Internal Waves and the Poincaré Equation
Numerical Computation and Laboratory Experiments

Interne Golven en de Poincaré Vergelijking
Numerieke Berekening en Laboratorium Experimenten

(met een samenvatting in het Nederlands)

Proefschrift

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE UNIVERSITEIT UTRECHT OP GEZAG VAN DE RECTOR MAGNIFICUS, PROF. DR. W. H. GISPEN, INGEVOLGE HET BESLUIT VAN HET COLLEGE VOOR PROMOTIES IN HET OPENBAAR TE VERDEDIGEN OP WOENSDAG 11 JULI 2007 DES MIDDAGS TE 12:45 UUR

door

Adriaan Nicolaas Swart

geboren op 30 september 1977, te Woudrichem

Promotoren:	Prof. dr.	H. A. van der Vorst
	Prof. dr.	Leo R. M. Maas
Co-promotor:	Dr.	Gerard L. G. Sleijpen

ISBN: 978–90–393–4570–2

Contents

Chapter 1

Introduction

1.1 Overview of the thesis

This thesis deals with the so called Poincaré equation. This equation has great importance in the field of fluid mechanics, as it models the behaviour of waves in fluid volumes: internal waves. Internal waves, in turn, are of interest to oceanographers studying oceanic processes, or biologists, who are interested in nutrients that are brought into suspension by internal waves breaking on a sloping bottom. Mathematically, solving the Poincaré equation on a bounded domain is interesting because it consists of the unusual combination of a hyperbolic partial differential equation with a bounded domain, on which the equation is to be solved. It is the combination of Poincaré's equation with a bounded domain that is challenging, so actually the Poincaré boundary value problem would be more fit. However, we will use the term Poincaré equation throughout this thesis to indicate the boundary value problem.

We will be particularly interested in domains that lack symmetries, since we know that highly symmetrical domains like the rectangle or circle are in some sense exceptional cases (see Section 1.4). In this thesis we will put forth numerical methods for efficient approximation of solutions to the Poincaré equation, set in two spatial dimensions. We believe such methods are of importance, since previously there has been a lack of numerical methods that take into account the rather special properties of the Poincaré equation. The reader will find an overview of previously proposed numerical methods in Chapter 2. In this introduction we will briefly outline the contents of this thesis, after which we discuss the role of internal waves in the field of geophysical fluid dynamics. This overview also serves as context for Chapter 4, where we report on a laboratory experiment aimed at detecting mixing processes caused by in-

ternal waves. We begin now by introducing the main equation of this thesis, Poincaré's equation. Later on in this chapter we will derive this equation from the Navier-Stokes equations.

Elliptic partial differential equations, such as the Poisson equation $\Delta u = f$ or Laplace equation $\Delta u = 0$ are routinely solved and efficient numerical methods are available. Also, hyperbolic equations on half open domains, such as the wave equation $u_{xx} - u_{tt} = 0$, for $x \in [0, 1]$ and $t \in [0, \infty)$ are extensively studied. The Poincaré equation, however, has received less attention in mathematics and computational science.

The Poincaré boundary value problem may be posed as, find eigenfunctions $\Psi(x, z)$ and eigenvalues λ^2 on a bounded domain $D \subset \mathbb{R}^2$ that satisfy

$$\begin{aligned} \Psi_{xx} - \lambda^2 \Psi_{zz} &= 0 \qquad \text{in } D, \\ \Psi &= 0 \qquad \text{at } \partial D. \end{aligned}$$

The motivation for using the term 'Poincaré equation', instead of 'wave equation' stems from the fact that the behaviour radically differs when introducing a bounded domain.

Perhaps the most challenging property of the Poincaré problem is its ill-posedness. We will use the term ill-posedness to indicate that solutions to a problem do not vary continuously upon the parameters. The Poincaré problem is both ill-posed[1] and under determined. However, in order to fully specify the problem, extra boundary conditions may be posed on the 'fundamental intervals', certain subsets of the boundary of the domain (see Section 1.4). In Section 1.6 we will return to the subject of ill-posedness and discuss our definition in more detail.

Another noteworthy feature is the occurrence of a fractal structure in the solution. At certain values of λ parts of the solution will reproduce on increasingly smaller scales. The limiting orbit to which the fractal structure tends[2] is a closed path in the fluid domain, called the 'wave attractor'. A ray-traced image of a solution in terms of the stream function is shown in figure 1.1. The occurrence of this fine structure is another obstacle in the way of an accurate numerical approximation.

The thesis contains four parts, each studying a mathematical (Poincaré equation) or physical (internal waves) aspect. Firstly, the introduction will define the basic problem and discuss the mathematical peculiarities of the problem. Also, a brief overview of the field of fluid mechanics will be given, with special attention to the position of internal waves. We also present some examples

[1] This is a simplification, the Poincaré equation is not ill-posed for all parameter values. There often exist regimes where the equation is well-posed. We will return to this in detail in Chapter 2, but see also the figures in Section 1.7.4

[2] This limiting orbit can also consist of one point only, which then acts as a 'sink' to which waves are attracted.

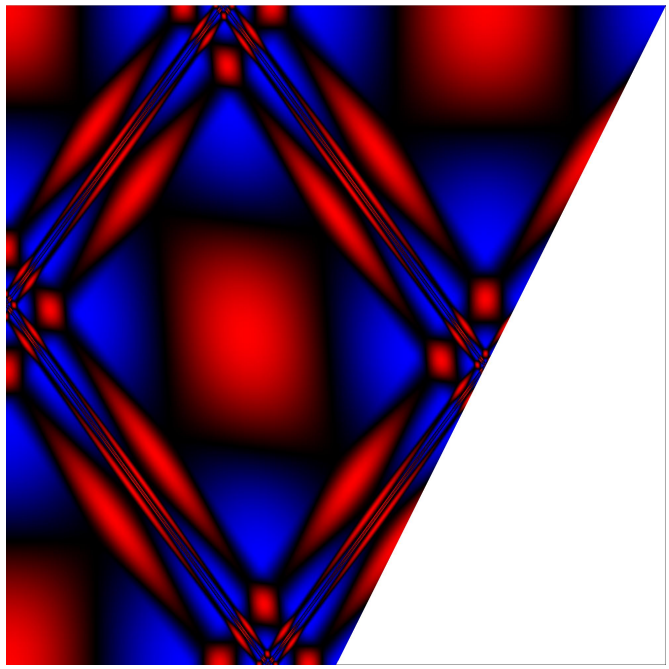


FIGURE 1.1: The stream function in a trapezoidal geometry. Constructed using a ray tracing technique, as described in Maas & Lam (1995).

of geometries in which we can solve the Poincaré equation and examples of geometries for which no closed form solution is known.

Chapter 2 concerns numerical approximation of the Poincaré equation. This chapter is mainly based on Swart *et al.* (2005a). We give an extensive historical overview of numerical methods for internal wave motion. Next we review the mathematical properties of the problem, specifying the necessary conditions for the existence of solutions and pinpointing the origin of the ill-posedness. The proposed numerical method consists of two parts. We present an efficient discretisation, resulting in sparse matrices of dimensions $(n+m) \times (n+m)$, where n and m are the number of grid points along the coordinate axes. Standard finite element or finite difference methods would result in a matrix with dimensions $(nm) \times (nm)$. Secondly, we propose a regularisation scheme that deals with the ill-posedness. This scheme aims at an optimal balance between smoothness and accuracy of the solution. The regularisation is based on a minimisation of a weighted sum of a residual and a measure for the energy. As an added bonus, we do not need to prescribe additional boundary conditions at

the fundamental intervals. The solutions we find are in good agreement with solutions found before (see Maas & Lam, 1995). It is however remarkable that a solution with internal shear layers is found to be the preferred solution, instead of solutions with fractal structure. This might explain why the fractal structure was not observed in laboratory experiments: the smoother solution is energetically preferred.

Also, Chapter 2 contains some additional material not present in the original work. Section 2.4 extends the theory to forced wave motion and Section 2.4.1 describes how to allow for inflow at parts of the boundary. Furthermore, numerical results for the rectangle have been added in Section 2.9.1.

Chapter 3 is based on Swart & Loghin (2005b). Here we examine the Poincaré equation in the presence of a viscous term (highest derivatives pre-multiplied with a small parameter). Physically, this equation is a more acceptable model for internal wave motion, since it incorporates the resistance of the fluid to shearing motion. The viscosity will regularize the discontinuity at the attractor location and transform it to an internal boundary layer. However, numerical problems arise since the resolution of the grid needs to be high in order to properly resolve this boundary layer. Numerically we have the problem, as with the inviscid case, that eigenvalues are close together. Especially for low viscosities where the viscous term is merely a small correction to the Poincaré equation. We are interested in the eigenvalue behaviour, viewing the viscosity as a parameter to be varied. In particular the limit for vanishing viscosity is interesting, in this limit the finite number of eigenvalues of the elliptic viscous problem should degenerate into the dense spectrum of the hyperbolic inviscid problem.

After introduction of the problem in Section 3.2, and a review of the mechanism behind the ill-posedness in Section 3.3, we describe a discretisation using a Finite Element Method in Section 3.4. The Finite Element Method used is fairly standard, the resulting discretisation matrices will, however, be used in subsequent sections in analysing the properties of the problem and error analysis.

In section 3.6 we study the properties of the discretised viscous Poincaré problem. This problem may be posed as an eigenvalue problem. The real and complex parts of the eigenvalue (frequency and damping) together with the viscous parameter were found to trace closed curves in three dimensional space. More specifically, we found ellipses tilted around a common axis (viscosity and damping equal to zero). Using perturbation techniques we show that the angle of the tilt is closely related to the eigenvalue of the inviscid equation, and in general the parameters defining the ellipses may to good approximation be determined from solutions to the inviscid equation.

Next we turn to the stability of eigenvalues and eigenvectors. We are inter-

ested in how the eigensystem responds to small perturbations. Since we know that the problem is ill-posed we expect that small perturbations in the matrices of the discretised problem may yield dramatic changes in the eigenvectors. For both the inviscid and the viscous equation we derive inequalities that supply an upper bound for the distance between solutions of the discretised eigenproblem and solutions of a perturbed eigenproblem. The spectral gap is a key element in these error bounds. As expected, viscosity regularises the solution and error bounds for eigenvectors are tighter. Mainly, this is a consequence of the increased spectral gap, eigenvalues lie further apart since they move away on the aforementioned ellipses. It is then proposed to use the expression for the error bound as an indicator for the reliability of computed solutions.

We conclude with numerical experimentation in Sections 3.8.1 and 3.8.2 where we considered the eigenvalues and the eigenvectors, respectively. The eigenvalues of the discretised problem are calculated as a function of the viscosity using a numerical continuation package (Dhooge *et al.*, 2004). We find the ellipses as predicted before, to good accuracy. Next, we compute the eigenvectors of the viscous Poincaré problem. Selection of reliable vectors is done using the indicator obtained before. For higher values of the viscosity the results are satisfactory. At low viscosity however, most solutions seem to suffer from spurious oscillations and solutions are unstable. At higher grid resolutions the ill-posedness of the problem is even more evident.

Although viscosity has a regularising property, we find that the effect is insufficient at low values of the viscosity. We believe that an additional regularisation procedure will be needed in such circumstances, perhaps an energy-minimising approach as advocated in Chapter 2. Such a method also allows for tuning of a regularisation parameter (e.g. via L-curves as proposed in Section 2.7.1), while the viscosity is in principle a *given* physical quantity.

At this point it is worth comparing the results of the numerical computations from Chapter 2 and Chapter 3. The former approach was aimed at solving the inviscid equation, using a regularisation based on an energy-minimisation. The latter approach relies on the regularising properties of a viscous term. Although the philosophy behind these approaches is quite different, the resulting discretised systems are similar. The results were however quite different. In Chapter 2 we found close agreement to ray-traced solutions of the inviscid equation, or solutions with a step discontinuity. In Chapter 3 we typically find solutions where there is structure around the attractor location, which disappears further away from the attractor. It is likely that difference in character between a minimal energy (containing first derivatives of the streamfunction) and a viscous term (containing fourth order derivatives of the streamfunction) is responsible for the dissimilar results.

The final chapter deals with the description of an experiment carried out

at the Coriolis laboratory in Grenoble. This laboratory contains the world's largest rotating platform, having a diameter 13 meters. The platform is used to study the effect of rotation on fluids, mimicking the rotation of the earth. The aim of our experiment was to test the hypothesis that internal waves, and more specifically internal wave attractors, carry the potential for efficient mixing of the fluid density. This mixing takes place at the boundaries of the fluid domain, where waves are focused. In oceanography, mixing processes are of great importance, they are the means of redistributing hot and cold water, thus driving large scale flows through density differences which they imply. In the deep ocean an amount of mixing takes place which is, as yet, unexplained. Although controversial, some authors argue that internal wave breaking may be the mechanism behind this mysterious mixing. Results of the analysis of the data gathered during the experiment strongly supports this hypothesis.

In our experiment a rotating, annulus-shaped basin, mounted on a turntable, is filled with water. In the fluid a density stratification was created, setting the conditions for gravito-inertial waves. On top of the steady rotation a modulation is applied, at the frequency where we theoretically expect a wave attractor. Measurements were made using two techniques. Firstly, probes were injected into the fluid. These probes record densities, either as time series at a fixed height, or as a function of depth during vertical movement. Secondly, measurements were performed using a PIV technique (Particle Image Velocimetry). This method uses a laser sheet, illuminating a part of the interior of the fluid, in our case a radial cross section. Suspended in the fluid are particles, that are illuminated when crossing the laser sheet. Due to the density stratification, the particles are homogeneously distributed. The illuminated particles are photographed using a high speed camera, and computer software processes the images to create vector field representations of the velocity field.

By inspection of probe data, we find that the density profiles that were established prior to the experiment are completely destroyed after several hours of modulation. This is a surprising result, since usually a stable density stratification lasts on a much larger time scale. We attribute the density modifications to internal wave beam focusing and subsequent mixing, and indeed wave attractors were identified in radial cross-sections of the annulus. Consistent phase behaviour and high amplitudes were found at the attractor location in images obtained by a harmonic analysis of the vorticity field (a measure of the local rotation or shear of fluid particles).

Another interesting point concerns the generation of the waves. Since we do not perturb the fluid in any direct way, nor are we using a topography that breaks the rotational symmetry, where are the waves generated? The answer might well lie in the physics of the boundary layer: a thin layer of fluid where the fluid frictionally adjusts to zero velocity at the boundary. Extending results

of Garrett *et al.* (1993), we propose a model that predicts a boundary layer eruption. At certain critical points at the boundary of the fluid domain the height of the boundary layer can 'explode', thereby periodically pumping fluid into and out of the fluid interior. This is exactly what we observe. We use complex EOF's (Empirical Orthogonal Functions), a decomposition of the signal into variation-minimizing eigenfunctions. The component that contributes most to the signal clearly shows a shear layer being spawned at a critical corner.

1.2 Geophysical Fluid Dynamics

The Poincaré equation has significance in the field of geophysical fluid dynamics, as the equation describing internal waves in a fluid volume. These may be either internal gravity waves or inertial waves, restored by buoyancy and Coriolis forces, respectively. This section will give a brief overview of wave phenomena in fluid dynamics, in order to make clear the position of internal waves in a broader setting. The overview is followed by a derivation of the Poincaré equation from the Navier-Stokes equation in Section 1.2.4.

Furthermore, this section serves as background for Chapter 4 where the importance of internal waves in mixing processes is experimentally examined. The single most important set of equations in fluid dynamics are arguably the *Navier-Stokes equations* for the dynamics of an incompressible fluid,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{\nabla p}{\rho} + \nu \Delta \mathbf{u} + \frac{\mathbf{F}}{\rho}.$$

These equations should be posed on some domain, and supplemented with suitable initial and boundary conditions. In the Navier-Stokes equations $\mathbf{u}$ is the velocity of a fluid particle, ρ is the density and p is the pressure. The function $\mathbf{F}$ is a prescribed body force. Usually, $\mathbf{F}$ contains the gravitational force only, $\mathbf{F} = \mathbf{g}\rho$, with $\mathbf{g} = -g\mathbf{e}_3$ the gravitational acceleration. The small parameter ν is the kinematic viscosity, which is of the order of $10^{-6} m^2/s$ for water. In the case that $\nu = 0$ the resulting inviscid equations are referred to as the *Euler equations*. The Navier-Stokes equations are supplemented by the *continuity equation*,

$$\nabla \cdot \mathbf{u} = 0,$$

following from local mass conservation in a fluid parcel. We will be primarily interested in the case of a rotating system. When we describe the equations from an inertial frame of reference, rotating with angular velocity $\mathbf{\Omega}$ (where we will add a prime to the variables) it is well known that two fictitious forces arise. We see this by examining the convective derivative, which may be writ-

ten in cylindrical coordinates as

$$\frac{D\mathbf{u}}{Dt} \equiv \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = \frac{D\mathbf{u}'}{Dt} + \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}) + 2\mathbf{\Omega} \times \mathbf{u}',$$

where $\mathbf{r}$ is the position vector in cylindrical (r, θ, z)-coordinates. The second term is the *centrifugal force*, the third term is the *Coriolis force*. The centrifugal force can be expressed as a gradient, and may be included in a modified pressure, the *reduced pressure*,

$$P' \equiv p - \frac{1}{2}\Omega^2(x^2 + y^2),$$

where we used the notation $\Omega = \|\mathbf{\Omega}\|_2$. The reduced pressure enables us to write the rotational Navier-Stokes equations in a co-rotating frame as

$$\frac{\partial \mathbf{u}'}{\partial t} + \mathbf{u}' \cdot \nabla \mathbf{u}' + \mathbf{f} \times \mathbf{u}' = -\frac{\nabla P'}{\rho'} + \nu \Delta \mathbf{u}' + \frac{F'}{\rho'}, \tag{1.1}$$

where $\mathbf{f} \equiv 2\mathbf{\Omega}$ is known as the *Coriolis parameter*. We continue our discussion of the Navier-Stokes equation by assuming that we deal with small amplitude motion, and we neglect the non-linear terms in (1.1), giving

$$\frac{\partial \mathbf{u}'}{\partial t} + \mathbf{f} \times \mathbf{u}' = -\frac{\nabla P'}{\rho'} + \nu \Delta \mathbf{u}' + \frac{F'}{\rho'}. \tag{1.2}$$

Note that there are circumstances in which linearisation is not justified. In such a case the *Oseen approximation* (Batchelor, 1967, section 4.10) is often adopted. This approximation consists of replacing the non-linear term $\mathbf{u}' \cdot \nabla \mathbf{u}'$, with the linear term $\mathbf{u}_0' \cdot \nabla \mathbf{u}'$, where $\mathbf{u}_0'$ is some suitably chosen constant vector.

The Navier-Stokes equations are usually studied by assuming some terms to dominate over others, the relative effects of the various terms being measured using several parameters. As yet these parameters are hidden in the physical quantities, they arise when we write the equations in dimensionless form by selecting appropriate scales. For example, if the position variable $\mathbf{x}$ is of relevance on scales of L meters, we write $\tilde{\mathbf{x}} = \mathbf{x}/L$. We have differentiated between dimensionless and dimensional variables by means of a tilde. Let us also set the following scales

$$\begin{aligned} \mathbf{u}' &= U\tilde{\mathbf{u}}', \\ t &= \tilde{t}L/U. \end{aligned}$$

Note that derivatives scale as $\nabla = L^{-1}\tilde{\nabla}$ and $\frac{\partial}{\partial t} = U/L\frac{\partial}{\partial \tilde{t}}$. The dimensionless Navier-Stokes equations may then be written as

$$\frac{\partial \tilde{\mathbf{u}}'}{\partial \tilde{t}} + \frac{Lf}{U}\mathbf{k} \times \tilde{\mathbf{u}}' = -\tilde{\nabla}\tilde{P}' + \frac{\nu}{LU}\tilde{\Delta}\tilde{\mathbf{u}}' + F' \tag{1.3}$$

Parameter	Name	Measures importance of
$\mathrm{Ro} = U/Lf$	Rossby number	Coriolis force over inertial acceleration
$\mathrm{Ek} = \nu/(L^2 f)$	Ekman number	viscous over Coriolis force
$\mathrm{Re} = LU/\nu$	Reynolds number	inertial over viscous force

TABLE 1.1: Some commonly used parameters in fluid dynamics.

Only two dimensional parameters remain: the *Reynolds number* is given by $\mathrm{Re} = \frac{LU}{\nu}$, and we see that it measures the relative importance of the inertial force (the left hand side) over the viscous force. We can also identify the *Rossby number* $\mathrm{Ro} = \frac{U}{Lf}$, measuring the importance of the the inertial acceleration over the Coriolis force. The ratio $\mathrm{Ro}/\mathrm{Re} = \mathrm{Ek}$, known as the *Ekman number*, measures measures the importance of viscous force over the Coriolis force. In table 1.1 we summarize the most commonly used parameters. The relative magnitudes of these parameters determine the type of solution dominant in the system. We will be primarily interested in the case where $Ek \ll 1$ and $Re \gg 1$, i.e. we consider nearly inviscid fluids. Furthermore we restrict ourselves to wave solutions, the nature of which is the topic of the following section.

1.2.1 Waves

A wave may be defined as the mechanism employed by nature to transport energy, without significant transport of mass (Tolstoy, 1973). Waves commonly manifest themselves as sinusoidal oscillations of particles around an equilibrium point. In order for waves to exist, a medium (a fluid) and a restoring force are needed. As pointed out before, there are several restoring forces, giving rise to various waves. Before discussing the various types of waves, we first like to introduce a few key concepts characterizing waves. However, since various types of wave exist, it is not possible to present one concise mathematical prototypical wave. For simplicity, and because this type of wave features prominently throughout this thesis, we will discuss wave properties of a monochromatic (i.e. having one frequency) linear wave:

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{A}\Re(e^{i(\omega t - \sigma + \mathbf{k}\cdot\mathbf{x})}).$$

For this type of wave we may easily characterise several properties:

- **Amplitude** The amplitude $\|\mathbf{A}\|$ measures the half-distance between crest and trough.
- **Frequency** The number of oscillations per second, $\omega \in \mathbb{R}$.

- **Period** The period, $T = 2\pi/\omega$, is the time it takes for the wave to perform an oscillation.
- **Wave number** The wave number measures the number of wave cycles that fit in one meter. In higher dimensions a wave number is assigned to each of the directions, yielding a *wave vector* $\mathbf{k} = (k, l, m)$.
- **Wavelength** As a wave propagates, the wavelength $\lambda = 2\pi/|\mathbf{k}|$ is the distance covered in one period.
- **Phase** The phase, σ measures the position of the wave in the oscillation. Phase is always relative, a gauge must be agreed on.
- **Phase velocity** The phase velocity $\mathbf{c}_p = \frac{\omega}{\kappa^2}\mathbf{k}$, with $\kappa = |\mathbf{k}|$, measures the velocity of a crest or trough.

There is one more concept, the *group velocity*, but before introducing this, it is convenient to define dispersion. A functional relation $\omega = f(\mathbf{k})$ between the frequency and the wave vector is called a *dispersion relation*. Waves are classified as being dispersive or non-dispersive according to $f'(\mathbf{k})$ being non-constant or constant. This implies a non-constant and a constant phase velocity (with respect to the wave number) respectively. If a wave is dispersive, then a combination of a number of waves of differing wave number will yield a complicated ensemble. The velocity of this ensemble is the group velocity (velocity of energy propagation) given by

$$\mathbf{c}_g = (\frac{\partial\omega}{\partial k}, \frac{\partial\omega}{\partial l}, \frac{\partial\omega}{\partial m}).$$

Note that it is assumed here that ω depends continuously on $\mathbf{k}$. Later we will see that in the case of internal waves the concept of wave number loses some of its significance.

1.2.2 Some common waves

Many waves occur as internal waves, i.e. waves that attain their maximum amplitude in the interior of the fluid volume. Surface waves have their maximal displacement at the surface, where we have a density interface. Internal waves have the *buoyancy force*, $F_b = -g\Delta\rho$ as a restoring mechanism. The buoyancy force is a consequence of the Archimedean principle that an object immersed in a fluid experiences an upward force proportional to the volume of the object. The buoyancy force is dependent on the *stratification*, the variation in density $\Delta\rho$ with depth.

In geophysical fluid dynamics one regularly encounters the term barotropic. If $\nabla p \times \nabla \rho = 0$ the fluid is called barotropic. This term is often used in the meaning 'surface' or 'external'. Also, it is frequently used to express lack of structure in the vertical direction. The opposite of barotropic is baroclinic, commonly used in the sense of 'internal', or expressing variations in vertical structure.

In principle waves are classified according to the restoring force that is dominant in the dimensionless Navier-Stokes equations (1.3). The following sections will, very briefly, discuss some of the predominant waves. Special attention is given to the dispersion relation, which allows us to sketch a picture later on in which the position of internal gravity waves is indicated.

Inertial Oscillations

When both $Re \to \infty$ and $Ek \ll 1$ and the pressure gradient force is assumed to be zero, the linearized Navier Stokes equation (1.2) reduces to

$$\mathbf{u}_t + \mathbf{f} \times \mathbf{u} = 0. \tag{1.4}$$

The solutions to this system is easily seen to be clockwise (anticyclonic – against the frame's rotation sense) circular motion in the horizontal plane, with a period of oscillation of $2\pi/f$. This result is obtained by combining the first two equations into $u_{tt} = -f^2 u$, which has the solutions

$$u = A\cos(ft) + B\sin(ft), \tag{1.5}$$

for constants A and B. The second equation of (1.4) then gives us an expression for v. Finally, we conclude that $u^2 + v^2 = A^2 + B^2$, which represents circular motion. The dispersion relation can be induced from (1.5) and reads,

$$\omega = f.$$

This solution is called the *inertial oscillation*. Since it represents a non-propagating solution, not transporting any energy, one could argue that this is not a wave by our definition. Nonetheless such a solution is usually called a 'standing wave'. We mention the inertial oscillation since it is of importance in practice, measurements repeatedly reveal waves at the inertial frequency.

Surface Gravity Waves

Surface gravity waves exist on interfaces between two fluids when there is a balance between acceleration and pressure gradient, plus gravity in the vertical direction. In this case, rotation is not important and $Ro \to \infty$ and $Ek \ll 1$.

Type	Examples	Dispersion relation
Shallow water waves	Tides, tsunamis	$\omega = \sqrt{gH}\kappa$
Deep water waves	Ocean wind waves	$\omega = \sqrt{g\kappa}$

TABLE 1.2: Classification of surface gravity waves.

Since we consider a surface wave, there is no density variation and the density is equal to a constant reference density: $\rho = \rho^*$. This gives the equations

$$\begin{aligned} u_t &= -P_x/\rho^*, \\ v_t &= -P_y/\rho^*, \\ w_t &= -P_z/\rho^* - g. \end{aligned}$$

One may solve these equations, under certain assumptions, distinguishing between deep and shallow water. In both cases, vertical excursions of water parcels decrease exponentially with depth, justifying the term 'surface wave'. In the vertical plane, motions are circular (deep water) or elliptic (shallow water), eventually rectilinear close to the bottom.

In table 1.2 some results are summarized, where the following symbols are used

- H, the water depth,
- k, l, the wave numbers in x and y directions, and $\kappa = \sqrt{k^2 + l^2}$ the length of the wave vector.

In fact, the interface does not need to be of the type water-air. We can also study general two-layer fluids, having a density jump at the interface of $\Delta\rho$. In this case, the results of Table 1.2 remain valid, if we replace the gravity g with the *reduced gravity*

$$g' = g\Delta\rho/\rho^*.$$

The corresponding force $F_b = g\Delta\rho$ is called the *buoyancy force*.

Poincaré and Kelvin Waves

Poincaré and Kelvin waves (or 'gravito-inertial waves') are high frequency, rotationally modified, gravity waves. The Ekman number must also be small, which enables us to neglect the viscous terms. The equations for the velocities take the form

$$\begin{aligned} u_t - fv &= -P_x/\rho^*, \\ v_t + fu &= -P_y/\rho^*, \\ w_t &= -P_z/\rho^* - g. \end{aligned}$$

In view of our investigation of internal waves in bounded geometries, it is interesting to solve the system in a rectangle of dimensions $H \times L$, at the boundary of which we need the velocities to vanish. The component v is eliminated from the system. When wave motion with frequency ω is assumed, we can then solve the system (Pedlosky, 1992) and find the dispersion relation,

$$\omega = \sqrt{f^2 + g'H((n\pi/L)^2 + k^2)}. \tag{1.6}$$

Thus the waves are quantised in the cross-channel direction according to the mode number $n = 1, 2, 3 \ldots$. These waves are called *Poincaré waves*. Poincaré wave exist as both internal and external form. They have a typical asymmetrical sinusoidal-like profile, with highest amplitudes away from the boundaries. There is a second type of solution, where the wave number is not quantised (taking the role of the $n = 0$ solution). These waves are called *Kelvin waves*. These waves are typically attached to the boundary, and decay exponentially into the interior. Kelvin waves have the dispersion relation

$$\omega = \sqrt{g'H\kappa^2}.$$

Note how the dispersion relation is identical to that of shallow water waves. However, it is a quite different type of wave, bound to the rightmost boundary of the channel and exponentially decreasing.

Rossby wave

A Rossby wave is a low frequency wave ($\omega \ll f$), where the same forces are at work as in the previous section, except that an extra mechanism comes into play. Rossby waves are either restored by variations in the local rotational frequency f, or by changes in depth (topographic Rossby waves). In other words, f and H are no longer assumed to be constant. We assume that f and H only vary in the y direction and define the following quantities,

- β is the local average value of $\frac{df}{dy}$,
- α is the local average value of $\frac{f}{H}\frac{dH}{dy}$,

The dispersion relation is then given by

$$\omega = \frac{(\alpha - \beta)k}{k^2 + l^2 + L^{-2}}.$$

See e.g. Batchelor (1967) or Pedlosky (1992) for a derivation. Here the quantity L is the *Rossby radius* given by

$$L = \sqrt{g'H}/f.$$

The gravitational factor can be either the reduced gravity for internal waves, or the usual gravitational constant for surface waves.

Internal waves

The waves that form the main topic of this thesis are internal waves, these are waves having their maximum amplitude in the interior of the fluid. We will discuss these waves in detail in paragraph 1.2.4, but list some main properties here in order to facilitate comparison with other waves.

The main driving forces are the buoyancy force, induced by a continuous stratification, and the Coriolis force . We will show how this leads to a dispersion relation of the form

$$\tan^2\theta = \frac{\omega^2 - f^2}{N^2 - \omega^2},$$

where N^2 is a measure of the strength of the continuous stratification and θ is the angle between the group velocity vector and the horizontal plane. Note that this dispersion relation implies that internal waves exist only in the frequency band where ω^2 is in between f^2 and N^2. In the ocean, N is generally larger than f, however in certain deeper parts of the ocean the stratification is nearly neutral and f may exceed N.

1.2.3 The big picture

An attempt of sketching the dispersion relation of the waves discussed before is given in figure 1.2. We have set $N^2 = 0$ in order make a fair comparison between internal waves and the other wave types possible. In the absence of stratification, the resulting internal waves are called *inertial waves*. One characteristic of inertial waves is that (potentially) they live in a continuous frequency window, in figure 1.2 represented by the shaded area. Later on we will see how the topography can also constrain the allowed frequencies for inertial waves. At the upper limit of allowable frequencies we find the inertial oscillations, while Rossby waves and deep water waves share (part of) the frequency band with inertial waves. At this point it is worthwhile to note that inertial Poincaré waves, Kelvin waves and Rossby waves can very well be considered inertial waves. After all, the same restoring forces are at work, the only difference lies in the typical scales involved and the boundary conditions. Therefore, in principle, the theory of the next section and the resulting Poincaré equation should be able to produce these waves.

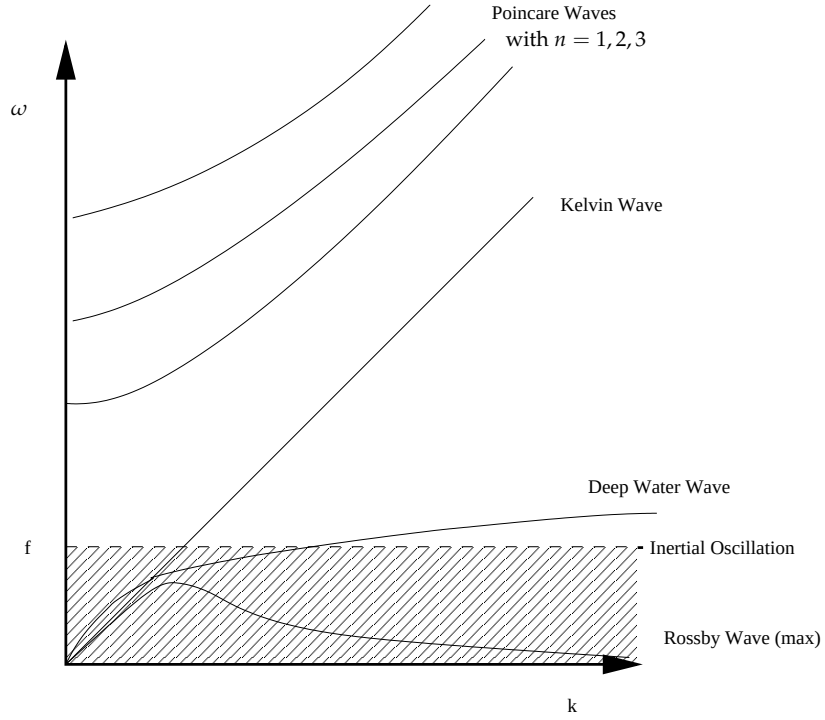


FIGURE 1.2: Classification of waves according to their frequency. The number n for the graphs of Poincaré waves refers to equation (1.2.2)

1.2.4 Internal Waves

We will now continue by studying the Navier-Stokes equation (1.2) in the presence of rotation and stratification. Thus we specialize the force F to be the buoyancy force. A thorough physical justification of the various approximations will be given. We note that, in principle, Rossby waves and Poincaré waves also fit in the following description. After derivation of the governing equations, Section 1.3 will further analyze internal waves in closed domains. We will derive the Poincaré equation, which is the main object of study of this thesis.

Taking into account the pressure gradient, gravitational force, Coriolis force and buoyancy force we can write down the *rotating Euler equations*,

$$\rho(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \mathbf{f} \times \mathbf{u}) = -\nabla p - \rho g \mathbf{e}_3 \tag{1.7}$$

Here ρ stands for density, g is the gravitational acceleration, $\mathbf{u} = (u, v, w)$ is the velocity, p is pressure and $\mathbf{f} = (0, 0, f)$ is a vector containing the Coriolis parameter $f = 2\Omega$. Here $\Omega\mathbf{k}$ stands for the angular velocity of the geometry in which the internal gravity wave problem is considered. The rotation axis is taken to be in the $\mathbf{e}_3$ direction, which is upwards. For a derivation of (1.7) see LeBlond & Mysak (1978). Note that this is an inviscid equation, friction effects have not been incorporated. In Chapter 3 we will consider the viscous equations, we will however not derive the viscous equations from basic principles.

Now, we need a second equation which states that *mass is conserved*,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0. \tag{1.8}$$

Furthermore, it is known from thermodynamics that a relation of the following type (equations of state) exists

$$\rho = \rho(p, S, s) \tag{1.9}$$

Which means that we can express ρ in terms of pressure, entropy (S) and salinity (s). The next equations posed are assumptions and assert absence of sources and sinks,

$$\begin{aligned} \frac{\partial S}{\partial t} + \mathbf{u} \cdot \nabla S &= 0 \\ \frac{\partial s}{\partial t} + \mathbf{u} \cdot \nabla s &= 0 \end{aligned}$$

These equations assert that the entropy and salinity of a small advected fluid element do not change in time. The equations stated so far are in principle closed, yet very impractical. Therefore some more manipulations are in order.

From (1.9) we have that we can write the derivatives

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial p}\frac{\partial p}{\partial t} + \frac{\partial \rho}{\partial s}\frac{\partial s}{\partial t} + \frac{\partial \rho}{\partial s}\frac{\partial s}{\partial t} \tag{1.10}$$

$$\mathbf{u} \cdot \nabla \rho = \frac{\partial \rho}{\partial p}\mathbf{u} \cdot \nabla p + \frac{\partial \rho}{\partial s}\mathbf{u} \cdot \nabla s + \frac{\partial \rho}{\partial S}\mathbf{u} \cdot \nabla S \tag{1.11}$$

Using (1.10)–(1.11) and the relation $\frac{\partial \rho}{\partial p} = c_s^{-2}$ now yields

$$\frac{\partial \rho}{\partial t} + \mathbf{u} \cdot \nabla \rho = \frac{1}{c_s^2}\left(\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p\right) \tag{1.12}$$

The quantity c_s stands for the (non-constant) speed of sound. Furthermore, pressure and density can be split in a static and dynamic part

$$p = p_0(z) + p'(t,x,y,z) \quad (1.13)$$
$$\rho = \rho_0(z) + \rho'(t,x,y,z) \quad (1.14)$$

The dynamic parts can be seen as a perturbation. The static part of the pressure is due to gravity and stratification

$$\frac{dp_0}{dz} = -\rho_0 g \quad (1.15)$$

This relation is, for obvious reasons, called the *hydrostatic balance*. The relations (1.13),(1.14) and (1.15) enable us to write (1.7), (1.8) and (1.12) as:

$$(\rho_0 + \rho')(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u} + \mathbf{f}\times\mathbf{u}) = -\nabla p' - \rho' g \mathbf{e}_3 z \quad (1.16a)$$
$$\frac{\partial \rho'}{\partial t} + \nabla\cdot((\rho_0+\rho')\mathbf{u}) = 0 \quad (1.16b)$$
$$\frac{\partial \rho'}{\partial t} + \mathbf{u}\cdot\nabla(\rho_0+\rho') = \frac{1}{c_s^2}(\frac{\partial p'}{\partial t} + \mathbf{u}\cdot\nabla(p_0+p') \quad (1.16c)$$

Since the speed of sound c_s is not a constant, these equations are not yet a closed system. In the next section, the system will be further simplified, using several approximation techniques.

1.2.5 Approximations

There are several approximations to be made to the system (1.16) to make them suitable for mathematical analysis. The assumptions made and the applied techniques will only be briefly outlined in this section.

- *Linearisation*
  Products of perturbations, or products of perturbations and static variables can be neglected. One can show that a sufficient condition for which this assumption is valid is

  $$|\mathbf{u}| \ll c \ll c_{surf}$$

  In plain English this says that the velocity of the fluid particles must be much smaller than the phase velocity of the internal waves (c), which in turn must be much smaller than the phase velocities of the (relatively long) surface waves (c_{surf}). As a result of the linearisation, the speed of sound c_s has become a constant ($c_s = c_s(p_0, \rho_0, s_0)$). The system (1.16) is now closed.

- *Quasi incompressibility*
  Equations (1.16b and 1.16c) can be combined and then approximated by the *continuity equation* $\nabla \cdot \mathbf{u} = 0$, neglecting the effects of compressibility on the velocity. This needs the assumptions $c \ll c_s$ and $c_{surf} \ll c_s$. In equation (1.16c) the term $-\frac{1}{c_s^2}\frac{\partial p'}{\partial t}$ will be neglected, this is safe if $c_{surf} \ll c_s$. This results in an equation where the compressibility of the fluid *does* have an effect, namely on the density. Taking into account the effects on the density, while neglecting kinematic effects of the compressibility, leads to the term 'quasi-incompressible'.

- *Boussinesq approximation*
  The density in equation (1.16a) plays two roles. In the right-hand side the density perturbation affects the force exerted on a fluid element. In the left-hand side the static density affects the resulting acceleration. The latter effect has a negligible effect on the acceleration of the fluid elements.[3] This motivates the replacement of ρ_0 by a constant, depth-averaged, value ρ^*. In the ocean ρ^* will typically differ from ρ_0 by about 2%. This simplification also introduced the desirable effect that the partial differential equation now has constant coefficients.

The conditions used above can be summarized by

$$|\mathbf{u}| \ll c \ll c_{surf} \ll c_s$$

1.2.6 Approximated equations

The linearisation, quasi incompressibility and Boussinesq approximation lead to

$$\begin{aligned} \rho^*(\mathbf{u}_t + \mathbf{f} \times \mathbf{u}) &= -\nabla p' - \rho' g \mathbf{e}_3, \\ \nabla \cdot \mathbf{u} &= 0, \\ \frac{\partial \rho'}{\partial t} - w\frac{\rho^*}{g}N^2 &= 0 \end{aligned} \tag{1.17}$$

The factor N^2 is given by

$$N^2(z) \equiv -\frac{g}{\rho^*}\{\frac{d\rho_0}{dz} + \frac{\rho_0 g}{c_s^2}\}.$$

Stability of the stratification demands $N^2 > 0$, which says that denser liquid must be below liquid of lesser density, in order for the density distribution to

[3] This is similar to the role of inertial and gravitational mass in classical mechanics. The inertial mass resists acceleration and the gravitational mass determines the gravitational force. The inertial effect determines the system to a much lesser extent than the gravitational effect.

be stable.The factor N is called the Brunt-Väisälä frequency, it acts as an upper bound on internal wave frequencies (Greenspan, 1968).

1.3 The wave equation in two dimensions

We are now in the position to derive the two-dimensional Poincaré equation. Starting from the equations (1.17), we first take the time derivative of the first equation and eliminate the dynamic density using the third equation. For notational convenience we switch to subscript notation for partial derivatives and have

$$\begin{aligned} \mathbf{u}_{tt} + \mathbf{f} \times \mathbf{u}_t &= -\frac{\nabla p_t}{\rho^*} - wN^2\mathbf{e}_3, \\ \nabla \cdot \mathbf{u} &= 0. \end{aligned}$$

The pressure may now be removed by taking the curl of the first equation. After this we have, in component form

$$\begin{aligned} (w_y - v_z)_{tt} + fu_{zt} &= -N^2 w_y, \\ (u_z - w_x)_{tt} + fv_{zt} &= w_x N^2, \\ (v_x - u_y)_{tt} - (fu_x + v_y)_t &= 0, \\ u_x + v_y + w_z &= 0. \end{aligned}$$

Here f is the vertical component of the rotation vector $\mathbf{f} = (0, 0, f)$. We now restrict ourselves to two-dimensional domains. One way of accomplishing this is setting all derivatives with respect to y to zero. Physically, this describes a domain infinite in length with invariant behaviour in the y direction. In the (x, z)-plane we can now introduce the stream function defined by $(\Psi_z, -\Psi_x) = (u, w)$, the continuity equation is then automatically satisfied and we obtain

$$\begin{aligned} -v_{ztt} + f\Psi_{zzt} &= 0, \\ \Delta\Psi_{tt} + fv_{zt} &= -N^2\Psi_{xx}, \\ v_{xtt} - f\Psi_{zxt} &= 0. \end{aligned}$$

Taking the derivative of the second equation with respect to t and inserting the first equation then finally gives us an equation for the stream function

$$\Delta\Psi_{ttt} + f^2\Psi_{zzt} = -N^2\Psi_{xxt}.$$

Since the equation is linear, solutions can be superposed at will. We therefore only consider monochromatic wave behaviour in time, with frequency ω, which is expressed by

$$\Psi(x, z, t) = e^{i\omega t}\Psi(x, z). \tag{1.18}$$

Given solutions for certain values of ω, functions with a more general time dependence can be formed by a Fourier analysis of these monochromatic wave solutions. Using the equation (1.18) gives us the two dimensional Poincaré equation,

$$\Psi_{xx} - \lambda^2 \Psi_{zz} = 0.$$

with

$$\lambda^2 = \frac{f^2 - \omega^2}{\omega^2 - N^2}.$$

We want to find solutions to this partial differential equation on some domain D bounded by ∂D. We take the geometry to be completely bounded. At the boundary it is natural to demand that the flow is parallel to the boundary. This is mathematically expressed as $\mathbf{u} \cdot \mathbf{n} = 0$, where $\mathbf{n}$ is a normal to the boundary ∂D. In terms of the stream function this becomes $\Psi = c$, for some arbitrary constant c. Without loss of generality we choose $c = 0$ and our equation plus boundary condition becomes

$$\Psi_{xx} - \lambda^2 \Psi_{zz} = 0 \qquad \text{in } D, \tag{1.19}$$

$$\Psi = 0 \qquad \text{at } \partial D. \tag{1.20}$$

The above problem is also an eigenvalue problem, only for specific values of λ solutions are possible (depending on the domain D). Chapter 2 gives more details on the solvability of the Poincaré equation. The eigenvalues are divided in two classes

- $\min(f^2, N^2) < \omega^2 < \max(f^2, N^2)$, the hyperbolic case.
- $0 < \omega^2 < \min(f^2, N^2)$, the elliptic case.

The special transition case $N^2 = f^2$ is more complicated and must be addressed separately. This issue is addressed in detail in Friedlander & Siegmann (1982*b*), but we will ussume $N^2 \neq f^2$ for simplicity.

We are interested in the hyperbolic regime, where the internal waves are found. Here we may further analyze the equation by introducing characteristic coordinates. Before we present this analysis we firstly consider the dispersion relation.

1.3.1 The dispersion relation

A dispersion relation gives a relation between the frequency and the wave vector $\mathbf{k} = (k, m)$. First a spatially infinite medium will be considered. In this case the dispersion relation is easy to obtain, because we can assume solutions of the form

$$\Psi = \Psi_0 e^{i(kx+mz)} \tag{1.21}$$

With $k, m \in \mathbb{R}$. This is a plane wave, with amplitude Ψ_0. Inserting this solution in (1.19) gives

$$\omega^2 = \frac{N^2k^2 + f^2m^2}{k^2 + m^2}.$$

More insight is gained by rewriting the wave vector in polar coordinates $\mathbf{k} = \kappa(\cos(\theta), \sin(\theta))$. This represents a vector with an angle θ with the horizontal and a length of κ. The dispersion relation becomes

$$\omega^2 = N^2 \cos^2\theta + f^2 \sin^2\theta, \tag{1.22}$$

or in terms of λ,

$$\lambda = \pm\frac{1}{\tan(\theta)}.$$

This is a fundamental result: the frequency ω determines the angle θ of the wave vector with the horizontal: $\omega = \omega(\theta)$. In contrast to this, elliptic waves typically have $\omega = \omega(|\mathbf{k}|)$, see also Maas (2005). Upon reflection, the incident and reflected waves will be confined to a fixed angle relative to the vertical. In later chapters the profound implications of this property will become clear.

The dispersion relation also implies that the group velocity vector is perpendicular to the wave vector. Energy therefore propagates parallel to the phase lines, under an angle θ with the vertical.

1.4 The method of characteristics

The method of characteristics is a useful method for solving hyperbolic equations. This technique will provide us with a method of predicting the general 'shape' of the solution to (1.19) in any domain by a geometric ray tracing algorithm. This algorithm is useful for testing the validity of the solutions obtained by numerical approximation. We will apply the characteristic method to the two-dimensional Poincaré equation.

The method of characteristics is based on the identification of characteristic curves on which partial differential equations are described by ordinary differential equations. For hyperbolic differential equations of the form

$$a(x,z)u_{xx} + b(x,z)u_{xz} + c(x,z)u_{zz} = 0,$$

there are two families of characteristic curves,

$$\begin{aligned} \xi(x,z) &= c_1, \\ \eta(x,z) &= c_2. \end{aligned}$$

Where $\xi(x,z)$ and $\eta(x,z)$ are obtained from solving the following ordinary differential equation

$$\frac{dz}{dx} = \frac{1}{2a(x,z)}[b(x,z) \pm \sqrt{b^2(x,z) - 4a(x,z)c(x,z)}].$$

In the case studied here we have $a(x,z) = 1$, $b(x,z) = 0$ and $c(x,z) = -\lambda^2$, this gives the following families of characteristic curves (which are now characteristic lines),

$$\xi(x,z) = x - |\lambda|^{-1}z, \tag{1.23}$$
$$\eta(x,z) = x + |\lambda|^{-1}z. \tag{1.24}$$

The smallest angle θ these lines make with the vertical is in fact the same angle found in the dispersion relation (1.22). Thus it can be concluded that the waves propagate in the direction of the characteristics. Two characteristics joining on a point on the boundary can in this light be viewed as a reflection of an internal wave.

In characteristic coordinates, the Poincaré equation reads simply $\Psi_{\xi\eta} = 0$ and the general solution can be obtained by integration,

$$\Psi(x,z) = \mathcal{F}(\xi) + \mathcal{G}(\eta) \tag{1.25}$$

where the functions $\mathcal{F}$ and $\mathcal{G}$ are arbitrary.

Consider a domain D with boundary ∂D. Suppose that every characteristic has at most two intersections with the boundary. This constraint is known as *characteristic convexity* and makes sure that we can carry out integrations inside the domain D. Equation (1.25) is then also valid inside D.

If we consider a point $P_1 = (\xi_1, \eta_1)$ on the boundary then we have by boundary conditions that $\mathcal{F}(\xi_1) = -\mathcal{G}(\eta_1)$. Following a characteristic in the ξ-direction we find a new intersection point $P_2 = (\xi_2, \eta_1)$ where $\mathcal{F}(\xi_2) = -\mathcal{G}(\eta_1)$, thus $\mathcal{F}(\xi_2) = \mathcal{F}(\xi_1)$. This procedure may be carried through to give $\mathcal{F}(\xi_1) = \mathcal{F}(\xi_2) = \mathcal{F}(\xi_3) = \ldots$ and $\mathcal{G}(\eta_1) = \mathcal{G}(\eta_2) = \mathcal{G}(\eta_3) = \ldots$. This is a powerful result, because when setting (prescribing) the pressure at a point P_i this also sets the pressures at all reflection points that can be reached by subsequent reflections. This mechanism also restricts the freedom in prescribing the boundary conditions, since inconsistencies can occur if different values for the pressure are assigned at coupled boundary points. These considerations lead to the concept of a *fundamental interval*. This is defined as a set of closed intervals on the boundary such that prescription of boundary conditions on these intervals determines the pressure (in a consistent way) on the whole of the boundary, as well as in the interior. The largest connected subset of the fundamental interval is called the *primary fundamental interval*. Unicity of the solution

is guaranteed if and only if all the fundamental intervals are prescribed, either by the application of boundary conditions or the influence of another fundamental interval. Unicity follows since all characteristics are in that case set to a specific value, which in turn gives the solutions in the interior by equation (1.25). For certain geometries the fundamental intervals can be expressed in terms of the properties of the geometry and the frequency ω by constructing explicit mappings for successive surface intersections. This was demonstrated before in Maas & Lam (1995).

1.5 The wave attractor

The construction described in the previous section, tracing characteristics from boundary point to boundary point, leads to a dynamical system on the boundary. For the moment we only present a simplified summary here, and leave the details for Chapter 2.

Let us write the transformation acting on a boundary point $s \in \partial D$ as $T : \partial D \rightarrow \partial D$. The map T has (independently of the starting point s) the following properties (John, 1941):

- Every point is a periodic point of order n, thus for all $s \in \partial D$ we have $T^{(n)}(s) = T(s)$.
- Every point tends to an element of some limit set $\omega(T)$ under the map T. Independently of s we then have that there exists a $k \in \omega(T)$ for which $\lim_{n\rightarrow\infty} T^{(n)}(s) = k$.
- For every s and $n > 1$ we have $T^{(n)}(s) \neq T(s)$.

Each of the above cases has its own specific type of solution to the Poincaré equation attached to it. In a sense, the iteration points of the map T describe a 'skeleton' of the solution. The first case describes *modal solutions*, smooth solutions without singularities.

The second case corresponds to solutions containing *wave attractors*, features in the solution then reproduce towards smaller scales. The cluster points of the map T correspond to a singularity in the solution. A prototypical example was shown in figure 1.1, where we have four limit-points. The fractal structure in the solutions presents us with a challenge, numerical approximation of the small scale structure in the solution will require special consideration.

Finally, the third case above is the *ergodic case*, no point is mapped to itself. When continuous solutions are sought, only the trivial solution is possible in this case. This follows since a boundary condition set at one point of the boundary implies that a dense subset of boundary also is assigned this value. From (1.25) we then find the trivial solution in the entire domain.

For the first two cases it is known that fundamental intervals exist. There exists an $M \subset \partial D$ such that

$$\cup_{n=0}^{\infty} T^{(n)}(M) = \partial D.$$

In the case of wave attractors, all iterates are disjoint,

$$\cap_{n=0}^{\infty} T^{(n)}(M) = \emptyset,$$

while for modal solutions the fundamental interval is for some $m > 0$ mapped to itself,

$$T^{(m)}(M) = T(M).$$

Finally, in the ergodic case the fundamental interval consists of a set of measure zero on the boundary.

1.6 Ill-posedness

It turns out that the mapping T described above is highly sensitive to perturbations in the boundary of the domain. An arbitrarily small perturbation can, for example, turn a wave attractor into a modal solution. We will describe in Chapter 2 how we can attach a value, the *winding number* or *rotation number*, to the boundary of the domain. This winding number classifies in which of the cases described above we are. The crucial observation is now that *the winding number does not depend continuously on the geometry of the boundary*. This is the source of the *ill-posedness* of the solution, which we take to mean that solutions do not vary continuously with the model parameters. In Section 1.7.4 we present limit point diagrams of a few selected geometries which clearly illustrate the ill-posedness of the problem.

Note that often parameter regimes exist where the solution is robust under perturbation, this robustness underlies the physical realisability of the wave attractor (Maas *et al.*, 1997).

The ill-posedness has consequences for the solvability of the Poincaré equation in a numerical sense. Solutions are unstable in the sense that a slight modification of the computational grid may destroy a solution. In essence solutions are so close together (in some sense) that numerically solutions are indistinguishable. In Chapter 2 we will give examples of numerically obtained solutions, that are unstable under even a small modification of the grid. At this

point it is worthwhile to mention that our definition of ill-posed is slightly different from the classical definition due to Hadamard (Hadamard, 1923). Mathematical problems are *well-posed in the sense of Hadamard* if

1. the solution exists and is non-trivial
2. the solution is unique
3. the solution depends continuously on the data.

A problem is ill-posed if it is not well-posed. The Poincaré problem would then be called ill-posed since boundary data at the fundamental intervals must be specified in order to obtain a unique solution *and* it the solution is sensitive a perturbation in the parameters. We will however use the term ill-posed in the 'discontinuous response to continuous variation' meaning. Whenever the solution is not uniquely determined for given boundary data we will use the term 'under-determined'. Thus, the Poincaré problem is both ill-posed and under-determined. The justification for this non-standard nomenclature lies in the techniques used to deal with the ill-posedness and under-determinedness of the problem. The sensitive nature with respect to variations in the parameters is effectively dealt with using a form of regularisation. In this thesis we use a variation on Tikhonov regularisation (Tikhonov & Arsenin, 1977) in Chapter 2 and a viscous regularisation in Chapter 3. On the other hand, we obtain unique solutions simply by supplying boundary conditions where those are needed, on the fundamental intervals that were introduced in Section 1.4

1.7 Examples

The Poincaré equation is notoriously difficult to solve analytically. Among the few geometries for which solutions were obtained (in two dimensions) are: the circle (Barcilon, 1968), the triangle (Franklin, 1972), and the square (Bourgin & Duffin, 1939). For the case of three-dimensional rotating flows, analytical solutions were obtained for the axial cylinder (Kelvin, 1880) and the oblate spheroid (Bryan, 1889). The following examples will show the possible solutions for some selected geometries. This will give some insight in the type of solution we can expect. Where no solutions in closed form are known, we revert to a ray tracing technique in order to identify wave attractors in parameter space.

1.7.1 The Strip

Suppose that ∂D, in characteristic coordinates, is bounded by the graphs of the two functions $h^t(\xi)$ (top) and $h^b(\xi)$ (bottom). In this case the boundary

condition becomes

$$\mathcal{F}(\xi) = -\mathcal{G}(h^t(\xi)) = -\mathcal{G}(h^b(\xi)). \tag{1.26}$$

The Poincaré problem on the 2D domain has now become a 1D problem, the determination of a function $\mathcal{G}$ for which it holds that $\mathcal{G}(h^t(\xi)) = \mathcal{G}(h^b(\xi))$. It should be noted that once a function $\mathcal{G}$ has been found, an infinity of solutions is available by constructing $f \circ \mathcal{G}$ for some function f.

