Universitat Autònoma de Barcelona

Numerical Computation of Invariant Objects with Wavelets

David Romero i Sànchez

UNIVERSITAT AUTÒNOMA DE BARCELONA

Departament de Matemàtiques

PhD THESIS

Numerical Computation of Invariant Objects with Wavelets

Thesis submitted by **David Romero i Sànchez** for the degree of Philosophæ Doctor by the Universitat Autònoma de Barcelona under the supervision of Prof. **Lluís Alsedà i Soler**.

Author & defended by:
David Romero i Sànchez.

Thesis Advisor:
Lluís Alsedà i Soler.

Certifico que la present memòria per aspirar al títol de Doctor en Matemàtiques ha estat elaborada per en David Romero i Sànchez, sota la direcció del Dr. Lluís Alsedà i Soler.

Signat: Dr. Lluís Alsedà i Soler a Bellaterra (Cerdanyola del Vallès), Setembre de 2015.

Agraïments

¿Saben aquel que *diu*? que es un tío que pasa por una ebanistería y ve un cartel que *posava*: falta oficial de primera. El payo entra y *diu*:
–Buenas, venía por el anuncio este.
–¿Usted es oficial de ebanistería? –*que li diu l'encarregat*.
–De toda la vida. – que le dice el tío.
–Le tendré que hacer una prueba.
Total que *l'encarregat* le saca un tronco *així de gran*.
–Sáqueme un San José de aquí.
Y el tío empieza a picar el tronco *i al cap de vuit hores* el tronco lo ha convertido en un palillo *així de prim*. Llega el jefe y *diu*:
–¿Qué, ha salido el San José ya?
–Todavía no, pero no se preocupe, que si esta aquí dentro este tío sale, oiga, este tío sale!!!

Ara és l'hora, diuen. Doncs sí, ara és el moment d'agrair a tots aquells que d'alguna manera o una altra heu contribuït en l'el.laboració d'aquesta Tesi. Amb tot, és difícil ser original i divertit per tal de fer un escrit que, com a totes les tesis, serà el que més cops es llegirà. Sigui n aquest número el qual com a mínim és ú. Ara bé, observi's que, ara mateix, n ha incrementat i per tant $n > 1$: gràcies! Havent fet la gracieta, jo i la meva inspiració, mirarem de treure un altre *San José* d'aquest *tronco* així de gran.

En primer lloc, les formalitats cal seguir-les. Així doncs, agrair al Departament de Matemàtiques de la Universitat Autònoma de Barcelona permetre'm el gaudi de les seves infraestructures i del voyeurisme amb el @totsdpt. També, agrair al Grup de Sistemes Dinàmics el suport i, alhora, fer-me entretingudes les sobretaules dels dilluns. Fins aquí ho tenia relativament fàcil.

És evident que em podria reduir al cas anterior i, abusant de tòpics, anar complint l'expedient. Ara bé, i ara parlant seriosament, els que veniu ara (i els d'abans també, no us enfadeu!) tenen i han tingut un paper essencial. Com no podia ser d'una altra manera el primer element cabdal és el meu director: en Lluís Alsedà. Fa poc em deia que dels directors se n'acostuma aprendre alguna cosa. Bé, puc garantir que amb en Lluís n'he après un munt de coses. De la vida, de consells, de plans de fugida, d'informàtica, de concerts, de fotografia i, és clar, de matemàtiques. Gràcies per tot Lluís.

Tampoc em puc descuidar d'una part crucial del Departament: la secretaria. Gràcies a tots i a totes (visca el llenguatge políticament correcte però no ecològic). Ara bé, especialment a la M. José i l'Ignasi. Sense ells, amb la seva gran feina de #formiguetes com ara fotocòpies estranyes, política, aire(s) acondicionat(s) que no van, lluita de classes, bitllets, ordinadors,... Per mi, fer aquesta Tesi sense ells hagués resultat encara més complicat.

Parlant de la lluita de classes, resulta evident que en qualsevol escrit d'agraïment

no pot faltar una salutació a tots els companys de despatx haguts i per haver: salut companys! Amb tot, una abraçada a *Paca La Coneja* i un fortíssim Ogayrop! pels *Barracks!!!* ja que sense vosaltres el grup subversiu *espirituós* d'ascendència euskaldun no hagués estat possible. Precisament parlant d'impossibles, hi ha qui diu que m'és impossible expressar-me sobre certs assumptes. Potser té raó. Però aquí em nego a fer servir el recurs de la llagrimeta fàcil. Així doncs, després d'una llarga temporada on la vida ha tingut una entropia massa alta: pare, mare, Jordi, germans (Pau, Júlia i Berta), iaia, àvia, tiets i cosines (amb les respectives parelles i descendència) tots hi teniu, potser sense saber-ho, algun tipus de participació important en la manufacturació d'aquest *San José*.

I és que en general, i com diria aquella, *tu, tot, tothom* hi ha estat benvingut. L'handbol per disfrutar, de manera anònima, jugar davant de moltíssima gent, la muntanya per ajudar-me a gestionar els límits, la política per lluitar per un entorn més just i, com que el tot és molt llarg, més val que ho deixi aquí. Si mirem el tothom, juntament amb el totdon, també és d'allò més vast. Així que hi posaré desitjos amb alguns dels elements representatius com ara anar collir ceps amb en Carles, parlar hores i hores d'handbol amb en Parra, apostes a la porteria amb en Cowboy, fer aquell Aneto pendent amb l'Aurora, l'expedició Ebre – Monegros amb en Josep, descobrir l'autèntica Eivissa amb la Neus, anar a aquell concert dels Guardian (abans de les onze) amb en Joan, aprendre a donar abraçades amb la Nere o regraciar na Maria per deixar-me en Bosa de tant en quant. I en Bosa. Agraït per tot, Bosa, algun dia de Salàs en farem Ítaca.

Ei, que et pensaves que m'havia oblidat de tu... doncs no! Simplement gràcies per haver caigut als braços de l'efecte papallona una nit indeterminada de color negre. Seguirem fent, somrients i tossudament alçats, la revolta.

La Garriga, Setembre de 2015.

Introduction

The main purpose of this Thesis is to give an interface between dynamical systems and analysis topics by means of the development of a software. Hence, the framework of this Thesis is Numerical Analysis. In particular, such developed interface focuses into obtain an analytical approximation of some invariant objects. From such approximation, and due to the difficulty to make calculations in an explicit way, we want to try to assess some properties of such invariant object. This is, roughly speaking, the main topic of the present dissertation.

In the following we will develop, in a not in-depth way, the motivation and the underlying problems. To this end, let us introduce the main subject of our study: a family of pinched skew products which are defined on the Cartesian product of $\mathbb{S}^1$ and $\mathbb{R}^+ = [0, \infty)$. They are of the type

$$\begin{pmatrix} \theta_{n+1} \\ x_{n+1} \end{pmatrix} = \mathfrak{F}_{\sigma,\varepsilon}(\theta_n, x_n) = \begin{pmatrix} R_\omega(\theta_n) \\ F_{\sigma,\varepsilon}(\theta_n, x_n) \end{pmatrix}, \tag{1}$$

where $R_\omega(\theta) = \theta + \omega$, $\omega \in \mathbb{R}\backslash\mathbb{Q}$ and $\sigma, \varepsilon \in \mathbb{R}$. Precisely, in these classes of dynamical systems invariant sets with a strange geometry appear. From now on, we will call the invariant φ, which depends on σ and ε by construction. Now, let us motivate our study.

Obtaining analytical approximation of this "*weird*" φ the use of wavelets instead of "*Fourier approach*" naturally arises. Indeed, the Fourier techniques, which are widely studied and used, tries to get expansions as:

$$\varphi \sim a_0 + \sum_{n=1}^{\infty} \left(a_n \cos(n\theta) + b_n \sin(n\theta)\right),$$

which are not convenient for our case. Therefore, the aim of this Thesis is to describe an efficient algorithm for the semi-analytical computation of the invariant object using wavelets. However, we want to mention that the methodology developed strongly depends on the Lyapunov exponent on φ. Recall that the Lyapunov exponent must be understood as the mean growth rate of the distance between neighboring initial point trajectories. The approximation is based on the computation of

$$\left\langle \varphi, \psi^{\mathrm{PER}}_{-j,n} \right\rangle = \int_{\mathrm{supp}(\psi^{\mathrm{PER}}_{-j,n})} \varphi(\theta)\psi^{\mathrm{PER}}_{-j,n}(\theta)\, d\theta$$

since we want to approximate a function from $\mathscr{L}^2(\mathbb{S}^1)$ by a wavelet expansion of the type

$$\varphi \sim a_0 + \sum_{j=0}^{\infty} \sum_{n=0}^{2^j-1} \left\langle \varphi, \psi^{\mathrm{PER}}_{-j,n} \right\rangle \psi^{\mathrm{PER}}_{-j,n}.$$

Observe that the unknown is, precisely, $\langle \varphi, \psi_{-j,n} \rangle$ because

$$\psi_{j,n}^{\mathrm{PER}}(x) = \sum_{\ell \in \mathbb{Z}} \psi_{j,n}(x+\ell) = 2^{-j/2} \sum_{\ell \in \mathbb{Z}} \psi\left(\frac{(x+\ell) - 2^j n}{2^j} \right)$$

and $\psi(x)$ is a given wavelet. Namely, $\psi(x)$ is a function such that its dyadic translations and its dilations by powers of two form an orthonormal basis of $\mathscr{L}^2(\mathbb{R})$.

The aim for this exercise is twofold. From one side, to study bifurcations and, perhaps, to give information of the dynamics of the object. On the other side, to estimate the regularity of φ. Notice that the study of this regularity, depending on parameters, provides another point of view to the "*fractalization routes*" as it is described, for instance, in [Nis96, JT08]. Therefore we need to calculate as well as possible the coefficients $\left\langle \varphi, \psi_{-j,n}^{\mathrm{PER}} \right\rangle$. We will compact the notation using $\mathrm{D}^{\mathrm{PER}}$ as the vector, which can be infinite dimensional, of wavelet coefficients.

Finally, we want to make a comment concerning to the wavelet coefficients $\mathrm{D}^{\mathrm{PER}}$ and the regularity pf φ. Leaving aside the classification of φ in terms of its regularity, we have used the regularity to decide, in some sense, the quality of the "*numerically obtained*" $\mathrm{D}^{\mathrm{PER}}$. Indeed, since, by construction, $\mathrm{D}^{\mathrm{PER}}$ can always restore φ then, how we can control if they are good enough? The answer is the regularity in those cases where the regularity is known.

An overview

In order to calculate the desired coefficients, let us make one step backwards. As it is described in Chapter 3 and 5 one of the main ingredients, besides of how to compute $\psi_{-j,n}^{\mathrm{PER}}(\theta)$, is the solution of a (non)-linear system of equations. Indeed, the Equation (1) can be understood as

$$\begin{array}{rccc} \mathfrak{F}_{\sigma,\varepsilon}: & \mathbb{S}^1 \times \mathbb{R} & \longrightarrow & \mathbb{S}^1 \times \mathbb{R} \\ & (\theta, x) & \longmapsto & (R_\omega(\theta), F_{\sigma,\varepsilon}(\theta, x)), \end{array}$$

and the invariant object, φ, is a solution of the invariance equation

$$\mathrm{Inv}(\theta) = F_{\sigma,\varepsilon}(\theta, \varphi(\theta)) - \varphi(R_\omega(\theta)) = 0.$$

An underlying subject related to the above equation is the Transfer Operator $\mathcal{M}$. Let us skip such concepts, which will be clearly explained in Chapter 3, and focus on something masked behind the invariance equation. Indeed, we turn our attention to

$$F_{\sigma,\varepsilon}(\theta, \varphi(\theta)) - \varphi(R_\omega(\theta)) = (F_{\sigma,\varepsilon} \circ \varphi)(\theta) - (\varphi \circ R_\omega)(\theta) = 0 \Leftrightarrow (F_{\sigma,\varepsilon} \circ \varphi)(\theta) = (\varphi \circ R_\omega)(\theta)$$

Working on this point of view, the invariance equation becomes similar to a cohomological equation. Precisely, the nature of the solutions of the cohomological equation,

in terms of regularity, is widely studied in many cases. Moreover, it arises as a good starting point to understand the reducibility methods in a general context and systems depicted in [FR11, Jor01, HdlL06a, HdlL06b, HdlL07]. These are used to undertake actions avoiding the problematic regions, which depend on σ and ε.

We point out that the methods used along the memoir are slightly different from such techniques. Nevertheless, due to the "*simplicity*" of the environment space, $\mathbb{S}^1 \times \mathbb{R}$, the Lyapunov exponent is the link between the notion of reducibility and our point of view, as it can be seen in [JT08].

Moving to the regularity ideas, remark that the idea of the Transfer Operator described along this Thesis will have two meanings. As a first instance, the Transfer Operator must be understood as the projection on the second component of the system given by Equation (1) (see [AM08, Kel96]):

$$\mathbf{T}(\varphi)(\theta) = F_{\sigma,\varepsilon}(\theta, \varphi(R_\omega(\theta))).$$

On the other side, when one uses the reducibility language to the Transfer Operator, the Invariance equation becomes an operator:

$$\mathbf{T}(\varphi)(\theta) = \varphi(R_\omega(\theta)) - F_{\sigma,\varepsilon}(\theta, \varphi(\theta)).$$

Notice that, both are essentially the same and they are used to derive regularity properties (see e.g [dlLO99, HdlL06a, HdlL06b, HdlL07, Kel96]). The main difficulty in carrying out the regularity assessment with these techniques, depending on the parameters σ and ε, is to plug them into wavelets theory. This is, precisely, one of the main purposes of the developed software.

