only. We show the two constructions in figures 3.13,3.14 and figures 3.15,3.16 respectively. Both can be obtained by considering the action of the unstable manifolds on a rectangular region that contains an unstable segment, keeping in mind that the manifolds are invariant so that the boundary of the rectangle is forced to follow the evolution of the unstable segment.

Finally, the number of iterations necessary to have a complete horseshoe can be expressed in terms of the signatures $\bar{m}$ and $\bar{m}'$ introduced above, since the dynamics of the unstable manifold intersecting the turnstile lobe is (partially) determined by these numbers. For instance, for the transport horseshoe of figure 3.13, the image $F^n L_{1,0}$ would cut through the lower box B^- as a horizontal strip for $\bar{m}+1 \leq n \leq \bar{m}'$, where $\bar{m} = 2$, $n = 4$ for the particular case depicted in the figure. As indicated in the figure, the lobe $L_{1,0}$ would then "drag" the upper box B^+ along to intersect B^- in a horizontal strip. We notice at this point that this construction does not by itself constitute a proof of existence of an invariant set on which the dynamics is chaotic. The proper estimates for the rates of stretching of fluid elements have to be established, and for the case of non-slip boundary conditions the proof becomes quite technical. We will report about this on a separate paper.

3.2.5 The Melnikov method and analytical estimates of the lobe areas

So far, no explicit use has been made of the fact that ϵ is small, and the previous results hold with only mild requirements on the size of ϵ, in order to have each of the turnstile lobes entirely contained in one single region. If we assume $\epsilon \to 0$ however, then it is possible to compute the first order term of the Taylor series expansion in ϵ of the distance between $W^s_\epsilon(p^\pm_j)$ and $W^u_\epsilon(p^\mp_j)$, along the direction normal to the unperturbed (heteroclinic) orbit, without solving (3.2). Denoting this distance (with sign) by $d(\tau,\epsilon)$, where $\tau \in \mathbf{R}$ parametrizes the vertical heteroclinic orbit, i.e., $(x(-\tau), z(-\tau)) \in W^s_0(p^\pm_j) \bigcup W^u_0(p^\mp_j)$, it can be shown that (see Appendix)

$$d(\tau,\epsilon) = \frac{M(\tau)}{\|DH_0(x(-\tau), z((-\tau))\|}\epsilon + O(\epsilon^2). \tag{3.19}$$

Here D is the (x,z) gradient, $\|\cdot\|$ is the usual norm in $\mathbb{R}^2$, and $M(\tau)$ is the Melnikov function[16],

$$M(\tau) = \int_{-\infty}^{\infty} \{H_0(x(t),z(t)), H_1(x(t),z(t),t+\tau)\}\, dt\ , \tag{3.20}$$

with $\{\cdot,\cdot\}$ denoting the Poisson bracket

$$\{H_0, H_1\} \equiv \frac{\partial H_0}{\partial x}\frac{\partial H_1}{\partial z} - \frac{\partial H_0}{\partial z}\frac{\partial H_1}{\partial x}\ . \tag{3.21}$$

A glance at (3.19) suggests that at the zeros $\{\tau_i\}$ of $M(\tau)$, the manifolds get very close, $O(\epsilon^2)$. In fact, an application of the Implicit Function Theorem shows that if M changes sign there, i.e., $M(\tau_i) = 0$ and $\frac{\partial M}{\partial \tau}(\tau_i) \neq 0$, then $W^s_\epsilon(p^\pm_j)$ and $W^u_\epsilon(p^\mp_j)$ intersect transversely in a ϵ-neighborhood of $(x(-\tau), z(-\tau))$. Thus, the approximate location of pip's can be computed along the unperturbed heteroclinic orbit. Once an estimate for the distance between the manifolds and the location of pip's is obtained,

it is easy to derive an expression for the lobe areas[16],[15] valid to order ϵ:

$$\mu(L) = \epsilon \left| \int_{\tau_1}^{\tau_2} M(t)dt \right| + O(\epsilon^2), \tag{3.22}$$

where τ_1 and τ_2 two consecutive zeros of the Melnikov function and L stands for any of the turnstile lobes.

We now proceed in calculating the Melnikov function for our problem. When the Poincaré section is chosen as in § 3.2.1, the symmetry properties (3.5),(3.6) assure that the Melnikov function is independent of the particular heteroclinic connection $W_0^s(p_j^{\pm}) \bigcup W_0^u(p_j^{\mp})$, so that we may take $j = 0$ in the following. Substituting H_0 and H_1 in (3.20) with their expressions (3.2), and noting that the heteroclinic orbit from p_0^- to p_0^+ is simply

$$x(t-\tau) = 0, \quad z(t-\tau) = \frac{1}{\pi} \arcsin\left[\operatorname{sech}\left(\pi A(t-\tau)\right)\right] , \tag{3.23}$$

the integral for the Melnikov function (3.20) can be evaluated in closed form by the method of residues

$$\begin{aligned} M(\tau) &= A\omega \sin(\omega\tau) \int_{-\infty}^{+\infty} \operatorname{sech}(\pi A t) \cos(\omega t) dt \\ &= \omega \sin(\omega\tau) \operatorname{sech}(\frac{\omega}{2A}) . \end{aligned} \tag{3.24}$$

This shows explictly that M has two simple zeros per period, i.e., only one extra pip between a pip and its image, and that the manifolds intersect at a point ϵ-close to $x = 0, z = \frac{1}{2}$. The lobe area is readily evaluated to be

$$\mu(L_{1,0}) = 2\epsilon \operatorname{sech}(\frac{\omega}{2A}) + O(\epsilon^2) . \tag{3.25}$$

We notice that it increases monotonically as $\omega \downarrow 0$, $A \uparrow$, and it does not depend on the wavelength λ of the convection rolls. This immediately implies that the flux

	$\epsilon = 0.1$		$\epsilon = 0.01$	
ω	Melnikov	numerical	Melnikov	numerical
0.6	0.019865	0.019858	0.001986	0.001986
0.4	0.053160	0.052916	0.005316	0.005315
0.24	0.11045	0.11035	0.011045	0.011043

Table 3.1: Comparison between the lobe area $\mu(L_{1,0})$ estimated by Melnikov function and numerically, with $A = 0.1$

of tracer across a roll boundary is independent of λ, in agreement with the findings of Solomon and Gollub. Furthermore, the linear dependence in ϵ signifies that the flux scales linearly with the strength of the oscillation, another fact well observed experimentally.

We also provide a comparison, in Table 3.1, between the area measures predicted by the estimate (3.25) and the ones obtained numerically, for various values of ω. We note that the agreement is quite good, and hence one can predict analytically the amount of fluid that can be exchanged between the rolls in one period of oscillation, according to the model, for a wide range of parameter values.

The information in (3.25) would be enough for determining the value of the enhanced diffusion coefficient D^* introduced by Solomon and Gollub[4], according to the Fickian law

$$\Phi(x,t) = -D^*(x,t)\frac{\partial \bar{C}(x,t)}{\partial x} \ , \tag{3.26}$$

where $\Phi(x,t)$ is the flux of tracer past a roll boundary at x at time t and $\frac{\partial \bar{C}(x,t)}{\partial x}$ is the difference, between two adjacent rolls, of the roll-averaged concentration $\bar{C}(x,t)$. The quantity $\bar{C}(x,t)$ is a coarse-grained concentration profile along the array of rolls,

and is obtained from the local concentration, say $\theta(x,z,t)$, by integrating over a roll region[4],[13]. In reference[4], D^* is evaluated from the variation in the average concentration between a roll initially containing all the dye and the adjacent one, during one period of oscillation, i.e., the volume of fluid corresponding to $L_{1,0}$ in our notation. Unfortunately, we cannot provide a direct comparison with the data provided in reference[4], as the authors do not report the values of the parameters ω and A at which they were collected. As we have seen, these parameters can be even more important than the strength of the perturbation.

The foregoing discussion has focussed exclusively on the case of slip boundary conditions. It is also of interest at this point to observe the effects of non-slip boundary conditions on the lobe areas. In this case we have to replace $\sin(\pi z)$ in (3.2) with the function[17]

$$V(z) = \cos(q_0\bar{z}) - a_1\cosh(q_1\bar{z})\cos(q_2\bar{z}) + a_2\sinh(q_1\bar{z})\sin(q_2\bar{z}), \tag{3.27}$$

where q_0, q_1, etc., are positive constants whose value can be determined numerically, and $\bar{z} = z - \frac{1}{2}$. With this function, it seems that the trajectory along the unperturbed separatrix can no longer be found in closed form, and we have to evaluate the Melnikov integral numerically. However, after some manipulations, (3.20) can be written as

$$M(\tau) = 2A\omega\sin(\omega\tau)\int_{\frac{1}{2}}^{1} dz\cos(\omega t(z)) \tag{3.28}$$

where $t(z)$ is the function

$$t(z) = \int_{\frac{1}{2}}^{z}\frac{dz'}{V(z')}. \tag{3.29}$$

Since the function $V(z)$ vanishes at $z = 0, 1$ together with its derivative, the motion towards the stagnation point along the stable manifold is no longer exponential, just algebraic, like t^{-1}. From this and (3.23),(3.28), by looking at the zeros of the

integrand, it can be argued that in the limit of small ω the lobe area for non-slip boundary conditions is always smaller than the corresponding case with stress free boundary. The opposite situation would occur for large ω, as shown by figure 3.20. We also provide a comparison of the lobe areas as obtained by the Melnikov function for the two types of boundary conditions, for a few values of interest of ω, in Table 3.2.

	$\epsilon = 0.1$	
ω	slip	no-slip
0.6	0.019865	0.023956
0.4	0.053160	0.051797
0.24	0.11045	0.09253

Table 3.2: Comparison between the lobe area $\mu(L_{1,0})$ estimated by Melnikov function for the case of slip and no-slip boundary conditions, with $A = 0.1$

We note that the algebraic, rather than exponential, convergence to the fixed point introduced by the non-slip boundary conditions can cause some additional term to arise in the expression (3.19) for the distance between the manifolds by means of the Melnikov function[20]. As this is the situation likely to occur generically in a fluid mechanical context, we discuss this point in greater detail in the Appendix, where we show that (3.19) is the correct expression in the case of non-slip boundary conditions as well.

3.2.6 The structures and transport within a roll

The theory outlined in the preceding sections leads to a scenario which is completely different from the time independent case. In fact, the fluid particles for stationary con-

vection rolls follow the streamlines, and these are *closed* in this case, so that transport would only be possible, and still essentially governed, by molecular diffusivity[13]. It is then natural to ask whether this behaviour of fluid particles in the time independent case is completely wiped out as a result of the time dependent perturbation, or remnants of the stationary flow are left in some part of the roll region. In the following, we show that, for ϵ not too large, the second alternative is correct.

In the absence of molecular diffusivity, the only fluid particles that participate in the transport from roll to roll are the ones contained in the lobes and their images. We thus can define a (noncompact) transport region by just taking the union of the lobes their images,

$$R^T \equiv \bigcup_{k=0}^{\infty} \bigcup_j F^k \left[L_{j,j-1} \bigcup L_{j-1,j} \right]. \tag{3.30}$$

Clearly, in two dimensions this region cannot include the interior of any invariant closed curve contained in a R_j, and for ϵ not too large it is well known that such curves will be provided by KAM tori and island bands. Being impenetrable by the unstable manifold, and hence by the lobe images, these structures would effectively constitute a barrier to transport via lobes and prevent mixing inside a roll.

In analogy with the time independent case, transport across the largest KAM torus would be possible by molecular diffusivity only, and the region encircled, from now on referred to as "core," will in general constitute the largest part of a roll not subject to lobe transport. From the above remarks, it is clear that as $\epsilon \to 0$ the core tends to occupy the whole region R_j. The m elliptic fixed points associated with a $\frac{m}{n}$ resonance will in general be surrounded by their own KAM tori, thereby contributing m extra forbidden regions, or island chain, whenever these lie outside the core. We notice that in general the measure of the portion of phase space occupied

by an m-island chain decreases exponentially with m, as we show below for our particular case. Furthermore, the stable and unstable manifolds associated with the m hyperbolic fixed points between the islands would regulate the transport of fluid from one side of the island chain to the other by a mechanism completely analogous to the one described above for the main (inter-roll) manifolds[11]. However, we note that this mechanism is in general much less effective than the one from roll to roll, as the areas of turnstiles associated with island chains are in general many orders of magnitude smaller than the main ones[21]. We also note that a similar role in regulating the transport of fluid in the inner region would be played by the cantori, i.e., quasi periodic orbits that do not fill a torus, but rather a torus with gaps on a Cantor set[10].

Once again, in the limit of small ϵ, it is possible to provide some analytical estimates for the size and location of islands and KAM tori for the Poincaré map induced by (3.2), by means of averaging techniques[16]. Denoting by $T(\kappa)$ the period of revolution of a fluid particle along a (closed) streamline for the unperturbed flow, where $\kappa \in [0,1)$ parametrizes the family of streamlines, it can be shown that

$$T(\kappa) = \frac{4}{\pi A} K(\kappa) \; . \tag{3.31}$$

Here $K(\kappa)$ is the complete elliptic integral of first kind, κ being the elliptic modulus, $\kappa = 0$ corresponds to the roll centre, while $\kappa \to 1^-$ for orbits close to the heteroclinic cycle, thus explicitly showing how their period tends to infinity approaching the cycle. We note that the period is monotonic in κ, and hence for a given frequency of the time periodic perturbation the resonant orbit would be unique. Suppose $\bar{\kappa}$ identifies the orbit whose period satisfies the resonant condition $\frac{T(\bar{\kappa})}{T} = \frac{m}{n}$, i.e., $\bar{\kappa}$ is the solution

of

$$K(\kappa) = \frac{A\pi^2 m}{2\omega n}. \tag{3.32}$$

Then, in analogy with § 3.2.5, the simple zeros of the subharmonic Melnikov function

$$M^{\frac{m}{n}}(\tau) \equiv \int_0^{mT} \{H_0(x_{\bar{\kappa}}(t), z_{\bar{\kappa}}(t)), H_1(x_{\bar{\kappa}}(t), z_{\bar{\kappa}}(t), t+\tau)\} dt \tag{3.33}$$

would correspond to periodic points for the Poincaré map located within a $O(\epsilon)$-neighbourhood of $(x(-\tau), z(-\tau))$ on the unperturbed $\bar{\kappa}$-orbit. It can be shown in general that these points occur in even number with alternating, hyperbolic and elliptic, stability type[16].

Although the method of averaging cannot resolve the fine details of the tangle of manifolds associated with the hyperbolic points, as these would appear to be connected by heteroclinic orbits in the averaged flow, it nonetheless provides an upper bound estimate of the area of the "core region" around the elliptic points, just as $\mu(R_j)$ does for the main core. Denoting the measure of the area enclosed by the "averaged" heteroclinic cycle with $\mu(I_m)$, one can show[16] that the coefficient of the leading order term in an expansion in ϵ depends on $M^{\frac{m}{n}}$,

$$\mu(I_m) = O\left(\epsilon^{\frac{1}{2}} \sqrt{\max_{\tau} M^{\frac{m}{n}}(\tau)}\right) + O(\epsilon)\ . \tag{3.34}$$

The function $M^{\frac{m}{n}}$ can be explicitly computed for the vector field (3.2),

$$M^{\frac{m}{n}}(\tau) = \begin{cases} 2\omega \text{ sech } \left(\frac{K(\sqrt{1-\bar{\kappa}^2})\omega}{\pi A}\right) \sin(\omega\tau)\ ; & \text{for } n \neq 1 \\ \\ 0\ ; & \text{for } n = 1,\ m \text{ even}, \end{cases} \tag{3.35}$$

showing that only the $m:1$ resonances with m odd have a nonzero $O(\epsilon^{\frac{1}{2}})$ leading order term, and would hence be the most important island contribution to the forbidden

region inside a roll. Furthermore, the estimate (3.34) is in this case

$$\mu(I_m) = \epsilon^{\frac{1}{2}} \frac{32}{m^{\frac{3}{2}}\pi^2 k^{\frac{1}{2}} A^{\frac{1}{2}}} \left[\frac{\bar{\kappa}}{\bar{\kappa}'} K^3(\bar{\kappa}) \left(\frac{dK}{d\kappa} \right)^{-1} \left(\max_{\tau} M^{\frac{m}{1}}(\tau) \right) \right]^{\frac{1}{2}} + O(\epsilon), \qquad (3.36)$$

with $\kappa' = \sqrt{1-\kappa^2}$, thus showing that as m increases, i.e., $\kappa \to 1$ and the resonant orbits approach the heteroclinic cycle, the size of the islands decreases exponentially as the coefficient of the $O(\epsilon^{\frac{1}{2}})$ term goes to zero. We note that in this limit the subharmonic Melnikov function reduces to (3.24) (apart from a factor 2 which takes into account the contribution from both the heteroclinic connections pertaining to each region R_j). Table 3.3 provides a comparison between the estimate (3.36) and a numerical evaluation of the area of an island, based on the largest identifiable KAM torus, for the 3:1 resonance band, for various values of ϵ, $\omega = 0.6$ fixed. The poor agreement when $\epsilon = 0.1$ can easily be explained by the observation that, for an ϵ this large, the splitting of the heteroclinic connection in the averaged system is relatively large and a considerable portion of the island would actually appear as chaotic.

	$\omega = 0.6$	
ϵ	averaging	numerical
10^{-1}	0.1397	0.033956
10^{-2}	0.04386	0.04386
10^{-3}	0.01397	0.01400

Table 3.3: Comparison between the island area $\mu(I_3)$ estimated by averaging and numerically for decreasing ϵ, $\omega = 0.6$, with $A = 0.1$.

We show the geometry of the islands and KAM tori by numerically computing the images under F of 10 initial conditions on a segment at roll mid-height $z = \frac{1}{2}$, using a 4-th order Runge-Kutta scheme. Figure 3.17 shows the result for $\omega = 0.6$,

$\epsilon = 0.1$ for 1000 iterations. For this ω, the period is about half the minimum period $T(0)$ in (3.31). The 3 : 1 resonance band can be clearly seen and seems to be the only relevant structure besides the core region, the next (5 : 1) band being too close to the manifold tangles and hence unobservable, as the elliptic periodic points are stripped of almost all their closed orbits and the manifolds of the hyperbolic points intersect the ones from the inter-roll homoclinic tangle and are forced to follow their dynamics. To further pursue this point, we reduce ϵ to 0.01 and the result is shown in figure 3.18. The 3 : 1 resonance is now surrounded by KAM tori and the 5 : 1 is clearly visible outside a core that has enlarged as much as to almost occupy the whole roll. By reducing ω one can bring the 1 : 1 resonance into play, as in figure 3.19, where $\omega = 0.24$ and $\epsilon = 0.1$. It can be seen that the transport region now deeply penetrates into the centre of the roll, so that this situation should favour a quicker homogenization of the tracer concentration.

As a final remark, we notice that the relevance of the inner structures described above (or their equivalent for more refined 2-D models) would depend on the degree to which the two dimensional idealization of the flow is realized in practice. Even for the time independent case, for instance, the experiment shows[4] that the tracer appears to invade unpolluted rolls by diffusing inward from the boundary and outward from the centre, because of a weak boundary induced 3-D flow that convects tracer directly onto the region corresponding to the roll axis. However, this 3-D component is orthogonal to the motion of the roll boundaries and therefore is not expected to significantly affect the inter-roll transport mechanism. A qualitative comparison between the visual observations of the interface between rolls in the time dependent experiment[4] performed by Solomon and Gollub[5], and the lobe structures for the

model suggests that this is indeed the case.

3.2.7 The relative time scales of chaotic advection and molecular diffusion

The transport theory outlined in the preceding sections refers to the purely convective case, but of course in any realistic situation the tracer will always have some, albeit small, molecular diffusivity. The applicability of the theory as it stands would then be limited to the time scales where the effects of diffusion are negligible. We remark here that in presence of molecular diffusivity, the spreading of a passive tracer would be governed by an advection-diffusion equation,

$$\frac{\partial\theta(x,z,t)}{\partial t} + \{\psi(x,z,t),\theta(x,z,t)\} = \nu\left(\frac{\partial^2\theta(x,z,t)}{\partial x^2} + \frac{\partial^2\theta(x,z,t)}{\partial z^2}\right), \qquad (3.37)$$

where ν is the diffusion coefficient and $\theta(x,z,t)$ is the tracer concentration[18]. Setting $\nu = 0$, the resulting equation can then be "solved" by the method of characteristics, which is of course the approach we have been following so far, since the equations for the characteristics would be (3.2). However, mathematically the limit $\nu \to 0$ is singular, since in this way the terms containing the higher order derivatives in (3.37) cancel and the structure of the partial differential equation would be completely altered. Thus, we cannot expect to uniformly approximate the solution of (3.37) with the one for $\nu = 0$ for all times, no matter how small ν is, a situation akin to the well known (and much more complicated) case of the Euler and Navier-Stokes equations.

One way of providing a criterion of applicability for the purely convective limit naturally suggests itself. In fact, the time scale for tracer to diffuse across a distance of the order of the turnstile width should be long compared to the time it would take a lobe to be mapped across the boundary of a region, i.e., one period. Thus, denoting

this time scale by T_d we have

$$T_d = \frac{(\bar{d}(\epsilon))^2}{\nu}, \tag{3.38}$$

where $\bar{d}(\epsilon)$ is the maximum on $\tau \in [0, T]$ of the distance function defined in (3.19), and we require

$$T_d >> T \ . \tag{3.39}$$

For the cases considered in Section 3.3, taking $\nu = 5.0 \times 10^{-6}\text{cm}^2/\text{sec}$ which corresponds to the diffusivity for the methylene blue tracer used in the experiment by Solomon and Gollub, one would have the following estimates for T_d:

$\omega = 0.6$, $\epsilon = 0.1$

$$\bar{d} = 0.123 \Rightarrow T_d \simeq 2000 \simeq 200T \ , \tag{3.40}$$

$\omega = 0.6$, $\epsilon = 0.01$

$$T_d \simeq 2T \ , \tag{3.41}$$

$\omega = 0.24$, $\epsilon = 0.1$

$$\bar{d} = 0.56 \Rightarrow T_d \simeq 300T. \tag{3.42}$$

The transport theory of § 3.2.2 can therefore be expected to perform well only in the first and third case, at least within the typical total number of iterations (total time) to which the computations of Section 3.3 are carried out, which is about 20 periods. We will come back to this point in § 3.3.3, where results from numerically simulating the tracer transport in presence of diffusivity are reported.

The relative importance of lobe vs. diffusive transport can immediately be conveyed by introducing a nondimensional number by taking the ratio of the two time scales T and T_d,

$$\frac{(\bar{d}(\epsilon))^2}{T\nu} = \frac{\left(\epsilon\frac{\omega}{A}\,\mathrm{sech}(\frac{\omega}{2A})\cosh(\frac{\pi^2 A}{2\omega})\right)^2}{T}\cdot\frac{1}{\nu}. \tag{3.43}$$

so that the applicability criterion of the purely convective theory can simply be summarized by the requirement that this number be large.

