Universiteit Utrecht

Department
of Mathematics

Numerical solution of the two dimensional Poincaré equation

by

Arno Swart, Gerard L.G. Sleijpen, Leo R.M. Maas, and Jan Brandts

Preprint

nr. 1320

Februari 28, 2005

Numerical solution of the two dimensional Poincaré equation*

Arno Swart[†], and Gerard L.G. Sleijpen[†], and Leo R.M. Maas[‡], and Jan Brandts[§]

Februari 28, 2005

Abstract

This paper deals with numerical approximation of the two dimensional Poincaré equation that arises as a model for internal wave motion in enclosed containers. Inspired by the hyperbolicity of the equation we propose a discretisation particularly suited for this problem, which results in matrices whose size varies linearly with the number of grid points along the coordinate axes. Exact solutions are obtained, defined on a perturbed boundary. Furthermore, the problem is seen to be ill-posed and there is need for a regularisation scheme, which we base on a minimal-energy approach.

Key words. Poincaré equation; Ill-posed problems; regularisation; internal waves.

AMS Subject Classification 65N22, 65F22, 67B55.

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)),

$$\begin{array}{rcl} \frac{\partial^2\tilde{\Psi}(x,z)}{\partial x^2} - \lambda^2 \frac{\partial^2\tilde{\Psi}(x,z)}{\partial z^2} & = & 0 \quad \text{in } \tilde{\Omega} \\ \tilde{\Psi}(x,z) & = & 0 \quad \text{at } \partial\tilde{\Omega}. \end{array} \tag{1}$$

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

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

*2nd March 2005

[†]Departement of Mathematics, Utrecht University, P.O. Box 80.010, NL-3508 TA Utrecht, The Netherlands. E-mail: swart@math.uu.nl, sleijpen@math.uu.nl.

[‡]maas@nioz.nl

[§]brandts@science.uva.nl

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 recognized Poincaré as the author who firstly described the equation in Poincaré (1885). Although (1) is a reduction of the original three dimensional problem, we will refer to it as the Poincaré equation. Some authors prefer the term 'wave equation', but we like to stress the absence of a time 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 (1) in sections 2 and 3. In section 4 we develop an efficient discretisation scheme. We continue in section 5 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 6), to obtain meaningful solutions. We conclude by presenting some results of the numerical approximation and regularisation, applied to a representative model geometry.

Previous work on numerical determination of internal waves include Chuang & Wang (1981) and Henderson & Aldridge (1992) who use finite differences respectively finite element methods. Later attempts use the characteristics of the differential equation, see for example Craig (1987*a*), Craig (1987*b*) or Cushman-Roisin *et al.* (1989). 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.

Our approach is different, we allow a large class of polynomial domains. Furthermore, we minimise an energy, which is also 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), 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, 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.

This preprint is an extended version of Swart *et al.* (2005a), this version includes appendices that were not in the original paper.

2 The Poincaré equation

In this section we will analyze 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 (1) by recognising the hyperbolic nature and introducing characteristic coordinates (ξ, η) defined by

$$(\xi(x,z), \eta(x,z)) = \Xi(x,z) = (x - \lambda z, x + \lambda z).$$

This transforms the domain into $\Xi\tilde{\Omega} = \Omega$ and it transforms (1) to

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

The solvability of (1) is different from (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) are also solutions to (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) 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,

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

where L is the total Euclidean length of $\partial\Omega$. We now suppose the following

- 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. *We call the four points* $\bar{\Omega} \cap \bar{R}$ *the* extreme vertices.

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

Theorem 2 (Unique extension of the solution). *Every solution to the Poincaré equation on* $\bar{\Omega} = \Omega \cup \partial\Omega$ *is uniquely extendible to* $\bar{R} = [\xi^-, \xi^+] \times [\eta^-, \eta^+]$.

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

$$\nabla\Psi = 0 \quad \text{at a corner point.} \tag{3}$$

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

$$\|\Psi\|^2_{L_2(\Omega)} = \int_\Omega \Psi^2(\xi, \eta)\, d\xi\, d\eta = < \mathcal{F} + \mathcal{G}, \mathcal{F} + \mathcal{G} >_{L_2(\Omega)} = 1. \tag{4}$$

This constraint also rules out 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

$$< \mathcal{F}, 1 >_{L_2([\xi^1, \xi^+])} = 0. \tag{5}$$

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

$$\begin{array}{lcl} \mathcal{F}(\xi) + \mathcal{G}(\eta) & = & 0 \quad \text{at } \partial\Omega, \\ \mathcal{F}'(\xi) = \mathcal{G}'(\eta) & = & 0 \quad \text{at a corner point,} \\ < \mathcal{F}, 1 >_{L_2(\Omega)} & = & 0, \\ < \mathcal{F} + \mathcal{G}, \mathcal{F} + \mathcal{G} >_{L_2(\Omega)} & = & 1, \end{array} \tag{6}$$

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 (6) induces a map from the boundary to itself (John (1941), 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 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

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

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

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

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 has a finite number of attracting fixed points. We call solutions in this case* wave attractors.