Zoom in

In Chapter 5 we describe the vector $\mathrm{D}^{\mathrm{PER}}$ using the invariance equation and the Transfer Operator, we find. To do so, we deal with the Newton's method. In such situation, we solve the following (non)-linear system

$$\left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right)(\Theta)\mathrm{D}_N^{\mathrm{PER}} = -\mathrm{Inv}(\Theta),\ \Theta \in \mathbb{S}^1 \overbrace{\times \cdots \times}^{N} \mathbb{S}^1$$

for a certain N-variable functions $\Delta = \frac{\partial F_{\sigma,\varepsilon}(\theta,\varphi(\theta))}{\partial x}$ and Ψ^{PER}, related to $F_{\sigma,\varepsilon}$ and the wavelets ψ^{PER} respectively. Hence, in order to find the solution, the Newton-Kantorovich hypothesis must be verified. Having said that, we warn that we will not go further in this topic because we have applied the idea of if Newton's method converges, converges.

However, we can go a little bit in-depth in the above equation if we restrict ourselves to the Haar's case. Indeed, see Chapter 5 for an exhaustive explanation, our equation

becomes

$$\left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right)(\Theta)\mathrm{D}_N^{\mathrm{PER}} = \left(\mathsf{Id} - \Delta_N P^{\top}\right)(\Theta)\mathrm{D}_N^{\mathrm{PER}} = -\mathrm{Inv}(\Theta),$$

using a permutation matrix given by

$$P_{i,j} = \begin{cases} 1 & \text{if } (j + \lfloor 2^j \omega \rfloor) \mod N, \\ 0 & \text{otherwise.} \end{cases}$$

The idea behind this change of variables is the preconditioning techniques for linear systems. Observe that $P^{\top}$ is a matrix whose image is the same point θ translated a certain quantity $\widetilde{\omega}$. Moreover, the permutation matrix $P^{\top}$, when j tends to∞, becomes *the* rotation $R_{\omega}(\theta)$. Indeed, for all $x \in \mathbb{R}$ it follows that $x - 1 \leq \lfloor x \rfloor \leq x$. Using such inequality it can be shown that $\lim_{j\to\infty} \dfrac{\lfloor 2^j \omega \rfloor}{2^j} = \omega$.That is, $P_{\infty}^{\top} = R_{\omega}(\theta)$. Namely, we have translated the Transfer Operator to another one more understandable. Indeed, we can find the vector $\mathrm{D}^{\mathrm{PER}}$ iteratively solving

$$\mathbf{T}_{\sigma,\varepsilon}(\mathrm{D}_k^{\mathrm{PER}}) = (\mathsf{Id} - \Delta_{\sigma,\varepsilon} \circ R_{\omega})(\Theta)\mathrm{D}_k^{\mathrm{PER}} = -\mathrm{Inv}(\Theta), \tag{2}$$

with a given initial seed $\mathrm{D}_0^{\mathrm{PER}}$. Hence, the "*contraction - invertible properties*" of $\mathbf{T}_{\sigma,\varepsilon}$ will decide the convergence towards $\mathrm{D}^{\mathrm{PER}}$. Taking $\mathbf{T}_{\sigma,\varepsilon}$ as an (infinite) matrix will help us. Actually, $\mathbf{T}_{\sigma,\varepsilon}$ as a matrix will be a contractive matrix if its spectral radius, $\varsigma(\mathbf{T}_{\sigma,\varepsilon})$, is less than one. Such condition is equivalent to demand $\varsigma(\Delta_{\sigma,\varepsilon} \circ R_{\omega}) < 1$, and this is equivalent to have the Lyapunov exponent of the system given by Equation (1) less than zero (see Chapter 5 for more details). In view of that, it arises three "*natural*" situations:

(a) If the Lyapunov exponent is less than zero, then the operator is invertible and Equation (2) converges to $\mathrm{D}^{\mathrm{PER}}$ hopefully.

(b) If the Lyapunov exponent is close to zero, then the operator nearby not invertible and we have kernel. Hence, other techniques must be applied.

(c) If the Lyapunov exponent is positive, then system given by Equation (1) has a repellor and we do not consider such case.

In other words, what will control in which of the above cases we are will be the parameters σ and ε. Notice that it is linked to the question whether those parameters, specially ε, "*control*" the regularity properties of φ or the regularity is inherent to φ.

Zoom out

From a general point of view, wavelet coefficients determine the regularity of a function in the same way as in the Fourier world. Hence, in our case, $\mathrm{D}^{\mathrm{PER}}$ will decide if φ is on a

certain regularity space. Let us denote this space by $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, where $s \in \mathbb{R}$. Roughly speaking, $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ are the generalization of Hölder spaces (under certain constraints) for all real value of s (see e.g [Tri83, Tri06, Tri10]). This range (even negative!) seems a bit weird; however there are systems where φ it is not a graph of a (usual) function, as it can be guessed in Figure 1. On the contrary, an invariant object with a strange geometry at a first glance may become a nice object after some manipulations (see e.g [AM08, Jor14]). Thus, all range of s must be allowed and other regularity spaces may be considered.

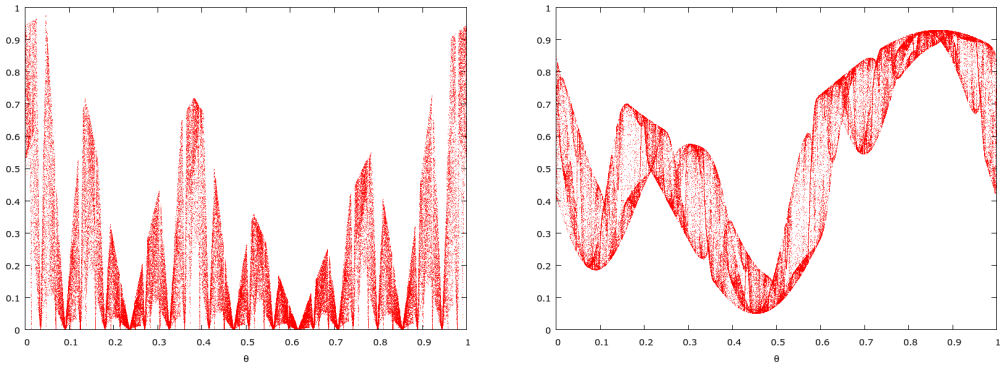


Figure 1: On the left picture it is shown an attractor *with* area [AM08]. On the right picture it is displayed the Nishikawa-Kaneko model with $\sigma = 3.0$ and $\varepsilon = 0.18$ [Nis96].

However, we will have two important and related drawbacks. The first one is linked to the three "*natural*" situations described at the end of the last paragraph. For certain values of σ and ε the Lyapunov exponent is relatively close to zero. In these cases, the vector $\mathrm{D}^{\mathrm{PER}}$ is very hard to calculate, in CPU time consuming. We use standard continuation techniques on the way of solving such problem. The other disadvantage is that "*a priori*" Equation (2) does not knows, in some sense, if $\mathrm{D}^{\mathrm{PER}}_k$ will have some nice vanishing property. That is, using this methodology is mandatory to converge towards $\mathrm{D}^{\mathrm{PER}}$ to be able to classify φ in one of these spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$.

Organization and contributions of this Thesis

This dissertation is divided into two parts. The first one (Chapters 1, 2and 3) is devoted to recall and introduce all the theory and the methodology which is used into the second part (Chapters 4, 5 and 6). In Chapter 1 a self-contained and general introduction to wavelets it is done. However, since we will use Daubechies wavelets with $p \geq 1$ vanishing moments, we will focus in such wavelets (in $\mathbb{R}$ or $\mathbb{S}^1$). Actually, along this thesis we have performed the translation from the $\mathbb{R}$-Daubechies wavelets language

to $\mathbb{S}^1$. Following the generic notation, in the literature we have call it ψ^{PER}. Also, in this first Chapter a crash course on the notion of regularity, in terms of the functional spaces $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$ and $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, and the relationship between them and the wavelets coefficients, $\left\langle f, \psi^{\mathrm{PER}}_{-j,n} \right\rangle$, is done.

Precisely, the main topic of Chapter 2 is the calculation of the wavelets coefficients using two different techniques; namely, Fast Wavelet Transform and the solution of (non)-linear systems of equations. In a more concrete terms, we have performed an algorithm to calculate, in an efficient and fast way, ψ^{PER} on a (really big!) mesh of points of $\mathbb{S}^1$. The method is based on the Daubechies-Lagarias algorithm (see [DL91, DL92]). Such computation, which is a key point of this disquisition, will be the main part of the Chapter 2.

Finally, the last chapters of this first part are devoted to give a short compilation of the theoretical framework where this dissertation is dealt. In Section 3.2 it is shown the machinery and also we try to characterize the mechanisms involved in the geometric properties of a particular family of quasi periodically forced skew products on the cylinder. As a matter of fact, combining them with ideas from [AM08, Har12] it is possible to extend the results of [Sta97, Sta99] to a more weird class of functions. Also we derive "*theoretically*" the regularity, in terms of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, for φ of the Keller-GOPY model. Despite of this two remarkable facts, such effort is to justify the use of the software in other cases of SNA's, as those ones in [AM08, Nis96], among the study of them in terms of regularity.

Moving to the second part of this Thesis, two different exercises are done. The first one, following [dlLP02], is the development of an algorithm to estimate regularities, in terms of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, for $s \in \mathbb{R}$. Such algorithm is the main goal of Chapter 4, and it is used in many situations, but not only with the Fast Wavelet Transform. Indeed, in Chapter 5 we perform the same kind of regularity assessment with a different methodology to obtain $\left\langle \varphi, \psi^{\mathrm{PER}}_{-j,n} \right\rangle$.

Certainly, in the last Chapter(s), we have focused to solve the Invariance Equation for several situations and dynamical systems. The solvers, which are iterative, are the main contribution of such Chapter(s). Due to the good properties of the Daubechies wavelet family, such as compact support or vanishing moments among others (see e.g [Dau92, HW96, Tri06]), we have derived two iterative strategies to find $\left\langle \varphi, \psi^{\mathrm{PER}}_{-j,n} \right\rangle$. Both of them are based in the same argument but, due to the simplicity of the Haar wavelet ($p = 1$), the first strategy give us a close to explicit method calculate the Haar wavelet coefficients. As mentioned, beyond the "*numerical approximation of the invariant objects*", we have estimated the regularity, in terms of $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$ and $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, of a certain models of skew products using the Haar wavelet. Both goals have been repeated with other Daubechies wavelet. Nevertheless, we would like to remark that, when $p > 1$, the core of the iterative method is the aforementioned massive evaluation of ψ^{PER} and, also, the left conditioned discretization of the Transfer Operator. As an

extra point, a numerical exploration of the Lyapunov exponent of a particular instance of the kind of systems in [AM08] it is done.

Loose ends of this Thesis

It is worth pointing out that in each of the chapters there are some open problems and questions which we summarize at the end as a (possible) future work. Concretely, the above strategies open the door too study the operator in terms of Newton-Kantorovich theorem. This would guarantee the stability of the iterative method. In this way, there are some open questions such as which is the limit operator, when one uses a precondition technique, and how the convergence of the iterative method is affected by the Lyapunov exponent. Of course, for a well suited norm, Newton-Kantorovich theorem must may be used to detect, "*a priori*", the lack of regularity of φ. Moving to the continuation methodology, since we have convergence, it seems useful to find, in a more general context, such strips of convergence towards the desired case of σ and ε using the operator framework.

Also, the developed software it might be updated to work on high dimensions and some qualitative-quantitative properties of φ. For example, a modification of it can produce rigorous numerical estimation of φ's traits, such as the Hausdorff Dimension or its length. Moreover, theoretical bounds on the aforesaid traits could be done using the Haar wavelet because of its simplicity. Nonetheless, as it is described in Chapter 5 and 6, the bug of the precision must be completely understood.

Contents

Part I

Tools and Problems

Chapter 1

An Introduction to Wavelets

Consider a square integrable function, f, of some Hilbert space H. Suppose that a "*sufficiently good*" approximation is needed. Such an approximation must be given in terms of an orthonormal basis of the aforementioned Hilbert space,$\{e_i(x)\}_{i\in\Lambda}$ (being $\Lambda \subseteq \mathbb{Z}$ a set of indexes). The problem is solved by determining some coefficients, $d_i \in \mathbb{R}$. Such quantities, are uniquely determined by the inner product inherited from the space H. That is,

$$f \sim \sum_{i\in\Lambda} \langle f, e_i\rangle\, e_i(x) = \sum_{i\in\Lambda} d_i e_i(x).$$

This chapter is devoted to recall basic results and fix notation that we will use in the following chapters. Also we will introduce an orthonormal basis of some useful Hilbert spaces: the wavelets. Wavelets are functions that verifies certain requirements and they are well suited to face various problems.

In the forthcoming, we will be focused on dynamical systems such that they exhibit (or they have) some kind of invariance (or invariant object). In the literature there are several techniques [HdlL06a, HdlL06b, HdlL07, Jor01] of how one can deal with the following problem

> *"Given a dynamical system exhibiting periodic or quasi-periodic behavior, obtain a sufficiently good approximation of the invariant object of the dynamical system".*

But since invariant objects can exhibit a complicated shape, obviously the shape will depend on the system considered, then finite wavelet expansions can be used. Such basis can capture different frequencies at different regions of the space allowing us to perform better strategies to solve the problem. In other words, wavelets seems well suited to approximate data with strange geometry.