Next, we will solve the Poincaré problem on the infinite strip of height b tilted by an angle α, which is described by

$$\begin{aligned} \tilde{h}^t(x) &= sx + b, \\ \tilde{h}^b(x) &= sx, \end{aligned}$$

with $s = \tan(\alpha)$. Using the inverse of the transformation (1.23) these equations can be translated to the (ξ, η) coordinate system,

$$\begin{aligned} h^t(\xi) &= \frac{1+\lambda s}{1-\lambda s}\xi + \frac{2\lambda b}{1-\lambda s}, \\ h^b(\xi) &= \frac{1+\lambda s}{1-\lambda s}\xi. \end{aligned}$$

The condition (1.7.1) now reads

$$\mathcal{G}(\frac{1+\lambda s}{1-\lambda s}\xi + \frac{2\lambda b}{1-\lambda s}) = \mathcal{G}(\frac{1+\lambda s}{1-\lambda s}\xi), \forall \xi \in \mathbb{R}.$$

Making the substitution $\xi' = \frac{1+\lambda s}{1-\lambda s}\xi$ one obtains

$$\mathcal{G}(\xi') = \mathcal{G}(\xi' + \frac{2\lambda b}{1-\lambda s}).$$

In other words: all functions $\mathcal{G}$ that have $T = \frac{2\lambda b}{1-\lambda \tan(\alpha)}$-periodicity generate solutions.

1.7.2 The wedge

In this section we will examine the wedge defined by

$$D = \{(x, z) | \, 0 < \alpha_2 x \leq z \leq \alpha_1 x\}.$$

This example is important since it models a corner in a domain. In characteristic coordinates we describe the boundary of the domain by

$$\begin{aligned} h^t(\xi) &= \frac{1+\alpha_1\lambda}{1-\alpha_1\lambda}\xi, \\ h^b(\xi) &= \frac{1+\alpha_2\lambda}{1-\alpha_2\lambda}\eta. \end{aligned}$$

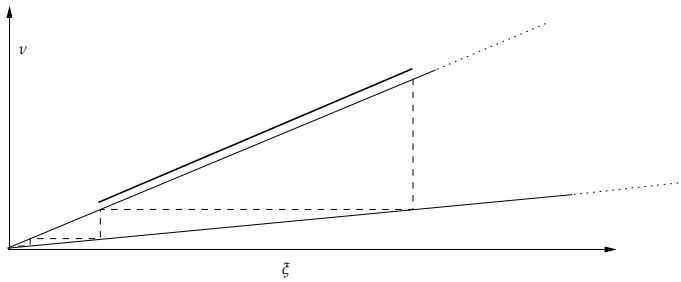


FIGURE 1.3: Schematic picture of characteristics in a wedge. Note that the characteristics tend to the apex of the wedge, while the fundamental intervals shrink.

Next we examine the consequences of $\mathcal{G}(h^t(\xi)) = \mathcal{G}(h^b(\xi))$. Substituting the expressions for the coordinates we obtain

$$\mathcal{G}(\frac{1+\alpha_1\lambda}{1-\alpha_1\lambda}\xi) = \mathcal{G}(\frac{1+\alpha_2\lambda}{1-\alpha_2\lambda}\xi).$$

After the coordinate transform $\xi' = \frac{1+\alpha_1\lambda}{1-\alpha_1\lambda}\xi$ we derive the relation

$$\mathcal{G}(\xi) = \mathcal{G}(\frac{1+\alpha_2\lambda}{1-\alpha_2\lambda}\frac{1-\alpha_1\lambda}{1+\alpha_1\lambda}\xi) \equiv \mathcal{G}(\alpha\xi),$$

with

$$\alpha = \frac{1+\alpha_2\lambda}{1-\alpha_2\lambda}\frac{1-\alpha_1\lambda}{1+\alpha_1\lambda} = \frac{1-\alpha_1\alpha_2\lambda^2+(\alpha_2-\alpha_1)\lambda}{1-\alpha_1\alpha_2\lambda^2+(\alpha_1-\alpha_2)\lambda}.$$

For any $k \in \mathbb{Z}$, the functions must satisfy

$$\mathcal{G}(x) = \mathcal{G}(\alpha^k x).$$

Suppose that $\alpha > 0$. The solution consists of parts that are repeated while shrinking in scale towards the origin. This is easy to see by evaluating $\mathcal{G}$ in the logarithm of its argument, yielding

$$\mathcal{G}(\log(x)) = \mathcal{G}(k\log(\alpha) + \log(x)),$$

thus $\mathcal{G}$ is $\log(\alpha)$-periodic when using a logarithmic x-axis, see Figure (1.3) for a clarification of this process. It is easy to identify the fundamental interval for this geometry by noting that $(\alpha^i, \alpha^{i+1}]$ is exactly the interval that is repeated for $\mathcal{G}$. The fundamental intervals for $\mathcal{G}$ are given by

$$M_i = (\alpha^i, \alpha^{i+1}],$$

for any i. We can now uniquely solve the Poincaré equation by choosing a value of j and applying the Dirichlet boundary condition

$$\mathcal{G}(\eta) = h_j(\eta), \text{ on } M_j.$$

Note that the function $\mathcal{F}$ is found by $\mathcal{F}(\xi) = -\mathcal{G}(h^t(\xi))$ and can always be constructed after $\mathcal{G}$ has been found.

If we want a solution that is of class $C^n(\mathbb{R}^+ \times \mathbb{R}^+)$ we need

$$\lim_{\xi \leftarrow \alpha^j} h_j^{(n)}(\xi) = h_{j-1}(\alpha_j) = h_j^{(n)}(\alpha^{i+1}).$$

Note however that $\mathcal{G}(0)$ is still undetermined. For continuity at the origin we need in addition

$$\mathcal{G}(0) = \lim_{\eta \downarrow 0} \mathcal{G}(\eta),$$

or in other words, for all $\epsilon > 0$ there is a $\delta > 0$ such that

$$x < \delta \rightarrow |\mathcal{G}(x) - \mathcal{G}(0)| < \epsilon.$$

Suppose that we have set boundary conditions $h_0(x)$ on M_0, then $\mathcal{G}(x) = h_k(x) = h_0(\alpha^{-k}x)$ for $x \in M_k$ and let us choose $\epsilon = \alpha^{-k}$. In that case

$$x < \delta,\ x \in M_k \rightarrow |h_0(\alpha^{-k}x) - \mathcal{G}(0)| < \alpha^{-k}.$$

Now, the values that $h_0(\alpha^{-k}x)$ attains on M_k are independent of k, while α^{-k} can become arbitrarily small. Suppose that $\min_{x \in M_k} |h_0(\alpha^{-k}x) - \mathcal{G}(0)| = c$, then we can always choose k such that $\alpha^k < c$, except when $c = 0$. This implies that only the constant solution is possible if the solution is to be continuous at zero.

1.7.3 The Rectangle

Consider the $a \times b$ rectangle in the xz-plane. We want to solve the Poincaré equation on this domain. Solutions on this domain have been classified before in Bourgin & Duffin (1939) and Maas & Lam (1995). They show that separation of variables easily yields the general solution

$$\Psi(x,z) = \sum_{j=0}^{\infty} a_j \sin(\alpha_j x) \sin(\beta_j z), \tag{1.27}$$

with $\alpha_j = \frac{\pi n j}{a}$, $\beta_j = \frac{\pi m j}{b}$, where $n, m \in \mathbb{N}$ and the additional constraint

$$\lambda^2 = (\frac{mb}{na})^2. \tag{1.28}$$

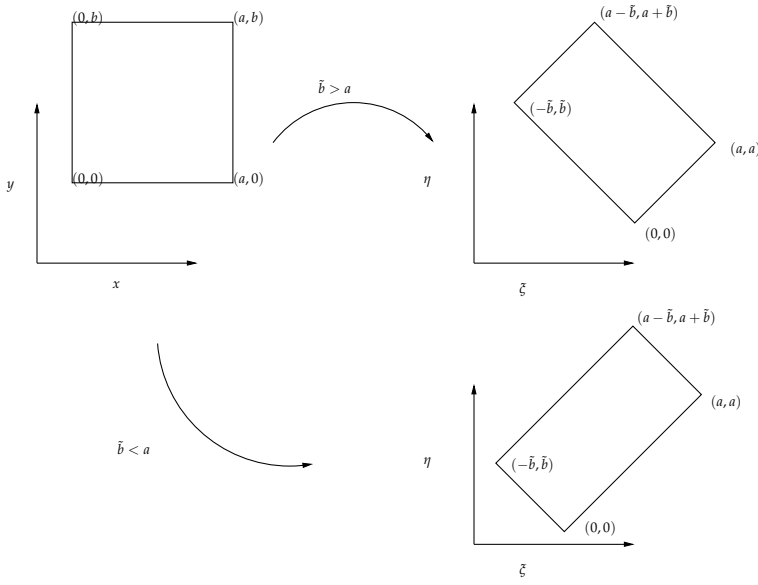


FIGURE 1.4: The $a \times b$ rectangle and its image in the (ξ, η)-plane.

Given a value of λ we will take m and n as the smallest possible wave numbers that satisfy (1.28). Solutions exist only if $\lambda a/b$ is a rational number. This is related to the rotation number being rational, as will become clear later on. Another feature which will be a recurring theme is the repetitive nature of the solution. For example, in the x-direction the solution repeats its structure in parts of maximum size $2a/n$. One could say that the coefficients a_j allow one to freely specify an interval which is then repeated. This, of course, is a manifestation of the fundamental interval.

The result above may also be obtained by considering the mappings of the characteristics. For convenience, we make the substitution $\tilde{b} = \lambda^{-1} b$. As can be seen in Figure (1.4) the geometry is bounded by four straight lines and we can define the transformed rectangle by four functions,

$$\begin{aligned} h^{t1}(\xi) &= \xi + 2\tilde{b} && ,\xi \in [-\tilde{b}, a - \tilde{b}], \\ h^{t2}(\xi) &= -\xi + 2a && ,\xi \in [a - \tilde{b}, a], \\ h^{b1}(\xi) &= -\xi && ,\xi \in [-\tilde{b}, 0], \\ h^{b2}(\xi) &= \xi && ,\xi \in [0, a]. \end{aligned}$$

In view of Figure (1.4) two cases must be distinguished. First look at $\tilde{b} < a$

where we have

$$\begin{aligned}
\mathcal{G}(h^{t1}(\xi)) &= \mathcal{G}(h^{b1}(\xi)) \quad \xi \in [-\tilde{b}, 0], \\
\mathcal{G}(h^{t1}(\xi)) &= \mathcal{G}(h^{b2}(\xi)) \quad \xi \in [0, a-\tilde{b}], \\
\mathcal{G}(h^{t2}(\xi)) &= \mathcal{G}(h^{b2}(\xi)) \quad \xi \in [a-\tilde{b}, a],
\end{aligned}$$

which gives,

$$\begin{aligned}
\mathcal{G}(\xi + 2\tilde{b}) &= \mathcal{G}(-\xi) \quad , \xi \in [-\tilde{b}, 0], \\
\mathcal{G}(\xi + 2\tilde{b}) &= \mathcal{G}(\xi) \quad , \xi \in [0, a-\tilde{b}], \\
\mathcal{G}(-\xi + 2a) &= \mathcal{G}(\xi) \quad , \xi \in [a-\tilde{b}, a].
\end{aligned}$$

This may be rewritten into

$$\mathcal{G}(\xi) = \begin{cases} \mathcal{G}(2\tilde{b} - \xi) & \xi \in [0, \tilde{b}], \\ \mathcal{G}(\xi + 2\tilde{b}) & \xi \in [0, a-\tilde{b}], \\ \mathcal{G}(2a - \xi) & \xi \in [a-\tilde{b}, a]. \end{cases}$$

In order to unclutter notation we will define the linear transformation

$$T^{c,d}_{a,b}(\xi) : [a, b] \rightarrow [c, d]$$

that linearly maps a to c and b to d for $\xi \in [a, b]$. It is the identity map for $\xi \notin [a, b]$. Using this notation we have

$$\mathcal{G}(\xi) = \mathcal{G}(T^{2\tilde{b},\tilde{b}}_{0,\tilde{b}}\xi) = \mathcal{G}(T^{2\tilde{b},a+\tilde{b}}_{0,a-\tilde{b}}\xi) = \mathcal{G}(T^{a+\tilde{b},a}_{a-\tilde{b},a}\xi).$$

These transformations can be viewed as several translation and reflection symmetries between subdomains that must be simultaneously satisfied. These domains can now be conveniently read from the sub and superscripts of the transformation T. A symbolic representation of these transformations can be seen in Figure (1.5). A question that now comes to mind is to determine the maximum interval on which the function $\mathcal{G}(\xi)$ can be prescribed. The rest of the interval being determined by the transformations. This interval is evidently not unique, there will be a set of intervals to choose from. Fortunately, the situation is not too complicated for the rectangle. Suppose we fill the interval $[0, \tilde{b}]$, the small side of the rectangle. Following the transformations we find that firstly the interval $[\tilde{b}, 2\tilde{b}]$ is filled with a mirrored copy. Secondly, the transformation $T^{2\tilde{b},a+\tilde{b}}_{0,a-\tilde{b}}$ shifts a copy of $[0, 2\tilde{b}]$ to $[2\tilde{b}, 4\tilde{b}]$. Next, the same transformation yields copies $[2k\tilde{b}, 4k\tilde{b}]$, for $k = 2, 3, 4, \ldots$. For a certain value of k the intervals will reach the point a. Now, we see that the interval $[0, a]$ must be

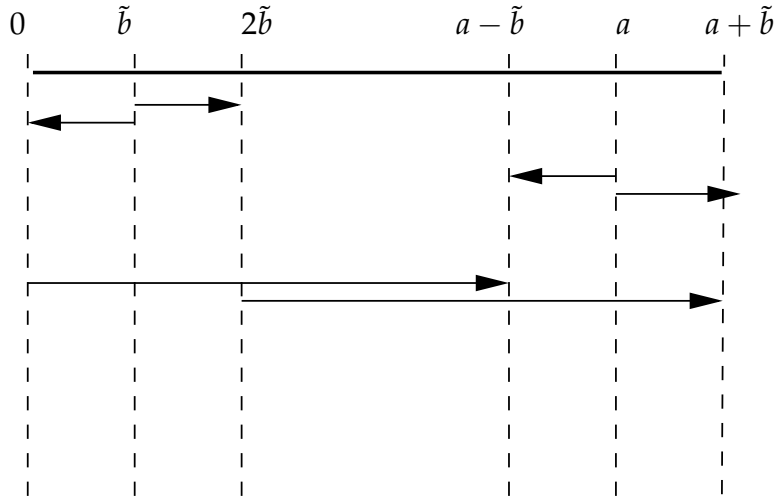


FIGURE 1.5: The transformations on the range of $\mathcal{G}(\xi)$.

filled exactly with k copies of $[0, 2\tilde{b}]$, otherwise the transformation $T^{a+\tilde{b},a}_{a-\tilde{b},a}$ will introduce an inconsistency. Thus

$$a = 2k\tilde{b}.$$

Another possibility consistent with the transformations is filling the entire interval $[0, a-\tilde{b}]$ with copies of $[0, 2\tilde{b}]$. This gives

$$(a+\tilde{b}) = 2k\tilde{b}.$$

Together these equations imply

$$a = k\tilde{b}.$$

Additionally, inside the interval $[0, \tilde{b}]$ we can consistently prescribe n copies of intervals with length $\tilde{b}/n$,

$$\mathcal{G}(\xi) = \mathcal{G}(\xi + 1/n), \text{ for } \xi \in [lb/n, (l+1)b/n) \text{ with } l = 0, \ldots, n-2.$$

So, we see that instead of k copies of $[0, \tilde{b}]$, we may also use m copies of $\tilde{b}/n$ to fill the interval $[0, a]$. Combining this result with $\tilde{b} = \lambda^{-1}b$ gives us the expression

$$\lambda = \frac{mb}{na},$$

which corresponds with the known result. The case where $\tilde{b} > a$ is completely analogous to the given analysis.

1.7.4 Some ray-tracing results

We have seen several geometries on which the Poincaré equation is analytically solvable. This is however exceptional, due to the simplicity or symmetry of the previous examples. In this section we give three examples of geometries that are not accessible to an analytical solution, in the sense that closed form solutions are not available. The cause of this difficulty is ultimately the complicated mapping induced by the characteristics.

Figure 1.6 presents the *tilted square* (see also Kopecz, 2006), the *trapezoid* (Maas *et al.*, 1997) and *the parallelogram*. The figures were obtained by means of the algorithm presented in Appendix 1.A. We vary two parameters: the angle of the characteristics and some parameter that deforms the geometry. The resulting length of the attractor is assigned a color and plotted in a so-called limit point diagram. The angle of characteristics with the vertical is given by $\lambda = \tan\theta$, we choose $f = 0$ and $N = 1$ so that $\lambda^2 = \omega^2/(1-\omega^2)$. The parameter ω is varied between zero and one.

For the geometry parameter s we varied the following:

- **the tilted square** $s \in [0, \pi/2]$ is the angle with the horizontal.
- **the trapezoid** $s \in [0, 1]$ is the x-coordinate of the intersection of the sloping wall with the horizontal.
- **the parallelogram** $s \in [0, \pi/2]$ is the angle with the horizontal.

What is directly clear from these limit point diagrams is the sensitive nature of the solution with respect to parameters, a manifestation of the ill-posedness of the problem. In between areas of low-length attractors are areas of complicated attractors. In between those are again more complicated attractors to be found. The 'tongue' shaped features are called *Arnol'd tongues*. In fact, the tongues are dense in parameter space, yet the ray-tracing algorithm can only calculate a finite number of reflections. For an infinite number of reflections the limit point diagram would appear to be completely filled.

1.7.5 Construction of domains for selected solutions

It is remarkable that the number of domains on which (1.3) is analytically solvable is very limited. The circle, triangle and square have been tackled, yet more complicated geometries need disproportional effort to solve. However, in this section we describe a method that makes an infinity of computational domains available.

First, we parametrize the boundary using

$$\gamma : [0,1] \to \partial D, \gamma(0) = \gamma(1), \gamma(t) = (x(t), y(t)).$$

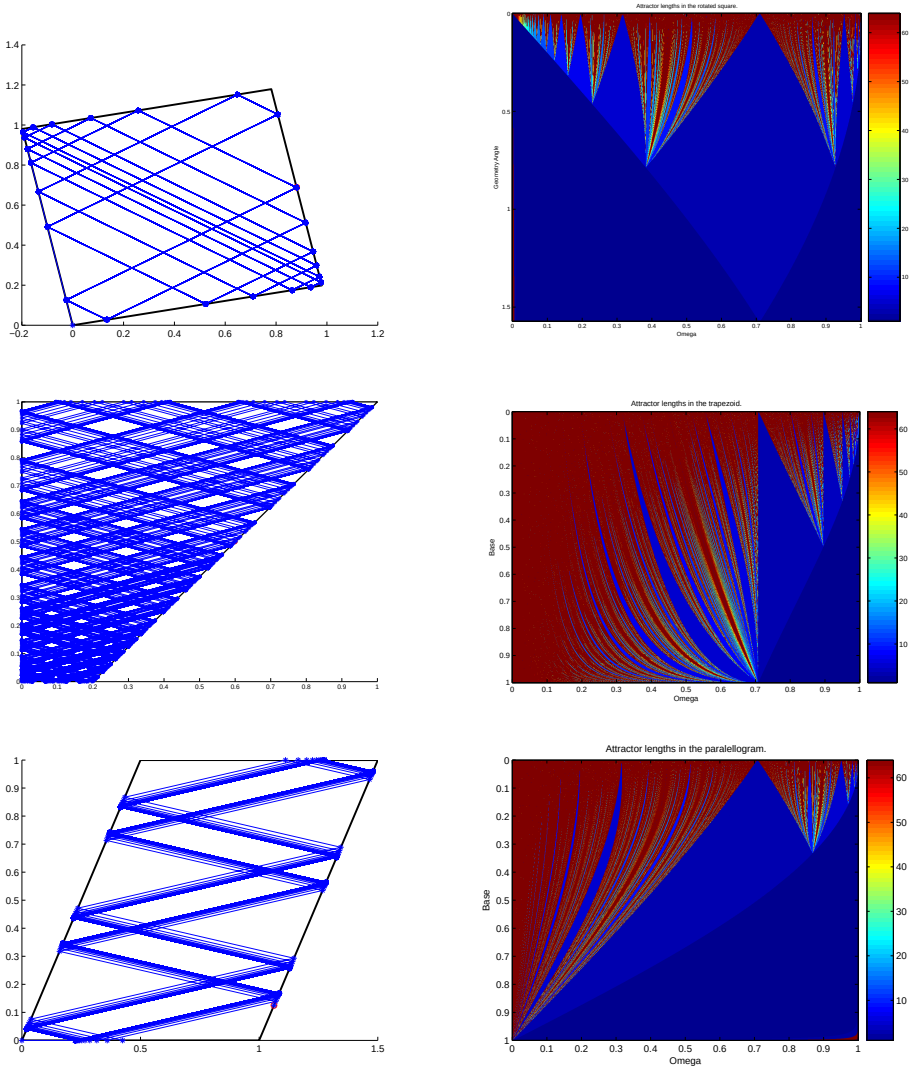


FIGURE 1.6: Some geometries with ray-traced characteristics (leftmost column) and corresponding limit-point diagrams (rightmost column). On the horizontal axis ω varies from zero to one, and thereby the angle of characteristics from zero to $\pi/2$. The vertical coordinate varies some aspect of the geometry, as explained in the main text.

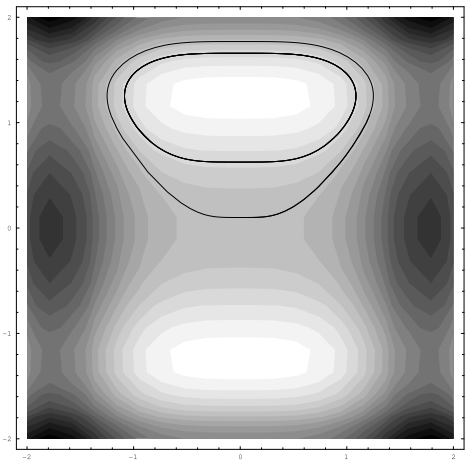


FIGURE 1.7: We have taken $F(\xi) = \cos(\xi^2)$ and $G(\eta) = \sin(\eta^2)$. Shown is a contour plot of the stream function with a few selected boundaries overlayed.

In the transformed domain $\mathbf{l} = (\dot{\xi}(t), \dot{\eta}(t))$ is a vector tangent to the boundary at time t. The boundary condition $\Psi = 0$ may also be written $\nabla\Psi \cdot \mathbf{l} = 0$ and thus

$$(\mathcal{F}'(\xi), \mathcal{G}'(\eta)) \cdot (\dot{\xi}, \dot{\eta}) = 0 \quad \text{on} \quad \partial D,$$

or $\dot{\xi}\mathcal{F}'(\xi) = -\dot{\eta}\mathcal{G}'(\eta)$. We can write down a system which has solutions that conform to the boundary conditions,

$$\begin{aligned} \dot{\eta} &= F'(\xi), \\ \dot{\xi} &= -G'(\eta). \end{aligned}$$

Now note that this is a Hamiltonian system

$$\begin{aligned} \dot{\eta} &= \frac{\partial H(\xi,\eta)}{\partial \xi}, \\ \dot{\xi} &= -\frac{\partial H(\xi,\eta)}{\partial \eta}. \end{aligned}$$

with Hamiltonian $H(\xi, \eta) = F(\xi) + G(\eta)$ equal to the stream function. Given the functions $\mathcal{G}$ and $\mathcal{F}$ we have a Hamiltonian and given the Hamiltonian we can construct boundaries for domains using the above system. An example is shown in figure 1.7. It is a well known fact that two dimensional Hamiltonian systems with one degree of freedom admit only saddles or center points (the flow is volume preserving). The saddle points give rise to possible closed domains with corners.

Appendix to Chapter 1

1.A Ray tracing

In this section we will describe an algorithm for calculating the paths of the characteristics in a two dimensional domain with piecewise linear boundary. Suppose that the vertices $(x_0, z_0), \ldots, (x_n, z_n)$ are given. For convenience we transform each coordinate to characteristic coordinates, since in this coordinate system the characteristics are simply horizontal and vertical lines.

$$\begin{aligned} \xi_i &= x_i + \lambda^{-1} z_i, \\ \eta_i &= x_i - \lambda^{-1} z_i. \end{aligned}$$

Define

$$\begin{aligned} \Delta\xi_i &= \begin{cases} \xi_{i+1} - \xi_i & \text{for } 0 \le i < n, \\ (\xi_0 - \xi_i) & \text{for } i = n. \end{cases} \\ \Delta\eta_i &= \begin{cases} \eta_{i+1} - \eta_i & \text{for } 0 \le i < n, \\ (\eta_0 - \eta_i) & \text{for } i = n. \end{cases} \end{aligned}$$

The line pieces that make up the domain may then be given in parametrized form by

$$(\xi_i(s), \eta_i(s)) = (\xi_i + \Delta\xi_i s, \eta_i + \Delta\eta_i s).$$

The internal wave ray is mapped from boundary point to boundary point. Denote the position at the boundary in step j by (R_j^ξ, R_j^η). After refection the orientation of the ray switches from horizontal to vertical. Set the initial orientation at $j = 0$ to be horizontal, then the line through (Rx_j, Ry_j) is parametrized by

$$(R_j^\xi(t), R_j^\eta(t)) = \begin{cases} (R_j^\xi, R_j^\eta) + (t, 0) & \text{for } j \text{ even}, \\ (R_j^\xi, R_j^\eta) + (0, t) & \text{for } j \text{ odd}. \end{cases}$$

An intersection between a ray at step j and a boundary line piece i is characterized by the solution (s, t) to the simultaneous equations $R_j^\xi(t) = \xi_i(s)$ and $R_j^\eta(t) = \eta_i(s)$. We have an intersection if $s \in [0, 1]$ and $t \neq 0$. Positive (negative) t indicates a ray traveling in the upward (downward) or rightward (leftward) direction. If the domain is characteristically convex then there is only one such solution. When the domain is not characteristically convex, then a horizontal or vertical line possibly has multiple intersections with the boundary of the domain. Also, a ray should not leave the domain. If the vertices of the domain are specified in counter clockwise order, then the outward normal

to line piece i may be described by $\mathbf{n}_i = (\Delta\eta_i, -\Delta\xi_i)$ and for each line piece we need to check the following

$$\begin{aligned} \operatorname{sign}(t) &= -\operatorname{sign}(\Delta\eta_c) = \operatorname{sign}(\Delta\eta_j) \quad \text{for } j \text{ even}, \\ \operatorname{sign}(t) &= \operatorname{sign}(\Delta\xi_c) = -\operatorname{sign}(\Delta\xi_j) \quad \text{for } j \text{ odd}, \end{aligned}$$

where c is the index into the current line piece and j is the index into the line piece which we want to check for an intersection.

The equations for s and t are easily seen to be

$$\begin{aligned} t &= \begin{cases} (\xi_i - R_j^\xi) + \frac{\Delta\xi_i}{\Delta\eta_i}(\eta_i - R_j^\eta) & \text{for } j \text{ even and } \Delta\eta \neq 0, \\ \frac{1}{\Delta\xi_i}(\xi_i - R_j^\xi) & \text{for } j \text{ odd and } \Delta\xi \neq 0. \end{cases} \\ s &= \begin{cases} \frac{1}{\Delta\eta_i}(\eta_i - R_j^\eta) & \text{for } j \text{ even and } \Delta\eta \neq 0, \\ (\eta_i - R_j^\eta) + \frac{\Delta\eta_i}{\Delta\xi_i}(\xi_i - R_j^\xi) & \text{for } j \text{ odd and } \Delta\xi \neq 0. \end{cases} \end{aligned}$$

Typically we need several reflections while not varying the geometry and it makes sense to precalculate some invariant quantities:

$$\begin{aligned} A_i &\equiv 1/\Delta\xi, & B_i &\equiv 1/\Delta\eta, \\ C_i &\equiv \Delta\xi/\Delta\eta, & D_i &\equiv \Delta\eta/\Delta\xi, \\ E_i &\equiv \xi_i - C_i\eta_i & F_i &\equiv \eta_i - D_i\xi_i. \end{aligned}$$

This yields

$$\begin{aligned} t &= \begin{cases} E_i - R_j^\xi + C_i R_j^\eta & \text{for } j \text{ even and } \Delta\eta_i \neq 0, \\ A_i(R_j^\xi - \xi_i) & \text{for } j \text{ odd and } \Delta\xi_i \neq 0. \end{cases} \\ s &= \begin{cases} B_i(R_j^\eta - \eta_i) & \text{for } j \text{ even and } \Delta\eta \neq 0, \\ F_i - R_j^\eta + D_i R_j^\xi & \text{for } j \text{ odd and } \Delta\xi \neq 0. \end{cases} \end{aligned}$$

Note that this only needs cheap multiplications, no divisions. Remembering that there can be only one edge for which $t > 0$ we arrive at the efficient algorithm 1. The efficiency, in terms of execution time of the algorithm, lies in the fact that many quantities are pre-calculated and only a few additions and multiplications are needed to find a new intersection of a characteristic with the domain. Also, working in characteristic coordinates, where characteristics are aligned with the coordinate axes, has simplified the calculations considerably.

Algorithm 1 Ray tracing in a characteristically convex domain.

Set tolerance ϵ
for $i = 1$ to n **do**
  $\xi_i \leftarrow x_i + \lambda^{-1} z_i$ and $\eta_i \leftarrow x_i - \lambda^{-1} z_i$
end for
for $i = 1$ to $n - 1$ **do**
  $\Delta\xi_i \leftarrow \xi_{i+1} - \xi_i$ and $\Delta\eta_i \leftarrow \eta_{i+1} - \eta_i$
end for
$\Delta\xi_n = \xi_n - \xi_0$ and $\Delta\eta_n = \eta_n - \eta_0$
for $i = 1$ to n **do**
  Precalculate $A_i, \ldots, F_i$
end for
Set start position (R_0^ξ, R_0^η)
for $j \leftarrow 0$ to nrTraces-1 **do**
  $i \leftarrow 0$
  found $\leftarrow$ false
  while not found **do**
    if $\Delta\eta \neq 0$ **then**
      $t \leftarrow E_i - R_j^\xi + C_i R_j^\eta$
      if $t > \epsilon$ **then**
        $s \leftarrow B_i(R_i^\eta - \eta_i)$
        if $0 \leq s \leq 1$ **then**
          $R_{j+1}^\xi \leftarrow R_j^\xi + t$
          found $\leftarrow$ true
        end if
      end if
    end if
    $i \leftarrow i + 1$
  end while
  $i \leftarrow 0$
  found $\leftarrow$ false
  while not found **do**
    if $\Delta\xi \neq 0$ **then**
      $t \leftarrow A_i(R_j^\xi - \xi_i)$
      if $t > \epsilon$ **then**
        $s \leftarrow F_i - R_j^\eta + D_i R_j^\xi)$
        if $0 \leq s \leq 1$ **then**
          $R_{j+1}^\eta \leftarrow R_j^\eta + t$
          found $\leftarrow$ true
        end if
      end if
    end if
    $i \leftarrow i + 1$
  end while
end for

Chapter 2

Numerical Solution of the Poincaré Equation

The research in this chapter is published as part of

Numerical Solution of the two Dimensional Poincare Equation
A.N. Swart and G.L.G. Sleijpen and L.R.M. Maas and J. Brandts,
Journal of Computational and Applied Mathematics,
Volume 200, Issue 1, March 2007, Pages 317–341

2.1 Introduction

This paper deals with efficient numerical computation of internal wave phenomena in two dimensional enclosed domains. In contrast to the usual surface waves, internal waves are a type of wave that propagate through a fluid volume. The mechanisms enabling these waves are either rotation or stratification. Efficient numerical models would be of use in various areas of (geophysical) fluid mechanics. One may think of the rotating core of the earth, or the oceans with their density stratification.

The governing equation can be conveniently expressed in terms of the stream function $\tilde{\Psi}$ (see Maas & Lam, 1995),

$$\begin{array}{lcll} \frac{\partial^2\tilde{\Psi}(x,z)}{\partial x^2} - \lambda^2 \frac{\partial^2\tilde{\Psi}(x,z)}{\partial z^2} & = & 0 & \text{in } \tilde{\Omega}, \\ \tilde{\Psi}(x,z) & = & 0 & \text{at } \partial\tilde{\Omega}. \end{array} \tag{2.1}$$

In this equation $\lambda \in \mathbb{R}$ is a parameter corresponding to the frequency at which the waves are forced. The stream function is defined by

$$(u(x,z), v(x,z)) \equiv (-\frac{\partial\tilde{\Psi}(x,z)}{\partial z}, \frac{\partial\tilde{\Psi}(x,z)}{\partial x}),$$

where $(u(x,z), v(x,z))$ is the velocity vector at a point (x,z). The fluid flows along streamlines, curves of constant magnitude of the stream function. The boundary $\partial\tilde{\Omega}$ is a simple closed curve in the (x,z) plane. As customary in fluid dynamics, the z-axis points upwards.

In a three dimensional setting one may derive the equation $\Psi_{xx} + \Psi_{yy} - \lambda^2\Psi_{zz} = 0$, this equation is commonly called the Poincaré equation, after Cartan (1922) who recognised Poincaré as the author who firstly described the equation in Poincaré (1885). Although (2.1) is a reduction of the original three dimensional problem, we will still refer to it as the Poincaré equation. Some authors prefer the term 'wave equation', but we like to stress the absence of a time-like coordinate. This dramatically changes the nature of the problem, instead of an initial value problem we have a boundary value problem.

In this paper we analyse the solvability of (2.1) in sections 2.3 and 2.5. Section 2.4.2 extends the problem to non-zero values on the boundary. In Section 2.6 we develop an efficient discretisation scheme. We continue in Section 2.7 where we will show that the differential equation plus boundary conditions constitute an *ill-posed* problem. We propose a regularisation procedure for the discretised problem, based on minimisation of the energy (Section 2.8), to obtain meaningful solutions. We conclude by presenting some results of the numerical approximation and regularisation, applied to a representative model geometry.

2.2 Previous numerical work

Most work on internal waves has been in the context of internal tides. One is usually interested in behaviour in the vicinity of a continental shelf. At the lateral boundaries a radiation condition is mostly used. One of the first analytical models is that of Rattray Jr. (1960) who considers a step topography in conjunction with a two layer system. In a later paper (Rattray Jr. *et al.*, 1969) the model was extended to include continuous stratification. The generation of internal wave beams traveling seaward is mentioned. Viscosity is added to model in the work of Prinsenberg *et al.* (1974). A special form of the solution is sought, in terms of standing and traveling waves. The dissipation damps the modes, higher modes are more suppressed. It is found that a large number of modes is needed to describe the beams. Nowadays there is strong evidence that a modal framework might not be appropriate to describe the peculiar features of internal waves.

The work of Baines leads to one of the first numerical models. In Baines (1971) reflection of internal waves a from supercritical boundary is studied. This leads to a Fredholm integral equation to be solved. This work is continued in Baines (1973) where more realistic domains are considered. The Fredholm integral equation is solved using an iteration technique. The author needs to smoothen some boundary points in order to obtain a smooth solution. The 'Fredholm integral' approach is dropped in Baines (1974), where functional relations are formed by identifying characteristics at reflection points on the boundary. Closed loops or limit-cycles are not allowed which places restrictions on the geometry. The author simultaneously solves a set of equations and the functional relations, using an complex iteration procedure. The stratification is allowed to be non-linear with depth, giving a non-constant Brunt-Väisälä frequency $N^2(z)$. Unfortunately the author assumes a unique solution, probably the solution is dependent on the initial guess to starting the iteration.

The work by Sandstrom (1976) is along similar lines. The author constructs transformations T from the boundary (given by $z = 0$ and $z = h(x)$) to itself and postulates the existence of a 'phase function' S such that $TS(x) = S(x) + 1$. The construction of such a transformation is nowadays the method of choice for mathematicians, see for example John (1941) or Lyashenko & Smiley (1995) and references therein. The function $f(S(x))$ that constitutes the solution is then periodic and can be solved using Fourier techniques. The author uses an iteration technique. The phase function S can be freely chosen, this is clearly a manifestation of the fundamental interval. Sandstrom notes that further *physical constraints* are needed to specify the solution, which is also the viewpoint that we adopt. In the obtained solution discontinuities are found which are removed by suppressing high wavenumber modes. We now know that rapidly

varying features often occur in solutions and suppressing them does not seem justified.

In Chuang & Wang (1981) a finite difference method is used to find solutions to the internal wave problem. The depth h is assumed to be a function of the x-coordinate. In order to be able to do finite differences the domain is transformed to a rectangle by the transformation $s = z/h(x)$. The drawback is that in physical space the grid is non-uniform, regions with steep slopes have less grid points. Also the lower boundary must have an analytical expression, or be sufficiently smooth, to ensure existence of the mapping. The main concern is the governing equation, it is cast into an awkward form by the coordinate transform.

In the method of Craig (1987*a*) and Craig (1987*b*) an algebraic system is formed by analytical integration over characteristics. The method allows for a depth-dependent frequency $N^2(z)$, but has the drawback that the only domains allowed are those described by a monotone height function:

$$\Omega = \{(x,z)|z \in [h(x), H], x \in [0,1], h'(x) > 0\}.$$

The integration is only possible because of open boundaries at $x = 0, 1$. The greatest concern lies in the grid generation, it is time consuming and only applicable for a small range of geometries.

Of particular interest is Henderson & Aldridge (1992) where a finite-element procedure is used for solving the inertial (i.e. three dimensional, rotational) wave problem in a closed geometry. Rotational symmetry is assumed which renders the problem two dimensional. The results are compared with experimental data by Beardsley (1970). The finite-element method is especially suited for efficient handling of boundary conditions, without resort to a stepped approximation. The authors view only closing characteristics (resonance) as acceptable, at some points wave attractors are mistaken for resonance. Furthermore the authors report 'spurious eigenvalues' whose number increases when the grid size increases. We now know that we can attribute this behaviour to the non-uniqueness of the problem.

The method of Cushman-Roisin *et al.* (1989) is also of interest. Here the hyperbolic equation is solved in a completely enclosed geometry. The application in this case is the modeling of internal waves in fjords. The method is based on recasting the equations in the form of a dissipative artificial-time dependent system and integrating to the point of equilibrium. The grid is aligned with the characteristics and a finite difference approach is taken. For this reason domains are only allowed to have horizontal and vertical segments. The steps introduced in this manner have a distorting effect on the solution as was first noticed by Wotherspoon (1996). It is interesting that the authors note that the system is ill-posed, the solution depends on the initial guess at time zero. They

relate this ill-posedness to the closing of ray paths. Unfortunately almost all ray paths close artificially in this method while non-closing orbits are in fact more rule than exception. In Tverberg *et al.* (1991) the method is extended for non-uniform stratification by a local stretching of the depth. A radiation condition at the boundaries is introduced and the model is applied to a realistic situation. The authors need to introduce a small viscosity to obtain useful solutions. Recent work on coastal type settings is the thesis Wotherspoon (1995). The author introduces the 'c-conformal mapping' (closely related to classical conformal mapping) which leaves the governing equation intact. At the same time the domain is transformed to a stepped geometry where finite differences can be easily applied. Apart from the coastal setting, there has also been internal wave research in the field of geophysical fluid dynamics. We would like to mention the papers Rieutord (1991) and Rieutord & Valdettaro (1997) where the authors consider inertial waves in the liquid core of the earth. Spherical harmonics are used, as they are well suited to the domain, a spherical shell. Numerical calculations show attractors as the general type of solution. Viscosity is thought of as regularising the singularities, attractors turn into shear layers. In later works, Dintrans *et al.* (1999) and Rieutord *et al.* (2001), there is more emphasis on the wave rays as a dynamical system. Poincaré plots and Lyapunov exponents are used to gain insight in the behaviour. The evolution of the eigenvalues and eigenvectors of the system with varying viscosity is discussed. The evolution of eigenvalues in the complex plane was found to exhibit complex behaviour. For zero viscosity an infinite number of eigenvalues, each corresponding to an infinite eigenspace exist. For nonzero viscosity a finite number of eigenvalues, each having a finite eigenspace, survive. With increasing viscosity eigenvalues will shift and get close together, eigenvectors corresponding to such eigenvectors will influence and 'contaminate' each other. We will return to this type of behaviour in Chapter 3.

Other recent works are that of Rieutord *et al.* (2001) or Rieutord *et al.* (2002), who examine the special case of the spherical shell and Maas & Lam (1995) where a ray-tracing approach is taken. The existence of *wave attractors* and ill-posedness of the problem is recognized. Another recent paper, describing a method that makes use of the characteristics of the equation, is that of Harlander & Maas (2006). Their method is also applicable to several variations of the Poincaré equation.

Our approach is different from previous numerical work, we allow a large class of polynomial domains. Furthermore, we minimise an energy, which acts as a regularisation technique that deals with the ill-posedness of the problem. Previously, authors relied on the inclusion of viscosity as a regularisation technique. The algorithmical approach of Maas & Lam (1995) does not use regularisation to obtain smooth functions, but instead it gives a recipe for obtaining

function values at selected points. Also, this method requires the determination of fundamental intervals (see Section 2.3), for which there is no general method available. We avoid the use of fundamental intervals. Finally, we believe that in a three dimensional setting the use of a regularisation method is of even greater importance. We expect the problem to be ill-posed to a higher degree, i.e. more sensitive to disturbances in the relevant parameters, since the behaviour of characteristics is more complex. Also, in principle our fundamental interval-free discretisation method can be generalised to three dimensions. We no longer have a convenient separation of variables, but the principle of discretisation on the boundary only, by considering the mapping of characteristics, is still viable.

Another regularisation technique is inclusion of viscosity in the governing equations. In a physical context this viscosity is a given parameter which is usually very small, which calls for a fine mesh. When the discretisation is refined, there will be more solutions that can be represented on the grid. The existence of many nearby solutions (nearby in parameter space) will still render the problem ill posed, perhaps not formally, but certainly from a numerical point of view. These issues will be addressed in the forthcoming paper Swart & Loghin (2005b). Also the delicate dependence of the solution on the shape of the boundary, even in idealised domains, calls for a regularisation technique that is more sophisticated than a fixed viscosity. An energy minimising regularisation procedure like we suggest may perform that function.

This chapter is an extended version of Swart *et al.* (2007a), including appendices that were not in the original paper.

2.3 The Poincaré equation

In this section we will analyse the Poincaré problem and summarise some previous work. Most importantly, the nature of the solution is seen to be highly dependent on the shape of boundary. There may be no solution at all, or an infinite number of solutions. Furthermore, almost every solution exhibits a fractal structure.

Let us first simplify the equations (2.1) by recognising the hyperbolic nature and introducing characteristic coordinates (ξ, η) defined by

$$(\xi(x,z), \eta(x,z)) = \Xi(x,z) \equiv (x - \lambda^{-1}z, x + \lambda^{-1}z).$$

This transforms the domain into $\Omega = \Xi\tilde{\Omega}$ and it transforms (2.1) to

$$\begin{array}{rcll} \frac{\partial^2 \Psi(\xi,\eta)}{\partial\xi\partial\eta} & = & 0 & \text{on } \Omega, \\ \Psi(\xi,\eta) & = & 0 & \text{at } \partial\Omega. \end{array} \tag{2.2}$$

The solvability of (2.1) is different from (2.2), there are differing smoothness requirements on solutions. We will however be concerned with the latter equation and will not address the issue whether solutions of (2.2) are also solutions to (2.1). Let us now define a class of domains.

Definition 1 (Characteristically convex domains)**.** *We call a domain* Ω characteristically convex*, or convex with respect to the characteristics, if every line in the* ξ *or in the* η *direction intersects the boundary* $\partial\Omega$ *in at most two points.*

Note that convexity in the usual sense can be expressed by asking that every line intersects the boundary in at most two points. In this light characteristic convexity is a weaker constraint, every convex domain is characteristically convex. We can now state the following theorem.

Theorem 1. *Every solution of the Poincaré equation (2.2) on a characteristically convex domain* Ω *is of the form* $\Psi(\xi,\eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta)$.

Proof. Choose the origin somewhere in Ω and integrate

$$\int_0^\xi \int_0^\eta \frac{\partial^2 \Psi(r,s)}{\partial r \partial s}\, dr\, ds = 0.$$

This yields

$$\Psi(\xi,\eta) - \Psi(\xi,0) - \Psi(0,\eta) + \Psi(0,0) = \tilde{\mathcal{F}}(\xi) + \tilde{\mathcal{G}}(\eta),$$

for arbitrary $\tilde{\mathcal{G}}$ and $\tilde{\mathcal{F}}$. The integration can be carried out, since by characteristic convexity the rectangle $(0,\xi)\times(0,\eta)$ is in Ω if the origin and (ξ,η) are in Ω. If $\tilde{\mathcal{F}}(\xi)$ and $\tilde{\mathcal{G}}(\eta)$ are arbitrary then also $\mathcal{F}(\xi) = \tilde{\mathcal{F}}(\xi) + \Psi(\xi,0) - \Psi(0,0)$ and $\mathcal{G}(\eta) = \tilde{\mathcal{G}}(\eta) + \Psi(0,\eta)$ are arbitrary and we have

$$\Psi(\xi,\eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta).$$

□

Let $c(t) : [0, L] \to \partial\Omega$ be a piecewise C^k, globally continuous, parametrisation of the boundary,

$$\partial\Omega = \{(\xi(t),\eta(t)) | (\xi(t),\eta(t)) = c(t), 0 \le t < L, c(0) = c(L)\},$$

where L is the total Euclidean length of $\partial\Omega$. We consider domains with the following properties

- We only allow simply connected, domains Ω, with closed boundary $\partial\Omega$ as described above, that are convex with respect to the characteristics.

- Segments of the boundary are not allowed to be parallel to the coordinate axes, i.e. the sets $\{t|c'(t)^T e_1 = 0\}$ and $\{t|c'(t)^T e_2 = 0\}$ have measure zero within $\partial\Omega$.

Next we define the vertices of the domain,

Definition 2 (Corners)**.** *If* $\lim_{t\uparrow a} c'(t) \neq \lim_{t\downarrow a} c'(t)$ *we call the point* $c(a)$ *a* corner point *or* vertex*. Define the* characteristic rectangle $R = (\xi^-, \xi^+) \times (\eta^-, \eta^+)$ *as the smallest rectangle containing* Ω*. Denote the closure of a domain by a bar, i.e* $\bar{\Omega} = \Omega \cup \partial\Omega$ *and* $\bar{R} = [\xi^-, \xi^+] \times [\eta^-, \eta^+]$*. We call the four points* $\bar{\Omega} \cap \bar{R}$ *the* extreme vertices*.*

The solution to the Poincaré equation can be extended to the characteristic rectangle.

Theorem 2 (Extension of the solution)**.** *Every solution to the Poincaré equation on* $\bar{\Omega}$ *is extensible to* $\bar{R}$*.*

Proof. For every point $(\xi, \eta) \in \bar{R} \setminus \bar{\Omega}$ it is possible to find ξ' and η' such that $(\xi', \eta) \in \bar{\Omega}$ and $(\xi, \eta') \in \bar{\Omega}$. These points are inside the domain and we have $\Psi(\xi', \eta) = \mathcal{F}(\xi') + \mathcal{G}(\eta)$ and $\Psi(\xi, \eta') = \mathcal{F}(\xi) + \mathcal{G}(\eta')$. Thus $\mathcal{F}(\xi)$ and $\mathcal{G}(\eta)$ exist and can be used to define $\Psi(\xi, \eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta)$. □

We have one extra constraint from the physics of the internal wave problem. The boundary condition $\Psi = 0$ comes from the fact that there should be no flow through the boundary, expressed by $(u, v) \cdot n = 0$ where n is the outward unit normal to the boundary. At a corner point the normal is not defined, but physically the no-flow condition dictates that $(u, v) = 0$. We will add this as an extra constraint. In terms of the stream function this becomes

$$\nabla\Psi = 0 \quad \text{at a corner point.} \tag{2.3}$$

Since we extended the domain to $\bar{R}$ the gradient operator is well defined except for the at most four extreme vertices where we may use one sided derivatives for $\mathcal{F}$ or $\mathcal{G}$. We know that the stream function Ψ may be scaled, multiplicatively, by any constant, thereby scaling the velocities. We do not care about the difference between Ψ and $c\Psi$, and we normalise the stream function as

$$\|\Psi\|^2_{L_2(\Omega)} = \int_\Omega \Psi^2(\xi, \eta)\, d\xi\, d\eta = < \mathcal{F} + \mathcal{G}, \mathcal{F} + \mathcal{G} >_{L_2(\Omega)} = 1. \tag{2.4}$$

This constraint also rules out the trivial solution $\Psi = 0$, and by boundary conditions the constant solution $\Psi = c$. Furthermore, we want a unique representation of Ψ in terms of $\mathcal{F}$ and $\mathcal{G}$. Suppose there is another pair $\tilde{\mathcal{F}}, \tilde{\mathcal{G}}$, for which

$$\mathcal{F} + \mathcal{G} = \tilde{\mathcal{F}} + \tilde{\mathcal{G}} \implies \mathcal{F} - \tilde{\mathcal{F}} = \tilde{\mathcal{G}} - \mathcal{G}.$$

Since ξ and η may be independently varied, we have that $\mathcal{F} = \tilde{\mathcal{F}} + c$ and $\mathcal{G} = \tilde{\mathcal{G}} - c$ for an arbitrary constant c. In order to have a unique constant we normalise

$$< \mathcal{F}, 1 >_{L_2([\xi^-,\xi^+])} = 0. \tag{2.5}$$

The Poincaré problem is now reduced to finding $\mathcal{F}$ and $\mathcal{G}$ for which

$$\begin{array}{lcll} \mathcal{F}(\xi) + \mathcal{G}(\eta) & = & 0 & \text{at } \partial\Omega, \\ \mathcal{F}'(\xi) = \mathcal{G}'(\eta) & = & 0 & \text{at a corner point,} \\ < \mathcal{F}, 1 >_{L_2(\Omega)} & = & 0, & \\ < \mathcal{F} + \mathcal{G}, \mathcal{F} + \mathcal{G} >_{L_2(\Omega)} & = & 1, & \end{array} \tag{2.6}$$

with $\mathcal{F} \in C^1 : [\xi^-, \xi^+] \to \mathbb{R}$ and $\mathcal{G} \in C^1 : [\eta^-, \eta^+] \to \mathbb{R}$.

We will show that the first requirement in (2.6) induces a map from the boundary to itself. This was firstly noted by John (1941) and later constructively applied in Maas & Lam (1995). This can be observed by considering function values $\mathcal{G}(\eta_1) = -\mathcal{F}(\xi_1)$ at a boundary point (ξ_1, η_1). When tracing a line in the ξ-direction a new boundary intersection (ξ_2, η_2) is found. Of course $\eta_2 = \eta_1$ and we find $\mathcal{F}(\xi_2) = -\mathcal{G}(\eta_2) = -\mathcal{G}(\eta_1) = \mathcal{F}(\xi_1)$. Carrying through with this procedure we find that one function value on the boundary determines function values at boundary points $(\xi_1, \eta_1), \ldots, (\xi_n, \eta_n)$. Much work has been done in studying the Poincaré problem by viewing the above construct as a dynamical system. We will review the main results on existence of solutions in the case that the boundary satisfies the assumption of characteristic convexity. In this case we can follow (John (1941)) and introduce the following homeomorphisms:

- T^+ assigns to a boundary point the unique boundary point with the same η-coordinate, with the exception that the top and bottom extremal vertices are mapped onto themselves.
- T^- assigns to a boundary point the unique boundary point with the same ξ-coordinate, with the exception that the left and right extremal vertices are mapped onto themselves.
- $F = T^- \circ T^+ : \partial\Omega \to \partial\Omega$.

Note that F is an orientation preserving map. We now define an orbit as

$$O(P) \equiv \{P, T^+P, FP, (T^+ \circ F)P, F^2P, \ldots\}, \text{ where } P \in \partial\Omega.$$

Let $S(P) \equiv c^{-1}(P)|_{[0,L)}$ be the coordinate (by arc length distance) of a point P on $\partial\Omega$. By characteristic convexity of the boundary, we can lift the homeomorphism F to a continuous increasing function $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(s + L) = f(s) + L, \quad s \in \mathbb{R},$$

and

$$S(FP) = f(S(P)) \pmod{L}, \quad P \in \partial\Omega.$$

If we now set $f_k(s) = f(f_{k-1}(s))$ where $f_1(s) = f(s)$ we can define the *rotation number* $\alpha(F)$ of F by

$$\alpha(F) \equiv \lim_{n\to\infty} \frac{f_n(s)}{n} \in [0, L). \tag{2.7}$$

The rotation number can be shown to exist and to be independent of s (de Melo & van Strien, 1994). A classical theorem due to John (1941) is

Theorem 3. *The homeomorphism F has three separate types of behaviour in a characteristically convex domain*

A. *We have $\alpha(F) = m/n$ for some $m, n \in \mathbb{N}$ and $F^n = I$. This case is often called a* resonance, *or* modal solution.

B. *We have $\alpha(F) = m/n$ for some $m, n \in \mathbb{N}$, but $F^n \neq I$. The dynamical system defined by F has a finite number of attracting fixed points. We call solutions of the Poincaré equation in this case* wave attractors.

C. *$\alpha(F) \in \mathbb{R} \setminus \mathbb{Q}$, then F^k has no fixed point for any k. There is only one orbit that fills the entire boundary. This is the* ergodic case.

Note that for a given $\alpha(F)$ the values of n and m are not unique in the first two cases. Multiples km and kn of m and n will yield the same rotation number and will not change the solution type. The solvability of the problem (2.2) depends on the case in which the rotation number falls. This delicate dependence on the rotation number, and thus on the geometry of the domain, makes the problem ill-posed. This will be of concern when considering a numerical approximation technique. To proceed, we need the concept of *fundamental interval* (Maas & Lam, 1995; Lyashenko, 1993). The fundamental interval can be defined as the set $M \subset \partial\Omega$ for which every orbit has at most one point in M and all orbits that have a point in M constitute the entire boundary. The following definition by Lyashenko & Smiley (1995) gives a more constructive approach:

Definition 3 (Fundamental interval)**.** *Denote by $(P, Q)_{\partial\Omega}$ the open set of points in between P and Q (using anti clockwise orientation) on the boundary. Set $(P, Q]_{\partial\Omega} \equiv (P, Q)_{\partial\Omega} \cup Q$, $[P, Q)_{\partial\Omega} \equiv (P, Q)_{\partial\Omega} \cup P$ and $[P, Q]_{\partial\Omega} \equiv (P, Q)_{\partial\Omega} \cup Q \cup P$. Denote the vertices by $P_0, \ldots, P_3$, with P_0 the uppermost extreme vertex and $P_1, \ldots, P_3$ the rest of the vertices in anti clockwise order. Now define, for the same n as used in Theorem (3), in the case A, the following*

- *for n even: P^* as the point from the finite set $O(P_1) \cap (P_0, P_1]_{\partial\Omega}$ such that $O(P_1) \cap (P_0, P^*)_{\partial\Omega} = \emptyset$,*

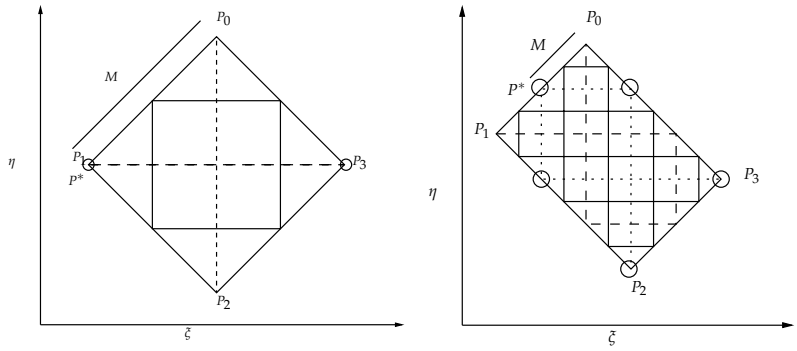


FIGURE 2.1: This figure clarifies the fundamental interval as introduced in Definition 3. At the right is shown a geometry with $n = 5$, which is easiest to see by following the closed orbit (solid line). Of course one may also consider the dashed and dotted lines that connect $O(P_1) = O(P_0)$, respectively $O(P_2) = O(P_3)$. Keep in mind that $T^+P_0 = P_0$, $T^-P_1 = P_1$, etc. The orbit $O(P_2)$ is indicated by circles. The left picture is a case where $n = 2$. In both pictures the set $M = [P_0, P^*)$ is the fundamental interval.