3.3 Numerical simulations for three "canonical" cases

In this Section we report the results of computations based on the lobe dynamics described in the previous section, for three sets of parameter values, (i) $\epsilon = 0.1$, $\omega = 0.6$, (ii) $\epsilon = 0.1$, $\omega = 0.24$, and (iii) $\epsilon = 0.01$, $\omega = 0.6$, with $A = 0.1$, $\lambda = \pi$ in all cases. These choices are within the experimental values reported in Solomon and Gollub and are motivated by the fact that they effectively illustrate the consequences of varying the two crucial parameters in the model, ϵ and ω (A can always be scaled away and included in ω through $t \to At$ and $\omega \to \frac{\omega}{A}$). Specifically, keeping ϵ fixed and decreasing ω, i.e., going from case (i) to (ii)), not only can lead to a dramatic change in the roll inner structure, as the central elliptic point undergoes a bifurcation, but also has the effect of increasing (nonlinearly) the lobe area (see (3.25) and Table 3.1) and of changing the signatures m, m' for the turnstile intersections (from $m = m' = 3$ to $m = m' = 1$). Keeping ω fixed and decreasing ϵ, on the other hand, i.e., going from case (i) to (iii), has comparatively milder consequences, as the lobe area decreases linearly in ϵ and the locations of pips on the tangle of $W^s_\epsilon(p_j^{\pm})$ and $W^u_\epsilon(p_j^{\mp})$ remain (almost) the same (see (3.24)). Accordingly, the core region becomes larger with higher order resonance bands becoming visible, and the signatures m, m' change slightly, from $m = m' = 3$ to $m = m' = 4$ in this case.

As shown in § 3.2.2, the transport rates can be obtained once the intersection measures of images of just one turnstile lobe of R_1 with the ones of each region R_j are known. In all of the cases considered above, we choose to operate on $L_{1,0}$, covering it with a grid of points and iterating it numerically, the areas of intersection being then given by the number of points falling in each of the turnstile lobes. The use of the lobe dynamics enables us to drastically reduce the amount of computation

time with respect to simply covering the whole region of interest, in our case R_1, with a mesh of the same size. For instance, with the typical grid size adopted in case (i), $1. \times 10^{-3}$ equivalent to 19850 grid points in $L_{1,0}$, integration of (3.2), using a vectorized 4th order Runge-Kutta code on a CRAY X-MP 48 machine, results in about 55 min. of CPU time for 22 periods (with 10^{-2} for the integration time step). Even when invariant regions are identified and therefore taken away from the domain of computation, this CPU time has to be multiplied by a factor of about 50 for the direct approach with the same grid size (the number of initial conditions would be about 9.2×10^5), which brings it to the limits of feasibility of the current computational power. The lobe dynamics approach is even more advantageous when the lobe areas and hence the transport rates are small, as in case (iii). To achieve an accuracy comparable with the one of case (i), we had to use a grid step of 2.5×10^{-4} resulting in 31760 points. Although the core region is greatly increased, in these units the area outside the largest identifiable KAM torus would still amount to more than $4. \times 10^6$ points or a factor of 150 for the CPU time.

In order to check for the accuracy of the computation, we have tested the numerical results pertaining to case (i) and (ii) versus the symmetry properties (3.5),(3.6) and described in more detail in the Appendix. We typically find errors in the most significant digit for some of the lobe intersections after 20 iterations. Reducing the integrator step size or even changing the integration scheme altogether, by using an adaptive step size predictor-corrector method, has almost no influence, the error being confined to at most a difference of one point in the counting for some of the intersections, which would amount to less than a 0.5% in the intersection area, after 18 and up to 22 iterations. We present results from these computations in § 3.3.1,

where they will be compared to the predictions offered by a simple model recently proposed by several authors[10]. In § 3.3.2 we show how to derive a formula for the measure of the transport region, and discuss the mixing of the tracer in a particular situation. We conclude the Section with the results of introducing a term representing molecular diffusivity in the equations of motion (3.2).

3.3.1 Roll concentration of tracer and comparison with a Markov chain model

As we have seen in Section 3.2, the lobe area measures the amount of fluid exchanged in one period of oscillation between two neighbouring rolls. This behaviour is analogous to the transport of particles through cantori as described by MacKay et al.[10], and indeed the term turnstile was first introduced by them. These authors also propose to model the transport of species among regions connected by turnstiles as a Markov chain, in which the states represent the average concentration of species in each region, and the transition probabilities are proportional to the area of the turnstile lobes.

Specifically, in our context, let us denote by $R_j{}^T$ the portion of roll R_j which participates in the transport, i.e., according to §3.2.6, the region outside the largest KAM torus and island bands, and let r_T be its measure. The subscript j is redundant for the measure, since by the symmetries (3.5) and (3.6) the transport region will have the same size for each R_j roll. If one assumes that the fluid transported across a roll boundary quickly homogenizes over the transport region of the invaded roll, in fact instantaneously if one looks at the discrete time n denoting the number of oscillation cycles (or the iterate of the Poincaré map), the change of R_1-species in the $j-$th roll

at time n can be written as

$$\begin{aligned} T_j(n) - T_j(n-1) &= \mu(L_{j+1,j})C_{j+1}(n-1) + \mu(L_{j-1,j})C_{j-1}(n-1) \\ &\quad - [\mu(L_{j,j+1}) + \mu(L_{j,j-1})]\, C_j(n-1) \end{aligned} \tag{3.44}$$

where $C_j(n)$ is the concentration (uniform by assumption) of R_1-species in the $j-$th roll at time n, i.e., $C_j(n) \equiv \frac{T_j(n)}{r_T}$. Thus, the change in T_j can be distinguished into the increment due to the amount of tracer coming from the neighbouring rolls $j-1, j+1$, i.e., ($R_{j\pm1}$ concentration) $\times$ (volume of fluid transported in R_j), and the decrement due the tracer transported from R_j to $R_{j\pm1}$. Since the lobe areas are the same for any turnstile, we can simplify (3.44) as

$$T_j(n) - T_j(n-1) = \alpha\,(T_{j+1}(n-1) + T_{j-1}(n-1)) - 2\alpha T_j(n-1) \tag{3.45}$$

where $\alpha \equiv \frac{\mu(L_{1,0})}{r_T}$ can be regarded as the probability for a fluid particle of being transported across a roll boundary. Although very simple, the model relies heavily on the knowledge of the transition probability. As we have seen, the area of the lobe can actually be determined analytically with great accuracy, but there is apparently no way of improving the analytical estimate for r_T beyond the one of a mere upper bound.

A more fundamental problem for the applicability of the Markov model is the fact that the fluid just transported across a roll boundary *does not* homogenize rapidly once inside a roll region, and indeed the results of section 3.2.3 illustrate how one can derive some "long time" consequences from the knowledge of a few initial features of the manifold tangles (and lobe dynamics). This problem is not directly related to the size of the turnstile lobes, as we will see for case (iii), i.e., $\omega = 0.6$, $\epsilon = 0.01$, which reduces the lobe measure by one order of magnitude with respect to case (i),

	$\epsilon = 0.1$		$\epsilon = 0.01$	
ω	r_T	α	r_T	α
0.6	0.619	0.03209	0.115	0.0173
0.24	1.135	0.09723		

Table 3.4: The numerical estimates for r_T, α, with $A = 0.1$

i.e., $\omega = 0.6$, $\epsilon = 0.1$.

In order to compare the results of the lobe dynamics with the Markov chain model (3.45), we have computed the size of the transport region directly, by covering a region R_j with a grid of step size 5×10^{-3} and removing the areas inside the clearly identifiable KAM tori to reduce the total number of points of the grid. Counting the points left inside the region after 100 iterations of the Poincaré map leads to Table 3.4 for the estimate of r_T in the three cases under consideration, and correspondingly we also exhibit the transition probabilities α. Keeping track of the number of iterations required by each grid point to escape the roll region and defining a color code for representing this number, we obtain figure 3.21.a and 3.21.b, in which the transport and core regions can clearly be seen, together with (a part of) the lobe images. The 1:1 island structure visible in figure 3.21.b greatly enhances r_T with respect to the case $\omega = 0.6$, almost by a factor 2 as quantified by the numerical value reported in the table.

According to the considerations in § 3.2.2, the initial condition for $T_j(n)$, the content of R_1-species in the $j-$th roll, is $r_T\delta_{1,j}$. One can then solve (3.45) for $T_j(n)$ at any later time n and compare with the results from lobe dynamics. This is done in figure 3.22, 3.23 and 3.24, for case (i), (ii) and (iii) respectively. For each of these figures, the solid lines represent the exact computation by lobe dynamics, while the

dashed lines refer to the predictions offered by the model. Each line originating from the $t-$ axis is a plot of the content $T_j(n)$ of R_1-species in the $j-$ th roll versus time, for $j = 0, -1, \ldots, -5$, i.e., for the five rolls R_j next to the "source" roll R_1. As can be seen, the general trend of the model is to over-estimate the content of the region next to the source roll while under-estimating it for the distant regions, i.e., the lateral spreading of the tracer is not as fast as in the exact calculation (where it is linear in time, see § 3.2.3). Furthermore, the oscillations of $T_j(n)$ in time, exhibited by the (iii) case for $j = 3, 4$ and 5, cannot of course be represented by the model, and actually the Markov chain description performs worse in this case of small lobe area, or small transition rates. This is in contrast to the hope that small lobe areas would be the optimal situation for the applicability of the Markov chain approach[10].

The model can be slightly improved by taking into account the correlations introduced by the lobe dynamics, which are related to the signatures $\bar{m}$ and $\bar{m}'$ discussed in § 3.2.3. For instance, each time step of the Markov chain approach can be made to correspond to the $\bar{m}-$ th iterate of the map, rather than just one iterate, and transition probabilities connecting non-neighbouring regions R_{j-2}, R_{j+2} can be defined, based on the the measure of the intersection of $F^{\bar{m}}L_{1,0}$ with the adjacent turnstile lobe $L_{0,-1}$. However, stopping at the first signature is not sufficient to obtain a significant improvement, implying that the hypothesis of loss of memory of the fluid transported via lobes, implicit in the Markov chain approach, can be too slow for the assumptions of the model to apply, at least for the cases considered.

As a final remark, we notice that the computation time for obtaining an estimate of the transport region area can be larger than the CPU time required by the lobe dynamics approach. Although the grid need not be as refined as the one covering

the lobes, for the cases we have considered we typically have to use about twice the number of lobe grid points. Furthermore, in order to identify with some certainty the points belonging to the transport region, one has to use a large number of iterations (100 in our case). For case (i), for instance, this results in a factor of about 5 for the overall CPU time. Since the determination of r_T has an interest besides the Markov model, e.g., when one wants to achieve higher "mixing," a more efficient way of evaluating r_T is desirable. An alternative approach, using the results by lobe dynamics, is discussed in the next section.

3.3.2 The case of compact phase space

Due to the periodicity in x, the vector field (3.2) on the strip $\mathbb{R} \times [0,1]$ can be considered as the lift of one with an identical mathematical expression but defined on a cylinder, by interpreting the x variable as an angle, see figure 3.25. Viewed in this way, the phase space would no longer correspond to the physical space, but we will see that this approach simplifies the arguments leading to a formula for the measure of the transport region. Furthermore, in reality the physical space will necessarily be compact, and the cylindrical phase space can be interpreted as a (possibly very rough) approximation to a finite size Rayleigh-Bénard cell, with only two convection rolls. In order to avoid ambiguities, we will relabel the lobes $L_{0,1}$ and $L_{1,0}$ belonging to the turnstile along the separatrix along $x = 0$ as $E_{0,1}$ and $E_{1,0}$ respectively. Similarly, for the separatrix along $x = \frac{\pi}{k}$, we will denote the turnstile lobes by $D_{1,0}$ and $D_{0,1}$. The definitions of the regions, in this case only two, R_0 and R_1, and their boundaries according to § 3.2.2 are of course the same (see figure 3.25).

Once again, let us suppose that the dyed fluid is initially contained in R_1 only.

If, according to the notation introduced in § 3.2.2, we denote by $E^1_{1,0}(n)$ the amount of R_1-species contained in lobe $E_{1,0}$ and entering R_1 during cycle n, we then have

$$T_1(n) - T_1(n-1) = \mu(E^1_{0,1}(n)) + \mu(D^1_{0,1}(n)) - \mu(E^1_{1,0}(n)) - \mu(D^1_{1,0}(n)) \qquad (3.46)$$

In analogy to (3.8) and (3.9), one can show that

$$\begin{aligned} E^1_{1,0}(n) \;=\; \mu(E_{1,0}) \quad & -\sum_{k=1}^{n-1}[\mu(F^k E_{1,0} \bigcap E_{0,1}) + \mu(F^k E_{1,0} \bigcap D_{0,1}) \\ & -\mu(F^k E_{1,0} \bigcap E_{1,0}) - \mu(F^k E_{1,0} \bigcap D_{1,0})], \end{aligned} \qquad (3.47)$$

and similarly for the other terms in (3.46). Furthermore, using the symmetries (3.5), (3.6), and the area preserving property of the map, (3.46) can be rewritten simply as

$$\begin{aligned} T_1(n) - T_1(n-1) = -2\mu(E_{1,0}) \quad & +4\sum_{k=1}^{n-1}[\mu(F^k E_{1,0} \bigcap E_{0,1}) + \mu(F^k E_{1,0} \bigcap D_{0,1}) \\ & -\mu(F^k E_{1,0} \bigcap E_{1,0}) - \mu(F^k E_{1,0} \bigcap D_{1,0})]. \end{aligned} \qquad (3.48)$$

Hence,

$$\begin{aligned} \frac{1}{2}\,(T_1(n) - T_1(0)) = \quad & \sum_{l=0}^{n-1}\{-\mu(E_{1,0}) + 2\sum_{k=1}^{l}[\mu(F^k E_{1,0} \bigcap E_{0,1}) + \mu(F^k E_{1,0} \bigcap D_{0,1}) \\ & -\mu(F^k E_{1,0} \bigcap E_{1,0}) - \mu(F^k E_{1,0} \bigcap D_{1,0})]\}. \end{aligned} \qquad (3.49)$$

If we assume that $\lim_{n\to\infty} T_1(n)$ exists, then the series on the right-hand side of (3.48) converges, since the flux goes to zero, i.e., the series has value

$$\begin{aligned} \mu(E_{1,0}) = \quad & 2\sum_{k=1}^{\infty}[\mu(F^k E_{1,0} \bigcap E_{0,1}) + \mu(F^k E_{1,0} \bigcap D_{0,1}) \\ & -\mu(F^k E_{1,0} \bigcap E_{1,0}) - \mu(F^k E_{1,0} \bigcap D_{1,0})]. \end{aligned} \qquad (3.50)$$

The converse is also true, i.e., if the series converges the flux goes to zero, since if it were not so, (3.49) would lead to a divergent $T_1(n)$, which violates mass conservation. We have checked relation (3.50) numerically, from the knowledge of the lobe

intersections up to 20 iterations for the cases (i)-(iii) considered above, and the result is shown in figure 3.26-3.28 respectively. The fastest convergence is achieved from case (ii), as can be expected due to the large size of the turnstile lobe, which has the effect of producing non-empty intersections at low k values for the negative terms in (3.49). The vanishing flux would mean that the system is reaching an "equilibrium," since no matter how large n is, the rolls always exchange the same volume of fluid per period, i.e., $2\mu(E_{1,0})$, and hence this fluid must have the same amount of tracer regardless of the direction of motion, whether from or to R_1. Since, by definition, only the fluid contained in the transport region can leave the R_1 roll, this would imply that the the tracer is reaching some sort of homogenization on the transport portion of $R_0 \bigcup R_1$. Thus, in the limit $n \rightarrow \infty$,

$$T_1(\infty) = T_0(\infty) + (\mu(R_1) - \mu(R_1^T)), \tag{3.51}$$

and by mass conservation (3.11), $T_1(\infty) = \mu(R_1) - T_0(\infty)$, so that

$$T_0(\infty) = \frac{1}{2}\mu(R_1^T), \quad T_1(\infty) = \mu(R_1) - \frac{1}{2}\mu(R_1^T). \tag{3.52}$$

Substituting this relation in (3.49), and recalling that $T_1(0) = R_1$, an expression for the size of the transport region can be obtained,

$$\begin{aligned} \mu(R_1^T) = \quad & 4\sum_{l=0}^{\infty}\{\mu(E_{1,0}) - 2\sum_{k=1}^{l}[\mu(F^k E_{1,0} \bigcap E_{0,1}) + \mu(F^k E_{1,0} \bigcap D_{0,1}) \\ & -\mu(F^k E_{1,0} \bigcap E_{1,0}) - \mu(F^k E_{1,0} \bigcap D_{1,0})]\}. \end{aligned} \tag{3.53}$$

Unfortunately, the remainder of the inner summation is now magnified by the action of the second sums (as can be seen by switching the two summations), and although the convergence of (3.50) is good in the three cases considered, the same cannot be said for (3.53), as Table 3.5 shows. However, the numerical techniques we have used

are very elementary, and there is considerable room left for improvements, the usage of an adaptive grid method being among the most promising. On the other hand, it is not clear how the direct approach for the evaluation of the transport region can be improved.

The subset of R_1 that so far has been referred to as the "transport region" is also known as the "chaotic layer" in the literature. Since we have talked about homogenization, it is tempting to speculate that the flow is (mathematically) mixing on this domain. We recall[23] that the map F can be called mixing on the invariant domain $R^T = R_1^T \bigcup R_0^T$, with respect to the normalized invariant measure $\frac{\mu}{\mu(R^T)} = \frac{\mu}{2r_T}$, if for any two subset $\mathcal{M}$ and $\mathcal{N} \subset R^T$

$$\lim_{n\to\infty} \frac{\mu(\mathcal{M} \bigcap F^n \mathcal{N})}{2r_T} = \frac{\mu(\mathcal{M})\mu(\mathcal{N})}{4r_T^2}. \tag{3.54}$$

We cannot establish (3.54) for general $\mathcal{M}$ and $\mathcal{N}$, but we can check if this property is satisfied for the special case of $\mathcal{M}$ and $\mathcal{N}$ coinciding with a turnstile lobe. In particular, the convergence of the series (3.50) would imply that the following limits exist,

$$\lim_{k\to\infty} \left[\mu(F^k E_{1,0} \bigcap E_{0,1}) + \mu(F^k E_{1,0} \bigcap D_{0,1})\right] =$$

	$\epsilon = 0.1$		$\epsilon = 0.01$	
ω	numerical	series	numerical	series
0.6	0.619	0.45	0.115	0.0573
0.24	1.135	1.2103		

Table 3.5: Comparison between the numerical estimate of r_T, and the one provided by the series, with $A = 0.1$

$$= \lim_{k\to\infty} \left[\mu(F^k E_{1,0} \bigcap E_{1,0}) + \mu(F^k E_{1,0} \bigcap D_{1,0})\right], \qquad (3.55)$$

although we cannot compute the limits in closed form nor rule out the possibility that they be zero. We plot the normalized measure of the intersection of $E_{1,0}$ with all the lobes vs. the number of iterations k in figure 3.29-3.31 for cases (i)-(iii) respectively, keeping the sign in front of the "1,0" lobes, as in the argument of (3.50), for better readability. Although the final k is too low for the oscillatory behaviour to die out, and the tails around ± 1 cannot be clearly identified, there seems to be a tendency to a limit in case (i) and (ii). Case (iii) seems to be far from a convergence, which should constitute an indication of still poor mixing after 21 iterations. This is in agreement with the prediction for the measure of the transport region offered by the series method (3.53), which is worse for case (iii).

3.3.3 The effects of molecular diffusivity

From the theory of transport of a passive tracer presented in the previous sections, one element is still missing from the physics of the problem, namely molecular diffusivity, which, as discussed in § 3.2.7, can be neglected only on a short time scale. The meaning of "short" here is made precise in § 3.2.7 by introducing a diffusion time scale, based on the Melnikov estimate for the distance between manifolds. In this section, we want to explicitly demonstrate the effects of molecular diffusivity, still keeping the Lagrangian point of view, by numerically integrating the vector field (3.2), with an extra term representing the Brownian motion that a tracer particle would exhibit in presence of molecular diffusivity. This motion can be described by a generalized Langevin equation[22],

$$\dot{x} = -\frac{\partial \psi}{\partial z} + \eta(t)$$

$$\dot{z} = \frac{\partial \psi}{\partial x} + \zeta(t) \tag{3.56}$$

where ψ is the stream function (3.2) and $\eta(t)$ and $\zeta(t)$ are random variables with a Gaussian probability distribution, such that their correlations are

$$\langle \eta(t)\eta(t') \rangle = \langle \zeta(t)\zeta(t') \rangle = 2\nu\delta(t-t'), \quad \langle \eta(t)\zeta(t') \rangle = 0, \tag{3.57}$$

and ν is the diffusivity value, chosen to be $\nu = 5. \times 10^{-6}$, which is close to the experimental value determined by Solomon and Gollub for the methylene blue tracer.

Choosing an initial configuration corresponding to the set-up described in § 3.2.2, i.e., covering the whole region R_1 with a grid of step size 10^{-5}, we integrate (3.56) for 21 periods of oscillations, using the parameters of case (i) and (iii), i.e., keeping $\omega = 0.6$ fixed and decreasing ϵ by one order of magnitude, from $\epsilon = 0.1$ to $\epsilon = 0.01$. This reduces the nondimensional number (3.43) and, accordingly, the diffusion time scale T_d by two orders of magnitude, and should therefore exhibit the transition, within $21T$, from a transport dominated by chaotic advection to one dominated by molecular diffusivity.

The results are shown in figures 3.32 and 3.33, where we plot the tracer content of each roll R_j, $j = 0, \ldots -4$ vs. time, for case (i) and (iii) respectively, the dashed lines referring to the case with molecular diffusivity, and the solid to the purely convective case, as in figure 3.22 and 3.24. The comparison shows that the lateral spreading of the tracer, for $j < 0$, is severely reduced in case (iii), while for (i) there is very little difference. The tracer content of each region R_j shows the general trend of being higher than the corresponding case with no molecular diffusivity for the rolls closer to the "source" roll R_1, and lower for rolls far away. This is especially evident for R_0, i.e., the region next to R_1, as can can be expected, since tracer can now cross the

boundary between R_1 and R_0, where the concentration gradient is initially large, by a (slow) diffusion process, even without being transported by a lobe. For the rolls not adjacent to R_1, the invading tracer will be almost exclusively the one propagating via lobes, since the spreading by diffusion is evolving on a slower time scale. However, as the volume of fluid corresponding to a lobe is stretched through regions of clear fluid, and the interface elongates, its net tracer content is going to be decreased by molecular diffusion. Tracer particles are in fact now able to cross the manifolds, smoothing the gradient in concentration which would be present at the lobe boundary. Hence, in the cases where the turnstile lobe is small as in (iii), the fluid corresponding to a lobe is going to be virtually depleted of tracer in a few iterations, and will enter the far lobes practically as clear fluid, thus effectively inhibiting the lateral spreading of dye.