1.1 Wavelets: a shortcut

Our aim is to approximate a certain class of functions, which in a natural way they live in $\mathbb{S}^1$, by means of wavelets. Recall that a standard approach used in the literature to compute and work with invariant objects of systems exhibiting periodic or quasi-periodic behavior is to use finite Fourier approximations (trigonometric polynomials),

namely functions of the form

$$\varphi(\theta) \sim a_0 + \sum_{n=1}^{N} \left(a_n \cos(n\theta) + b_n \sin(n\theta) \right).$$

But finite wavelet expansions could be used instead of the above Equation, namely:

$$\varphi(\theta) \sim \sum_{j=0}^{N} \sum_{n=0}^{N_j} \langle f, \psi_{j,n} \rangle \, \psi_{j,n}(\theta),$$

where $\psi_{j,n}(\theta)$ is obtained by translation and dilation of a mother wavelet $\psi(x)$. To this end, let us start by introducing an orthonormal basis of $\mathscr{L}^2(\mathbb{R})$. Firstly,recall that a *Hilbert space* H is a real (or complex) inner product space that is also a complete metric space with respect to the distance function induced by the inner product. Also, remember that a *complete metric space* is defined by forcing that every Cauchy sequence in such space converges in the space. For us, the space H will be either $\mathscr{L}^2(\mathbb{R})$ or $\mathscr{L}^2(\mathbb{S}^1)$ and hence one can define

$$\langle f, g \rangle = \int_H f(x) g(x) \, d\eta$$

as the inner product and $d\eta$ will be the usual Lebesgue measure. Also

$$\|f\|_2 = \left(\int_H |f(x)|^2 \, d\eta \right)^{1/2}$$

is a norm in $\mathscr{L}^2(H)$. Moreover, if we restrict $\| \ \|_2$ over a closed subspace F of $\mathscr{L}^2(H)$ then F is a Hilbert space also. After that, we can start the (historical) construction of the wavelets.

1.1.1 The historical construction

There are several ways to introduce wavelets and wavelet bases. For our purposes we will present and construct them by means of the Multi-resolution Analysis. From such point of view will appear, in a natural way, a crucial sequence of numbers for our calculations: the scaling filter. We refer the reader to [HW96, Mal98] for more detailed and comprehensive expositions, but let us mention that for "*historical*" reasons we will start out introduction to the wavelet stuff in the $\mathbb{R}$ case. After that, we will be able to make the translation to the $\mathbb{S}^1$ in a more feasible way. With this comments in mind, we are ready to introduce the notion of Multi-resolution Analysis.

Definition 1.1.1. A sequence of closed subspaces $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ of $\mathscr{L}^2(\mathbb{R})$ is a *Multi-resolution Analysis* (or simply a *MRA*) if it satisfies the following six properties:

(a) $\{0\} \subset \cdots \subset \mathcal{V}_1 \subset \mathcal{V}_0 \subset \mathcal{V}_{-1} \subset \cdots \subset \mathscr{L}^2(\mathbb{R})$.

(b) $\{0\} = \bigcap_{j\in\mathbb{Z}} \mathcal{V}_j$.

(c) $\operatorname{clos}\left(\bigcup_{j\in\mathbb{Z}} \mathcal{V}_j\right) = \mathscr{L}^2(\mathbb{R})$.

(d) There exists a function ϕ whose integer translates, $\phi(x-n)$, form an orthonormal basis of $\mathcal{V}_0$. Such function is called the *scaling function*.

(e) For each $j \in \mathbb{Z}$ it follows that $f(x) \in \mathcal{V}_j$ if and only if the *dyadic translations* verifies that $f(x - 2^j n) \in \mathcal{V}_j$ for each $n \in \mathbb{Z}$.

(f) For each $j \in \mathbb{Z}$ it follows that $f(x) \in \mathcal{V}_j$ if and only if the *dilations* verifies that $f(x/2) \in \mathcal{V}_{j+1}$.

∎

If we fix an MRA, we know that $\mathcal{V}_j \subset \mathcal{V}_{j-1}$, for every j, and that $\mathcal{V}_j$ has an orthonormal basis $\{\phi_{j,n}\}_{n\in\mathbb{Z}}$, for every j, where

$$\phi_{j,n}(x) = 2^{-j/2}\phi\left(\frac{x - 2^j n}{2^j}\right).$$

Now define the subspace $\mathcal{W}_j$ as the orthogonal complement of $\mathcal{V}_j$ on $\mathcal{V}_{j-1}$. That is,

$$\mathcal{V}_{j-1} = \mathcal{W}_j \oplus \mathcal{V}_j. \tag{1.1}$$

Therefore, by the inclusion property of the spaces $\mathcal{V}_j$ we have

$$\mathscr{L}^2(\mathbb{R}) = \operatorname{clos}\left(\bigoplus_{j\in\mathbb{Z}} \mathcal{W}_j\right) = \operatorname{clos}\left(\mathcal{V}_0 \oplus \bigoplus_{j=-\infty}^{0} \mathcal{W}_j\right). \tag{1.2}$$

Now, define the *mother wavelet* $\psi \in \mathcal{W}_0$ to be the function whose Fourier transform is

$$\widehat{\psi}(\xi) = \frac{1}{\sqrt{2}} e^{-i\xi}\widehat{h}^*(\xi + \pi)\widehat{\phi}(\xi) \tag{1.3}$$

where $\widehat{h}^*(\xi)$ is the complex conjugate of

$$\widehat{h}(\xi) = \sum_{n\in\mathbb{Z}} h[n]e^{-in\xi}$$

and under mild conditions it is verified that $\widehat{\phi}(0) = 1$ and, by definition, one can take $h[n] = \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right), \phi(x-n)\right\rangle$. For completeness let us recall that the *Fourier Transform* of a function f is given, as usual, by

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-i\xi x}\, dx.$$

Definition 1.1.2. Let $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ be a Multi-resolution Analysis and let $\phi(x)$ be its scaling function. Define

$$h[n] := \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right), \phi(x-n)\right\rangle \tag{1.4}$$

The sequence $h[n]$ is called the *scaling filter* (or the *low pass filter*) of the Multi-resolution Analysis. We define the support of $h[n]$, denoted by $\mathrm{supp}(h)$, as the minimum subset $\mathfrak{I}$ of $\mathbb{Z}$ such that $\mathfrak{I} = \{\ell, \ell+1, \ldots, \ell'\}$ is a set of consecutive integers and

$$h[n] = 0 \text{ for every } n \in \mathbb{Z}\backslash\mathfrak{I}.$$

■

Remark 1.1.3. From the above definitions, it is easy to see the following helpful equations

$$\frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right) = \sum_{n\in\mathbb{Z}} h[n]\phi(x-n) \quad \text{and} \quad \frac{1}{\sqrt{2}}\psi\left(\frac{x}{2}\right) = \sum_{n\in\mathbb{Z}} g[n]\phi(x-n). \tag{1.5}$$

■

For abbreviate, and when it is not possible to induce a confusion, we will refer to the sequence $h[n]$ simply by filter instead of scaling filter (or low pass filter). Such list of coefficients has two important properties which will be helpful in the forthcoming.

Proposition 1.1.4. *Let $h[n]$ be a filter of the Multi-resolution Analysis and $\phi(x)$ its scaling function associated. Then it is verified that*

(a) The normalization *property*

$$\sum_{n\in\mathbb{Z}} h[n] = \sqrt{2}.$$

(b) The double shift orthogonality *property*

$$\sum_{n\in\mathbb{Z}} h[n]h[n-2l] = \delta_{0,l} \text{ for any } l \in \mathbb{Z},$$

where $\delta_{0,l}$ stands for the Dirac's delta.

(c) It is verified that

$$\sum_{n\in\mathbb{Z}} h[2n] = \sum_{n\in\mathbb{Z}} h[2n+1].$$

Proof. In order to prove the first property, just consider

$$\frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right) = \sum_{n\in\mathbb{Z}} h[n]\phi(x-n).$$

Then, using a change of variables,

$$\int_{\mathbb{R}} \phi(x)\, dx = \frac{1}{2}\int_{\mathbb{R}} \phi\left(\frac{t}{2}\right)\, dt = \sqrt{2}\int_{\mathbb{R}} \sum_{n\in\mathbb{Z}} h[n]\phi(x-n) = \sqrt{2}\sum_{n\in\mathbb{Z}} h[n]\int_{\mathbb{R}} \phi(x)\, dx.$$

and the result follows since $\widehat{\phi}(0) = 1$.
For the second item, recall that the function $\phi(x)$ verifies that $\phi(x-n)$ form an orthonormal basis and

$$\phi(x-l) = \sqrt{2}\sum_{n\in\mathbb{Z}} h[n]\phi(2(x-l)-m).$$

Therefore,

$$\begin{aligned}
\delta_{0,l} &= \langle \phi(x), \phi(x-l)\rangle \\
&= 2\sum_{n\in\mathbb{Z}} h[n]\left(2\sum_{m\in\mathbb{Z}} h[m]\int_{\mathbb{R}} \phi(2x-n)\phi(2(x-l)-m)\right) \\
&= \sum_{n\in\mathbb{Z}}\sum_{m\in\mathbb{Z}} h[n]h[m]\delta_{n,2l+m} \\
&= \sum_{n\in\mathbb{Z}} h[n]h[n-2l]
\end{aligned}$$

as we wanted to show.
Finally, for the last point we must say that for all $\xi \in \mathbb{R}$ then, see [Mal98, Theorem 7.2],

$$|\widehat{h}(\xi)|^2 + |\widehat{h}(\xi+\pi)| = 2$$

and is a 2π-periodic function. Therefore, $\widehat{h}(\pi)$ must be zero since $\widehat{h}(2\pi) = \widehat{h}(0) = \sqrt{2}$. In other words, we have

$$0 = \widehat{h}(\pi) = \sum_{n\in\mathbb{Z}} h[n]e^{-in\pi}$$

The proof follows taking into account that $e^{in\pi} = \cos(n\pi) + i\sin(n\pi) = (-1)^n$. □

Let us make another comment concerning on the filter. Observe that, since $h[n]$ is defined in terms of an equation which involves two functions then either the support of $h[n]$ is infinite, that is $\mathrm{supp}(h) = \mathbb{Z}$, or the support of $h[n]$'s is finite. We will focus on a family of $h[n]$ with finite support: *Daubechies wavelets*.

The following theorem (see [Mal98, Theorem 7.3]) allows to obtain the wavelet basis from the scaling function and hence we will see that $\psi(x)$ is completely characterized by h:

Theorem 1.1.5 (Mallat, Meyer). *The mother wavelet given by Equation (1.3) verifies that, for each integer j, the family $\{\psi_{j,n}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $\mathcal{W}_j$, where:*

$$\psi_{j,n}(x) = 2^{-j/2}\psi\left(\frac{x-2^j n}{2^j}\right).$$

As a consequence, the family $\{\psi_{j,n}\}_{(j,n)\in\mathbb{Z}\times\mathbb{Z}}$ is an orthonormal basis of $\mathscr{L}^2(\mathbb{R})$.

Thus, taking into account (1.2) and the above theorem, every map $f \in \mathscr{L}^2(\mathbb{R})$ can be written as

$$\sum_{j\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} \langle f, \psi_{j,n}\rangle\, \psi_{j,n}(x) = \sum_{n\in\mathbb{Z}} \langle f, \phi_{0,n}\rangle\, \phi_{0,n}(x) + \sum_{j=0}^{\infty}\sum_{n\in\mathbb{Z}} \langle f, \psi_{-j,n}\rangle\, \psi_{-j,n}(x),$$

which is the translation of the Fourier language to the wavelets one. We want to emphasize that there are many families of wavelets and each one of them it is well suited for a kind of problems (see [Mal98]). We will be focused on Daubechies wavelets which are appropriated for our purposes in the forthcoming. Also, we will recall and prove some properties concerning to the Haar wavelet. Such wavelet, due to its simplicity, can be considered as a particular instance of the Daubechies wavelet family (among others). Of course, such kind of orthonormal basis appears before all the wavelet theory was introduced. However, it fulfills the requisites to be a wavelet.

1.1.2 Daubechies wavelets

From Theorem 1.1.5, it is clear that the discrete filter $h[n]$ characterizes the wavelet $\psi(x)$. As we have said before, we will work with a family of wavelets called Daubechies wavelets and consequently we must explain how $h[n]$ appears for such family of wavelets. The Daubechies family appears by combining, at least, two tools: the vanishing moments and the support of the wavelet. Let us start by defining the first one.

Definition 1.1.6. Let $\psi(x)$ be a wavelet from a Multi-resolution Analysis $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$. We say that $\psi(x)$ has p-*Vanishing Moments* if the integer p is the minimum number such that

$$\int_{\mathbb{R}} x^k\psi(x)\, dx = 0 \text{ for } 0 \le k < p.$$

■

Notice that Definition 1.1.6 means that $\psi(x)$ is orthogonal to any polynomial of degree $p-1$ i.e the integral is zero. Hence, if a function is locally as a polynomial then, if one uses a wavelet with sufficiently enough vanishing moments, the coefficients $\langle f, \psi_{j,n}\rangle$ will become, in absolute value, closer to zero. Therefore, we have an indicator for the localization of possible function singularities once, given threshold, *big* wavelet coefficients have been located.