- *for n odd: P^* as the point from the finite set $O(P_2) \cap (P_0, P_1]_{\partial\Omega}$ such that $O(P_2) \cap (P_0, P^*)_{\partial\Omega} = \emptyset$.*

The fundamental interval [1] *is then given by $M = [P_0, P^*)$. Define M_ξ and M_η by orthogonal projection of M on the ξ and η axis.*

In order to clarify the above definition an example has been constructed in Figure 2.1. It was proved for the resonant case that the fundamental interval must be supplied with boundary conditions (for either $\mathcal{F}$ or $\mathcal{G}$) in order to obtain an unique solution. It is likely that the concept of the fundamental interval is also useful in the case (B), the attractor regime. In Maas & Lam (1995) a procedure is suggested for several geometries. In the following we will work with the assumption that the fundamental intervals also exist in the wave attractor case.

When in the following discussion we speak of Dirichlet boundary conditions, we mean choosing a function $h(\xi)$ and setting $\mathcal{F}(\xi) = h(\xi)$ on M_ξ. This implies the value for $\mathcal{G}$ on M_η by the boundary condition $\mathcal{F}(\xi) + \mathcal{G}(\eta) = 0$. The solvability of (2.2), together with this additional boundary condition, is different for the three cases in Theorem 3. Some results are summarised in the following theorem.

Theorem 4 (Solvability in L_2)**.** *The solvability of the Poincaré problem (2.2) in $L_2(\Omega)$ is dependent on the cases, and the value of n, from Theorem 3 as follows*

[1] In Lyashenko & Smiley (1995) the fundamental interval is called *generating set.*

A. *If Dirichlet data $h(\xi)$ for $\mathcal{F}(\xi)$ are supplied on M_ξ, then we have an unique solution (Lyashenko, 1993, Theorem 6). In this case there exist piecewise C^k coordinate transforms $\xi \to p(\xi)$ and $\eta \to q(\eta)$ that transform the domain to a rotated rectangle (Lyashenko, 1994; Lyashenko & Smiley, 1995). In (x,z) coordinates this corresponds to a non-rotated rectangle for which the solutions are known. The solutions on the original domain are then found using inverse coordinate transforms.*

B. *According to Fokin (1983, Theorem 2, Remark 2), the problem is solvable.*

C. *From Fokin (1983, Theorem 5, Remark 5), there exists only the trivial solution in this case.*

The conclusions in this theorem can be extended to the situation where an inhomogeneous term $f(\xi,\eta,\Psi)$ is added to (2.2). The conclusions do have to be modified, for example, solutions are possible in the case (C) if there exists a $C(\alpha) > 0$ such that $|\alpha - m/n| \geq C(\alpha)/n^2$, for any rational number m/n where α is the rotation number (Fokin, 1983).

2.4 Forced waves

Thus far only free wave motion was considered, i.e. natural modes of the system. A physically more complete model would need to incorporate some form of forcing, which excites the modes of the system. One could compare this with the example of a vibrating string, fixed at two positions, say $x = 0$ and $x = L$. The equation for the excursion of string from it's initial position is given by

$$u_{tt} = c^2 u_{xx},$$

for some constant c. Separation of variables in the form $u(x,t) = U(x)W(t)$ yields

$$W''(t)/W(t) = c^2 W(t)U''(x)/U(x).$$

Both the left and right hand side may be varied independently and we infer they must both equal a constant λ^2. The PDE is turned into two ordinary differential equations, we list them with boundary conditions,

$$\begin{aligned} W'' &= \lambda W, \\ U'' &= (\lambda/c^2)U, \\ U(0) = U(L) &= 0, \end{aligned}$$

This system is solved by the infinitude of solutions

$$\begin{aligned} U(x) &= a\sin(\lambda x/c), \\ W(t) &= b\sin(\lambda t) + d\cos(\lambda t), \end{aligned}$$

with the additional constraint (from $U(L) = 0$) that $\lambda = c^2 k\pi/L$ for integers k. It follows that the general solution for u is the infinite series

$$u(x,t) = \sum_{k=0}^{\infty} \sin(k\pi x/L)[a_i \sin(ck\pi t/L) + b_i \cos(ck\pi t/L)]. \tag{2.8}$$

Assume that the string has a given initial state $u(x,0) = f(x)$, this determines the coefficients b_i by

$$u(t,0) = \sum_{k=0}^{\infty} b_i \sin(k\pi x/L). \tag{2.9}$$

We have in mind the following identification:

- The general solution (2.8) to the vibrating string problem corresponds to the general solution $\Psi(\xi,\eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta)$ of the Poincaré equation.
- Boundary conditions are needed, setting $u(0) = u(L) = 0$ for the vibrating string corresponds to specifying $\Psi = 0$ at the boundary in the poincaré problem.
- Both problems are still under-determined. In the Poincaré problem we may prescribe either $\mathcal{F}$ or $\mathcal{G}$ on the fundamental interval, the vibrating string problem may be given an initial velocity $u_t(x,t)|_{t=0}$ to determine the coefficients a_i.

What we have now described are *free vibrations*, natural modes of the system. In order for such a solution to exist, it has to be driven. The motion of a string is set by its initial conditions and, if present, an external forcing. For internal waves the situation is likewise, the initial velocity of the string corresponds to supplying boundary conditions on the fundamental interval. External forces, however, can also be added, and moreover they are required in a viscous setting in order to maintain the waves. For the vibrating string equation, a right-hand side containing the forcing (denoted by $s(x,t)$ may be added,

$$u_{tt} - c^2 u_{xx} = s(x,t).$$

Suppose we have initial data,

$$\begin{aligned} u(x,0) &= f(x), \\ u_x(x,0) &= g(x). \end{aligned}$$

A wave equation of this form may be solved using d'Alemberts formula, it is based on an integration over the region of influence of a point (x,t), and is given by

$$u(x,t) = \frac{f(x+ct)+f(x-ct)}{2} + \int_{x-ct}^{c+ct} g(\tilde{x})\,d\tilde{x} + \int_0^t \int_{x-c(t-\tilde{t})}^{x+c(t-\tilde{t})} s(\tilde{x},\tilde{t})\,d\tilde{x}d\tilde{t}.$$

For the Poincaré equation the situation is similar, we may add a driving force to the momentum equations from which the equation was derived.

There are various types of force that may be considered. We consider two cases, leading to an inhomogeneous PDE, like the vibrating string example sketched above, and a PDE with inhomogeneous boundary conditions. If one adds a driving force $\mathbf{F} = (F^u, F^w)$ to the momentum equations, then one may derive the inhomogeneous PDE

$$\Psi_{xx} - \frac{\omega^2}{N^2-\omega^2}\Psi_{zz} = \frac{i}{\omega}(F^u_z - F^w_x).$$

Note that is immediately clear that any force for which $\nabla \times \mathbf{F} = 0$ has no effect on the solutions. The following are examples of relevant forces (for example in fluid dynamics experiments),

1. If the rotation is time dependent this induces an Euler body force $F_E = \rho(yf_t, -xf_t, 0)$ in the momentum equations. Since $\nabla \times F_E = 0$ this force has no effect.

2. If the fluid container is disturbed in the z direction, then this is equivalent to a perturbation of gravity. Therefore, we may consider a gravitational force which is time and z dependent.

Let us now seek solutions to the inhomogeneous Poincaré equation, we add the forcing $F(x,z) = (F^u_z - F^w_x)$. In characteristic variables the equation transforms to

$$\Psi_{\xi\eta} = \frac{i}{4\omega}F((\xi+\eta)/2, \lambda(\eta-\xi)/2) \equiv \frac{i}{\omega}\tilde{F}(\xi,\eta),$$

where we introduced $\tilde{F}$ to represent the transformed function F. Upon integration this yields

$$\Psi(\xi,\eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta) + \frac{i}{\omega}\int_0^\xi \int_0^\eta \tilde{F}(u,v)\,dvdu.$$

When we introduce the notation $\bar{F}$ for the integral at the right hand side, then the equation plus boundary conditions becomes

$$\begin{aligned} \Psi(\xi,\eta) &= \mathcal{F}(\xi) + \mathcal{G}(\eta) + \bar{F}(\xi,\eta), \text{ in } \Omega, \\ \mathcal{F}(\xi) + \mathcal{G}(\eta) &= -\bar{F}(\xi,\eta), \text{ at } \partial\Omega. \end{aligned}$$

In principle, we are now in a situation where we may apply the boundary-based technique developed in this chapter. The only difference being the introduction of a right hand side in the discretised equations. This observation greatly increases the applicability of our method. Section 2.4.2 will discuss the implications of forced problems further, but first we will briefly touch upon the topic of defining inflow of fluid at selected parts of the boundary.

2.4.1 Defining inflow at the boundary

This section describes how to specify inflow of fluid at a part of the boundary, which is relevant considering the roots of the Poincaré equation in physics. Denote the pieces of the boundary where we want to prescribe inflow by $\partial\Omega_i$ and suppose that these are curves parametrised by $\mathbf{c}_i(t)$, with $t \in [0,1]$. At $\partial\Omega_i$ we want to specify some velocity,

$$(u, v) = (U_i, V_i), \text{ at } \partial\Omega_i.$$

Now use line integration and the definition of the streamfunction to obtain

$$\Psi(\mathbf{c}_i(t)) - \Psi(\mathbf{c}_i(0)) = \int_0^t \nabla\Psi \cdot \mathbf{c}'(s)\, ds = \int_0^t (-V_i, U_i) \cdot \mathbf{c}'(s)\, ds.$$

Continuity of the streamfunction at the endpoints of the curves gives us additional constraints on the velocities U_i and V_i.

We proceed by considering a piecewise linear boundary and we use continuity of the streamfunction ($\Psi(\mathbf{c}_i(0) = \Psi(\mathbf{c}_i(1) = 0$). This yields

$$\Psi(\mathbf{c}_i(t)) = \int_0^t -\Delta\xi V_i(\mathbf{c}(t)) + \Delta\eta U_i(\mathbf{c}(t))\, dt.$$

Let us take the example $U_i(\mathbf{c}_i(t)) = t - 2t$ and $V_i = 0$, in this particular case

$$\Psi(\mathbf{c}_i(t)) = \Delta\eta t(1-t).$$

As will be shown later, inflow regions induce a non-zero righthand side in the discretised equations. Let $\mathbf{c}_i(t)$ parametrise $\partial\Omega_i$, and consider $(\mathcal{F}(\xi) + \mathcal{G}(\eta))$ as a function mapping $\mathbb{R}^2$ to $\mathbb{R}$. The problem may then be posed in general as

$$\begin{aligned} (\mathcal{F}(\xi) + \mathcal{G}(\eta)) \circ \mathbf{c}_i(t) &= f_i(\mathbf{c}_i(t)), \\ f_i(\mathbf{c}_i(1)) &= f_j(\mathbf{c}_j(0)) \text{ if } \mathbf{c}_i(1) = \mathbf{c}_j(0). \end{aligned}$$

Note how this defines a time independent streamfunction. The time dependence is supposed to be in form of an oscillation with frequency ω. The inflow parts of the boundary will therefore also oscillate between inflow and outflow and no fluid is added to the system. The in and outflow regions could model forcing by periodic movement of a piston, or forcing by oceanic tides.

2.4.2 Boundary forcing

The Poincaré problem as posed before describes *free vibrations*, i.e. eigenmodes of the system. In this case however the term 'eigenmode' may be inappropriate since the spectrum often has continuous parts, or the spectrum might be dense in an interval. The spectral properties of the Poincaré problem are more extensively investigated in Ralston (1973).

We present two approaches for selecting relevant solutions from the infinitude of possibilities. Firstly we present an energy minimizing regularisation procedure for the Poincaré equation

$$\begin{array}{ll} \Psi_{xx} - \lambda^2 \Psi_{zz} = 0 & \text{in } \Omega \\ \Psi = 0 & \text{at } \partial\Omega. \end{array} \tag{2.10}$$

This leads to a minimisation problem of the form

$$x_\tau = \operatorname{argmin} \left(\|Ax\|_2^2 + \tau^2 \|Lx\|_2^2 \right),$$

where A discretises the Poincaré problem and L measures the energy of the system. Sections 2.6 and 2.7 deal with this regularisation problem. Note that the energy of the system diverges due to the presence of attractors, nonetheless, in Section 2.8 we argue that the discretised measure $\|Lx\|_2^2$ for the energy is still a meaningful quantity.

One may also hold the view that the eigenmodes that occur in nature are determined by some kind of forcing. The simplest case is a forcing operating on the boundary of the domain only. Examples include tidal forcing on ocean surfaces or some mechanical forcing. The problem to be solved is in this case

$$\begin{array}{ll} \Psi_{xx} - \lambda^2 \Psi_{zz} = 0 & \text{in } \Omega, \\ \Psi = h(\xi, \eta) & \text{at } \partial\Omega, \end{array} \tag{2.11}$$

leading to a minimisation problem of the form

$$x_\tau = \operatorname{argmin} \left(\|Ax - b\|_2^2 + \tau^2 \|Lx\|_2^2 \right), \tag{2.12}$$

where b is the discrete discretised right hand side h. Note that any solution of (2.10) may be arbitrarily added to a solution of the forced problem. In Section 2.3 it was shown that the characteristics lead to functional relationships between $\mathcal{F}$ and $\mathcal{G}$. The forced problem similarly yields functional dependencies, involving the function h. We add a bar to solutions of the forced problem from here on. In terms of $\bar{\mathcal{F}}$ and $\bar{\mathcal{G}}$ the problem reads

$$\bar{\mathcal{F}} + \bar{\mathcal{G}} = h(\xi, \eta), \text{at } \partial\Omega. \tag{2.13}$$

The restrictions on $\mathcal{G}$ and $\bar{\mathcal{G}}$ induced by the characteristic connecting (ξ_1, η_1) with (ξ_1, η_2) look like

$$\begin{array}{ll} \bar{\mathcal{G}}(\eta_1) - \bar{\mathcal{G}}(\eta_2) = h(\xi_1, \eta_1) - h(\xi_1, \eta_2) & \text{for forced problem,} \\ \mathcal{G}(\eta_1) - \mathcal{G}(\eta_2) = 0 & \text{for the unforced problem,} \end{array} \tag{2.14}$$

and we see that indeed $\mathcal{G} + \bar{\mathcal{G}}$ solves the forced problem. The only open question is the role of the fundamental interval and the manner in which the fractal nature of solutions presents itself in the forced setting. In the homogeneous problem we would, for example, conclude from (2.14) that $\mathcal{G}(\eta_1)$ must be equal to $\mathcal{G}(\eta_2)$. Further analysis, as sketched above, revealed that there exist fundamental intervals where it is necessary and sufficient to prescribe $\mathcal{G}$ (or $\mathcal{F}$). Any homogeneous solution that we add to the forced problem will have fundamental intervals and associated fractal structure. Working out the functional relations for the forced problem reveals that

$$\begin{array}{rcl} \bar{\mathcal{F}}_i + \bar{\mathcal{G}}_i & = & h_{2i} \\ \bar{\mathcal{F}}_i + \bar{\mathcal{G}}_{i+1} & = & h_{2i+1}, \end{array} \tag{2.15}$$

where $\mathcal{F}_i = \mathcal{F}(\xi_i)$, $\mathcal{G}_i = \mathcal{G}(\eta_i)$ and the h_i are the function h, evaluated at the points visited while tracing the characteristics. For $\mathcal{F}_i$ and $\mathcal{G}_i$ this implies

$$\begin{array}{rcl} \bar{\mathcal{G}}_{i+1} & = & \bar{\mathcal{G}}_i + h_{2i+1} - h_{2i}, \\ \bar{\mathcal{F}}_{i+1} & = & \bar{\mathcal{F}}_i + h_{2i+2} - h_{2i+1} \end{array}$$

for the relation between the function values separated by one step. Suppose we know $\bar{\mathcal{F}}_0, \bar{\mathcal{G}}_0$, we may then solve the recursion and obtain

$$\begin{aligned} \bar{\mathcal{G}}_i &= \bar{\mathcal{G}}_0 + \sum_{j=0}^{2i-1} (-1)^{j+1} h_j \\ \bar{\mathcal{F}}_i &= \bar{\mathcal{F}}_0 + \sum_{j=1}^{2i} (-1)^j h_j, \end{aligned}$$

which tells us that boundary values are alternately added and subtracted from the initial values $\mathcal{G}_0$ and $\mathcal{F}_0$. This behaviour was noted before in Arnold & Khesin (1999). Note that by adding together the expressions we recover (2.15), as required. If the prescribed function h is non zero in the neighbourhood of the attractor, then the limit for $i \to \infty$ is likely to be non-existent. Just like in the unforced case we have fundamental intervals where values of $\bar{\mathcal{F}}$ or $\bar{\mathcal{G}}$ may be arbitrarily prescribed. The forced problem is still under determined and sensitive to the shape of the domain, thus it is still an ill-posed problem.

Note that with the inclusion of forcing terms we have resolved a peculiarity in the original model. In the unforced problem, boundary data was supplied on the fundamental intervals. This data was then reproduced, in a fractal manner, along the boundary, with the corner points of the attractor as limiting points. In reality such boundary conditions are obviously absurd. With the extension presented here, arbitrary boundary conditions, that do not have to conform to the fundamental intervals, can be applied and our model has become physically more acceptable.

2.5 Structure of solutions

Except for the classification of solutions into attractors and resonances we can get further information by viewing the Poincaré equation as a Hamiltonian system. Choose any closed curve $(\xi(t), \eta(t)) \in R$. Denote the derivative with respect to time by a dot. Take the stream function as the Hamiltonian,then

$$\dot{\Psi}(\xi, \eta) = \mathcal{F}'(\xi)\dot{\xi} + \mathcal{G}'(\eta)\dot{\eta}, \tag{2.16}$$

At curves $\{(\xi(t), \eta(t)) | \Psi = 0\}$ the stream function Ψ is constant. Since $\frac{\partial \Psi}{\partial \xi} = \mathcal{F}'$ and $\frac{\partial \Psi}{\partial \eta} = \mathcal{G}'$ we see from (2.16) that this is the case when

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \begin{pmatrix} \mathcal{G}'(\eta) \\ -\mathcal{F}'(\xi) \end{pmatrix}. \tag{2.17}$$

This system is in Hamiltonian form. We proceed by linearising $\mathcal{F}$ and $\mathcal{G}$ around $(\xi, \eta) = (0, 0)$, giving

$$\begin{aligned} \mathcal{F}'(\xi) &= \mathcal{F}'(0) + \xi\mathcal{F}''(0) + \mathcal{O}(\xi^2), \\ \mathcal{G}'(\eta) &= \mathcal{G}'(0) + \eta\mathcal{G}''(0) + \mathcal{O}(\eta^2). \end{aligned}$$

We suppose that $\mathcal{F}$ and $\mathcal{G}$ are non-degenerate at the origin, i.e. $\mathcal{F}(0)'' \neq 0$ and $\mathcal{G}(0)'' \neq 0$. The system (2.17) is now transformed to

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \begin{pmatrix} 0 & \mathcal{G}''(0) \\ -\mathcal{F}''(0) & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \begin{pmatrix} \mathcal{G}'(0) \\ -\mathcal{F}'(0) \end{pmatrix} + \begin{pmatrix} \mathcal{O}(\eta^2) \\ \mathcal{O}(\xi^2) \end{pmatrix}.$$

The coordinates may be shifted, $(x, y) = (\xi + a, \eta + b)$ to make the problem homogeneous. The choices $a = -\mathcal{F}'(0)/\mathcal{F}''(0)$ and $b = \mathcal{G}'(0)/\mathcal{G}''(0)$ lead to

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \mathcal{O}(y^2) \\ \mathcal{O}(x^2) \end{pmatrix}, \text{ where } A = \begin{pmatrix} 0 & \mathcal{G}''(0) \\ -\mathcal{F}''(0) & 0 \end{pmatrix}.$$

The eigenvalues of A are $\lambda_{1,2} = \pm\sqrt{-\mathcal{F}''(0)\mathcal{G}''(0)}$ which makes the level set of $\Psi(x, y)$ either of saddle or of center type. We call points where $\nabla\Psi = 0$ critical points, this happens for example at corner points. At these points we have potentially a local extremum for Ψ and the shift is $a = b = 0$.

However, if the origin is not critical, but degenerate ($\mathcal{F}''(0) = \mathcal{G}''(0) = 0$), then it is easy to see that

$$\xi = \mathcal{G}'(0)t + t\mathcal{O}(\eta^2) + c_1,$$
$$\eta = -\mathcal{F}'(0)t + t\mathcal{O}(\xi^2) + c_2,$$

for arbitrary constants c_1 and c_2. Close to the chosen origin the level sets of Ψ are approximately straight lines. If we have

$$\mathcal{F}'(0) = \mathcal{G}'(0) = 0 \text{ and } \mathcal{F}''(0) = 0 \text{ or } \mathcal{G}''(0) = 0, \tag{2.18}$$

then we need to look at higher order expansions. We will not consider this situation here. We summarise the results,

- Suppose $\Psi(0,0) = c$ and $\mathcal{F}''\mathcal{G}'' < 0$, then the level set of $\Psi = c$ is a hyperbola close to the critical point where $\nabla\Psi = 0$, this is a saddle point.
- Suppose $\Psi(0,0) = c$ and $\mathcal{F}''\mathcal{G}'' = 0$, then the level set of $\Psi = c$ is a curve with slope $\mathcal{F}'(0)/\mathcal{G}'(0)$ at the origin which is not a critical point.
- Suppose $\Psi(0,0) = c$ and $\mathcal{F}''\mathcal{G}'' > 0$, then the level set of $\Psi = c$ is an ellipse around the extremum where $\nabla\Psi = 0$. This extremum can only occur in the interior of the domain.

All results are valid for $(\xi, \eta) \in R$, which excludes the extreme vertices. The invariance of $\mathcal{G}$ and $\mathcal{F}$ along, respectively, the ξ and η coordinate axes leads to the following theorems.

Lemma 1. *If we are not in the case (2.18), then every corner point is a saddle point, including the extreme vertices if they are corner points.*

Proof. Every corner point is a critical point, since $\nabla\Psi = 0$ by (2.3). Take one sided derivatives at the extreme vertices. Center points are excluded since we are on the boundary. Since we are not in the case (2.18) the only option is a saddle point. □

Now, let $l(P)$ be a tangent vector to the boundary at a point $P \in \partial\Omega$, let l be multi valued at corner points.

Theorem 5. *If a point P at the boundary is a saddle point, then the points in the sequences $(T^+P, T^-T^+P, \ldots)$ and $(T^-P, T^+T^-P, \ldots)$ are also saddles. If a sequence has a finite number of distinct members, then the final point Q is such that $l^T(T^-Q)e_2 = 0$ or $l^T(T^+Q)e_1 = 0$.*

Proof. If P is a saddle then $\mathcal{F}' = 0$ and $\mathcal{F}'' \neq 0$. At the boundary point $D = T^-(P)$ this also holds. Furthermore, at this point $\nabla\Psi(D) \cdot l(P) = 0$. If the η component of $l(D)$ is not zero, it must be the case that $\mathcal{G}' = 0$. We thus have $\nabla\Psi = 0$, the point is a critical point. Since center points cannot exist on the boundary the point must be a saddle. If the η component of $l(D)$ were zero the procedure stops. If not, then we continue with the boundary point with the same η coordinate as D. □

Theorem 6. *The points (ξ_1, η_1) and (ξ_2, η_2) in R are critical points if and only if (ξ_1, η_2) and (ξ_2, η_1) are critical points in R.*

Proof. Again, use the invariance of $\mathcal{F}$ and $\mathcal{G}$ in the η and ξ directions to find that $\nabla\Psi = 0$ at (ξ_1, η_2) and (ξ_2, η_1) and that $\mathcal{F}''\mathcal{G}'' \neq 0$ at these points. □

How these results help us is perhaps best demonstrated by some examples.

Example 1 (The circle $\xi^2 + \eta^2 = 1$)**.** *The simplest solution would be one center in the middle. The extremal points with the same ξ and η coordinates have $l^T e_1 = 0$ or $l^T e_2 = 0$ and thus need not be saddle points. If we allow two critical points at the boundary, they must be at opposite extremal points or we have four critical points by Theorem 6. Of each saddle, one of the branches is part of the boundary, while the other branch crosses the boundary. Connecting each saddle to itself yields a critical point on the boundary that makes the boundary non-smooth. Therefore, the two zero level set curves that point into the circle must be connected. The circle is now divided in two cells, a center point must be inside each cell. One may continue in this fashion to obtain more possibilities. These solutions correspond topologically to the solutions given by Barcilon (1968).*

Example 2 (The rotated rectangle)**.** *Consider a rectangle of width a and length $2a$, rotated over 45 degrees. By Theorem 5 the saddles at the corner points induce saddles at the midpoints of the long sides of the rectangle. By Theorem 6 two critical points must exist inside the domain. They can be centers, giving rise to a pattern with two cells. They can not be saddles, the stable and unstable manifold would have no option but forming homoclinic connections. Inside the loop there have to be centers which in turn induce saddles on the boundary. This cannot be done without violating the flow directions.*

By placing extra saddles on the boundary and centers in the domain one obtains $j \times 2j$ cells, in accordance with the theory (Maas & Lam, 1995).

2.6 Discretisation

In this section we will develop a discretisation for equation (2.13). The discretisation will be posed in the (ξ, η)-coordinate frame, since here the Poincaré

equation becomes separable. Approximations $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{G}}$ of $\mathcal{F}$ and $\mathcal{G}$ will be sought in the the spaces $V_\xi \subset H^1([\xi^-, \xi^+])$ and $V_\eta \subset H^1([\eta^-, \eta^+])$, spanned by piecewise polynomial basis functions $\phi_i(\xi)$ respectively $\psi_j(\eta)$. In this basis we can write

$$\begin{array}{rcll} \tilde{\mathcal{F}}(\xi) & = & \sum_{i=1}^n f_i \phi_i(\xi) & \in V_\xi, \\ \tilde{\mathcal{G}}(\eta) & = & \sum_{j=1}^m g_j \psi_j(\eta) & \in V_\eta. \end{array} \tag{2.19}$$

We now define our approximation $\tilde{\Psi}$ to Ψ by

$$\tilde{\Psi}(\xi, \eta) = \tilde{\mathcal{F}}(\xi) + \tilde{\mathcal{G}}(\eta). \tag{2.20}$$

We propose a Galerkin orthogonality condition of the system (2.13) and try to find $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{G}}$ such that the residual $H = \mathcal{F} + \mathcal{G} - h$ is orthogonal to all test functions $v_1 \in V_\xi$ and $v_2 \in V_\eta$ on the boundary $\partial\Omega$. First define on $\partial\Omega$ the line integral

$$(u, v) = \int_{\partial\Omega} uv \, dl.$$

The orthogonality relation becomes

$$(H, v) \quad = \quad (\tilde{\mathcal{F}} + \tilde{\mathcal{G}}, v) - (h, v) \quad = \quad 0, \quad \forall v \in V,$$

where $V = \{w(\xi) + v(\eta) | w \in V_\xi, v \in V_\eta\}$. Equivalently, the orthogonality relation can be written

$$\begin{array}{rcll} (\tilde{\mathcal{F}}, v_1) + (\tilde{\mathcal{G}}, v_1) & = & (h, v_1) & \forall v_1 \in V_\xi, \\ (\tilde{\mathcal{F}}, v_2) + (\tilde{\mathcal{G}}, v_2) & = & (h, v_2) & \forall v_2 \in V_\eta. \end{array} \tag{2.21}$$

We have to test against all functions in V_ξ and V_η, we choose the basis functions ϕ_i and ψ_j. This yields

$$\begin{array}{rcll} \sum_{i=1}^n f_i(\phi_i, \phi_j) + \sum_{i=1}^m g_i(\psi_i, \phi_j) & = & (h, \phi_j), & j = 1, \ldots, n, \\ \sum_{i=1}^n f_i(\phi_i, \psi_j) + \sum_{i=1}^m g_i(\psi_i, \psi_j) & = & (h, \psi_j), & j = 1, \ldots, m. \end{array} \tag{2.22}$$

After choosing suitable spaces V_ξ and V_η we can work out the integrations in (2.22). This results in a matrix vector equation which we need to solve for $f = (f_1, \ldots, f_n)^T$ and $g = (g_1, \ldots, g_m)^T$:

$$\begin{pmatrix} A_1 & A_2 \\ A_2^T & A_3 \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} = A \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \tag{2.23}$$

After discretisation, the dimension of the matrix A will be much smaller than the dimension that would have been obtained by using, for example, standard finite element or finite difference methods on (2.6). For a resolution of $m \times n$

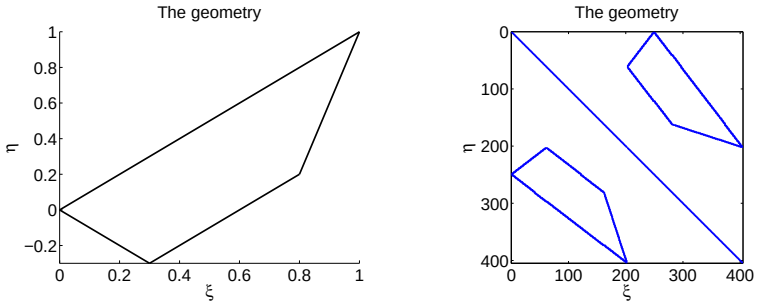


FIGURE 2.2: This figure shows a sample geometry in the left panel. To the right is a plot of the non zero elements of the matrix A. The geometry stands out clearly in the blocks A_2 and A_2^T (see (2.23)). The sub-matrices A_1 and A_3 are tri-diagonal since standard piecewise linear basis functions were used.

grid points we will only require matrices of dimension $m + n$, whereas standard methods yield a dimension of order mn. It is nice to see that the nonzero pattern of A_2 mimics the shape of the boundary. This is clear, since if $(A_2)_{ij} \neq 0$ then the intersection of the supports of ϕ_i and ψ_j contain the boundary. The sub-matrices A_1 and A_3 are banded since only the supports of $\psi_{i-N}, \ldots \psi_{i+N}$ overlap the support of ψ_i, with N depending on the specific basis functions used. See Figure 2.2 for an example.

The calculation of the matrix A is detailed in Appendix A, for the case of piecewise linear functions and a polygonal domain Ω.

The constraint $< \mathcal{F}, 1 >_{L_2([\xi^-,\xi^+])} = 0$ is easily discretised as

$$(\sum_{i=1}^{n} e_i^T) f = 0.$$

We will express this in matrix form as $Cx = 0$, with $C \in \mathbb{R}^{2(n+m)}$ and x the concatenation of the vectors f and g. The constraint $\|\Psi\|_{L_2(\Omega)} = 1$ can be discretised as $\|Nx\|_2^2 = 1$. We do not need N explicitly, since if we have a solution x with $\|Nx\| = 1$, then there is a constant c such that $\|cx\|_2^2 = 1$. The precise value of c is immaterial, as long as the norm of the solution is fixed at some value, therefore we fix the norm using

$$\|x\|_2^2 = 1. \tag{2.24}$$

Finally, we need to discretise the constraint at corner points; $\nabla\Psi = 0$. We use

finite differences as follows

$$\tilde{\mathcal{F}}'(\xi_i) = \begin{cases} \frac{f_{i+1}-f_i}{\xi_{i+1}-\xi_i} & \text{if } \xi_i = \xi^-, \\ \frac{f_i-f_{i-1}}{\xi_i-\xi_{i-1}} & \text{otherwise.} \end{cases}$$

The derivative of $\tilde{\mathcal{G}}(\eta)$ is defined accordingly. We write a matrix vector equation,

$$Px = 0. \tag{2.25}$$

Rows of P express that $f_i = f_{i-1}$ (or $f_i = f_{i+1}$ at $\xi_i = \xi^-$) or $g_j = g_{j-1}$ (or the exception $g_j = g_{j+1}$ at $\eta_j = \eta^-$) if (ξ_i, η_j) is a corner point.

Note that solutions to the discretised problem are *exact* solutions to the Poincaré equation, since they are of the form $\mathcal{F}(\xi) + \mathcal{G}(\eta)$. If the residual is non-zero, then the zero level set of this solution will not coincide with the boundary $\partial\Omega$. We obtain an exact solution on a modified boundary. Unfortunately the perturbation of the zero level set of $\mathcal{F} + \mathcal{G}$ is not necessarily small, even though $\mathcal{F} + \mathcal{G}$ is small on $\partial\Omega$. The boundary might 'fold open' if the zero level set was at a saddle point and follows nearby level sets in the discretised case.

2.7 Ill-Posedness and uniqueness

After discretisation of the Poincaré equation with $\Psi = 0$ at the boundary, the problem reduces to solving $Ax = 0$. The ill-posedness of the Poincaré equation is reflected in the matrix A. It has singular values rapidly decreasing towards zero. There are many singular values close to zero, their number increasing with increasing grid size. To these singular values correspond singular vectors that are close to the null space of A in the sense that the residual $\|Av\|_2^2$ is small for such a singular vector v. We exclude the trivial solution and we pose the problem as a minimisation of $\|Ax\|_2^2$. There is a large sensitivity of the solution x to small perturbations in the matrix A. This is the discrete analogue of the delicate dependence of the solution to changes in the boundary. We will find solutions close to solutions from the infinite dimensional solution space of the Poincaré equation, yet we cannot control which discrete vectors x we obtain. The specific vectors found are dependent on the discretisation. When the discretisation is slightly changed, for example by adding some grid points, completely different valid solutions may emerge (see Fig. 2.3).

Typically, continuous ill-posed problems have no gaps in their spectrum. The discretised problem tends to inherit this property, and a truncation of the singular values in order to decide what does or does not belong to the kernel of A becomes infeasible, since there is no clear cutoff point.

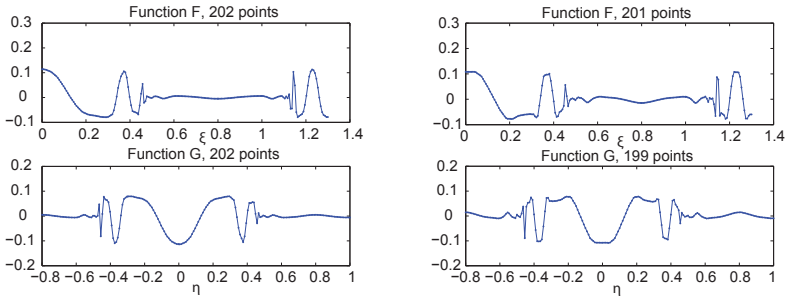


FIGURE 2.3: This figure illustrates that a small change in the discretisation can have substantial effects on the solution. Plotted are the approximations to $\mathcal{F}$ and $\mathcal{G}$. The panel to the left has 200 grid points in both the ξ and η direction, the right panel has one grid point less in both direction. The two solutions are quite different.

Also, we know that there is a delicate dependence of the solvability on the rotation number of the domain. Slight changes in the parameter λ might change the nature of the solution. We anticipate the possibility of contamination of the numerical null space by unwanted components of solutions corresponding to nearby values of λ.

We use a regularisation method that in some sense relaxes the requirement that $\mathcal{F} + \mathcal{G} = 0$ at the boundary. We find solutions where $\mathcal{F} + \mathcal{G}$ is small at the boundary. Equivalently we can say that we solve the Poincaré equation on a domain that is slightly perturbed from Ω, the level curve of $\Psi = 0$ is likely to be close to $\partial\Omega$. Instead of asking that $\mathcal{F} + \mathcal{G}$ be zero at boundary we try to minimise this quantity, while at the same time we try to obtain a smooth solution by minimising the energy. The balance between these two goals is tuned using a regularisation parameter. This parameter is picked using a tool called the *L-curve* (see e.g. Hansen, 2001), which we discuss in Section 2.7.1. The important point is that a small value of the regularisation parameter indicates that little smoothing was required, the value of λ is probably very close to a resonance. If in contrast the value is rather high, then we must be far away from a resonance. Maybe we are in the attractor case where high energy is induced by the fractal structure, or maybe the chosen value of λ does not yield solutions at all. Visual inspection of the solutions and interpretation of the L-curve usually gives information on these issues.

There exists the freedom to specify boundary conditions on the fundamental interval, which has not been incorporated in our method of discretisation and solution. We will rely on our numerical method to fill the fundamental interval. The first reason for this is that fundamental intervals are a property of the specific geometry under consideration and need to be established when-

ever a new geometry is considered. A second reason is that fundamental intervals are not unique. If M is a fundamental interval, then any $F^k M$ is also a valid fundamental interval. The question arises to which interval Dirichlet boundary conditions should be applied. The interval that contains the largest number of grid points would be a sensible choice, yet it is not clear if the solution is now unique in any numerical sense.

We will assume that the state possessing *minimal kinetic energy* is the physically relevant solution. Additionally the buoyancy energy or potential energy could be included, we will not do this for ease of presentation. One can think of the energy as an ordering principle of the solution space. We define the total kinetic energy operator $\bar{T} : C^1(\Omega) \to \mathbb{R}$ in the (x, y) system by

$$\bar{T}\bar{\Psi} = \int_{\bar{\Omega}} \nabla\bar{\Psi} \cdot \nabla\bar{\Psi}\, dx\, dy,$$

which is, in characteristic coordinates

$$\begin{aligned} T\Psi &= \int_\Omega \left[(\frac{\partial}{\partial\eta} + \frac{\partial}{\partial\xi})^2 + \lambda^2 (\frac{\partial}{\partial\eta} - \frac{\partial}{\partial\xi})^2 \right] \Psi^2 \frac{1}{2\lambda}\, d\xi d\eta \\ &= \frac{1}{2\lambda} \int_\Omega (1+\lambda^2) \left[(\mathcal{F}')^2 + (\mathcal{G}')^2 \right] + (1-\lambda^2)\mathcal{F}'\mathcal{G}'\, d\xi\, d\eta. \end{aligned}$$

We can eliminate the cross term by using integration by parts in the form

$$\int_\Omega (\partial_i u) v\, dA = \int_\Omega (\partial_i v) u dA + \int_{\partial\Omega} uvn_i\, dl,$$

where ∂_i is the i-th partial derivative and n_i the i-th component of the normal to the boundary. For the cross term we can write either

$$\int_\Omega \mathcal{F}'\mathcal{G}'\, dA = \int_{\partial\Omega} \mathcal{G}'\mathcal{F} n_1\, dl \quad \text{or} \quad \int_\Omega \mathcal{F}'\mathcal{G}'\, dA = \int_{\partial\Omega} \mathcal{F}'\mathcal{G} n_2\, dl.$$

For their sum we have

$$2\int_\Omega \mathcal{F}'\mathcal{G}'\, dA = \int_{\partial\Omega} \mathcal{F}'\mathcal{G} n_2 + \mathcal{G}'\mathcal{F} n_1\, dl,$$

and since $\mathbf{n} = \pm(\mathcal{F}', \mathcal{G}')$, and $\mathcal{F} + \mathcal{G} = 0$ at the boundary, the integral equals zero. The energy is now simply

$$T\Psi = \frac{(1+\lambda^2)}{2\lambda} \int_\Omega (\mathcal{F}'(\xi))^2 + (\mathcal{G}'(\eta))^2\, d\xi\, d\eta. \tag{2.26}$$

Note that the energy is zero if the solution Ψ is constant. These solutions are excluded by (2.4), which also makes sure that the stream function can not be

scaled to get arbitrarily small energy. We also incorporate the corner constraint (2.3). We formulate a minimisation problem:

$$\begin{array}{ll}
\text{minimise } T\Psi, & \text{in } \Omega, \text{ and} \\
\text{minimise } \Psi, & \text{at } \partial\Omega, \text{ subject to} \\
\Psi(\xi,\eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta) & \text{in } \Omega \\
< \mathcal{F}, 1 >_{L_2([\xi^-,\xi^+])} = 0, & \\
\|\Psi\|_{L_2(\Omega)} = 1, & \\
\mathcal{F}' = \mathcal{G}' = 0 & \text{at a corner point.}
\end{array}$$

As we saw in Section 2.6 this will lead to,

$$\begin{array}{ll}
\text{minimise } \|Lx\|_2^2 & \text{energy minimisation,} \\
\text{minimise } \|Ax\|_2^2 & \text{function minimisation at the boundary,} \\
\text{subject to } Cx = 0 & \text{linear 'unique representation' constraint,} \\
\text{and } \|x\|_2 = 1 & \text{normalisation,} \\
\text{and } Px = 0 & \text{linear 'corner point' constraint.}
\end{array} \tag{2.27}$$

The calculation of the matrix L is given in Appendix B. By analysing the singular value decomposition of a matrix corresponding to an ill-posed problem one can conclude that solutions corresponding to small singular values have many sign changes, and therefore possess high frequency components (see Hansen, 1994*b*) and the corresponding manual, which is also a good introduction to regularisation techniques). Some smoothing will be required. The smoothing is in the energy matrix L which contains discretisations of derivatives. If we have sufficient minimisation of the energy, the solution will be smooth. However, $\|Ax\|_2^2$ and $\|Lx\|_2^2$ cannot be minimised simultaneously. Also, by ill-posedness, there might exist a matrix close to A with much better properties in the sense that the energy $\|Lx\|_2^2$ is much lower while the residual $\|Ax\|_2^2$ is only slightly higher. Taking these considerations into account we propose to balance the residual $\|Ax\|_2^2$ and energy $\|Lx\|_2^2$ using a parameter τ. The regularised solution x_τ is then defined as

$$x_\tau = \text{argmin}(\|Ax\|_2^2 + \tau^2\|Lx\|_2^2), \text{ with } \|x\|_2^2 = 1, Cx = Px = 0, \tag{2.28}$$

The L-curve approach requires the solution of (2.28) for a range of parameters τ. In practice we choose for τ the numbers from the sequence $10^k, 10^{k-1}, \ldots, 0$ with k approximately -10 or -15 (depending on the resolution of the grid). For lack of a better method of incorporating the $Px = Cx = 0$ constraint we add them to the matrix A with a very large multiplicative factor a,

$$x_\tau = \text{argmin}(\| \begin{pmatrix} A \\ aP \\ aC \end{pmatrix} x\|_2^2 + \tau^2\|Lx\|_2^2). \tag{2.29}$$

In this way the constraints will certainly be satisfied, or the residual will become very large. Note that the value of a depends on the magnitude of the elements of A, which makes it dependent on the grid resolution. Solving the minimisation problem (2.29) is then equivalent to finding the normalised singular vector corresponding to the smallest singular value of

$$B = \begin{pmatrix} A \\ aP \\ aC \\ \tau L \end{pmatrix}. \tag{2.30}$$

In this way the constraints will certainly be satisfied, or the residual $\|A'\|_2^2$ will become very large.

2.7.1 The L-Curve

The choice of the value of the parameter τ is made using the *L-curve*. This is a continuous curve parametrised by τ and given by $\{(x(\tau), y(\tau)) | x(\tau) = \|Ax_\tau\|_2^2,\ y(\tau) = \|Lx_\tau\|_2^2\}$ where x_τ solves (2.29). The typical shape is that of an 'L', as shown in Figure 2.4. This is an idealisation, in reality the L-curve will often be much less pronounced. The ideal value of τ is found in the 'elbow' of the curve. Increasing τ would lead to a larger residual with only a slightly smaller energy (and thus a little gain in smoothness). Decreasing τ would lead to a slightly smaller residual, while the energy becomes very large, and we have less smoothness. In our experiments (see Section 2.9) we often find L-curves which have a less sharp corner, the ideal value of τ is not directly clear. Also, it might happen that the L-curve is not 'L' shaped at all. In these cases the curve can still be interpreted, and give us information on the solution. For example some curves start horizontal, only to go down for very large values of τ. A mostly flat curve tells us that the value of the regularisation parameter is insubstantial, the solution has the same energy for every reasonable τ. The specific geometry under consideration did not give us an ill-posed problem. This happens often in the resonant case.

2.8 Energy minimisation

In this section we will show that, in the presence of attractors, the energy diverges. Nevertheless, we argue that the total energy is still a useful quantity. We can give qualitative descriptions of the solutions that minimise the energy. These solutions will turn out to be continuous and differentiable, which shows that minimising the energy is a useful idea. Firstly we clarify the procedure by a one dimensional model.

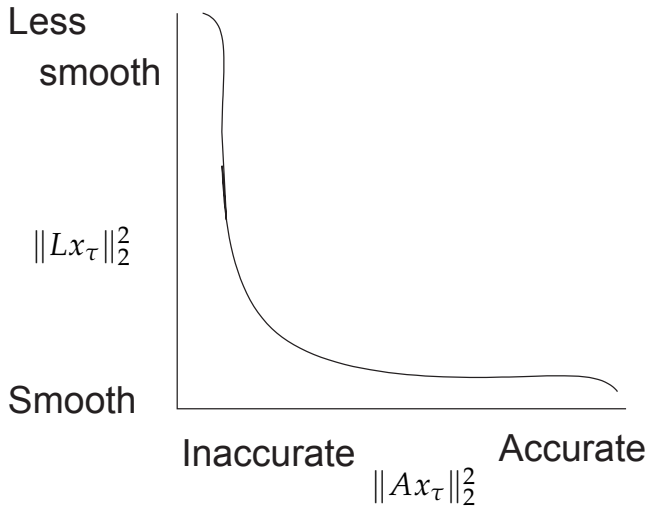


FIGURE 2.4: This figure shows a schematic L-curve parametrised by τ. On the horizontal axis is the residual, representing the accuracy of the solution. On the vertical axis is the energy, representing the smoothness of the solution.

2.8.1 A one-dimensional example

As we will show later, there are features of the minimal energy solution to the hyperbolic equation that can be predicted. The generic two dimensional case is rather complicated, for this reason this section will give a two dimensional example that reduces to a one dimensional model. The model will however exhibit the features that are also found in the full two dimensional case. Some steps may seem rather trivial, yet they closely follow the more complicated steps in the two dimensional case treated in Appendix C. Consider the wave equation $\Psi_{xx} - \lambda^2 \Psi_{zz} = 0$ on a domain Ω with $\Psi = 0$ at the boundary $\partial\Omega$. We also set the normalisation constraint $< \Psi, \Psi >= 1$. The Poincaré equation is separable in any rectangle, for illustration we use the square domain $\Omega = [0,1]^2$. In this case the Poincaré equation with the conditions stated above can be reduced to

$$\frac{f_{xx}(x)}{f(x)} = \lambda^2 \frac{g_{zz}(z)}{g(z)}, \text{ in } \Omega, \tag{2.31}$$

$$g(0) = g(1) = f(0) = f(1) = 0, \tag{2.32}$$

$$\int_0^1 f^2(x)\,dx \int_0^1 g^2(z)\,dz = 1. \tag{2.33}$$

We would like to find the solution that minimises the energy

$$\begin{aligned} E &= \int_\Omega \nabla\Psi \cdot \nabla\Psi \, dxdy \\ &= \int_0^1 (f_x)^2 \, dx \int_0^1 g^2 \, dy + \int_0^1 f^2 \, dx \int_0^1 (g_y)^2 \, dy. \end{aligned}$$

Use the normalisation constraint (2.33) to see that minimising E is equivalent to minimising

$$\frac{\int_0^1 (f_x)^2 \, dx}{\int_0^1 f^2 \, dx} + \frac{\int_0^1 (g_y)^2 \, dy}{\int_0^1 g^2 \, dy}.$$

Now this is really two separate minimisation problems, one for $f(x)$ and one for $g(y)$. We will concentrate on the function f and put $\delta = \int_0^1 (g_y)^2 \, dy$. Since we can vary x and y independently we can suppose the right hand side of (2.31) equals γ^2. The simplified model now reads

$$\begin{aligned} &f_{xx}(x) = \gamma^2 f(x), \\ &\|f(x)\|^2_{L_2(\Omega)} = \delta, \\ &f(0) = f(1) = 0. \end{aligned} \tag{2.34}$$

We need to find the solution that minimises

$$\tilde{E} = \delta \int_0^1 (f_x)^2 \, dx.$$

The differential equation plus boundary conditions (2.34) is solved by

$$f(x) = \sum_{k=0}^{\infty} a_k \sin(\gamma_k x),$$

with $\gamma_k = k\pi$. By (2.33) the $(a_i) \in \ell^2$ are subject to the constraint $\sum_{i=0}^{\infty} a_i^2 \neq 0$. Consequently the trivial solution is excluded, yet there are still infinitely many solutions. In order to pick one solution out of the infinite set of degenerate solutions we minimise the energy. Write the energy as

$$\tilde{E} = \delta \int_0^1 (f_x)^2 \, dx = < f, -\delta \frac{d^2}{dx^2} f >_{L_2([0,1])} .$$

The Courant-Fischer theorem now gives us that the minimum of $\tilde{E}$ is given by the eigenfunction belonging to the smallest eigenvalue of the operator $L = -\delta \frac{d^2}{dx^2}$, i.e. the function f corresponding to the smallest eigenvalue μ that solves

$$f_{xx}(x) = -\frac{\mu}{\delta} f(x). \tag{2.35}$$

This equation is of the same form as our differential equation (2.34), which is unfortunate from a didactical point of view: in the full two dimensional model the situation is less trivial, and the Courant-Fisher theorem is a useful tool. The eigenvalue problem (2.35) has solution

$$f(x) = \alpha \cos(\sqrt{\frac{\mu}{\delta}}x) + \beta \sin(\sqrt{\frac{\mu}{\delta}}x).$$

From boundary conditions (2.32) we see that α must be zero and μ is quantised as $\mu_k = \delta k^2 \pi^2$ with $k = 1, 2, 3, \ldots$. Zero is excluded since the trivial solution is excluded. The eigenfunction f belonging to μ_k is then given by

$$f(x) = \beta_k \sin(k\pi x).$$

The constant β_k is found by solving $\int_0^1 f^2(x)\,dx = 1/d$ and we have a unique solution for given k. Considering solutions that belong to $\mu_1, \mu_2, \ldots$ we see that the energy orders the solutions by decreasing smoothness (in the sense of number of oscillations). The energy therefore has a regularising property.

Furthermore we see that we find continuous and differentiable solutions 'for free', without explicitly enforcing these properties. These features carry over to the two dimensional situation. However, when dealing with the two dimensional hyperbolic equation in an arbitrary non-rectangular domain we have two additional issues

- The domain is arbitrary and we no longer have $\Psi(x, y) = f(x)g(y)$. However in characteristic coordinates we do know $\Psi(\xi, \eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta)$.

- Attractors will occur and the fractal behaviour of the solution introduces high energies.

Nonetheless, it will be shown that energy minimisation still leads to smooth regularised solutions. We will first give a one-dimensional analogue of the energy in the presence of an attractor and consider the general situation in Appendix C.

2.8.2 Behaviour near an attractor

Close to the attractor features of the solution are repeated at increasingly smaller scales. This makes the energy blow up, but we argue that we nonetheless still have a valid minimisation problem. We consider a one dimensional model. For

an arbitrary function $f \in H^1([0,1])$ define:

$$\begin{aligned} \Psi(x) &= \sum_{i=0}^{\infty} f_i(x), \\ f_i(x) &= \begin{cases} f(\frac{x-\alpha_i}{\alpha_{i+1}-\alpha_i}), & \text{for } x \in [\alpha_i, \alpha_{i+1}], \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Here $[\alpha_0, \alpha_1]$ models the fundamental interval, with $\alpha_2, \alpha_3, \ldots$ iterates of its endpoints under the mapping T, as defined in Section 2.3. Usually the contents of an iterate of the fundamental interval is mirrored and scaled by the action of the mapping. We will not model the mirroring effect here, but suppose that an interval $[\alpha_i, \alpha_{i+1}]$ contains both the original function and its mirrored copy. The differences $\alpha_{i+1} - \alpha_i$ are steadily decreasing and there is an 'attractor' at $\lim_{i\to\infty} \alpha_i = \alpha$. The functions f_i are scaled and translated copies of f, this models the replication of the fundamental interval towards the attractor. The energy on the interval $[\alpha_0, \alpha_N]$ is modeled by

$$\begin{aligned} T([\alpha_0, \alpha_N)] &= \int_{\alpha_0}^{\alpha_N} (\Psi_x)^2 \, dx \\ &= \int_0^1 f'(x)^2 \, dx \sum_{i=0}^{N} \frac{1}{\alpha_{i+1} - \alpha_i}. \end{aligned}$$

Here we used the disjoint supports of f_i in order to write the square of sums as a sum of squares. There are two ways in which the energy may become infinite. Firstly the derivative of the function f itself may not be square integrable, but such functions do not minimise the energy. Secondly the sum in the second factor diverges, but for every finite value of the number of steps towards the attractor N, the expression is finite. Moreover, minimisation of the energy yields *the same* function f independently of N. Therefore this function f may also be considered a solution for $N \to \infty$. Thus, even though we have a diverging total energy it is still useful for minimisation purposes.

From a numerical point of view, we propose that discrete solutions on a grid can be compared to the continuous solution on $[\alpha_0, \alpha_N]$ for finite N. Equivalently one can say we always keep a finite distance to the attractor. The fine scale structure which causes the energy to diverge close to the attractor at limit point α is simply replaced with one basis function in the discretisation. Since solutions are independent of N our numerical solutions will correspond to solutions of the continuous system. In this way it makes sense to minimise the energy on a fixed grid, however the energy of discrete solutions will increase with increasing grid resolution in the presence of attractors.

2.8.3 The two dimensional case

In the two dimensional case the calculation is more tedious, we will only state the results of the calculation here and give the technical details in Appendix C. Before giving the solution we introduce some notation. The α_i and β_i variables give ξ and η coordinates of the iterates of the fundamental interval. The index i counts the number of iterates towards the attractor and $k \leq K$ counts the fundamental intervals. The index $l \leq L$ takes care of the possibility that iterates of endpoints of a fundamental interval may approach a limit cycle with period L. For fixed k and l the sequence $\alpha_i^{l,k}$ approaches one fixed point of the attractor from one side, $\cap_i(\alpha_i^{l,k}, \alpha_{i+1}^{l,k}) = \emptyset$ and $\cup_i[\alpha_i^{l,k}, \alpha_{i+1}^{l,k}]$ has no holes. The general solution can be described using wavenumber-like numbers (m, n) and arbitrary constants $A_n^{k,l}$ and $B_m^{k,l}$. It is given by

$$\Psi_{m,n}(\xi, \eta) = \mathcal{F}_n(\xi) + \mathcal{G}_m(\eta),$$

with

$$\begin{aligned} \mathcal{F}_n = &\quad \sum_{i=0}^{N}\sum_{k=0}^{K}\sum_{l=0}^{J} A_n^{l,k} I_{\xi\in[\alpha_i^{l,k},\alpha_{i+1}^{l,k}]} \cos(n\pi \frac{\xi - \alpha_i^{l,k}}{\Delta\alpha_i^{l,k}}), \\ \mathcal{G}_m = &-\sum_{i=0}^{M}\sum_{k=0}^{K}\sum_{l=0}^{J} B_m^{l,k} I_{\eta\in[\beta_i^{l,k},\beta_{i+1}^{l,k}]} \cos(m\pi \frac{\eta - \beta_i^{l,k}}{\Delta\beta_i^{l,k}}), \end{aligned} \tag{2.36}$$

where I denotes the indicator function. The above formulas describe a solution which is built from arbitrary Fourier expansions on the fundamental intervals, which are reproduced in smaller scales towards the attractor. Note that $A_n^{k,l}$ can not all be zero. The numbers N, M determine the size of neighborhoods of the attractors that we leave out of the domain. For $N, M \to \infty$ we face a diverging energy, the above solutions are not valid. However, for every large *finite* values N, M we do have the solutions given above. With increasing N, M the solution does not change, except it's domain of definition grows, towards the attractor. For more details we refer to Appendix C.