C. *$\alpha(F) \in \mathbb{R} \setminus \mathbb{Q}$, 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. The solvability of the problem (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 that such that 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 $(P, Q]_{\partial\Omega} \equiv (P, Q)_{\partial\Omega} \cup Q$ 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 $O(P_1) \cap (P_0, P^*)_{\partial\Omega} = \emptyset$,*
- *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 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 a 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), together with this additional boundary condition, is different for the tree 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) in $L_2(\Omega)$ is dependent on the cases, and the value of n, from Theorem 3 as follows*

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.*

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

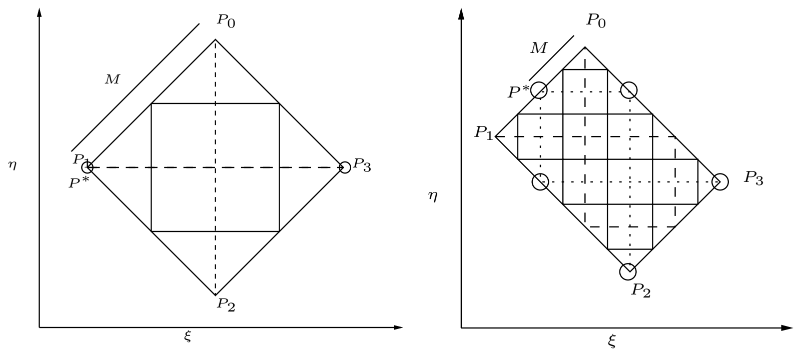

Figure 1. This figure clarifies the fundamental interval as introduced in Definition 3. At the right is shown a geometry with $n = 5$, which is easiest to see by following the closed orbit (solid line). Of course one may also consider the dashed and dotted lines that connect $O(P_1) = O(P_0)$, respectively $O(P_2) = O(P_3)$. 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 leftmost picture is a case where $n = 2$. In both pictures the set $M = [P_0, P^*)$ is the fundamental interval.

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). The conclusions do have to be modified, for example, solutions are possible in the case (C) if there exists a $C(\alpha)$ for any rational number m/n such that $|\alpha - m/n| \geq C(\alpha)/n^2$ where α is the rotation number (Fokin (1983)).

3 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

$$\dot{\Psi}(\xi, \eta) = \mathcal{F}'(\xi)\dot{\xi} + \mathcal{G}'(\eta)\dot{\eta}, \tag{7}$$

At curves $\{(\xi(t), \eta(t))|\dot{\Psi} = 0\}$ 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 (7) that this is the case when

$$\begin{pmatrix} \dot{\eta} \\ \dot{\xi} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -\mathcal{F}'(\xi) \\ \mathcal{G}'(\eta) \end{pmatrix} \tag{8}$$

The factor 1/2 is for future convenience. This system is in Hamiltonian form. We proceed by linearising Ψ around $(\xi, \eta) = (0, 0)$, giving

$$\Psi(\xi, \eta) = \mathcal{F}(0) + \mathcal{G}(0) + \xi\mathcal{F}'(0) + \eta\mathcal{G}'(0) + \frac{\xi^2\mathcal{F}''(0)}{2} + \frac{\eta^2\mathcal{G}''(0)}{2} + \mathcal{O}(\xi^3) + \mathcal{O}(\eta^3),$$

around the origin. From this expression we find

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

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 (8) 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

$$\begin{aligned}&\xi = \mathcal{G}'(0)t + t\mathcal{O}(\eta^2) + c_1,\\ &\eta = \mathcal{F}'(0)t + t\mathcal{O}(\xi^2) + c_1,\end{aligned}$$

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

$$\mathcal{F}'(0) = \mathcal{G}'(0) = 0 \text{ and } \mathcal{F}''(0) = 0 \text{ or } \mathcal{G}''(0) = 0, \tag{9}$$

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 hyperbole close to the critical point where $\nabla\Psi = 0$.
- 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 the ξ and η coordinate axes leads to the next theorems.

Lemma 1. *If we are not in the case (9), 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 (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 (9) 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 is Q with $l(T^-Q)e_2 = 0$ or $l(T^+Q)e_1 = 0$.*

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 5. The two zero level set curves of the saddles that point into the circle must be connected, there is no alternative. Connecting each saddle to itself would induce a center point in the region bounded by the homoclinic orbit. This in turn would induce more saddle points on the boundary. 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)).

4 Discretisation

In this section we will develop a discretisation for equations (6). The following section may easily be extended to allow for nonzero Dirichlet data $h(\xi, \eta)$ on the boundary, or boundary conditions on a subset $\Gamma \subset \partial\Omega$ only. The latter allows for, i.e., ocean shelf simulations.
A 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 basis functions $\phi_i(\xi)$ respectively $\psi_j(\eta)$. In this basis we can write

$$\begin{array}{rcll} \tilde{\mathcal{F}}(\xi) & = & \sum_{i=1}^n f_i\phi_i(\xi) & \in V_\xi, \\ \tilde{\mathcal{G}}(\eta) & = & \sum_{j=1}^m g_j\psi_j(\eta) & \in V_\eta. \end{array} \tag{10}$$

Write $f = (f_1,\ldots,f_n)$ and $g = (g_1,\ldots,g_m)$. We now define our approximation $\tilde{\Psi}$ to Ψ by

$$\tilde{\Psi}(\xi,\eta) = \tilde{\mathcal{F}}(\xi) + \tilde{\mathcal{G}}(\eta). \tag{11}$$

We propose a Galerkin orthogonality condition of the system (6) and try to find $H = \tilde{\mathcal{F}} + \tilde{\mathcal{G}}$ on the boundary $\partial\Omega$ for which all test functions $v_1 \in V_\xi$ and $v_2 \in V_\eta$ are orthogonal to H. First define on $\partial\Omega$ the line integral

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

The orthogonality relation becomes

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

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

$$\begin{array}{rcll} (\tilde{\mathcal{F}},v_1) + (\tilde{\mathcal{G}},v_1) & = & 0 & \forall v_1 \in V_\xi, \\ (\tilde{\mathcal{F}},v_2) + (\tilde{\mathcal{G}},v_2) & = & 0 & \forall v_2 \in V_\eta. \end{array} \tag{12}$$

Write the functions $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{G}}$ as a linear combination of basis functions (10). 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

$$\begin{array}{rcll} \sum_{i=1}^n f_i(\phi_i,\phi_j) + \sum_{i=1}^m g_i(\psi_i,\phi_j) & = & 0, & j = 1,\ldots,n \\ \sum_{i=1}^n f_i(\phi_i,\psi_j) + \sum_{i=1}^m g_i(\psi_i,\psi_j) & = & 0, & j = 1,\ldots,m. \end{array} \tag{13}$$

After choosing suitable spaces V_ξ and V_η we can work out the integrations in (13). This results in a matrix vector equation which we need to solve for f and g:

$$\begin{pmatrix} A_1 & A_2 \\ A_2^T & A_3 \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} = A \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{14}$$

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 (6). For a resolution of $m \times n$ 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 for an example.