Let us introduce the other tool involved in the creation of Daubechies wavelets: the size support of $\psi(x)$. It plays an important role in order to be able to make finite (and exact) calculations. Let us give its definition.

Definition 1.1.7. Let $\psi(x)$ be a wavelet from a Multi-resolution Analysis $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$. We define the *support* of $\psi(x)$, and we will denote it by $\operatorname{supp}(\psi(x))$, as the minimum interval of the real line such that $\psi(x) = 0$ for every x which does not belongs to the support of $\psi(x)$. The size of $\operatorname{supp}(\psi(x))$ is defined as the diameter of such interval. ∎

Before continuing the explanation we should point why we to keep in mind the size support of a wavelet $\psi(x)$ besides the finite calculations. To this end, let us consider an isolated singularity on x_* of a function f and assume that x_* belongs to the support of some $\psi_{j,n}(x)$. Then the coefficient $\langle f, \psi_{j,n}\rangle$ may be big or may be not. But in any case, there exists some translations of $\psi_{j,n}(x)$ whose support contains such point x_* which is probably problematic and, in terms of finite numerical calculations, it can be a error source. Consequently, it is interesting to minimize, as much as possible, the size support of the wavelet in order to avoid this possibly problematic overlapping.

Proposition 1.1.8 (Proposition 7.2, [Mal98]). *Let $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ be an Multi-resolution Analysis which has $\phi(x)$ as its scaling function, $h[n]$ as its associated discrete filter and $\psi(x)$ its associated wavelet. In such conditions, $\phi(x)$ has compact support if and only if $h[n]$ has compact support. Moreover, their supports are equal. Let $[C_1, C_2]$ be the support of $\phi(x)$, then $\left[\frac{C_1-C_2+1}{2}, \frac{C_2-C_1+1}{2}\right]$ is the support of $\psi(x)$.*

Now we can state a theorem due to I. Daubechies that relate the, *a priori* independent, notions of vanishing moments and the size support of a wavelet $\psi(x)$.

Theorem 1.1.9 (Daubechies, Theorem 7.3, [Mal98]). *Let h be a real Conjugate Mirror filter from a Multi-resolution Analysis $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ such that $\widehat{h}(\omega)$, has $\omega = \pi$ a root with multiplicity p. Then its associated wavelet has p-vanishing moments and its support size is larger than or equal $2p-1$. Moreover such a wavelet has a minimum size support equal to $[-p+1, p]$.*

Remark 1.1.10. The proof of the above Theorem it is not done here and it can be found at [Mal98] for example, but what it is important to notice is that the proof is constructive and from it, the coefficients $h[n]$ can be, at least numerically, computed but it is important to remark that other methods can be used (using the Pascal's triangle). Wavelets constructed following such proof are a family of compactly supported orthonormal wavelets with the minimum size of support and they are obtained from a constraint conditions over the $h[n]$. In order to minimize the approximation error, these coefficients can be also constructed by means of parameterizations, as we will see in the Remark 1.1.12. ∎

Hence, by the above remark, we are allowed to define the Daubechies wavelets family as follows.

Definition 1.1.11. Let $\psi(x)$ be a wavelet constructed under the hypothesis of Theorem 1.1.9. Such a wavelet is called *Daubechies wavelet* with p vanishing moments. ∎

As it has been mentioned before, wavelets are completely determined by $h[n]$ and, for example, if one wants to use one of the Daubechies wavelets family they are given by list of coefficients, $h[n]$, instead of a explicit formulation for $\psi(x)$ (see [Mal98, Table 7.2]). This last comment will be very important for us since besides the case $p = 1$ which is the Haar wavelet which recall that it is given by

$$\psi(x) = \begin{cases} 1 & 0 \leq x < 1/2, \\ -1 & 1/2 \leq x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

wavelets do not have a closed expression, in contrast to the trigonometric polynomials, besides this trivial, but useful, equality

$$\psi(x) = \begin{cases} \psi(x) & 1 - p \leq x < p, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, it is difficult to evaluate a wavelet at a given point x_*. The following chapter, in Section 2.1, there is a method devoted to explain a way of use of $h[n]$ to get values of $\psi(x_*)$. Before continuing the explanation we want to remark one problem related with the accuracy of the values $h[n]$.

Remark 1.1.12 ([RS05, Reg07]). Let us consider the case of Daubechies wavelet with four filter coefficients, being all them real. Using the above techniques one has the following system of equations

$$\begin{aligned} h[0] + h[1] + h[2] + h[3] &= 2 \\ h[0] - h[1] + h[2] - h[3] &= 0 \\ h[1] + 2 \cdot h[2] + 3 \cdot h[3] &= \iota_1 \\ h[0]h[2] + h[1]h[3] &= 0 \end{aligned}$$

which corresponds to the normalization, first sum rule, one discrete moment ι_1 and the double shift orthogonality respectively. Hence, one can consider the solution of the linear part in terms of $h[0]$ and using it in the non-linear part

$$-2h[0]^2 + (5 - \iota_1)h[0] - \frac{1}{4}\iota_1^2 + 2\iota_1 - \frac{15}{4} = 0$$

we get

$$h[0] = \frac{45 - \iota_1 \mp \sqrt{-\iota_1^2 + 6\iota - 5}}{4}.$$

Taking into account that we are looking for a real $h[n]$ for $n = 0, \ldots, 3$ this implies that $\iota_1 \in [1, 5]$. Now, in order to have the parameter values symmetrically to zero, we can apply the transformation $\iota_1 = \alpha + 3$ and rewrite the solutions with this transformation

$$\begin{array}{rcl} h[0] & = & \frac{1}{2} - \frac{1}{4}\alpha - \frac{1}{4}w \\ h[1] & = & \frac{1}{2} - \frac{1}{4}\alpha + \frac{1}{4}w \\ h[2] & = & \frac{1}{2} + \frac{1}{4}\alpha + \frac{1}{4}w \\ h[3] & = & \frac{1}{2} + \frac{1}{4}\alpha - \frac{1}{4}w \end{array}$$

with $w = \sqrt{4 - \alpha^2}$ and $v = \iota_1 - 3$. Now, Daubechies wavelet has two vanishing moments, then using the sum rule

$$2h[0] - h[1] + h[3] = 0$$

and solving, one has that $\alpha = 3 \pm \sqrt{3}$. This values of α yields to the well known 4 tap Daubechies filter if $\alpha = 3 - \sqrt{3}$:

$$h[0] = \frac{1 + \sqrt{3}}{4} \qquad h[1] = \frac{3 + \sqrt{3}}{4},$$
$$h[2] = \frac{3 - \sqrt{3}}{4} \qquad h[3] = \frac{1 - \sqrt{3}}{4}.$$

One may think, what a rude way to present the filter (instead of the usual way). What it is done here is a kind of "parametrization", again instead of the "*explicit*" root of a polynomial, which can be useful in order to minimize the error and also allows the use of a symbolic calculator. This way, it is not usual but it can be useful for example when one tries to make Gaussian quadratures. ∎

Now, it only remains the translation from $\mathbb{R}$ to $\mathbb{S}^1$ of the wavelets setting.

1.1.3 Wavelets on the unit circle

As we have said at the beginning of the present chapter our natural scenario will be $\mathbb{S}^1$. And so then, we need to translate some technical results from the previous section to the unit circle setting. To this end, recall that $\mathscr{L}^2(\mathbb{S}^1)$ can be understood as the space of the 1-periodic functions defined on $\mathbb{R}$ such that

$$\|f\|_{\mathscr{L}^2([0,1])} = \int_0^1 f(x)\, dx < \infty.$$

There are several ways to work with wavelets on $\mathbb{S}^1$ but all of them have in common that whichever it is the point of view, one has that the space generated by the translation of the j-generating wavelet is finite dimensional. This is contrary to the corresponding Theorem 1.1.5. Let us try to justify such comment.

In Section 4.5 from [HW96] wavelets and MRA's are constructed to be an orthonormal basis of $\mathscr{L}^2(\mathbb{S}^1)$ using the "*periodization*" of those ones in $\mathscr{L}^2(\mathbb{R})$. Let us define them, in the same way:

$$\psi_{j,n}^{\mathrm{PER}}(x) = \sum_{\ell\in\mathbb{Z}} \psi_{j,n}(x+\ell) = 2^{-j/2} \sum_{\ell\in\mathbb{Z}} \psi\left(\frac{(x+\ell)-2^j n}{2^j}\right) \tag{1.6}$$

which are, just by simple inspection, 1-periodic functions belonging to $\mathscr{L}^1(\mathbb{S}^1)$. Now, we are ready to state the Theorem that has raised up the initial comment of this section: the finiteness of the translations.

Theorem 1.1.13 (Theorem 5.9, [HW96]). *Let $\psi(x)$ be an orthonormal wavelet from a $\mathbb{R}$-MRA. Then the system given by*

$$\{1, \psi_{j,n}^{\mathrm{PER}} \text{ with } j \leq 0 \text{ and } n = 0, 1, \ldots, 2^{-j}-1\}$$

is an orthonormal basis of $\mathscr{L}^2(\mathbb{S}^1)$.

Recall that from Equation (1.2) we have the $\mathbb{R}$-expansion of $f \in \mathscr{L}^2(\mathbb{R})$ by linear combinations of wavelets of the form

$$f = \sum_{n\in\mathbb{Z}} \langle f, \phi_{0,n}\rangle \, \phi_{0,n} + \sum_{j=0}^{\infty} \sum_{n\in\mathbb{Z}} \langle f, \psi_{-j,n}\rangle \, \psi_{-j,n} \in \mathcal{V}_0 \oplus \bigoplus_{j=0}^{\infty} \mathcal{W}_{-j}$$

but in view of Theorem 1.1.13 the above expression, in $\mathbb{S}^1$, becomes

$$f = \text{constant} + \sum_{j=0}^{\infty} \sum_{n\in\mathbb{Z}} \left\langle f, \psi_{-j,n}^{\mathrm{PER}} \right\rangle \psi_{-j,n}^{\mathrm{PER}} \in \mathcal{V}_0 \oplus \bigoplus_{j=0}^{\infty} \mathcal{W}_{-j}^{\mathrm{PER}}$$

where, in the spirit of such Theorem, for $j \geq 0$ we define

$$\mathcal{V}_0^{\mathrm{PER}} := \langle 1 \rangle \quad \text{and} \quad \mathcal{W}_{-j}^{\mathrm{PER}} := \left\langle \psi_{-j,0}^{\mathrm{PER}}, \psi_{-j,1}^{\mathrm{PER}}, \ldots, \psi_{-j,2^j-1}^{\mathrm{PER}} \right\rangle.$$

Notice that, as usual, $\langle f_1, f_2, \ldots, f_n\rangle$ denotes the subspace of $\mathscr{L}^2(\mathbb{S}^1)$ generated by the functions $f_1, f_2, \ldots, f_n$, where $f_i \in \mathscr{L}^2(\mathbb{S}^1)$ with $i = 1, \ldots, n$. In view of that, one can consider to perform the same strategy done in $\mathbb{R}$ to get an expression for the periodization of Daubechies wavelet which is a wavelet with compact support but now in $\mathbb{S}^1$. Recall that one we have defined the *support* of $\psi(x)$, Definition 1.1.7, as the minimum interval of the real line such that $\psi(x) = 0$ for every x which does not belongs to the support of $\psi(x)$.

Proposition 1.1.14. *Let $\psi(x)$ be an $\mathbb{R}$-MRA orthonormal wavelet with compact support: $\mathrm{supp}(\psi(x)) = [a, b]$. Then*

$$\psi^{\mathrm{PER}}(x) = \sum_{\ell\in\Lambda_x} \psi(x+\ell),$$

where $\Lambda_x \subset \mathbb{Z}$ is a finite set.

Proof. Recall that $\psi^{\mathrm{PER}}(x)$ is a 1-periodic function over $\mathbb{R}$ which belongs to $\mathscr{L}^2([0,1]$ and, therefore, its values are copies of those ones in $[0,1]$. On the other side, since $\mathrm{supp}(\psi(x)) = [a,b]$ then $\psi(x) \neq 0$ if $a \leq x \leq b$. Hence, to know the values of $\psi^{\mathrm{PER}}(x)$, ℓ must verify that $a \leq x + \ell \leq b$ with $x \in [0,1]$. That is $\ell \in [a-x, b-x]$. Now, to conclude the proof, just set Λ_x to be the minimum subset of $\mathbb{Z}$ such that $\ell \in [a-x, b-x]$ with $(x \in [0,1])$. That is, Λ is a set of consecutive integers $\Lambda = \ell, \ell+1, \ldots, \ell'$ and such that $\Lambda_x \subset [\lceil a-x \rceil, \lfloor b-x \rfloor]$, where $\lfloor x \rfloor$ is the largest integer not greater than x and $\lceil x \rceil$ is the smallest integer not less than x. □

Moreover, as in the $\mathbb{R}$ case, the vanishing moments are allowed "*with the*" sense of the next Proposition which is a slightly different from the usual sense. However, we want to emphasize that the $\mathbb{S}^1$-vanishing moments also holds if the $\mathbb{R}$-Daubechies wavelet has $p \geq 1$ vanishing moments.

Proposition 1.1.15. *Let $\psi(x)$ be an $\mathbb{R}$-Daubechies wavelet with $p > 1$ vanishing moments. Then*

$$\psi_{j,n}^{\mathrm{PER}}(x) = \sum_{\ell \in \Lambda_x} \psi_{j,n}(x+\ell),$$

where $0 \leq 1 \leq n \leq 2^j - 1$ and $\Lambda_x \subset [\lceil 2^j(1-p+n) - x \rceil, \lfloor 2^j(p+n) - x \rfloor]$.