Some words on the space in which the functions $\mathcal{F}$ and $\mathcal{G}$ live are in order. In Section 2.3 we proposed $\mathcal{F} \in L_2([\xi^-, \xi^+])$ and $\mathcal{G} \in L_2([\eta^-, \eta^+])$, giving also $\Psi \in L_2(\Omega)$. The solution is discontinuous at the attractor, this follows easily from the definition of continuity and the fact that attractors are limit cycles of the transformation F. Theorem 4 established that solutions exist, and thus that the discontinuity is square integrable. In this section we need to calculate an energy, the suitable space would seem to be $H^1(\Omega)$. However, in this space the energy diverges when we insist on solving on the entire domain Ω (i.e. $N, M \to \infty$). The proper space seems to be $H^1(\Omega \setminus (\mathcal{A} \cup \mathcal{B})$, where $\mathcal{A}$ consists of a neighborhood of the attractors. The set $\mathcal{A}$ is in our approach left out of the

domain by using finite N and M as shown above. It could also be the case that $\mathcal{F}(\xi)$ or $\mathcal{G}(\eta)$ are discontinuous. The set $\mathcal{B}$ consists of horizontal and vertical lines corresponding to these discontinuities. This is only possible if $N, M \to \infty$ in (2.36), and we do not need to worry about excluding $\mathcal{B}$ from the domain since we only consider finite values.

It is important to realise that $\mathcal{A}$ and $\mathcal{B}$ are completely different in nature. Attractors are a feature of the problem, but we know how to deal with this. Discontinuities in $\mathcal{F}$ and $\mathcal{G}$ away from the attractor will not occur because the minimising functions (2.36) are automatically continuous.

2.8.4 Minimal energy of the discretised system

Consider the behaviour of the solution to the discrete system. When piecewise smooth basis functions are used, then across the attractor there will be a piecewise smooth approximation of the self similar structure. We suggest that this discrete system corresponds, to good approximation, to solving the continuous system where we only minimise the energy on the domain given by a finite number of iterates of the fundamental interval. This corresponds to taking N and M to be finite numbers. We have now taken away the contributions to the energy that caused it to diverge, and have in effect removed a small strip around the attractors. Still we know that the minimal energy solution we find is independent of the values of M and N, if $M \gg 1$ and $N \gg 1$. If the grid has sufficient resolution, then we will find a meaningful solution to the Poincaré equation. The energy will increase with decreasing grid size, yet for a fixed grid we can safely minimise the energy. If differences between calculated solutions on grids of increasingly higher resolution become smaller, then we can have confidence that we are converging towards a minimal energy solution of the continuous problem.

2.9 Results

This section will give some solutions of the Poincaré equation using the discretisation from Section 2.6 and the regularisation procedure outlined in Section 2.7. Firstly, we briefly consider the solutions in the rectangle, which is a suitable geometry for testing our algorithm. Next, we study the trapezoid and we consider the case where the stream function is zero on the entire boundary, i.e. the unforced setting. After this we apply a non-zero boundary condition for the stream function and find that the scenario sketched in Section 2.4.2 indeed occurs.

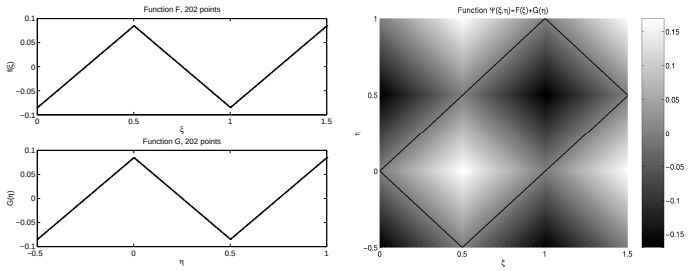


FIGURE 2.5: In the absence of regularisation the solution converges to a solution that contains discontinuities in the gradients.

2.9.1 Results for the rectangle

The rectangle is an excellent test object, since we know all there is to know about the solutions. We perform an experiment with $\lambda = 1/2$, giving the $(n, 2n)$ modes for $n = 1, 2, \ldots$, the mode having the lowest energy is the $(1, 2)$-mode. The number of grid points is $(n, m) = (202, 202)$. It turns out that if no regularisation is performed, that one obtains a solution as in Figure 2.5, which has kinks in the stream function. In terms of the velocities this would mean an unacceptable instantaneous change of direction. We plot the solution $\Psi(\xi, \eta)$ in the (ξ, η) frame, this is the coordinate system in which it is computed. Also the solution is plotted outside of the domain (indicated by black lines). Figure (2.6) shows the singular value distribution of A which has the hallmark of an ill-posed problem, values slowly decreasing towards zero. The same figure also shows the L-Curve which has in this case a clear optimum. Figure 2.7 gives results for various regularisation parameters. Experiments have shown that the value for the regularisation parameter is fairly independent of the mesh size. Solutions also converge with decreasing grid spacing.

There is a special case for which regularisation is even more important. It may happen that the number of grid points is such that the coordinates of corner points coincide with the grid points. In this case no extra points have to be introduced to handle the corners. Upon closer inspection of these cases it turns out that the orbits become decoupled. They are decoupled in the continuous case if the rotation number is rational, now they are also numerically independent. See Figure 2.8 for examples of a representative situation. This decoupling leads to a non zero dimension of the null space of A. Regularisation becomes even more important since the unregularised solutions will look meaningless. Inspection of the L-curve in Figure 2.9 learns that we may pick any small value for τ. Results are shown in Figure 2.10.

We see that even for the relatively simple case of the rectangle we need

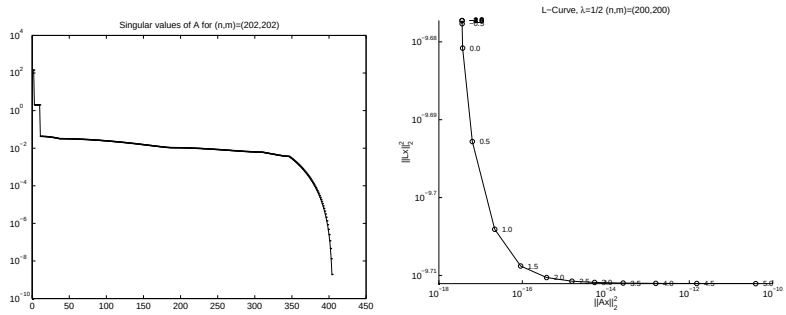


FIGURE 2.6: Singular values and the L-Curve for the square with $\lambda = 1/2$ and $(m, n) = (202, 202)$. The singular values decrease towards zero. The L-curve has the characteristic 'L' shape.

regularisation techniques. It seems that regularisation is essential for obtaining relevant solutions to this ill-posed problem.

Energy of solutions in the rectangle

We will now investigate the the minimal kinetic energy solution. We use the general solution (1.27). The energy can be computed as

$$\begin{aligned} T(\Omega) &= \int_\Omega u^2 + v^2 \, dS \\ &= \int_\Omega \nabla\Psi \cdot \nabla\Psi \, dS \\ &= \int_\Omega (\sum_{j=0}^{\infty} a_j \alpha_j \cos(\alpha_j x) \sin(\beta_j z))^2 \\ &+ (\sum_{j=0}^{\infty} a_j \beta_j \sin(\alpha_j x) \cos(\beta_j z))^2 \, ds. \end{aligned}$$

Since the sine and cosines are orthogonal in the sense that

$$\begin{aligned} \int_0^a \sin(\alpha_i x) \sin(\alpha_j x) \, dx &= a\delta_{ij}/2, \\ \int_0^b \cos(\beta_i z) \cos(\beta_j z) \, dz &= b\delta_{ij}/2, \end{aligned}$$

the expression can be simplified after some calculation to yield

$$T = \frac{ab\pi^2}{4} (\frac{n^2}{a^2} + \frac{m^2}{b^2}) \sum_{j=1}^{\infty} a_j^2 j^2.$$

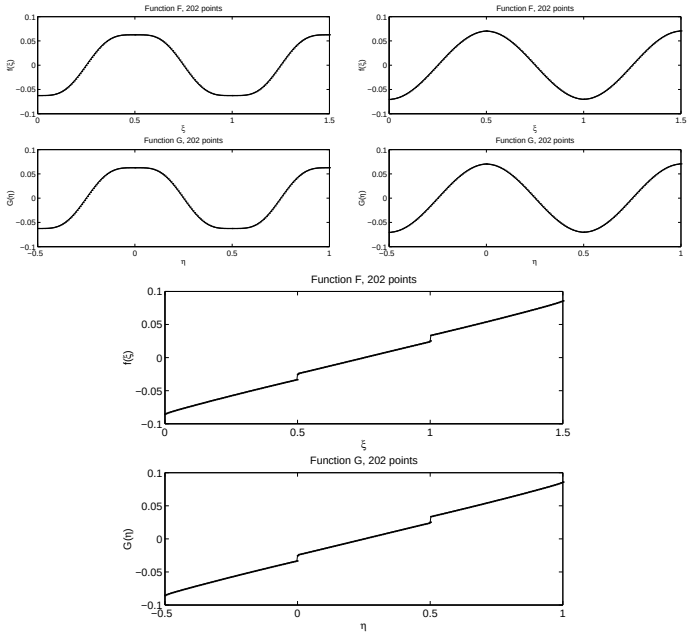


FIGURE 2.7: From the curve of Figure(2.6) three values of τ were chosen. The first picture shows an under regularised solution for $\tau = 10^{-8}$, the second picture has a value of $\tau = 10^{2}$ which is in the elbow of the curve and the final picture is heavily over-regularised with $\tau = 10^{5}$.

Without the additional constraint $||\Psi||_{L_2} = 1$ the energy would not attain a non-trivial minimum, now however it gives us the extra equation $\sum_{j=1}^{\infty} a_j^2 = 1$ and the minimum is at $a_1 = 1$ with all other a_j's equal to zero. This shows that the smallest possible mode is indeed the preferred solution when minimizing the energy.

2.9.2 Free oscillations

We will discuss the trapezoid, a rectangle with one sloping side as shown in (x, z) coordinates in Figure 2.16. This geometry is interesting since it is a simple geometry for which attractors exist. The complex distribution of attractor regimes separated by modal solutions (a manifestation of the ill-posed nature of the problem) is extensively described in Maas *et al.* (1997). For $\lambda = 0.75$ we are in the attractor case, indeed Figure 2.16 (as later discussed) clearly shows features that reproduce in increasingly smaller scales towards a limit cycle. The

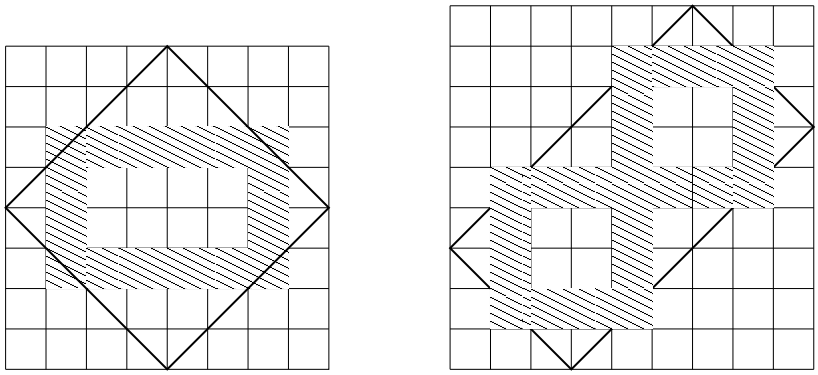


FIGURE 2.8: This picture shows the rectangle for $\lambda = 1$ and $\lambda = 1/2$. Equidistant grids are overlayed that coincide with the corner points. It is possible to track closed orbits having the thickness of the grid spacing. Even in a numerical sense the solution is at resonance. The number of grid points required is $2k + 3$ for the $(1, 1)$ mode and $3k + 4$ for the $(2, 1)$ mode. It is even possible to read of the dimension of the null space of the corresponding matrix A: 7 respectively 3 for the cases above.

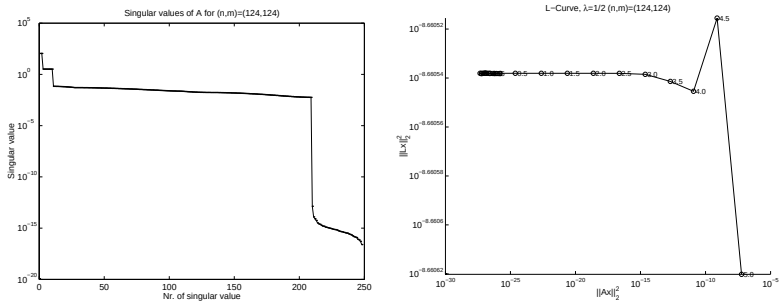


FIGURE 2.9: Singular values and the L-Curve for the square with $\lambda = 1/2$ and $(m, n) = (124, 124)$. The L-curve is strikingly different from the previous case. The regularisation parameter can be chosen as small as one wishes (but nonzero). The precise value has no large effect on the smoothness. The last 39 singular values are the numerical null space.

behaviour of the dynamical system induced by F was investigated by a ray-tracing technique, which enables us to find the values of λ where the various attractors exist.

We take 300 grid points in both the ξ and η direction and plot L-curves of the first four solutions. This means that we plot the L-curves of the first few solutions that minimise (2.29) best. This simply amounts to calculating the four smallest singular values of the matrix (2.30). The curves are approximated by

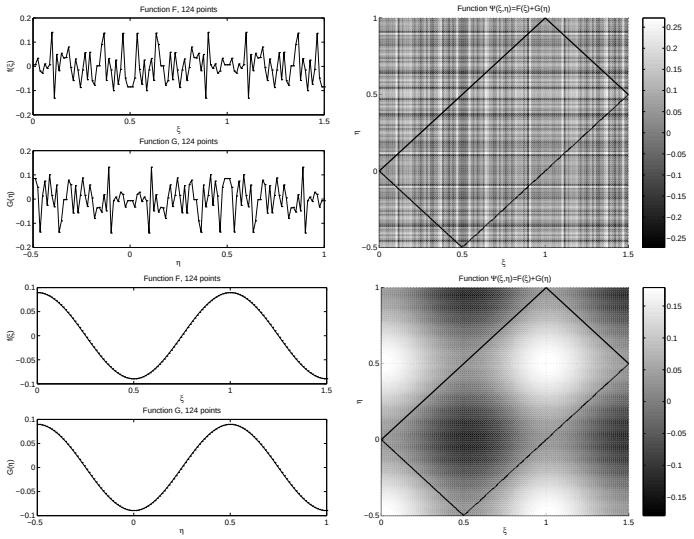


FIGURE 2.10: In contrast to the previous case it is found that regularisation is crucial for meaningful solutions. The top picture shows a random solution from the 39 dimensional solution space. The bottom picture shows a regularised solution with $\tau = 10^{-4}$, it is clearly more acceptable.

taking a sequence of values for τ, we choose $\{10^j | j = -10, -9.8, -9.6, \ldots, -0.8, -1\}$. The L-curves are given in Figure 2.11, Figure 2.12 gives enlargements of the parts of the curves where we pick the regularisation parameter.

We will discuss the solutions for the different L-curves and point to noteworthy features of the solutions.

The 'plateau' solution

From the bottom curve of Figure 2.11, the top-left curve of Figure 2.12, representing the minimising solution, we choose a value of $\tau \approx 10^{-5}$. We plot the functions $\mathcal{F}$ and $\mathcal{G}$ and their sum in Figure 2.13. We observe a solution with a large step at the attractor, featuring small oscillations. Two questions come to mind, firstly if such a solution is allowed. Secondly if the small oscillations are a numerical artifact (reminiscent of the Gibbs phenomenon) or a real feature of the solutions. The latter question is answered by observing the structure of the oscillations: they are mapped towards the attractor like the oscillations in figures 2.14 and 2.17. We conclude that the structure is 'mixed in' from another solution, the regularisation is not perfect. The question of the validity is more

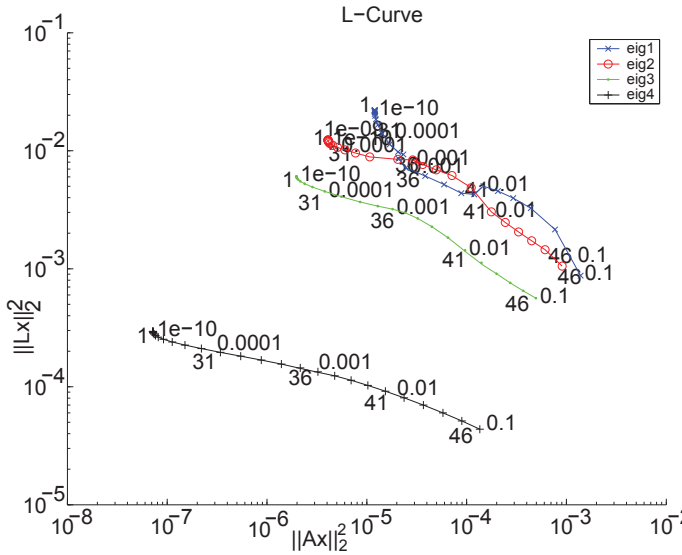


FIGURE 2.11: Four L-curves for the trapezoid. The 'elbows' of the curves do not look very pronounced when plotted in one picture. One would have expected a longer vertical 'leg', instead we find an accumulation point. A number of points on the curves have been calculated and labeled with the value of the regularisation parameter τ and an index. Enlargements of relevant curves are presented in Figure 2.12.

complicated. The solution is of the form (2.36), with large coefficients for the constant term and very small coefficients for other terms. In the domain outside of the attractor the values $\mathcal{F}$ plus $\mathcal{G}$ cancel, while inside of the attractor they add up to a nonzero value. Physically this represents a solution where (almost) all energy is located at the attractor. A problem is that the analytical analogue of this solution is not from the space $H^1(\Omega)$, the energy would be infinite. As argued in Section 2.8 we ought to compare discrete solutions in Ω to analytical solutions in Ω minus a neighborhood of the attractor, and in this sense the solution is valid. In Swart & Loghin (2005b) it is shown that this type of solution is also found upon inclusion of a viscous (dissipative) term in the Poincaré equation. This is an indication that this might also physically be a relevant solution. The experiment described in Maas *et al.* (1997) is a case in point, the existence of a wave attractor was established, but no evidence of fractal structure was found.

Wavenumber 1 solution

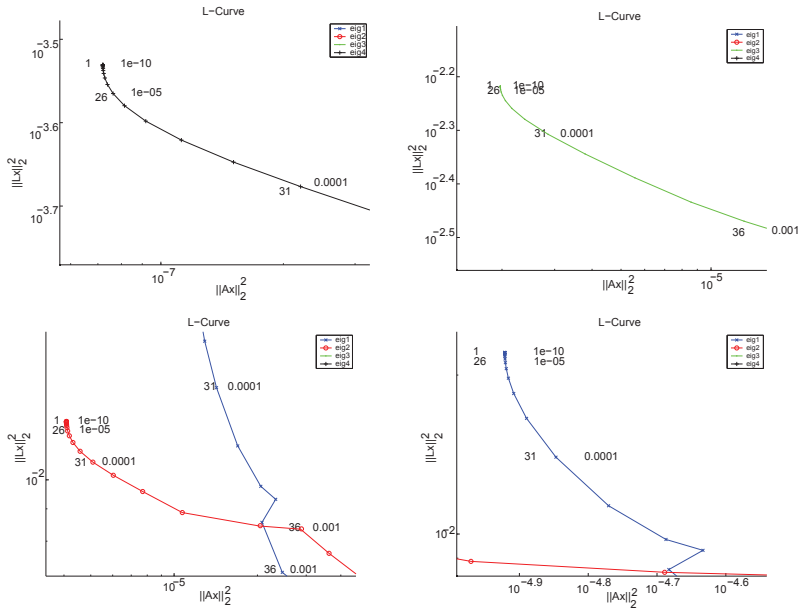


FIGURE 2.12: These four panels show the L-curves corresponding to the best four solutions. The indices of the regularisation parameters are (in clockwise order): 26,31,31,31, corresponding to $\tau \approx 10^{-5}, 10^{-4}, 10^{-4}, 10^{-4}$.

The next-best solution we find is shown in Figure 2.14, where we plot the solution corresponding to $\tau \approx 10^{-4}$ on the second curve (Figure 2.12). It clearly consists of one empty fundamental interval plus one fundamental interval with half a cosine. Referring to (2.36) we call this a wavenumber 1 solution for this fundamental interval. We will now demonstrate the effect of over- and under-regularisation by picking inappropriate values of τ. The result of this practice is shown in Figure 2.15. Other experiments have revealed even more sensitive dependence of the solution on the choice of the regularisation parameter. When the grid has sufficient resolution one may also observe oscillations that are introduced or suppressed when changing the regularisation parameter. These observations are ultimately the justification for the use of a regularisation procedure.

Combined solutions

Figure 2.11 shows the curves corresponding to the four best minimizing solu-

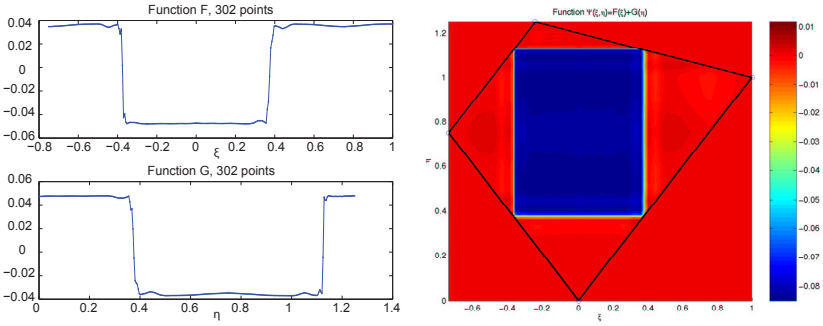


FIGURE 2.13: This figure shows the functions $\mathcal{F}$ and $\mathcal{G}$ and a plot of $\Psi = \mathcal{F} + \mathcal{G}$, in the (ξ, η) coordinate frame. Darker colors indicate higher values, gray is at the zero level.

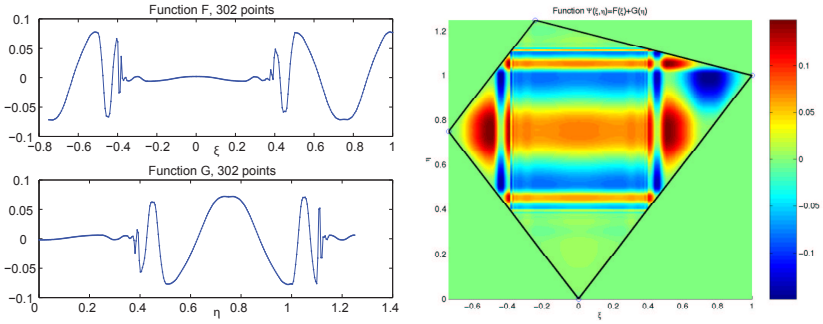


FIGURE 2.14: This figure shows the functions $\mathcal{F}$ and $\mathcal{G}$ and a plot of $\Psi = \mathcal{F} + \mathcal{G}$, in the (ξ, η) coordinate frame. One fundamental interval shows a smooth function (half the period of a cosine), the other one is zero.

tions. The top two curves cross each other twice, which might seen strange at first sight, but there is no reason why a under regularised n^{th} best solution cannot have a higher energy than an over-regularised $(n+1)^{th}$ best solution. The top-most crossover of the curves is in fact an artifact of the way the L-curves are plotted. For each value of τ we store a column of coordinates $(\|Ax_\tau\|_2^2, \|Lx_\tau\|_2^2)$ in a matrix, we then plot the rows of this matrix. The right panel of Figure 2.11 shows that the eigenvalue curves and L-curves switch at index 35.

At eigenvalue curve three we find the solution where the empty fundamental interval has switched places with the wavenumber 1 interval. If we add both solutions (see Figure 2.16), then we fill both fundamental intervals and recover the solution that was presented in Maas & Lam (1995).

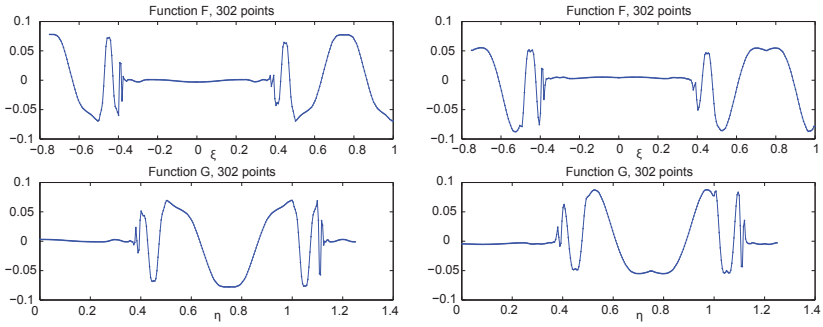


FIGURE 2.15: These graphs show what happens if one regularises improperly. In the left panel a value of $\tau \approx 3 \times 10^{-3}$ was chosen, yielding a smooth but inaccurate solution. The extrema are too flat. The right panel corresponds to $\tau \approx 10^{-5}$, the residual is smaller but the energy is higher. We see cusps appearing, giving a discontinuity in the derivative.

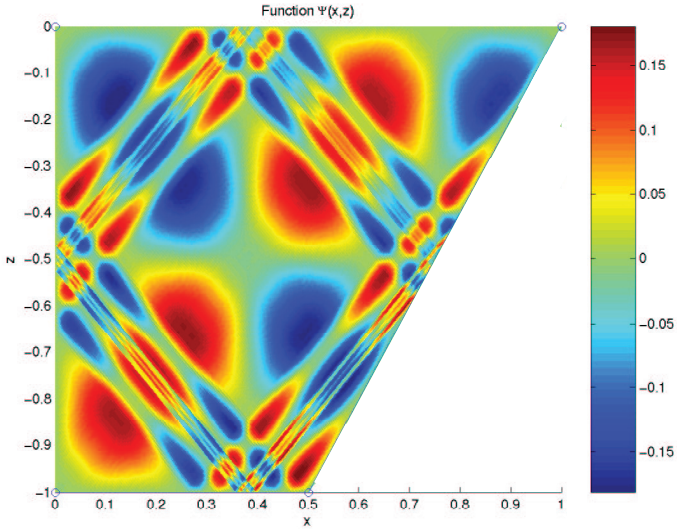


FIGURE 2.16: This figure shows the addition of the smoothest solutions found on each of the fundamental intervals. The grid resolution is 300 grid points in the ξ and η directions. Shown is the stream function $\tilde{\Psi}$, in the (x, z) frame. Darker colors indicate higher values, gray is at the zero level. The attractor, and the fractal structure are nicely visible.

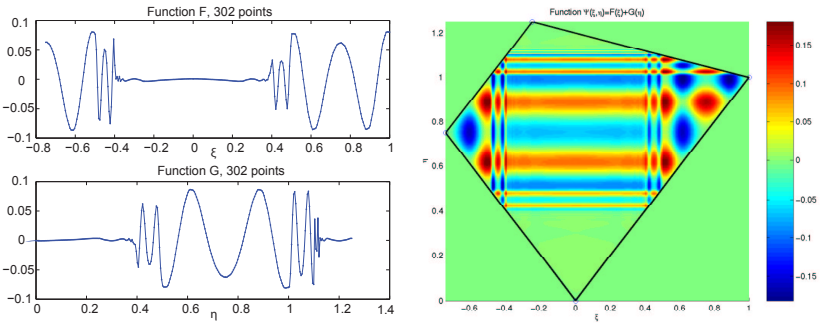


FIGURE 2.17: In this figure, again one fundamental interval shows a smooth function, but now a cosine of higher periodicity. The other interval is (almost) zero.

Wavenumber 2 solution

The solution we find, at $\tau \approx 10^{-4}$, curve four of Figure 2.12, is a solution where again one of the fundamental intervals is empty. The other one however is now filled with a complete period of a cosine (Figure 2.17). We remark that the order in which we find solutions is not a priori clear. In this case we first found both the 'wavenumber 1' solutions, we could also have found a 'wavenumber 2' solution first. The exact order is dependent on the energy of the solution, which is (among other things) determined by the sizes of the fundamental intervals.

We conclude that we find the solutions predicted by equation (2.36), where the $\mathcal{F}_n$ and $\mathcal{G}_m$ parts are indeed decoupled. The regularisation scheme is indispensable in obtaining meaningful solutions to the ill-posed Poincaré equation.

2.9.3 Boundary forcing

In this section we apply a boundary condition $\Psi = 1$ to the left side of the trapezoid. We have to deal with equation (2.11), which leads to the regularisation problem (2.12). In contrast to the unforced case, with zero right hand side b, this is a standard problem which may be tackled using existing methods. We draw the L-curve and calculate solutions using the *Regularization Tools*, described in Hansen (1994*a*). In this section we focus on the piecewise constant solution, since we learned in the previous section that this is energetically the preferred solution.

We calculate a solution at $\lambda = 0.75$, which is in the regime where the square attractor lives. The number of grid points is 512 in both the ξ and η direction.

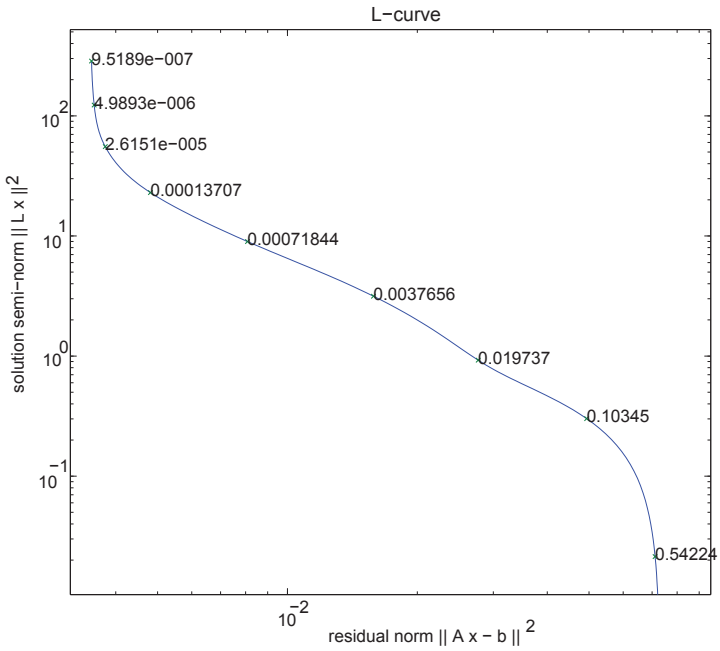


FIGURE 2.18: The L-curve for the Poincaré problem with non-zero boundary condition on part of the boundary.

The L-curve calculated for this problem is shown in Fig. 2.18. This L-curve has a clearer optimal point than the curves presented before, although it is still necessary to try a few values around the bend in the curve before deciding on a 'best solution'. Results for the functions $\mathcal{F}$ and $\mathcal{G}$ are shown in Figure 2.19, where the effects of over-regularisation and under-regularisation are nicely visible.

Finally we present a figure of the complete solution $\Psi(\xi,\eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta)$ in Figure 2.20. The theoretically predicted behaviour (Section 2.4.2) is clearly observed.

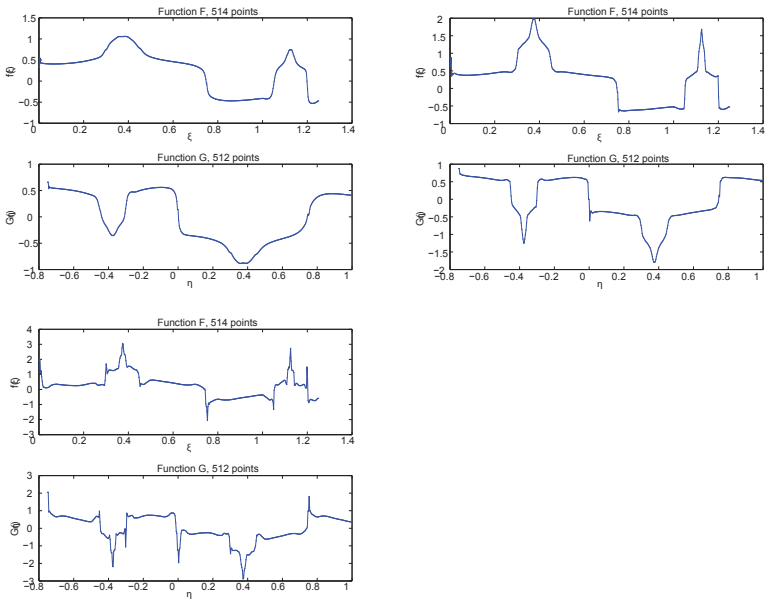


FIGURE 2.19: These panels show the functions $\mathcal{F}$ and $\mathcal{G}$ for three different values of the regularisation parameter. To the left we see under-regularisation, a value of $\tau = 5 \times 10^{-4}$ was used, which is below the optimum (see also Fig. 2.18). The graphs are clearly too smooth. A value of $\tau = 8 \times 10^{-5}$ seems close to optimal and yields the graphs in the middle panel. Note how the functions $\mathcal{F}$ and $\mathcal{G}$ have jumps of ± 1 as predicted in the theory. Over-regularisation is shown in the right panel, at $\tau = 1 \times 10^{-5}$, the functions suffer from too many sharp peaks.

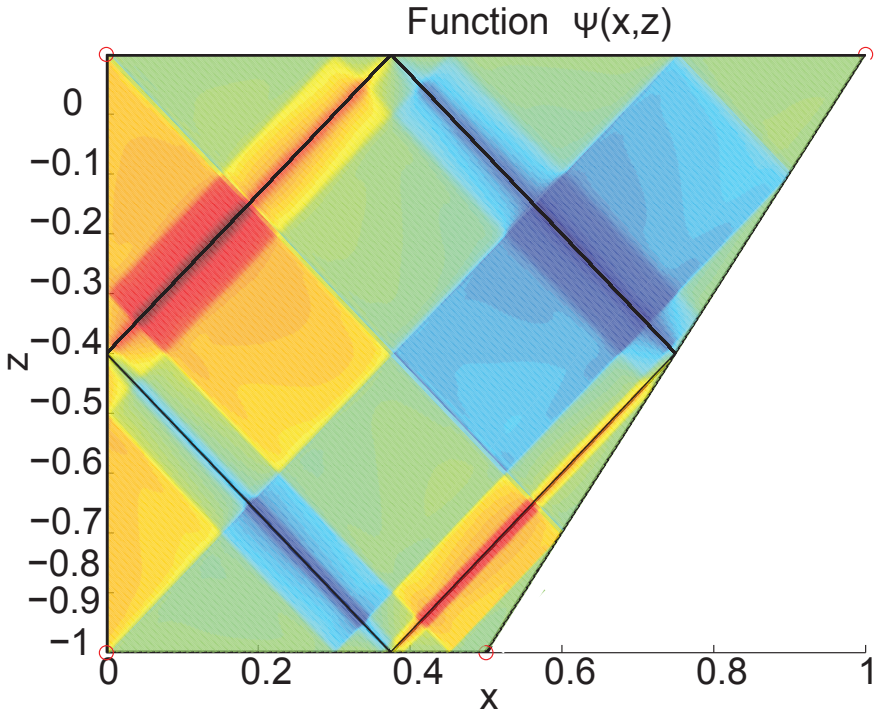


FIGURE 2.20: A plot of the stream function for the forced problem. At the vertical left boundary $\Psi = 1$ is assigned, the remaining boundary has $\Psi = 0$. It is nicely visible how the characteristics 'pick up' values of ± 1 when visiting the side where $\Psi = 1$. Also, we clearly see focusing towards the attractor, where values of $|\Psi|$ get increasingly higher due to continued visits to the left boundary.

Appendices to Chapter 2

2.A Calculation of the matrix A

This appendix gives the calculation of the matrix A for the special case of a polygonal domain Ω and piecewise linear basis functions. The aim is to develop a fast algorithm. We first make some definitions, intended to facilitate the conversion from formula to Matlab code. For this reason indices start at one. We try casting expressions in matrix-equation form to benefit from Matlab's coarse granularity. First of all we introduce the grid spacings

$$\begin{aligned} \Delta\xi &= (\xi_2, \ldots, \xi_n)^T - (\xi_1, \ldots, \xi_{n-1})^T \in \mathbb{R}^{n-1}, \\ \Delta\eta &= (\eta_2, \ldots, \eta_m)^T - (\eta_1, \ldots, \eta_{m-1})^T \in \mathbb{R}^{m-1}. \end{aligned}$$

Denote the element wise product of matrices A and B by $A \cdot B$. If the quotient of two matrices is taken, element wise division is implied. We will also use the notation $\mathrm{diag}(A)$ for the diagonal of a square matrix and $\mathrm{diag}(a)$ for the matrix with the vector a on it's diagonal. For square matrices we will also use $\mathrm{lowerdiag}(A)$ and $\mathrm{upperdiag}(A)$. Also introduce $\mathbf{1}_n = \sum_{i=1}^{n} e_i$. Furthermore we will need convenient abbreviations for appending zeros to matrices and removing rows and columns. These are needed to incorporate the boundaries.

Definition 4. *In the following the zero vector matches the size of the identity matrix I_{m-1}. The matrix A is $m \times n$ and the notation $[A, B]$ signifies concatenation of B to the right of A.*

$$\begin{array}{lrcl} \textit{Remove column 1,} & A[0, I_{m-1}]^T & \equiv & AI_l', \\ \textit{Remove column n,} & A[I_{m-1}, 0]^T & \equiv & AI_r', \\ \textit{Remove row 1,} & [0, I_{n-1}]A & \equiv & I_t'A, \\ \textit{Remove row n,} & [I_{n-1}, 0]A & \equiv & I_b'A, \\ \textit{Append zeros left,} & A[0, I_m] & \equiv & AI_l, \\ \textit{Append zeros right,} & A[I_m, 0] & \equiv & AI_r, \\ \textit{Append zeros top,} & [0, I_n]^T A & \equiv & I_tA, \\ \textit{Append zeros bottom,} & [I_n, 0]^T A & \equiv & I_bA. \end{array}$$

The basis functions will be split in ascending and descending parts, they

are described by

$$\begin{aligned}
\phi_i^-(\xi) &= \begin{cases} \frac{\xi-\xi_{i-1}}{\Delta\xi_{i-1}} & \text{if } \xi \in [\xi_{i-1},\xi_i] \text{ and } 2 \le i \le n, \\ 0 & \text{otherwise.} \end{cases} \\
\phi_i^+(\xi) &= \begin{cases} \frac{\xi_{i+1}-\xi}{\Delta\xi_i} & \text{if } \xi \in [\xi_i,\xi_{i+1}] \text{ and } 1 \le i \le n-1, \\ 0 & \text{otherwise.} \end{cases} \\
\psi_j^-(\eta) &= \begin{cases} \frac{\eta-\eta_{j-1}}{\Delta\eta_{j-1}} & \text{if } \eta \in [\eta_{j-1},\eta_j] \text{ and } 2 \le j \le m, \\ 0 & \text{otherwise.} \end{cases} \\
\psi_j^+(\eta) &= \begin{cases} \frac{\eta_{j+1}-\eta}{\Delta\eta_j} & \text{if } \eta \in [\eta_j,\eta_{j+1}] \text{ and } 1 \le j \le m-1, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Now define the rectangles R_{ij} and strips P_j and S_i by

$$R_{ij} = [\alpha_i,\alpha_{i+1}] \times [\beta_j,\beta_{j+1}].$$

Suppose the polygonal domain Ω consists of segments P_i for $1 \le i \le l$. It is now very convenient to have grid points ξ_i and η_j for each corner of the boundary. In this way the intersection of a segment P_i of the boundary with R_{ij} is also a segment. We now need parametrisations of these segments in order to evaluate the integrations. Let the unit parametrisation of $R_{ij} \cap P_l$ be given by

$$c_{ij}^l(t) = \begin{cases} (F_{ij}^\xi + tD_{ij}^\xi, F_{ij}^\eta + tD_{ij}^\eta) & \text{if } R_{ij} \cap P_l \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Also define the lengths L_{ij}^l by

$$L_{ij}^l = \|{c_{ij}^l}'(t)\|_2 = \sqrt{\left(D_{ij}^{l,\eta}\right)^2 + \left(D_{ij}^{l,\xi}\right)^2}.$$

Given corner points $(\hat{\xi}_1,\hat{\eta}_1),\ldots,(\hat{\xi}_l,\hat{\eta}_l)$ of the domain and the grid points it is easy to calculate the matrices $D^{k,\xi}$, $D^{k,\eta}$ and L^k. See Algorithm 2 for this

procedure. We now ready to proceed with the calculation of the matrix A_2

$$\begin{aligned}
(A_2)_{ij} &= \int_{\partial\Omega} \phi_i \psi_j \, dl \\
&= \int_{\partial\Omega} \phi_i^- \psi_j^- + \phi_i^- \psi_j^+ + \phi_i^+ \psi_j^- + \phi_i^+ \psi_j^+ \, dl \\
&= \int_0^1 \sum_{\substack{k=1 \\ i,j\neq 1}}^{l} L_{i-1,j-1}^k (\phi_i^- \psi_j^-) \circ c_{i-1,j-1}^k(t) + \sum_{\substack{k=1 \\ i\neq 1, j\neq m}}^{l} \int_0^1 L_{i-1,j}^k (\phi_i^- \psi_j^+) \circ c_{i-1,j}^k(t) \\
&+ \sum_{\substack{k=1 \\ i\neq n, j\neq 1}}^{l} \int_0^1 L_{i,j-1}^k (\phi_i^+ \psi_j^-) \circ c_{i,j-1}^k(t) + \sum_{\substack{k=1 \\ i\neq n, j\neq m}}^{l} \int_0^1 L_{i,j}^k (\phi_i^+ \psi_j^+) \circ c_{ij}^k(t) \, dt.
\end{aligned}$$

Using the definition of the basis functions we obtain

$$\begin{aligned}
(A_2)_{ij} &= \sum_{\substack{k=1 \\ i,j\neq 1}}^{l} \frac{L_{i-1,j-1}^k}{6\Delta\xi_{i-1}\Delta\eta_{j-1}} \Big\{ 6(F_{i-1,j-1}^{k,\xi} - \xi_{i-1})(F_{i-1,j-1}^{k,\eta} - \eta_{j-1}) \\
&\quad + 3\Big[D_{i-1,j-1}^{k,\xi}(F_{i-1,j-1}^{k,\eta} - \eta_{j-1}) + D_{i-1,j-1}^{k,\eta}(F_{i-1,j-1}^{k,\xi} - \xi_{i-1}) \Big] \\
&+ 2D_{i-1,j-1}^{k,\xi} D_{i-1,j-1}^{k,\eta} \Big\} \\
&- \sum_{\substack{k=1 \\ i\neq 1, j\neq m}}^{l} \frac{L_{i-1,j}^k}{6\Delta\xi_{i-1}\Delta\eta_{j}} \Big\{ 6(F_{i-1,j}^{k,\xi} - \xi_{i-1})(F_{i-1,j}^{k,\eta} - \eta_{j+1}) \\
&\quad + 3\Big[D_{i-1,j}^{k,\xi}(F_{i-1,j}^{k,\eta} - \eta_{j+1}) + D_{i-1,j}^{k,\eta}(F_{i-1,j}^{k,\xi} - \xi_{i-1}) \Big] + 2D_{i-1,j}^{k,\xi} D_{i-1,j}^{k,\eta} \Big\} \\
&- \sum_{\substack{k=1 \\ i\neq n, j\neq 1}}^{l} \frac{L_{i,j-1}^k}{6\Delta\xi_{i}\Delta\eta_{j-1}} \Big\{ 6(F_{i,j-1}^{k,\xi} - \xi_{i+1})(F_{i,j-1}^{k,\eta} - \eta_{j-1}) \\
&\quad + 3\Big[D_{i,j-1}^{k,\xi}(F_{i,j-1}^{k,\eta} - \eta_{j-1}) + D_{i,j-1}^{k,\eta}(F_{i,j-1}^{k,\xi} - \xi_{i+1}) \Big] + 2D_{i,j-1}^{k,\xi} D_{i,j-1}^{k,\eta} \Big\} \\
&+ \sum_{\substack{k=1 \\ i\neq n, j\neq m}}^{l} \frac{L_{i,j}^k}{6\Delta\xi_{i}\Delta\eta_{j}} \Big\{ 6(F_{i,j}^{k,\xi} - \xi_{i+1})(F_{i,j}^{k,\eta} - \eta_{j+1}) \\
&\quad + 3\Big[D_{i,j}^{k,\xi}(F_{i,j}^{k,\eta} - \eta_{j+1}) + D_{i,j}^{k,\eta}(F_{i,j}^{k,\xi} - \xi_{i+1}) \Big] + 2D_{i,j}^{k,\xi} D_{i,j}^{k,\eta} \Big\}.
\end{aligned}$$

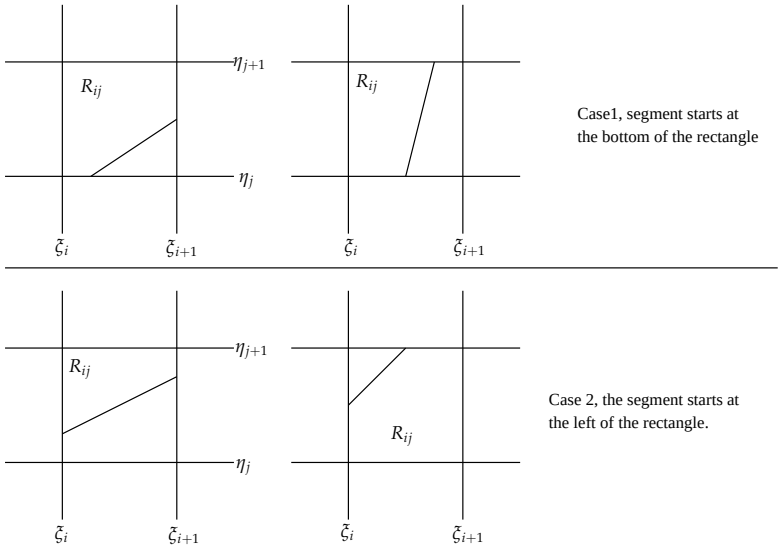


FIGURE 2.21: The four ways in which a segment with positive slope can intersect the rectangle R_{ij}

After some reordering, we write down the expression for A_2

$$
\begin{aligned}
A_2 &= \sum_{k=1}^{l} I_t \frac{L^k}{\Delta\xi\Delta\eta^T} \cdot \left\{ ((F^{k,\xi} - I'_b\xi\mathbf{1}_{m-1}^T) + \frac{1}{2}D^{k,\xi}) \cdot ((F^{k,\eta} - \mathbf{1}_{n-1}\eta^T I'_r) + \frac{1}{2}D^{k,\eta}) \right\} I_l \\
&- I_b \frac{L^k}{\Delta\xi\Delta\eta^T} \cdot \left\{ ((F^{k,\xi} - I'_b\xi\mathbf{1}_{m-1}^T) + \frac{1}{2}D^{k,\xi}) \cdot ((F^{k,\eta} - \mathbf{1}_{n-1}\eta^T I'_l) + \frac{1}{2}D^{k,\eta}) \right\} I_l \\
&- I_t \frac{L^k}{\Delta\xi\Delta\eta^T} \cdot \left\{ ((F^{k,\xi} - I'_t\xi\mathbf{1}_{m-1}^T) + \frac{1}{2}D^{k,\xi}) \cdot ((F^{k,\eta} - \mathbf{1}_{n-1}\eta^T I'_r) + \frac{1}{2}D^{k,\eta}) \right\} I_r \\
&+ I_b \frac{L^k}{\Delta\xi\Delta\eta^T} \cdot \left\{ ((F^{k,\xi} - I'_t\xi\mathbf{1}_{m-1}^T) + \frac{1}{2}D^{k,\xi}) \cdot ((F^{k,\eta} - \mathbf{1}_{n-1}\eta^T I'_l) + \frac{1}{2}D^{k,\eta}) \right\} I_r \\
&+ (I_t - I_b) \frac{L^k \cdot D^{k,\xi} \cdot D^{k,\eta}}{12\Delta\xi\Delta\eta^T} (I_r - I_l).
\end{aligned}
$$

Next we turn to the matrix A_1 containing integration of the ϕ basis func-

Algorithm 2 This algorithm finds all intersections of a boundary segment with rectangles R_{ij}. We consider the case of positive slope s. One starts at the beginning of a line segment and increments the ξ or η coordinate. Depending on the size of R_{ij} and s there are two cases (see Figure 2.21) to consider, distinguished by $c = 1$ or $c = 2$. While traversing the segment the matrices containing the parametrisation are built.

Initialise $c \leftarrow 1$
Find (i, j) for which (ξ_i, η_j) is start of segment
Initialise $(\xi, \eta) \leftarrow (\xi_i, \eta_j)$
Initialise $F^{\xi}, F^{\eta}, D^{\xi}, D^{\eta}, L$ as empty matrices
while not end of line piece **do**
  $(F^{\xi}_{ij}, F^{\eta}_{ij}) \leftarrow (\xi, \eta)$
  if c==1 **then**
    if $s < \Delta\eta_j/(\xi_{i+1} - \xi)$ **then**
      $\eta \leftarrow \eta + s(\xi_{i+1} - \xi)$
      $\xi \leftarrow \xi_{i+1}$
      $c \leftarrow 2$
    else
      $\xi \leftarrow \xi + \Delta\eta_j/s$
      $\eta \leftarrow \eta_{j+1}$
      $c \leftarrow 1$
    end if
  else if $s < (\eta_{j+1} - \eta)/\Delta\xi_i$ **then**
    $\eta \leftarrow \eta + s\Delta\xi)$
    $\xi \leftarrow \xi_{i+1}$
    $c \leftarrow 2$
  else
    $\xi \leftarrow \xi + (\eta_{j+1} - \eta)/s$
    $\eta \leftarrow \eta_{j+1}$
    $c \leftarrow 1$
  end if
  $(D^{\xi}_{ij}, D^{\eta}_{ij}) \leftarrow (\xi, \eta) - (F^{\xi}_{ij}, F^{\eta}_{ij})$
  $L_{ij} \leftarrow \|(D^{\xi}_{ij}, D^{\eta}_{ij})\|_2^2$
end while

tions,

$$\begin{aligned}(A_1)_{ij} &= \int_{\partial\Omega} \phi_i \phi_j \, dl \\ &= \int_{\partial\Omega} (\phi_i^+ + \phi_i^-)(\phi_j^+ + \phi_j^-) \, dl \\ &= \int_{\partial\Omega} \phi_i^+ \phi_j^+ + \phi_i^+ \phi_j^- + \phi_i^- \phi_j^+ + \phi_i^- \phi_j^- \, dl\end{aligned}$$

Since basis function ϕ_i overlaps only ϕ_{i-1} and ϕ_{i+1} it follows that A_1 is a tridiagonal matrix.

$$\begin{aligned}(A_1)_{ii} &= \int_{\partial\Omega} \phi_i^+ \phi_i^+ + \phi_i^- \phi_i^- \, dl, \\ (A_1)_{i,i-1} &= \int_{\partial\Omega} \phi_i^- \phi_{i-1}^+ \, dl, \\ (A_1)_{i,i+1} &= \int_{\partial\Omega} \phi_i^+ \phi_{i+1}^- \, dl.\end{aligned}$$

We now need a parametrisation $\tilde{c}_i^l(t)$ of $\{P_l \cap [\xi_i, \xi_{i+1}]\}$. Fortunately, we are only interested in the first component which is simply $\xi_i + t\Delta\xi_i$. Another needed quantity, the total length of the intersection of the strip with a boundary segment is also easy to obtain, as the row sum of row i of the length matrix L^k

$$\int_0^1 \|\tilde{c}_i^{k\prime}(t)\|_2^2 \, dt = \sum_{j=1}^{m} e_i^T L^k e_j = \tilde{L}_i^k.$$

This gives

$$\begin{aligned}(A_1)_{ii} &= \sum_{\substack{k=1 \\ i\neq n}}^{l} \int_0^1 (\phi_i^+ \phi_i^+) \circ \tilde{c}_i^k \tilde{L}_i^k \, dt + \sum_{\substack{k=1 \\ i\neq 1}}^{l} \int_0^1 (\phi_i^- \phi_i^-) \circ \tilde{c}_{i-1}^k \tilde{L}_{i-1}^k \, dt \\ &= \frac{1}{3} \sum_{\substack{k=1 \\ i\neq n}}^{l} \tilde{L}_i^k + \frac{1}{3} \sum_{\substack{k=1 \\ i\neq 1}}^{l} \tilde{L}_{i-1}^k, \\ (A_1)_{i,i-1} &= \sum_{\substack{k=1 \\ i\neq 1}}^{l} \int_0^1 (\phi_i^- \phi_{i-1}^+) \circ \tilde{c}_{i-1}^k \tilde{L}_{i-1}^k \, dt = \frac{1}{6} \sum_{\substack{k=1 \\ i\neq 1}}^{l} \tilde{L}_{i-1}^k, \\ (A_1)_{i,i+1} &= \sum_{\substack{k=1 \\ i\neq n}}^{l} \int_0^1 (\phi_i^+ \phi_{i+1}^-) \circ \tilde{c}_i^k \tilde{L}_i^k \, dt = \frac{1}{6} \sum_{\substack{k=1 \\ i\neq n}}^{l} \tilde{L}_i^k.\end{aligned}$$

We can compactly write the matrix A_1 if we introduce the vector $\tilde{L}^k$ with components $\tilde{L}_i^k$. in this case

$$A_1 = \frac{1}{6} \sum_{k=1}^{l} \text{tridiag}(\tilde{L}^k, 2\tilde{L}^k I_r + 2\tilde{L}^k I_l, \tilde{L}^k).$$

For A_3 we have by a similar calculation

$$A_3 = \frac{1}{6} \sum_{k=1}^{l} \text{tridiag}(\bar{L}^k, 2\bar{L}^k I_r + 2\bar{L}^k I_l, \bar{L}^k),$$

with $\bar{L}_j^k \sum_{i=1}^n e_i^T L^k e_j$.

2.B Calculation of the energy norm matrix

A measure of the total kinetic energy of the system will be used to regularise the solution. We aim to solve

$$x_\tau = \text{argmin}\{\|Ax\|_2^2 + \tau^2 \|Lx\|_2^2\},$$

for a range of parameters τ. This section describes the construction of the matrix L. It is such that $\|Lx\|_2^2$ is the total kinetic energy of the discretised solution as defined by (2.26). Since we minimise the energy, we can set $\lambda = 1$ without loss of generality.