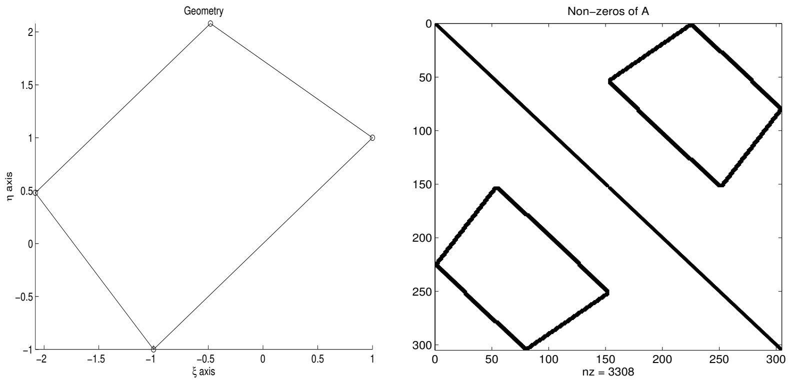

FIGURE 2. This figure shows a sample geometry in the left panel. To the right is a plot of the non zeros of the matrix A. The geometry stands out clearly in the blocks A_2 and A_2^T (see (14)). The sub-matrices A_1 and A_3 are tri-diagonal since standard piecewise linear basis functions were used.

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

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

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 it is fixed at some value, therefore we fix the norm using

$$\|x\|_2^2 = 1. \tag{15}$$

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,

$$Px = 0. \tag{16}$$

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.
It is of 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.

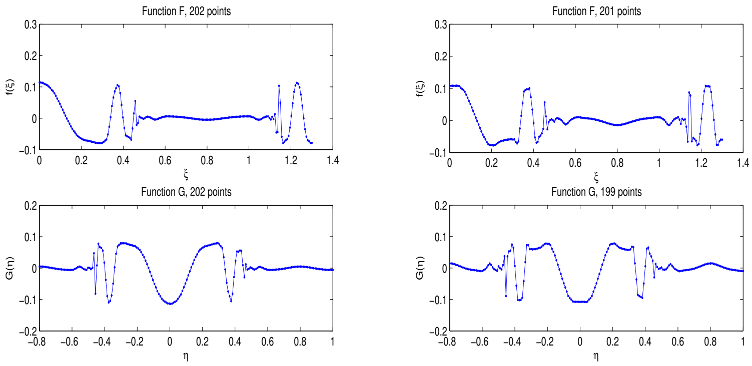

FIGURE 3. This figure illustrates that a small change in the discretisation can have substantial effects on the solution. Plotted are the approximations to $\mathcal{F}$ and $\mathcal{G}$. The panel to the left has 200 grid points in both the ξ and η direction, the rightmost panel has one grid point less in both direction. The two solutions are quite different.

5 Ill-Posedness and uniqueness

After discretisation (see Section 4) 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 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. 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.

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 5.1. 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 whenever 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 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

$$\bar{T}\bar{\Psi} = \int_{\bar{\Omega}} \nabla\tilde{\Psi} \cdot \nabla\tilde{\Psi}\, dx\, dy, \tag{17}$$

which is, in characteristic coordinates

$$\begin{aligned} T\Psi &= \int_\Omega \left[(\frac{\partial}{\partial\eta} + \frac{\partial}{\partial\xi})^2 + \lambda^2 (\frac{\partial}{\partial\eta} - \frac{\partial}{\partial\xi})^2 \right] \Psi^2 \frac{1}{2\lambda}\, d\xi d\eta \\ &= \frac{1}{2\lambda} \int_\Omega (1+\lambda^2) \left[(\mathcal{F}')^2 + (\mathcal{G}')^2 \right] + (1-\lambda^2)\mathcal{F}'\mathcal{G}'\, d\xi\, d\eta. \end{aligned} \tag{18}$$

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

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

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

$$T\Psi = \frac{(1+\lambda^2)}{2\lambda} \int_\Omega (\mathcal{F}'(\xi))^2 + (\mathcal{G}'(\eta))^2\, d\xi\, d\eta. \tag{19}$$

Note that the energy is zero if the solution Ψ is constant. These solutions are excluded by (4), which also makes sure that the stream function can not be scaled to get arbitrarily small energy. We also incorporate the corner constraint (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 4 this will lead to,

$$
\begin{array}{ll}
\text{minimise } \|Lx\|_2^2 & \text{energy minimisation,} \\
\text{minimise } \|Ax\|_2^2 & \text{function minimisation at the boundary,} \\
\text{subject to } Cx = 0 & \text{linear 'unique representation' constraint,} \\
\text{and } \|x\|_2 = 1 & \text{normalisation,} \\
\text{and } Px = 0 & \text{linear 'corner point' constraint.}
\end{array}
\tag{20}
$$

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 possesses high frequency components (see Hansen (1994) 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

$$x_\tau = \operatorname{argmin}(\|Ax\|_2^2 + \tau^2\|Lx\|_2^2), \text{ with } \|x\|_2^2 = 1, Cx = Px = 0, \tag{21}$$

The L-curve approach requires the solution of (21) 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,

$$x_\tau = \operatorname{argmin}(\| \begin{pmatrix} A \\ aP \\ aC \end{pmatrix} x\|_2^2 + \tau^2\|Lx\|_2^2). \tag{22}$$