Proof. It is clear by Proposition 1.1.14 and Theorem 1.1.9. □

Remark 1.1.16. As usual, the domain of $\psi_{j,n}^{\mathrm{PER}}(\cdot)$ it is defined as the minimal set $\Omega_{j,n}^{\mathrm{PER}} = \{\theta \in \mathbb{S}^1 \,:\, \psi_{j,n}^{\mathrm{PER}}(\theta) \neq 0\}$. But, by definition $\psi_{j,n}^{\mathrm{PER}}(\theta) \neq 0$ if and only if $[2^j(1-p+n) - \theta, 2^j(p+n) - \theta] \cap \mathbb{Z} \neq \emptyset$. Therefore, we define the domain of $\psi_{j,n}^{\mathrm{PER}}(\cdot)$ as follows

$$\Omega_{j,n}^{\mathrm{PER}} = \{\theta \in \mathbb{S}^1 \,:\, [2^j(1-p+n) - \theta, 2^j(p+n) - \theta] \cap \mathbb{Z} \neq \emptyset\}.$$

■

Summarizing, up to now we have introduced Daubechies wavelets which are an orthonormal basis of $\mathscr{L}^2(\mathbb{R})$ and $\mathscr{L}^2(\mathbb{S}^1)$. With this methodology in mind, we can derive some properties of the Haar wavelet easily. Indeed, the Haar wavelet's mother wavelet function $\psi(x)$ (over $\mathbb{R}$) is given by

$$\psi(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2}, \\ -1 & \frac{1}{2} \leq x < 1, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \phi(x) = \begin{cases} 1 & 0 \leq x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

stands for its scaling function $\phi(x)$. One of disadvantages of the Haar wavelet is that it is not continuous but, this lack of continuity can be an advantage for the study of

functions with sudden transitions. Such wavelet verifies Definition 1.1.1 (for $\phi(x)$) and Theorems 1.1.5 and 1.1.9 with

$$h[n] = \begin{cases} \frac{1}{2} & n = 0, 1, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, we can apply Theorem 1.1.13 and Proposition 1.1.14 to get the following Corollary which will be very useful to determine the wavelet value over $\mathbb{S}^1$ of the Haar basis.

Corollary 1.1.17. *Let $\{\psi_{j,n}(x)\}$ be the standard Haar basis given by Equation (1.7). Then ℓ is equal to zero and for all $j > 0$ we have that $\psi_{j,n}^{\mathrm{PER}}(\theta) = \psi_{j,n}(x)$.*

Now, in this quick tour, we want to go a little bit further since we will need more technical results on wavelets setting since Daubechies wavelets have a "*good behavior*" in a certain functional spaces: Besov spaces. Such spaces, roughly speaking, generalizes the order and the "*strength*" of differentiability of a function. Also, a concrete space where the usage of Haar wavelet makes sense will be introduced.

1.2 Regularity through Besov spaces

In this section we will describe, in two steps, the functional spaces that define the notion of regularity in a more general setting since we have to deal what it is called "*non-positive regularities*". The framework to define these regularity values is given by the Besov spaces (see [BL76, Tri83]). First, following [Tri83], we start by defining Besov spaces on the real line. Next, as we have done along this memory, we will recall the extension of such definition to $\mathbb{S}^1$.

1.2.1 Basic notions

The space of all real valued rapidly decreasing infinitely differentiable functions is called the *(real) Schwartz space* and it is denoted by $\mathcal{S}(\mathbb{R})$. The topological dual of $\mathcal{S}(\mathbb{R})$ is the space of *tempered distributions* which we will denote it by $\mathcal{S}'(\mathbb{R})$. Also $\mathcal{C}_{\mathrm{c}}(\mathbb{R})$ will denote the space of continuous functions with compact support having continuous derivatives of any order. In such setting, for $f \in \mathcal{S}'(\mathbb{R})$, $\widehat{f}(\xi)$ denotes the *Fourier transform* of f and $f^{\vee}(x)$ stands for the *inverse Fourier transform* (with the usual integrals). Recall, also, that the essential supremum is defined as

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}} f(x) = \inf\{a \in \mathbb{R} : \mu(\{x \in \mathbb{R} : f(x) > a\}) = 0\},$$

where μ is a measure (in our case the usual Lebesgue measure). Let $\varphi_0 \in \mathcal{S}(\mathbb{R})$ be such that

$$\varphi_0(x) := \begin{cases} 1 & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| \geq 3/2 \end{cases}$$

and set

$$\varphi_j(x) := \varphi_0(2^{-j}x) - \varphi_0(2^{-j+1}x)$$

for $j \in \mathbb{N}$. It verified that, independently of the choice of φ_0, $\sum_{j=0}^{\infty} \varphi_j(x) = 1$ for all $x \in \mathbb{R}$. Each of the families $\{\varphi_j\}_{j=0}^{\infty}$ is called a *Dyadic Resolution of Unity* in $\mathbb{R}$.

Definition 1.2.1. Let $\varphi = \{\varphi_j\}_{j=0}^{\infty}$ be a Dyadic Resolution of Unity and $s \in \mathbb{R}$. For $f \in \mathcal{S}'(\mathbb{R})$ we define the quasi-norm

$$\|f\|_{\infty,\infty,\varphi,s} = \sup_{j\geq 0} 2^{js} \left(\operatorname*{ess\,sup}_{x\in\mathbb{R}} \left| (\varphi_j \widehat{f})^{\vee}(x) \right| \right).$$

Then, we define the *Besov Spaces* by

$$\mathscr{B}^s_{\infty,\infty}(\mathbb{R}) := \left\{ f \in \mathcal{S}'(\mathbb{R}) : \ \|f\|_{\infty,\infty,\varphi,s} < \infty \right\}.$$

■

As it can be seen in [Tri83, Remark 2 of Section 2.3], the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ are, in fact, independent of the chosen dyadic resolution of unity φ. Therefore, we can remove the subscript φ from $\|f\|_{\infty,\infty,\varphi,s}$. So, in what follows we will write $\|f\|_{\infty,\infty,s}$ instead of $\|f\|_{\infty,\infty,\varphi,s}$. The spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ are a particular case of the Generalized Besov Spaces $\mathscr{B}^s_{p,q}(\mathbb{R})$ defined also, for example, in [Tri83] and one has the inclusion property. That is, if $s < s'$, then $\mathscr{B}^{s'}_{p,q}(\mathbb{R}) \subset \mathscr{B}^s_{p,q}(\mathbb{R})$. Let us go a little bit in-depth. Define the quasi-norm $\|f\|_{p,q,s}$ by

$$\left(\sum_{j=0}^{\infty} 2^{jsq} \left(\|(\varphi_j \widehat{f})^{\vee}\|_p \right)^q \right)^{1/q},$$

The *Generalized Besov Spaces* are defined as follows

$$\mathscr{B}^s_{p,q}(\mathbb{R}) = \{ f \in \mathcal{S}'(\mathbb{R}) \colon \|f\|_{p,q,s} < \infty \}.$$

For the special case of $p = q = 2$, the above definition coincides with the usual *Sobolev space*. That is, define the norm

$$\|f\|_{k,2} = \left(\sum_{i=0}^{k} \|f^{i)}\|_p^p \right)^{1/p}$$

and define the Sobolev space to be

$$\mathscr{H}^k := \left\{ f \in \mathcal{S}'(\mathbb{R}) : \ \|f\|_{k,s} < \infty \right\}.$$

Notice that in the literature there are several ways to denote the aforementioned Sobolev space such as $\mathscr{W}^{k,2}$. Such notation is due to the fact that Sobolev spaces must be understood as the all locally summable functions such that its weak derivative is in $\mathscr{L}^p$. In view of that, and for sake of simplicity, when we want to refer it will use $\mathscr{B}^s_{2,2}(\mathbb{R})$. Finally, for completeness, let us recall the following definition

Definition 1.2.2. Given an integer $k \geq 1$, we define *distributional derivative of order* k *of* $f \in \mathscr{L}^1_{\mathrm{loc}}$ to be the linear functional

$$\delta_{D^k_f}(\rho) = (-1)^k \int_{\mathbb{R}} f(x) D^k \rho(x)\, dx.$$

If there exists a locally integrable function g

$$\int_{\mathbb{R}} g\rho(x)\, dx = (-1)^k \int_{\mathbb{R}} f(x) D^k \rho(x)\, dx \text{ for all } \rho \in \mathcal{C}_{\mathrm{c}}(\mathbb{R})$$

then we say that g is the *weak derivative of order* k of f. In other words, g is the weak derivative of order k of f if $\delta_{D^k_f} = \delta_g$ for all continuous functions with compact having continuous derivatives of every order ρ. ∎

On the other side, for $s > 0$ the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ coincide with the Hölder-Zygmund (or Lipschitz) spaces and it is natural to extend the notion of regularity to $s \leq 0$ through $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ in the above way (we refer to [Ste70, Tri83] for a more complete explanation). However, before we define what will be our notion of regularity let us make a final comment concerning on the inclusions of spaces. What we have set to be an extension of a Hölder space $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ it does not coincide with the space of continuous and differentiable functions, $\mathcal{C}^1$. That is, we have the following inclusion

$$\mathcal{C}^1(\mathbb{R}) \subset \mathscr{B}^s_{\infty,\infty}(\mathbb{R}).$$

Moreover, $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ either coincides with the functions that satisfy the Lipschitz condition (Hölder exponent equals 1), Lip. That is, as before we have

$$\mathcal{C}^1 \subset \mathrm{Lip} \subset \mathscr{B}^1_{\infty,\infty}$$

since for example $|x| \notin \mathcal{C}^1$ but $|x| \in \mathrm{Lip}$ and, on the other side $\sqrt{x} \in \mathscr{B}^1_{\infty,\infty}$ but $\sqrt{x} \notin \mathrm{Lip}$ because it does not have bounded derivative around the origin. In view of that, we are allowed to give the following definition.

Definition 1.2.3. We say that a map f *has regularity* $s \in \mathbb{R}$ if $f \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$. ∎

Example 1.2.4. The following examples help us to clarify this regularity notion.

(a) $\cos(2\pi x)$ is an analytic function whereas $|\cos(2\pi x)|$ is only a Lipschitz since it is differentiable almost everywhere and its derivative it is bounded. Moreover an upper semi-continuous function does not verify a s-Hölder condition on the jump discontinuities for $s \in (0, 1)$.

(b) Let us recall that the Hölder exponent for a functions $f \in \mathscr{L}^\infty$ one has the "*usual*" definition

$$\sup_{|y|>0} \left(\sup_{x \in \mathbb{R}} \frac{|f(x+y) - f(x)|}{|y|^\alpha} \right).$$

Now define,

$$f(x)=\begin{cases}0 & \text{if } x=0,\\ \sqrt{x} & \text{if } x\in(0,1/2],\\ \sqrt{1-x} & \text{if } x\in(1/2,1),\\ 0 & \text{if } x=1.\end{cases}$$

Such a function verifies, in a neighbourhood of $x=0$ or $x=1$, the Hölder condition for $0<\alpha\leq 1/2$. This assertion follows from the unimodality of $f(x)$ and therefore $f(x+y)\leq f(x)+f(y)$. Hence,

$$\begin{aligned}\sup_{|y|>0}\left(\sup_{x\in\mathbb{R}}\frac{|f(x+y)-f(x)|}{|y|^{\alpha}}\right) &= \sup_{|y|>0}\left(\sup_{x\in\mathbb{R}}\frac{|\sqrt{x+y}\cdot(1-\sqrt{x+y})}{|y|^{\alpha}}-\right.\\ &\quad\left.\frac{\sqrt{x}\cdot(1-\sqrt{x})|}{|y|^{\alpha}}\right)\\ &\leq \sup_{|y|>0}\left(\sup_{x\in\mathbb{R}}\frac{|\sqrt{x}\cdot(1-\sqrt{x})}{|y|^{\alpha}}+\right.\\ &\quad\left.\frac{\sqrt{y}\cdot(1-\sqrt{y})-\sqrt{x}\cdot(1-\sqrt{x}))|}{|y|^{\alpha}}\right)\\ &\leq \sup_{|y|>0}\left(\frac{|\sqrt{y}\cdot(1-\sqrt{y})|}{|y|^{\alpha}}\right)\end{aligned}$$

and the assertion follows.

(c) Moreover, one can prove that if f verifies a Hölder condition α, then its primitive has exponent $\alpha+1$. On the contrary, it is not true that f' verifies a Hölder condition $\alpha-1$.