The discretised solution is piecewise linear, which means that the derivatives are piecewise constants given by

$$\begin{aligned} \frac{d\tilde{\mathcal{F}}(\xi)}{d\xi} &= \frac{f_{i+1} - f_i}{\Delta\xi_i} \equiv f_i^\xi && \text{for } \xi \in [\xi_i, \xi_{i+1}], \\ \frac{d\tilde{\mathcal{G}}(\eta)}{d\eta} &= \frac{g_{j+1} - g_j}{\Delta\eta_j} \equiv g_j^\eta && \text{for } \eta \in [\eta_j, \eta_{j+1}]. \end{aligned}$$

Also define the vectors

$$\begin{aligned} f^\xi &= (f_1^\xi, \dots f_{n-1}^\xi)^T, \\ g^\eta &= (g_1^\eta, \dots, g_{m-1}^\eta)^T, \end{aligned}$$

and write

$$\begin{aligned} f^\xi = Ff = (\frac{1}{\text{diag}(\Delta\xi)} I_l - \frac{1}{\text{diag}(\Delta\xi)} I_r) f \quad \text{with} \quad F \in \mathbb{R}^{n-1 \times n}, \\ g^\eta = Gg = (\frac{1}{\text{diag}(\Delta\eta)} I_l - \frac{1}{\text{diag}(\Delta\eta)} I_r) g \quad \text{with} \quad G \in \mathbb{R}^{m-1 \times m}. \end{aligned}$$

A discretised measure for the energy is then

$$\tilde{T}(\Omega) = \sum_{j=1}^{m-1} \sum_{i=1}^{n-1} O_{ij} \left[(g_j^\eta)^2 + (f_i^\xi)^2 \right],$$

where O_{ij} is the area of $\Omega \cap R_{ij}$. We want to write the expression for the energy in the form

$$\tilde{T}(\Omega) = \|Lx\|_2^2 = \begin{pmatrix} f \\ g \end{pmatrix}^T \begin{pmatrix} F^T \text{diag}(O\mathbf{1}_m) F & 0 \\ 0 & G^T \text{diag}(\mathbf{1}_n^T O) G \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}.$$

It turns out that we do not need O explicitly, only its column and row sums. These are the areas of the strips P_j and S_i, defined by

$$\begin{aligned} S_i = &\quad ([\alpha_i, \alpha_{i+1}] \times [\eta^+, \eta^-]) \cap \Omega, \\ P_j = &\quad ([\xi^+, \xi^-] \times [\beta_j, \beta_{j+1}]) \cap \Omega. \end{aligned} \tag{2.37}$$

These area vectors, denote them by h and v, are easily calculated using the matrices D_{ij}^{ξ} and D_{ij}^{η}, which we already have from Algorithm 2. Using these vectors we can write for the matrix L

$$L = \begin{pmatrix} \sqrt{\text{diag}(v)}F & 0 \\ 0 & \sqrt{\text{diag}(h)}G. \end{pmatrix}$$

2.C Two dimensional energy minimisation

From (2.26) we have

$$T(\Omega) = A \int_{\Omega} (\mathcal{F}'(\xi))^2 + (\mathcal{G}'(\eta))^2 \, d\xi d\eta, \tag{2.38}$$

where $A = \frac{(1+\lambda^2)}{2\lambda} \in \mathbb{R}$. In the following we assume that there exist $J+1$ fundamental intervals at which $\mathcal{F}$ or $\mathcal{G}$ must be specified. The boundaries of the fundamental intervals are iterated by the map F defined in Section 2.3. Consider the projection of the iterates onto the ξ axis. For a fundamental interval l these will approach a fixed point of certain period K, associated with an attractor. We can reorder them to obtain for each fundamental interval K sequences of coordinates, converging to the K fixed points. Let them occur as

$$\begin{aligned} &\alpha_0^{l,k}, \alpha_1^{l,k}, \ldots \quad \text{for the } \xi \text{ coordinate, with} \quad \lim_{i\to\infty} \alpha_i^{l,k} = a^{l,k}, \\ &\beta_0^{l,k}, \beta_1^{l,k}, \ldots \quad \text{for the } \eta \text{ coordinate, with} \quad \lim_{j\to\infty} \beta_j^{k,l} = b^{l,k}, \end{aligned}$$

for $l = 0, \ldots, J, k = 0 \ldots K$. See Figure 2.22 for an example domain. The iterated fundamental intervals cover the entire domain, except for the attractor:

$$\begin{aligned} \cup_{i,l,k} [\alpha_i^{l,k}, \alpha_{i+1}^{l,k}) &= [\xi^-, \xi^+] \setminus \cup_{l,k} \{a^{l,k}\}, \\ \cup_{j,l,k} [\beta_j^{l,k}, \beta_{j+1}^{l,k}) &= [\eta^-, \eta^+] \setminus \{b^{l,k}\}. \end{aligned}$$

Furthermore, fundamental intervals have the property that they do not overlap and also that the above sequences define intervals that do not overlap. Mathematically this is expressed by

$$\begin{aligned} (\cup_{k,i} \alpha_i^{l_1,k}) \cap (\cup_{k,i} \alpha_i^{l_2,k}) &= \emptyset \text{ for } l_1 \neq l_2, \\ (\cup_i \alpha_i^{l,k_1}) \cap (\cup_i \alpha_i^{l,k_2}) &= \emptyset \text{ for } k_1 \neq k_2. \end{aligned}$$

In order to avoid notational clutter the following discussion will be for one fundamental interval and one sequence only. When we introduce a function or variable it is understood that we need J times K functions and variables. Thus we write for the iterates

$$\begin{aligned} &\alpha_0, \alpha_1, \ldots \quad \text{for the } \xi \text{ coordinate, with} \quad \lim_{i\to\infty} \alpha_i = a, \\ &\beta_0, \beta_1, \ldots \quad \text{for the } \eta \text{ coordinate, with} \quad \lim_{j\to\infty} \beta_j = b. \end{aligned}$$

Define functions $f, g : [0,1] \to \mathbb{R}$, these play the role of the arbitrary functions that must be prescribed. We can define copies of the function f on $[\alpha_i, \alpha_{i+1}]$ and g on $[\beta_j, \beta_{j+1}]$, these will be dilated and translated versions of the original function. The functions reverse their orientation at every iteration[2] , and copies $f(x)$ and $f(1-x)$ alternate. We have for $f_i : [\alpha_i, \alpha_{i+1}] \to \mathbb{R}$ and $g_j : [\beta_j, \beta_{j+1}] \to \mathbb{R}$ the expressions

$$\begin{aligned} f_i(x) &= \begin{cases} f(\frac{x-\alpha_i}{\Delta\alpha_i}), & \text{if } i \text{ is even,} \\ f(\frac{\alpha_{i+i}-x}{\Delta\alpha_i}), & \text{if } i \text{ is odd,} \end{cases} \\ g_j(x) &= \begin{cases} g(\frac{x-\beta_j}{\Delta\beta_j}), & \text{if } j \text{ is even,} \\ g(\frac{\beta_{j+1}-x}{\Delta\beta_j}), & \text{if } j \text{ is odd,} \end{cases} \end{aligned} \tag{2.39}$$

where $\Delta\alpha_i = \alpha_{i+1} - \alpha_i$, $\Delta\beta_j = \beta_{j+1} - \beta_j$. When integrating the distinction is not needed since

$$\int_{\alpha_i}^{\alpha_{i+1}} f(\frac{x-\alpha_i}{\Delta\alpha_i})\, dx = \int_{\alpha_i}^{\alpha_{i+1}} f(\frac{\alpha_{i+i}-x}{\Delta\alpha_i})\, dx.$$

Fundamental intervals and their iterates can never overlap, thus in the end we can set the functions $\mathcal{F}$ and $\mathcal{G}$ to be of the form

$$\begin{aligned} \mathcal{F}(\xi) &= \sum_{k=0}^{K} \sum_{l=0}^{J} \sum_{i=0}^{N(k,l)} f_i^{k,l}(\xi), \\ \mathcal{G}(\eta) &= \sum_{k=0}^{K} \sum_{l=0}^{J} \sum_{j=0}^{M(k,l)} g_j^{k,l}(\eta), \end{aligned} \tag{2.40}$$

where M and N are integers depending on k and l. Note that *at* the attractor we have a problem, the solution will become multi valued. This could be circumvented by defining f and g on open intervals, but this would render the

[2]This effect can be observed by tracing a characteristic from the endpoint of the fundamental interval. By construction this orbit will end in a corner point. Starting slightly to the left of this endpoint, one gets close to the corner where it acts as a mirror and reflects the characteristic back to the right of the fundamental interval.

functions $\mathcal{F}$ and $\mathcal{G}$ undefined in an infinite number of points. If f and g are continuous, then also $\mathcal{F}$ and $\mathcal{G}$ are continuous (at least for finite M, N). The choice of M and N determines the support of $\mathcal{F}(\xi) + \mathcal{G}(\eta)$ and we can use this freedom to remove neighborhoods of attractors. For finite N, M we have left out a strip around the attractor. To calculate the energy on the whole of Ω we let N and M approach infinity.

We have to take care at $x = \alpha_i$ and $x = \alpha_{i+1}$, the derivative of f_i is not defined. We define the derivative in such a way that $\mathcal{F}' = (\sum_{i=0}^{\infty} f_i)'$ holds. Note that we have continuity at the iterates of the endpoints of the fundamental interval since they reverse orientation at every step. Because $\mathcal{F}' = \mathcal{G}' = 0$ at a corner point, and also at every point reached from a corner point by iterating the map F, we have that the derivative of $\mathcal{F}$ at the endpoints of iterates of fundamental intervals is zero (boundaries of fundamental intervals always trace back to corner points). Thus it holds that $f'(0) = f'(1) = g'(0) = g'(1) = 0$. We define

$$\begin{aligned} \frac{df_i(x)}{dx} &= \begin{cases} 0 & \text{if } x = \alpha_{i+1} \text{ or } x = \alpha_i, \\ \frac{1}{\Delta\alpha_i} f'(\frac{x-\alpha_i}{\Delta\alpha_i}) & \text{otherwise.} \end{cases} \\ \frac{dg_j(x)}{dx} &= \begin{cases} 0 & \text{if } x = \beta_{j+1} \text{ or } x = \beta_j, \\ \frac{1}{\Delta\beta_j} g'(\frac{x-\beta_i}{\Delta\beta_j}) & \text{otherwise.} \end{cases} \end{aligned} \tag{2.41}$$

Now plug in the series (2.40) in the energy (2.38) to get

$$T_{NM}(\Omega) = A(\sum_{i=0}^{N} < \frac{\partial}{\partial\xi} f_i, \frac{\partial}{\partial\xi} f_i >_{L_2(\Omega)} + \sum_{j=0}^{M} < \frac{\partial}{\partial\eta} g_j \frac{\partial}{\partial\eta} g_j >_{L_2(\Omega)}), \tag{2.42}$$

where the true energy is the limit $T(\Omega) = \lim_{N,M\to\infty} T_{NM}(\Omega)$. Now recall the definitions (2.37). We will add superscripts k, l when needed. When those strips are transformed to $[0,1]$ in their ξ coordinate (for S_{ij}) or η coordinate (for P_{ij}) we add a tilde. Under the assumption that the domain is characteristically convex we have that the strips $\tilde{S}_i$ and $\tilde{P}_j$ are bounded by functions as follows

$$\begin{aligned} \mathcal{L}_j(\eta) : [0,1] &\to \mathbb{R} \quad \text{describes boundary at the left of } \tilde{P}_j, \\ \mathcal{R}_j(\eta) : [0,1] &\to \mathbb{R} \quad \text{describes boundary at the right of } \tilde{P}_j, \\ \mathcal{T}_i(\xi) : [0,1] &:\to \mathbb{R} \quad \text{describes boundary at the top of } \tilde{S}_i, \\ \mathcal{B}_i(\xi) : [0,1] &:\to \mathbb{R} \quad \text{describes boundary at the bottom of } \tilde{S}_i. \end{aligned}$$

However, in the end we only need the total width and height at certain η and ξ coordinates. For these functions we write

$$\begin{aligned} \mathcal{W}_j(\eta) &= \mathcal{R}_j(\eta) - \mathcal{L}_j(\eta), \\ \mathcal{H}_i(\xi) &= \mathcal{T}_i(\xi) - \mathcal{B}_i(\xi). \end{aligned}$$

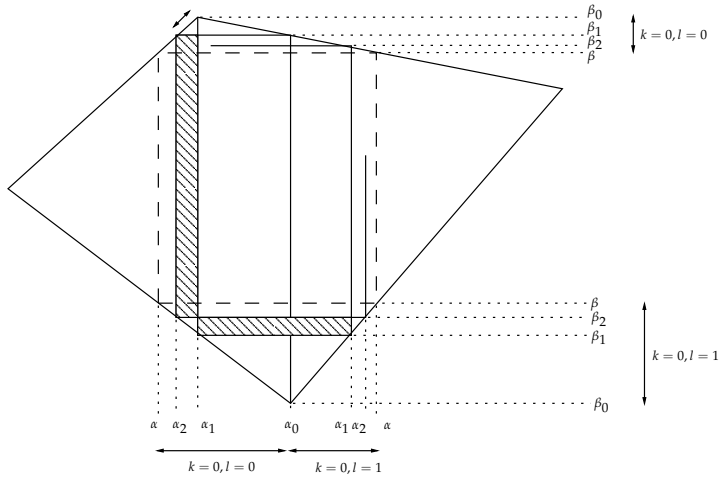


FIGURE 2.22: An example domain Ω. The endpoints of a fundamental interval, with (arbitrarily chosen) index $k = 0$, indicated by the thick double arrow, have been iterated a small number of times. In this case the fundamental interval induces two sequences in both the ξ and η direction, both with indices $l = 0$ and $l = 1$. For $i \to \infty$ these sequences approach the attractor (dashed line). The numbering of α and β is indicated, except that k, l are not written as superscript. Only a few α and β coordinates were plotted for clarity. The horizontal strip is $P_1^{0,1}$, the vertical strip $S_1^{0,0}$.

Now we have the tools to proceed with the calculation of the terms in (2.42),

$$< \frac{\partial}{\partial\xi} f_i(\xi), \frac{\partial}{\partial\xi} f_i(\xi) >_{L_2(\Omega)} = \frac{1}{\Delta\alpha_i} < f'(\xi), \mathcal{H}_i(\xi) f'(\xi) >_{L_2([0,1])} .$$

We rewrite the equation using partial integration and get

$$\begin{aligned} < f'(\xi), \mathcal{H}_i(\xi) f'(\xi) >_{L_2([0,1])} &= \left[(f^2(\xi))' \mathcal{H}_i(\xi) \right]_0^1 \\ &+ < f(\xi), -\mathcal{H}_i'(\xi) f'(\xi) - \mathcal{H}_i(\xi) f''(\xi) >_{L_2([0,1])} . \end{aligned}$$

The first term at the right hand side is zero by the corner point constraint. Perform a similar calculation for the term involving $\mathcal{G}$ to find

$$T_{NM}(\Omega) = < f(\xi), L_{1,N} f(\xi) >_{L_2([0,1])} + < g(\eta), L_{2,M} g(\eta) >_{L_2([0,1])}, \quad (2.43)$$

with

$$L_{1,N} = \sum_{i=0}^{N} -\frac{\mathcal{H}_i'(\xi)}{\Delta\alpha_i}\frac{\partial}{\partial\xi} - \frac{\mathcal{H}_i}{\Delta\alpha_i}\frac{\partial^2}{\partial\xi^2}, \tag{2.44}$$

$$L_{2,M} = \sum_{j=0}^{M} -\frac{\mathcal{W}_j'(\eta)}{\Delta\beta_j}\frac{\partial}{\partial\eta} - \frac{\mathcal{W}_j}{\Delta\beta_j}\frac{\partial^2}{\partial\xi^2}, \tag{2.45}$$

Finally, the minimisation of the energy becomes two decoupled minimisation problems for $\mathcal{F}$ and $\mathcal{G}$. The minimisation problems are very much alike, from now on we will concentrate on the $\mathcal{F}$ function. Let us first examine the minimisation of $< f, L_{1,N} f >_{L_2([0,1])}$. By the Courant-Fisher theorem the minimising function is the eigenfunction belonging to the smallest eigenvalue of the eigenvalue problem

$$L_{1,N} f = \lambda f. \tag{2.46}$$

For $N \to \infty$ we may consider the asymptotic behaviour of $L_{1,N}$. In this regime, close to the attractor, it holds that the function $\mathcal{H}_i$ gets increasingly flat with increasing i, since in the ξ coordinate the strips that are transformed to $[0,1]$ get smaller and smaller as we approach the attractor. Suppose that $\mathcal{H}$ can be well approximated by a linear function for large i and that the boundary makes an angle θ with the vertical at the attractor. Then it holds that $\mathcal{H}_i(x) \to \mathcal{H}_i(0) + \Delta\alpha_i \tan(\theta)^{-1} x$ and the terms in the summation for large i are well approximated by

$$-\tan(\theta)^{-1}\left[\frac{\partial}{\partial x} + x\frac{\partial^2}{\partial x^2}\right] - \frac{\mathcal{H}_i(0)}{\Delta\alpha_i}\frac{\partial^2}{\partial x^2}.$$

It is clear that the second term will dominate for large i and the eigenvalue problem may be written

$$-c_i\frac{\partial^2}{\partial x^2} f = \lambda f,$$

with $c_i = \frac{\mathcal{H}_i(0)}{\Delta\alpha_i}$. For every finite value of i we have the same solution,

$$f(x) = \sum_{k=0}^{\infty} A_k \sin(\sqrt{\frac{\lambda}{c_i}}x) + B_k \cos(\sqrt{\frac{\lambda}{c_i}}x).$$

Boundary conditions $f'(0) = f'(1) = 0$ dictate that all A_k must be zero. This also quantises the eigenvalues to yield

$$f(x) = \sum_{k=0}^{\infty} B_k \cos(\pi k x), \tag{2.47}$$

which is *independent of i*. By the normalisation constraint and $< \mathcal{F}, 1 >= 1$, the B_k are not all zero. For the function $\mathcal{G}$ we can find a similar expression

$$g(x) = \sum_{k=0}^{\infty} C_k \cos(\pi k x).$$

From the normalisation constraint $\|\Psi\| = 1$ we have that not all B_k and C_k are zero. Note also that c_i is diverging, and *the total energy (2.43) diverges* for $i \to \infty$. For any finite N the energy is meaningful, and minimisation yields a solution independent of N.

If we substitute the solutions obtained above in the expansion (2.40) and rewrite using (2.39) we find

$$\begin{aligned}
\Psi_{nm} &= \mathcal{F}_n + \mathcal{G}_m, \\
\mathcal{F}_n(\xi) &= \sum_{k=0}^{K} \sum_{l=0}^{J} \sum_{i=0}^{N(k,l)} B_n^{l,k} I_{\xi \in [\alpha_i^{l,k}, \alpha_{i+1}^{l,k}]} \cos(n\pi \frac{\xi - \alpha_i^{l,k}}{\Delta \alpha_i^{l,k}}), \\
\mathcal{G}_m(\eta) &= \sum_{k=0}^{K} \sum_{l=0}^{J} \sum_{j=0}^{M(k,l)} C_m^{l,k} I_{\eta \in [\beta_j^{l,k}, \beta j+1^{l,k}]} \cos(m\pi \frac{\eta - \beta_j^{l,k}}{\Delta \beta_j^{l,k}}).
\end{aligned}$$

In these expressions I denotes the indicator function. Also, the boundary condition $\Psi = \mathcal{F} + \mathcal{G} = 0$ implies that $B_n^{l,k} = -C_m^{l,k}$, if $m = n$. Therefore we need only the coefficient $B_n^{l,k}$. The minimisation problem is solved by $\Psi_{0,0}$, with the coefficients $B_0^{l,k}$ and $C_0^{l,k}$ determined by $\|\Psi\| = 1$ and $< \mathcal{F}, 1 >= 1$. In this case $\mathcal{F}$ and $\mathcal{G}$ are piecewise constants and the solution has a jump at the attractor, see also Section 2.9.

The next best solutions are less trivial to find. Small values of n, m will yield small energies, yet it is not a priori clear how the energies are ordered with respect n and m, this depends on α_i and β_j.

Even if we do not consider asymptotic behaviour we have information on the behaviour of (2.46) by noting that it is of Sturm-Liouville type

$$\frac{d}{dx}(p(x)\frac{df}{dx}) + [q(x) + \lambda \rho(x)]f = 0.$$

Write (2.46) as

$$P(x)f''(x) + Q(x)f'(x) - \lambda f = 0,$$

then we have Sturm-Liouville type if we set

$$\begin{aligned}
q(x) &= 0, \\
p(x) &= e^{\int_0^x \frac{Q(y)}{P(y)} dy}, \\
\rho(x) &= -e^{\int_0^x \frac{Q(y)}{P(y)} dy} / P(x).
\end{aligned}$$

We have boundary conditions $f'(0) = f'(1)$, which makes the problem of *periodic Sturm-Liouville* type. Standard Sturm-Liouville theory gives us that there are a countably infinite number of eigenvalues, the eigenvalues are real and have a smallest member, in absolute value. Therefore our minimisation problem attains a minimum. The eigenfunctions are orthogonal, and by the Sturm separation theorem they oscillate more rapidly with increasing eigenvalue.

Chapter 3

The Viscous Equation

The research in this chapter will be published as part of

A.N. Swart and D. Loghin and G.L.G. Sleijpen, 2007

3.1 Introduction

This paper deals with the numerical approximation of the Poincaré equation with viscosity. This equation arises as a model for internal waves, i.e., waves that propagate within a fluid volume instead of at its surface (Greenspan, 1968). Two notable occurrences of internal waves are waves in the oceans (LeBlond & Mysak, 1978) or in the liquid outer core of the earth (Rieutord *et al.*, 2001, 2002). In Section 3.2 we will present the physical model describing these waves.

Two challenging properties affecting the numerical solution of the internal wave problem are its ill-posedness and the occurrence of *wave attractors* (Maas & Lam, 1995). The ill-posedness makes the solutions highly sensitive to small perturbations of the boundary, or changes in the frequency of the wave. In a two dimensional model, wave attractors are closed orbits in the fluid domain to which the waves are attracted. In the inviscid case one observes a fractal structure, with arbitrarily small features and wave motion with diverging energy at the attractor.

The purpose of this paper is twofold. Firstly, there is currently a lack of numerical solution packages for internal wave phenomena that take into account the ill-posedness and occurrence of wave attractors. A numerical solution method and regularisation procedure for the inviscid equation was proposed in Chapter 2. This method is however limited to the Poincaré equation and not easily extended to more complex equations. The spectral collocation method of Rieutord *et al.* (2002) also deals with viscous internal waves, but is limited by the type of domain, spherical shells. Recently, Harlander & Maas (2006) proposed two novel numerical approaches to solve several variations on the Poincaré equation. Although the method may be used to tackle a wide range of equations, it requires information on the fundamental intervals and is only applicable to characteristically convex domains. The Finite Element Method proposed in this chapter can deal with any polygon. The method is also easily extended for solving related problems, e.g. the addition of nonlinearity.

The second main theme of this paper is the investigation of the effect of viscosity on the solutions. We investigate, using perturbation techniques, to what degree the viscosity regularises the solution. For small viscosity, the problem may be well-posed in principle, however numerically the solutions are not stable. Here we use the word 'stable' in the sense that small perturbations of the viscosity or the grid may dramatically alter the solution, or the set of solutions that is obtained numerically. Although we have no analytical solution for comparison, the stability is often clear from visual inspection. The quality of the solutions is often apparent from the observed smoothness and persistence of the solution upon grid refinement. Apart from visual inspection, we also find

criteria, in terms of the matrices and computed eigenvectors, that enable us to assess which numerically computed solutions can be considered reliable.

We start with a description of the physical model in Section 3.2, resulting the object of our study: the viscous Poincaré equation. The viscous Poincaré equation is an eigenvalue problem, where the complex eigenvalue contains the frequency and damping of the associated eigenvector. Section 3.3 reviews the basic properties of the inviscid equation, most notably the occurrence of wave attractors and the ill-posedness of the problem. We expect that these features also play a role for low viscosities. Section 3.4 describes the finite element model used to discretise the Poincaré equation. We then continue in Section 3.5 by describing the resulting eigenproblem. The properties of the eigenvalues are then studied in Section 3.6. We consider the evolution of eigenvalues depending on the viscosity and find a complicated scenario. If we denote the strength of the viscosity by γ and split eigenvalue in real and complex parts as $\omega = \omega^R + i\omega^I$ then we find that solutions only exist on certain curves in $(\omega^R, \omega^I, \gamma)$-space. In physical terms: only certain combinations of frequency (ω^R), damping $(\omega^I)$and viscosity (γ) are possible. More specifically, we find that the curves are ordered as tilted ellipses, with bifurcation points located at the extreme points. The curves may get close together for small γ which is an indication that numerically the problem is ill-posed. Furthermore we use perturbation techniques and express the sensitivity of the eigenvalues and eigenvectors in terms of solutions to the inviscid equation. Viscosity is found to have a beneficial effect on the stability of the solutions, and a criterion is formulated that expresses the stability of solutions with respect to perturbations in the system.

Equipped with the knowledge that solutions exist on certain curves only, we describe a method for continuation of the eigenvalues in Section 3.7. This method enables us to start a solution for large γ, and follow the solution to low values of γ. By virtue of the results from Section 3.6 we know how to find the eigenvalue at high γ that will converge to a target eigenvalue at zero viscosity.

Finally, sections 3.8.1 and 3.8.2 describe results of numerical experiments for the eigenvalues respecively eigenvectors. We use the stability criterion from Section 3.6 to isolate reliable solutions. Also we find modal solutions by considering the damping of the solutions. It is known that viscosity has a smoothing effect on the solutions. The small scale features in the neighbourhood of wave attractors are smoothed into viscous shear layers predicted to be of a thickness determined by the viscosity to power one third and one fourth (Dintrans *et al.*, 1999; Rieutord & Valdettaro, 1997). These shear layers are also present in our results (although we cannot confirm the precise scaling). Modal solutions are thought to be relatively invariant with respect to changes in the viscosity (see e.g. Bourgin, 1940), however, we find a mode without an inviscid counterpart.

3.2 Physical Model

In this section we describe a model for internal waves, which are waves traveling through a fluid volume. These waves have remarkable properties, compared to ordinary surface waves. For example, the dispersion relation states that the phase velocity vector is perpendicular to the group velocity vector, instead of being parallel. Also, the frequency of the wave determines the direction of the wave vector, as opposed to surface waves where the frequency sets the length. This has far-reaching consequences, as will be shown later. We start with considering the three dimensional situation, later we will simplify and work in two dimensions. See also section 1.2 for more details on the physical model.

Under the Boussinesq approximation, in a quasi-compressible, stably stratified, rotating fluid, the linearised equations of motion for the perturbation velocities may be written (Greenspan, 1968):

$$\rho^*(\mathbf{u}_t + \mathbf{f} \times \mathbf{u}) = -\nabla p - g\rho\mathbf{k} + \gamma\Delta\mathbf{u}, \tag{3.1}$$

$$\rho_t - \frac{N^2(z)\rho^*}{g}\mathbf{u}\cdot\mathbf{k} = 0, \tag{3.2}$$

$$\nabla\cdot\mathbf{u} = 0. \tag{3.3}$$

Here the given constant $N^2 \equiv -\frac{g}{\rho^*}\{\frac{d\rho_0}{dz} + \frac{\rho_0 g}{c_s^2}\}$ is the buoyancy frequency, which is dependent on the gravitation constant g and the static density stratification ρ_0, which is a function of the depth z only. Furthermore, the unknown ρ is the density and ρ^* is a given constant reference density. The pressure is denoted by p and $\mathbf{u} = (u, v, w)$ is the velocity vector. The Coriolis parameter f is half the angular velocity and $\mathbf{f} = (0, 0, f)$. The dynamic viscosity is denoted by γ. Let all quantities be defined in a closed domain $D \subset \mathbb{R}^3$ with boundary ∂D.

We are interested in solutions of the form

$$\mathbf{u}(x, y, z, t) = \Re(e^{i\omega t}\tilde{\mathbf{u}}(x, y, z)), \text{ with } \omega \in \{a + bi \in \mathbb{C} | b > 0\}, \tag{3.4}$$

i.e. we look for monochromatic waves. The appearance of the imaginary unit in the above expression is customary in fluid dynamics. The real part of ω now represents a frequency of oscillation, while the imaginary part represents a damping. It is also convenient that $\omega \in \mathbb{R}$ for $\gamma = 0$ (see also Section 3.3). The reason we are interested in the upper half of the complex plane only is that this is the regime where waves are damped. In the lower half-plane we find solutions that blow up in time, which is physically not acceptable. Note that equation (3.4) implies $\mathbf{u}_t = i\omega\mathbf{u}$ and we can insert equation (3.2) for ρ into

the momentum equations (3.1). We now eliminate the pressure p from (3.1) by taking the curl. Also, we introduce the operator $\mathcal{L} = i\omega - \gamma\Delta$, then

$$\begin{aligned} \mathcal{L}\nabla \times \mathbf{u} - \nabla \times \mathbf{u} \times \mathbf{f} = & \quad -\nabla \times [\frac{N^2(z)}{i\omega}(\mathbf{u} \cdot \mathbf{k})\mathbf{k}], \\ \nabla \cdot \mathbf{u} = & \quad 0. \end{aligned} \tag{3.5}$$

In order to fully specify the problem we apply the following customary boundary conditions,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{n} = 0 \quad & \text{if } \gamma = 0, \\ \mathbf{u} = \mathbf{0} \quad & \text{if } \gamma \neq 0, \end{aligned} \tag{3.6}$$

with $\mathbf{n}$ the outward normal to the boundary. The two-dimensional version of these equations, with boundary conditions (3.6) will be the object of our study.

3.2.1 The 2D Model

One simple reduction to a two-dimensional model is obtained by dropping the y-dependencies in the equations. This amounts to setting all derivatives with respect to y to zero in equations (3.5). Physically this represents a channel with infinite length along the y-axis. This view however is somewhat difficult to reconcile with rotation. Another model is obtained by interpreting (x, y, z) as cylindrical coordinates. Then the equations (3.5) are a model for internal wave motion in a rotationally symmetric geometry of large radius, as discussed in Section 4.A. This model has significance in physics, in modeling waves in atmosphere and ocean, or in the equatorial region of the outer liquid core of the earth (e.g. Rieutord *et al.*, 2002, Section 4). To proceed, we introduce the stream function given by $\mathbf{u} = \operatorname{curl} \Psi$, which amounts to $(u, w) = (\Psi_z, -\Psi_x)$ in component form. Note that, when defining $\mathbf{u}$ via the stream function, we automatically satisfy $\nabla \cdot \mathbf{u} = 0$. We obtain after some calculation the following expressions for (3.5),

$$\begin{aligned} \partial_z(\mathcal{L}v + f\Psi_z) &= 0, \\ \partial_x(\mathcal{L}v + f\Psi_z) &= 0, \\ \mathcal{L}\Delta\Psi - fv_z &= -\frac{N^2(z)}{i\omega}\Psi_{xx}. \end{aligned}$$

We list a few special cases of the above system,

$$\mathcal{L}[(\mathcal{L}\Delta\Psi + \frac{N^2(z)}{\omega}\Psi_{xx}] = -f^2\Psi_{zz}, \text{ in the general case,} \tag{3.7}$$

$$\omega^2\Delta\Psi - f^2\Psi_{zz} = -2i\omega\gamma\Delta^2\Psi, \text{ if } N^2(z) = 0 \text{ and } \gamma \ll 1, \tag{3.8}$$

$$\omega^2\Delta\Psi - N^2(z)\Psi_{xx} = -i\gamma\omega\Delta^2\Psi, \text{ if } f = 0, \tag{3.9}$$

In the second equation we have neglected a term proportional to γ^2, which is reasonable given that $\gamma \ll 1$ (the dynamic viscosity of water is about $1.004 \times 10^{-6}\, m^2/s$ at $20\,^\circ C$). We designed a finite element method which handles equations of the form

$$\omega^2 \Delta \Psi - a(z)\Psi_{xx} = -ib\omega\Delta^2\Psi, \tag{3.10}$$

which may be applied to equations (3.8) and (3.9). In equation (3.10) we suppose $b > 0$ is a small parameter, $a(z)$ a given function and (ω, Ψ) is an eigenpair, which is to be determined. The depth dependent $a(z)$ is of significance in oceanography, it allows for an arbitrary density profile. For ease of discussion, we will restrict ourselves to a constant frequency N^2, which is justified in several practical situations, e.g. the deep oceans. The method however, was implemented for arbitrary z-dependent profiles. Also, we non-dimensionalise the equations in order to obtain comparable quantities and reduce the number of parameters. In the density-stratified case we choose the typical scales $\Psi = NL^2\Psi'$, $\omega = N\omega'$ and $\gamma = NL^2\gamma'$ where L stands for a typical length scale. Similarly, the rotational case is non-dimensionalised by setting $\Psi = fL^2\Psi'$, $\omega = f\omega'$ and $\gamma = NL^2\gamma'/2$. In the oceans, the value of N varies strongly, ranging in orders of magnitude from 10^{-1} to 10^{-3}. Length scales may range from a few hundred meters to a few kilometers. Relevant values for γ' for oceanographic purposes are therefore roughly between $\gamma' = 10^{-5}$ and $\gamma' = 10^{-8}$. If we drop the primes this reduces the equations to

$$\omega^2 \Delta \Psi - \Psi_{xx} = -i\gamma\omega\Delta^2\Psi, \tag{3.11}$$

$$\omega^2 \Delta \Psi - \Psi_{zz} = -i\gamma\omega\Delta^2\Psi. \tag{3.12}$$

Since there is no essential difference between the two equations, we choose the first equation as the object of our study. From (3.6) we have that boundary conditions for $\gamma > 0$ read

$$\Psi = 0 \text{ and } \nabla\Psi \cdot \mathbf{n} = 0.$$

Another special case is the inviscid equation, $\gamma = 0$. The behaviour of this equation was extensively studied in Chapter 2 and will be briefly reviewed in the following section. This is of interest since the inviscid hyperbolic problem seems to dictate the behaviour of the elliptic viscous problem to a large extent.

3.3 Ray dynamics and ill-posedness

Many properties of the viscous equation seem to stem from the inviscid equation, also known as the Poincaré equation. In this section we briefly review some important facts on the Poincaré equation. Some of the material was also

presented in sections 1.4 and 2.3. We refer the reader to these sections for a more detailed exposition.We set $\gamma = 0$ in (3.7) and the governing equation becomes

$$\omega^2 \Delta \Psi - f^2 \Psi_{zz} - N^2(z) \Psi_{xx} = 0.$$

In Friedlander & Siegmann (1982*a*) it was shown for the inviscid linear case that $\omega \in \mathbb{R}$ and furthermore, that the equation is of hyperbolic type for

$$\omega^2 < \max(\sup N^2, f^2). \tag{3.13}$$

The inviscid equation is now more compactly written as

$$\Psi_{xx} - \lambda^2 \Psi_{zz} = 0, \tag{3.14}$$

with $\lambda^2 = \frac{\omega^2 - f^2}{N^2 - \omega^2}$. The boundary conditions in terms of the stream function read

$$\Psi = 0 \quad \text{at } \partial D,$$

The characteristics of equation 3.14 determine the structure of the solution. This approach has been discussed before (e.g. John, 1941; Maas & Lam, 1995) but we will briefly restate the main results here. The characteristic lines of (3.14) are given by

$$\begin{aligned} \xi(x,z) &= x - \lambda^{-1} z = \text{const.}, \\ \eta(x,z) &= x + \lambda^{-1} z = \text{const.}. \end{aligned}$$

Note that these lines have a *fixed* smallest angle of $\arctan(\lambda^{-1})$ with the vertical. Next, we show that the characteristics determine a map from the boundary to itself. This is most conveniently done in characteristic coordinates, where (3.14) reduces to $\Psi_{\xi\eta} = 0$. Now suppose that the domain is *characteristically convex*, i.e. each characteristic intersects the boundary ∂D in at most two points. In this case we may integrate to obtain the general solution in terms of arbitrary functions $\mathcal{F}$ and $\mathcal{G}$,

$$\Psi(\xi, \eta) = \mathcal{F}(\xi) + \mathcal{G}(\eta).$$

There are however constraints on $\mathcal{F}$ and $\mathcal{G}$, induced by the boundary condition $\Psi = 0$. Consider a point $(\xi_0, \eta_0) \in \partial D$ where $\mathcal{F}(\xi_0) = c_0$ and $\mathcal{G}(\eta_0) = -c_0$. Since $\mathcal{F}$ is a function of ξ only there is a point $(\xi_0, \eta_1) \in \partial D$ where still $\mathcal{F}(\xi_0) = c_1$. But this again implies that $\mathcal{G}(\eta_1) = -c_0$ and we conclude that characteristics transport a quantity, the difference between $\mathcal{F}$ and $\mathcal{G}$, from boundary point to boundary point.

A unique solution may be obtained by specifying $\mathcal{G}$ (or $\mathcal{F}$) on a set $M \subset \partial D$. This set is a finite union of disjunct parts and is called the *fundamental*

interval (Maas & Lam, 1995) or *generating set* (Lyashenko & Smiley, 1995). The Poincaré problem is easily over-determined or under-determined by using a domain $M' \subset M$ respectively $M' \supset M$. In the viscous case there is no concept of fundamental interval, no matter how small the viscosity. It is an interesting question whether some aspect of the fundamental interval has a counterpart in the viscous case. Our experiments (Section 3.8.2) show no evidence that the fundamental interval is retained in some form in the viscous setting.

The type of solution one obtains is highly dependent on λ and the shape of the boundary. The latter effect may be characterised by the rotation number of the map from the boundary to itself (defined in equation 2.7), which is induced by the characteristics. These statements are made more precise in Section 2.3. There are three cases to consider (see also John, 1941),

- Every characteristic retraces back to its starting point. This is called a *modal solution*. In this case, the rotation number is rational.
- Every characteristic approaches the same limit cycle. This is called a *wave attractor*; this solution displays fractal structure, features are reproduced in smaller scales towards the attractor. In this case, the rotation number is rational.
- All orbits are dense on the boundary. This is the *ergodic case*. In the class of continuous functions, only the trivial solution exists, and the rotation number is irrational. The fundamental interval consists of one or more isolated points.

See Fig. 3.1 for illustrations of these cases.

The *ill-posedness* of the problem stems from the fact that arbitrarily small perturbations in λ or the rotation number can change the type of solution. This statement is nicely illustrated by plotting the lengths of the attractors versus the parameter ω in Fig. 3.2. One must realise however, that this is a delicate calculation. Due to the finite arithmetic of the computations, and the large number of reflections, complicated attractors and resonances will not be found correctly. However, the overall structure will still be representative, since smaller scale attractors are more easily found. Note that attractors appear in parameter intervals, while modal solutions exist at points between the intervals. These intervals get increasingly small with larger attractor length. Modal solutions will probably not be found by our algorithm, since they appear at specific isolated points, while the algorithm tests a discrete number of frequencies.

It was this figure that was used in selecting the values of ω used in Figure 3.1. It is easy to calculate that for $\omega \in [1/\sqrt{5}, 1]$ every characteristic ends up in the upper right corner. This interval was not plotted. Physically this corner is a point of very high energy and we expect that solutions exist in this interval

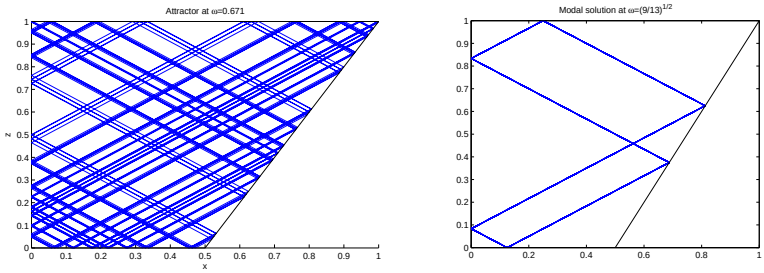


FIGURE 3.1: These panels show the different types of orbits made by the characteristics. On the left is an attractor, found at $\omega = 0.671$. Every characteristic will end up on the same attractor. The right panel is a modal solution, at $\omega = \sqrt{2/5}$. In this case every characteristic traces back to itself. Different initial points lead to different orbits, with the same number of intersections with the boundary (except for the corner points, which trace to other corner points).

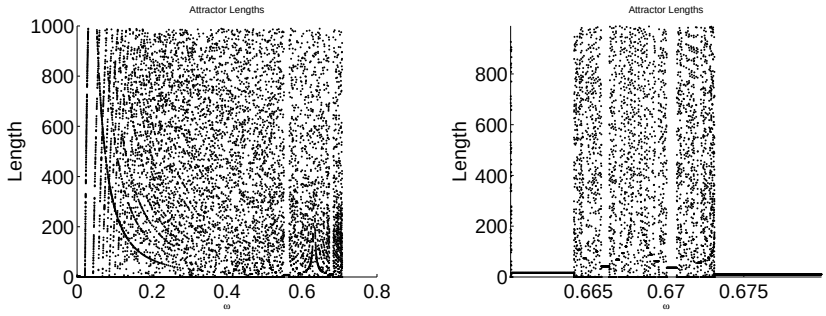


FIGURE 3.2: These panels show attractor lengths as a function of ω. The left panel shows the interval $[0, 1/\sqrt{2}]$. For $\omega \in [1/\sqrt{5}, 1]$ every characteristic ends up the top-right corner, while in $\omega \in [1/\sqrt{2}, 1/\sqrt{5}]$ the $(1,1)$-attractor lives. Outside these intervals the ordering is rather complicated. The rightmost figure shows the region $\omega \in [0.66, 0.68]$. We see that there is a self-similar structure in these plots.

only for large viscosity. Also not shown is the interval $\omega \in [1/\sqrt{2}, 1/\sqrt{5}]$ where the 'square' attractor lives (i.e, the attractor with one reflection at each wall segment).

3.4 The Finite Element Method

This section will describe a finite element discretisation for the viscous Poincaré equation (3.11). Because of the high order derivatives it is convenient to use a mixed method, this avoids using high order elements. The method described in this section is the standard finite element method, see e.g. Ciarlet (2002) for an overview of finite element methods, or Henderson & Aldridge (1992) for a finite element method applied to internal waves in a frustum. However the application of a Finite Element Method to the viscous Poincaré equation is a novel application. We start by introducing $-\Delta\Psi = \Phi$ in the viscous model (3.9), this allows us to formulate the following problem.

Given $\gamma \in \mathbb{R}$, find eigenfunctions $\Psi \in H_0^1(D)$, $\Phi \in H^1(D)$ and eigenvalues $\omega \in \mathbb{C}$, such that

$$\begin{aligned} \omega^2\Delta\Psi - \Psi_{xx} &= i\gamma\omega\Delta\Phi, \\ -\Delta\Psi &= \Phi. \end{aligned} \tag{3.15}$$

We define the L_2 inner product on D in the usual way as

$$(u,v)_{L_2(D)} = \int_D u \cdot v \, dxdz,$$

and introduce the following notation for integration over the boundary,

$$< f,g >_{\partial D} = \int_{\partial D} fg \, dl.$$

We proceed by multiplying with functions $\xi \in H_0^1(D)$ and $\eta \in H^1(D)$ and integrate over the domain D using Green's formula in the form

$$(f,\Delta g)_{L_2(D)} = -(\nabla f, \nabla g)_{L_2(D)} + < f, \nabla g \cdot \mathbf{n} >_{\partial D} .$$

Equations (3.15) can now be written

$$\begin{aligned} &-\omega^2(\nabla\Psi,\nabla\xi)_{L_2(D)} + (\Psi_x,\xi_x)_{L_2(D)} + i\gamma\omega(\nabla\Phi,\nabla\xi)_{L_2(D)} \\ &= i\gamma\omega < \xi, \nabla\Phi\cdot\mathbf{n} > - < \xi, \nabla\Psi\cdot\mathbf{n} >_{\partial D} + < \xi, \Psi_x n_x >_{\partial D}, \end{aligned}$$

$$(\nabla\Psi,\nabla\eta)_{L_2(D)} - (\Phi,\eta)_{L_2(D)} = < \eta, \nabla\Psi\cdot\mathbf{n})_{\partial D}.$$

The boundary terms in the first equation equal zero since $\xi \in H_0^1(D)$. Also, since the test function can be varied on the domain and at boundary independently, the second equation implies both $\nabla\Psi\cdot\mathbf{n} = 0$ at the boundary and $(\nabla\Psi,\nabla\gamma) = (\Phi,\gamma)$ in the domain. These considerations yield an equivalent form of the problem (3.9),

Given $\gamma \in \mathbb{R}$, *find eigenfunctions* $\Psi \in H_0^1(D)$, $\Phi \in H^1(D)$ *and eigenvalues* $\omega \in \mathbb{C}$, *such that*

$$\begin{aligned} -\omega^2(\nabla\Psi, \nabla\xi)_{L_2(D)} + (\Psi_x, \xi_x)_{L_2(D)} &= -i\gamma\omega(\nabla\Phi, \nabla\xi)_{L_2(D)}, \\ (\nabla\Psi, \nabla\eta)_{L_2(D)} &= (\Phi, \eta), \end{aligned}$$

for all $\xi \in H_0^1(D)$ *and* $\eta \in H^1(D)$.

Recall that the domain D is assumed to be a polygon. It may thus be conformably triangulated, yielding a triangulation $\mathcal{T}_h$ with mesh parameter h. The triangulation consists of a collection of non-overlapping triangles K_i that share their vertices. Denote the vertices in the triangulation by N_j, with $j = 1, \ldots, n$, and let the vertices be numbered such that for $j = 1, \ldots, m < n$ the nodes N_j are the interior nodes, while for $j = m+1, \ldots n$ the nodes are located at the boundary. We discretise the problem by looking for solutions in suitable subspaces $\mathcal{V} \subset H_0^1(D)$ and $\mathcal{W} \subset H^1(D)$. Introduce the continuous and piecewise linear nodal (with respect to $\mathcal{T}_k$) basis functions ξ_j satisfying $\xi_j(N_i) = \delta_{ij}$. Using this basis, we define the discretised function spaces as

$$\begin{aligned} \mathcal{W} &= \text{span}\{\xi_1, \ldots, \xi_n\}, \\ \mathcal{V} &= \text{span}\{\xi_1, \ldots, \xi_m\}. \end{aligned}$$

We will write $\Psi_h \in \mathcal{V}$ and $\Phi_h \in \mathcal{W}$ for the finite element approximations to Ψ and Φ. We may now write the finite element approximations as a linear combination of the basis functions,

$$\begin{aligned} \Psi_h(x,z) &= \sum_{j=1}^{m} v_j \xi_j(x,z) \in \mathcal{V}, \\ \Phi_h(x,z) &= \sum_{j=1}^{n} w_j \xi_j(x,z) \in \mathcal{W}, \end{aligned}$$

where the v_j and w_j are the coordinates with respect to the spaces $\mathcal{V}$ and $\mathcal{W}$. The discretised problem can now be stated as follows.

Find the $v_j, w_i \in \mathbb{R}$ *and* $\omega \in \mathbb{C}$ *such that*

$$\sum_{j=1}^{m} v_j[\omega^2(\nabla\xi_j, \nabla\xi_i) - (\frac{\partial\xi_j}{\partial x}, \frac{\partial\xi_i}{\partial x})] = i\gamma\omega \sum_{j=1}^{n} w_j(\nabla\xi_j, \nabla\xi_i), \qquad i \leq m,$$

$$\sum_{j=1}^{m} v_j(\nabla\xi_j, \nabla\xi_i) = \sum_{j=1}^{n} w_j(\xi_j, \xi_i),$$

Define the vectors $x(\gamma) = (v_1, \ldots, v_m)^T$ and $z(\gamma) = (w_1, \ldots, w_n)^T$. We have stressed the dependence on γ for future convenience, since we will be examining solutions as a function of γ. Let us for this reason also write $\omega(\gamma)$ instead

of ω from here on. The introduction of the vectors $x(\gamma)$ and $z(\gamma)$ allows us to write our problem as the system

$$\begin{pmatrix} -\omega^2(\gamma)L + H & i\gamma\omega(\gamma)\tilde{L} \\ \tilde{L}^T & -M \end{pmatrix} \begin{pmatrix} x(\gamma) \\ z(\gamma) \end{pmatrix} = 0, \tag{3.16}$$

where we introduced the matrices $L \in \mathbb{R}^{m\times m}$, $H \in \mathbb{R}^{m\times m}$, $\tilde{L} \in \mathbb{R}^{m\times n}$ and $M \in \mathbb{R}^{n\times n}$ defined by $L_{ij} = (\nabla\xi_j, \nabla\xi_i)$, $\tilde{L}_{ij} = (\nabla\xi_j, \nabla\xi_i)$, $H_{ij} = (\frac{\partial\xi_j}{\partial x}, \frac{\partial\xi_i}{\partial x})$ and $M_{ij} = (\xi_j, \xi_i)$ is the mass matrix. The tilde on the $\tilde{L}$ matrix indicates that it differs from L at entries corresponding to the boundary. For future convenience we also define $V \in \mathbb{R}^{m\times m}$, given by $V_{ij} = (\frac{\partial\xi_j}{\partial z}, \frac{\partial\xi_i}{\partial z})$. Note that we have the relation $L = H + V$.

3.5 The Eigenvalue Problem

The system (3.16) is equivalent to a quadratic eigenproblem for $\omega(\gamma)$ and the eigenvector $x(\gamma)$. This is easily seen from the equations for $z(\gamma)$ in (3.16) which imply $z(\gamma) = M^{-1}\tilde{L}^T x(\gamma)$. Note that the inverses of M, H, V and L exist since they are Gram matrices. The equations for $x(\gamma)$ now give

$$(-\omega^2(\gamma)L + H + i\gamma\omega(\gamma)\tilde{L}M^{-1}\tilde{L}^T)x(\gamma) = 0.$$

We simplify the notation by defining $K = \tilde{L}M^{-1}\tilde{L}^T$. This yields

$$(\omega^2(\gamma)L - H - i\gamma\omega(\gamma)K)x(\gamma) = 0. \tag{3.17}$$

One may cast the problem in the form of a generalised eigenvalue problem by using an auxiliary vector $y(\gamma) = \omega(\gamma)x(\gamma)$, this enables us to write (3.17) as

$$\omega(\gamma)\begin{pmatrix} i\gamma K & -L \\ I & 0 \end{pmatrix}\begin{pmatrix} x(\gamma) \\ y(\gamma) \end{pmatrix} = \begin{pmatrix} -H & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} x(\gamma) \\ y(\gamma) \end{pmatrix}.$$

Now multiplication with the inverse of the left-hand side matrix gives us an ordinary eigenvalue problem:

$$\begin{pmatrix} 0 & I \\ L^{-1}H & i\gamma L^{-1}K \end{pmatrix}\begin{pmatrix} x(\gamma) \\ y(\gamma) \end{pmatrix} = \omega(\gamma)\begin{pmatrix} x(\gamma) \\ y(\gamma) \end{pmatrix}. \tag{3.18}$$

We are interested in computing a number of eigenvalues and eigenvectors of this system close to a target eigenvalue. For large and sparse matrices iterative methods are often a good choice (Bai *et al.*, 2000; Golub & Van Loan, 1989). The sparsity allows for storing only the non-zero entries. Although the matrix L^{-1}

in the above problem generally not sparse, the resulting algorithm below only deals with sparse matrices.

It is known that eigenvalues at the extremes of the spectrum are approximated best (they are well approximated earlier in the iterative process), while eigenvalues in the interior need a larger number of iterations. In order to approximate eigenvalues in the interior, one may use a *shift-and-invert* technique, which transforms eigenvalues near a target σ to the outer parts of the spectrum. A shift-and-invert approach for calculating the eigenpairs corresponding to eigenvalues closest to σ amounts to calculating the largest eigenvalues of $(A - \sigma I)^{-1}$. Applied to our problem (3.18), we need the largest eigenvalues of the matrix

$$\begin{pmatrix} -\sigma I & I \\ L^{-1}H & i\gamma L^{-1}K - \sigma I. \end{pmatrix}^{-1} = -\begin{pmatrix} (Q^{-1}H + I)/\sigma & Q^{-1}L \\ Q^{-1}H & \sigma Q^{-1}L \end{pmatrix},$$

with $Q = \sigma^2 L - \sigma i\gamma K - H$. For use in an iterative solver, we repeatedly need multiplications of the above matrix with a vector. This product can be evaluated more efficiently than it might seem. This can be seen by writing out the multiplication with a vector $(p, q)^T$,

$$\begin{pmatrix} a \\ b \end{pmatrix} = -\begin{pmatrix} (Q^{-1}H + I)/\sigma & Q^{-1}L \\ Q^{-1}H & \sigma Q^{-1}L \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} = -\begin{pmatrix} (Q^{-1}p_2 + p)/\sigma \\ Q^{-1}p_2. \end{pmatrix},$$

with $p_2 = Hp + \sigma Lq$. An efficient scheme is now

1. Solve $Qb = -p_2$ for b
2. Set $a = (b - p)/\sigma$

We see that we need only one linear system solve and two matrix-vector multiplications. Also, note that all matrices involved are sparse.

Other than calculation of a number of eigenvalues and eigenvectors for fixed γ, we also want to calculate eigenvalues and eigenvectors for a range of values of γ, in order to study the evolution of eigenvalues and eigenvectors while varying the viscosity. To this end we use a continuation technique, which will be the topic of Section 3.7. Before we turn to eigenvalue continuation we first examine the properties of the eigenvalue problem (3.17) in the next section.

3.6 Analysis

This section will analyse the behaviour of the eigenvalues and eigenvectors of both the viscous and the inviscid problem. We will show that the inviscid equation is very ill-conditioned, however the viscous problem improves on the

conditioning of the eigenvalue problem. Also, properties of the eigenvalues of the viscous problem will be shown to be well-approximated by eigenvalues of the inviscid problem. This allows us to predict the behaviour of the viscous solution using properties of the matrices of the inviscid equation.

3.6.1 Properties and Notation

Whenever we consider the eigenvectors of the viscous equation we will use the vectors $x_i(\gamma) \in \mathbb{C}^m$ and $y_i(\gamma) \in \mathbb{C}^n$, eigenvectors of the inviscid equation are denoted by $x_i \in \mathbb{C}^n$. Correspondingly, we write $\omega_i(\gamma)$ and ω_i for the eigenvalues of the viscous respectively inviscid equation. We suppose that the eigenvectors have been enumerated in some fashion, and have added indices to indicate that we consider the i-th eigenvector (with $1 \leq i \leq n+m$). We assume the eigenvectors to be nomalised. For eigenvectors $x_i(\gamma)$, we define the following quantities,

$$\begin{aligned} l_{ij}(\gamma) &\equiv x_i^*(\gamma) L x_j(\gamma), \\ h_{ij}(\gamma) &\equiv x_i^*(\gamma) H x_j(\gamma), \\ v_{ij}(\gamma) &\equiv x_i^*(\gamma) V x_j(\gamma), \\ k_{ij}(\gamma) &\equiv x_i^*(\gamma) K x_j(\gamma). \end{aligned}$$

If $\gamma = 0$ we write x_i, l_{ij}, etc. Also, for all quantities defined above, we use the shorthand notation $l_{ii} \equiv l_i$, $h_{ii} \equiv h_i$, etc. The matrices L and H are symmetric positive definite and therefore

$$\begin{aligned} x_j^*(\gamma) H x_j(\gamma) \equiv h_j(\gamma) > 0, \\ x_j^*(\gamma) V x_j(\gamma) \equiv v_j(\gamma) > 0. \end{aligned}$$

From $L = H + V$ we have $l_j = h_j + v_j$.

3.6.2 The inviscid problem

We begin with a study of the inviscid equation, i.e. equation (3.17) with $\gamma = 0$,

$$(H + \nu_j^2 V) x_j = 0. \tag{3.19}$$

with $\nu_j^2 = \omega_j^2/(\omega_j^2 - 1)$. Premultiplication with x_j^* yields $h_j + \nu_j^2 v_j = 0$, which implies $\nu_j^2 = -v_j/h_j = (1 - l_j/h_j)$. From $l_j > h_j$ it now follows

$$\nu_j \in (-i, i) \rightarrow \omega \in (-1, 1). \tag{3.20}$$

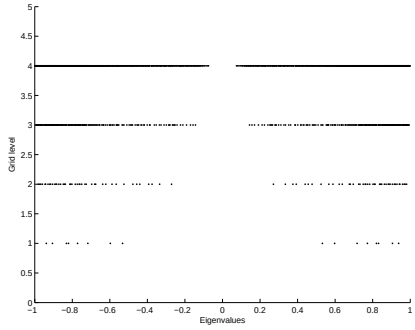


FIGURE 3.3: This figure depicts the eigenvalues (on the horizontal axis) as a function of the grid resolution. The eigenvalues are plotted at heights $1, \ldots, 4$ corresponding to dimensions matrix A in equation (3.20) of $8, 43, 197$ and 841. With increasing grid resolution the eigenvalues increasingly fill the interval $(-1, 1)$.