Acknowledgements

This research was partially supported by an NSF Presidential Young Investigator Award and an ONR Young Investigator Award.

Appendix B

In this appendix we show in some detail how the use of the symmetries pertaining to each Poincaré section can significantly reduce the amount of computation required for evaluating (3.9). According to this formula (and (3.10)), we can evaluate the content of R_1-species in any region R_j once we a) know the (forward) images under F of the four turnstile lobes of region R_1, and b) determine the measure of their intersections with each of the turnstile lobes of a region R_j. These two steps have to

be carried out numerically, and of them a) is certainly the most expensive in terms of CPU time (see §3). Therefore, it is very important to try to reduce the number of lobes whose images are needed in (3.9) by identifying the largest possible number of identities for the lobe intersections. In the following, for definiteness and with no loss of generality, we will assume j to be an even negative integer or zero . We will show that the images of only one lobe, $L_{1,0}$ for instance, need to be evaluated.

First of all we introduce some notation. In the following S_x^R, S_z^R will denote the reflection operators with respect to the $x = 0$ and $z = \frac{1}{2}$-axis respectively, $S_l^{T_x}$ the translation operator by a distance l along the x-axis. Thus, for a point $(x, z) \in \mathbb{R}^2$,

$$S_x^R\,(x,z) = (-x,z), \quad S_z^R\,(x,z) = (x,1-z), \quad S_l^{T_x}\,(x,z) = (x+l,z). \tag{B.1}$$

It is obvious from the definitions (B.1) that these maps are one-to-one and area preserving. Furthermore, they commute with each other, and $(S_x^R)^2 = (S_z^R)^2 = I$, with I the identity mapping. We also recall that if an ordinary differential equation is left invariant by a symmetry operation, integral curves are transformed into each other under the symmetry[25]. Thus, for any point (x_0, z_0) in $\mathbf{R} \times [0,1]$, if S is the symmetry map, and $g_{t_1}^{t_2}$ is the t-advancing map[26],

$$g_{t_1}^t(x_0,z_0) = (x(t,t_1,x_0,z_0), z(t,t_1,x_0,z_0))\,, \tag{B.2}$$

then

$$g_{t_1}^t S\,(x_0,z_0) = S g_{t_1}^t(x_0,z_0). \tag{B.3}$$

This in particular holds for t discretized on the set $t = Tn$, $n \in \mathbb{Z}$, that is, for the Poincaré map (3.4). In the following we will always take $t_1 = 0$ and omit the subscript t_1.

We begin by noticing that the symmetries (3.6) and (3.7) enable us to reduce to the images of only one couple of turnstile lobes, say the one near $x = 0$. In fact, by reflection about the line $z = \frac{1}{2}$ and translation of $\frac{\lambda}{2}$ along x we have, using (B.3),

$$\begin{aligned} g^t L_{0,1} &= S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, g^t (S^{T_x}_{+\frac{\lambda}{2}} S^R_z \, L_{0,1}) \\ &= S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, g^t L_{1,2} \end{aligned}$$

and so

$$S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j,j+1} \bigcap F^k L_{1,2}) = L_{j-1,j} \bigcap F^k L_{0,1}. \tag{B.4}$$

Similarly,

$$\begin{aligned} S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j,j+1} \bigcap F^k L_{2,1}) &= L_{j-1,j} \bigcap F^k L_{1,0} \\ S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j+1,j} \bigcap F^k L_{1,2}) &= L_{j,j-1} \bigcap F^k L_{0,1} \\ S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j+1,j} \bigcap F^k L_{2,1}) &= L_{j,j-1} \bigcap F^k L_{1,0} \end{aligned} \tag{B.5}$$

and

$$\begin{aligned} S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j-1,j} \bigcap F^k L_{1,2}) &= L_{j-2,j-1} \bigcap F^k L_{0,1} \\ S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j-1,j} \bigcap F^k L_{2,1}) &= L_{j-2,j-1} \bigcap F^k L_{1,0} \\ S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j,j-1} \bigcap F^k L_{1,2}) &= L_{j-1,j-2} \bigcap F^k L_{0,1} \\ S^{T_x}_{-\frac{\lambda}{2}} S^R_z \, (L_{j,j-1} \bigcap F^k L_{2,1}) &= L_{j-1,j-2} \bigcap F^k L_{1,0}. \end{aligned} \tag{B.6}$$

Thus, only $L_{0,1}$ and $L_{1,0}$ need to be iterated and their intersections with the turnstile lobes near $x = j\frac{\lambda}{2}, (j-1)\frac{\lambda}{2}, (j-2)\frac{\lambda}{2}$ determined for evaluating all the terms in (3.9).

We will now take Poincaré sections at $t_0 = \frac{T}{4}$ and $t_0 = \frac{T}{2}$. As remarked in Section 3.2, in general each Poincaré section may have its own symmetries, and, in addition, symmetries might exist between the phase portraits of different cross

sections. We notice that a Poincaré map with $t_0 \neq 0$ can be obtained from the velocity field (3.2) by introducing t_0 as a phase in the argument of $f(t)$, i.e., in our case $\cos(\omega t + \omega t_0)$.

Let us denote by $\tilde{\mathcal{M}}$, $\tilde{\tilde{\mathcal{M}}}$ the images of a region $\mathcal{M}$ of phase space under the action of $g^{\frac{T}{4}}$ and $g^{\frac{T}{2}}$ respectively. It can be seen that, for the $t_0 = \frac{T}{4}$ section, besides the overall translational symmetries along the x-axis, (3.6) and (3.7), one also has invariance of equation (3.2) (with $f(t) = -\sin(\omega t)$), with respect to $S_x^R S_z^R$ and $t \to -t$. Thus, the unstable manifold $\tilde{W}_\epsilon^u(p_0^-)$ can be obtained from $\tilde{W}_\epsilon^s(p_0^+)$ by the symmetry $z \to 1 - z$ and $x \to -x$. As usual, using (3.6), the tangle corresponding to the unperturbed position $x = \frac{\lambda}{2}$ can be obtained from the one for $p_0^\pm$ by reflection S_z^R and translation $S_{\frac{\lambda}{2}}^{T_x}$.

The Poincaré section at $t_0 = \frac{T}{2}$ can be obtained from the velocity field (3.2), with $\cos(\omega t + \omega t_0) = -\cos\omega t$ in place of $\cos\omega t$. Thus, we will have the same symmetries as for the section $t_0 = 0$, and in addition for any orbit $\{F^n(x,z)\}$, $n \in \mathbb{Z}$, for the Poincaré section at $t_0 = 0$, $S_x^R\{F^n(x,z)\}$ will be an orbit of $F_{\frac{T}{2}}$, where (x,z) is any point $\in \mathbb{R} \times [0,1]$. In particular, this implies that the manifolds can be mapped into each other by S_x^R, e.g., $\tilde{\tilde{W}}_\epsilon^s(p_0^+) = S_x^R W_\epsilon^s(p_0^+)$. Figure 3.6 shows the structure of the manifold tangles for the Poincaré sections at $t_0 = 0, \frac{T}{4}, \frac{T}{2}$ and identifies the turnstile lobes.

Focusing on the section $t_0 = \frac{T}{4}$ first, we first notice that, following how pip's are mapped under $F_{\frac{T}{4}}$ and $S_x^R S_z^R$,

$$F_{\frac{T}{4}} \tilde{L}_{0,1} = S_x^R S_z^R \tilde{L}_{0,1}. \tag{B.7}$$

Using the symmetries mentioned above, and (B.3), one can see that

$$g^{-t}\tilde{L}_{j,j+1} \;=\; S_{j\frac{\lambda}{2}}^{T_x} S_x^R S_z^R g^t (S_x^R S_z^R \tilde{L}_{0,1})$$

$$= S^{T_x}_{j\frac{\lambda}{2}} S^R_x S^R_z g^t (F_{\frac{T}{4}} \tilde{L}_{0,1}),$$

and so

$$(S^{T_x}_{-j\frac{\lambda}{2}} S^R_x S^R_z)\ (F^{k+1}_{\frac{T}{4}} \tilde{L}_{0,1} \bigcap \tilde{L}_{j+1,j}) = F^{-k}_{\frac{T}{4}} \tilde{L}_{j,j+1} \bigcap \tilde{L}_{1,0}. \tag{B.8}$$

Area preservation then implies

$$\mu\left(L_{j+1,j} \bigcap F^k L_{0,1}\right) = \mu\left(L_{j,j+1} \bigcap F^{k-1} L_{1,0}\right). \tag{B.9}$$

Similarly, for the section at $t_0 = \frac{T}{2}$, we can write

$$L_{j,j-1} \bigcap F^{k+1} L_{0,1} = S^R_x\ (\tilde{\tilde{L}}_{j-1,j} \bigcap F^k_{\frac{T}{2}} \tilde{\tilde{L}}_{1,0}), \tag{B.10}$$

from which, due to area preservation of the Poincaré map,

$$\mu\left(L_{j,j-1} \bigcap F^k L_{0,1}\right) = \mu\left(L_{j-1,j} \bigcap F^{k-1} L_{1,0}\right). \tag{B.11}$$

Thanks to these relations, two of the four terms in (3.9) involving images of $L_{0,1}$ can be eliminated in favor of terms containing only $L_{1,0}$. The remaining two terms involving $L_{0,1}$ are $L_{j-1,j} \bigcap F^k L_{0,1}$ and $L_{j,j+1} \bigcap F^k L_{0,1}$. Using the translational symmetry $x \to x + \lambda q$, $q \in \mathbb{Z}$ and (3.6), we have, for $q = -j$,

$$S^{T_x}_{-j\lambda} S^R_z\ F^{k+1} L_{0,1} = F^{-k} L_{-j+1,-j}, \tag{B.12}$$

where we have used

$$F L_{0,1} = S^R_z L_{1,0} \tag{B.13}$$

which once again can be obtained by looking at how the pip's defining the turnstile lobes are mapped by F and S^R_z. We can then write

$$S^{T_x}_{-j\lambda} S^R_z\ (F^{k+1} L_{0,1} \bigcap L_{j,j+1}) = F^{-k} L_{-j+1,-j} \bigcap F L_{1,0}. \tag{B.14}$$

Similarly, for $q = -j + 1$,

$$S^{T_x}_{-j+1\lambda} S^R_z \left(F^{k+1} L_{0,1} \bigcap L_{j-1,j}\right) = F^{-k} L_{-j+2,-j+1} \bigcap F L_{1,0} \tag{B.15}$$

The usual area preservation argument then implies

$$\begin{aligned} \mu(L_{j-1,j} \bigcap F^k L_{0,1}) &= \mu(L_{-j+2,-j+1} \bigcap F^k L_{1,0}) \\ \mu(L_{j,j+1} \bigcap F^k L_{0,1}) &= \mu(L_{-j+1,-j} \bigcap F^k L_{1,0}). \end{aligned} \tag{B.16}$$

Using the relations (B.5-B.16) one can determine each term in (3.9) using (forward) images of $L_{1,0}$ only. The summation in (3.9) can be rewritten as

$$\begin{aligned} &T_j(n) - T_j(n-1) = \\ &= (\delta_{j,2} + \delta_{j,0})\, \mu(L_{1,0}) + \\ &\quad \sum_{k=1}^{n-1} \Big\{ 2\mu\left(L_{j-1,j} \bigcap F^k L_{1,0}\right) - 2\mu\left(L_{-j+2,-j} \bigcap F^k L_{1,0}\right) \\ &\qquad -2\mu\left(L_{j,j-1} \bigcap F^k L_{1,0}\right) + 2\mu\left(L_{j-1,j} \bigcap F^{k-1} L_{1,0}\right) \\ &\qquad +\mu\left(L_{j+1,j} \bigcap F^k L_{1,0}\right) - \mu\left(L_{j,j+1} \bigcap F^{k-1} L_{1,0}\right) \\ &\qquad -\mu\left(L_{j,j+1} \bigcap F^k L_{1,0}\right) + \mu\left(L_{-j+1,-j} \bigcap F^k L_{1,0}\right) \\ &\qquad +\mu\left(L_{-j+2,-j+1} \bigcap F^k L_{1,0}\right) - \mu\left(L_{j-2,j-1} \bigcap F^k L_{1,0}\right) \\ &\qquad -\mu\left(L_{j-2,j-1} \bigcap F^{k-1} L_{1,0}\right) + \mu\left(L_{j-1,j-2} \bigcap F^k L_{1,0}\right) \Big\} \end{aligned} \tag{B.17}$$

Appendix C

In this appendix we show how the Melnikov approximation (3.19), (3.20), for the distance between the stable and unstable manifold is also valid in presence of no-slip boundary conditions. Under this assumption any rigid boundary in the flow will correspond to a curve of fixed points. This situation can certainly be expected to

occur generically in problems with a fluid mechanical interpretation, but is almost neglected in the literature on dynamical systems, and there is a lack of standard terminology for this case.

We will refer to the set of points which reach the boundaries $z = 0$, $z = 1$ asymptotically with forward or backward iterations of the Poincaré map as centre stable manifold and centre unstable manifold respectively[19]. By definition these curves are invariant under the action of the Poincaré map. The Jacobian of the vector field (3.2) at the fixed points $p^{\pm}_{j,0}$ vanishes identically, and information on local behaviour of invariant manifolds can no longer be obtained by linearizing the vector field around the fixed points. In fact, the convergence to the fixed points on the boundaries is only algebraic in the case of non-slip boundary conditions, as opposed to exponential for the slip (hyperbolic) case. However, in the unperturbed case the manifolds are explicitly known, merging into the separatrix between two rolls, and we will then assume that for small ϵ the perturbed manifolds exist[24], going on to examine the question of their mutual distance.

Referring to orbits lying on $W^u_\epsilon(p^{\pm}_j)$ and $W^s_\epsilon(p^{\mp}_j)$ as

$$q^{u,s}_\epsilon \equiv \begin{pmatrix} x^{u,s}_\epsilon(t,\tau) \\ z^{u,s}_\epsilon(t,\tau) \end{pmatrix} \tag{C.1}$$

a distance (with sign) between the manifolds at time $t = 0$ can be introduced as

$$d(\tau,\epsilon) = \frac{DH_0(q_0(-\tau)) \cdot (q^u_\epsilon(0,\tau) - q^s_\epsilon(0,\tau)}{\|DH_0(q_0(-\tau))\|} \tag{C.2}$$

where $q_0(t-\tau) = (0, z(t-\tau))$ is the unperturbed orbit on the heteroclinic given by (3.29), the dot is the usual scalar product in $\mathbf{R}^2$, and otherwise the notation of (3.19) is used. Therefore, this distance is the projection of the separation $q^u_\epsilon - q^s_\epsilon$ along the unit vector normal to the unperturbed heteroclinic orbit (whose tangent is in the

direction of $(-\partial_z H_0, \partial_x H_0)$, i.e., the vector field).

If we further assume that the manifolds be differentiable with respect to the parameter ϵ, we can look for an approximation to $d(\tau, \epsilon)$ by Taylor expanding (C.3),

$$d(\tau,\epsilon) = \epsilon \frac{DH_0(q_0(-\tau)) \cdot (\partial_\epsilon q_\epsilon^u(0,\tau) - \partial_\epsilon q_\epsilon^s(0,\tau))\Big|_{\epsilon=0}}{\|DH_0(q_0(-\tau))\|} + O(\epsilon^2) \ . \qquad \text{(C.3)}$$

The Melnikov trick would then consist in deriving a linear differential equation for the function at the numerator in (C.3), i.e.,

$$\Delta^{u,s}(t,\tau) \equiv DH_0(q_0(t-\tau)) \cdot \frac{\partial q_\epsilon^{u,s}(t,\tau)}{\partial \epsilon}\Big|_{\epsilon=0} \qquad \text{(C.4)}$$

by restoring the time dependency in q_0, $q_\epsilon^{u,s}$. Since $\frac{\partial q_\epsilon^{u,s}(t,\tau)}{\partial \epsilon}\Big|_{\epsilon=0}$ satisfies the first variational equation

$$\frac{d}{dt}\frac{\partial q_\epsilon^{u,s}(t,\tau)}{\partial \epsilon}\Big|_{\epsilon=0} = JD^2H_0(q_0(t-\tau))\frac{\partial q_\epsilon^{u,s}(t,\tau)}{\partial \epsilon}\Big|_{\epsilon=0} + JDH_1(q_0(t-\tau),t) \ , \qquad \text{(C.5)}$$

where J is the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and taking into account that q_0 solves the unperturbed equation 3.2, we have

$$\begin{aligned} \dot{\Delta}^{u,s}(t,\tau) &= \left[\text{trace} JD^2H_0(q_0(t-\tau))\right] \Delta^{u,s} + DH_0(q_0(t-\tau)) \cdot JDH_1(q_0(t-\tau),t) \\ &= \{H_0(q_0), H_1(q_0,t)\}, \end{aligned} \qquad \text{(C.6)}$$

where we have used the Poisson bracket defined in § 3.2.5 and the fact that the trace is identically zero. Hence,

$$\Delta^{u,s}(t_1,\tau) - \Delta^{u,s}(0,\tau) = \int_0^{t_1} \{H_0(q_0), H_1(q_0,t)\} dt \ . \qquad \text{(C.7)}$$

and the integrand can be recognized as the one appearing in the Melnikov function (3.20). By taking the limits $t_1 \to \infty$ and $t_1 \to -\infty$ for the stable and unstable parts

respectively, one can obtain the first term in (C.3) as

$$\begin{aligned}\Delta^u(0,\tau) - \Delta^s(0,\tau) &= \int_{-\infty}^{+\infty} \{H_0(q_0(t)), H_1(q_0(t), t+\tau)\} dt \\ &+ \lim_{t_1 \to -\infty} \Delta^u(t_1,\tau) - \lim_{t_1 \to \infty} \Delta^s(t_1,\tau) \end{aligned} \tag{C.8}$$

provided the limits and the integral exist.

In our case, the question about the integral has already been resolved, since it can be expressed in the form (3.28). The integral term is of course the Melnikov function, and in order to show that the distance is given by (3.19) at first order in ϵ, the extra terms given by the limits have to vanish. In the hyperbolic case, this can always shown to be true[15], but, as already remarked, in the case under consideration the convergence of $q_0(t)$ to the fixed points is only algebraic. For instance, in the limit of large $|t|$, the expression (3.27) for $V(z)$ and the differential equation satisfied by z

$$\dot{z} = AV(z) \tag{C.9}$$

shows that

$$z(t-\tau) = \frac{\beta}{t}\left(1 + \frac{Q(t)}{t}\right) \tag{C.10}$$

with β constant and Q bounded as $t \to -\infty$.

From

$$DH_0(q_0(t)) = \begin{pmatrix} 0 \\ V(z(t-\tau)) \end{pmatrix}, \tag{C.11}$$

and from (C.4), it can be seen that all one has to check in order for the second and third term in (C.8) to vanish is (by the symmetry (3.5), we need to analyze one term only, the unstable one say)

$$\lim_{t \to -\infty} V(z(t-\tau)) \frac{\partial q^u_{2,\epsilon}(t,\tau)}{\partial \epsilon}\Big|_{\epsilon=0} = 0, \tag{C.12}$$

with $q^u_{2,\epsilon}$ the second ('z') component of q^u_ϵ. Denoting $\frac{\partial q^u_\epsilon(t,\tau)}{\partial \epsilon}\Big|_{\epsilon=0}$ by $\begin{pmatrix} y_1(t,\tau) \\ y_2(t,\tau) \end{pmatrix}$, it is easy to see that the first variational equation, evaluated on the heteroclinic orbit, reduces to

$$\begin{aligned} \dot{y}_1 &= -AV'(z(t-\tau))y_1 + Af(t)V'(z(t-\tau)) \ , \\ \dot{y}_2 &= AV'(z(t-\tau))y_2 \ , \end{aligned} \tag{C.13}$$

where the prime denotes differentiation of the function V with respect to z. Recalling that $\dot{z} = AV(z)$, a solution for y_2 is simply

$$y_2(t,\tau) = \text{const.} \times \dot{z}(t-\tau) = \text{const.} \times V(z(t-\tau)). \tag{C.14}$$

Similarly, the solution for y_1 is

$$y_1(t,\tau) = \frac{\text{const.}}{V(z(t-\tau))} + \frac{1}{V(z(t-\tau))}\int_\tau^t f(t')V'(z(t'-\tau))V(z(t'-\tau))dt'. \tag{C.15}$$

Hence, the limit (C.12) reduces to

$$\lim_{t\to\infty} V(z(t-\tau))y_2(t,\tau) = \lim_{t\to\infty} V(z(t-\tau))V'(z(t-\tau)) = 0 \ , \tag{C.16}$$

since $\lim_{t\to\infty} z(t) = 0$, and $V(0) = V'(0) = 0$.

References

[1] Ottino, J.M. 1989 *The Kinematics of Mixing: Stretching, Chaos, and Transport.* Cambridge Univ. Press, Cambridge.

[2] Busse, F.H. 1972 The oscillatory instability of convection rolls in a low Prandtl number fluid. *J. Fluid Mech.* **52**, 97-112.

[3] Clever, R.M. & Busse, F.H. 1974 Transition to time-dependent convection. *J. Fluid Mech.* **65**, 625-645.

[4] Solomon, T.H. & Gollub, J.P. 1988 Chaotic particle transport in time-dependent Rayleigh-Bénard convection. *Phys. Rev. A* **38**, 6280-6286.

[5] Solomon, T.H. & Gollub, J.P. 1988 Passive transport in steady Rayleigh-Bénard convection. *Phys. Fluids* **31**, 1372-1379.

[6] Gollub, J.P. & Solomon, T.H. 1988 in *Proceedings of the Fritz Haber International Symposium* (I. Procaccia ed.), Plenum, New York .

[7] Moses, E. & Steinberg, V. 1988 Mass transport in propagating patterns of convection. *Phys. Rev. Lett.* **20**, 2030-2033.

[8] Cardoso, O. & Tabeling, P. 1989 Anomalous diffusion in a linear system of vortices. *Eur. J. Mech. B/Fluids* **8**, 459-470.

[9] Easton, R.W. 1986 Trellises formed by stable and unstable manifolds. *Trans. Am. Math. Soc.* **294**, 719-732.

[10] MacKay, R.S., Meiss, J.D., Percival, I.C. 1984 Transport in Hamiltonian systems. *Physica D* **13**, 55-81.

[11] Rom-Kedar, V., Leonard, A. & Wiggins, S. 1989 An analytical study of

transport, mixing and chaos in an unsteady vortical flow. *J. Fluid Mech.* In press.

[12] Rom-Kedar, V. & Wiggins, S. 1990 Transport in two dimensional maps. *Arch. Rat. Mech. Anal.* **109**, 239-298.