(d) A Hölder continuous non-differentiable function belongs to $\mathscr{B}^{s}_{\infty,\infty}(\mathbb{R})$ with $s\in(0,1)$. For instance, the *Weierstraß function* defined by

$$\mathfrak{W}_{A,B}(x):=\sum_{n=1}^{\infty}A^{n}\sin(B^{n}x), \tag{1.8}$$

where $A,B\in\mathbb{R}$ are such that $B^{-1}<A<1<B$ is Hölder continuous and nowhere differentiable. Moreover, it has regularity $-\log_{B}(A)$; that is: from our point of view $\mathfrak{W}_{A,B}\in\mathscr{B}^{-\log_{B}(A)}_{\infty,\infty}(\mathbb{R})$. Indeed, there exists $C>0$ such that

$$|\mathfrak{W}_{A,B}(x)-\mathfrak{W}_{A,B}(y)|\leq C|x-y|^{-\frac{\log A}{\log B}}$$

holds for all $x,\ y\in\mathbb{R}$. Observe that $|\sin(B^{n}x)-\sin(B^{n}y)|\leq 2$ and also, by the Mean Value Theorem, one has that

$$|\sin(B^{n}x)-\sin(B^{n}y)|\leq|\cos(B^{n}\xi)|B^{n}|x-y|\leq B^{n}|x-y|,$$

where ξ lies between x and y. Now fix $N = \left[\frac{-\log(|x-y|)}{\log(B)}\right]$ and notice that

$$B|x-y| \leq B^2|x-y| \leq \ldots \leq B^N|x-y| \leq 1 < B^{N+1}|x-y|.$$

Therefore, by the choice of N,

$$\begin{aligned}
|\mathfrak{W}_{A,B}(x) - \mathfrak{W}_{A,B}(y)| &= \left|\sum_{n=1}^{\infty} A^n \sin(B^n x) - \sum_{n=1}^{\infty} A^n \sin(B^n y)\right| \\
&= \left|\sum_{n=1}^{N} A^n(\sin(B^n x) - \sin(B^n y)) \right. \\
&\quad \left. + \sum_{n=N+1}^{\infty} A^n(\sin(B^n x) - \sin(B^n y))\right| \\
&\leq \sum_{n=1}^{N} A^n B^N |x-y| + 2\sum_{n=N+1}^{\infty} A^n \\
&\leq B^N|x-y|\frac{A - A^{N+1}}{1-A} + 2\frac{A^{N+1}}{1-A} \\
&\leq \frac{A - A^{N+1}}{1-A} + 2\frac{A^{N+1}}{1-A} \\
&\leq CA^N,
\end{aligned}$$

for an appropriate C. Moreover, since

$$A^N = e^{N\log A} = e^{N\log B \frac{\log A}{\log B}} = (B^N)^{\frac{\log A}{\log B}} \leq \left(\frac{1}{|x-y|}\right)^{\frac{\log A}{\log B}},$$

it follows that

$$|\mathfrak{W}_{A,B}(x) - \mathfrak{W}_{A,B}(y)| \leq C|x-y|^{-\frac{\log A}{\log B}}$$

as we wanted to show.

(e) The function

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{-1}{\log(x)} & \text{if } x \in (0, 1/2], \\ \frac{-1}{\log(1-x)} & \text{if } x \in (1/2, 1), \\ 0 & \text{if } x = 1 \end{cases}$$

does not satisfy the Hölder condition for any $\alpha > 0$ on neighborhoods of 0 and 1 whereas on the rest of the interval is Lipschitz. The proof, for $x = 0$, is done by contradiction, if it were an $\alpha > 0$, then there would exist $C > 0$ and α such that $|0 - \frac{1}{\log x}| \leq C|x|^\alpha$ for all $x \in (0, 1/2]$. But, arranging the terms, one also

has that $C|x|^{\alpha}|\log(x)| \geq 1$. And this is impossible since for $\alpha > 0$ one has that $\lim_{x\to 0^+} C|x|^{\alpha}|\log(x)| = 0$. In particular $f(x) \in \mathscr{B}^{0}_{\infty,\infty}$.

(f) As a matter of fact, [RS96, Section 2.3, Example 1] is devoted to give examples of functions on such kind of spaces. Indeed, the prototypical example is to consider $\alpha^2 + \beta^2 > 0$, with $\beta > 0$, and define

$$f_{\alpha,\beta}(x) = \upsilon(x)|x|^{\alpha}(-\log|x|)^{-\beta}$$

where $\upsilon(x)$ is a smooth cut-off function with $\operatorname{supp}\upsilon \subset \{x \in \mathbb{R} : |x| \leq \delta\}$ and $\delta > 0$. That is, $\upsilon(x)$ has the support near the origin and the singularity is located near the origin.

(g) It is known that $\delta(x) \in \mathscr{B}^{-1}_{\infty,\infty}(\mathbb{R})$ where $\delta(x)$ stands for Dirac's delta function. In view of that, $\delta(x)$ can be considered as the second derivative of the continuous function

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } x \geq 0. \end{cases}$$

■

Remark 1.2.5. As it can be seen in Proposition 6 from [Ste70, Chapter 4.4], for $0 < s < 1$ every $f \in \mathscr{B}^{s}_{\infty,\infty}(\Omega)$ may be modified on a set of measure zero so that it becomes continuous. We want to emphasize that this can be done by how the $\|f\|_{\infty}$ is chosen. That is, since we are defining the spaces in terms of a measure the "usual" condition of continuity, in terms of ε, can fail on a zero measure set. ■

Recalling that we will have to deal with functions that live in $\mathbb{S}^1$, we need to extend the Besov spaces to $\mathbb{S}^1$ and we will do it following [BL76, Ste70]. Indeed, given $f \in \mathcal{S}'(\mathbb{S}^1)$ (the space of tempered distributions on $\mathbb{S}^1$) it is known that

$$f = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{inx}.$$

Definition 1.2.6. Let $\varphi = \{\varphi_j\}_{j=0}^{\infty}$ be a dyadic resolution of unity (on $\mathbb{R}$). We define the *Besov Spaces* on $\mathbb{S}^1$ by

$$\mathscr{B}^{s}_{\infty,\infty}(\mathbb{S}^1) := \left\{ f \in \mathcal{S}'(\mathbb{S}^1) : \|f\|_{\infty,\infty,s} < \infty \right\}$$

where

$$\|f\|_{\infty,\infty,s} = \sup_{j\geq 0} 2^{js} \left(\operatorname*{ess\,sup}_{x\in\mathbb{R}} \left| \sum_{n\in\mathbb{Z}} \varphi_j(n)\widehat{f}(n)e^{inx} \right| \right)$$

is a quasi-norm for the quasi-Banach space $\mathscr{B}^{s}_{\infty,\infty}(\mathbb{S}^1)$. ■

As in Definition 1.2.3 we say that a circle map f *has regularity* $s \in \mathbb{R}$ if the function is in $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$. The following lemma shows that the regularity of a circle map coincides with the regularity of its real extension which we define as follows. Given $f \in \mathcal{S}'(\mathbb{S}^1)$ there exists a unique $f^{\mathrm{PER}} \in \mathcal{S}'(\mathbb{R})$ such that f^{PER} is 1-periodic and the restriction of f^{PER} over $[0,1)$ coincides with f (such an f^{PER} can be defined as $f(\{\cdot\})$, where $\{\cdot\}$ denotes the fractional part function). This lemma is usually omitted and used implicitly but we include here for completeness.

Lemma 1.2.7. *For every $f \in \mathcal{S}'(\mathbb{S}^1)$ it follows that $f^{\mathrm{PER}} \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ if and only if $f \in \mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$.*

Proof. Since f^{PER} is 1-periodic and $f^{\mathrm{PER}}\big|_{[0,1]} = f$,

$$\widehat{f^{\mathrm{PER}}}(n) = \int_0^1 f^{\mathrm{PER}}(x) e^{-2\pi i n x}\, dx = \int_0^1 f(x) e^{-2\pi i n x}\, dx = \widehat{f}(n).$$

Hence,

$$\begin{aligned}\operatorname*{ess\,sup}_{x\in\mathbb{R}} \left| \sum_{n\in\mathbb{Z}} \varphi_j(n) \widehat{f}(n) e^{inx} \right| &= \operatorname*{ess\,sup}_{x\in\mathbb{R}} \left| \sum_{n\in\mathbb{Z}} \varphi_j(n) \widehat{f^{\mathrm{PER}}}(n) e^{inx} \right| \\ &= \operatorname*{ess\,sup}_{x\in\mathbb{R}} \left| \sum_{n\in\mathbb{Z}} (\varphi_j \widehat{f^{\mathrm{PER}}})(n) e^{inx} \right| = \operatorname*{ess\,sup}_{x\in\mathbb{R}} \left| (\varphi_j \widehat{f^{\mathrm{PER}}})^\vee(x) \right|.\end{aligned}$$

That is, $\|f^{\mathrm{PER}}\|_{\infty,\infty,s} = \|f\|_{\infty,\infty,s}$ and, hence, we have that $f^{\mathrm{PER}} \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ if and only if $f \in \mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$. □

A classical tool for measuring the fact of $f \in \mathscr{B}^s_{\infty,\infty}$, either $\mathbb{R}$ or $\mathbb{S}^1$, is to look at the asymptotic decay of its Fourier transform but, also, with wavelets can be done. This will be the main topic of the following section.

1.2.2 Wavelets and regularity

Until now we have recalled the notion of the regularity of a function through the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ and $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$. Also, we have introduced the wavelet expansions of a given function in $\mathscr{L}^2(\mathbb{R})$. Next, we want to show the relationship between this notion of regularity and the wavelet coefficients. The main tool for this will be the Daubechies wavelets, because, as we have said on Section 1.1, they are orthonormal bases on $\mathscr{L}^2(\mathbb{R})$ (see [Mal98]) and, depending on the number of vanishing moments, they are well adapted to the functional spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ (see [HW96, Tri06]).

Recall that the Daubechies wavelets are a family of wavelets with compact support that has an element with p vanishing moments for each $p \geq 1$ (see Section 1.1.2 and

[Mal98] for a definition and construction respectively). In order to motivate the following results, let us consider f to be such that verifies a Hölder condition. Then, since a wavelet has zero mean we have

$$\left|\int f(x)\psi_{j,n}(x)\,dx\right| = \left|\int f(x)\psi_{j,n}(x)\,dx - \int f(y)\psi_{j,n}(x)\,dx\right| \leq \int |x-y|^{\alpha}|\psi_{j,n}(x)|\,dx.$$

That is, using a "*good enough*" wavelets one can describe scaling properties of the function f and the converse i also true. In view of that and from [HW96, Theorem 7.16], [Coh03, Theorem 3.8.1], [Tri06, Theorem 1.64] we can state the following theorem, in the spirit of [dlLP02, Theorem 5.10] and [Mey01, Chapter 3].

Theorem 1.2.8. *Let $s \in \mathbb{R}\setminus\{0\}$ and let ψ be a mother Daubechies wavelet with more than $\max(s, 5/2-s)$ vanishing moments. Then $f \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ if and only if there exists $C>0$ such that*

$$\sup_{n\in\mathbb{Z}} |\langle f, \psi_{j,n}\rangle| \leq C2^{\tau j} \qquad \textit{with} \qquad \tau = \begin{cases} s+\frac{1}{2} & \textit{if } s>0, \\ s-\frac{1}{2} & \textit{if } s<0, \end{cases}$$

for all $j \leq 0$.

Remark 1.2.9. Since the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ satisfy the inclusion property described on Section 1.2.1, $f \in \mathscr{B}^0_{\infty,\infty}(\mathbb{R})$ if and only if either there exist $C>0$ and $\tau \in [-\frac{1}{2}, \frac{1}{2}]$ such that $\sup_{n\in\mathbb{Z}} |\langle f, \psi_{j,n}\rangle| \leq C2^{\tau j}$ for $j \leq 0$ or the numbers $\sup_{n\in\mathbb{Z}} |\langle f, \psi_{j,n}\rangle|$ do not have an upper bound of the form $C2^{\tau j}$ with $C>0$ and $\tau \in \mathbb{R}$. ∎

Remark 1.2.10. From [Mal98, Section 6.4 (Chapter 6)] we know that when f is self-similar then so it is the wavelet transform. Moreover, the numbers $\sup_{n\in\mathbb{Z}} |\langle f, \psi_{j,n}\rangle|$ should decay exponentially with respect to j. That is,

$$s_j := \log_2\left(\sup_{n\in\mathbb{Z}} |\langle f, \psi_{j,n}\rangle|\right) = \tau j + \log_2(C).$$

This tells us that, in this case, to compute the value of regularity s we can make a linear regression to estimate the slope of the graph of the pairs (j, s_j) and get the correct value of s from this slope. The Pearson correlation coefficient controls the degree of linear correlation between the variables j and s_j and, hence, it is a test of the self-similarity of f (and so, of the validity of the estimated s). ∎

The above theorem and remarks motivate the following definition and rewriting of Theorem 1.2.8 in view of [Mey01, Chapter 3]. For $t \in \mathbb{R}$ we set

$$\mathrm{R}(t) = \begin{cases} t-\frac{1}{2} & \text{if } t>\frac{1}{2}, \\ t+\frac{1}{2} & \text{if } t<-\frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 1.2.11. *Let $f \in \mathscr{L}^2(\mathbb{R})$ and let ψ be a mother Daubechies wavelet with k vanishing moments. Assume that there exists $C > 0$ and $\tau \in \mathbb{R}$ such that*

$$\sup_{n\in\mathbb{Z}} |\langle f, \psi_{j,n}\rangle| \approx C2^{\tau j}$$

for all $j \leq 0$. Then, $f \in \mathscr{B}^{\mathrm{R}(\tau)}_{\infty,\infty}(\mathbb{R})$ provided that $k > \max(\mathrm{R}(\tau), \frac{5}{2} - \mathrm{R}(\tau))$.

Remark 1.2.12. In view of Lemma 1.2.7, to estimate the regularity of an $f \in \mathcal{S}'(\mathbb{S}^1)$ it is enough to use Theorem 1.2.11 for f^{PER}. ∎

Observe that the above Theorem is not valid for Haar wavelet since it does not verify the property of $k > \max(\mathrm{R}(\tau), \frac{5}{2} - \mathrm{R}(\tau))$. What it is done in the following section is a small survey of how it can be relaxed such hypotheses.