Furthermore, by adding to equation (3.19) its complex conjugate we see that we may assume $x \in \mathbb{R}^m$. Figure 3.3 confirms that $\omega \in (-1, 1)$, also we see that with increasing grid resolution the eigenvalues lie denser in the interval. Note that eigenvalues close to zero are only found at higher grid resolutions. This can be understood by considering that these eigenvalues correspond to very steep characteristics, the associated eigenfunctions cannot be represented well enough on a grid that lacks resolution.

We proceed with a perturbative error analysis, the discussion is standard (see e.g Golub & Van Loan, 1989), but we restate it here for completeness. The reader who wishes to skip it may proceed directly to the result, equations (3.25) and (3.27).

First, we rewrite the eigenvalue problem (3.19) as

$$Ax_j = \nu_j^2 x_j$$

by setting $A = V^{-1}H$, which is a symmetric matrix. Consider the perturbed system

$$(A + \epsilon E)x_j(\epsilon) = \nu_j^2(\epsilon)x_j(\epsilon). \tag{3.21}$$

The following discussion is standard (Golub & Van Loan, 1989, e.g), but we restate it here for completeness. The reader who wishes to skip it may proceed directly to the result, equations (3.25) and (3.27). We now assume that each eigenvalue is simple (has multiplicity one), this assumption is technical and avoids working with eigenprojections instead of eigenvectors in the following calculations. According to Kato (1995, Theorem 5.11), the $x_j(\epsilon)$ and $\nu_j^2(\epsilon)$ may

be written in the form of an asymptotic power series,

$$\begin{aligned} x_j(\epsilon) &= x_j + \epsilon z_j + \mathcal{O}(\epsilon^2), \\ \nu_j^2(\epsilon) &= \nu_j^2 + \epsilon\alpha_j + \mathcal{O}(\epsilon^2). \end{aligned}$$

The vector z_j may be written as a linear combination of eigenvectors of A,

$$z_j = \sum_{i=1}^{n} t_i x_i$$

and we find

$$x_i(\epsilon) = x_i + \epsilon \sum_{j=1}^{n} t_j x_j + \mathcal{O}(\epsilon^2).$$

Let us rescale in order to have a coefficient of one for x_i,

$$x_i(\epsilon) = x_i + \epsilon \sum_{j\neq i}^{n} \tilde{t}_j x_j + \mathcal{O}(\epsilon^2).$$

With these expansions, the eigenvalue equation becomes

$$(A+\epsilon E)(x_i + \epsilon \sum_{j\neq i}^{n} \tilde{t}_j x_j) = (\nu_i^2 + \epsilon\alpha_1)(x_i + \epsilon \sum_{j\neq i}^{n} \tilde{t}_j x_j) + \mathcal{O}(\epsilon^2),$$

which gives for the orders of magnitude in ϵ,

$$Ax_i = \nu_i^2 x_i, \tag{3.22}$$

$$(\alpha_1 - E)x_i = (A-\nu_i)\sum_{j\neq i}^{n} \tilde{t}_j x_j = \sum_{j\neq i}^{n} \tilde{t}_j(\nu_j^2 - \nu_i^2)x_j. \tag{3.23}$$

Let $\hat{x}_i$ stand for a normalised left eigenvector of A corresponding to eigenvalue ν_i^2. Introduce the *eigenvalue condition number* $s(\nu_i^2) = |\hat{x}_i^* x_i|$ and the quotient $\beta_{ij} = \hat{x}_i^* E x_j$, then premultiplication of (3.23) with $\hat{x}_i^*$ gives

$$\alpha_1 = \beta_{ii}/s(\nu_i^2)$$

and the error in the eigenvalues is

$$|\nu_i^2(\epsilon) - \nu_i^2| = \epsilon|\beta_{ii}|/s(\nu_i^2) + \mathcal{O}(\epsilon^2). \tag{3.24}$$

In our case, $A = A^T$ and therefore $x_i = \pm\hat{x}_i$ and $s(\nu^2) = 1$, thus

$$|\nu_i^2(\epsilon) - \nu_i^2| = \epsilon|\beta_{ii}| + \mathcal{O}(\epsilon^2). \tag{3.25}$$

For the eigenvectors we premultiply with $\hat{x}_j^*$, $i \neq j$, and obtain

$$|\tilde{t}_j| = \frac{|\beta_{ij}|}{|(\nu_j^2 - \nu_i^2)s(\nu_j^2)|},$$

and

$$\|x_i(\epsilon) - x_i\| \leq \epsilon \sum_{j \neq i}^{n} |\frac{|\beta_{ij}|}{|(\nu_j^2 - \nu_i^2)s(\nu_j^2)|} + \mathcal{O}(\epsilon^2). \tag{3.26}$$

Again, since $s(\nu_j^2) = 1$ this simplifies to

$$\|x_i(\epsilon) - x_i\| \leq \epsilon \sum_{j \neq i}^{n} |\frac{|\beta_{ij}|}{|\nu_j^2 - \nu_i^2|} + \mathcal{O}(\epsilon^2). \tag{3.27}$$

We conclude that for $\epsilon \rightarrow 0$ the eigenvalues converge much faster than the eigenvectors. This is evident from equation (3.27), the convergence depends on the *spectral gap*, $\min_{i,j} |\nu_j^2 - \nu_i^2|$, which may be very small in our case. For example, in the case of a square domain D one can show that the eigenvalues of L and V are approximately

$$\begin{aligned} \nu_k(H) &\approx k^2\pi^2, \\ \nu_l(V) &\approx l^2\pi^2. \end{aligned}$$

with the eigenvalues of A the quotient of these eigenvalues. Now, for example the mode $(k, l) = (20, 21)$ yields an eigenvalue of A that is very close to eigenvalues of A corresponding to $(k, l) = (1, 1)$. The finer the mesh, the more possibilities arise for a small spectral gap. Figure 3.4 shows that a disregarding a small spectral gap can indeed lead to an unsatisfactory solution, with many spurious oscillations. In Chapter 2 a remedy for this behaviour was proposed, by a regularisation technique, yielding smooth regularised solutions (see Section 2.9).

In contrast to tackling the inviscid equation, we here rely on the regularising properties of viscosity. The next section will study the viscous eigenvalue problem in more detail.

3.6.3 The viscous equation

This section will derive some properties of the viscous equation. We study the dependence of eigenvalues on γ and find that only specific combinations of $\omega_(\gamma)$ and γ are possible. Start with writing the viscous equation (3.17) in the form

$$((1 - \omega_j^2(\gamma))H - \omega_j^2(\gamma)V + i\gamma\omega_j(\gamma)K)x_j(\gamma) = 0. \tag{3.28}$$

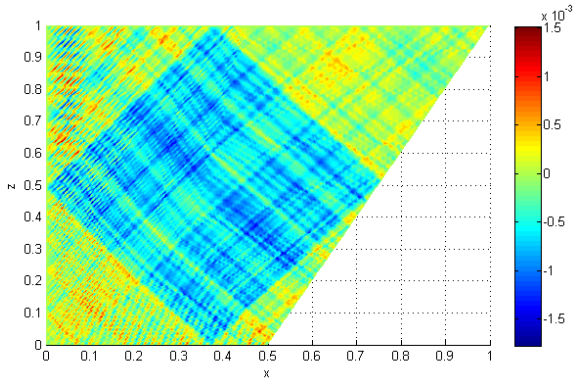


FIGURE 3.4: The solution of (3.19) shown in this figure, corresponds to an eigenvalue located in the regime where we expect a simple attractor. The result is not satisfactory, we see many high frequency oscillations.

Taking the complex conjugate (indicated by bar) we obtain,

$$((1-\bar{\omega}_j^2(\gamma))H - \bar{\omega}_j^2(\gamma)V - i\gamma\bar{\omega}_j(\gamma)K)\bar{x}_j(\gamma) = 0.$$

and we conclude that $-\bar{\omega}_j(\gamma)$ is an eigenvalue with eigenvector $\bar{x}_j(\gamma)$.

Further properties of the eigenvalue equation can be obtained by premultiplying (3.28) with $x_j^*(\gamma)$. We find

$$\omega(\gamma)^2 l_j(\gamma) - h_j(\gamma) - i\gamma k_j(\gamma)\omega_j(\gamma) = 0.$$

Split this equation in a real and imaginary part and write $\omega_j^R(\gamma)$ ($\omega_j^I(\gamma)$) for the real (imaginary) part of $\omega_j(\gamma)$. We then find

$$\begin{aligned}[(\omega_j^I)^2(\gamma) - (\omega_j^R)^2(\gamma)]l_j(\gamma) + h_j(\gamma) - \gamma k_j(\gamma)\omega_j^I(\gamma) &= 0,\\ \omega_j^R(\gamma)[\gamma k_j(\gamma) - 2\omega_j^I(\gamma)l_j(\gamma)] &= 0.\end{aligned}$$

We have three possible solutions,

$$\begin{aligned}\omega_j^R(\gamma) &= 0,\\ \omega_j^I(\gamma) &= \frac{1}{2l_j(\gamma)}(1 \pm \sqrt{1 - \frac{4l_j(\gamma)h_j(\gamma)}{\gamma^2 k_j^2(\gamma)}}),\end{aligned} \tag{3.29}$$

which is valid for $\gamma^2 \geq 4l_j(\gamma)h_j(\gamma)/k_j^2(\gamma)$. On the other hand, for $\gamma^2 \leq 2h_j(\gamma)l_j(\gamma)/k_j^2(\gamma)$,

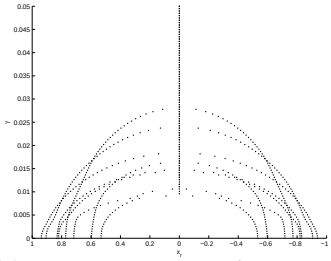
$$\begin{aligned} \omega_j^I(\gamma) &= \frac{k_j(\gamma)}{2l_j(\gamma)}\gamma, \\ \frac{l_j(\gamma)}{h_j(\gamma)}(\omega_j^R)^2(\gamma) + \frac{k_j^2(\gamma)}{4h_j(\gamma)l_j(\gamma)}\gamma^2 &= 1. \end{aligned} \tag{3.30}$$

Would the quantities $l_j(\gamma)$, $h_j(\gamma)$ and $k_j(\gamma)$ be independent of γ, then we would discern the equations for a straight line and an ellipse. These ellipses are what we find in the numerical experiments. For $\gamma = 0$ we find the eigenvalues of the inviscid equation, as required,

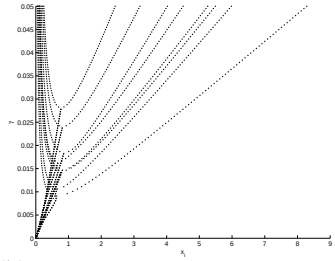
$$\begin{aligned} \omega_j^I &= 0, \\ (\omega_j^R)^2 &= \frac{h_j(0)}{l_j(0)} = \frac{h_j}{l_j}. \end{aligned}$$

Figure 3.5 shows that the behaviour sketched here indeed occurs, we clearly observe ellipses and straight lines (cf. equations (3.30)). For large γ, when the extremal point of the ellipse is reached, eigenvalues ω_j become purely imaginary, rapidly running to infinity or zero (cf. equations (3.29)). At the moment we only present the eigenvalues as discrete points, in Section 3.7 we will employ a continuation technique that enables us to track eigenvalues while changing the viscosity.

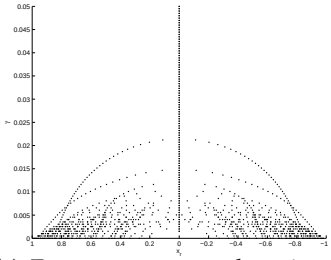
We are now interested in how well we can describe this behaviour using perturbation techniques. We consider γ to be a small parameter and expand the eigenvectors and eigenvalues of (3.17) with respect to γ. It is known that when eigenvalues do not cross, an expansion of the eigenvalues in integer powers of γ is permitted. Also, since the unperturbed problem deals with a symmetric matrix, while the perturbation is non-symmetric, the eigenvalues will extend, for $\gamma > 0$, into the complex plane (see Verhulst, 2005, Section 15.3). For the eigenvectors we also take an expansion in integer powers. However, it is a legitimate concern if such an expansion does justice to the structure of the solutions. One could employ other (fractional) powers of γ, motivated by the knowledge of $\mathcal{O}(\gamma^{1/3})$ and $\mathcal{O}(\gamma^{1/4})$ shear layers in the fluid (cf. Rieutord *et al.*, 2002). On the other hand the boundary layers are encoded in the numerical vector $x_i(\gamma)$ in a complicated way, and it is unclear how a proper scaling should be determined. We assume here that boundary layer and shear layer effects do not influence the order of the perturbation parameter. Motivated by the results obtained later (Section 3.8.1) we choose to employ a simple expan-



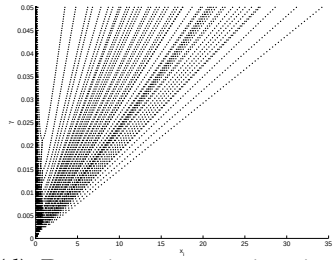
(a) Frequency versus damping, eight gridpoints.



(b) Damping versus viscosity, eight gridpoints.



(c) Frequency versus damping, 42 gridpoints.



(d) Damping versus viscosity, 42 gridpoints.

FIGURE 3.5: The above panels show all eigenvalues for a range of values of γ. Panels 3.5(a) and 3.5(c) plot ω^R versus γ, while panels 3.5(b) and 3.5(d) plot ω^I versus γ. The ellipses and straight lines predicted by (3.30) are clearly visible. Also we see solutions branching off at the top of the ellipses as expressed by (3.29).

sion in integer powers. The perturbation series are then given by

$$x_i(\gamma) = x_i + \gamma t_i + \gamma^2 s_i + \mathcal{O}(\gamma^3), \qquad (\gamma \to 0), \tag{3.31}$$

$$\omega_i(\gamma) = \omega_i + \gamma \alpha_i + \gamma^2 \beta_i + \mathcal{O}(\gamma^3), \qquad (\gamma \to 0). \tag{3.32}$$

The constant terms have been chosen in such a way that for $\gamma = 0$ we find solutions to the inviscid problem. Note that these eigenvectors x_i are real and the eigenvalues ω_i are real and contained in $(-1, 1)$. The values of α_i and β_i can be found by inserting the expansions into the eigenvalue equation (3.17) and equating the orders of magnitude in γ,

$$(\omega_i^2 L - H)x_i = 0, \tag{3.33}$$

$$(\omega_i^2 L - H)t_i + \omega_i(2\alpha_i L - iK)x_i = 0, \tag{3.34}$$

$$(\omega_i^2 L - H)s_i + \omega_i(2\alpha_i L - iK)t_i - [(\alpha_i^2 + 2\beta_i\omega_i)L - i\alpha_i K]x_i = 0. \tag{3.35}$$

The first equation states that x_i obeys the inviscid eigenvalue equation, as expected. From the second equation we like to obtain information on α_i and t_i. We premultiply the first equation with t_i^*, the second equation with x_i^T, and we insert the first equation. This yields

$$\alpha_i = \begin{cases} i\frac{k_i}{2l_i} & \text{if } \omega_i \neq 0, \\ i\frac{k_i}{l_i} \text{ or } 0 & \text{if } \omega_i = 0. \end{cases}$$

Using the third equation we find for $\omega_i \neq 0$,

$$\beta_i = -\frac{k^2}{8l^2 i\omega_i}.$$

When $\omega = 0$ the value of β_i can not be determined. The expression for the eigenvalues becomes

$$\omega_i(\gamma) = \omega_i + i\gamma\frac{k_i}{2l_i} - \gamma^2\frac{k_i^2}{8l_i^2\omega_i} + \mathcal{O}(\gamma^3). \tag{3.36}$$

The imaginary part is approximately linear in γ, for small γ, with the coefficient $k_i/2l_i$. We therefore recover the linear relationship, with $l_i \approx l_i(\gamma) \in \mathbb{R}$. The real part reads

$$\omega_i^R(\gamma) = \omega_i - \gamma^2\frac{k_i^2}{8l_i^2\omega_i} + \mathcal{O}(\gamma^3),$$

and thus, noting that $\omega_i^2 = h_i/l_i$, we obtain

$$(\omega_i^R)^2(\gamma) + \gamma^2\frac{k_i^2}{4l_i^2} = h_i/l_i + \mathcal{O}(\gamma^3).$$

Introduce the notation $a_i = \frac{l_i}{h_i}$ and $b_i = \frac{k_i^2}{4l_ih_i}$. Then equations in terms of ω read,

$$a_i(\omega_i^R)^2(\gamma) + b_i\gamma^2 = 1 + \mathcal{O}(\gamma^3), \tag{3.37}$$

$$\omega_i^I(\gamma) = \gamma\sqrt{b_i/a_i} + \mathcal{O}(\gamma^3). \tag{3.38}$$

The utility of the above equations lies in the fact that we may now calculate where to find a solution of the viscous equation, given as a starting point a solution of the inviscid equation. On the other hand, we can also calculate the converse, the inviscid eigenvalues and eigenvectors given a viscous solution. This can be done by calculating a_i from (3.37) with $\gamma = 0$, given $\omega_i^I(\gamma)$. Then $\omega_i = 1/a_i$.

3.6.4 Perturbation error analysis

In this section we will analyse the response of the solutions to the viscous eigenvalue problem (3.18) to perturbations in the matrix. We follow the same analysis as for the inviscid problem (Section 3.6.2), now using the matrix from (3.18), again adding a perturbation of the form ϵE. The eigenvectors of (3.18) contain vectors $x_i(\gamma)$ and $y_i(\gamma)$. We are interested in the part $x_i(\gamma)$, and rewrite the $y_i(\gamma)$ part in terms of $x_i(\gamma)$. We do the same for the left eigenvectors (distinguished from the right eigenvectors by a hat) and find

$$y_i(\gamma) = \omega_i(\gamma)x_i(\gamma), \tag{3.39}$$

$$\hat{y}_i^*(\gamma) = \bar{\omega_i}(\gamma)\hat{x}_i^*(\gamma)H^{-1}L. \tag{3.40}$$

From the eigenvalue equation (3.17) it follows that $\hat{x}_j(\gamma) = x_j(\gamma)$, since the matrices L, H and K are symmetric. The error in the eigenvalues is again proportional to $s^{-1}(\omega_i(\gamma))$, now given by

$$\begin{aligned} s(\omega_i(\gamma)) &= |(\hat{x},\hat{y})^T(x,y)| \\ &= |1+|\omega_i(\gamma)|^2\hat{x}_i^*(\gamma)H^{-1}L)x_i(\gamma)| \\ &= |1+|\omega_i(\gamma)|^2 l_i(\gamma)/h_i(\gamma)|. \end{aligned}$$

In principle we may now use equations (3.26) and (3.24) to estimate the perturbation error, which would lead to

$$|\omega_i(\gamma,\epsilon)-\omega_i(\gamma)| \leq \frac{\epsilon\beta_{ii}}{|1+|\omega_i(\gamma)|^2\frac{l_i(\gamma)}{h_i(\gamma)}|} + \mathcal{O}(\epsilon^2) \tag{3.41}$$

$$\|x_i(\gamma,\epsilon)-x_i(\gamma)\| \leq \sum_{j\neq i}^{n} \frac{\epsilon|\beta_{ij}|}{|(\omega_j(\gamma)-\omega_i(\gamma))(1+\frac{|\omega_i(\gamma)|^2 l_i(\gamma)}{h_i(\gamma)})|} + \mathcal{O}(\epsilon^2). \tag{3.42}$$

Here we introduced the notation $\omega(\gamma,\epsilon)$ for an eigenvalue of the perturbed viscous system. Correspondingly, we write $x_i(\gamma,\epsilon)$ for the eigenvectors. Already we see that viscosity has a beneficial effect. Firstly, the eigenvalue gap will be larger, since the eigenvalues spread out on the ellipses. Secondly, compared to (3.26) we have an extra factor $(1+|\omega_i(\gamma)|^2 l_i(\gamma)/h_i(\gamma)) > 1$ in the denominator.

However, for small values of γ the eigenvalue gap will become approximately equal to the eigenvalue gap for the inviscid problem. We calculate $|\omega_i(\gamma)|^2$ using the expansion (3.36), and find that several terms cancel to give us $|\omega_i(\gamma)|^2 = \omega_i^2 + \mathcal{O}(\gamma^3)$. For the eigenvectors we can also use the expansions in powers of γ. First notice that from (3.34) we find $t_i \in i\mathbb{R}^n$ and therefore

$t_i^* = -t_i^T$. We then have from (3.31) that $x_i^*(\gamma)x_i(\gamma) = x_i^T x_i + \mathcal{O}(\gamma^2)$. Therefore the expression $s(\omega_i(\gamma))$ simplifies to

$$s(\omega_i(\gamma)) = 2 + \mathcal{O}(\gamma^2).$$

Now we find for (3.24) and (3.26), with the above expression for $s(\omega_i(\gamma))$,

$$\begin{aligned} |\omega_i(\gamma,\epsilon) - \omega_i(\gamma)| &\leq \frac{\epsilon\beta_{ii}}{|2+\mathcal{O}(\gamma^2)|} + \mathcal{O}(\epsilon^2), \\ \|x_i(\gamma,\epsilon) - x_i(\gamma)\| &\leq \epsilon\sum_{j\neq i}^{n} \frac{|\beta_{ij}|}{|(\omega_j(\gamma)-\omega_i(\gamma))|(2+\mathcal{O}(\gamma^2)} + \mathcal{O}(\epsilon^2). \end{aligned}$$

For the denominator find

$$\begin{aligned} |(\omega_j(\gamma)-\omega_i(\gamma))s(\omega_j(\gamma))| &= |(\omega_j - \omega_i + i\gamma(\frac{k_j}{2l_j} - \frac{k_i}{2l_i})(2+\mathcal{O}(\gamma^2))| \\ &= 2\sqrt{(\omega_j-\omega_i)^2 - \gamma^2(\frac{k_j}{2l_j} - \frac{k_i}{2l_i})^2} + \mathcal{O}(\gamma^2). \end{aligned}$$

Assume that the eigenvalues and associated eigenvectors are ordered according to their real part. The error in the eigenvalues is dependent on the spectral gap (between the inviscid eigenvalues). With the eigenvalue ordering chosen above, the gap between ω_i and ω_j is small when i is close to j. In this case we can approximate the denominator by

$$2\gamma|\frac{k_j}{l_j} - \frac{k_i}{l_i}|.$$

Note that $k_i/l_i \equiv \gamma\sqrt{b_i}a_i$. This coefficient $\sqrt{b_i}a_i$ also represents the slope in the linear relation between γ and ω_i^I. Thus, only if neighbouring eigenvalues have a large difference in slope we can hope for a reasonable perturbation error. Since we calculate eigenvectors of the *v*iscous problem we approximate k_i by $k_i(\gamma)$ and l_i by $l_i(\gamma)$. An indicator for the perturbation error at small values of γ is now:

$$I_2 = |\frac{k_j(\gamma)}{l_j(\gamma)} - \frac{k_{j+1}(\gamma)}{l_{j+1}(\gamma)}| + |\frac{k_j(\gamma)}{l_j(\gamma)} - \frac{k_{j-1}(\gamma)}{l_{j-1}(\gamma)}|. \tag{3.43}$$

Note that we only compare values of $k_j(\gamma)/l_j(\gamma)$ with their direct neighbours at $j \pm 1$, since at $i = j \pm 1$ the value of $|\omega_i - \omega_j|$ is smallest.

3.7 Equilibrium Continuation

This section describes a continuation technique which will be used in the next section for the numerical experiments. It will allow us to follow solutions while varying the viscosity. We use a technique called *equilibrium continuation*, which is a technique for finding stationary points of a system of differential equations,

$$u'(t) = F(u, \alpha), \text{ with } u : \mathbb{R}^n \to \mathbb{R}^n, F : \mathbb{R}^{n+1} \to \mathbb{R}^n \text{ and } \alpha \in \mathbb{R}.$$

One is then interested in the equilibrium curve $F(u, \alpha) = 0$, where α is considered a parameter. In our case we want to compute the equilibrium curve corresponding to the system (3.17), with x and ω the unknowns and γ the parameter of the system. We split (3.17) and the variables in real and imaginary parts and obtain a function that maps $\mathbb{R}^{2m+3}$ to $\mathbb{R}^{2m}$, since the unknowns $x^I(\gamma), x^R(\gamma), \omega^I(\gamma), \omega^R(\gamma)$ and γ add up to $2m+3$ unknowns, while we have $2m$ equations. We therefore need two more constraints. We note that if (x, ω) solves (3.17), then the eigenpair $(re^{i\alpha}x, \omega)$ with $r, \alpha \in \mathbb{R}$ also solves the equation. We may resolve the degrees of freedom for r and α by adding the constraints

$$\|x\| = 1, \quad \Re(x)^T \Im(x) = 0. \tag{3.44}$$

Now $F(x, \gamma) = 0$ is equivalent with (3.18), supplemented with the above constraints, if we define

$$F : \mathbb{R}^{2m+3} \to \mathbb{R}^{2m+2}, \begin{pmatrix} x^R \\ x^I \\ \omega^R \\ \omega^I \\ \gamma \end{pmatrix} \to \begin{pmatrix} L_1(\gamma, \omega)x^R - L_2(\gamma, \omega)x^I \\ L_2(\gamma, \omega)x^R + L_1(\gamma, \omega)x^I \\ (x^I)^T x^R \\ (x^R)^T x^R + (x^I)^T x^I - 1 \end{pmatrix},$$

with

$$\begin{aligned} L_1 &= ((\omega^R)^2 - (\omega^I)^2)L - H + \omega^I \gamma K, \\ L_2 &= \omega^R[2\omega^I L - \gamma K]. \end{aligned}$$

Once an initial solution x^0 to $F(x, \gamma) = 0$ has been found, it is possible to calculate a new point x_1 on the equilibrium curve corresponding to the parameter value $x^0 + h$, where h is called the stepsize. We use the package Matcont (Dhooge *et al.*, 2004), a toolbox which runs under Matlab. Matcont implements Moore-Penrose continuation, a variation on Newton iteration. Furthermore the package detects critical points along the curve. Of special interest in our case are *branching points*, points where two equilibrium curves intersect.

The Moore-Penrose process however needs the Jacobian of F, which is given by

$$J = \begin{pmatrix} L_1 & -L_2 & D_1 x^R - D_2 x^I & -D_2 x^R - D_1 x^I & D_3 x^R - D_4 x^I \\ L_2 & L_1 & D_1 x^I + D_2 x^R & -D_2 x^I + D_1 x^R & D_3 x^I + D_4 x^R \\ (x^I)^T & (x^R)^T & 0 & 0 & 0 \\ 2(x^R)^T & 2(x^I)^T & 0 & 0 & 0 \end{pmatrix},$$

with

$$\begin{aligned} D_1 &= 2\omega^R L, \\ D_2 &= 2\omega^I L - \gamma K, \\ D_3 &= \omega^I K, \\ D_4 &= -\omega^R K. \end{aligned}$$

Now that we have the functions F and J we have a suitable system that we may continue using the Matcont software. The next section will provide results on numerical experiments using equilibrium continuation.

3.8 Numerical Experiments

3.8.1 Equilibrium Continuation

This section investigates the effect of viscosity on the eigenvalues, using the results of the previous sections. In particular we like to see to what extent viscosity has a regularising effect. We choose the trapezoid as our geometry, as shown in Section 3.3 it is a domain that allows for both wave attractors and modal solutions.

We started the investigation with the continuation of 5 eigenvalues. The initial parameters where chosen in such a way that at $\gamma = 0$ they would have real part close to 0.8 and imaginary part zero. This target value has been chosen in the interval where the $(1,1)$-attractor exists. An initial solution was found using the Matlab `eigs` function, which implements the Implicitly Restarted Arnoldi method (Lehoucq *et al.*, 1998). This method is suitable for a few eigenvalues of a sparse matrix near a target value. The result is presented in figure 3.6. Firstly we note that negative values of γ occur, this is physically unacceptable but convenient in the continuation process. The picture also confirms that the eigenvalues form closed curves in $(\omega^I, \omega^R, \gamma)$-space as anticipated in Section 3.6.3. Interestingly, at the points on the curve where γ and ω^I reach their maximum value (i.e. the extremes of the ellipses), we find branching points. At these points we intersect the solutions given by equation (3.29). We have

not switched branches at the branching points, the solutions with $\omega^R = 0$ are not interesting for our purposes. We are specifically interested in small values of γ, and fortunately we find that for small γ the match between the predicted positions of the ellipses and the calculated curves is quite good.

The drawback of the continuation technique is that it is very time consuming, for a large number of grid points it is no longer feasible to track several eigenvalues. However, we may still calculate a few eigenvalues at low viscosity using a sparse solver and use criterion (3.43) to asses the reliability of the solution. In the next section we will present solutions obtained using this technique.

3.8.2 Eigenvector Calculation

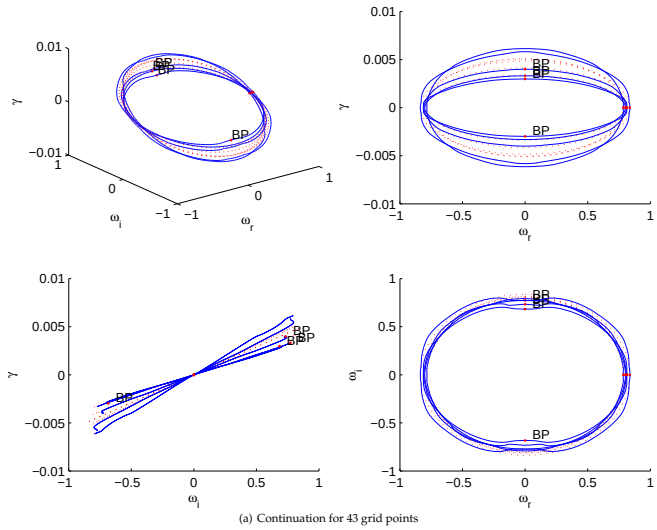
This section describes the results of the eigenvector calculation, using the criterion developed in Section 3.6.4. We describe two experiments that differ only in grid resolution. For each experiment we calculated 20 eigenvalues near an eigenvalue $\omega(\gamma)$ corresponding to $\omega = 0.8$, using equations (3.30). This eigenvalue is at a location where we expect wave attractors, at least from the inviscid theory (Section 3.3). The experiment is performed at three values of the viscosity, and using the criterion (3.43) the best and worst solutions are determined. Also, we compared results using (3.43) with results obtained by direct calculation of the right hand side of (3.42). Results were indistinguishable, and we continued using only (3.43) for assessment of the stability of the solutions.

The result of the calculation at a resolution of 3473 grid points is shown in Figure 3.7. The left column shows the best solutions, while the right column shows the worst solutions, both according to criterion (3.43). The viscosity is varied in the rows of the figure.

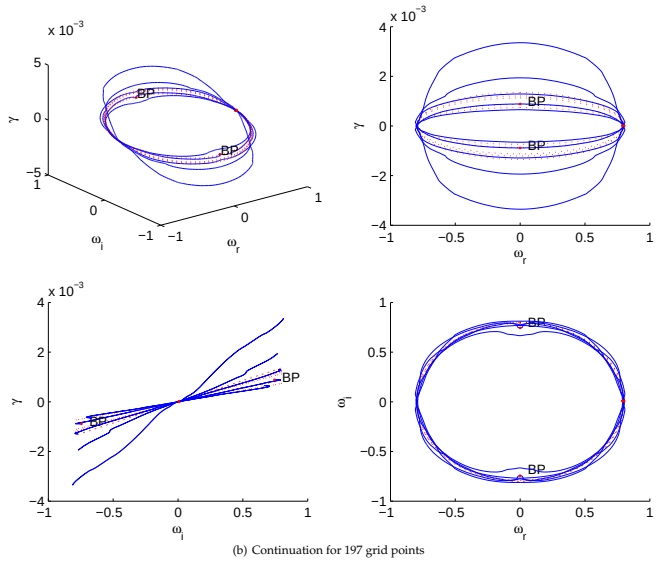
At the highest viscosity, top row, it seems that we successfully find a reliable solution. Of course we can only assess the quality by visual inspection, with the inviscid solutions obtained by ray tracing (e.g. Fig. 1.1) and regularised solutions from Chapter 2 (e.g. Fig. 2.16) as a reference. We find that we do not find any fractal structure, which is possibly smoothed out by the viscosity. However, the attractor is nicely visible.

The next row, at a viscosity of $\gamma = 10^{-6}$ shows narrower internal boundary layers around the attractor, which is to be expected. We also see that there are numerous high frequency oscillations in the figure. These could be genuine features of the solution, or spurious oscillations due to the ill-posedness. We have no means of deciding between the two, although the consistent structure seems to hint at this being a genuine feature of the solution. The difference in quality between the 'best' and 'worst' solutions is evident from the figures.

Finally, the bottom row gives results at $\gamma = 10^{-7}$. At this point the grid is

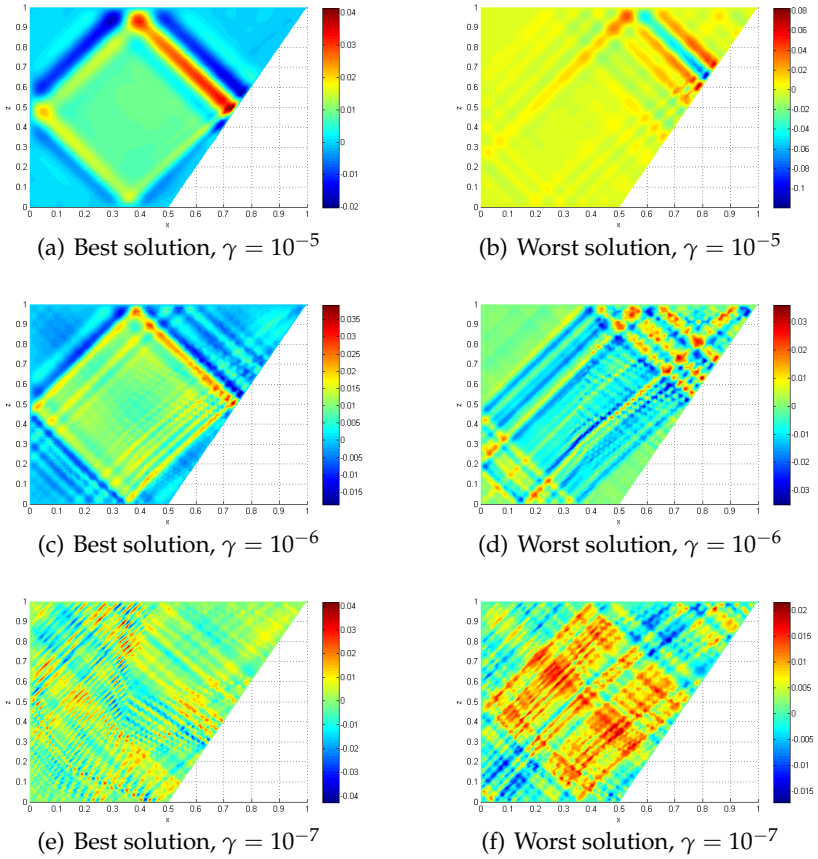


(a) Continuation for 43 grid points



(b) Continuation for 197 grid points

FIGURE 3.6: These panels show the curves on which the viscosity and real and imaginary parts of the eigenvalues lie. On the top-left is a three dimensional view, we see in blue the curves produced by the continuation software and in red the location of the curves as predicted by equations (3.30). The other two panels give orthogonal views on the (ω^I, γ)-plane, the (ω^R, γ)-plane and the (ω^R, ω^I)-plane. The label 'BP' indicates a branching point, where the solutions (3.29) come into existence (not plotted).



(a) Best solution, $\gamma = 10^{-5}$

(b) Worst solution, $\gamma = 10^{-5}$

(c) Best solution, $\gamma = 10^{-6}$

(d) Worst solution, $\gamma = 10^{-6}$

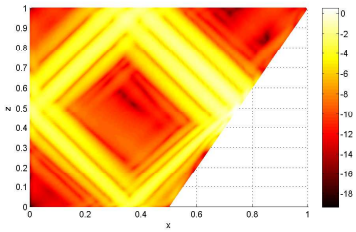
(e) Best solution, $\gamma = 10^{-7}$

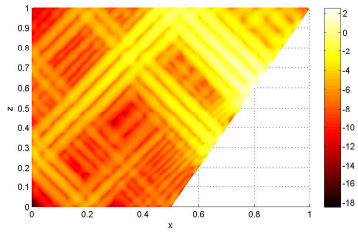
(f) Worst solution, $\gamma = 10^{-7}$

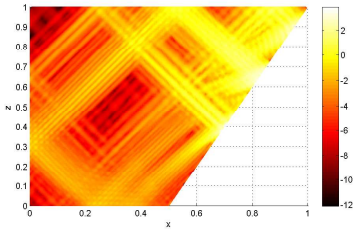
FIGURE 3.7: These panels show the best and worst solutions to the eigenvalue problem (3.18), determined using criterion (3.43). A resolution of 3473 grid points was used. In the rows we vary the parameter $\gamma \in \{10^{-5}, 10^{-6}, 10^{-7}\}$. Shown is a representation of the solutions in terms of the streamfunction.

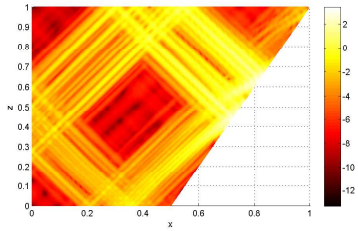
unable to resolve the required small scale boundary layer. Both panels in this row show irregular solutions with many suspicious high frequency patterns. At this point even the best solution found is not satisfactory.

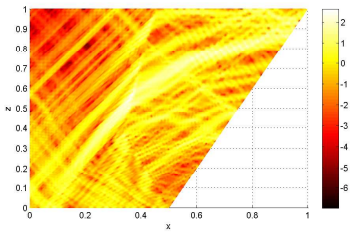
We continue with the second experiment, calculated at 14113 gridpoints. It was found that solutions were very similar to those shown in Figure 3.7. Therefore, in order to present an alternative view on the structure of the solutions, we display the solutions in terms of the logarithm of the energy instead


(a) Best solution, $\gamma = 10^{-5}$


(b) Worst solution, $\gamma = 10^{-5}$


(c) Best solution, $\gamma = 10^{-6}$


(d) Worst solution, $\gamma = 10^{-6}$


(e) Best solution, $\gamma = 10^{-7}$

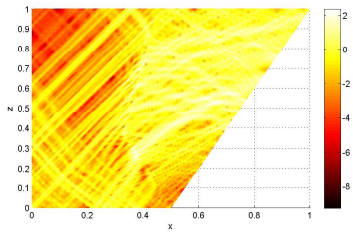

(f) Worst solution, $\gamma = 10^{-7}$

FIGURE 3.8: These panels show the best and worst solutions to the eigenvalue problem (3.18), determined using criterion (3.43). A resolution of 14113 grid points was used. In the rows we vary the parameter $\gamma \in \{10^{-5}, 10^{-6}, 10^{-7}\}$. Shown is the energy of the solution, in log-scale.

of the streamfunction. At locations where the streamfunction varies rapidly the energy will attain high values. The fact that results obtained at different grid resolutions were similar gives some confidence in the robustness of the criterion (3.43).

Firstly, we consider the first row of Figure 3.8. In terms of the streamfunction this solution was indistinguishable from the calculation in the first experiment. The representation in terms of the energy however gives us additional

insight. It is remarkable that the highest energy is concentrated at the attractor leg above the focusing point, while inviscid theory predicts an intensification of energy below the focusing point[1]. A possible explanation is that at viscosities this high the inviscid theory is no longer a suitable comparison. The focusing of rays could be inhibited by viscous effects. A second point of interest is that, at the lower leg below the focusing point, the shear layers are pinched towards the focusing point. To our knowledge, such behaviour was not reported before. We can only speculate, in a physical interpretation, that after wave motion was blocked at the focusing point, the wave needs to 'build up' again and regain energy.

In the second row, at a viscosity which is lower by a factor of ten, we find more detail and thinner shear layers. At this point we do find some intensification of energy at the focusing point, in both figures, although the energy above the focusing point is also rather high. The quality of the solutions can only be assessed visually, the figure to left does look better than the rightmost figure. The attractor stands out clearer than in the rightmost figure, although it is not possible to trace a closed orbit, which should be possible.

The bottom row represents a viscosity of $\gamma = 10^{-7}$, again lower by a factor of ten. Both the left and rightmost figure look unreliable at this point. The solution to the left now shows intensification of the energy below the reflection point, which is in line with inviscid theory. However the energy seems to fan out into the domain, and stop abruptly at about the middle of the domain. At this point we believe that the viscosity is too low, and solutions can no longer be resolved by the grid.

In fact we may quantify this statement a little more. It is known that shear layers of order $\gamma^{1/3}$ arise. We can then suppose that for small viscosities the shear layer width h_s behaves like $h_s \approx c\gamma^{1/3}$, with an unknown constant c. We may estimate this constant c from Fig. 3.8.2 by noting that at $\gamma = 10^{-5}$ we have $h_s \approx 0.05$, and thus $c \approx 0.05 \times 10^{5/3}$. Now, suppose that the number of grid points n and the mesh width h are related by $h = n^{-1/2}$ and that we need approximately 10 grid points along the shear layer. For $\gamma = 10^{-k}$ we then find for $k = 5, 6, 7$ grid sizes $n \approx 2000, 9000, 40.000$. This heuristic calculation shows that resolving the boundary layers at $\gamma = 10^{-5}$ is feasible, while doing this at $\gamma = 10^{-6}$ is doubtful and $\gamma = 10^{-7}$ is too low a viscosity. An attractive option (which have not tested) is to use local grid refinement along the attractor, whose location is known. This would lower the computational burden considerably.

Taking the above experiments into account we believe that the criterion (3.43) is suitable for selecting reliable solutions, as long as the grid has sufficient resolution. At viscosities below approximately $\gamma = 10^{-6}$ we would not

[1] Energy propagation is clockwise around the attractor.

recommend direct calculation, but argue in favour of an additional regularisation procedure.

The procedure of checking the values of (3.43) for several eigenvalues is not feasible for finding modal solutions. Those exist at one point at the spectrum (as opposed to wave attractors that live in an interval) and are almost certainly destroyed at perturbations of boundary. Therefore, they will be considered unstable by criterion (3.43). There is however an alternative for finding modal solutions: one may check the value of the imaginary part of the eigenvalue, representing the damping of the solutions. Physically, one expects that, since wave focusing is absent, viscosity will not damp the solutions much. We calculated, for $\gamma = 10^{-5}$ the value of $\omega(\gamma)$ corresponding to $\omega = \sqrt{2/5}$. This value corresponds to the skeleton shown in the rightmost panel of Figure 3.1. The calculation was performed at both 3473 and 14113 gridpoints, and results were identical, therefore we present the results of the highest resolution only. Looking for the first few lowest values of the damping we were able to find modal solutions (see Fig. 3.9), that seem to be represented rather well. The third solution is rather special, it is not predicted by inviscid theory. In the inviscid case one would expect n cells in the horizontal direction, accompanied by $2n$ cells in the vertical direction (see e.g. Maas & Lam, 1995). Figure 3.9(c) violates this principle and seems to be a solution without an inviscid counterpart.

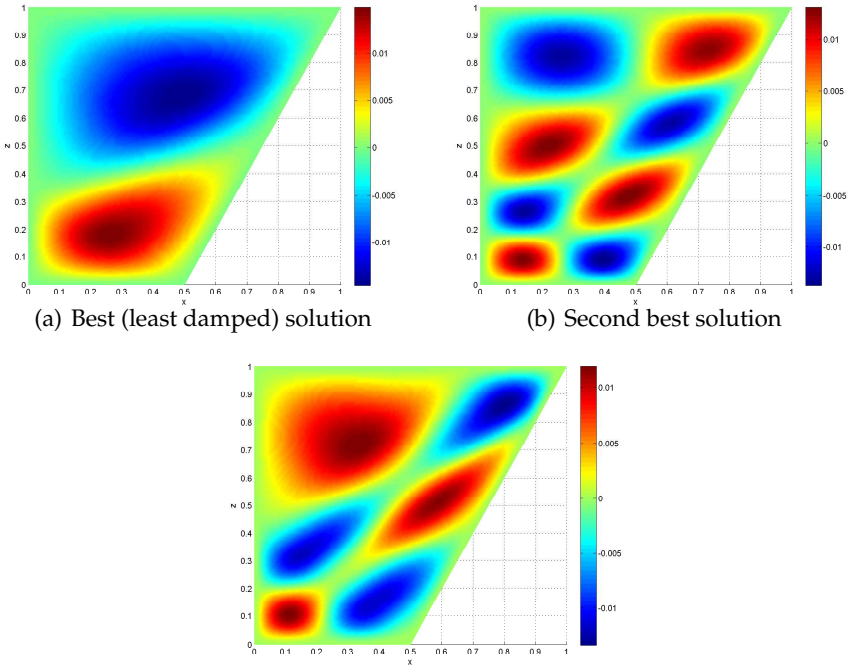
However, it must be noted that at lower viscosities, the modal solutions could often no longer be found. Instead complex attractor solutions were found. Again it seems that at low viscosity an additional regularisation procedure is in order. An energy minimising procedure based on the approach taken in Chapter 2 might be a good candidate.

3.9 Conclusion

In this Chapter we have put forth a methodology for numerical calculation of solution to the Poincaré equation in the presence of a viscous term. To this end an we used a finite element method, whose implementation was already available. Special attention has been paid to the ill-posedness of the problem, and the regularising effect of viscosity.

After derivation of the main problem and a discussion of the ill-posednedness in Sections 3.2 and 3.3, a Finite Element Method was established in Section 3.4. The discretisation by means of the Finite Element Method yields a discrete eigenvalue problem. Section 3.5 states the eigenvalue problem and indicates how matrix-vector products for use in a shift-and-invert strategy can be efficiently computed.

In Section 3.6 we analyse the properties of the eigenvalue equation and the sensitivity of the solutions to perturbations in the problem. The inviscid equa-



(a) Best (least damped) solution

(b) Second best solution

(c) Third best solution

FIGURE 3.9: This figure shows the three solutions with least damping, from a total of 20 solutions around $\omega = \sqrt{2/5}$, at 14113 gridpoints.

tion is discussed in Section 3.6.2. It is established that the discrete spectrum is contained in the interval $(-1, 1)$, thus it is within the bounds of the spectrum of the continuous problem. Furthermore, using a perturbative analysis we show that the eigenvectors are extremely sensitive to perturbations in the problem, and the error may become arbitrarily large for grids with high resolution. This behaviour was to be expected, from our previous discussion of the ill-posedness of the problem in Section 3.3 (see also Section 2.7), and the results presented for a rectangular geometry in Section 2.9.1, see in particular Fig. 2.10.

We continue in Section 3.6.3 with considering the eigenvalues of the viscous problem. We find that the eigenvalues are associated with tilted ellipses $(\omega^R(\gamma), \omega^I(\gamma), \gamma)$, parametrised by γ. The extremal points of the ellipses represent the maximum value of the viscosity γ that can be reached for the associated solution. Also, by considering γ to be a small parameter, and writing

eigenvalues and eigenvectors as a series expansion in γ we obtain descriptions for the ellipses in terms of solutions to the inviscid equation. It is interesting to note that in the (ω^I, γ) plane, the projection of the ellipses are straight lines. Thus we have a linear relation between damping and viscosity. More specific, the angle of the ellipses with the ω^I-axis is proportional to the eigenvalue of the inviscid equation.

Section 3.6.4 discusses the sensitivity of the viscous equation with respect to perturbations in the problem. We have seen that the situation is not very promising in the inviscid case. However, for the viscous equation we find that we may obtain reliable solutions at modest grid resolution. When we aim for an eigenvalue $\omega_i(\gamma)$ with eigenvector $x_i(\gamma)$, then the perturbation error depends on the distance between eigenvalues. We have seen in Section (3.6.3) that this spectral gap may become large, since eigenvalues lie on ellipsoidal orbits. Again using a perturbative approach we arrive at a simple criterion, equation 3.43 for assessing the quality of the solutions.

Finally, Sections 3.8.1 and 3.8.2 describe the results of numerical experiments. As long as the grid has sufficient resolution, the criterion (3.43) works well. However, for lower values of the viscosity we do not obtain reliable solutions. The spectral gap in this case is very small and we believe an additional regularisation procedure should be used.

We conclude with calculations of a few modal solutions, using the principle that modal solutions should be among the least damped eigenvectors. Surprisingly, we do not only find the expected modes, but also a mode that is a true viscous mode, without an inviscid counterpart.

Chapter 4

Internal Wave Focusing over Sloping Terrain

The research in this chapter will be published as part of

Experimental observation of strong mixing due to internal wave focusing over sloping terrain.
A.N. Swart and Leo Maas and Uwe Harlander and Astrid Manders, 2007

4.1 Introduction

The mixing of density in rotating stratified fluids is of great interest in the fields of geophysical fluid dynamics and oceanography. Particularly interesting is the fact that in the deep ocean there is an amount of mixing which is as yet unaccounted for. Internal wave breaking has been put forth as a possible source for this unexplained mixing (e.g. Ledwell *et al.*, 2000; Staquet & Sommeria, 2002). We have set up a laboratory experiment, aiming to collect evidence for the hypothesis that internal wave reflections can lead to strong localized mixing. Throughout the paper we will emphasize the relevance for oceanic processes.

In our experimental setup a stratified fluid is contained in a rotating annulus, with sloping inner wall. This setup mimics the topography of an ocean slope. In order to excite waves, we applied a weak modulation of the rotation rate, on top of the solid body rotation. This modulation rate was chosen at a frequency where we theoretically expect a *wave attractor*. Section 4.2 gives the theoretical background on wave attractors in rotating stratified fluids. Evidence for the manifestation of the attractor in our experiment is presented in Section 4.5. We argue that the modulation is responsible for a localized eruption of the boundary layer, which directly forces internal waves. We discuss a theoretical model which predicts these eruptions in Section 4.3 and we also present direct measurements of the boundary layer eruption.

Despite the fact that we do not disturb the fluid in any mechanical way, other than the weak modulation, we observe very strong mixing, much stronger than seen in similar experiments that lack the periodic modulation. We offer the hypothesis that the focusing action of internal waves at the sloping boundary is responsible for this mixing. Firstly, we present density profiles in Section 4.5.1 that clearly show that the density gradient has been heavily modified. Furthermore, we experimentally establish the existence of a wave attractor (Section 4.5.2) which could theoretically further enhance the mixing. The wave attractor was observed in a vertical plane. Section 4.5.3 studies the wave field in a horizontal plane. The results obtained are consistent with the data from the vertical plane, and furthermore show a remarkable spatial asymmetry.

While Section 4.5 focuses on direct measurements of mixing and internal waves, Section 4.6 presents measurements of waves and flows that are *caused* by the mixing. In Section 4.6.2, we discuss an observed mean flow that occurs over the sloping inner wall. Such a mean flow is indicative of internal wave breaking as discussed in Section 4.6.1. Additionally, in Section 4.6.3, we establish the occurrence of Rossby waves, which are also representative of local mixing.

Summarizing, we describe an experiment aimed at the validation of the following scenario

1. Boundary layer eruptions may generate gravito-inertial waves at certain critical angles

2. These waves locally mix the density profile upon reflection at the boundary

3. Wave attractors may form, leading to increased mixing efficiency

4. As a result of the mixing a measurable mean flow ensues and, being unstable, Rossby waves develop.

We believe that analysis of the collected data strongly supports our claims. Mixing, wave attractors, mean flow generation and Rossby waves have been observed and the observations fit well into our theoretical framework.

4.2 Wave attractors

In a channel-like geometry, characterized by cross-channel topographic variations, internal wave beams are focused upon reflection at a slope. Ultimately these reflections may lead to so-called wave attractors (Maas & Lam, 1995), see a review in Maas (2005) and also Sections 1.5 and 2.3 in this thesis. An oblique, free internal wavy boundary layer develops around the attractor location. Earlier exploratory laboratory studies have identified the existence and structure of such wave attractors both in uniformly-stratified fluids (Maas *et al.*, 1997; Lam & Maas, 2007), as well as in rotating homogeneous fluids (Maas, 2001; Manders & Maas, 2003, 2004). In the former case these were excited parametrically by vertical oscillation of the tank, in contrast to homogeneous rotating fluids, excited by modulation of the container's rotation speed. Here we explore their occurrence in a fluid that is both uniformly-stratified as well as rotating, again employing modulation of the rotation speed as an axisymmetric forcing mechanism.

4.2.1 Ray dynamics

This Section will recollect some properties of inertio-gravity waves and wave attractors. Also we define the parameters used in the experiments. We consider a quasi-compressible, stably stratified, rotating fluid. Furthermore we assume monochromatic wave motion with frequency ω. Since the domain D is rotationally symmetric around the z-axis (see fig. 4.2) we use a cylindrical (r, θ, z)-coordinate system. In appendix 4.A it is shown that in this case, in the

interior of the fluid domain, in a linearised description, the governing equations (where we neglect viscosity ν and diffusivity κ) may be written

$$\begin{array}{lcll} P_{rr} - \lambda^2 P_{zz} & = & F(r, z, P_r) & \text{in } D, \\ \mathbf{n} \cdot (P_z, 0, -r^{-1}(rP)_r) & = & 0 & \text{at } \partial D. \end{array} \tag{4.1}$$

where $P = r^{-1}\Psi$ and $(\Psi_z, -\Psi_r) = (ru, rw)$. In these equations P stands for pressure, Ψ is the stream function and u,w are velocities in the radial and vertical directions. All terms that contain only first derivatives were collected in a function F. The parameter λ^2 is defined by

$$\lambda^2 = \frac{\omega^2 - f^2}{N^2 - \omega^2}, \tag{4.2}$$

where the constants f and N are the Coriolis parameter (measuring the importance of rotation) and stability frequency (measuring the importance of stratification). In order for λ to be a real number it is needed that

$$\omega^2 \in [\min(N^2, f^2), \max(N^2, f^2)].$$

In our experiment, frequencies were set such that ω is in this frequency range. Note that the equation (4.1) is then of hyperbolic type.

Due to viscous effects at the boundary ∂D of the fluid domain D, the fluid must co-rotate with the reference frame, which implies $\mathbf{u} = 0$. However, we take into account the existence of a boundary layer where the fluid adjusts from the frictional influence of the boundary. Because of the presence of these boundary layers, we have to good approximation a frictionless boundary between the interior and the boundary layer region. This justifies posing boundary conditions $\mathbf{u} \cdot \mathbf{n} = 0$, with $\mathbf{n}$ the unit outward normal to the boundary.

The dispersion relation can be obtained by assuming wave motion $P(r, z) = \Re[\hat{P} e^{i(kr+mz)}]$. For simplicity we also set the right-hand side of (4.1) to zero, which is justified for $r \gg 1$. This simplification is not crucial (see e.g. Lyashenko & Smiley, 1995; Harlander & Maas, 2006), but it greatly simplifies the following discussion. This gives us $\lambda^2 m^2 = k^2$ or

$$\omega = \sqrt{\frac{N^2 k^2 + f^2 m^2}{k^2 + m^2}}. \tag{4.3}$$

We write the wave vector in polar coordinates, $\mathbf{k}_w = (k, m) = K(\sin(\alpha), \cos(\alpha))$, where the angle α is measured with respect to the vertical. This yields

$$\omega^2 = f^2 \cos^2(\alpha) + N^2 \sin^2(\alpha), \text{ and } \lambda = \pm \tan(\alpha).$$

The group velocity and phase velocity are now given by

$$\mathbf{c}_g = \nabla_{\mathbf{k_w}}\omega = \frac{km(N^2-f^2)}{K^3\sqrt{N^2k^2+f^2m^2}}(m,-k), \tag{4.4}$$

$$\mathbf{c}_p = \frac{\omega}{K^2}\mathbf{k}_w = \frac{\sqrt{N^2k^2+f^2m^2}}{K^3}(k,m). \tag{4.5}$$

Taking their dot product reveals that phase and energy propagation are mutually perpendicular. As shown in Maas & Lam (1995) we can construct a 'skeleton' of the solution by considering the characteristics of (4.1), namely

$$\begin{aligned} \xi(r,z) &= r-\lambda^{-1}z, \\ \eta(r,z) &= r+\lambda^{-1}z. \end{aligned} \tag{4.6}$$

The levelsets of these characteristics define two families of straight lines. The angle that the characteristic lines make with respect to the vertical is equal to the angle that the group velocity makes with the vertical, and is at right angles with the phase velocity vector. We conclude that wave motion is in the direction of the characteristic lines. The precise behaviour along a characteristic is determined by the function F, we are however only concerned with the singularity at the wave attractor. For numerical results in the case of non-zero F, see Harlander & Maas (2006). The location of the wave attractor can be found by a ray-tracing procedure (see also Maas & Lam, 1995).