In this way the constraints will certainly be satisfied, or the residual will become very large. Solving the minimisation problem (22) is then equivalent to finding the normalised singular vector corresponding to the smallest singular value of

$$B = \begin{pmatrix} A \\ aP \\ aC \\ \tau L \end{pmatrix} \tag{23}$$

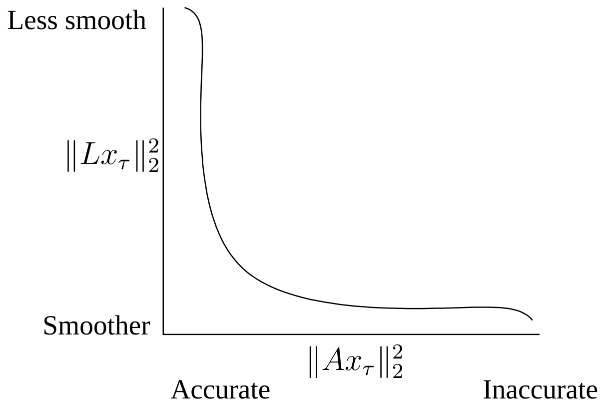

Figure 4. This figure shows a schematic L-curve parametrised by τ. On the horizontal axis is the residual, representing the accuracy of the solution. On the vertical axis is the energy, representing the smoothness of the solution.

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

5.1 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 (22). The typical shape is that of an 'L', as shown in Figure 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 7) 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.

6 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.

6.1 An 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 may be reduced to

$$\begin{aligned} \frac{f_{xx}(x)}{f(x)} &= \lambda^2 \frac{g_{zz}(z)}{g(z)}, \text{ in } \Omega, && (24) \\ g(0) = g(1) = f(0) = f(1) &= 0, && (25) \\ \int_0^1 f^2(x)\, dx \int_0^1 g^2(z)\, dz &= 1. && (26) \end{aligned}$$

We would like to find the solution that minimises the energy

$$\begin{aligned} E &= \int_\Omega \nabla\Psi \cdot \nabla\Psi \, dxdy \\ &= \int_0^1 (f_x)^2\, dx \int_0^1 g^2\, dy + \int_0^1 f^2\, dx \int_0^1 (g_y)^2\, dy. && (27) \end{aligned}$$

Use the normalisation constraint (26) 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 (24) equals γ^2. The simplified model now reads

$$\begin{aligned} & f_{xx}(x) = \gamma^2 f(x), \\ & \|f(x)\|^2_{L_2(\Omega)} = \delta, && (28) \\ & f(0) = f(1) = 0. \end{aligned}$$

We need to find the solution that minimises

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

The differential equation plus boundary conditions (28) is solved by

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

with $\lambda_k = k\pi$. By (26) 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

$$f_{xx}(x) = -\frac{\mu}{d} f(x). \tag{29}$$

This equation is of the same form as our differential equation (28), 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 (29) has solution

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

From boundary conditions (25) we see that α must be zero and μ is quantised as $\mu_k = dk^2\pi^2$ 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

$$f(x) = \beta_k \sin(k\pi x). \tag{30}$$

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.

6.2 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:

$$\Psi(x) = \sum_{i=0}^{\infty} f_i(x), \tag{31}$$

$$f_i(x) = \begin{cases} f(\frac{x-\alpha_i}{\alpha_{i+1}-\alpha_i}), & \text{for } x \in [\alpha_i, \alpha_{i+1}], \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

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. 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

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

Here we used the disjoint supports of f_i in order to write the sum of squares as a square of sums. There are two ways in which the energy may become infinite. Firstly the derivative of the function f itself may not be square integrable. We will see later that such functions do not minimise the energy. Secondly the sum in the second factor diverges, but 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 suggest 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.

6.3 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

$$\begin{aligned}\mathcal{F}_n &= \quad \sum_{i=0}^{N}\sum_{k=0}^{K}\sum_{l=0}^{J} A_n^{l,k} I_{\xi\in[\alpha_i^{l,k},\alpha_{i+1}^{l,k}]} \cos(n\pi\frac{\xi-\alpha_i^{l,k}}{\Delta\alpha_i^{l,k}}),\\ \mathcal{G}_m &= -\sum_{i=0}^{M}\sum_{k=0}^{K}\sum_{l=0}^{J} A_m^{l,k} I_{\eta\in[\beta_i^{l,k},\beta_{i+1}^{l,k}]} \cos(m\pi\frac{\eta-\beta_i^{l,k}}{\Delta\beta_i^{l,k}}),\end{aligned} \tag{33}$$

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 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 (33), 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 (33) are automatically continuous.

6.4 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 propose 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