1.2.3 Haar basis revised

As we have said, Haar wavelet does not have enough vanishing moments to be a basis of $\mathscr{B}^s_{\infty,\infty}$ with $s > 0$ and $s < -1$. Such wavelet can be readjusted, or turned to the Schauder basis[1], to be a desired basis of the Besov space. The problem is that for our forthcoming purposes this way is not useful. Therefore, we need to work a little bit more in order to be able to use the Haar wavelet intact.

To this end, we need to introduce some definitions which, in fact, are implicit in the previous lines. In what follows a *rearrangement* of $\mathbb{N}$ will be a one-to-one self map from $\mathbb{N}$ to itself.

Definition 1.2.13. Let $\mathscr{B}$ be a real (quasi)-Banach space. A basis $\{b_j\}_{j=1}^{\infty}$ of $\mathscr{B}$ is called *unconditional* if for any rearrangement σ of $\mathbb{N}$ the resultant, $\{b_{\sigma(j)}\}_{j=1}^{\infty}$, is again a basis of $\mathscr{B}$. Also for every $\{\lambda_j\}_{j=1}^{\infty} \subset \mathbb{R}$ it follows that

$$\sum_{j=1}^{\infty} \lambda_{\sigma(j)} b_{\sigma(j)} = \sum_{j=1}^{\infty} \lambda_j b_j.$$

Moreover, let the sequence $\{\lambda_j\}_{j=1}^{\infty}$ be such that $\lambda_j \in \mathscr{B}$ for all $j \in \mathbb{N}$. Then the series $\sum \lambda_j$ is called *unconditionally convergent* if for any rearrangement σ of $\mathbb{N}$ the series $\sum \lambda_{\sigma(j)}$ is convergent in $\mathscr{B}$. ∎

With such Definition, we can define the discrete version of the space $\mathscr{B}^s_{p,q}$ which is a (quasi-)Banach space. We will denote it by $\wp^s_{p,q}$.

[1]The Schauder basis (or the Faber-Schauder system) is the integral of the Haar basis. That is, for $n = 0, \ldots, 2^j - 1$ define

$$s_{j,n}(x) = 2^{1+j/2} \int_0^x \psi_{j,n}(\eta)\, d\eta$$

where $\{\psi_{j,n}\}_{\mathbb{Z}^+\times\mathbb{Z}}$ is the usual Haar basis.

Definition 1.2.14. Let $0 < p \leq \infty$, $0 < q \leq \infty$, $s \in \mathbb{R}$ and let $\{\lambda_{j,n}\}_{(j,n)\in\mathbb{Z}^+\times\mathbb{Z}}$ be a bi-indexed sequence of reals. Consider the norm

$$\|\lambda\|_{p,q,s} = \left(\sum_{j=0}^{\infty} 2^{j(s-1/p)q}\left(\sum_{n\in\mathbb{Z}}|\lambda_{j,n}|^p\right)^{q/p}\right)^{1/q}$$

with the usual modification if $p = \infty$ and/or $q = \infty$. In such conditions we define the *Discrete Generalized Besov Spaces* by

$$\wp^s_{p,q}(\mathbb{R}) = \{\lambda\colon \|\lambda\|_{p,q,s} < \infty\}.$$

■

Once we have introduced all this stuff related to the aforesaid functional spaces $\mathscr{B}^s_{p,q}$, we can return back to the Haar wavelet. Recall that the standard Haar basis gives rise to a simple basis in $\mathbb{S}^1$ just by considering to be $j > 0$. Having said that, and as in Theorem 1.2.11, we have the following characterization, in terms of Haar coefficients, of the belonging property on a concrete $\mathscr{B}^s_{p,q}$ of a function f.

Theorem 1.2.15 (Theorem 2.13 [Tri10]). *Let $\{\psi_{j,n}(x)\}_{\mathbb{Z}^+,\mathbb{Z}}$ be the standard Haar basis given by Equation (1.7). Consider $0 < p \leq \infty, 0 < q \leq \infty$ and $1/p - 1 < s < \min(1, 1/p)$. Let $f \in \mathcal{S}'([0,1])$. Then $f \in \mathscr{B}^s_{p,q}([0,1))$ if, and only if, it can be represented as*

$$f = \mu_0\phi + \sum_{j=0}^{\infty}\sum_{n=0}^{2^j-1}\mu_{j,n}2^{-j(s-1/p)}\psi_{j,n}$$

with $\mu_{j,n} \in \wp^s_{p,q}(\mathbb{R})$ and unconditional convergence being in $\mathscr{B}^\sigma_{p,q}$ with $\sigma < s$. The above representation is unique and given by

$$\begin{cases}\mu_0 = \int_0^1 f(x)\phi(x)\,dx,\\ \mu_{j,n} = 2^{j(s-1/p+1)}\int_0^1 f(x)\psi_{j,n}(x)\,dx, \quad j > 0 \text{ and } n = 0,\ldots,2^j-1.\end{cases}$$

If, in addition, $p < \infty$ and $q < \infty$ then the Haar basis is an unconditional basis of $\mathscr{B}^s_{p,q}([0,1])$.

That is, in our case we have that the Haar basis is a basis of $\mathscr{B}^s_{p,q}(\mathbb{S}^1)$ but when $p = q = 2$ will be, for us, the most interesting case. Indeed, in the same way of Theorem 1.2.8, we can use the decay of the Haar coefficients of a function f to relate them with the regularity, from the Sobolev point of view, of such function f. Let us state such comments in a more formal way.

Corollary 1.2.16. *Let $\{\psi_{j,n}(x)\}_{\mathbb{Z}^+,\mathbb{Z}}$ be the standard Haar basis given by Equation (1.7) and consider $s \in (-1/2, 1/2)$. Then $f \in \mathscr{B}^s_{2,2}(\mathbb{S}^1)$ if and only if there exists a constant $C > 0$ such that*

$$\sup_{0 \leq n \leq 2^j - 1} |\langle f, \psi_{-j,n} \rangle| < C 2^{-j(s-1/2)}$$

for all $j \geq$, where

$$\langle f, \psi_{-j,n} \rangle = 2^{j(s-1/2+1)} \int_0^1 f(x) \psi_{j,n}(x)\, dx, j > 0 \text{ and } n = 0, \ldots, 2^j - 1.$$

From Theorem 1.2.15, in the same way as in Remark 1.2.10 and Theorem 1.2.11, we have the appropriate rewriting of R for the Haar's case.

$$\mathrm{R}(\tau) = \tau - 1 + \frac{1}{p}, \tag{1.9}$$

where τ is the slope of the linear model. Concretely, in the case of the above corollary the formula becomes $\mathrm{R}(\tau) = \tau - 1/2$.

Coming to the end of this section, let us illustrate the idea of $\mathscr{B}^s_{2,2}$ with some instances in the same spirit of Example 1.2.4.

Example 1.2.17. The following examples help us to clarify the notion of Sobolev spaces and weak derivative.

(a) Consider the function

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } x \geq 0. \end{cases}$$

Then the *Heaviside function*

$$H(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

is the weak derivative of f. Indeed, the distributional derivative of f is

$$\delta_{D^k_f}(\rho) = -\int_0^\infty x\phi'(x)\, dx = -x\phi(x)|_{x=0}^{x=\infty} + \int_0^\infty \phi(x)\, dx = \int_{\mathbb{R}} H(x)\, dx.$$

(b) The function

$$f(x) = \begin{cases} x & \text{if } 0 \leq x < 1, \\ 2 & \text{if } 1 \leq x \leq 2. \end{cases}$$

does not have weak derivative. Indeed, let us assume that there exists $g \in \mathscr{L}^1_{\mathrm{loc}}(0,2)$ such that for all $\rho \in \mathcal{C}_{\mathrm{c}}(0,2)$

$$\int_0^2 g(x)\phi'(x)\, dx = -\int_0^2 g'(x)\phi(x)\, dx.$$

Then,

$$\begin{aligned}\int_0^2 g(x)\phi(x)\,dx &= -\int_0^1 x\phi'(x)\,dx - 2\int_1^2 \phi'(x)\,dx \\ &= \int_0^1 \phi(x)\,dx - x\phi(x)\big|_{x=0}^{x=1} + 2\phi(1) \\ &= \int_0^1 \phi(x)\,dx + \phi(1).\end{aligned}$$

To get the contradiction, let us take a sequence of $\phi_k(x) \in \mathcal{C}_c(0,2)$ such that for all positive integer k $\phi_k(1) = 1$ and $\phi_k(x) = 0$ as k goes to infinity. Then

$$1 = \lim_{k\to\infty} \phi_k(1) \neq \lim_{k\to\infty} \int_0^2 g(x)\phi_k(x)\,dx - \int_0^1 \phi_k(x)\,dx = 0.$$

This means that discontinuous functions does not have a weak derivative.

(c) The ternary Cantor function does not have a weak derivative on $\mathbb{R}$. To prove the assertion, let g to be the supposed weak derivative. Then, since the Cantor function is constant over the intervals, $g(x) = 0$ for all $x \in \mathbb{R}$. On the other side, take $\phi(x) \in \mathcal{C}_c(\mathbb{R})$ such that $\phi(x) = 1$ if $x \in [0,1]$ and $\phi(x) = 0$ for $x \geq a > 1$. We can join the ingredients to get a contradiction:

$$0 = \int_{\mathbb{R}} g(x)\phi(x)\,dx = -\int_{\mathbb{R}} f(x)\phi'(x)\,dx = -\int_0^1 f(x)\cdot 0\,dx - \int_1^a 1\cdot\phi'(x)\,dx = 1.$$

(d) Let $(a,b) \subset \mathbb{R}$ be an open interval and assume that $g \in \mathscr{L}^1_{\mathrm{loc}}(a,b)$ is the weak derivative of $f \in \mathscr{L}^1_{\mathrm{loc}}(a,b)$. Then there exists and absolutely continuous function $\tilde{f}$ such that for all $x \in (a,b)$ $\tilde{f}(x) = f(x)$ almost everywhere and

$$g(x) = \lim_{h\to 0} \frac{\tilde{f}(x+h) - \tilde{f}(x)}{h},$$

also for all $x \in (a,b)$. This means that each element of $\mathscr{B}^1_{2,2}(a,b)$ coincides almost everywhere with an absolutely continuous function f having derivative $f' \in \mathscr{L}^2(a,b)$.

■

In order to conclude this chapter let us make a final comment concerning to the wavelet and Haar coefficients. As we have said, wavelets are well posed to detect and characterize scaling behavior. But moreover, Theorem 1.2.11 gives us a method (see Section 4.1) to decide the quality of the coefficients $\langle f, \psi_{j,n}\rangle$. Or, in a less ambitious way, give a kind of intuition of how weird is the function f. The problem is how one can get the desired coefficients $\langle f, \psi_{j,n}\rangle$ and this topic will be faced in the following chapter. Of

course, due to Corollary 1.2.16, the same comments are also valid with the Haar basis. However, as we will see in Chapter 5 and we have already seen, when one have to deal with the Haar basis there are some advantages in the calculation but not in the Besov regularity.

Chapter 2

On the Computation of Wavelet Coefficients

Up to now, we have introduced orthonormal wavelet bases by means of the Multiresolution Analysis. Also, we have seen how from such structure appears the scaling filter which, recall that, plays an important role on how the wavelet is. Let us point out that we want to calculate

$$\langle f, \psi_{-j,n} \rangle = \int_{\mathrm{supp}(\psi_{-j,n}(x))} f(x)\psi_{-j,n}(x)\, dx$$

since if one wants to make approximations of a function from $\mathscr{L}^2$ by a wavelet expansions of the type

$$f \sim a_0 + \sum_{j=0}^{\infty} \sum_{n=0}^{2^j-1} \langle f, \psi_{-j,n} \rangle \, \psi_{-j,n}$$

the unknown is, precisely, $\langle f, \psi_{-j,n} \rangle$. The present chapter is devoted to present some feasible techniques to compute the coefficients $\langle f, \psi_{j,n} \rangle$ once we know how $h[n]$ is. Again, we will be specially focused on Daubechies wavelets (either in $\mathbb{R}$ or $\mathbb{S}^1$). This chapter must be understood as a natural continuation of Chapter 1 in the sense that, there are two sections (the first two ones in Section 2.3) which are standard in the wavelets literature. The third one is, as far as we know, a slight modifications of existent methods. Again, as in the previous chapter, we will show some specific properties of the Haar wavelet basis that will be useful in the following chapters. Moreover, the first section must be seen as the cornerstone of Chapter 5. Indeed, as we will see, we will need the Daubechies wavelet basis evaluated on a huge mesh of equally spaced points of $\mathbb{S}^1$. This implies that it is imperative to have a "*sufficiently fast*" algorithm to create such amount of data. Such topic must be taken as a one of the contributions of the present Thesis.

2.1 The wavelet value: a single point

In some methods to get the value of $\langle f, \psi_{-j,n} \rangle$ it is needed the value of the wavelet at a given point (see for example Chapter 5). If the wavelet does not have a closed

expression, as the trigonometric polynomials for example, one has to derive a method to get $\psi_{j,n}(x)$ in an efficient way. This is what it is done here. As usual, we will start with the $\mathbb{R}$ setting and after that we will state the main result of this section: the "*Daubechies - Lagarias algorithm*" on $\mathbb{S}^1$. First of all, let us mention that such algorithm is a way to get bounds of the regularity of the compactly supported wavelets (see Section 7.2 of [Dau92] to go further in such topic). In order to face the Daubechies - Lagarias algorithm let us start with a simple recursion which will derive and motivate the algorithm. This way of presentation is similar to the Section 6.3 of [SN96] but we will include it for sake of completeness.