[13] Shraiman, B.I. 1987 Diffusive Transport in a Rayleigh-Bénard convection cell. *Phys. Rev. A* **36**, 261-267.

[14] Young, W., Pumir A. & Pomeau, Y. 1989 Anomalous diffusion of tracer in convection rolls. *Phys. Fluids. A* **1** , 462-469.

[15] Wiggins, S. 1988 *Global Bifurcations and Chaos.* Springer-Verlag, New York.

[16] Guckenheimer, J. & Holmes, P. 1983 *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* Springer-Verlag, New York.

[17] Chandrasekhar, S. 1961 *Hydrodynamics and Hydromagnetic Stability.* Dover, New York.

[18] Aref, H. 1983 Stirring by chaotic advection. *J. Fluid Mech.* **143**, 1-21.

[19] Carr, J. 1980 *Applications of Centre Manifold Theory.* Springer Verlag, New York.

[20] Schecter, S. 1987 Melnikov's method at a saddle node and the dynamics of the forced Josephson junction. *SIAM J. Math. Anal.* **18**, 1699-1715.

[21] Holmes, P., Marsden J. & Scheurle, J. 1988 Exponentially small splitting of separatrices with application to KAM theory and degenerate bifurcations. *Contemp. Math. A* **81**, 213-244.

[22] Chandrasekhar, S. 1943 Stochastic problems in physics and astronomy. *Rev.*

Mod. Phys. **15**, 1-89.

[23] Lasota, A. & Mackey M. 1985 *Probabilistic Properties of Deterministic Systems.* Cambridge Univ. Press, Cambridge.

[24] Hirsch, M.W., Pugh, C.C., & Shub, M. 1977 *Invariant Manifolds.* Lecture Notes in Mathematics, Springer-Verlag, New York.

[25] Arnold, V.I. 1988 *Geometrical Methods in the Theory of Ordinary Differential Equations.* Grundlehren der Mathematischen Wissenschaften 250, Springer-Verlag, New York.

[26] Arnold, V.I. 1973 *Ordinary Differential Equations.* The MIT Press, Cambridge.

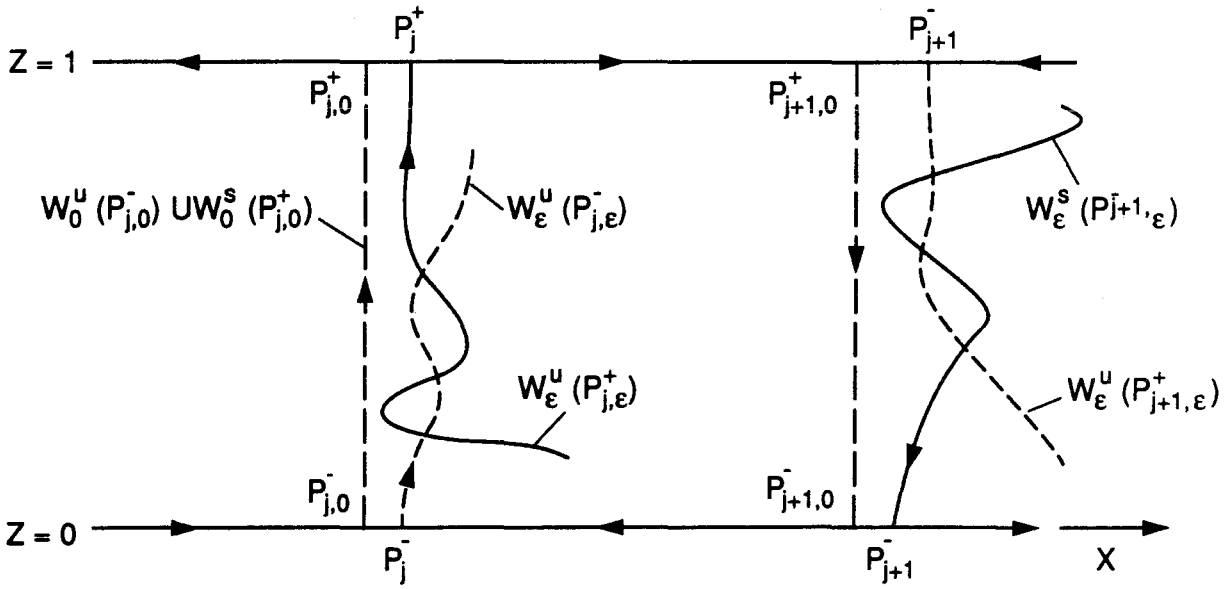

Figure 3.1. The splitting of the vertical heteroclinic connections for the Poincaré map.

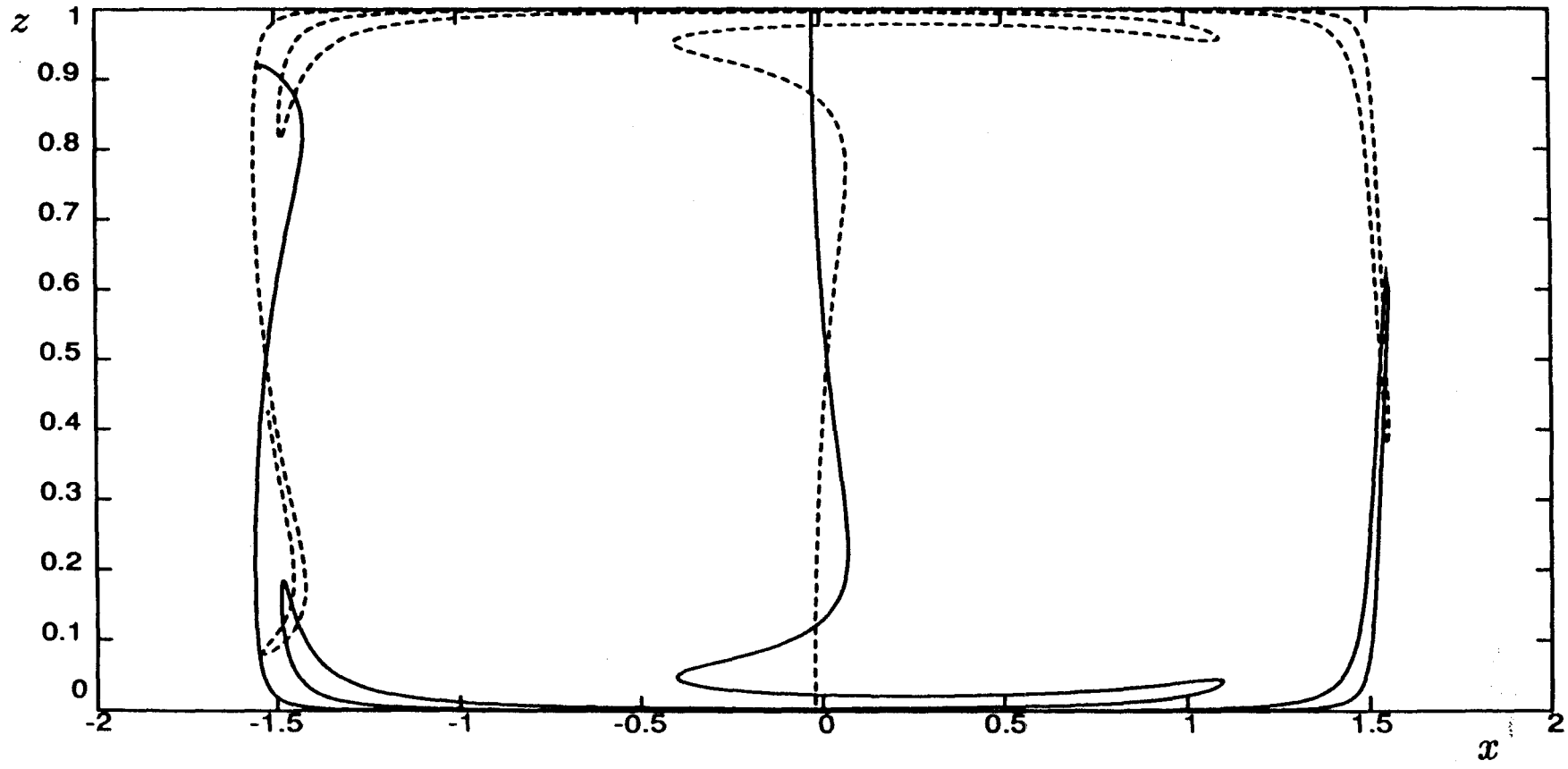

Figure 3.2. The tangle of the stable (solid) and unstable (dashed) manifold of, respectively, the points p_0^+ and p_0^-, for the Poincaré section $t_0 = 0$.

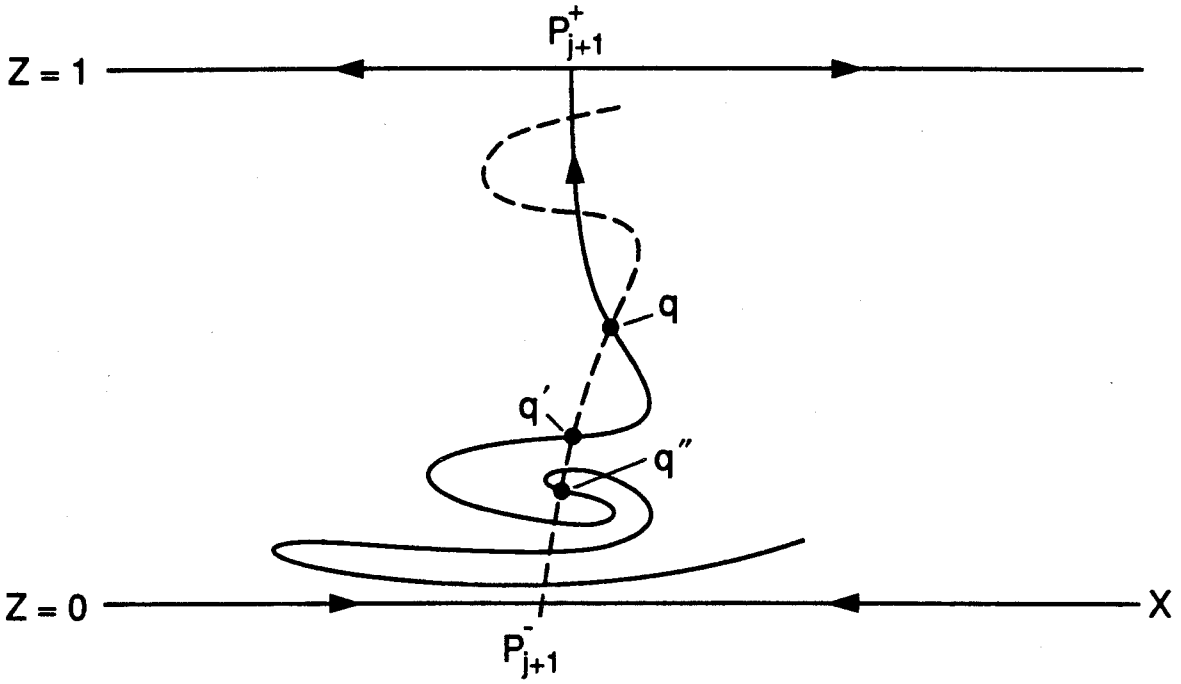

Figure 3.3. Example of heteroclinic points; q and q' are pips, q'' is not.

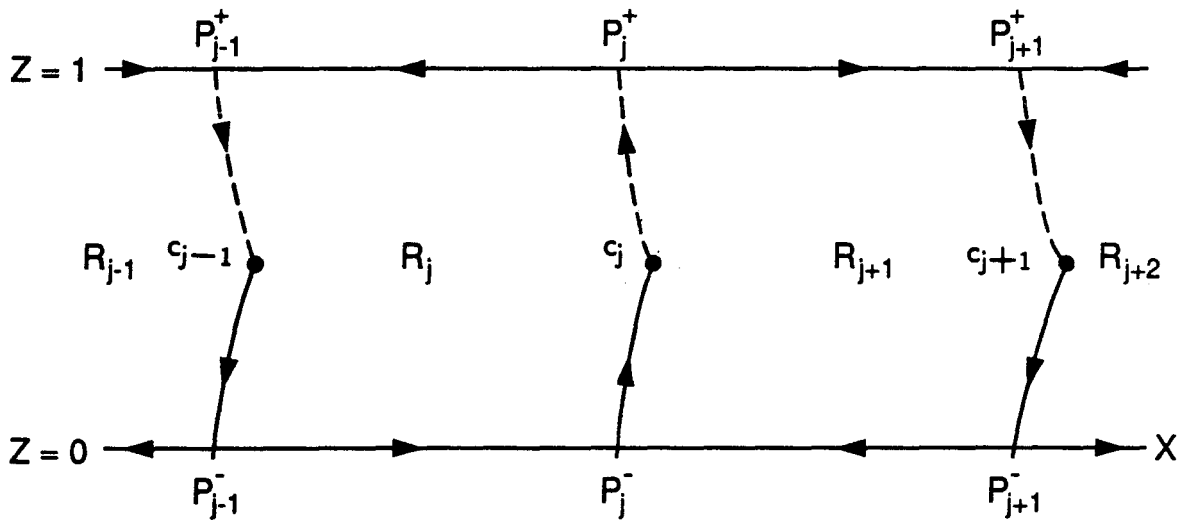

Figure 3.4. The definition of the time dependent analogue of the roll regions R_j.

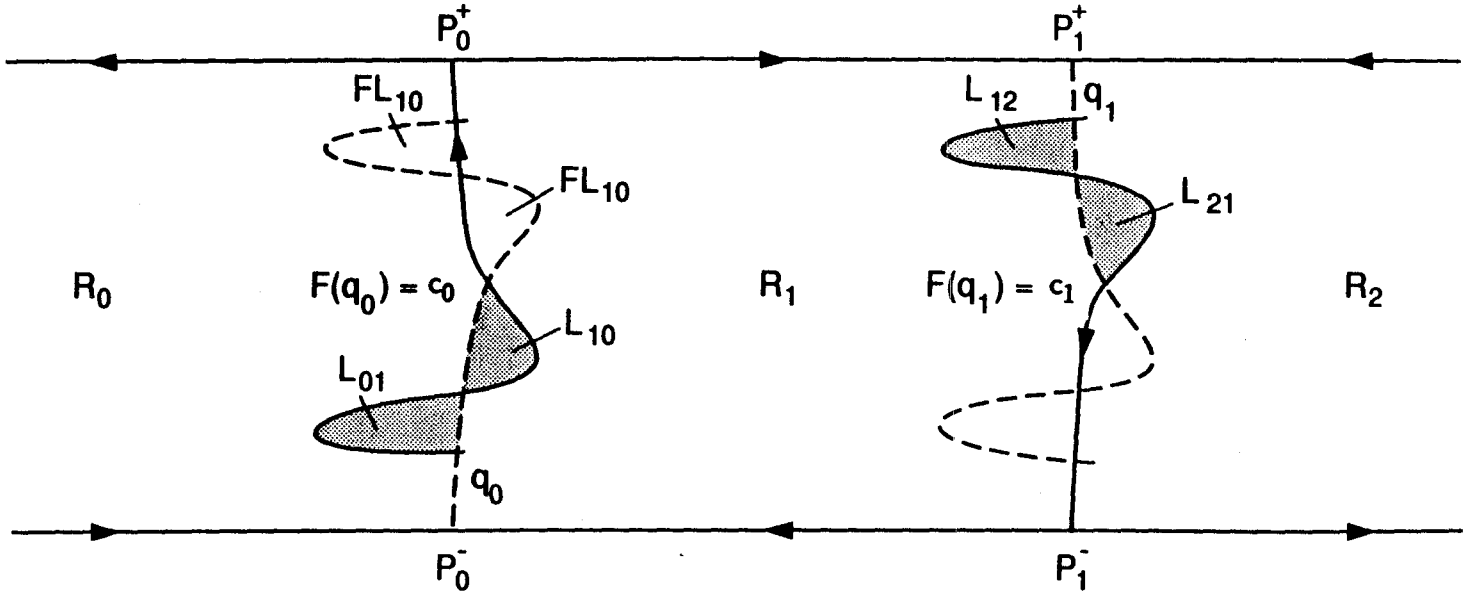

Figure 3.5. The turnstile lobes between region R_1 and R_0.

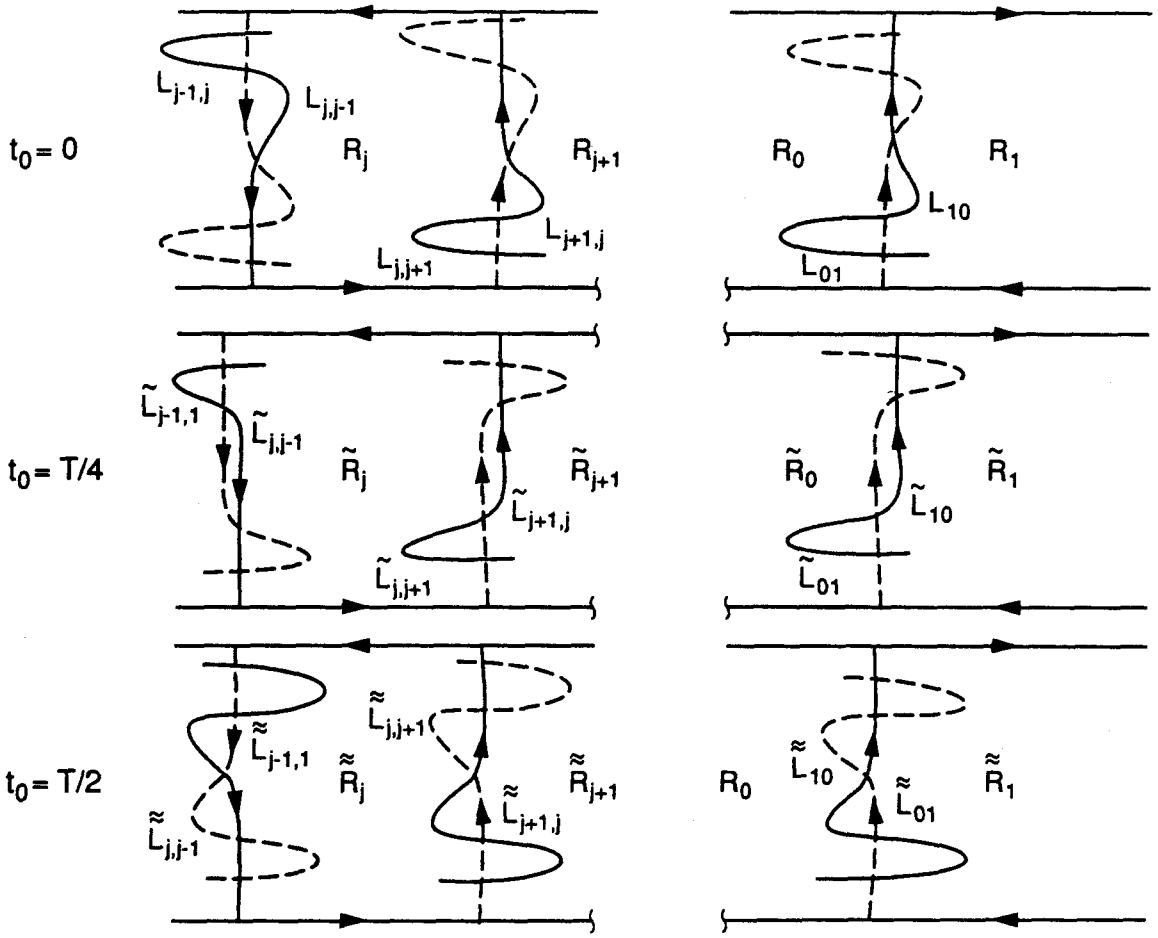

Figure 3.6. The Poincaré sections at $t_0 = 0$, $\frac{T}{4}$ and $\frac{T}{2}$.

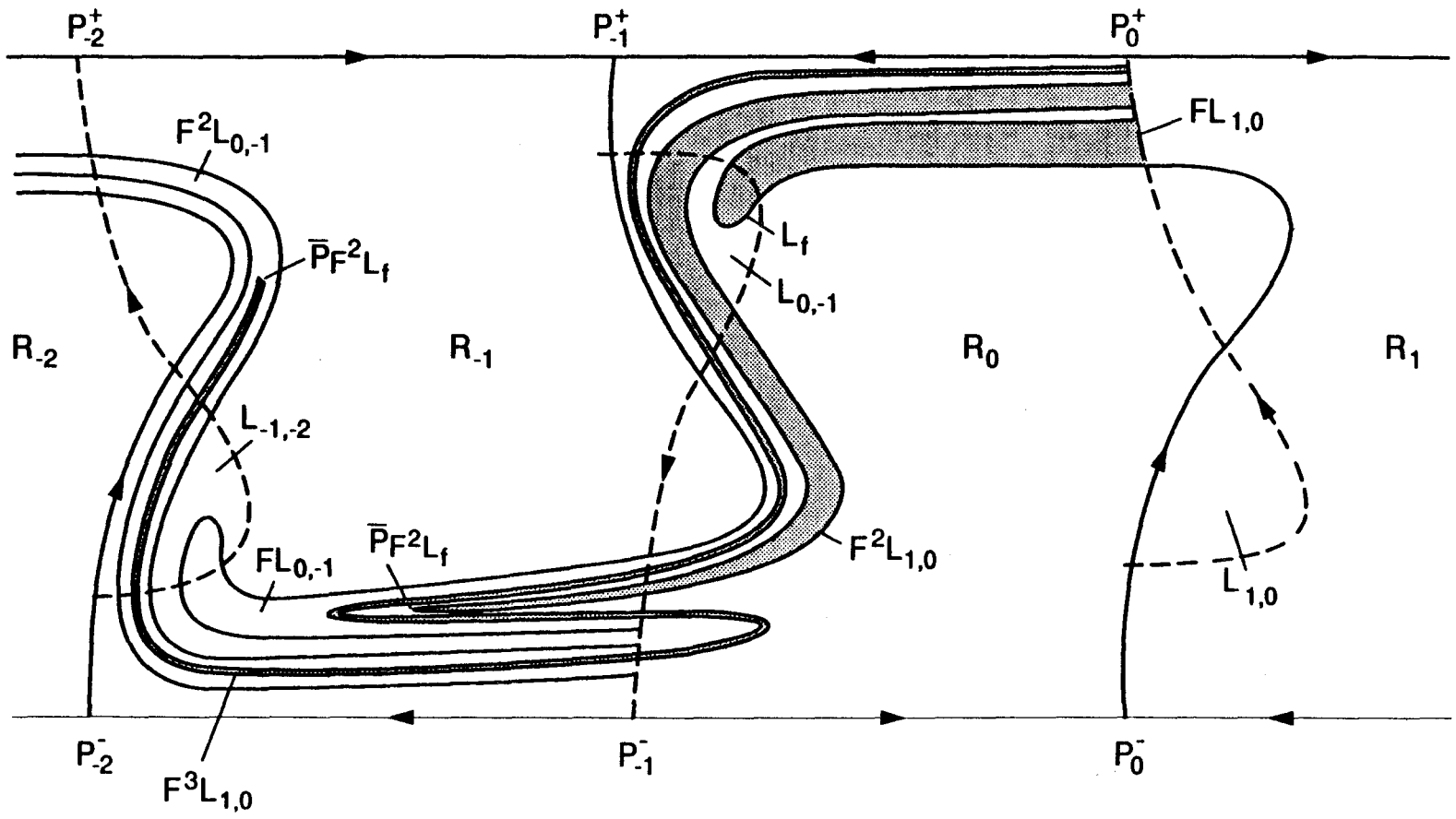

Figure 3.7. The stretching of $L_{1,0}$ for $\bar{m} = 1$, $\bar{m}'=2$.