If one starts at a boundary point one may trace straight lines from boundary point to boundary point. The fixed angle of the characteristics with the vertical, and the shape of the container, causes them to converge to a limit cycle termed the *wave attractor*. Figure 4.1 shows the web of characteristics for frequencies used in the experiment, listed in table 4.1. These frequencies were carefully selected from numerical ray tracing experiments in order to obtain a wave attractor with six intersections at the boundary (left panel) a complicated attractor (middle panel) and an attractor with four intersections with the boundary (right panel).

The inviscid solutions predict an infinite energy density along the attractor, since spatial scales will become arbitrarily small and the kinetic energy density becomes singular at the attractor(see Section 2.8. Viscosity (ν), diffusivity (κ) and non-linearity (not present in our model) will regularize the solution, creating high amplitude shear layers along the attractor. Note that at $\omega = 0.30$, corresponding to the middle panel of figure 4.1, based on the characteristics, we would expect a complicated attractor. However, its complexity makes it unrealizable in practice, since viscous and nonlinear processes will inhibit its formation. In fact, this frequency was selected to act as a 'counter-experiment' and we expect less efficient mixing here since an attractor will not be formed in practice.

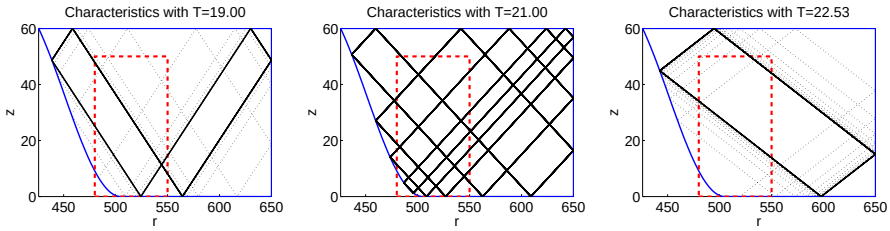


FIGURE 4.1: These panels show the skeleton of characteristics (4.6) for the frequencies in table 4.1. The dotted line shows the approach of the characteristics to the attractor, in solid black. The attractor in the middle panel is expected to be sufficiently complicated, it will be broken down by viscous and non-linear effects. The rectangle in dashed red indicates the field of observation.

T	ω	α	Description
19	0.33	60.40	$(1,2)$-attractor
21	0.30	68.09	'Anti-frequency', no simple attractor.
22.53	0.32	73.88	$(1,1)$-attractor

TABLE 4.1: Settings used in the experiments. Shown are the periods ($T = 2\pi/\omega$), frequencies (ω) and angles (α) of the characteristics with the horizontal.

4.2.2 Physical manifestation of the attractor

We will now turn to the question of how to detect the existence of the wave attractor in the measurements. Firstly, we expect at the attractor location a high amplitude of the waves. Secondly, around the attractor location we expect a 180 degree phase difference (anti-phase). For both phenomena there is numerical evidence, see also Sections 2.9 and 3.8.2 in this thesis, and it has also been observed in the laboratory (e.g. Maas *et al.*, 1997). This shearing motion implies that the absolute value of the curl field will be a good indicator of the presence of a wave attractor. A second indicator of the presence of an attractor is the generation of higher harmonics. Although we have neglected the usual nonlinear term $\mathbf{u} \cdot \nabla \mathbf{u}$ in our model equations, this term can become significant when $\nabla \mathbf{u}$ is large. We expect this to be the case at the attractor. It is easy to see why this term could induce second harmonics by considering a one dimensional toy model, where the non-linear term is modeled by $u(r)u_r(r)$ and we assume simple wave motion $u(r) = \sin(kr)$. Substitution and application of the double angle formula for the sine function yields $-k\sin(2k)$ for the non-linear term. A more elaborate model is needed to show that we also generate waves with a double frequency 2ω (e.g. Greenspan, 1968, page 215). See also Lam & Maas

(2007) for a discussion of generation of higher harmonics at attractor locations.

Note that the attractor is a *local* source of second harmonics. Waves are generated along the attractor. When these waves propagate we expect the angle of the characteristic lines to be modified. From (4.2) we see that λ^2 will grow, and from the dispersion relation $\lambda^2 = \tan^2(\alpha)$ we find that α will become larger. The angle α indicates the angle of the characteristic lines with the horizontal, thus we anticipate that characteristics and group velocity will lie steeper in the plane.

As a final remark we point out that mixing modifies the density stratification, and therefore also the characteristic lines. We expect that wave beams may follow more curved paths at locations where there is strong mixing. If the influence is strong enough, then the mixing process may alter the characteristics in such a way that the attractor is destroyed. However, if the mixing is modest, then the attractor may survive for some time. The local curvature of the characteristics can be removed and carried over into a local curvature of the domain by a change of variables. In terms of frequency, attractors live in a frequency interval. Also, when we can describe perturbation of the domain using some parameter τ, attractors also live in a τ-interval (this statement is made more precise in Sections 1.4–1.6 and Section 2.3). When choosing the frequency for use in our experiment, we made sure to pick one close to the center of the frequency interval where the attractor lives. Therefore we have good reason to believe that the attractor is relatively stable with respect to density variations.

4.3 Wave generation from boundary layers

We will now present the mechanism responsible for the forcing of internal waves in our experiments: eruptions of the boundary layer at critical angle. We start with an introduction of the physics, after which we present our mathematical model of the boundary layer.

4.3.1 Introduction

The concepts of mixing and wave attractors that are explored in our laboratory experiment are of relevance for understanding the physics of the oceans. In the oceans, internal tides are generated over continental slopes due to tidal motion in the cross-slope direction, that lifts and depresses isopycnals[1](Baines, 1982).

[1]Curves of equal density are commonly called *isopycnals*, while curves of equal topography height are termed *isobaths*. Curves of equal pressure are called *isobars*.

While the 'tidal' flow corresponds to a purely axisymmetric flow in our experiments, apparently lacking such a cross-isobath component, near the bottom it does lead to cross-slope oriented oscillatory currents due to frictional effects at the bottom (Thorade, 1928; Prandle, 1982; Maas & Haren, 1987). Close to the bottom, the azimuthal flow is retarded by frictional forces. Consequently, the flow is deflected by Coriolis forces, since the fluid has a nonzero velocity with respect to the co-rotating frame of reference. This deflection yields a fluid transport in the boundary layer into the radial direction. This boundary layer is known as the *Ekman layer*.

Far above the bottom the tidal flow is approximately linear, and can be seen as a superposition of two counter rotating circular currents of tidal frequency and of equal amplitude. The cyclonic one (co-rotating with the platform) however, decays to its zero value at the boundary over a much shorter distance than the anticyclonic one. The adaptation of these flows to the boundary gives rise to two nested Ekman layers. Over distances below the thickness of the anticyclonic Ekman layer, the flow is dominated by the cyclonic component and thus is able to alternatingly lift and depress isopycnals, directly generating internal tides.

There are two important effects. Firstly, it leads to an alternatingly stabilizing versus destabilizing influence on the density profile, and secondly it leads to internal wave generation in corner regions.

Potentially, the former process may directly lead to mixing as when, in the down-welling phase, the Ekman boundary layer becomes statically unstable (Trowbridge & Lentz, 1991). This might, in fact, lead to a complete shut-down of the Ekman layer, as shown for a steady current by MacCready & Rhines (1991) and Garrett *et al.* (1993), so that water on top of a mixed boundary layer flows uninhibited until the mixed water relaxes and the frictional influence of the bottom is felt afresh. Tidal modulation of the flow, as in our experiment, may alter this process in a periodic way (Ramsden, 1995; Middleton & Ramsden, 1996). Field indications of such instabilities have come from the intermittent appearance of spikes in the registration of some current meters, located in the deeper regions over the slope of the Faroe-Shetland Channel (Hosegood & van Haren, 2006) during down-welling favorable flows.

However, even when the Ekman layer does not turn unstable, Ekman transport over the slope may still lead to mixing in an indirect way. One theory asserts that the transport over the slope may be unable to blend in with the Ekman transport over the flat bottom or along the free surface, which results in mass ejection or suction into the bulk of the fluid from the connecting corner regions (similar to that observed in inertial wave experiments, see Beardsley (1970)). The periodic nature of this process generates internal waves which then propagate away and contribute to mixing elsewhere. Given the vast di-

mensions of the open ocean into which internal waves propagate, their fate, and the mixing these might produce, are generally unknown. Our model also predicts periodic pumping and suction at corner regions, but by a different mechanism. We will show in the next section that the boundary layer can erupt into the interior of the fluid at certain critical angles of the boundary. Note that the aforementioned hypothesis of non-matching flows does not occur in our model.

We believe that mixing by means of an unstable Ekman layer is not the primary source of mixing in our experiment. Firstly because of the scale of the mixing, and secondly because the mixing continues to be in effect, even after the modulation has been shut down.

Unfortunately, direct measurement of shut-down or instability of the Ekman layer is beyond our grasp in the current experimental setup. We will therefore focus on boundary layer eruptions at critical slopes of the boundary, which we have observed experimentally and which are the direct generators of internal wave beams.

There is however another effect that might result in mixing after the modulation was stopped. This effect is a result of the *baroclinic torque*. The baroclinic torque arises when the curl of the right-hand side of the momentum equations is taken

$$\nabla \times (\rho^{-1} \nabla p) = \rho^{-2} (\nabla \rho \times \nabla p).$$

When the fluid is baroclinic, $\nabla \rho \times \nabla p \neq 0$, the baroclinic torque generates vorticity, which in turn leads to mixing. However, in the interior of the fluid, a barotropic equilibrium is possible. This is an equilibrium where the isobars and isopycnals (see also Footnote 1) are parallel and slightly parabolically shaped[2]. On the other hand, when the isopycnals are parabolically shaped, their gradient can not be parallel to the boundary and a secondary flow may be driven. This flow might lead to mixing (by mechanisms analogous to the ones described above) and opens a possibility for persistent mixing after shut-down of the modulation. However, in previous experiments at the LEGI turntable, the stratification is typically kept in place for months, while in our experiment the stratification was destroyed in a matter of hours. This indicates that mixing via the baroclinic torque route is negligible compared to the mixing by means of reflecting internal wave beams.

The next section will describe a mathematical model that aims to explain the mechanism behind the generation of waves from critical points of the Ekman layer.

[2]Note that this requires a background density ρ_0 dependent on r and z, while our model uses a background density dependent on ρ_0 only.

4.3.2 Boundary Layer Analysis

In this section we consider the boundary layer behaviour of a viscous rotating fluid, at a sloping boundary. We wish to test the hypothesis of a boundary layer 'eruption', a singularity in the equation describing the boundary layer thickness. This eruption would lead to periodic ejection and suction of fluid from the boundary layer which consequently acts as a generator for internal wave beams (see also Hollerbach & Kerswell, 1995). This concept may be of relevance in the field of oceanography, where the possibility of internal wave generation from boundary layer eruptions has not yet been explored, although indications of them are found in numerical model results on internal wave generation Drijfhout & Maas (2007).

How does the boundary layer depth h (measured perpendicular to the boundary in a radial plane) behave as a function of the slope? In appendix 4.A we calculate this quantity for both non-oscillating (h_0) and oscillating (h_1) flows,

$$h_0(\phi) = \mathcal{O}\left(\frac{\nu^2}{\tilde{f}^2+\sigma\tilde{N}^2}\right)^{1/4}, \text{ as } \nu \downarrow 0 \text{ and } \kappa \downarrow 0, \tag{4.7}$$

$$h_1(\phi) = \mathcal{O}\left(\frac{\kappa}{\omega}\left(\frac{\sigma-\tilde{\lambda}^2}{\frac{\tilde{N}^2}{\tilde{N}^2-2\omega^2}-\tilde{\lambda}^2}\right)\right)^{1/2}, \text{ as } \nu \downarrow 0 \text{ and } \kappa \downarrow 0, \tag{4.8}$$

where $\sigma = \nu/\kappa$ is the Prandtl number, $\tilde{\lambda}^2 = (\omega^2 - \tilde{f}^2)/(\tilde{N}^2 - 2\omega^2)$ with $\tilde{N} = N\sin(\phi)$ and $\tilde{f} = f\cos(\phi)$. The angle ϕ is the angle between the horizontal and the tangent to the boundary. The classical Ekman depth, $h_0 = \mathcal{O}\sqrt{\nu/f}$ for stationary flow over a flat bottom ($\phi = 0$) is reproduced (e.g. Batchelor, 1967, page 198). For sloping walls our results differ from previous results. For example, for $\omega = 0$, Garrett *et al.* (1993) found that the boundary layer over a flat bottom is a factor $(1+\sigma N^2\sin^2(\phi)/f^2)^{-1/4}$ thinner as compared to a sloping bottom, while we find a factor $(1+\sigma(N^2-f^2)\sin^2(\phi)/f^2)^{-1/4}$.

The difference is insignificant for $N \gg f$, but for oscillating flow the results differ more. We can rewrite (4.8) as

$$h_1(\phi) = \mathcal{O}\left(\sqrt{\frac{\kappa}{\omega}\left(\frac{f^2+\sigma\tilde{N}^2-[\omega^2(1+2\sigma)+f^2\sin^2(\phi)]}{f^2+\tilde{N}^2-[\omega^2+f^2\sin^2(\phi)]}\right)}\right), \text{ as } \nu \downarrow 0 \text{ and } \kappa \downarrow 0.$$

The terms in square brackets are not present in the results cited above. Our equations have an important new property, the depth becomes singular for

$$\omega^2 = \tilde{N}^2 + \tilde{f}^2.$$

For the specific case treated in our laboratory experiment, we have that $N = 2f = 4\Omega$ and $\omega = 2\pi/T$ and the boundary layer may erupt when

$$\sin^2(\phi) = \frac{1}{3}\left[\frac{\Omega - T}{T}\right]^2.$$

For example, for $T = 22.53$ and $\Omega = 50$, the setting in our experiment when aiming for the $(1,1)$-attractor, we predict critical layers at $\phi_1 \approx 0.25\pi$, $\phi_2 = 2\pi - \phi_1$, $\phi_3 \approx 0.75\pi$ and $\phi_4 = 2\pi - \phi_3$. The angle ϕ_1 is present in our geometry, it might be argued that the other angles are present in the sense of being 'degenerately buried' in the non-smooth corner points. In Section 4.5.2 we will discuss further the presence of boundary layer eruptions in the experimental data. We clearly observe the eruption from the bottom-left corner (Fig. 4.4), where a slope of angle ϕ_1 is found, while occurrence of eruptions from the other corners is unclear.

4.4 Experimental setup

The LEGI turntable consists of a rotating basin, having a diameter of thirteen meters. On the inner rim of the platform a sloping topography is mounted, the outer rim is vertical. The resulting axi-symmetric annulus was filled with a stratified fluid. See figure 4.2 for a schematic representation of the turntable and its dimensions.

In the experiment we utilize two mechanisms that give rise to internal wave motion, density stratification and rotation. In addition to a solid body rotation, the angular velocity was modulated. This modulation sets the frequency of the generated internal waves, according to the mechanism as explained in Section 4.3.2. The amplitude of the modulation was chosen to be small compared to the amplitude of the solid rotation. The resulting angular frequency is

$$\Omega(t) = \Omega_0[1 + \epsilon \sin(\omega t)], \quad \text{for} \quad 0 < \epsilon \ll 1.$$

Initially, the fluid was homogeneously stratified, with $N = 2f = 4\Omega_0$. The amplitude of modulation has been set to have an azimuthal velocity of $v_{rim} = 6$ cm/s at the outer rim, relative to the solid body rotation. A convenient side effect of having a stratification is that it leads to homogeneously distributed particles, that were added to the fluid for visualisation purposes. The particles have slightly differing density, which makes them float on different isopycnal layers where gravity and buoyancy are in equilibrium. The stratification is stable, therefore due to waves the particles will execute vertical motion around their rest level.

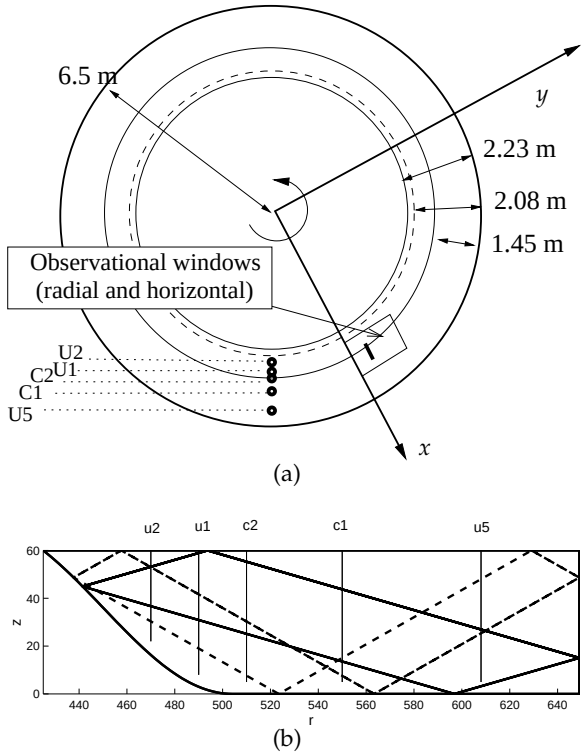


FIGURE 4.2: The left panel shows a schematic picture of the LEGI turntable. The outer rim is located at at 6.5 m. The inner circle represents the intersection of the water surface with the slope, while the middle solid circle is the start of the slope. The dashed line is the intersection of the attractor (see Fig. 4.1 rightmost panel) with the slope. Also indicated are the radial positions of the probes and the observational window. The right panel shows a radial cross-section of the basin. The attractors with $T = 22.53$ and $T = 19$ are indicated (solid and dashed lines) as well as the probe positions (thin vertical lines, the probes are by default at the lowest position but were also used to obtain $\rho(z)$ profiles. A forcing period of T=21 s does not yield a simple pattern and was therefore not drawn.

The velocities of the particles are measured using the *Particle Image Velocimetry* (PIV) technique. To this end the particles are illuminated using a laser sheet. The laser is mounted on the turntable in such a way as to produce either radial or horizontal sheets. The particles in the laser sheet are then photographed using a camera that is also mounted on the turntable. The camera sampling

rate has been set to obtain eight velocity fields per wave period. Each velocity field is created from four pictures, which are rapidly taken in a very small time interval (in the order of milliseconds), this is called the *burst cycle*. A time series of bursts is further processed by PIV software to create the vector fields of the velocities from the images. These velocities are our basic fields, which need further processing.

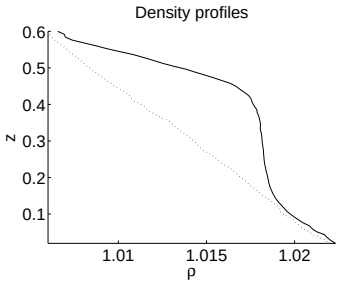
In addition to the camera field measurements, also velocity and conductivity probes were injected into the fluid. These probes provide density measurements at fixed radial positions and could be moved vertically to get density profiles (see Figure 4.2). Two of them were conductivity probes (C1 and C2), the other three were ultrasonic probes (U1,U2,U5).

4.5 Results

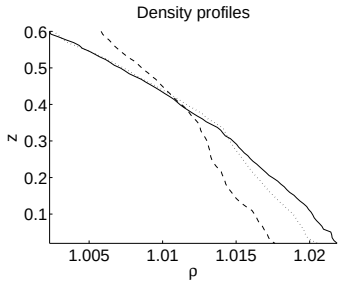
In the experiments very strong mixing was observed. This might seem remarkable at first sight, since we do not disturb the fluid in any direct way, merely the axi-symmetric turntable is rotated and modulated. No traditional source of mixing seems to be present. Yet, as described before in Section 4.3 we have a mechanism for wave generation plus the possibility of focusing reflection of internal waves at the sloping boundary, which leads to intensification and mixing. We will make plausible in the following sections that this scenario indeed occurred.

4.5.1 Mixing

Firstly we argue that we indeed observed mixing from the probe data. Figure 4.3 shows us the evidence. The leftmost panel represents a rather extreme experiment where modulation was on for six hours after which another profile was measured 63 hours later. The linear stratification was heavily modified. This is quite extraordinary as experience with previous experiments learned that the stratification could be maintained for periods up to one month (Sommeria, 2007, personal communication). The rightmost panel was an experiment where modulation was turned on for a shorter time. The profiles before and after one hour of modulation already show that mixing was in effect. But, even after modulation was stopped there was additional mixing as can be seen from the profile which was taken 3 hours after modulation. Of course internal wave beams continue to travel and reflect at sloping surfaces (which explains the ongoing mixing) until ultimately damped by viscous or nonlinear processes. The mechanisms leading to mixing discussed in the introduction (Section 4.3.1) may also contribute to the overall mixing, although they do not seem to explain the magnitude of the observed mixing.



(a) The dotted line represents the stratification before modulation, the solid line stands for the stratification after six hours of modulation and a subsequent 63 hours of solid body rotation.



(b) The solid line represents the stratification before modulation, after one hour of modulation the profile given by the dotted line was obtained. After this profile was taken the modulation was switched off and the dashed third profile was obtained after three hours of solid body rotation.

FIGURE 4.3: The profiles show the mixing that has taken place during two experiments. Both profiles were taken from probe C2, at $T = 22.53$ sec., aiming at the $(1, 1)$-attractor.

4.5.2 Formation of the attractor

This section presents an analysis of the formation of the attractor. We use PIV data from an experiment where we modulated at T=22.53 sec. The data are presented in the form of a harmonic analysis of the absolute value of the curl field. We expect that this representation emphasizes the main features of the attractor as discussed in Section 4.2.2. We segmented the data in time intervals of 50 pictures, corresponding to 147 seconds or approximately 6.3 wave periods. In the following we will briefly discuss the results for each interval.

0-147 seconds, Development of beam.

At the start of the modulation we clearly observe from figure 4.4 the development of a beam, emanating from the bottom-left corner. We estimate the beam to have a width of approximately 10 cm, maximum amplitudes increase from zero to 0.06 cm in six wave periods. From the crest-to-crest distance of the beam we can find the group velocity using equation (4.4), yielding about 17 cm/sec. The phenomenon of beams emanating from corners was noticed before by Beardsley (1970), in an experiment similar to ours. Since it is likely that similar beams could emanate from other corners of the basin we overlay

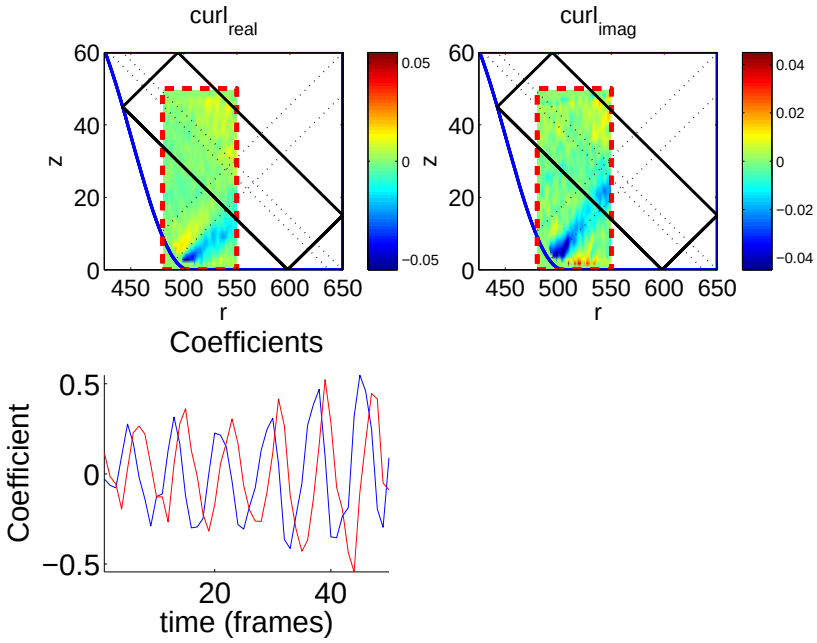


FIGURE 4.4: The lower panel shows the amplitudes of a Hilbert EOF (see Section 4.B) and corresponding real (upper left panel) and imaginary (upper right panel) fields. The signal is well represented by a superposition of both images with coefficients that evolve in time as in the lower panel. We see that the beam grows in amplitude while coming into existence in an oscillating manner, with a frequency equal to the modulation frequency.

the corresponding characteristics in the figures (eight dotted lines in Fig. 4.4 and following figures). However, even with this visual aid such beams are not visible in the pictures. Phase propagation is downwards, as would be observed by viewing a movie of the curl fields. This implies energy propagating upward and to the right. The directions of phase propagation and group velocity are consistent with (4.4) and (4.5).

148-249 seconds Fig. 4.5(a)

After approximately four minutes a second beam becomes visible. The location of the beam and the direction of phase propagation shows that it is likely that we see the attractor. It seems that we see the attractor developing, sooner than would be anticipated. One would expect, theoretically, that the attractor forms after several reflections of internal wave beams at the sloping side, at

which point the energy is focused. This theoretical model however disregards the influence of viscosity. Perhaps it is not necessary for the beams to focus repeatedly in order to form an attractor, since in a viscous fluid particle motion at a certain point will induce particle motion in a small neighborhood. In other words: in a hyperbolic system information is strictly carried along characteristics, our present viscous problem has elliptic traits and information is also diffused through the fluid interior. Second harmonics (not shown) were found to have amplitudes of a factor 10 less, and are insignificant at this point.

250-441 seconds. Fig. 4.5(b)

The harmonic analysis shows largely the same picture as before, but the amplitude at the attractor location has grown. Also the next interval (442 to 558 seconds) shows largely the same features as the current interval, we skip this interval.

589-735 seconds. Fig. 4.5(c)

The main feature in the first harmonics is a second beam, visible just above the bottom segment of the attractor. This could be a boundary layer eruption from the bottom-right corner of the domain. Second harmonics have now developed, mainly along the attractor location, with amplitude one fifth of the first harmonic. Most likely these are induced by nonlinear effects at the attractor location (see Section 4.2.2).

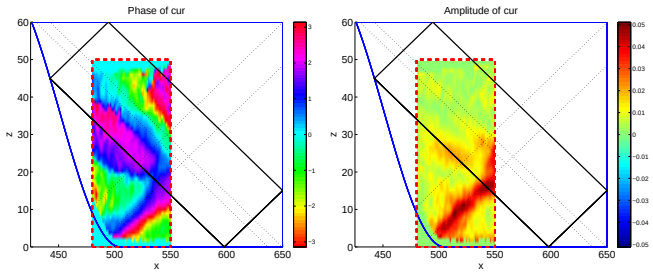
736-882 seconds. Fig. 4.6(a)

In this picture we observe a strong beam just above the slope, it might be the reflection of the beam from the bottom left corner, after two reflections. At the sloping wall it is focused, which would explain its increase in amplitude. The phase pictures are very clear, yet the amplitude pictures become difficult to interpret. The downward propagating phase at the top right corner of the observational window is consistent with the 'upper leg' of the attractor.We see that the angle differs from the angle at the attractor, as predicted from (4.3), (4.4) and (4.5). The amplitude of the second harmonics has grown, and is now comparable to that of the first harmonic. In the phase picture we see perfect anti-phase.

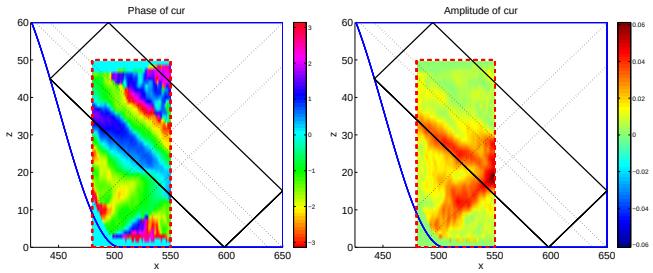
$>$ **3600 seconds.** Fig. 4.6(b)

From this point on it is too messy to draw any conclusions. In the long term the phase and amplitude lines do no longer follow the characteristic directions, but

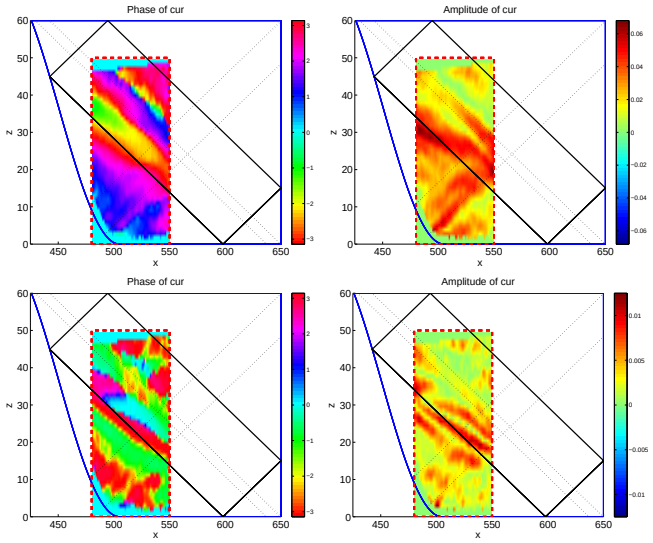
become flatter. It is likely that mixing has modified the stratification to such an extent as to modify the direction of wave rays and inhibit the further formation of the attractor. The beam above the slope however, was found to be present also for longer time scales. After one hour of modulation, and half an hour of solid body rotation the attractor was completely destroyed as can be seen in figure 4.6(b).



(a) Wave periods 7 to 12,. The direct beam has a high amplitude and the phase corresponds to the direction observed in a movie of the direct fields.

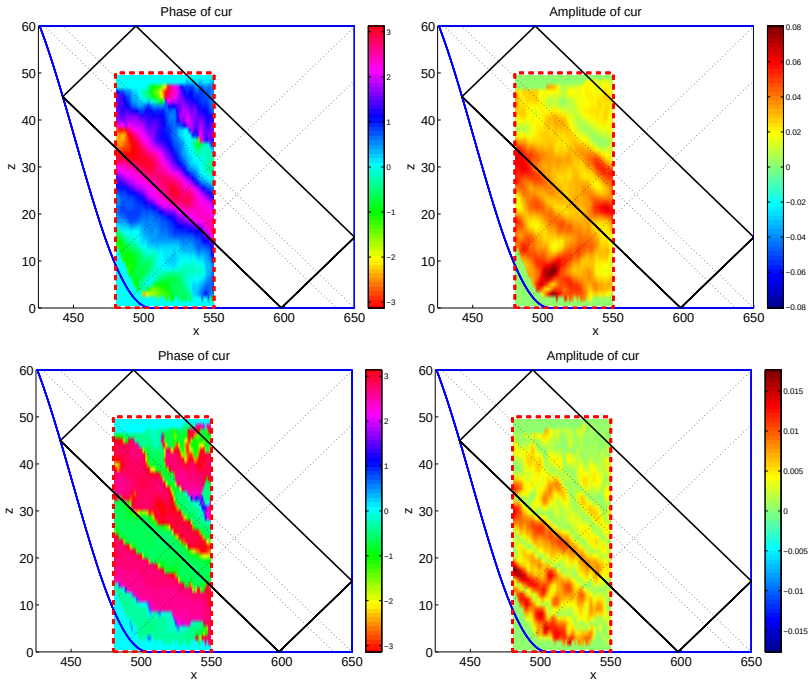


(b) Wave periods 13 to 18. Comparison to Fig. 4.5(a) reveals that the amplitude at the predicted attractor location has grown.

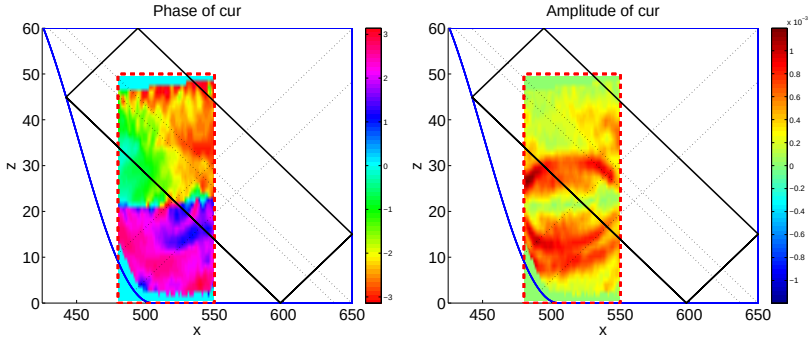


(c) Wave periods 25 to 30. The bottom row represents the second harmonic.

FIGURE 4.5: These panels show the phase field (left) and amplitude field (right) of the curl field, obtained by harmonic analysis (see Section 4.B).

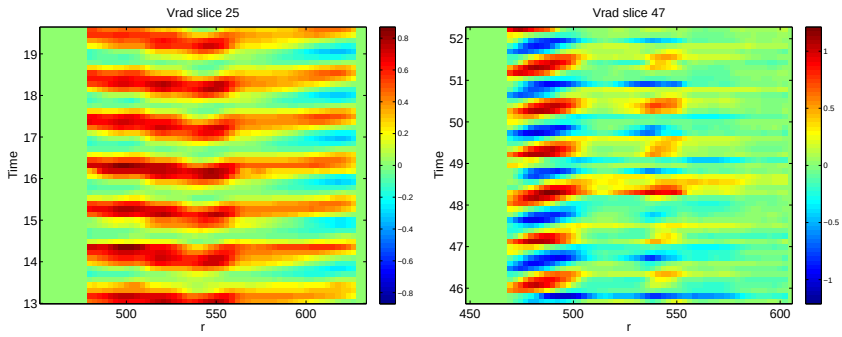


(a) Wave periods 31 to 36. There is now a clear beam above the slope, whose amplitude is comparable to that of the attractor. The bottom row represents the second harmonic.



(b) After approximately one hour of modulation and half an hour of solid body rotation the attractor was destroyed.

FIGURE 4.6: This figure continues figure 4.5.



(a) Laserlevel 21cm below surface. (b) Laserlevel 10cm above bottom.

FIGURE 4.7: Radius-time plots, The ordinate is labeled by t/T where t is time and $T = 22.53s$ is period. The abscissa shows distance in cm. The sloping bottom is at the left. Red corresponds to motion towards the slope, and blue indicates motion away from the slope.

4.5.3 The wave field in a horizontal plane

The previous section dealt with observations in a vertical plane. The experimental results were found to be in good agreement with the theory. For the behaviour of internal waves in three dimensions our knowledge is much more limited. In our discussion of the theory we assumed rotational invariance, although it is known from previous experiments that substantial spatial inhomogeneities may be present. In the following we report on our findings on the structure in the horizontal plane, unfortunately without a solid theoretical model for the structure of the internal waves which is as yet not fully understood in a three dimensional setting.

Firstly, we present so-called Hovmöller plots, showing the radial velocity at a fixed azimuthal angle and laser level height as a function of time. Such figures contain information about the radial phase velocity. From the fact that upward (downward) energy propagation corresponds with downward (upward) phase propagation we can determine whether the energy ray has an outward (i.e. off slope) or an inward component. This opens the possibility to connect observations from horizontal laser sheets to the vertical ones. Secondly, to shed light on the spatial structure of the wave field we will present figures obtained from a harmonic analysis, considering the component that oscillates harmonically with the forcing frequency. Fig. 4.7 comes from observations with a forcing period of 22.53 seconds, i.e. the case where the geometrically simplest attractor can be expected. In Fig. 4.7(a) the laser sheet was positioned close to the

fluid's surface (21cm below the surface). Note that the boundaries of the laser sheet do not coincide with the rotating tank's boundaries. This implies that the velocity is not necessarily zero along the borders of the laser sheet. Red (blue) indicates motion towards (away from) the slope. Velocity amplitude is about $1.1 cm\, s^{-1}$. Above the inner, sloping sidewall (to the left in the figure) an inward phase propagation with a phase speed of roughly $6 cm\, s^{-1}$ is visible. This is consistent with a wave beam pointing downward towards the slope. That beam might correspond with the upper-left segment of the wave attractor (see Fig. 4.1), which points downward towards the sloping boundary. Interestingly, we observe an outward phase propagation with a somewhat smaller phase speed near the vertical, outer rim of the tank. This can be interpreted as the signature of an energy ray which is emitted from the sloping boundary, pointing to the outer upper corner of the vertical cross section (see e.g. Fig. 4.4). The existence of such a ray is confirmed by observations where the laser sheet was positioned 10cm above the bottom of the rotating tank (Fig. 4.7(b)). In contrast to Fig. 4.7(a) the phase of a rather prominent and localized wave propagates away from the sloping sidewall with a phase speed of about $3 cm\, s^{-1}$ (the velocity amplitude in Fig. 4.7(b) is about $1.4 cm\, s^{-1}$). This observation is consistent with the one where a vertically oriented laser sheet was used. In Fig. 4.5 we can clearly identify an energy ray emitted upward from the foot of the sloping sidewall.

Another interesting feature can be observed in the center of both figures, that is at $r \approx 550 cm$ in Fig. 4.7(a) and Fig. 4.7(b). There appears to be another localized wave with a phase speed away from the slope. Concerning Fig. 4.7(b) it is likely that the local wave visible is related to the lower part of the wave attractor, where energy is propagating downward. Indeed, this view can be confirmed by experiments with a vertically oriented laser sheet (Fig. 4.5 and 4.6). In contrast, the wave perturbation in the center of Fig. 4.7(a) might stem from the upper left corner and energy is radiated downward (we can only speculate here because the vertically oriented laser sheets do not cover the upper 10cm of the fluid).

In order to get insight into the spatial structure of the oscillations shown in the figures, we performed a harmonic analysis (see Section 4.B for technical details) for the time-window displayed in the figures. The result (showing the amplitude of the first harmonic component with the forcing period of 22.53s) corresponding to Fig. 4.7(a) and 4.7(b) are given in Fig. 4.8.

An harmonic analysis of the fields corresponding to Fig. 4.7(a) shows that amplitude becomes very strong along the outer rim of the tank, whereas the oscillation vanishes above the slope. Further away the original pattern seems to recover. The phase (not shown) exhibits quite a complex pattern, too. Roughly speaking, we observe phase velocities with different signs at the inner and outer boundary, that is phase speeds pointing towards the boundaries. How-

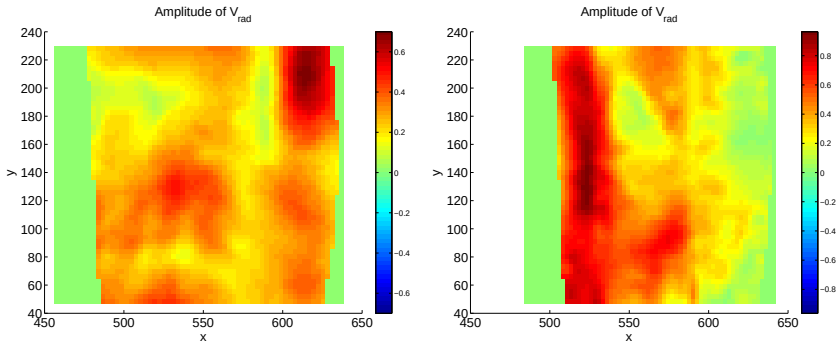


FIGURE 4.8: Amplitude of the time harmonic component of the motion $T = 22.53s$ corresponding to Fig. 4.7(a) and Fig. 4.7(b).

ever, we observe a strong asymmetry of the wave's phase above the slope.

In contrast, figure 4.7(b) shows large amplitudes above the slope and a weaker second maximum in the center of the picture. Remarkably, there is a sharp transition at $r \approx 590cm$, where the amplitude drops off. As said above, we think that this second maximum marks a regions where the waves are focused due to reflections from the sloping boundary. The phase plot (not shown) confirms our earlier findings that the phase speed above the slope points away from this region.

In summary we can roughly say that the results obtained by the harmonic analysis confirm the findings from the $r - t$-plots, with consistent phase behaviour. However, we observe an unexpected strong spatial asymmetry with patchy patterns of amplitude and phase. One explanation could be that wave amplitudes become too large for a simple linear (harmonic) description. Waves could break locally which might then lead to spatial inhomogeneity.

4.6 Secondary phenomena

We now proceed with the discussion of two consequences of the mixing process, the generation of a localized unstable mean flow and subsequent generation of Rossby waves. After an introduction, we continue with experimental evidence from our experiment.

4.6.1 Mean flow generation and Rossby waves

In a steadily rotating frame of reference, rotating with the average rotation rate, the modulation Ω_0 is perceived as a 'tidal' flow. Gravito-inertial waves can be forced by tidal Ekman transport convergence and divergence, or boundary layer eruptions.

In rotating fluids, tidal flow along continental slopes has an other important effect, it generates a mean flow by means of tidal rectification: advection of tidal perturbations by the barotropic tide (Huthnance, 1973). This is often cast in terms of vortex dynamics, in which topographic depth variation leads to compression or stretching of vortex tubes, which are advected by the barotropic tide and which is converted into tidal residual currents because of a time lag that arises due to frictional retardation (Zimmerman, 1978). Stratification inhibits vertical motion further away from the bottom, restricting the compression and stretching of vertically-aligned vortex tubes forming part of the rectification process, and thus also the mean flow generation, to the bottom vicinity (Maas & Zimmerman, 1989*b*,*a*). These authors show that the mean flow is accompanied by circulation cells in the vertical, the strongest of which predicts down-welling higher up on the slope and upwelling over its foot. For the present experimental set-up, where the inner wall is sloping, this mechanism produces an along-isobath, bottom-intensified retrograde flow. The establishment of such a sheared mean flow leads to a venue to another mechanism by means of which the fluid can be mixed. This flow may, after all, turn baroclinically or barotropically unstable, leading to the establishment of long topographic Rossby waves, indications of which have been found in the measurements.

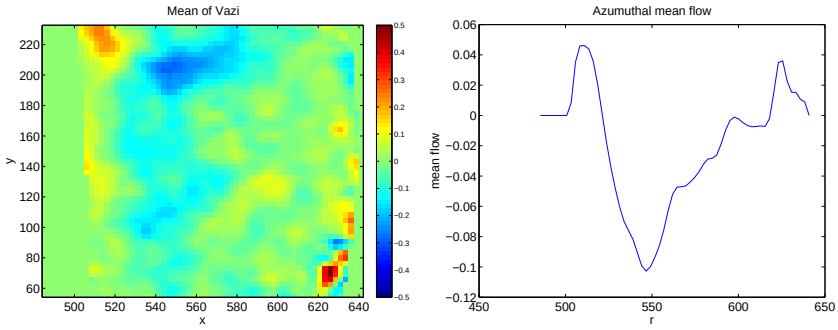
A mean flow is also generated by impinging internal waves (Thorpe, 1999). The momentum carried by them can be partly released, in a process likened to radiation stress divergence of surface waves (Dunkerton *et al.*, 1998). In a non-rotating frame these authors argue that the generated mean flow is in the same direction as the along-isobath component of the obliquely incident internal wave field. The frame's background rotation will break this symmetry and will lead to a preferential orientation of the mean flow. (This symmetry breaking is not explicitly discussed, but can be inferred from Thorpe (2000, Fig. 3b)). Zikanov & Slinn (2001) argue that as the mean flow develops, breaking extends further and further away from the boundary, being located at the position where mean flow equals phase speed (at a critical layer). Thorpe (1999) argues that over the slope also a mean Eulerian circulation is set up in the vertical plane, with up-slope bottom currents, contrasting those from the baroclinic tidal rectification.

4.6.2 Azimuthal mean flow

As was discussed in Section 4.6.1, tidal flows along sloping boundaries can generate an alongshore mean flow in rotating fluids. In this section we confirm such a flow in our measurements. In Fig. 4.9(a) we show the time mean of the azimuthal velocity field over the 6 periods shown in Fig. 4.7(b). Obviously, the azimuthal mean velocity shows a prominent band of negative (retrograde) velocity somewhat left to the center of the annulus, and a narrower band of positive (prograde) velocity above the inner, sloping boundary. This can be seen even clearer if we consider the azimuthal mean of the time mean azimuthal velocity field (see Fig. 4.9(b)). Note that the maximum of the mean flow is about 12% of the amplitude of the time harmonic component of internal waves with the forcing frequency. A (barotropic) multiple jet flow of the type shown in Fig. 4.9(b) is typically unstable. Thus the occurrence of such a mean flow could explain the existence of long topographic Rossby type of waves that will be discussed in Section 4.6.3. Indeed, in the upper layers the mean velocity field shows a wave-like structure over the slope, similar to that of Fig. 4.10. At the surface, the azimuthal mean flow is dominated by a strong pro-grade current along the outer rim of the annulus, with a maximum four times larger than the maximum in Fig. 4.9(b). Possibly, wind stress might affect the surface layer of the fluid in the rotating annulus. At intermediate depth, the structure of the flow is similar to the one shown in Fig. 4.9(b), but the amplitude is somewhat larger. It should be noted that for the upper layers the laser sheet shows a larger part of the fluid over the slope. It is thus understandable that the maximum of the retrograde flow increases with height. Nevertheless, the amplitude of the pro-grade flow shows the same behaviour.

The mechanisms that drive banded mean flows in rotating stratified fluids are still not fully understood. Recently, Read *et al.* (2004) created banded mean flows in laboratory experiments on the same rotating platform that we used for the experiments described here. In contrast to our experiment, they used a homogeneous fluid. Moreover, in their experiments the slope covered the full gap of the annulus and the mean flows were excited by convection, triggered by spraying a salty fluid onto the surface. In our experiment wave induced mixing could have a similar effect as convection and thus a large scale azimuthal mean flow could be driven primarily by an inverse energy cascade in geostrophic turbulence (Read *et al.*, 2004). On the other hand, the retrograde flow above the slope could also be explained by the mechanism given in Maas & Zimmerman (1989*b*) and Maas & Zimmerman (1989*a*). However, the retrograde flow has a rather barotropic structure, whereas the flow described by Maas and Zimmerman is strongest at the bottom. [3]

[3]In the paper by Read *et al.* (2004) experiments without a slope are described also. They find



(a) Spatial field (b) Azimuthal mean

FIGURE 4.9: Time mean of the azimuthal flow field.

4.6.3 Rossby waves

The probe data show low frequency oscillations too. These frequencies come from the slowly propagating topographic Rossby waves. In Fig. 4.10 we show the radial velocity field at a depth of 8cm, averaged over two periods. The azimuthal wavelength of the pattern is about 2.0m and the wave propagates with an azimuthal phase speed of about $0.8 \times 10^{-3} m\,s^{-1}$ in retrograde direction (towards smaller values of y).

There is a however a problem: a phase speed of $0.8 cm/s$ is too slow for topographic shelf waves. The radial velocity for such waves is proportional to $J_2(K)$, where J_2 indicates a Bessel function of the first kind, of order 2. The value K is an eigenvalue computed from boundary and matching conditions (see Pedlosky, 1992, p632). We estimate a value of K from the wavelength and the phase speed which is one order of magnitude too large.

Inspecting Fig. 4.9(b), we find a large mean flow, about two orders of magnitude larger than the observed phase speed. From the figure we estimate a mean flow close to the sloping inner wall of about $5 cm/s$ in which the wave propagates. Taking this mean flow into account, we find a consistent value for K. We have to keep in mind however that the mean flow is not truly constant, and the model we present here is a simplification.

Again considering Fig. 4.9(b), one observes that the second derivative of the mean flow changes sign at about $25 cm$ from the sloping wall (at a height of $8 cm$ from the fluid surface). At this location a strong shearing flow is present.

---

a azimuthal mean flow that has a somewhat similar structure to the one described here. The retrograde part seems to be broader and there is less jet structure. They call this a single-cell circulation and think it is a wind stress related feature.

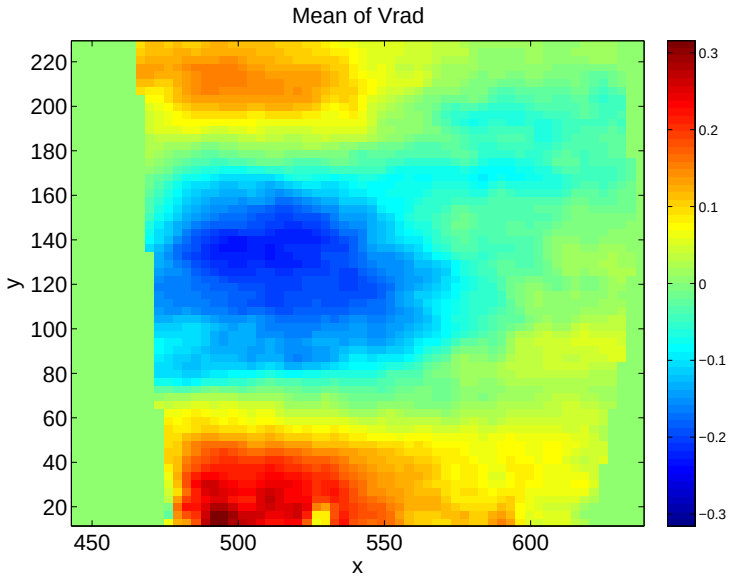


FIGURE 4.10: Mean radial velocity field, 8cm below the surface. The shelf is on the left side of the figure. The wave propagates in retrograde direction, towards small values of y.

As discussed in Section 4.6.1, it is this shearing flow which could cause the instabilities that give rise to the Rossby waves we observe. Furthermore, the shearing mean flow might also act as a barrier that traps the Rossby wave (see e.g. Harlander *et al.*, 2000). This is not inconsistent with the wave shown in figure 4.10, where we see the wave decaying exponentially at about $x = 25cm$.

The pattern does not depend much on depth, thus it is not a bottom trapped mode. The internal Rossby radius of deformation $L_D = N_s D/f$ ($N_s = 2f$, $D = 0.6m$) is approximately twice the wavelength of the wave shown in Fig. 4.10 which explains its quasi-barotropic structure (Pedlosky, 1992, p.411) and its similarity to Rossby β-waves.

4.7 Conclusions and Discussion

In this paper we collected experimental evidence for strong mixing in a stratified rotating basin. Internal wave reflection and the formation of a wave attractor were found to be the cause of this mixing of the density profile. As a forcing mechanism for the internal waves we identified a boundary layer eruption at

a corner of the domain. We will briefly review each of the mentioned steps.

The remarkably strong mixing of the density profile was discussed in Section 4.5.1. Since the fluid was not directly disturbed, no obvious turbulence is present, and the basin is axisymmetric, none of the usual mechanisms for mixing is present. This compels us to examine other avenues, and the remainder of the paper was devoted to the detection of internal waves and wave beams that act as a mixing device.

We firstly discussed the boundary layer, and the possibility of its eruption into the interior. Theoretically the boundary layer eruption was made plausible, by modification of the theory by MacCready & Rhines (1991), opening the possibility for singularities in boundary layer depth in Section 4.3 and appendix 4.A. Experimentally, the beam was very clearly observed (Section 4.5.2). It is as yet unclear why only waves were generated at the bottom of the shelf. Of course, it is the only location at the boundary where the critical angle is truly existent. However, it was argued that the critical angle is supposed to be 'degenerately buried' in sharp corner points. Perhaps this view does not hold, further experimentation could shed some light on this issue.

Using harmonic analysis techniques we obtained phase and amplitude pictures in radial cross-sections. Wave beams were clearly observed. Furthermore, phase and amplitude strongly suggest the formation of a wave attractor. In Section 4.5.2 the formation of the wave attractor was presented. Also, diagrams of the flow in a radius-time plot show not only data consistent with the existence of a wave attractor (Section 4.5.3), but also the generation of mean flow (Section 4.6.1). Furthermore, we report in this section on Rossby waves, which may also be indicative of strong mixing (Solomon *et al.*, 1993). Considering the above, we believe that we have established the scenario sketched at the beginning of the introduction. Possibly this insight may be of use to explain mixing in oceans, where the combination of rotation, stratification and boundary topography can similarly lead to efficient localized mixing. This in turn would have consequences in climate research and would impact large scale ocean flow (Wunsch & Ferrari, 2004) and biology (Witbaard *et al.*, 2005).

4.8 Acknowledgments

Help with data acquisition and processing by F.Eijgenraam, and the Coriolis team (S. Vibaud, H. Didelle, O. Praud, A. Fincham, S. Mercier, E. Thivolle and J. Sommeria) is gratefully acknowledged. The access to the experimental facility of the Coriolis Laboratory was supported by the European Commission, Enhancing Access to Research Infrastructures action of the Improving Human Potential programme of FP5, under contract HPRI-CT-1999- 00006. A. N. Swart was funded by NWO as part of the project Numerical Determination of Wave

Attractors.

Appendices to Chapter 4

4.A Physical model

Equations of Motion

This section will introduce the physical model for describing gravito-inertial waves. Under the Boussinesq approximation the linearized equations of motion for the velocities $\mathbf{u} = (u, v, w)$, density ρ and pressure p may be written

$$\begin{aligned} \mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} &= -\nabla p - g\rho \mathbf{k} + \nu \Delta \mathbf{u}, \\ \rho_t - \frac{N^2(z)\rho^*}{g} \mathbf{u} \cdot \mathbf{k} &= \kappa \Delta \rho, \\ \nabla \cdot \mathbf{u} &= 0. \end{aligned} \tag{4.9}$$

The depth-dependent quantity $N^2 = -\frac{g}{\rho^*}\{\frac{d\rho_0}{dz} + \frac{\rho_0 g}{c_s^2}\}$ is the Brunt-Väisälä (or inertial) frequency, the constant ρ^* is a depth-averaged density, $\rho_0(z)$ is a static density profile, g is the gravitational constant. Furthermore, ν is the kinematic viscosity and κ is the thermal diffusivity.

We ignore the non-linear terms and transform the equations to a frame of reference that rotates along angular velocity vector Ω_0. This introduces two fictitious forces:

1. the Coriolis force, $2\Omega \times u$,
2. the centrifugal force, $\Omega \times \Omega \times r = -\Omega^2 \nabla r^2/2$,

where r and u denote the 3D position and velocity vectors respectively, as observed from the co-rotating frame. Note now that the centrifugal force may be absorbed in a reduced pressure P, as demonstrated before in Section 1.2. This gives us,

$$\begin{aligned} \mathbf{u}_t + \mathbf{f} \times \mathbf{u} &= -\nabla P - g\rho \mathbf{k} + \nu \Delta \mathbf{u}, \\ \rho_t - \frac{N^2(z)\rho^*}{g} \mathbf{u} \cdot \mathbf{k} &= \kappa \Delta \rho, \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned}$$

where f is twice the angular velocity Ω and $\mathbf{f} = (0, 0, f)$.

We have now transformed the equations to a frame of reference, rotating at a rate Ω_0, while in our experiment we have set a total rotation rate of $\Omega = \Omega_0 + \Omega'(t)$, consisting of a constant, Ω_0, and a time periodic part, $\Omega'(t)$. However, one can show that the effect of the perturbation $\Omega'(t)$ is felt in the boundary conditions only. Usually one would pose $(u, v, w) = 0$ boundary conditions in

a viscous context. In our case however, instead of $v = 0$, the time dependent modulation induces a boundary condition $v = \Omega' r$.