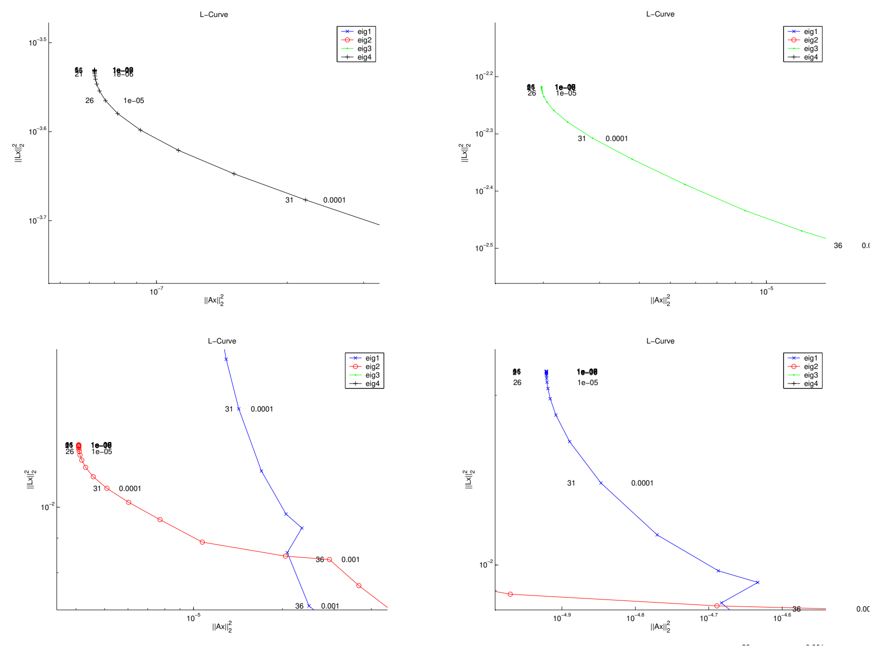

FIGURE 5. These four panels show the L-curves corresponding to the best four solutions. The indices of the regularisation parameters are (in clockwise order): 26,31,31,31, corresponding to $\tau \approx 10^{-5}, 10^{-4}, 10^{-4}, 10^{-4}$.

values of M and N, if $M >> 1$ and $N >> 1$. 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.

7 Results

This section will give some solutions of the Poincaré equation using the discretisation from section 4 and the regularisation procedure outlined in section 5. We will discuss the trapezoid, a rectangle with one sloping side as shown in (x, z) coordinates in Figure 10. This geometry is interesting since it is a simple geometry in 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 & Lam (1995). For $\lambda = 0.75$ we are in the attractor case, indeed figure 10 clearly shows features that reproduce in increasingly smaller scales towards a limit cycle.
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 (22) best. This simply amounts to calculating the four smallest singular values of the matrix (23). The curves are approximated by taking a sequence of values for τ, we choose $\{10^j | j = -10, -9.8, -9.6, \ldots, -0.8, -1\}$. Figure 5 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 11, first panel of Figure 5, representing the minimising solution, we choose a value of $\tau \approx 10^{-5}$. We plot the functions $\mathcal{F}$ and $\mathcal{G}$ and a plot of their sum in Figure 6. 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 Gibb's 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 7 and 8. We conclude that the structure is 'mixed in' from another solution, the regularisation is not perfect. The question of the validity is more complicated. The solution is of the form (33), 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 6 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.

Wavenumber 1 solution

The next-best solution we find is shown in Figure 7, where we plot the solution corresponding to $\tau \approx 10^{-4}$ on the second curve. It clearly consists of one empty fundamental interval plus one fundamental interval with half a cosine. Referring to (33) 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 9. 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 11 shows the curves corresponding to the four best minimizing solutions. The top two curves in 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 a 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 curve thusly obtained does not necessarily have to follow the L-curve of one solution. The rightmost panel of Figure 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 10),

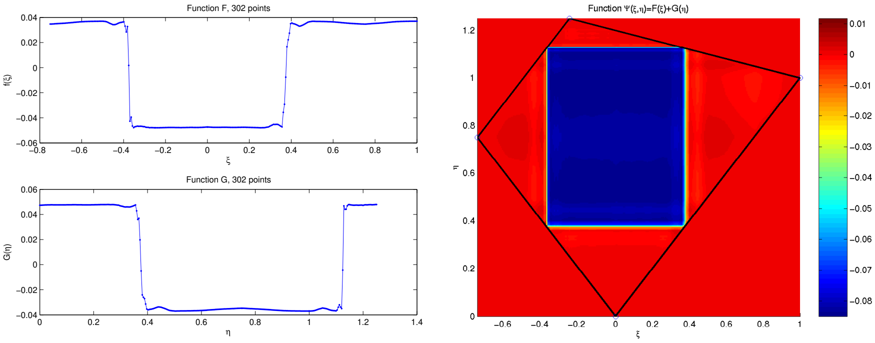

FIGURE 6. This figure shows the functions $\mathcal{F}$ and $\mathcal{G}$ and a plot of $\Psi = \mathcal{F} + \mathcal{G}$, in the (ξ, η) coordinate frame. Darker colors indicate higher values, gray is at the zero level.

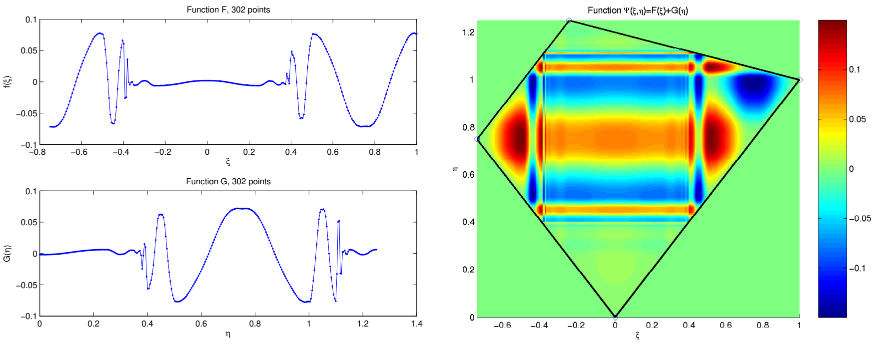

FIGURE 7. This figure shows the functions $\mathcal{F}$ and $\mathcal{G}$ and a plot of $\Psi = \mathcal{F} + \mathcal{G}$, in the (ξ, η) coordinate frame. One fundamental interval shows a smooth function (half the period of a cosine), the other one is zero.

then we fill both fundamental intervals and recover the solution that was presented in Maas & Lam (1995).

Wavenumber 2 solution

The solution we find, at $\tau \approx 10^{-4}$ 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 8). 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 we find the solutions predicted by equation (33), 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.