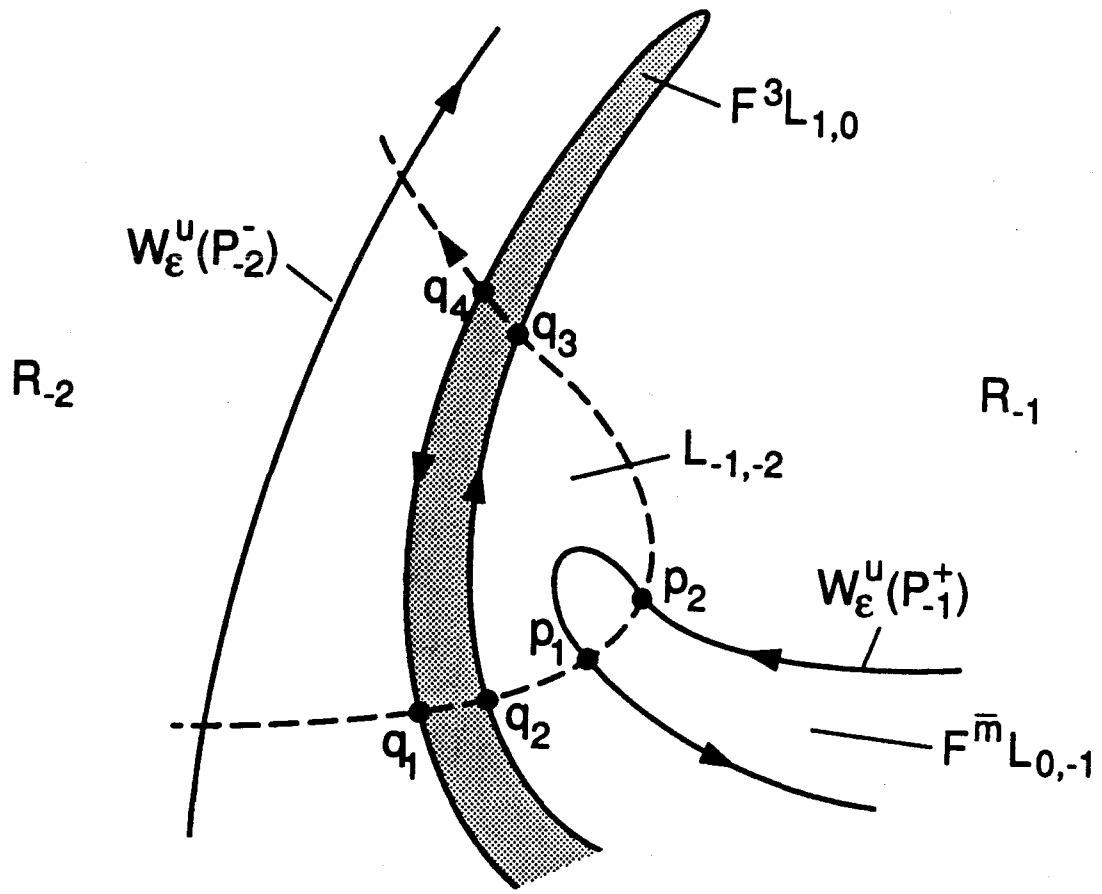

Figure 3.8. The tip of $F^{\bar{m}+\bar{m}'}L_{1,0} = F^3L_{1,0}$ completely "pushes" through $L_{-1,-2}$. The points q_1, q_2 are preceding p_1, q_3 and q_4 are following p_2 according to the arc length measure of $W^u_\epsilon(p^+_{-2})$.

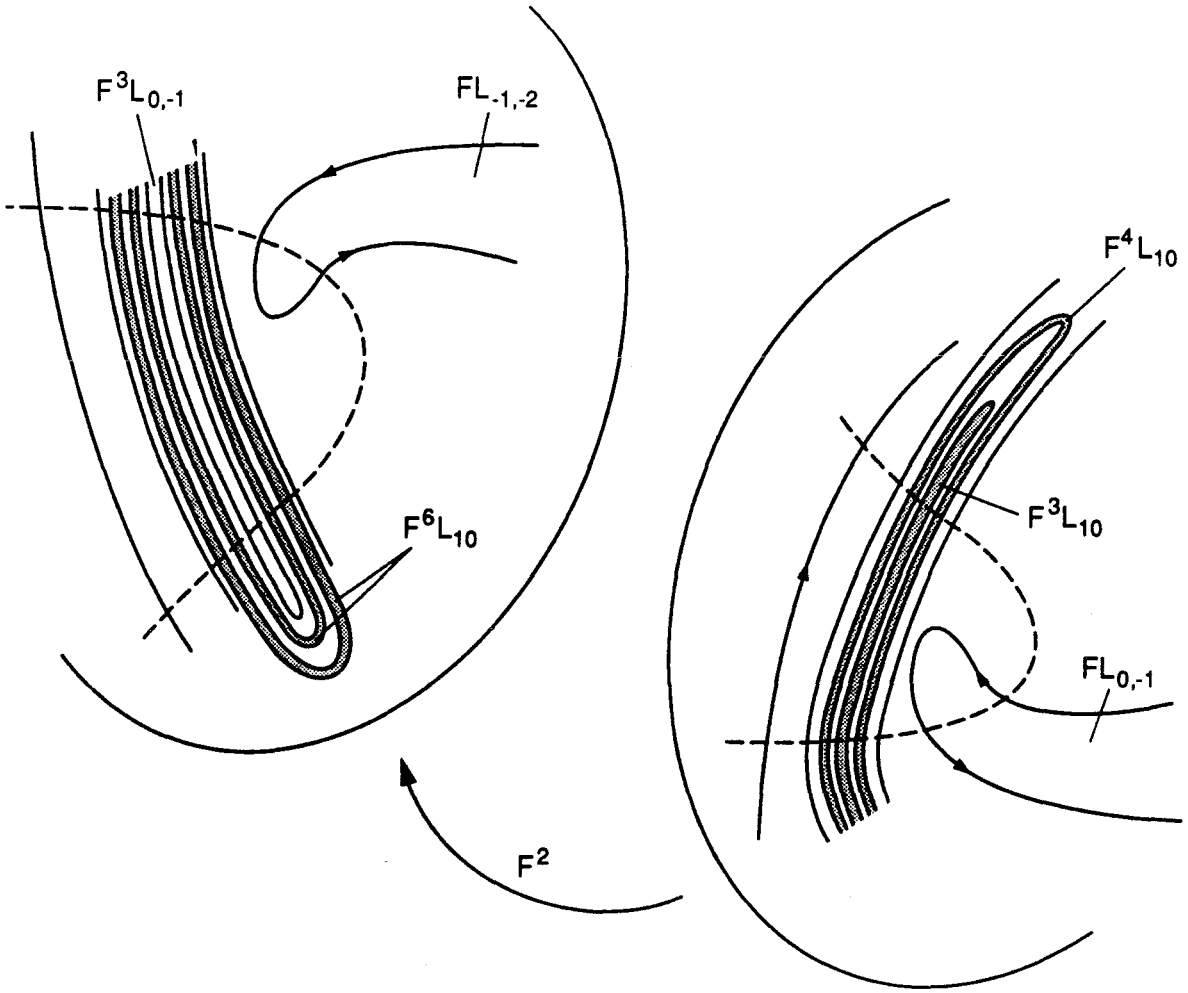

Figure 3.9. The action of the $\bar{m}'$-th iterate of F on the "trapped" part of $F^{\bar{m}+\bar{m}'+1}L_{1,0}$ for the case $\bar{m} = 1$, $\bar{m}' = 2$.

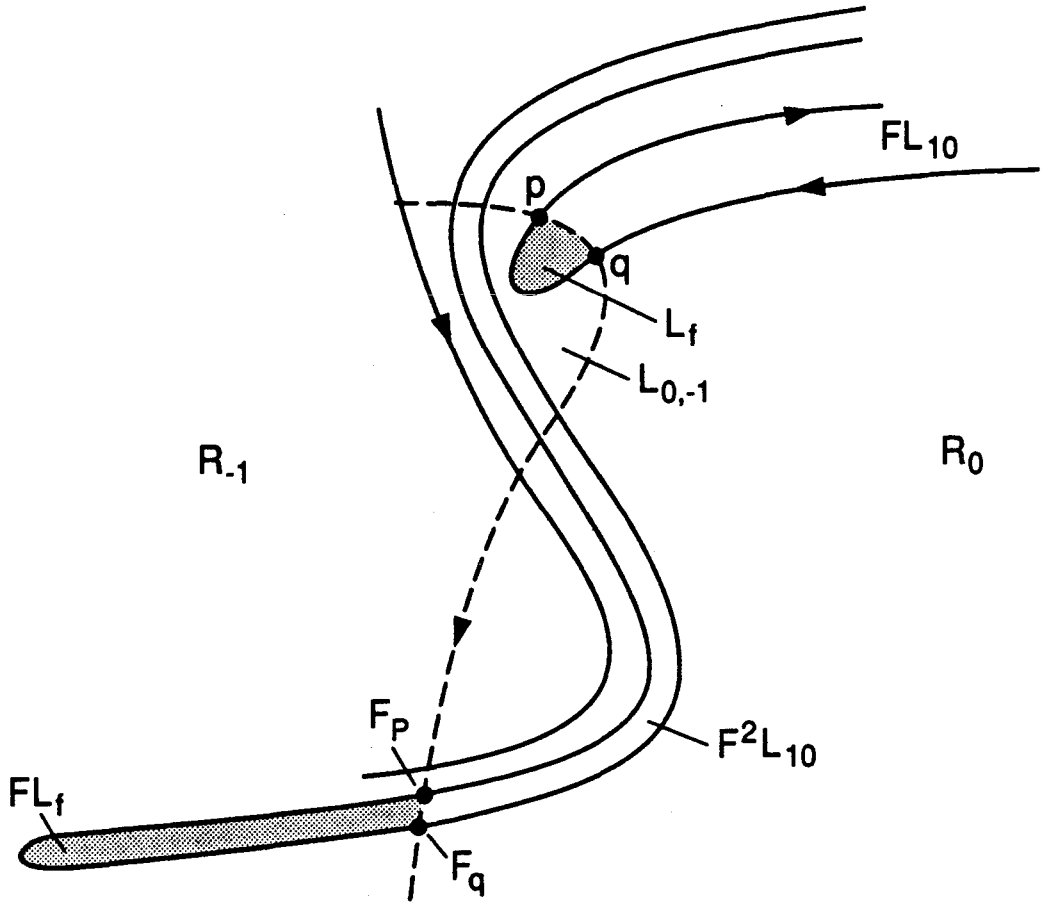

Figure 3.10. Intersection of $F^{\bar{m}}L_{1,0} = FL_{1,0}$ with $FL_{0,-1}$. At the next iteration, $\bar{m}+1 = 2$, the "tip" (shaded region) of $FL_{1,0}$ enters region R_1.

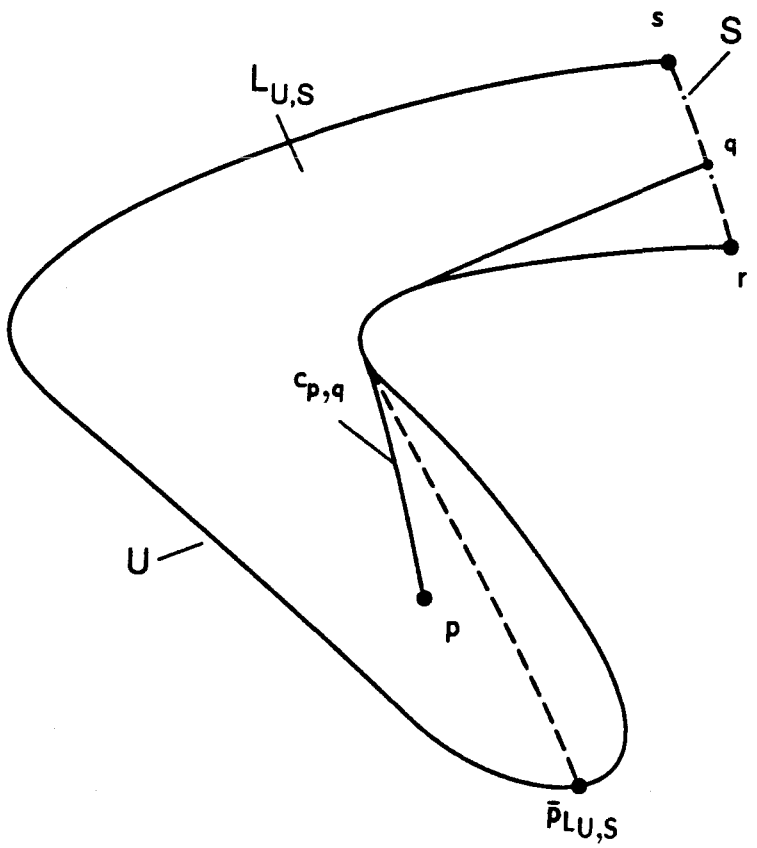

Figure 3.11. The definition of distance for a region bounded by a segment S of stable and U of unstable manifold between two intersection points s, r.

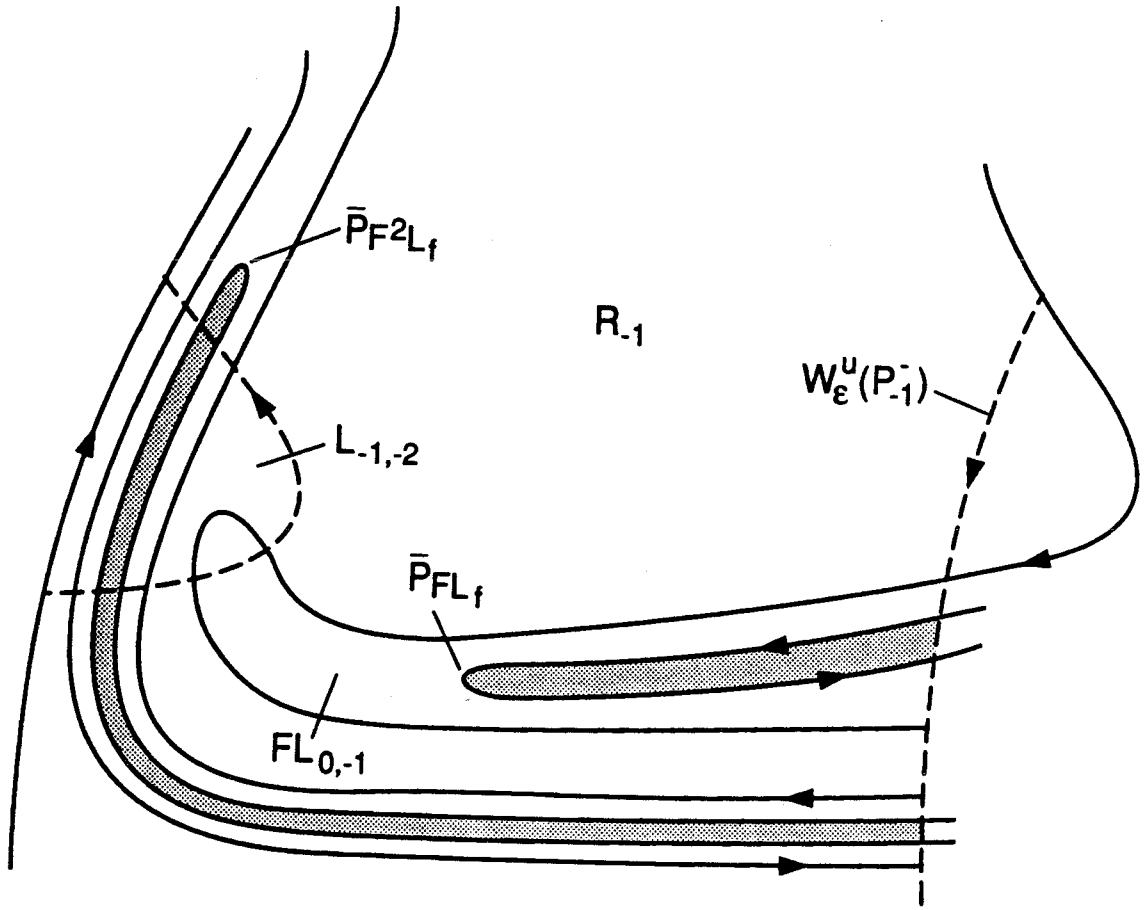

Figure 3.12. The points $\bar{p}_{FL_f}$, $\bar{p}_{F^2L_f}$ of maximum distance of FL_f, F^2L_f from the stable manifold $W^s_\epsilon(p^-_{-1})$.

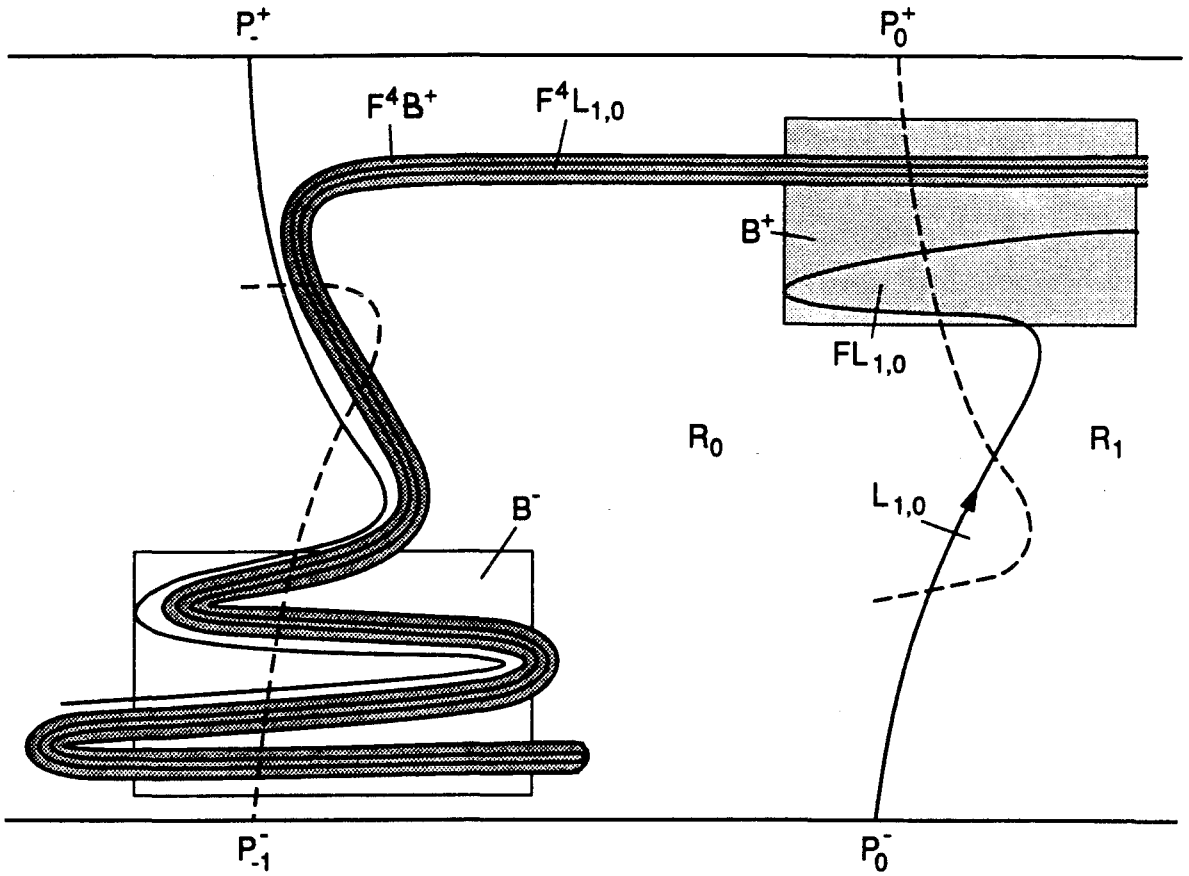

Figure 3.13. The evolution of the rectangular region B^+. The unstable manifold "drives" the stretching and folding of B^+ until it intersects the lower box B^-. B^- would follow a similar evolution, thus mapping a part of B^+ back onto itself.

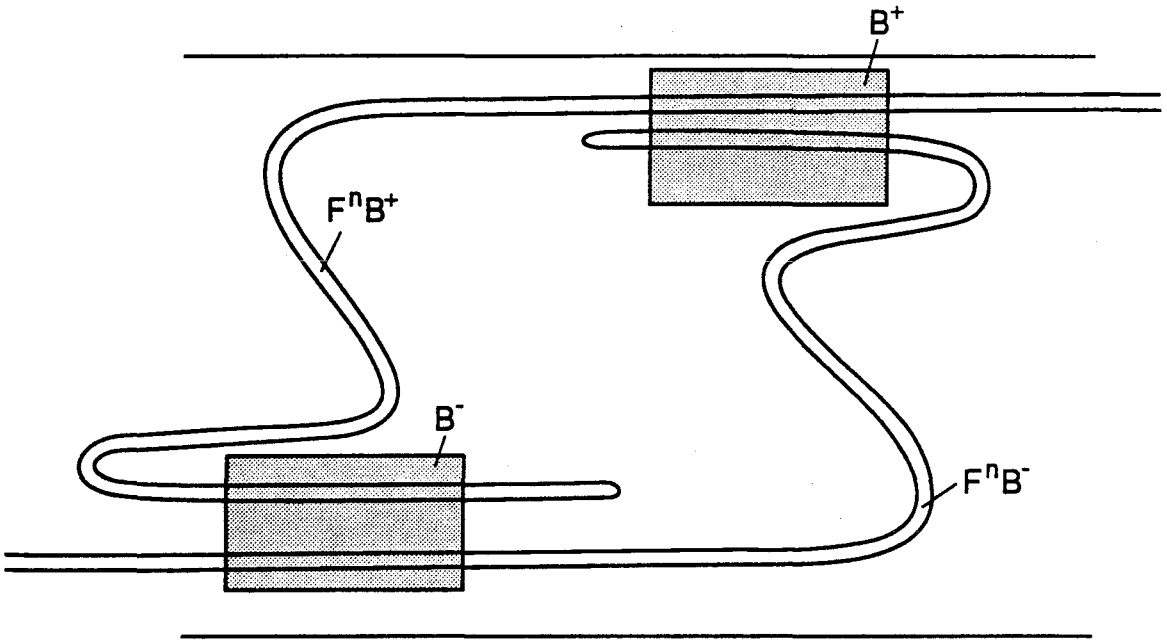

Figure 3.14. Schematic diagram showing the geometry of the mapping of the two rectangular regions B^+ and B^- onto each other.