We now perform some calculations on the governing equations, in order to obtain a manageable expression. Define $\mathcal{L} = \partial_t - \nu\Delta$ and $\mathcal{T} = \partial_t - \kappa\Delta$, with $\nu = \frac{\mu}{\rho^*}$. Then the governing equations can be written

$$\begin{aligned} \mathcal{L}\mathbf{u} + \mathbf{f} \times \mathbf{u} &= -\frac{\nabla P}{\rho^*} - \frac{g}{\rho^*}\rho\mathbf{k}, \\ \mathcal{T}\rho &= \frac{N^2(z)\rho^*}{g}\mathbf{u}\cdot\mathbf{k}, \\ \nabla\cdot\mathbf{u} &= 0. \end{aligned}$$

Apply $\mathcal{T}$ to the first equation. This allows us to insert the second equation, yielding

$$\mathcal{T}\mathcal{L}\mathbf{u} - f\mathcal{T}\mathbf{u}\times\mathbf{k} = -\frac{\mathcal{T}\nabla P}{\rho^*} - N^2(z)(\mathbf{u}\cdot\mathbf{k})\mathbf{k}.$$

We would like to eliminate the pressure by utilizing the vector identity $\nabla \times (\nabla p) = \mathbf{0}$. Note that $\mathcal{T}$ and $\mathcal{L}$ commute with the curl operator and we obtain

$$\mathcal{T}\mathcal{L}\nabla\times\mathbf{u} - f\mathcal{T}\nabla\times\mathbf{u}\times\mathbf{k} = -N^2(z)\nabla\times\mathbf{u}\cdot\mathbf{k}\mathbf{k}. \tag{4.10}$$

At this point we choose a cylindrical coordinate system with axes (r, θ, z). Furthermore we use rotational symmetry and set all derivatives with respect to the θ coordinate to zero [4] In this case the following identities hold

$$\begin{aligned} \nabla\cdot\mathbf{u} &= r^{-1}\partial_r(ru) + w_z, \\ \nabla\times\mathbf{u} &= (-v_z, u_z - w_r, r^{-1}\partial_r(rv)), \\ \Delta u &= u_{rr} + r^{-1}u_r + u_{zz}, \end{aligned}$$

In component form the equation (4.10) now reads

$$\mathcal{T}\mathcal{L}\begin{pmatrix} -v_z \\ u_z - w_r \\ r^{-1}\partial_r(rv)\end{pmatrix} - f\mathcal{T}\begin{pmatrix} u_z \\ v_z \\ -r^{-1}\partial_r(ru)\end{pmatrix} + N^2(z)\begin{pmatrix} 0 \\ w_r \\ 0\end{pmatrix} = 0;$$

We still need to satisfy the continuity equation $\nabla\cdot\mathbf{u} = 0$, which may be accomplished by introducing the stream function Ψ, satisfying $(\Psi_z, -\Psi_r) = (ru, rw)$.

[4] Instead of setting derivatives to zero, there are other possibilities. E.g., separation in the form $\mathbf{u}(r,\theta,z) = \mathbf{f}(r,z)\mathbf{g}(\theta)$, with a 2π-periodic function g, but this would complicate the discussion.

This casts the equations in the form

$$\mathcal{T}\mathcal{L}\begin{pmatrix} -v_z \\ r^{-1}(\Psi_{rr}+\Psi_{zz})-r^{-2}\Psi_r \\ r^{-1}\partial_r(rv)) \end{pmatrix} + N^2(z)\begin{pmatrix} 0 \\ r^{-1}\Psi_{rr}-r^{-2}\Psi_r \\ 0 \end{pmatrix} = f\mathcal{T}\begin{pmatrix} r^{-1}\Psi_{zz} \\ v_z \\ -r^{-1}\Psi_{zr} \end{pmatrix}$$

Eliminate the velocity v from the second equation by applying $\mathcal{L}$ to it and inserting the first equation,

$$\mathcal{L}^2\mathcal{T}(r^{-1}(\Psi_{rr}+\Psi_{zz})-r^{-2}\Psi_r)+\mathcal{L}N^2(z)(r^{-1}\Psi_{rr}-r^{-2}\Psi_r) = -f^2\mathcal{T}r^{-1}\Psi_{zz}.$$

Introduce $P = r^{-1}\Psi$ and rewrite the above equation,

$$\mathcal{L}(\mathcal{L}\mathcal{T}+N^2(z))(P_{rr}+P_{zz})+(\mathcal{T}f^2-\mathcal{L}N^2(z))P_{zz} = -\mathcal{L}(\mathcal{L}\mathcal{T}+N^2(z))(r^{-1}P)_r. \tag{4.11}$$

This general equation is much too formidable to solve directly and we will make some simplifications. We look for wave solutions to (4.11), i.e. we look for solutions of the form

$$P(r,z,t) = \Re[\tilde{P}(r,z)e^{-i\omega t}], \tag{4.12}$$

but we will drop the tilde for notational convenience.

In interior of the fluid we assume the viscosity and diffusivity to be negligible. This is justified, since viscosity and diffusion coefficients are small (Batchelor, 1967). In this case the operators $\mathcal{L}$ and $\mathcal{T}$ simplify to $\mathcal{L} = \mathcal{T} = i\omega$ and (4.11) reduces to

$$P_{rr}-\lambda^2 P_{zz} = -(r^{-1}P)_r,$$

with

$$\lambda^2 = \frac{\omega^2-f^2}{N^2(z)-\omega^2}.$$

The Boundary Layer

At the boundary of the fluid domain the oscillating fluid has to adjust to the boundary. We are interested in the boundary layer behaviour and consider a small segment of the boundary locally. We describe this segment *locally* as plane, inclined at an angle ϕ with the horizontal and parallel to the θ axis. We rotate the coordinate frame in order to have the z axis perpendicular to the slope, the θ axis cross slope and the r axis along-slope. We add a prime to quantities in the rotated system. See Figure 4.11 for a schematic picture of the axis rotation. We need the following transformations

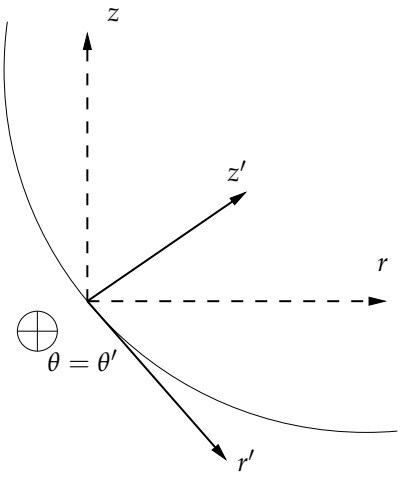


FIGURE 4.11: The local coordinate system (primed) for use in the boundary layer calculations. The curve represents the boundary of the domain. The inward normal is z', while r' is tangent to the boundary.

$$\begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} = \begin{pmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}.$$

The relevant differential operators transform as

$$\begin{aligned} \partial_{rr} &= \cos^2(\phi)\partial_{r'r'} + \sin^2(\phi)\partial_{z'z'} - 2\sin(\phi)\cos(\phi)\partial_{r'z'}, \\ \partial_{zz} &= \sin^2(\phi)\partial_{r'r'} + \cos^2(\phi)\partial_{z'z'} + 2\sin(\phi)\cos(\phi)\partial_{r'z'}, \\ \partial_{rr} + \partial_{zz} &= \partial_{r'r'} + \partial_{z'z'}. \end{aligned}$$

At this point we need some additional simplification. In our experiment $r \approx 6$ and $z \approx 0.6$, thus $r \gg z$. This justifies setting all terms of (4.11) involving r^{-1} to zero. Furthermore we assume that in the boundary layer derivatives with respect to r' are negligible compared to derivatives with respect to z'. This means that we assume rapid variation in the boundary layer z' direction (needed to adjust the interior of the fluid to the boundary) and slow variation along the boundary layer. This gives us

$$\begin{aligned} \partial_{rr} &= \sin^2(\phi)\partial_{z'z'} \\ \partial_{zz} &= \cos^2(\phi)\partial_{z'z'} \\ \Delta &= \partial_{z'z'}. \end{aligned}$$

The equation for the boundary layer becomes

$$\mathcal{L}(\mathcal{L}\mathcal{T} + N^2(z))P_{z'z'} + (\mathcal{T}f^2 - \mathcal{L}N^2(z'))\cos^2(\phi)P_{z'z'} = 0.$$

We simplify by assuming monochromatic wave motion, as in (4.12) and rewrite the equation in terms of the along-slope velocity $u = P_{z'}$:

$$i\omega[\tilde{f}^2 + \tilde{N}^2 - \omega^2]u^{(1)} + [\omega^2(2\nu + \kappa) - (\kappa\tilde{f}^2 + \nu\tilde{N}^2)]u^{(3)} + i\nu\omega[\nu + 2\kappa]u^{(5)} - \kappa\nu^2u^{(7)} = 0, \quad (4.13)$$

with $\tilde{f} = \cos(\phi)f$ and $\tilde{N} = \sin(\phi)N$. There are two cases of special interest to our setting, oscillating and non-oscillating flow. These are the topics of the next two paragraphs

Non-oscillating flow

When we have a steady flow, then $\omega = 0$ and (4.13) reduces to

$$\frac{\sigma\tilde{N}^2 + \tilde{f}^2}{\nu^2}u^{(3)} = -u^{(7)},$$

with $\sigma = \nu/\kappa$. The velocity is now a real valued function. The solution to this ordinary differential equation is given by

$$\begin{aligned} u(z) &= e^{\frac{1}{\sqrt{2}}\sqrt[4]{c}z}[b_1e^{i\frac{1}{\sqrt{2}}\sqrt[4]{c}z} + b_2e^{-i\frac{1}{\sqrt{2}}\sqrt[4]{c}z}] \\ &+ e^{-\frac{1}{\sqrt{2}}\sqrt[4]{c}z}[b_3e^{i\frac{1}{\sqrt{2}}\sqrt[4]{c}z} + b_4e^{-i\frac{1}{\sqrt{2}}\sqrt[4]{c}z}] + b_5 + b_6z + b_7z^2, \end{aligned}$$

with $c = \frac{\sigma\tilde{N}^2+\tilde{f}^2}{\nu^2}$ and constants $b_1, \ldots, b_7$. We consider the following boundary conditions,

$$\begin{aligned} u(0) &= -U, \\ \lim_{z\to\infty} u(z) &= 0. \end{aligned}$$

The given constant U represents the value of the interior solution at the boundary, which needs to be canceled by the boundary layer velocity. Far away from the boundary, the boundary layer velocity should vanish, which amounts to the second condition stated above. In order to satisfy the boundary condition at infinity it must be the case that $b_1 = b_2 = b_5 = b_6 = b_7 = 0$ and we find that the real part of the velocity is given by

$$u(z) = e^{-\frac{1}{\sqrt{2}}\sqrt[4]{c}z}[c_1\cos(\frac{1}{\sqrt{2}}\sqrt[4]{c}z) + c_2\sin(\frac{1}{\sqrt{2}\sqrt[4]{c}z})].$$

The boundary conditions at $z = 0$ then imply

$$u(z) = e^{-\frac{1}{\sqrt{2}}\sqrt[4]{c}z}[U\cos(\frac{1}{\sqrt{2}}\sqrt[4]{c}z) + c_2\sin(\frac{1}{\sqrt{2}}\sqrt[4]{c}z)]$$

for some unknown constant c_2. Note that, in principle, the coefficient may be determined by considering the vanishing of the velocity component v at the boundary. The viscosities (ν) and diffusivities (κ) are small, thus we are interested in the case $c \to \infty$. When c tends to infinity, the exponential function is relevant on spatial scales $z = \mathcal{O}(c^{-1/4})$. We conclude that the boundary layer height extends over

$$h = \mathcal{O}(\sqrt[4]{\frac{\nu^2}{\sigma\tilde{N}^2 + \tilde{f}^2}}), \text{ as } \nu \downarrow 0 \text{ or } \kappa \downarrow 0.$$

Oscillating flow

Since κ and ν are very small parameters, we also neglect their products and obtain the following differential equation

$$\frac{i\omega}{\kappa}\frac{\tilde{f}^2 + \tilde{N}^2 - \omega^2}{\sigma(\tilde{N}^2 - 2\omega^2) + (\tilde{f}^2 - \omega^2)}u^{(1)} = u^{(3)}$$

with $\sigma = \nu/\kappa$.

Recall that the oscillatory nature of the flow was introduced by $u(z,t) = \Re(e^{i\omega t}u(z))$. By definition of differentiation of complex valued functions $f_x(x) = \partial_x\Re(f) + i\partial_x\Im(f)$ we obtain a system of equations, from which we obtain the equation for the real part of u,

$$-\frac{\omega^2}{\kappa^2}\left(\frac{\tilde{N}^2/(\tilde{N}^2 - 2\omega^2) - \tilde{\lambda}^2}{\sigma - \tilde{\lambda}^2}\right)^2 u^{(1)} = u^{(5)},$$

with $\tilde{\lambda}^2 = (\omega^2 - \tilde{f}^2)/(\tilde{N}^2 - 2\omega^2)$. We proceed as in the non-oscillating case to find the physically acceptable boundary layer depth and find

$$h = \mathcal{O}(\sqrt{\frac{\kappa}{\omega}\left(\frac{\sigma - \tilde{\lambda}^2}{\tilde{N}^2/(\tilde{N}^2 - 2\omega^2) - \tilde{\lambda}^2}\right)}), \text{ as } \nu \downarrow 0 \text{ and } \kappa \downarrow 0.$$

4.B Methods of analysis

Vector field correction

The first basic correction to the vector fields was the application of a mask, to hide areas of the field of view. In the vertical view we mask the part of the

image that lies outside the fluid domain. Additionally we mask a strip at the top of the field where the laser intensity was insufficient.

After applying the masks, there are two major corrections which need to be applied to the vector fields. Firstly, the horizontal fields are measured in a Cartesian (x, y) frame, yet we would like velocities in the radial and azimuthal directions. The reason for this is that the modulation of the turntable induces a relative azimuthal flow [5]. We like to subtract this flow, which is most naturally done in a cylindrical frame of reference. Denote unit vectors in the (x, y) frame by $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ and unit vectors in the (r, θ) frame by $\hat{\mathbf{r}}$ and $\hat{\mathbf{\Theta}}$. Let θ be the angle between $\hat{\mathbf{x}}$ and $\hat{\mathbf{r}}$. The relation between the unit vectors can be easily seen to be

$$\begin{aligned} \hat{\mathbf{x}} &= \hat{\mathbf{r}}\cos(\theta) - \hat{\mathbf{\Theta}}\sin(\theta), \\ \hat{\mathbf{y}} &= \hat{\mathbf{r}}\sin(\theta) + \hat{\mathbf{\Theta}}\cos(\theta). \end{aligned}$$

The correspondence between the velocities $\mathbf{u} = (u, v)$ measured in the (x, y) frame and the velocities $\mathbf{u}' = (u_r, u_\theta)$ required in the (r, θ) frame can now be found by projection

$$\mathbf{u}' = (\mathbf{u} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}} + (\mathbf{u} \cdot \hat{\mathbf{\Theta}})\hat{\mathbf{\Theta}},$$

which yields

$$\begin{aligned} u_\theta &= v\cos(\theta) - u\sin(\theta) \\ u_r &= v\sin(\theta) + u\cos(\theta). \end{aligned}$$

The flow is a direct sum of the azimuthal flow induced by the modulation and a 'perturbation' flow which is of interest

$$u_\theta = u_{per} + u_{azi}.$$

In the experiment the rotation of the table is the sum of a steady rotation plus a modulation. We assume that the steady rotation has produced a solid body rotation that is perceived as zero motion relative to the moving frame of reference. The modulation of the azimuthal velocity induces the linear profile

$$v_{azi}(r) = v_{rim}\sin(\omega t - \omega_0)r/R, \tag{4.14}$$

with settings $v_{rim} = 6$ cm/sec, radius $R = 6.5m$ and frequencies ω as given in table 4.1). The phase ω_0 is easily estimated using an harmonic analysis (see Section 4.B).

[5] Although the camera is mounted on the turntable, we are not in an co-modulating inertial frame. Viscous friction at the boundary introduces a retardation, and therefore a phase difference with the modulation of the turntable.

Correction for laser sheet projection

The second correction stems from the finite thickness of the laser sheet combined with position and viewing angle of the camera. The camera projects motion perpendicular to the laser sheet in a direction parallel to the sheet. This creates artificial velocity components. This effect was found to be significant only in the case of vertical sheets, where the v_{azi} component is projected. We assume that v_{azi} is completely determined by the modulation of the turntable, not by waves propagating in the azimuthal direction (in accordance with our assumption of rotational symmetry of the field). We model the projection as follows, suppose the true velocity is $\mathbf{u}_{azi} = (0, v_{azi}, 0)$, yielding measured velocity $\mathbf{u}_m = (u_m, 0, w_m)$ which is the sum of a true velocity in the radial plane plus a projection:

$$\mathbf{u}_m = (u, 0, w) + (u_p, 0, w_p). \tag{4.15}$$

If we can find the expressions for the the projections, assumed to be stemming from the known azimuthal velocity, then we may subtract them from the measured velocities and obtain the corrected fields. Consider figure 4.12, where we model the geometry of the projections. The following equalities hold,

$$\mathbf{C} + \mathbf{\Delta} = (u_p, 0, w_p), \tag{4.16}$$

$$\mathbf{C} + a\mathbf{\Delta} = \mathbf{v}_{azi}. \tag{4.17}$$

The *fixed* vector C' points at the camera and is the sum of a vector $X = (x, 0, z)$, and a vector $C = (C_1, C_2, C_3)$ pointing to the camera from this position. For these vectors the relation $\mathbf{C} = \mathbf{C}' - \mathbf{X}$ holds. We estimate $\mathbf{C}' \approx (-15, 270, 11)/s$, where s is the sampling rate in *sec./picture*. We can solve the system for the unknowns u_p and w_p to obtain

$$u_p = \frac{v_{azi} C_1}{v_{azi} - C_2},$$

$$w_p = \frac{v_{azi} C_3}{v_{azi} - C_2}.$$

We remove this contribution, using (4.15), to find the corrected fields

$$u = u_m - u_p,$$

$$w = w_m - w_p.$$

We find however that the correction is not always perfect, often residual gradients are found. For this reason we often consider the *curl field* of the velocities. This works very well, since along an internal wave ray (or attractor) there is considerable shearing motion. The curl responds to this shearing motion, independently of the angle of the wave ray. Additionally, the curl does not respond

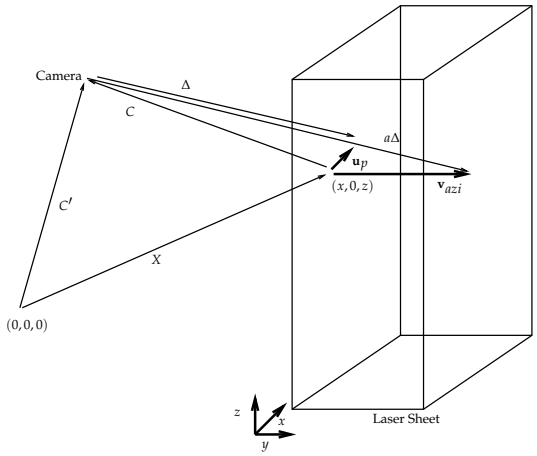


FIGURE 4.12: This figure shows the laser sheet with the camera and projections. In order to make the projection more clear, the dimension in the y direction is highly exaggerated. Thick arrows indicate velocity vectors. The fixed vector C' points towards the camera. Consider a point $(x, 0, z)$ at the 'front' of the laser sheet. The (non fixed) vector C points towards the camera. At this point we have a velocity $\mathbf{v}_{azi} = (0, v_{azi}, 0)$ that is projected to yield a velocity $\mathbf{u}_p = (u_p, 0, w_p)$. The projection is in the direction of Δ, that points to the tip of $\mathbf{u}_p$. We can scale Δ with an as yet unknown scalar a to have it point at the tip of $\mathbf{v}_{azi}$. Note how $C, \Delta, a\Delta, \mathbf{v}_{azi}$ and $\mathbf{u}_p$ are all in one plane.

to gradients in the vector field, thereby effectively filtering out the needed laser sheet correction.

Harmonic analysis

Harmonic analysis is a technique that can be used for estimating the amplitude and phase of components of a flow that are expected to behave in a periodic way, with a known period T. A signal f can be decomposed into its frequency components $k\omega \equiv k2\pi/T$, for $k = 0, 1, 2, \ldots$, as follows

$$f(t) = \sum_{j=0}^{\infty} A_j \cos(\omega t j) + B_j \sin(\omega t j). \tag{4.18}$$

We apply this decomposition to each spatial component of our field. In our case the fields are available at discrete times ,$t_k = k/s$, for $k = 0, 1, 2, \ldots N$, with s the sampling rate of the camera. Write $f(t_k) = f_k$, then the harmonic analysis for a component of the field can be represented in matrix-vector form

as

$$\begin{pmatrix} 1 & \cos(\omega/s) & \sin(\omega/s) & \cos(2\omega/s) & \sin(2\omega/s) & \dots \\ 1 & \cos(2\omega/s) & \sin(2\omega/s) & \cos(4\omega/s) & \sin(4\omega/s) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(N\omega/s) & \sin(N\omega/s) & \cos(2N\omega/s) & \sin(2N\omega/s) & \dots \end{pmatrix} \begin{pmatrix} A_0 \\ A_1 \\ B_1 \\ A_2 \\ B_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}.$$

We truncate the series and estimate the coefficients A_j and B_j using a least squares approach. Now note that the contribution of the term containing ω_j is given by

$$f_k = \mathrm{Re}(A_j - iB_j)e^{i\omega j t_k} \equiv a_k \mathrm{Re}(e^{(i\omega j - \psi_k)t_k}),$$

with, for $k > 0$,

$$a_j \quad = \quad \sqrt{A_j^2 + B_j^2}, \quad \psi_j \quad = \quad \tan^{-1}(\tfrac{B_j}{A_j}).$$

These quantities are the real-valued amplitude a_j and phase ψ_j corresponding to the j-th harmonic $j\omega$. Note that B_0 is undetermined since in (4.18) the term containing B_0 reads $B_0 \sin(\omega t_0 j)$ and $t_0 = 0$. However, A_0 is defined and gives information on the amplitude of the mean flow.

EOF analysis

The Empirical Orthogonal Function (EOF) analysis is most suitable for (i) compressing data, (ii) filtering out few dynamically significant components of a complicated system, (iii) filtering out unwanted noise from the data. The EOFs are orthogonal spatial patterns that can be seen as empirically derived basis functions. Frequently, the low-order EOFs are interpreted as the natural modes of variability of an observed system although some caution is called for since EOFs do not represent physical fields a priori. In the following we give a brief description of how EOFs can be found from observations (see e.g. Harlander, 1997; v. Storch & Zwiers, 1999). First, let p be the number of grid points and let q be the number of observations at different times. Enumerate the grid points in some arbitrary fashion to obtain a matrix of data values $\mathbf{F} \in \mathbb{R}^{p\times q}$, with $\mathbf{f}_{ik}$ equal to the value of gridpoint i at time k.

The ensemble mean is defined as

$$\bar{\mathbf{f}} = q^{-1}\mathbf{F}\mathbf{1}_q,$$

where $\mathbf{1}_q \in \mathbb{R}^q$ is the vector of ones. The deviation from the ensemble mean is defined as

$$\tilde{\mathbf{F}} = \mathbf{F} - \bar{\mathbf{f}}\mathbf{1}_q^T.$$

We now construct a covariance matrix by

$$\mathbf{C} = \tilde{\mathbf{F}}\tilde{\mathbf{F}}^*,$$

where the star indicates complex conjugate transpose. The eigenvectors $\mathbf{p}_i$ of the covariance matrix are the EOFs and they form an orthogonal basis. There is however a more efficient way to obtain the eigenvectors and eigenvalues. Let $\mathbf{FQ} = \mathbf{P\Sigma}$ be a singular value decomposition. It is easy to show that the left-singular vectors $\mathbf{p}_i$ in the columns of $\mathbf{P}$ are exactly the eigenvectors of $\mathbf{C}$. The eigenvalues λ_i of $\mathbf{C}$ are the square of the singular values σ_i of $\mathbf{F}$. This eliminates the need to form $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^*$ explicitly.

Note that, although called functions, the EOFs are actually vectors. Any discrete field at time k, which is the k-th column of F can be expanded in terms of the left singular vectors $\mathbf{p}_i$ by projection,

$$\mathbf{f}_k = \sum_{i=1}^{p} a_{ik}\mathbf{p}_i, \quad a_{ik} = \mathbf{p}_i^* \mathbf{f}_k, \tag{4.19}$$

where the a_{ik} are called 'the time dependent EOF coefficients' (sometimes also referred to as principal components). Note that the equal sign in (4.19) holds only if all p eigenvectors are considered. However, in practice the sum converges rapidly so that only a few eigenvectors will be taken into account, namely the ones corresponding to the largest eigenvalues. It can be shown that the eigenvalues λ_i of the covariance matrix are a measure for the amount of variance explained by the corresponding eigenvector $\mathbf{p}_i$.

The standard EOF analysis ignores the time evolution of the analyzed field. In contrast, Hilbert EOFs (sometimes also called complex EOFs) account for patterns that evolve in time. As a first step we perform an harmonic analysis on our signal $\mathbf{f}$. In appendix 4.B it was shown how to find the coefficient vectors $\mathbf{A}_k$ and $\mathbf{B}_k$, containing the coefficients for all elements of the field at frequencies $k\omega = 2k\pi/T$. Define

$$\hat{\mathbf{f}}_j(t) = \sum_{k=0}^{(q-1)/2} \mathbf{B}_k \cos(\omega k j) - \mathbf{A}_k \sin(\omega k j).$$

Hilbert EOFs are defined as an EOF analysis of the complexified field $\mathbf{F}^H$ given by

$$\mathbf{f}_j^H = \mathbf{f}_j + i\hat{\mathbf{f}}_j.$$

Real and imaginary part of the Hilbert EOFs show one complete cycle of the corresponding variance pattern: the real part shows the patterns at $t = 0$, the imaginary part shows it at $t = \pi/2$, the real part multiplied by -1 gives the

pattern at $t = \pi$, and the imaginary part multiplied by -1 shows the variance pattern at $t = 3\pi/2$. The period of the oscillating pattern can be found from the Hilbert EOF coefficient, which contains the information about the complex amplitude evolution in time.

Bibliography

ARNOLD, V. & KHESIN, B. 1999 *Topological Methods in Hydrodynamics*, *Applied Mathematical Sciences*, vol. 125. Springer Verlag.

BAI, Z., DEMMEL, J., DONGARRA, J., RUHE, A. & VAN DER VORST, H. 2000 *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. Philadelphia: SIAM.

BAINES, P. 1971 The reflection of internal/inertial waves from bumpy surfaces. *Journal of Fluid Mechanics* **46**, 273–291.

BAINES, P. 1973 The generation of internal tides by flat-bump topography. *Deep-Sea Research* **20**, 179–205.

BAINES, P. 1974 The generation of internal tides over steep continental slopes. *Phil. Trans. Roy. Soc. London* **277** (1263), 27–58.

BAINES, P. 1982 On internal tide generation methods. *Deep-Sea Research* **29** (3A), 307–338.

BARCILON, V. 1968 Axi-symmetric inertial oscillations of a rotating ring of fluid. *Mathematika* **15**.

BATCHELOR, G. 1967 *An Introduction to Fluid Dynamics*. Cambridge: Cambridge University Press.

BEARDSLEY, R. 1970 An experimental study of inertial waves in a closed cone. *Stud. Appl. Maths.* **49**, 187–196.

BOURGIN, D. & DUFFIN, R. 1939 The dirichlet problem for the vibrating string equation. *Bull. Am. Math. Soc.* **45**, 851–859.

BOURGIN, D. G. 1940 The dirichlet problem for the damped wave equation. *Duke Mathematical Journal* **7**, 97–120.

BRYAN, G. 1889 The waves on a rotating liquid spheroid of finite ellipticity. *Philosophical Transactions of the Royal Society of London. A* **180**, 187–219.

CARTAN, M. E. 1922 Sur les petites oscillations d'une masse fluide. *Bull. des Sciènces Math.* **46**, 317–369.

CHUANG, W. & WANG, D. 1981 Effects of density front on the generation and propagation of internal tides. *J. Phys. Oceanogr.* **11**, 1357–1374.

CIARLET, P. 2002 *The Finite Element Method for Elliptic Problems*. Philadephia: SIAM.

CRAIG, P. D. 1987*a* Numerical modelling of internal tides. *Numerical modelling: Aplications to marine systems* pp. 107–122.

CRAIG, P. D. 1987*b* Solutions for internal tidal generation over coastal topography. *Journal of Marine Research* **45**, 83–105.

CUSHMAN-ROISIN, B., TVERBERG, V. & PAVIA, E. G. 1989 Resonance of internal waves in flords: A finite difference model. *Journal of Marine Research* **47**, 547–567.

DHOOGE, A., GOVAERTS, W., KUZNETSOV, Y., MESTROM, W. & RIET, A. 2004 Cl matcont: A continuation toolbox in matlab.

DINTRANS, B., RIEUTORD, M. & VALDETTARO, L. 1999 Gravito-inertial waves in a rotating stratified sphere or spherical shell. *Journal of Fluid Mechanics* **398**, 271–297.

DRIJFHOUT, S. & MAAS, L. 2007 Impact of channel geometry and rotation on the trapping of internal tides. Submitted to J. of Phys. Oceanogr.; in press.

DUNKERTON, T., DELISI, D. & LELONG, M.-P. 1998 Alongslope current generation by obliquely incident internal gravity waves. *Geophys Res. Lett.* **25**, 3871–3874.

FOKIN, M. 1983 On the solvability of the dirichlet problem for the equation of the vibrating string. *Soviet Math. Dokl.* **28** (2), 455–459.

FRANKLIN, J. N. 1972 Axisymmetric inertial oscillations of a rotating fluid. *Journal of mathematical analysis and applications* **39**, 742–760.

FRIEDLANDER, S. & SIEGMANN, W. L. 1982*a* Internal waves in a contained rotating stratified fluid. *Journal of Fluid Mechanics* **114**, 123–156.

FRIEDLANDER, S. & SIEGMANN, W. L. 1982*b* Internal waves in a rotating stratified fluid in an arbitrary gravitational field. *Geophys. Astrophys. Fluid Dynamics* **19**, 267–291.

GARRETT, C., MACCREADY, P. & RHINES, P. 1993 Boundary mixing and arrested ekman layers: Rotating stratified flow near a sloping boundary. *Annu. Rev. Fluid Mech* **25**, 291–323.

GOLUB, G. & VAN LOAN, C. 1989 *Matrix Computations*, 2nd edn. Baltimore, MD: Johns Hopkins Press.

GREENSPAN, H. P. 1968 *The theory of rotating fluids.*. Cambridge University Press.

HADAMARD, J. 1923 *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press.

HANSEN, P. 1994*a* Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms* **6**, 1–35.

HANSEN, P. C. 1994*b* Regularization tools. *Numerical Algorithms* **6**, 1–34.

HANSEN, P. C. 2001 The L-curve and its use in the numerical treatment of inverse problems. In *Computational Inverse Problems in Electrocardiology*, *Advances in Computational Bioengineering* 5, pp. 119–142. Southampton: WIT Press.

HARLANDER, U. 1997 Flow climatology in the alpine region as simulated by a simple shallow water model. *Contr. Atmos. Phys.* **70**, 285–299.

HARLANDER, U. & MAAS, L. 2006 Two alternatives for solving hyperbolic boundary value problems of geophysical fluid dynamics. Accepted by Journal of Fluid Mechanics.

HARLANDER, U., SCHÖNFELD, H.-J. & METZ, W. 2000 Rossby waveguides in high-latitude shear flows with boundaries. *Journal of Geophysical Research* **105** (C7), 17,063–17,078.

HENDERSON, G. & ALDRIDGE, K. 1992 A finite-element method for inertial waves in a frustum. *Journal of Fluid Mechanics* **234**, 317–327.

HOLLERBACH, R. & KERSWELL, R. R. 1995 Oscillatory internal shear layers in rotating and precessing flows. *Journal of Fluid Mechanics* **298**, 327–339.

HOSEGOOD, P. & VAN HAREN, H. 2006 Sub-inertial modulation of semidiurnal currents over the continental slope in the faeroe-shetland channel. *Deep-Sea Res. I* **53**, 627–655.

HUTHNANCE, J. M. 1973 Tidal current asymmetries over the norfolk sandbanks. *Estuarine Coastal Mar. Sci.* **1**, 89–99.

JOHN, F. 1941 The dirichlet problem for a hyperbolic equation. *American Journal of Mathematics* **63** (1), 141–154.

KATO, T. 1995 *Perturbation Theory for Linear Operators*, 2nd edn. Heidelberg: Springer-Verlag.

KELVIN, L. 1880 Vibrations of a columnar vortex. *Phil. Mag* **10** (155).

KOPECZ, S. 2006 Fractal internal wave patterns in a tilted square. `http://www.nioz.nl/kopecz`.

LAM, F.-P. A. & MAAS, L. R. M. 2007 Internal wave focusing revisited; a reanalysis and new theoretical links. Submitted to Fluid Dyn. Res.

LEBLOND, P. H. & MYSAK, L. A. 1978 *Waves in the Ocean*. Elsevier.

LEDWELL, J., MONTGOMERY, E., POLZIN, K., LAURENT, L. S., SCHMIDT, R. & TOOLE, J. 2000 Evidence for enhanced mixing over rough topography in the abyssal ocean. *Nature* **403** (13).

LEHOUCQ, R., SORENSON, D. & YANG, C. 1998 *ARPACK Users Guides, Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Method.*. Philadelphia: SIAM.

LYASHENKO, A. A. 1993 On the Dirichlet problem for the nonlinear equation of the vibrating string. I. *J. Math. Kyoto Univ.* **33** (2), 543–570.

LYASHENKO, A. A. 1994 Mappings of domains connected with the dirichlet problem for the equation of the vibrating string. *J. Math. Kyoto Univ.* **34** (2), 329–352.

LYASHENKO, A. A. & SMILEY, M. W. 1995 The dirichlet problem for the semilinear vibrating string equation in a class of domains with corner points. *Journal of Mathematical Analysis and Applications* **189**, 872–896.

MAAS, L. R. M. 2001 Wave focusing and ensuing mean flow due to symmetry breaking in rotating fluids. *Journal of Fluid Mechanics* **437**, 13–28.

MAAS, L. R. M. 2005 Wave attractors: linear yet nonlinear. *International Journal of Bifurcation & Chaos* **15**, 2757–2782.

MAAS, L. R. M., BENIELLI, D., SOMMERIA, J. & LAM, F.-P. A. 1997 Observation of an internal wave attractor in a confined stably stratified fluid. *Nature* **388**, 557–561.

MAAS, L. R. M. & HAREN, J. J. M. V. 1987 Observations on the vertical structure of tidal and inertial currents in the central north sea. *Journal of Marine Research* **45**, 293–318.

MAAS, L. R. M. & LAM, F.-P. A. 1995 Geometric focusing of internal waves. *Journal of Fluid Mechanics* **300**, 1–41.

MAAS, L. R. M. & ZIMMERMAN, J. T. F. 1989*a* Tide-topography interactions in a stratified shelf sea i. basic equations for quasi-nonlinear internal tides. *Geophysical and Astrophysical Fluid Dynamics* **45**, 1–35.

MAAS, L. R. M. & ZIMMERMAN, J. T. F. 1989*b* Tide-topography interactions in a stratified shelf sea ii. bottom trapped internal tides and baroclinic residual currents. *Geophysical and Astrophysical Fluid Dynamics* **45**, 37–69.

MACCREADY, P. & RHINES, P. B. 1991 Buoyant inhibition of ekman transport on a slope and its effect on stratified spin-up. *J. Fluid Mech* **223**, 631–661.

MANDERS, A. & MAAS, L. 2003 Observations of inertial waves in a rectangular basin with one sloping boundary. *Journal of Fluid Mechanics* **493**, 59–88.

MANDERS, A. & MAAS, L. 2004 On the three-dimensional structure of the inertial wave field in a rectangular basin with one sloping boundary. *Fluid Dynamics Res.* **35**, 1–21.

DE MELO, W. & VAN STRIEN, S. 1994 *One-dimensional dynamics*. Springer-Verlag.

MIDDLETON, J. F. & RAMSDEN, D. 1996 Time-dependent boundary layers on the continental slope. *J. Geophys. Res.* **101**, 18061–18077.

PEDLOSKY, J. 1992 *Geophysical Fluid Dynamics*. Springer Verlag.

POINCARÉ, H. 1885 Sur l'equilibre d'une masse fluide animée d'un mouvement de rotation. *Acta Mathematica* **7**, 259–380.

PRANDLE, D. 1982 The vertical structure of tidal currents and other oscillatory flows. *Continental shelf research 1* **2**, 191–207.

PRINSENBERG, S., WILMOT, W. & JR., M. R. 1974 Generation and dissipation of coastal internal tides. *Deep-Sea Research* **21**, 263–281.

RALSTON, J. V. 1973 On stationary modes in inviscid rotating fluids. *Journal of Mathematical Analysis and Applications* **44**, 366–383.

RAMSDEN, D. 1995 Response of an oceanic bottom boundary layer on a slope to interior flow. part i: Time-independent interior flow. *J. Phys. Oceanogr.* **25**, 1672–1687.

RATTRAY JR., M. 1960 On the coastal generation of internal tides. *Tellus* **12**, 54–62.

RATTRAY JR., M., DWORSKI, J. G. & KOVALA, P. 1969 Generation of long internal waves at the continental slope. *Deep-Sea Research* **16**, 179–195.

READ, P., YAMAZAKI, Y., LEWIS, S., WILLIAMS, P., MIKI-YAMAZAKI, K., SOMMERIA, J., DIDELLE, H. & FINCHAM, A. 2004 Jupiter's and saturn's convectively driven banded jets in the laboratory. *Geophysical Research Letters* **31** (22).

RIEUTORD, M. 1991 Linear theory of rotating fluids using spherical harmonics part ii, time periodic flows. *Geophysical and Astrophysical Fluid Mechanics* **59**, 185–208.

RIEUTORD, M., GEORGEOT, B. & L. VALDETTO 2001 Inertial waves in a rotating spherical shell: attractors and asymptotic spectrum. *J. Fluid. Mech.* **435**, 103–144.

RIEUTORD, M. & VALDETTARO, L. 1997 Inertial waves in a rotating spherical shell. *J. Fluid Mech.* **341**, 77–99.

RIEUTORD, M., VALDETTARO, L. & GEORGEOT, B. 2002 Analysis of singular inertial modes in a spherical shell: the slender toroidal shell model. *J. Fluid. Mech.* **463**, 345–360.

SANDSTROM, H. 1976 On topographic generation and coupling of internal waves. *Geophysical Fluid Mechanics* **7**, 231–270.

SOLOMON, T., HOLLOWAY, W. & SWINNEY, H. 1993 Shear flow instabilities and rossby waves in barotropic flow in a rotating annulus. *Physics of Fluids* **5** (8).

SOMMERIA, J. 2007 Personal communication.

STAQUET, C. & SOMMERIA, J. 2002 Internal gravity waves: From instabilities to turbulence. *Annu. Rev. Fluid. Mech.* **34**, 559–593.

V. STORCH, H. & ZWIERS, F. W. 1999 *Statistical analysis in climate research*. Cambridge.

SWART, A. & LOGHIN, D. 2005b Viscous regularisation of the poincaré equation. *Paper in progress* .

SWART, A., SLEIJPEN, G., MAAS, L. & BRANDTS, J. 2005a Numerical solution of the two dimensional poincaré equation, with apppendices. Preprint 1320. Dep. Math., University Utrecht, Utrecht, the Netherlands.

SWART, A., SLEIJPEN, G., MAAS, L. & BRANDTS, J. 2007a Numerical solution of the two dimensional poincare equation. *Journal of Applied and Computational Mathematics* **200**, 317–341.

THORADE, H. 1928 Gezeitenuntersuchungen in der Deutschen Bucht der Nordsee. *Deutsche Seewarte* **46** (3).

THORPE, S. A. 1999 The generation of alongslope currents by breaking internal waves. *Journal of Physical Oceanography* **29**, 29–38.

THORPE, S. A. 2000 The effects of rotation on the nonlinear reflection of internal waves from a slope. *Journal of Physical Oceanography* **30**, 1901–1909.

TIKHONOV, A. & ARSENIN, V. 1977 *Solution of Ill-posed Problems*. Washington: Winston & Sons.

TOLSTOY, I. 1973 *Wave propagation*. McGraw-Hill.

TROWBRIDGE, J. H. & LENTZ, S. J. 1991 Asymmetric behavior of an oceanic boundary layer above a sloping bottom. *J. Phys. Oceanogr* **21**, 1171–1185.

TVERBERG, T., CUSHMAN-ROISIN, B. & SVENDSEN, H. 1991 Modeling of internal tides in fjords. *Journal of Marine Research* **49**, 635–658.

VERHULST, F. 2005 *Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics*, *Texts in Applied Mathematics*, vol. 50. Springer.

WITBAARD, R., DAAN, R., MULDER, M. & LAVALEYE, M. 2005 The mollusc fauna along a depth transect in the faroe shetland channel: Is there a relationship with internal waves? *Marine Biology Research 1* **3**, 186–201.

WOTHERSPOON, S. 1995 Internal tides and resonance. PhD thesis, University of Tasmania.

WOTHERSPOON, S. 1996 Finite differences and internal tides - representing the boundary. *Deep-Sea Research 1* **43** (6), 949–958.

WUNSCH, C. & FERRARI, R. 2004 Vertical mixing, energy and the general circulation of the oceans. *Annu. Rev. Fluid Mech* **36**, 281–314.

ZIKANOV, O. & SLINN, D. 2001 Along-slope current generation by obliquely incident internal waves. *J. Fluid Mech.* **445**, 235–261.

ZIMMERMAN, J. T. F. 1978 Topographic generation of residual circulation by oscillatory (tidal) currents. *Geophysical and Astrophysical Fluid Dynamics* **11**, 35–47.

Index

Samenvatting

Dit proefschrift beschrijft numerieke oplosmethoden voor het benaderd oplossen van de zogenaamde Poincaré vergelijking. Deze vergelijking beschrijft op een wiskundige manier het gedrag van *interne golven*.

Nu is een interne golf een type golf dat zich wat merkwaardig gedraagt, vergeleken met oppervlaktegolven zoals we die kunnen waarnemen aan het oppervlak van het water. We kunnen ons afvragen wat een golf eigenlijk is en dit lijkt lastig zonder het werkwoord 'golven' te gebruiken. Een goed hanteerbare definitie van een golf is: 'een manier om energie te verplaatsen, zonder dat het medium grote uitwijkingen vertoont'. Dit is waar voor golven op zee; een kurk zal (ongeveer) op dezelfde plaats blijven terwijl de golven er onderdoor lopen. De waterdeeltjes bewegen slechts op en neer, terwijl de golf lijkt te bewegen. Ook geluidsgolven gedragen zich op deze manier, wanneer we een gesprek met iemand voeren bewegen de luchtdeeltjes die de golven dragen zich slechts minimaal terwijl het geluid zich over lange afstanden kan verplaatsen.

Interne golven vormen ook een categorie van golven, zij bevinden zich niet tussen vloeistof en lucht zoals oppervlaktegolven, maar *in* het inwendige van een vloeistof. In bijvoorbeeld de oceanen bestaat een gestaag toenemende dichtheid van het water wanneer we richting bodem bewegen. Zulke dichtheidsverschillen kunnen golven dragen, zoals het dichtheidsverschil tussen water en lucht oppervlaktegolven mogelijk maakt. Een bijzondere eigenschap van interne golven is dat ze zich manifesteren in bundels met een bepaalde richting. Wanneer zo'n bundel een wand of bodem raakt weerkaatst de bundel en na meerdere reflecties kunnen zelfs complexe patronen van golfbundels optreden. Een bijzonder patroon is de golfaantrekker, een gesloten baan waar alle golven op den duur op belanden. Op en rond een golfaantrekker is de golfbeweging het hevigst, je kan zeggen dat de energie er wordt ingevangen.

Het eerste hoofdstuk van dit proefschrift beschijft de natuurkundige aspecten van interne golven, de voorwaarden waaronder ze optreden en hun gedrag, gevat in formules. De uitdaging is nu om oplossingen te vinden van het zo geformuleerde vraagstuk. Hoofdstuk twee beschrijft hoe we een nu-

merieke oplossing kunnen verkrijgen. Dat wil zeggen dat we een benaderde oplossing zoeken, door het probleem in een aantal deelproblemen te verdelen die vervolgens door de computer doorgerekend kunnen worden. Er zijn verschillende van zulke 'discretisatie methoden'. Wij gebruiken een variatie op de zogenaamde Eindige Elementen Methode, waarbij het domein waarop we willen rekenen wordt opgesplitst in een aantal deelgebieden (het 'rooster'). Maar voor het zover is, is het een goed idee om het probleem te analyseren.

Wiskundigen houden ervan om een gesteld probleem te vereenvoudigen totdat slechts de essentie overblijft, in dit geval kunnen we het hele probleem herleiden tot de zogenaamde Poincaré vergelijking. Het blijkt dat deze vergelijking uiteindelijk oneindig dunne golfstralen beschrijft, die weerkaatsen in een basin en zo diverse patronen kunnen volgen (zie bijvoorbeeld de linker kolom in Figuur 1.6). Blijkbaar hebben we in ieder geval het golfbundelgedrag goed gevangen in ons vereenvoudigde model.

Dit vereenvoudigde model verschaft ons verder inzicht over het gedrag van de oplossingen. De golven (lijnen in ons model) hebben een bepaalde hoek met de verticaal en reflecteren aan wanden. Het blijkt nu dat het uiteindelijk resulterende patroon heel sterk afhangt van de hoek van de golfstraal, en de precieze vorm van het basin. Wiskundigen hebben het in zo'n geval over een *slecht-gesteld probleem*; variaties in de oplossingen varieren niet continu met variaties in de parameters. Anders gezegd, een willekeurig kleine variatie in een eigenschap van de golf kan een willekeurig grote verandering in de oplossing teweeg brengen.

Dit heeft serieuze consequenties voor de numerieke oplosmethode, het blijkt dat de extreme gevoeligheid van het probleem ten opzichte van de omgeving direct impliceert dat numerieke oplossingen ook niet meer te vertrouwen zijn. Het discretiseren van het probleem met behulp van een nauwkeuriger rooster zorgt er al voor dat we totaal andere oplossingen vinden, terwijl je in zo'n geval juist zou willen dat er betere oplossingen gevonden worden. De remedie die we voorstellen is een *regularisatie* techniek. Grofweg gesproken zoeken we niet naar een oplossing onder bepaalde voorwaarden, maar we zoeken naar een bepaalde combinatie van oplossingen die zich bevinden bij naburige voorwaarden. Dit doen we op zodanige manier dat we een oplossing vinden die de energie minimaliseert. De energie is een maat voor de gladheid van de oplossing, en het blijkt dat we met behulp van deze regularisatie techniek inderdaad oplossingen vinden die er 'glad' uitzien en ook stabiel blijven wanneer we kleine veranderingen in de randvoorwaarden aanbrengen. Bijvoorbeeld Figuur 2.10 geeft duidelijk aan hoe regularisatie een acceptabele oplossing kan produceren.

Hoofdstuk drie beschrijft een alternatieve regularisatie techniek. In tegenstelling tot de mathematische aanpak in het vorige hoofdstuk gebruiken we

hier een fysische eigenschap van het water, de viscositeit. De term viscositeit staat voor de stroperigheid van een vloeistof. Tot nu toe beschouwden we een vloeistof met viscositeit nul, dat wil zeggen dat waterdeeltjes straffeloos langs elkaar kunnen schuiven. In werkelijkheid is een vloeistof altijd visceus, en een waterdeeltje zal bij beweging zijn naburige deeltjes meevoeren. Dit effect werkt ook regulariserend, het dwingt een bepaalde gladheid af op de golfpatronen. Het probleem met dit type regularisatie is dat het effect in zekere zin klein is, de grootheid die de sterkte van de viscositeit meet is in de orde van een duizendste kilogram per meter-seconde. Wanneer we ons vraagstuk benaderend oplossen, is dit kleine getal dan nog wel significant? Of, is voor het gediscretiseerde probleem de viscositeit effectief nul? We hebben deze vraag geprobeerd te beantwoorden door het uitvoeren van computer experimenten bij verschillende viscositeiten. Het blijkt inderdaad dat wanneer de viscositeit te laag wordt, dat het effect niet voldoende meer is. Een extra toegevoegde regularisatie methode is dan aan te raden. Verder bevat dit hoofdstuk schattingen van eigenschappen van de oplossingen door de viscositeit te beschouwen als een *perturbatie* ('storing') op een eenvoudiger probleem. De oplossingen van het verstoorde probleem zijn dan te beschouwen als kleine verstoringen van de oplossingen van het onverstoorde probleem, waar we al veel kennis over bezitten. Zo zijn uitspraken gedaan over de stabiliteit van de oplossing en er is aangetoond dat de uitdemping van de golven lineair afhangt van de sterkte van de viscositeit.

Het laatste hoofdstuk gaat in op een tot nu toe wat onderbelicht aspect, wat is het praktische nut van het bestuderen van de eigenschappen van interne golven? Deze golven bestaan in oceanen, maar zoals gezegd manifesteren ze zich in de vorm van sterk gelocaliseerde golfbundels. Hebben interne golven dan nog enig significant effect, of is de toegevoegde waarde van dit proefschrift puur theoretisch? Het blijkt dat interne golven praktisch nut hebben, omdat ze voor menging kunnen zorgen in vloeistoffen. Deze menging treedt op wanneer een golfstraal reflecteert aan de bodem van de oceaan of zee. Zo'n menging is interessant voor biologen, het zorgt ervoor dat nutriënten worden opgewoeld en kan daardoor zorgen voor voedselbronnen voor zeeleven. Een ander effect van menging is dat het diverse eigenschappen van een vloeistof herverdeelt. Menging egaliseert bijvoorbeeld dichtheid en energie van het water. Dit is van belang voor de grootschalige oceaancirculatie, die op zijn beurt weer bepalend is voor het klimaat op aarde. In de beschrijving van deze circulatie is een bepaalde hoeveelheid menging nodig, maar nog niet alle bijdragen aan de benodigde menging zijn verklaard. Het is zeer wel mogelijk dat reflecterende interne golven deze rol kunnen vervullen.

In het laboratorium hebben we metingen gedaan aan dichtheids-gelaagd water in een roterend basin. Rotatie is, naast dichtheidsverschillen, ook een

mechanisme dat interne golven mogelijk maakt. Aangezien de aarde roteert is het belangrijk dit aspect ook te modelleren in ons experiment. Na analyse van de verkregen data bleek dat inderdaad een zeer sterke mening had opgetreden. Het experiment was zo opgezet dat een golfaantrekker kon bestaan, ook deze was duidelijk zichtbaar in de experimenten. De kaft van dit proefschrift laat een dwarsdoorsnede van het basin zien, waar de (numeriek berekende) golfaantrekker zichtbaar is. Daar overheen is experimentele data geplot die mooi de vorm van de aantrekker volgt. Het valt op dat in de linker onderhoek van het basin een andere straal zichtbaar is, die niet in de numerieke oplossing te vinden is. Het blijkt dat de golven worden opgewekt in de hoeken van het basin.

De verklaring hiervoor moet gezocht worden in de *grenslaag*, een smalle laag water aan de wand waar de vloeistof zich aanpast aan de roterende wand. In dit hoofdstuk hebben we theoretisch aangetoond dat deze grenslaag zich in hoekpunten kan openen, waardoor een aandrijvend mechanisme voor de interne golven wordt aangeboden.

Hiermee is de uiteenzetting compleet: grenslaag erupties drijven interne golven aan, deze golven lopen naar een aantrekker en wanneer ze reflecteren mengen ze het water.

Dankwoord

Het schrijven van een proefschrift niet iets wat je alleen doet, het komt tot stand door de hulp en betrokkenheid van anderen. Ik wil hier graag degenen bedanken die het promotieonderzoek tot een aangename periode gemaakt hebben.

Allereerst wil ik mijn promotor Henk van der Vorst bedanken, die mij de mogelijkheid geeft gegeven om dit promotieonderzoek uit te voeren.

Bij de uitvoering van het onderzoek waren vooral Gerard Sleijpen en Leo Maas betrokken, met wie ik altijd met veel plezier heb samengewerkt. Die samenwerking begon al meer dan zes jaar geleden, toen ik naar Texel vertrok om daar aan mijn afstudeeronderzoek te werken.

Gerard heeft een feilloos talent om bij de kern van een probleem uit te komen, dwars door technische barieres heen. Dit is niet te onderschatten, vooral wanneer je even ook niet meer weet hoe verder te gaan. De afgelopen jaren ben ik ook regelmatig voor overleg met Leo teruggekeerd naar Texel. Leo, ik vond onze gesprekken altijd erg inspirerend, op de terugweg naar Utrecht had ik altijd genoeg stof tot nadenken om de treinreis door te komen.

Dan zijn er natuurlijk mijn collega's van het mathematisch instituut, die ik wil bedanken voor de gezelligheid en de vele leerzame discussies. Vooral de traditionele 'Bastaard' avonden waren een welkome manier om de week af te sluiten. In min of meer chronologische volgorde wil ik de volgende mensen bedanken: Menno, Mischa, Bob, Barara, Daan, Lennaert, Philipp, Theo, Jasper, Taoufik, Tammo-Jan en Arthur.

Verder wil ik hierbij de volgende personen bedanken voor hun vriendschap, steun en het geduldig aanhoren van mijn verhalen over promotieperikelen. Maartje, Sam, Elise, Remco, Arne, Sanne, Rianne, Jeroen, Peter, Rutger, Linda, Jop, Maaike, Joris, Laura, Matthijs, Willemijn, Michiel, Ciska, Eveline, Jorn en Tanya.

Als je bij een schermvereniging zit, dan kan je je onderzoek even afschudden door een floret ter hand te nemen. In het bijzonder wil ik Mascha, Ruud, Isabelle, Teije, Nathalie, en Willem bedanken voor de ontspannende scherm- en Heksenketel sessies.

Mijn ouders, Ad en Annie, en zussen Annet en Esther, hebben altijd vertrouwen in me gehad en hebben me gesteund op de weg die uiteindelijk naar dit proefschrift heeft geleid. Bedankt!

Arno Swart 28 juni 2007

Curriculum Vitae

Arno Swart was born on September 30, 1977, in Woudrichem, The Netherlands and grew up in Genderen. In June 1995 he obtained his VWO diploma at the Willem van Oranje College, Waalwijk.

Between 1995 and 2001 he studied Computational Science at Utrecht University and graduated in April 2001 under the supervision of dr. Gerard L.G. Sleijpen and dr. Leo R.M. Maas with the masters thesis "A Finite Element Method for Internal Gravity Waves". This masters thesis was largely prepared during a half year stay at NIOZ, the Royal Netherlands Institute for Sea Research, located on the island Texel.

In August 2001 he started his PhD research, again under supervision of dr. Sleijpen and dr. Leo Maas, with promotor prof. dr. Henk van der Vorst. In 2002 he attended the Geophysical and Environmental Fluid Dynamics summer school at Department of Applied Mathematics and Theoretical Physics at the Univerisity of Cambridge. Also, he has visited CERFACS (European Centre for Research and Advanced Training in Scientific Computation) in Toulouse, France, for a collaboration with dr. Daniel Loghin on the application of a Finite Element Method. Later, he returned to France for experiments using the 'Coriolis turntable' at LEGI (Laboratoire Des Ecoulements Géophysiques et Industriels) in Grenoble. The experiments were performed in collaboration with dr. Leo Maas and dr. Uwe Harlander. During the four years of his research, the results were presented at various occasions, most notably in May 2005 at the SciCade conference in Nagoya, Japan.

From August 2005 to August 2006 he was attached to the Mathematical Institute of the University of Utrecht as a teacher, being responsible for both bachelor and master courses. During this period he has re-implemented the masters course 'Scientific Computing Laboratory Class'.

The research performed between 2001 and 2006 has finally resulted in this thesis.

Colophon

This thesis was was writting using LaTeX, and set using the *Math Pazo* font. The cover illustration was designed using the desktop publishing software Scribus. The thesis was printed by Ipskamp PrintPartners, Enschede.

This research was supported by the Netherlands Organisation for Scientific Research (NWO) under project number 613.000.107, titled *Numerical Determination of Wave Attractors*. The research conducted for Chapter three of this thesis was supported by the NWO *van Gogh* fund under project number VGP 61-534.