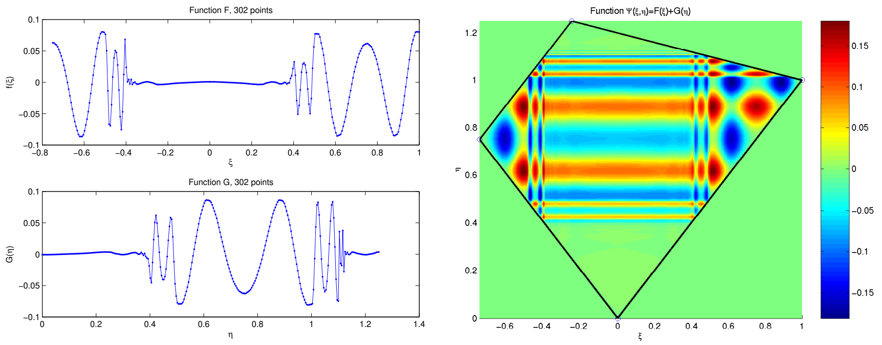

Figure 8. In this figure, one fundamental interval shows a smooth function (one period of a cosine, the other one is (almost) zero.

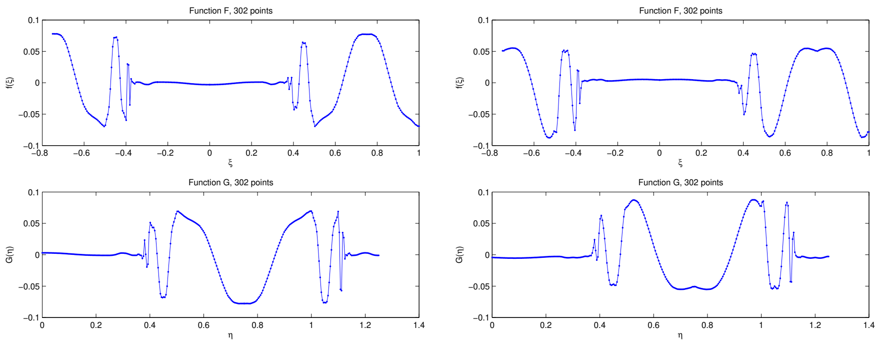

Figure 9. These graphs show what happens if one regularises improperly. In the leftmost panel a value of $\tau \approx 3 \times 10^{-3}$ was chosen, yielding a smooth but inaccurate solution. The extrema are too flat, and there is not a proper cosine profile. The rightmost panel corresponds to $\tau \approx 10^{-5}$, the residual is smaller but the energy is higher. We see cusps appearing, giving a discontinuity in the derivative.

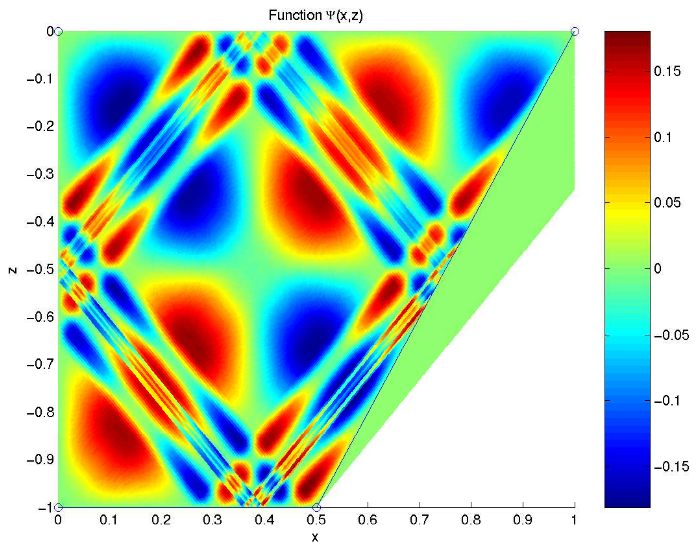

Figure 10. This figure shows the addition of the smoothest solutions found on each of the fundamental intervals. The grid resolution is 300 grid points in the ξ and η directions. Shown is the stream function $\tilde{\Psi}$, in the (x, z) frame. Darker colors indicate higher values, gray is at the zero level. The attractor, and the fractal structure are nicely visible.

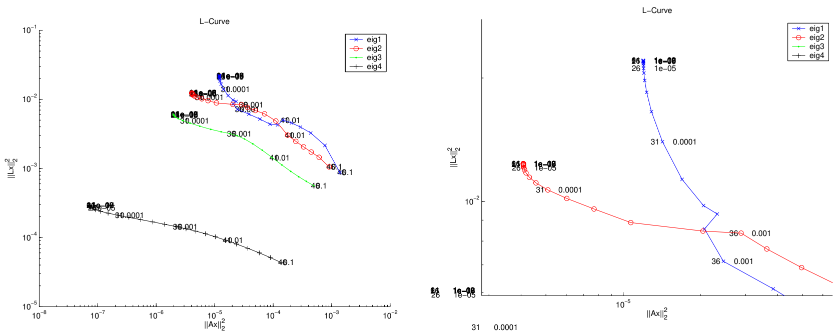

FIGURE 11. Four L-curves for the trapezoid. The 'elbows' of the curves do not look very pronounced when plotted in one picture. One would have expected a longer vertical 'leg', instead we find an accumulation point. A number of points on the curves have been calculated and labeled with the value of the regularisation parameter τ and an index. To the right is an enlargement of the crossover of the top two curves

Appendix 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

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

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.*

$$\begin{array}{lccc} \textit{Remove column 1,} & A[0, I_{m-1}]^T & \equiv & AI'_l, \\ \textit{Remove column n,} & A[I_{m-1}, 0]^T & \equiv & AI'_r, \\ \textit{Remove row 1,} & [0, I_{n-1}]A & \equiv & I'_tA, \\ \textit{Remove row n,} & [I_{n-1}, 0]A & \equiv & I'_bA, \\ \textit{Append zeros left,} & A[0, I_m] & \equiv & AI_l, \\ \textit{Append zeros right,} & A[I_m, 0] & \equiv & AI_r, \\ \textit{Append zeros top,} & [0, I_n]^TA & \equiv & I_tA, \\ \textit{Append zeros bottom,} & [I_n, 0]^TA & \equiv & I_bA, \end{array}$$