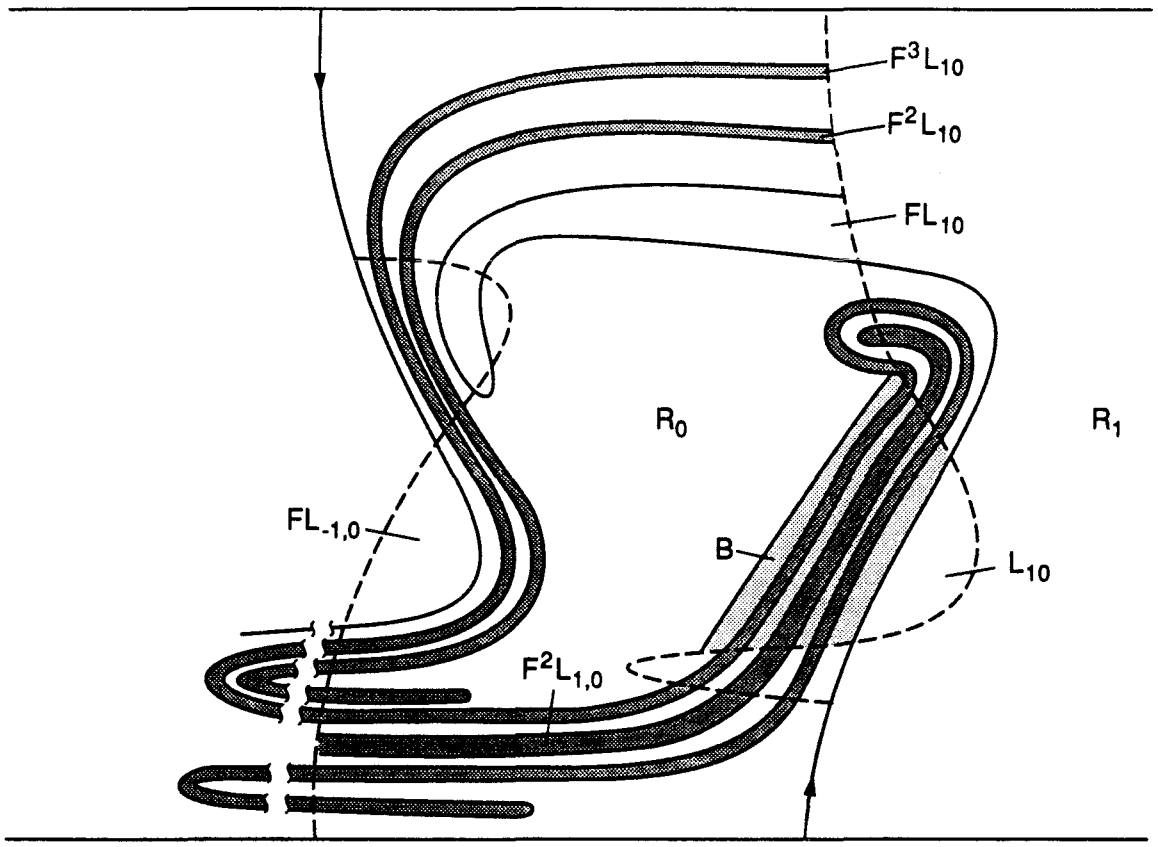

Figure 3.15. The evolution of the rectangular region B for the internal horseshoe. The unstable manifold "drives" the stretching and folding of B until it intersects itself.

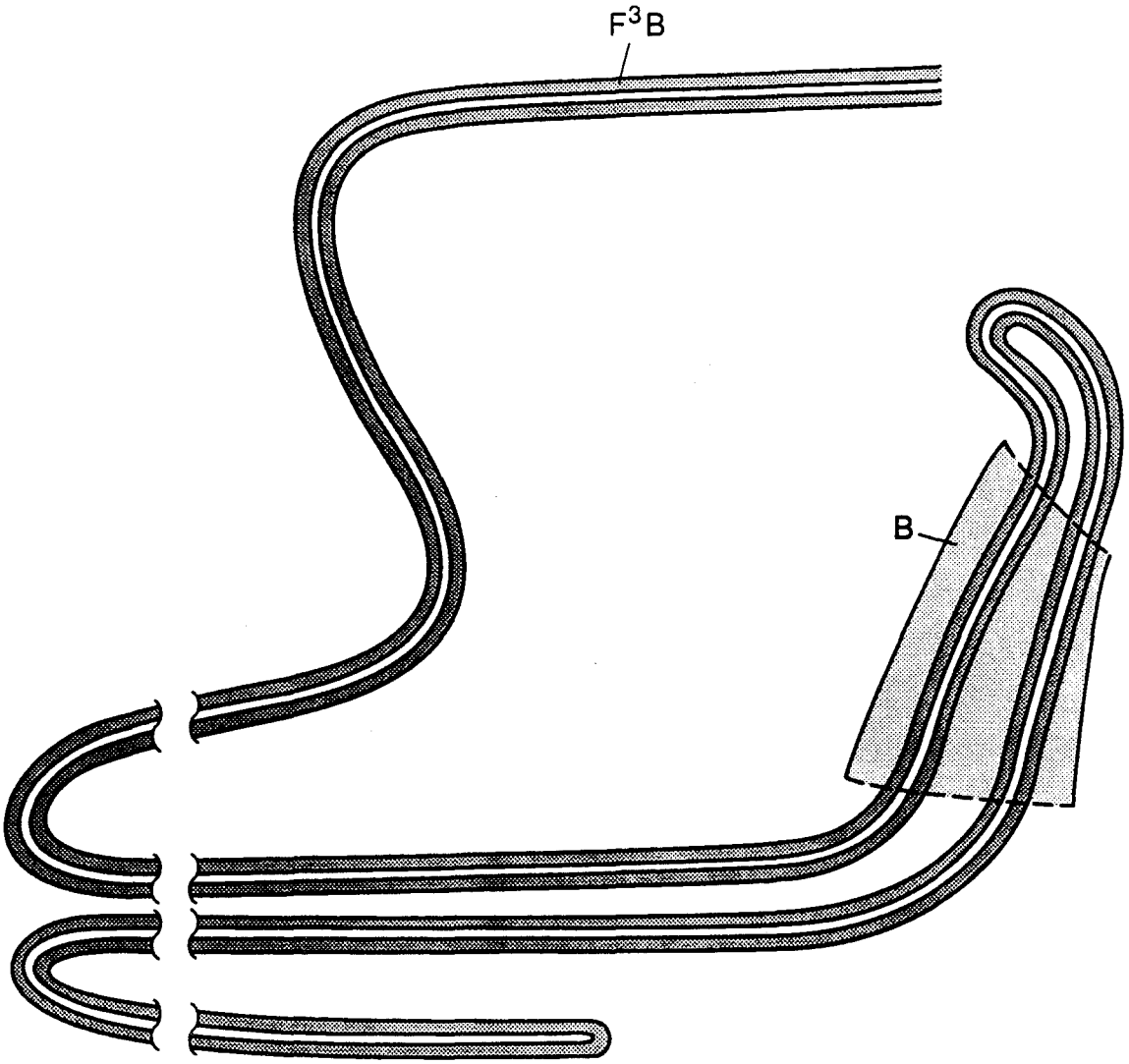

Figure 3.16. Schematic diagram showing the geometry of the mapping of the rectangular region B onto itself.

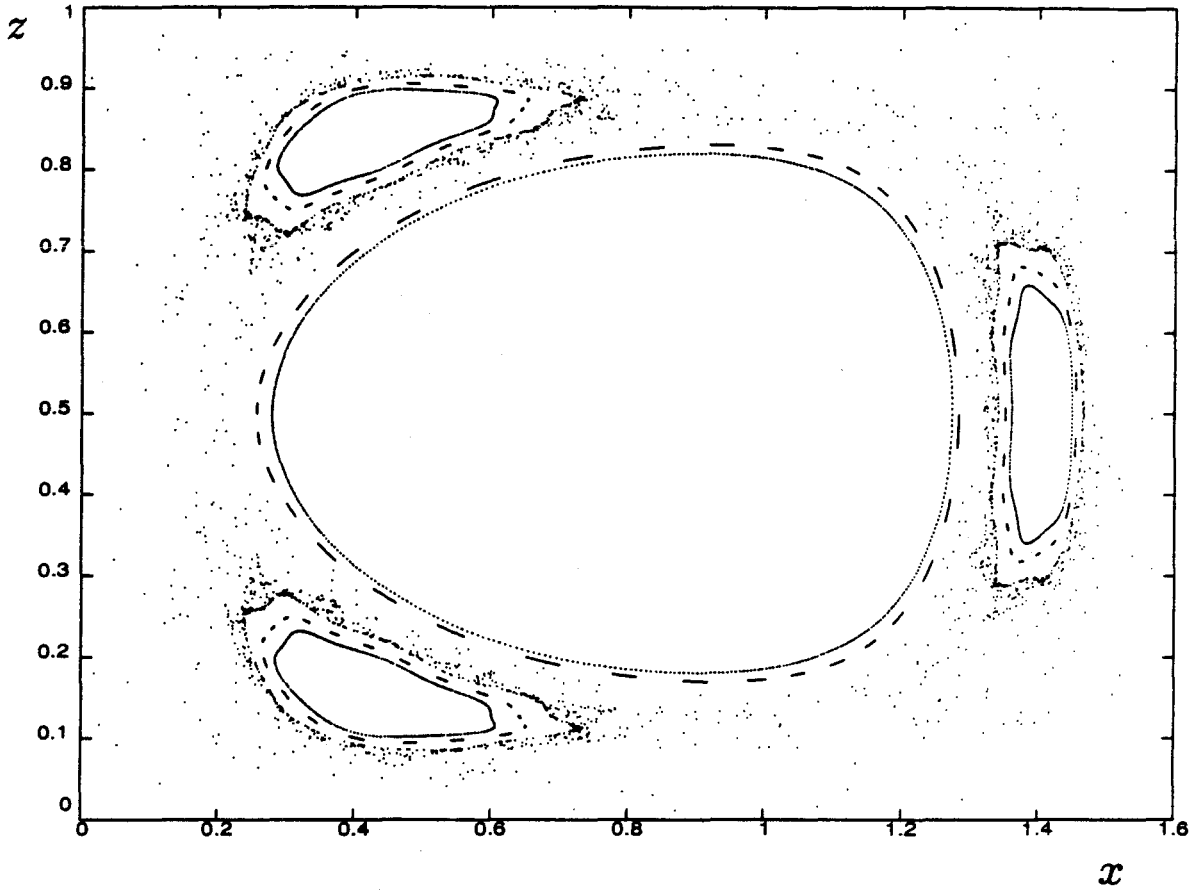

Figure 3.17. The 3:1 resonance band for $\omega = 0.6$, $\epsilon = 0.1$ and $A = 0.1$.

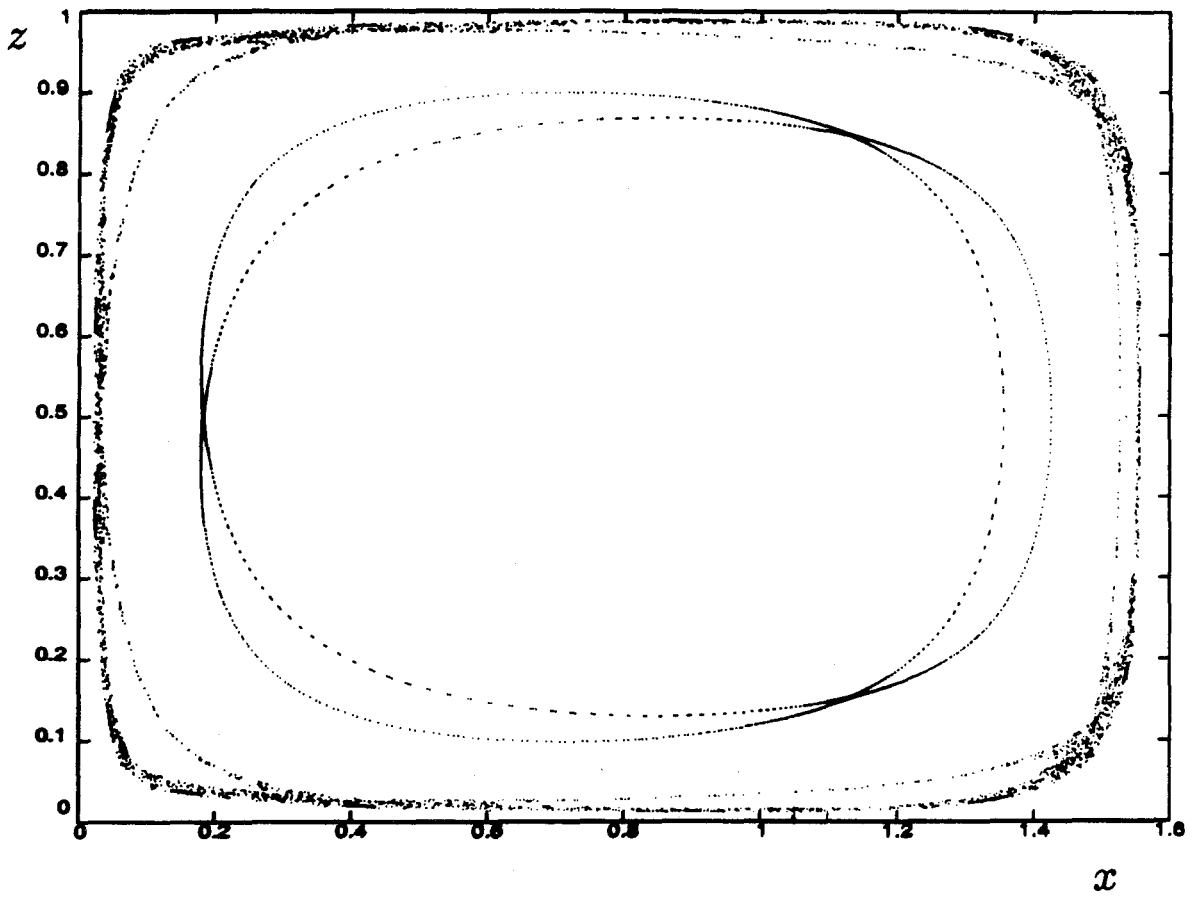

Figure 3.18. The 5:1 and 3:1 resonance bands for $\omega = 0.6$, $\epsilon = 0.01$ and $A = 0.1$.

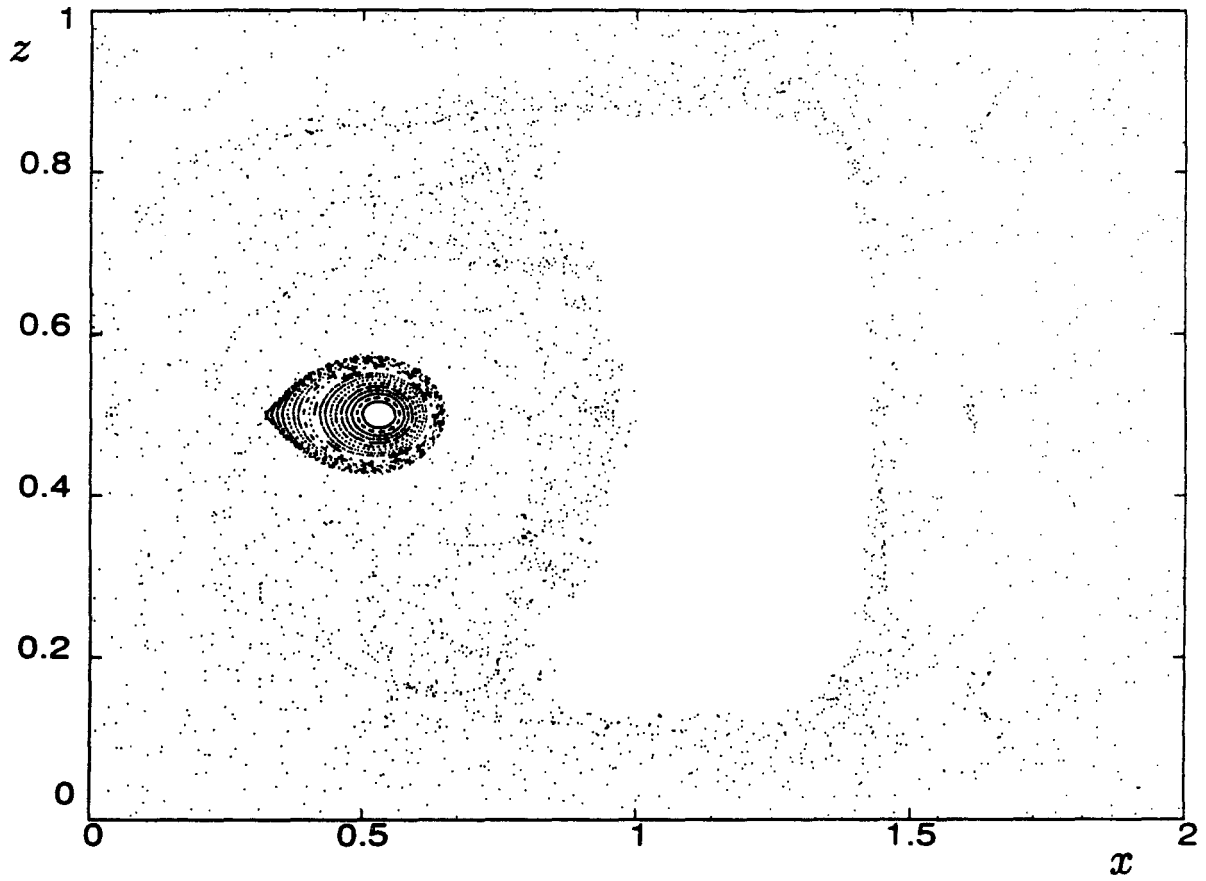

Figure 3.19. The 1:1 resonance band for $\omega = 0.24$, $\epsilon = 0.1$ and $A = 0.1$.

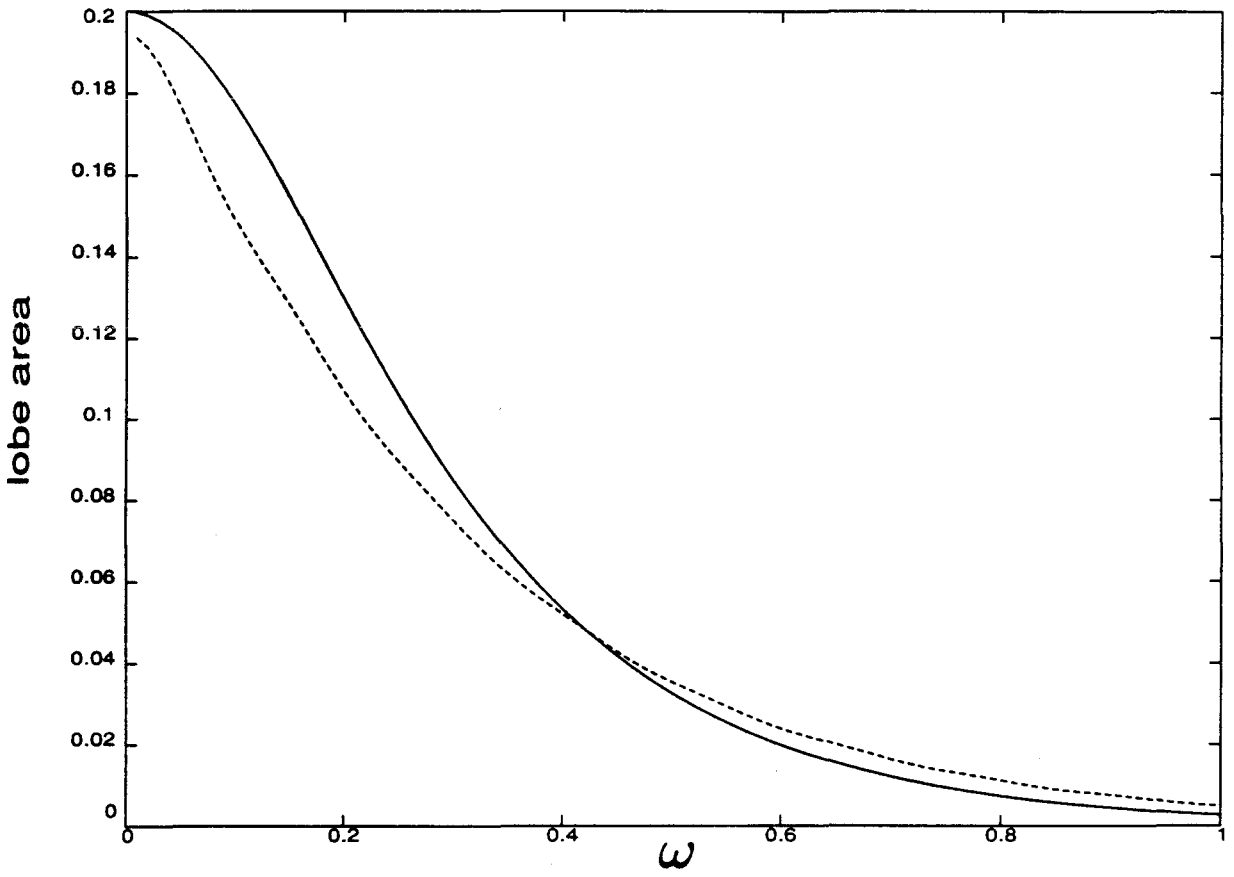

Figure 3.20. The lobe area $\mu(L_{1,0})$ for $\omega \in (0, 1]$ using the slip (solid) and non-slip (dashed) boundary conditions, with $\epsilon = 0.1$ and $A = 0.1$.

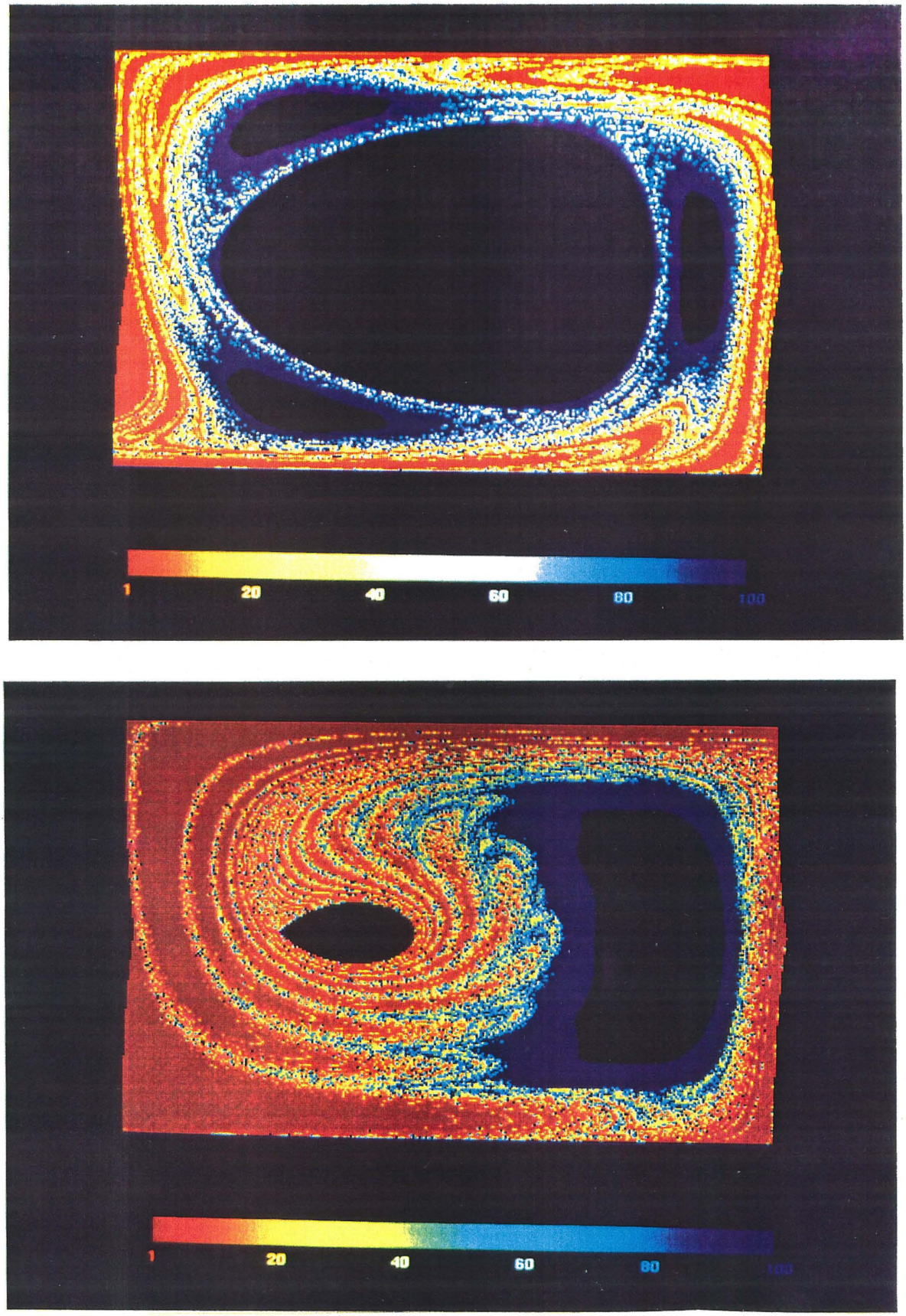

Figure 3.21. The distribution of escaping times for the fluid particles in a roll from $n = 1$ (red) to $n = 100$ (blue). For the black areas $n = \infty$. a) $\epsilon = 0.1$, $\omega = 0.6$ and $A = 0.1$. b) $\epsilon = 0.1$ $\omega = 0.24$ and $A = 0.1$.

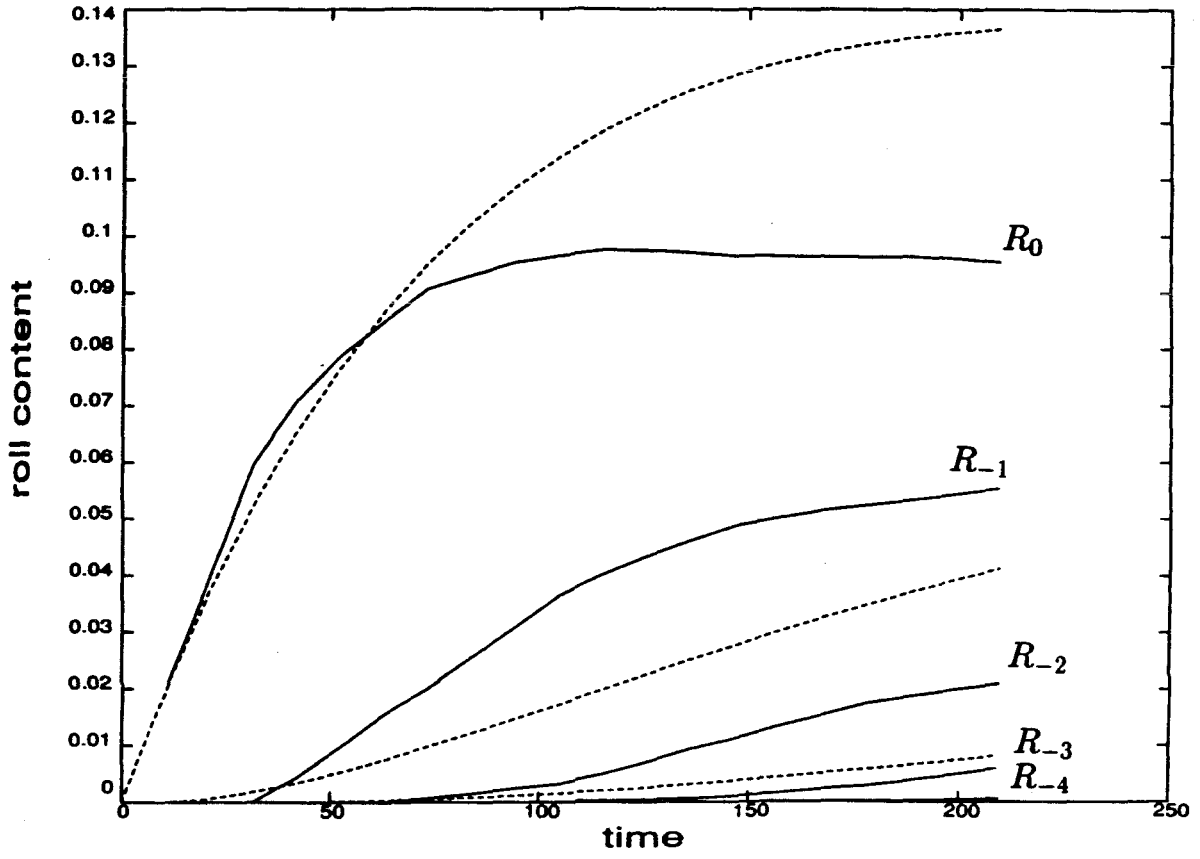

Figure 3.22. Comparison between the exact result (solid) and the Markov model prediction (dashed) for the j-th roll content of R_1-species vs. time, $j = 0, \ldots, -4$, with $\epsilon = 0.1$ $\omega = 0.6$ and $A = 0.1$.

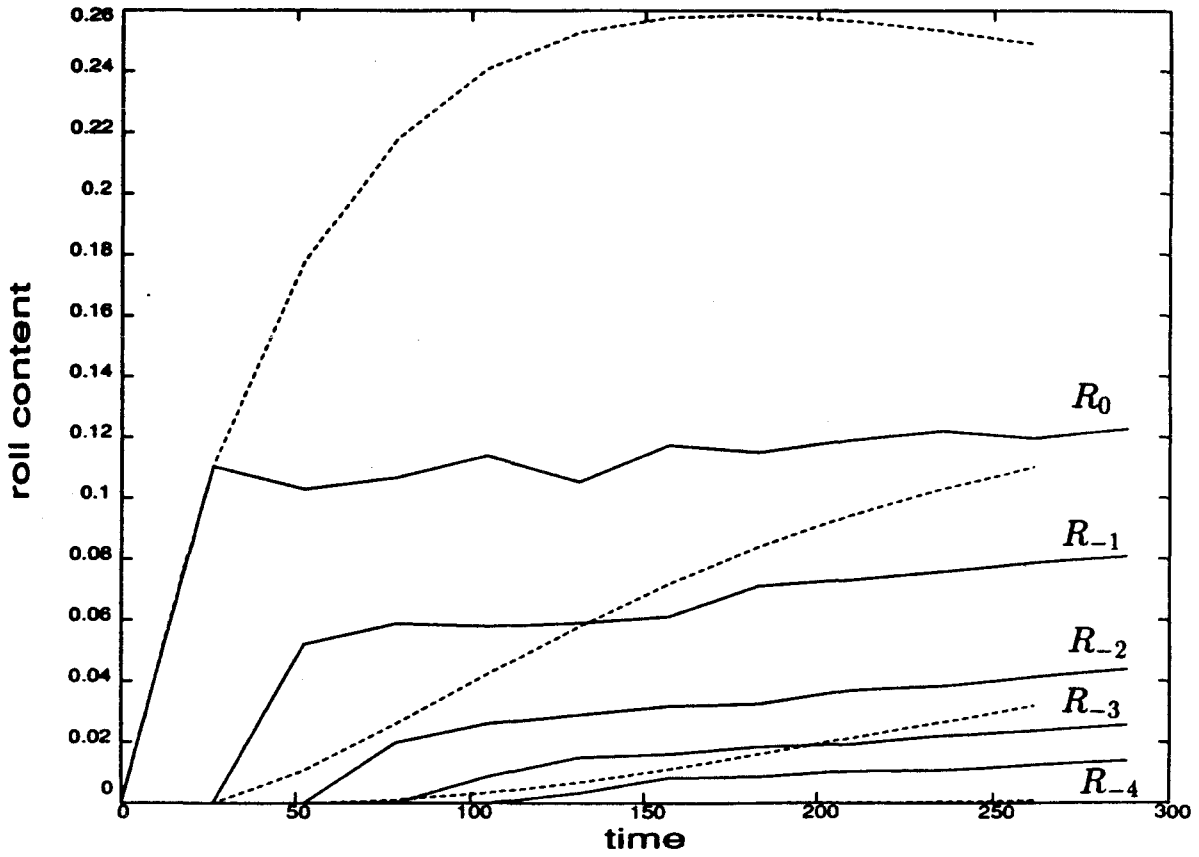

Figure 3.23. Comparison between the exact result (solid) and the Markov model prediction (dashed) for the j-th roll content of R_1-species vs. time, $j = 0, \ldots, -4$, with $\epsilon = 0.1$ $\omega = 0.24$ and $A = 0.1$.

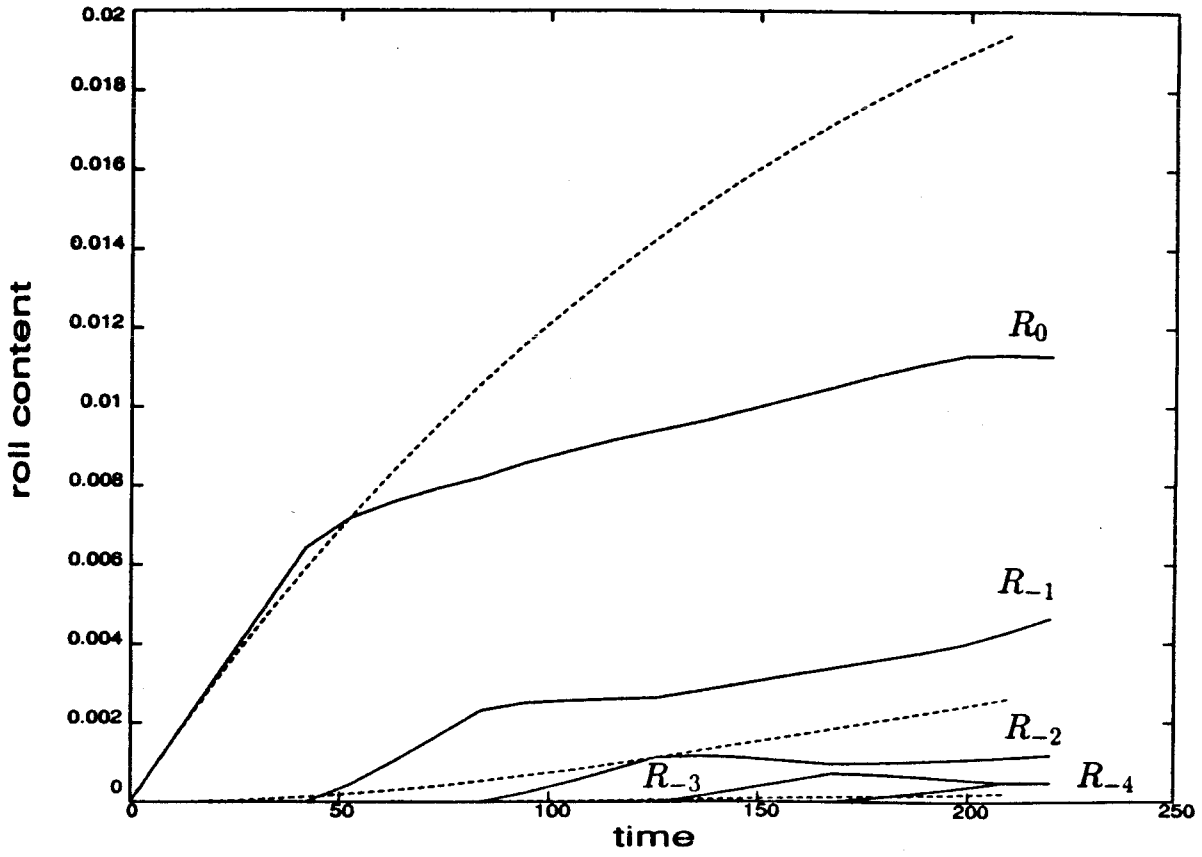

Figure 3.24. Comparison between the exact result (solid) and the Markov model prediction (dashed) for the j-th roll content of R_1-species vs. time, $j = 0, \ldots, -4$, with $\epsilon = 0.01$ $\omega = 0.6$ and $A = 0.1$.

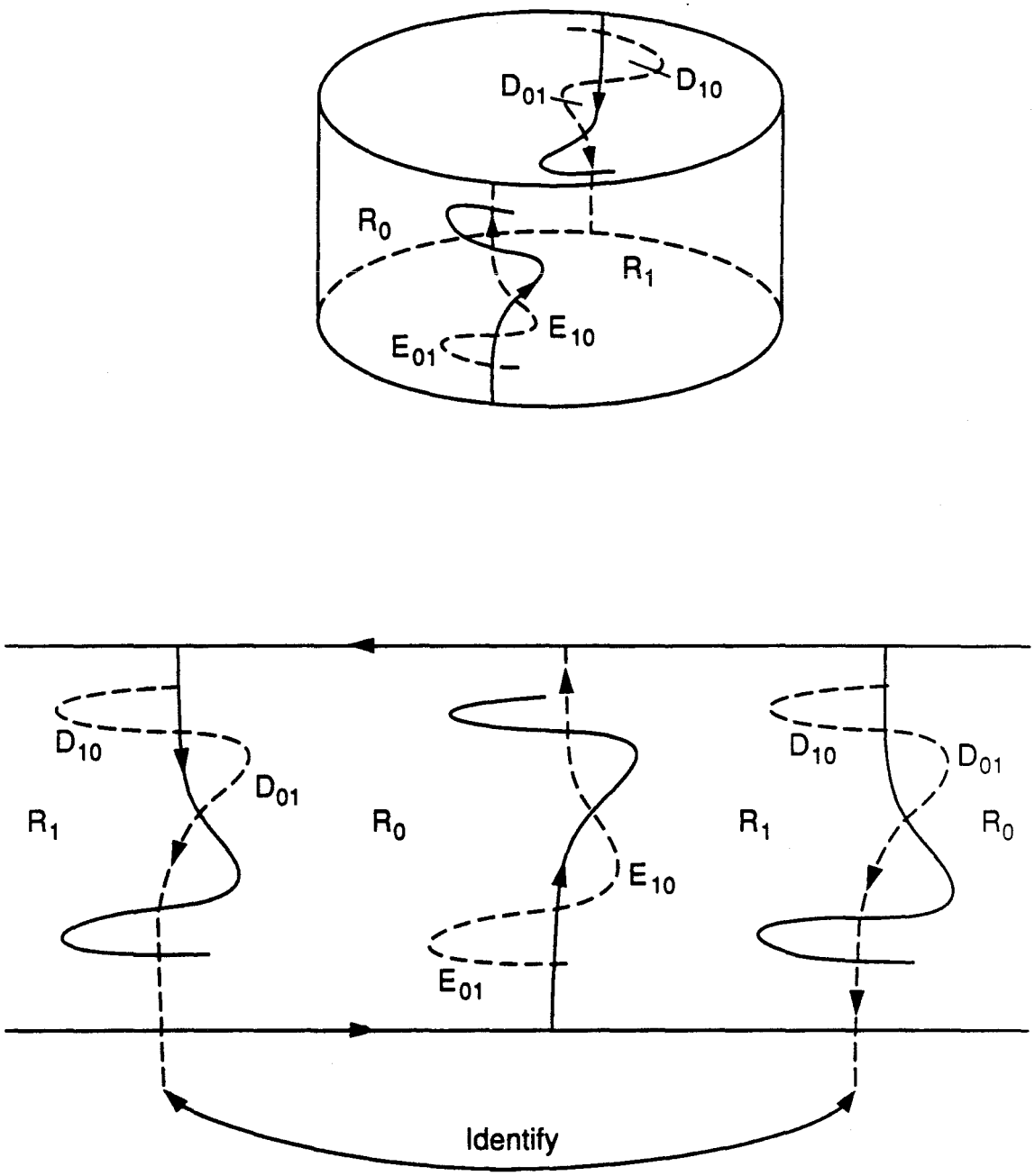

Figure 3.25. The manifolds and regions for the vector field (3.2) on a cylinder

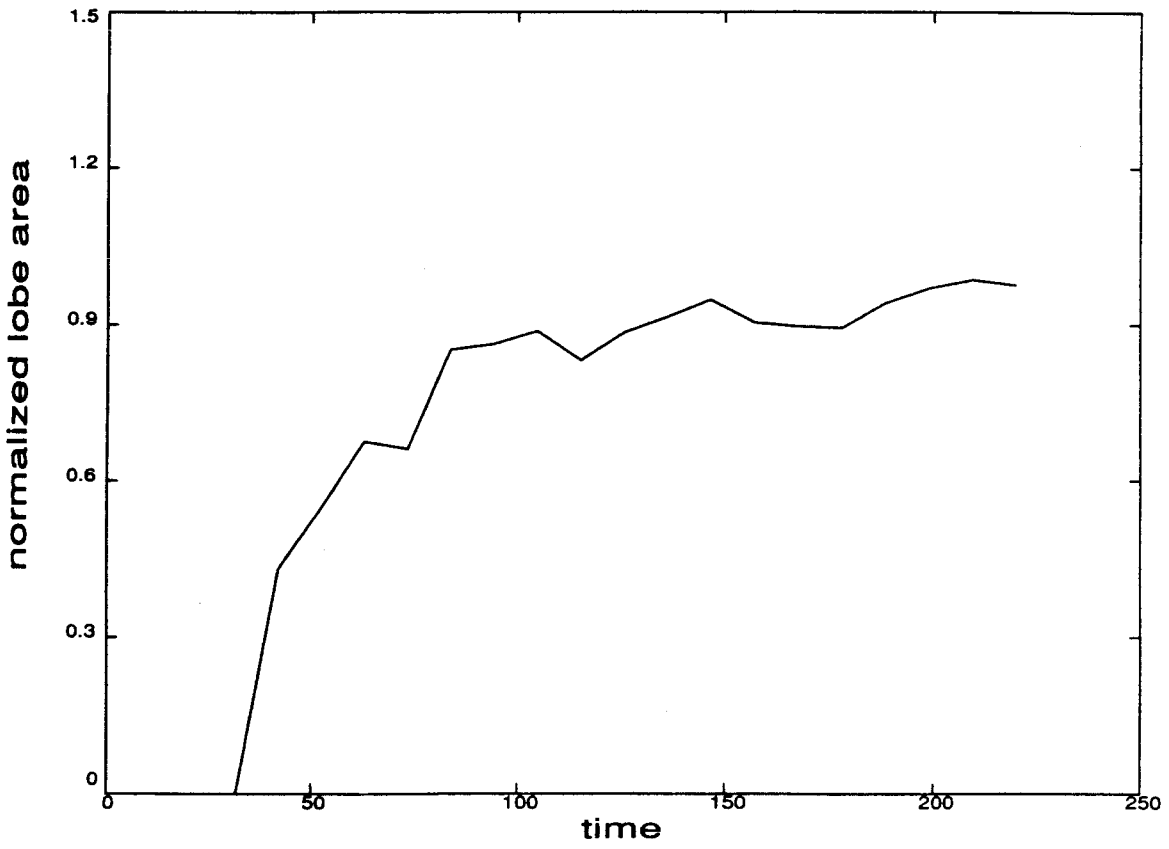

Figure 3.26. The behaviour of the partial sums of the series (3.53), $\epsilon = 0.1$, $\omega = 0.6$ and $A = 0.1$.

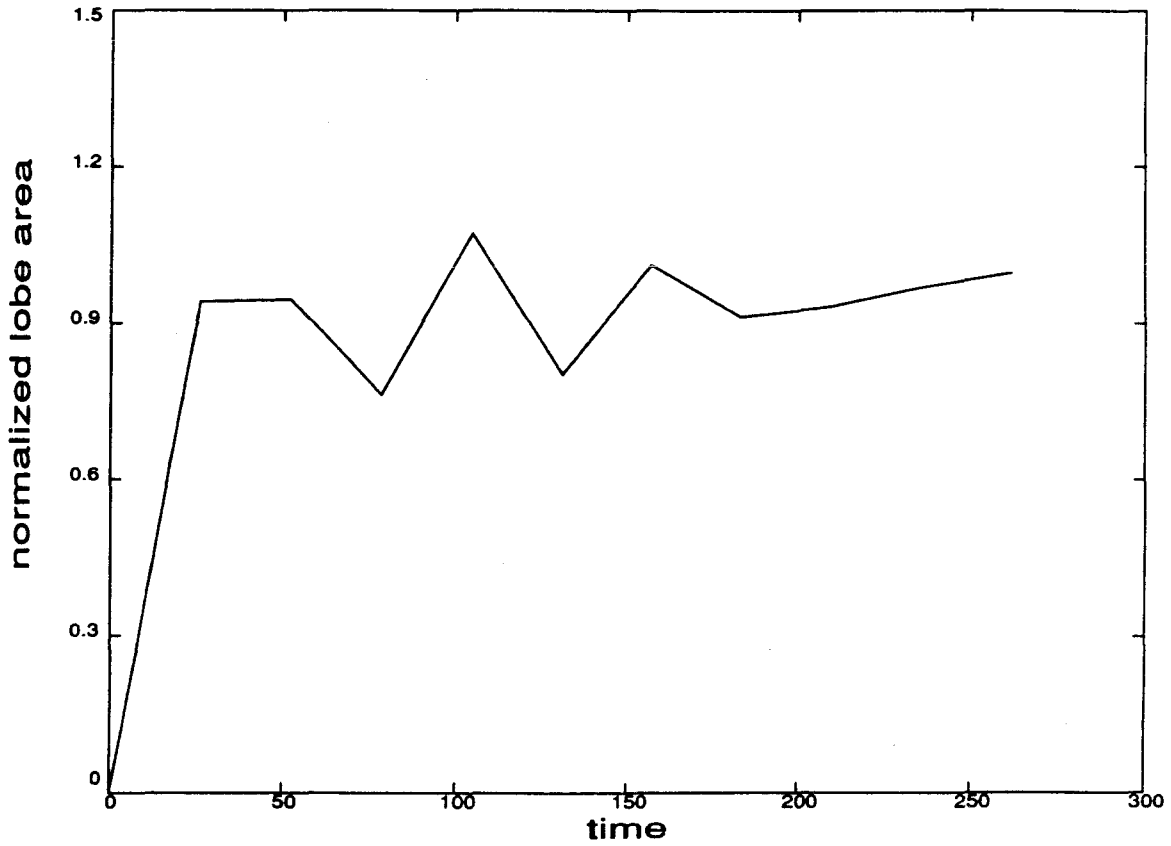

Figure 3.27. The behaviour of the partial sums of the series (3.53), $\epsilon = 0.1$, $\omega = 0.24$ and $A = 0.1$.