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

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

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 1 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

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

After some reordering, we write down the expression for A_2

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

```
Initialise c ← 1
Find (i, j) for which (ξ_i, η_j) is start of segment
Initialise (ξ, η) ← (ξ_i, η_j)
Initialise F^ξ, F^η, D^ξ, D^η, L as empty matrices
While not end of line piece
    (F^ξ_ij, F^η_ij) ← (ξ, η)
    if c==1
        if s < Δη_j/(ξ_{i+1} − ξ)
            η ← η + s(ξ_{i+1} − ξ)
            ξ ← ξ_{i+1}
            c ← 2
        else
            ξ ← ξ + Δη_j/s
            η ← η_{j+1}
            c ← 1
    else
        if s < (η_{j+1} − η)/Δξ_i
            η ← η + sΔξ)
            ξ ← ξ_{i+1}
            c ← 2
        else
            ξ ← ξ + (η_{j+1} − η)/s
            η ← η_{j+1}
            c ← 1
    (D^ξ_ij, D^η_ij) ← (ξ, η) − (F^ξ_ij, F^η_ij)
    L_ij ← ||(D^ξ_ij, D^η_ij)||_2^2
```

ALGORITHM 1. 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 12) to consider, distinguished by $c = 1$ or $c = 2$. While traversing the segment the matrices containing the parametrisation are built.

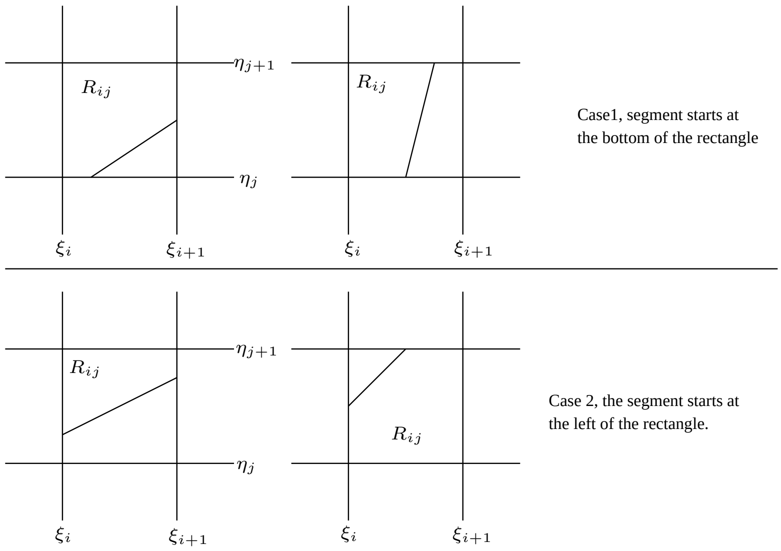

FIGURE 12. The four ways in which a segment with positive slope can intersect the rectangle R_{ij}

Next we turn to the matrix A_1 containing integration of the ϕ basis functions.

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

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

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

For A_3 we have by a similar calculation

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

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

Appendix 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

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

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 (19). 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

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

Also define the vectors

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

and write

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

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

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

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

$$\begin{aligned} S_i &= \ ([\alpha_i, \alpha_{i+1}] \times [\eta^+, \eta^-]) \cap \Omega, \\ P_j &= \ ([\xi^+, \xi^-] \times [\beta_j, \beta_{j+1}]) \cap \Omega. \end{aligned} \tag{34}$$

These area vectors, call h and v, are easily calculated using the matrices D_{ij}^ξ and D_{ij}^η, which we already have from Algorithm 1. Using these vectors we can write for the matrix L

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

Appendix C: Two dimensional energy minimisation

From (19) we have

$$T(\Omega) = A \int_{\Omega} (\mathcal{F}'(\xi))^2 + (\mathcal{G}'(\eta))^2 \, d\xi d\eta \tag{35}$$

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. 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 13 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

$$\begin{aligned} f_i(x) &= \begin{cases} f(\frac{x-\alpha_i}{\Delta\alpha_i}), & \text{if } i \text{ is even,} \\ f(\frac{\alpha_{i+i}-x}{\Delta\alpha_i}), & \text{if } i \text{ is odd,} \end{cases} \\ g_j(x) &= \begin{cases} g(\frac{x-\beta_j}{\Delta\beta_j}), & \text{if } j \text{ is even,} \\ g(\frac{\beta_{j+1}-x}{\Delta\beta_j}), & \text{if } j \text{ is odd,} \end{cases} \end{aligned} \tag{36}$$

[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.

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

$$\begin{aligned} \mathcal{F}(\xi) = & \sum_{k=0}^{K}\sum_{l=0}^{J}\sum_{i=0}^{N(k,l)} f_i^{k,l}(\xi), \\ \mathcal{G}(\eta) = & \sum_{k=0}^{K}\sum_{l=0}^{J}\sum_{j=0}^{M(k,l)} g_j^{k,l}(\eta), \end{aligned} \tag{37}$$

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 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

$$\begin{aligned} \tfrac{df_i(x)}{dx} &= \begin{cases} 0 & \text{if } x = \alpha_{i+1} \text{ or } x = \alpha_i, \\ \frac{1}{\Delta\alpha_i} f'(\frac{x-\alpha_i}{\Delta\alpha_i}) & \text{otherwise.} \end{cases} \\ \tfrac{dg_j(x)}{dx} &= \begin{cases} 0 & \text{if } x = \beta_{j+1} \text{ or } x = \beta_j, \\ \frac{1}{\Delta\beta_j} g'(\frac{x-\beta_i}{\Delta\beta_j}) & \text{otherwise.} \end{cases} \end{aligned} \tag{38}$$

Now plug in the series (37) in the energy (35) to get

$$T_{NM}(\Omega) = A(\sum_{i=0}^{N} < \frac{\partial}{\partial\xi} f_i, \frac{\partial}{\partial\xi} f_i >_{L_2(\Omega)} + \sum_{j=0}^{M} < \frac{\partial}{\partial\eta} g_j \frac{\partial}{\partial\eta} g_j >_{L_2(\Omega)}), \tag{39}$$

where the true energy is the limit $T(\Omega) = \lim_{N,M\to\infty} T_{NM}(\Omega)$. Now recall the definitions (34). 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

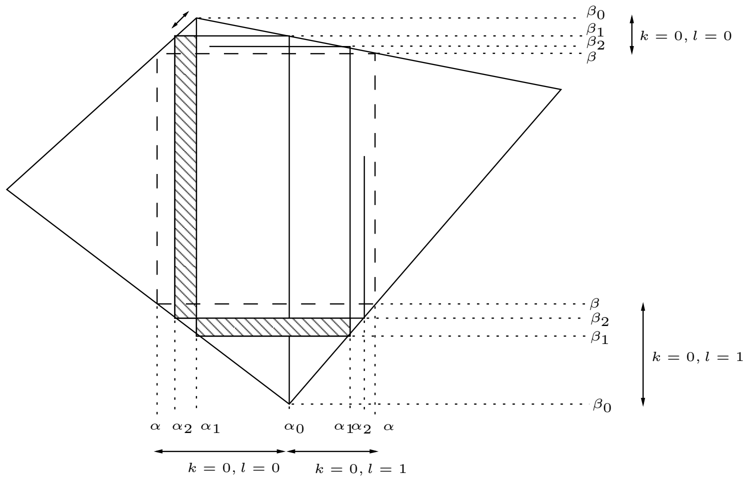

FIGURE 13. An example domain Ω. The endpoints of a fundamental interval, with (arbitrarily chosen) index $k = 0$, indicated by the thick double arrow, have been iterated a small number of times. In this case the fundamental interval induces two sequences in both the ξ and η direction, both with indices $l = 0$ 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 $S_1^{0,0}$.

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

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

Now we have the tools to proceed with the calculation of the terms in (39),

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

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

$$T_{NM}(\Omega) =< f(\xi), L_{1,N} f(\xi) >_{L_2([0,1])} + < g(\eta), L_{2,M} g(\eta) >_{L_2([0,1])} \tag{40}$$

with

$$L_{1,N} = \sum_{i=0}^{N} -\frac{\mathcal{H}_i'(\xi)}{\Delta\alpha_i}\frac{\partial}{\partial\xi} - \frac{\mathcal{H}_i}{\Delta\alpha_i}\frac{\partial^2}{\partial\xi^2}, \tag{41}$$

$$L_{2,M} = \sum_{j=0}^{M} -\frac{\mathcal{W}_j'(\eta)}{\Delta\beta_j}\frac{\partial}{\partial\eta} - \frac{\mathcal{W}_j}{\Delta\beta_j}\frac{\partial^2}{\partial\xi^2}, \tag{42}$$

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

$$L_{1,N}f = \lambda f. \tag{43}$$

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

$$f(x) = \sum_{k=0}^{\infty} B_k \cos(\pi kx), \tag{44}$$

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

$$g(x) = \sum_{k=0}^{\infty} C_k \cos(\pi kx). \tag{45}$$

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 (40) 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 (37) and rewrite using (36) we find

$$\begin{aligned} \Psi_{nm} &= \mathcal{F}_n + \mathcal{G}_m, \\ \mathcal{F}_n(\xi) &= \sum_{k=0}^{K}\sum_{l=0}^{J}\sum_{i=0}^{N(k,l)} B_n^{l,k} I_{\xi\in[\alpha_i^{l,k},\alpha_{i+1}^{l,k}]} \cos(n\pi\frac{\xi-\alpha_i^{l,k}}{\Delta\alpha_i^{l,k}}), \\ \mathcal{G}_m(\eta) &= \sum_{k=0}^{K}\sum_{l=0}^{J}\sum_{j=0}^{M(k,l)} C_m^{l,k} I_{\eta\in[\beta_j^{l,k},\beta j+1^{l,k}]} \cos(m\pi\frac{\eta-\beta_j^{l,k}}{\Delta\beta_j^{l,k}}). \end{aligned} \tag{46}$$

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 7.
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 (43) 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 (43) 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.

References

Barcilon, V. 1968 Axi-symmetric inertial oscillations of a rotating ring of fluid. *Mathematika* **15**.

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.

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 fiords: A finite difference model. *Journal of Marine Research* **47**, 547–567.

Fokin, M. 1983 On the solvability of the dirichlet problem for the equation of the vibrating string. *Soviet Math. Dokl.* **28** (2), 455–459.

Hansen, P. C. 1994 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.

Henderson, G. & Aldridge, K. 1992 A finite-element method for inertial waves in a frustum. *Journal of Fluid Mechanics* **234**, 317–327.

John, F. 1941 The dirichlet problem for a hyperbolic equation. *American Journal of Mathematics* **63** (1), 141–154.

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. & Lam, F. P. A. 1995 Geometric focusing of internal waves. *Journal of Fluid Mechanics* **300**, 1–41.

de Melo, W. & van Strien, S. 1994 *One-dimensional dynamics*. Springer-Verlag.

Poincaré, H. 1885 Sur l'equilibre d'une masse fluide animée d'un mouvement de rotation. *Acta Mathematica* **7**, 259–380.

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. & Georgeot, B. 2002 Analysis of singular inertial modes in a spherical shell: the slender toroidal shell model. *J. Fluid. Mech.* **463**, 345–360.

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. *Submitted* .