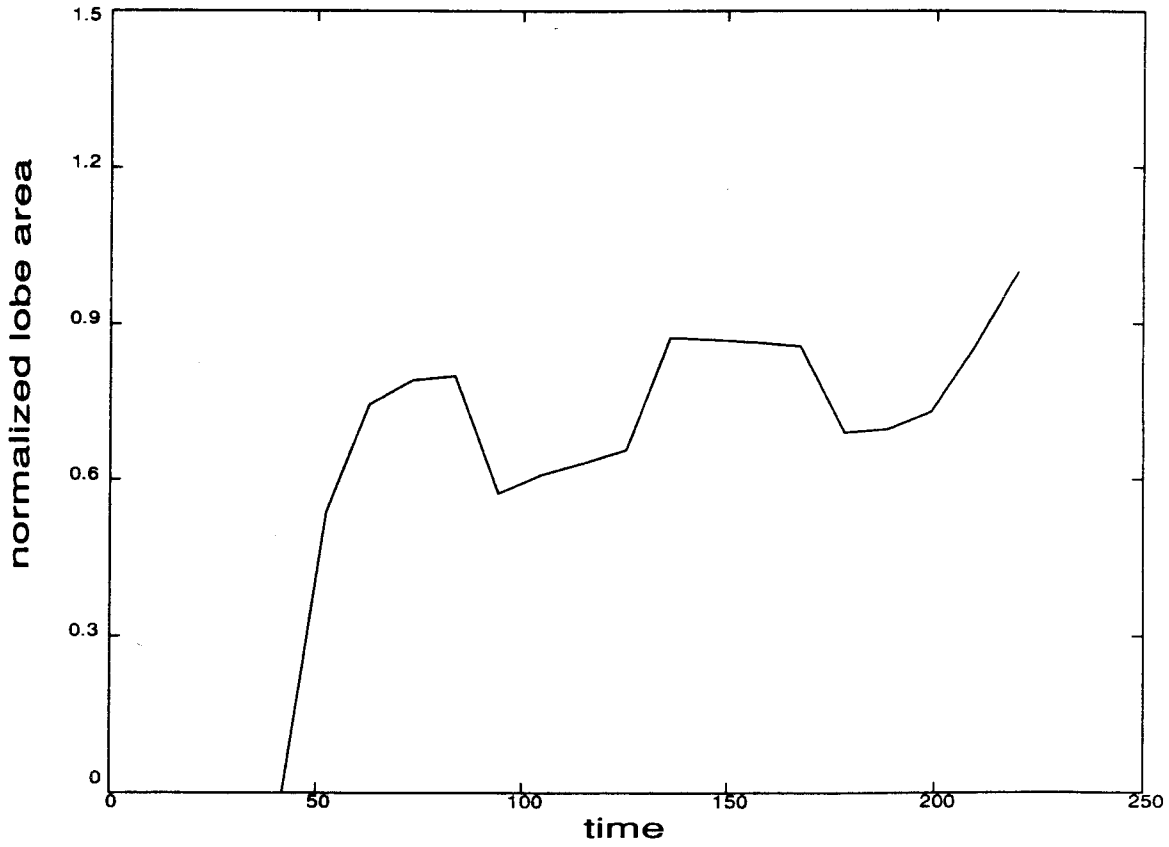

Figure 3.28. The behaviour of the partial sums of the series (3.53), $\epsilon = 0.01$, $\omega = 0.6$ and $A = 0.1$.

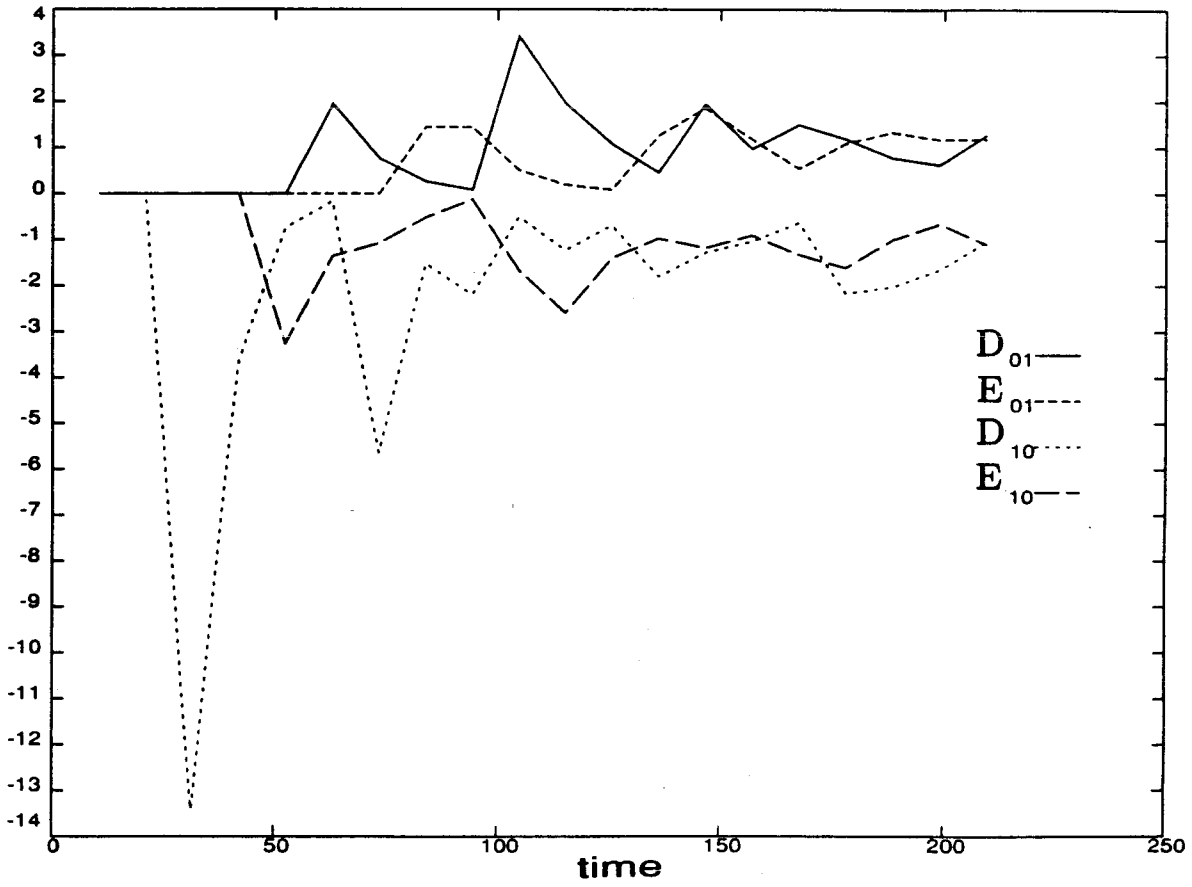

Figure 3.29. The areas of intersections of $E_{1,0}$ with each of the turnstile lobes, normalized with respect to $2r_T/(\mu(E_{1,0}))^2$, $\epsilon = 0.1$, $\omega = 0.6$ and $A = 0.1$.

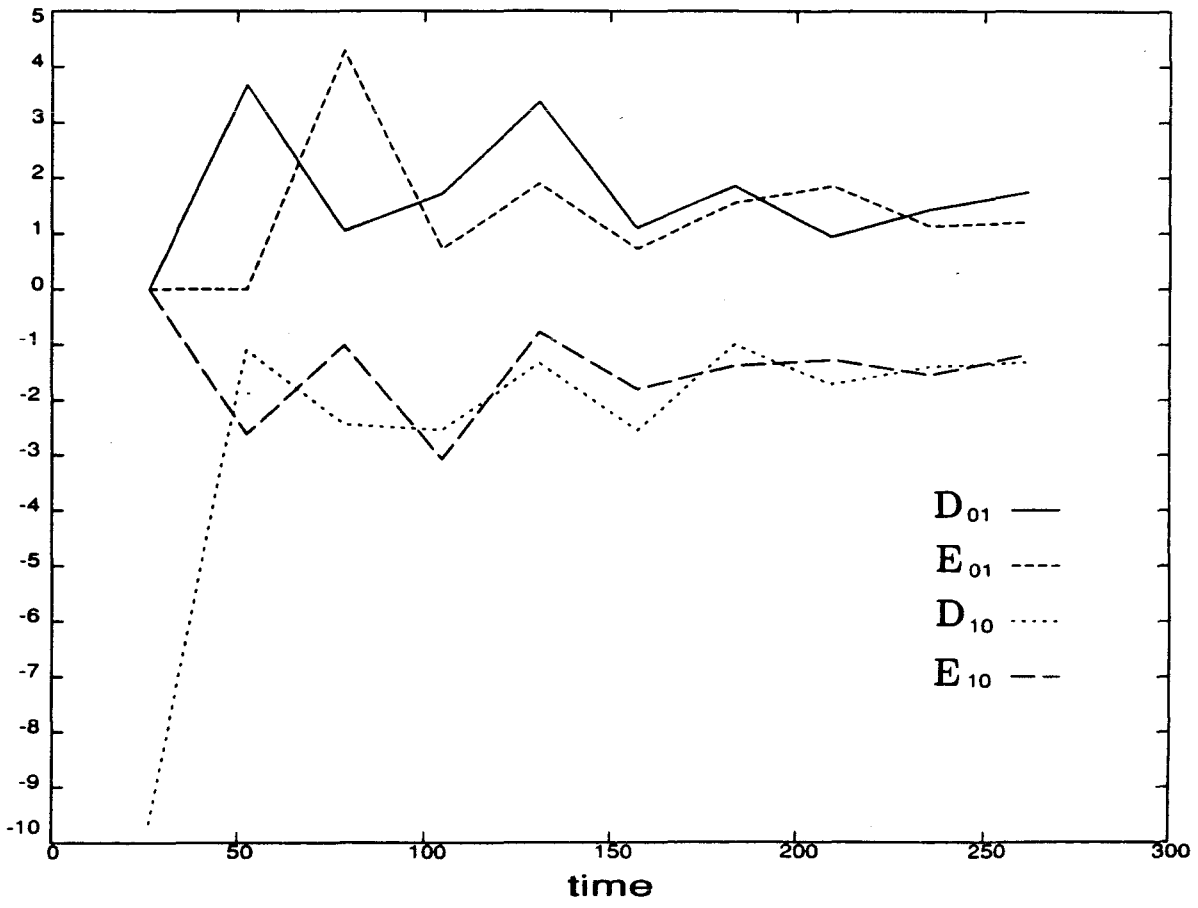

Figure 3.30. The areas of intersections of $E_{1,0}$ with each of the turnstile lobes, normalized with respect to $2r_T/(\mu(E_{1,0}))^2$, $\epsilon = 0.1$, $\omega = 0.24$ and $A = 0.1$.

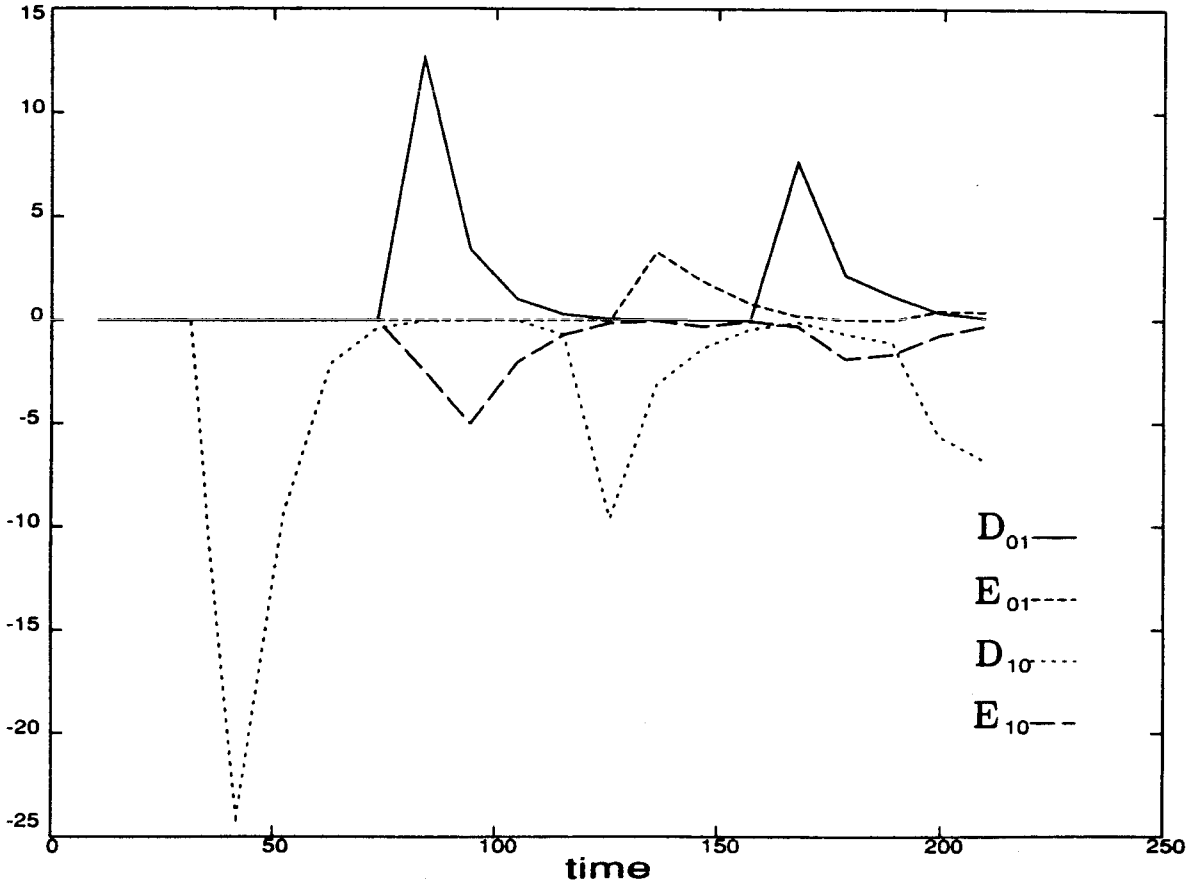

Figure 3.31. The areas of intersections of $E_{1,0}$ with each of the turnstile lobes, normalized with respect to $2r_T/(\mu(E_{1,0}))^2$, $\epsilon = 0.01$, $\omega = 0.6$ and $A = 0.1$.

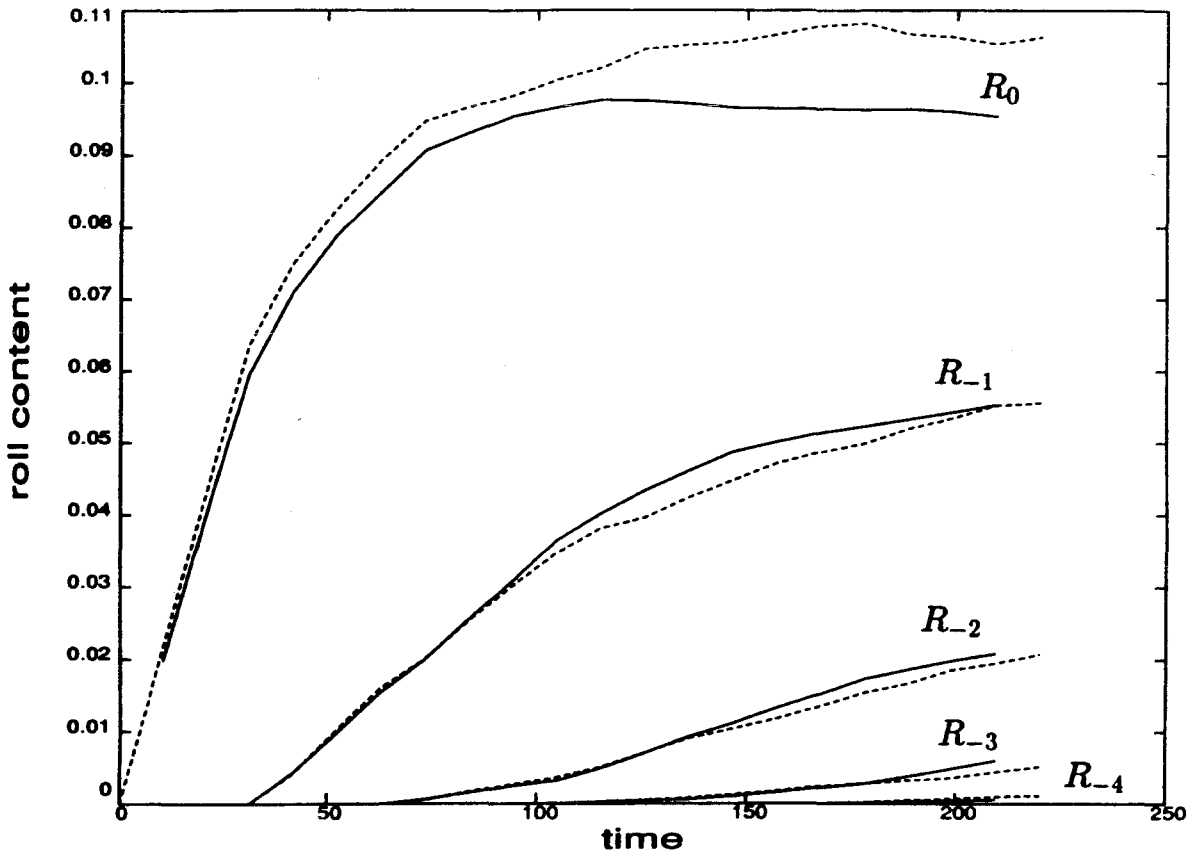

Figure 3.32. Comparison between the exact results (solid) and the ones simulating numerical diffusivity (dashed) for the j-th roll content of R_1-species vs. time, $j = 0, \ldots, -4$, with $\epsilon = 0.1$ $\omega = 0.6$ and $A = 0.1$.

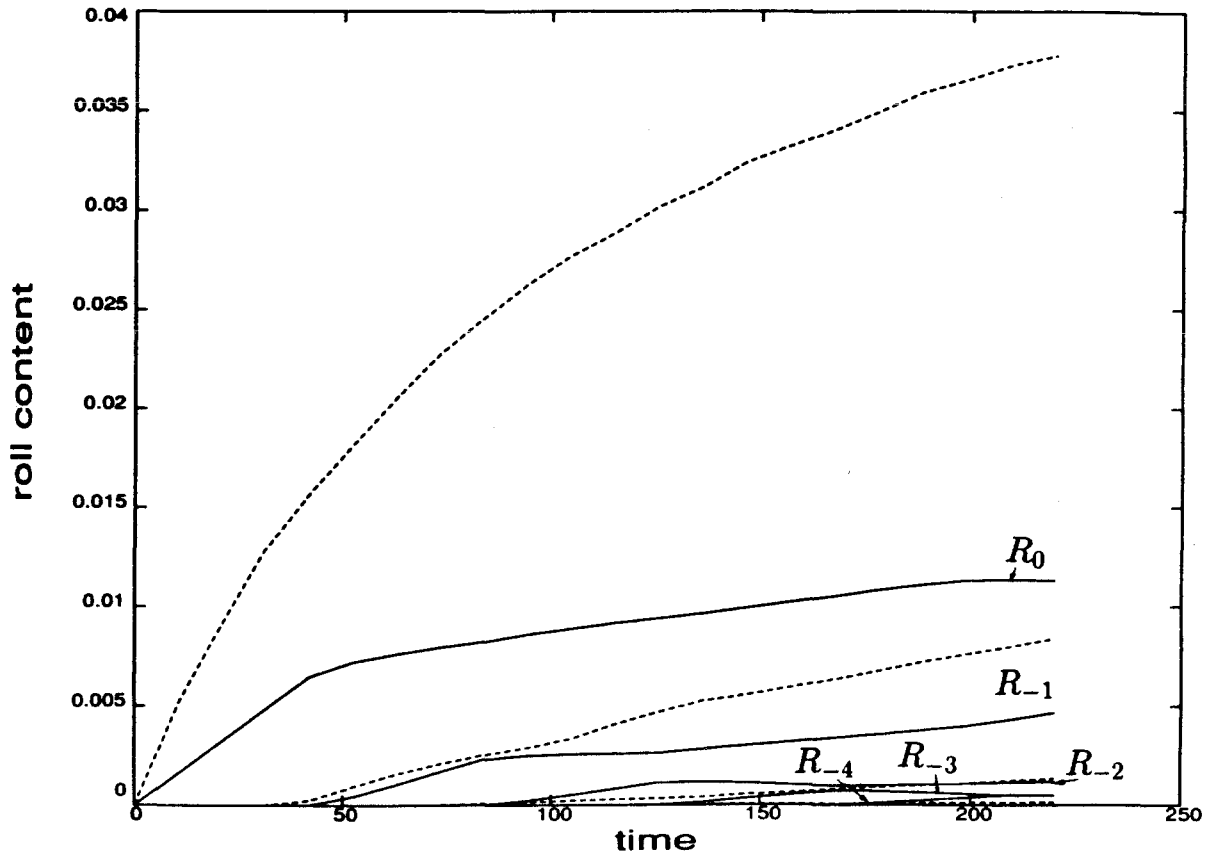

Figure 3.33. Comparison between the exact results (solid) and the ones simulating numerical diffusivity (dashed) for the j-th roll content of R_1-species vs. time, $j = 0, \ldots, -4$, with $\epsilon = 0.01$ $\omega = 0.6$ and $A = 0.1$.