The starting point is the inclusion property, the first property of Definition 1.1.1, with the scaling function (which leads us to Equation (1.5))

$$\frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right) = \sum_{n\in\mathbb{Z}} h[n]\phi(x-n).$$

Such equation, in the case of the Daubechies wavelets with $p > 1$ vanishing moments becomes

$$\phi(x) = \sqrt{2}\sum_{n=0}^{2p-1} h[n]\phi(2x-n).$$

Remark 2.1.1. Notice that we are forcing $p > 1$. This is because the corresponding Daubechies wavelet for $p = 1$, as we have already said, is the Haar wavelet. Such wavelet has an explicit expression and the forthcoming methodology becomes sterile. However, it still works correctly even for $p = 1$. ∎

Now, let $\Phi(0) = [\phi(0), \phi(1), \ldots, \phi(2p-1)]$ be a vector with $2p$ components, the vector of the half-integers values $\Phi(1/2) = [\phi(1/2), \phi(3/2), \ldots, \phi((2p-1)/2)]$ and define the following matrices, $\mathbf{M}_0$ and $\mathbf{M}_1$, whose entries are

$$m^0_{i,j} = \sqrt{2}h[2i-j-1] \; 1 \le i,j \le 2p-1 \text{ for } \mathbf{M}_0, \tag{2.1}$$

$$m^1_{i,j} = \sqrt{2}h[2i-j] \; 1 \le i,j \le 2p-1 \text{ for } \mathbf{M}_1. \tag{2.2}$$

Remark 2.1.2. Notice that the even columns of the matrices $\mathbf{M}_0$ and $\mathbf{M}_1$ are the even components of the coefficients $h[n]$. The odd ones also. On the other side, due to Proposition 1.1.4, the sum of the odd columns and the even columns must be one. This means that the row vector of ones is a common left eigenvector of $\mathbf{M}_0$ and $\mathbf{M}_1$ for the eigenvalue 1. ∎

With this new language we have the following translation:

$$\phi(x) = \sqrt{2}\sum_{n=0}^{2p-1} h[n]\phi(2x-n) \Leftrightarrow \Phi(x) = \mathbf{M}_0\Phi(2x) + \mathbf{M}_1\Phi(2x-1).$$

That is, the scaling equation can be written in terms of the vector $\Phi(x)$ and, hence, since from Proposition 1.1.8 $\operatorname{supp}(\phi(x)) = [0, 2p-1]$ one can check that

$$\Phi(0) = \mathbf{M}_0\Phi(0) \quad \text{and} \quad \Phi(1/2) = \mathbf{M}_1\Phi(0)$$

because $\operatorname{supp}(\Phi(x)) = [0,1]$ where, as we will see, the above expressions are a linear systems *with a* solution. Moreover, using recursively what we have claimed before and after some manipulations for $x = 3/8$ it is true that

$$\Phi(3/8) = \mathbf{M}_0\Phi(3/4) = \mathbf{M}_0(\mathbf{M}_1\Phi(1/2)) = \mathbf{M}_0(\mathbf{M}_1(\mathbf{M}_1\Phi(0))).$$

Now notice that the binary expression of $3/8$ is 0.011 where the sequence $0, 1, 1,$ is the same sequence of matrices multiplications $\mathbf{M}_0 \cdot \mathbf{M}_1 \cdot \mathbf{M}_1$. Therefore, roughly speaking, we have a recursion to get the values of Φ at some points and, by definition of Φ, we can get the values of ϕ at some points of its support also. Such procedure can be extended, as we will see in the following, to all dyadic values. Indeed, let $\{d_1, d_2, \ldots, d_m\}$ be the set of 0-1 digits in the binary representation of a dyadic point x then

$$\Phi(.d_1, d_2, \ldots, d_m) = \mathbf{M}_{d_1}\Phi(.d_2, d_3, \ldots, d_m). \tag{2.3}$$

Notice that one problem is left since we want $\psi(x)$ and using the recursion given by Equation (2.3) we still are not able to calculate the desired value. The answer is given by using, again, the inclusion property, but now with the wavelet, in the same way as Equation (1.5):

$$\psi(x) = \sum_{n\in\mathbb{Z}} g[n]\phi(2x-n)$$

since we know the values of $\phi(x)$. Of course, this kind of construction must be rigorously stated and the convergence results (i.e the pass to the limit) for an arbitrary $x \in \mathbb{R}$ should be proven. This is done, mainly, in [DL91, DL92, Dau92, Vid99]. Indeed, in [DL91, DL92] solutions of some functional equations, like the dilation equation, are done and studied. Precisely, in the studies concerned to the (local and global) regularity of $\phi(x)$ (and $\psi(x)$) a procedure to get some bounds is presented: the *Daubechies – Lagarias algorithm*. Such idea will be crucial for our purposes and we will present it, in a general framework, in what follows. Indeed, let

$$f(x) = \sum_{n=0}^{N} c[n]f(2x-n) \tag{2.4}$$

be the functional equation, which is called *refinable*, that we want to solve. To this end, if $f(x)$ is a solution of above then, according to Proposition 2.1 of [DL92], one can define

$$f(n) = \begin{cases} a_n & \text{if } 0 \le k < N-1, \\ 0 & \text{otherwise,} \end{cases}$$

where a_n are the components of a right eigenvector of a matrix, related to the coefficients $c[n]$, with eigenvalue 1. Also, define the spline approximations f_j to f by

$$\begin{cases} f_0(x) & \text{linear on every } [n, n+1], \\ f_0(n) = f(n) & \text{as above and} \\ f_j = V^j f_0, \end{cases}$$

where V is the following linear operator $(Vg)(x) = \sum_{n=0}^{N} c[n]g(2x-n)$ and V^j stands for the j-th application of such operator. The above sequence of functions, as it is shown in Theorem 4.1 of [DL91], converges with some restrictions on the coefficients $c[n]$ to the function f which is, precisely the solution wanted. In other words, finding solutions of the above refinable equation is the same as finding fixed points for the above linear operator where g it is chosen such that has the same support of f: $[0, N]$. We are almost done, because we can build the rigorous setting of what we have pointed out at the beginning of the section. Indeed, define the vector valued function $\mathbf{w} : [0,1] \longrightarrow \mathbb{R}^N$ as $\mathbf{w}(x)_n = g(x+n-1)$ for $n = 1, \ldots, N$ and define the linear operator $(\mathbf{Vw})(x)_n = (\mathbf{V}g)(x+n-1)$ which can be written in the following explicit way

$$\mathbf{Vw}(x) = \begin{cases} \mathbf{M}_0\mathbf{w}(2x) & \text{if } 0 < x \leq 1/2, \\ \mathbf{M}_1\mathbf{w}(2x-1) & \text{if } 1/2 \leq x < 1, \end{cases}$$

where $\mathbf{M}_0$ and $\mathbf{M}_1$ the matrices are given by (2.1) and (2.2) respectively.

Remark 2.1.3. Since it will be important for our implementation, we want to emphasize that $\mathbf{M}_0$ and $\mathbf{M}_1$ have eigenvalue 1 and the vector whose entries are all ones as a left eigenvector. ∎

In order to reproduce Equation (2.3), consider for any $x \in [0,1]$ its binary expansion and the well known *shift operator* τ applied to x: $\tau x = \sum_{j=2}^{\infty} d_j 2^{-j+1}$ where the numbers are $d_j = 0$ or $d_j = 1$. Such operator can be, again explicitly, written as

$$\tau x = \begin{cases} 2x & \text{if } 0 \leq x < 1/2, \\ 2x-1 & \text{if } 1/2 < x \leq 1. \end{cases}$$

In such conditions observe that $\mathbf{Vw}(x) = \mathbf{M}_{d_1(x)}\mathbf{w}(\tau x)$ and in view of that, one can derive the analogous to Equation (2.3):

$$\mathbf{v}_j(x) = (\mathbf{V}^j\mathbf{v}_0)(x) = \mathbf{M}_{d_1(x)}\mathbf{M}_{d_2(x)}\cdots\mathbf{M}_{d_j(x)}\mathbf{v}_0(\tau^j x),$$

where $\mathbf{v}_0 = f_0$. The following Theorem makes clear all the ideas outlined until now.

Theorem 2.1.4 (Theorem 2.2 [DL92]). *Assume that the $c[n]$, $n = 0, \dots, N$ satisfy*

$$\sum_{n\in\mathbb{Z}} c[2n] = \sum_{n\in\mathbb{Z}} c[2n+1] = 1$$

and consider the matrices $\mathbf{M}_0$ and $\mathbf{M}_1$ given by Equation (2.1) and Equation (2.2). Define $\mathbf{E}_1$ to be the $(N-1)$-dimensional subspace orthogonal to $\mathbf{e}_1 = (1, \dots, 1)$, the common left eigenvector of $\mathbf{M}_0$ and $\mathbf{M}_1$ for the eigenvalue 1. Assume that there exist $\lambda < 1$ and $C > 0$ such that, for all $m \in \mathbb{N}$,

$$\max_{d_i \in \mathrm{dyad}(x,m)} \|\mathbf{M}_{d_1} \cdots \mathbf{M}_{d_m}|_{\mathbf{E}_1}\| \leq C\lambda^m.$$

Then the following hold:

(a) The eigenvalue 1 is of the $(N-1)\times(N-1)$-dimensional matrix M defined by $M_{i,j} = c_{2i-j}$ $1 \leq i, j \leq N-1$, is simple and there us an associated right eigenvector $\mathbf{a}$ with $\sum_{i=1}^{N-1} a_i = 1$.

(b) The vector-valued functions $v_j(x)$ defined above satisfy $\mathbf{e}_1 \cdot \mathbf{v}_j(x) = 1$ for all $j \in \mathbb{N}$ and $x \in [0,1]$.

(c) The corresponding functions f_j converge uniformly to a continuous function f with

$$\|f_j - f\|_\infty \leq C 2^{-j\frac{\ln\lambda}{\ln 2}}.$$

(d) the limit function f is an $\mathscr{L}^1$-solution to the refinable equation (2.4); it is normalized so that $\int_0^N f(x)\,dx = 1$ and it is Hölder continuous

$$|f(x) - f(y)| \leq C|x-y|^\alpha$$

with $\alpha = \frac{\ln\lambda}{\ln 2}$.

The above Theorem is valid for a large class of refinable functions (i.e those ones such that are the solution of a refinable equation) which it is not the case of wavelets. Having said that, wavelets are close to be refinable functions since they can be obtained from the scaling function $\phi(x)$ and, hence, it is necessary to apply Theorem 2.1.4 to $\phi(x)$. This is the spirit of the following Remark: written with the goodwill of make explicit the Daubechies - Lagarias theorem for our particular instances.

Remark 2.1.5 (Theorem 3.5.4 [Vid99]). The above limit must be understood, in the case of the Daubechies wavelet with p vanishing moments, as follows:

$$\lim_{n\to\infty} \mathbf{M}_{d_1} \cdots \mathbf{M}_{d_n} = \begin{pmatrix} \phi(x) & \phi(x) & \dots & \phi(x) \\ \phi(x+1) & \phi(x+1) & \dots & \phi(x+1) \\ \vdots & \vdots & \vdots & \vdots \\ \phi(x+2p-1) & \phi(x+2p-1) & \dots & \phi(x+2p-1) \end{pmatrix},$$

where $\phi(x)$ is the scaling function of the Daubechies wavelet with p vanishing moments. ∎

Thus, taking into account the above Remark is clear how one can get the value of $\phi(x)$ at an arbitrary point but recall that we want the value of $\psi(x)$. This value is given by the following Theorem which is based on Theorem 2.1.4.

Theorem 2.1.6 (Theorem 3.5.5 [Vid99]). *Let x be an arbitrary real number, $\psi(x)$ be an $\mathbb{R}$-Daubechies wavelet with $p \geq 1$ vanishing moments given by its filter coefficients and let u be a vector defined as*

$$u_i(x) = (-1)^{1-\lfloor 2x \rfloor} h[i + 1 - \lfloor 2x \rfloor] \text{ for } i = 0, \ \ldots, \ 2p-2.$$

If for some i the index $i+1-\lfloor 2x \rfloor$ is negative or larger than $2p-1$, then the corresponding component of u is 0.
Let v the vector given by

$$v(x,n) = \frac{1}{2p-1} 1' \prod_{i \in \mathrm{dyad}(\{2x\},n)} \mathbf{M}_i$$

where $1' = (1, 1, \ldots, 1)$ is the row-vector of ones and $\{\cdot\}$ stands for the fractional part function. Then,

$$\psi(x) = \lim_{n \to \infty} u(x)' v(x,n), \tag{2.5}$$

and the limit is constructive (with effective error bounds).

Remark 2.1.7. It is important to notice that the Equation (2.5) is a direct consequence of Equation (1.5) and the convergence given by the Theorem 2.1.4 to $\phi(x)$. In other words, Theorem 2.1.6 firstly calculates $\phi(x)$ and after that uses Equation (1.5) to get $\psi(x)$. However, the expression of $v(x,n)$ has no sense because the sum of the rows of M_∞ is identically one. That is, the right version must be

$$v(x,n) = \frac{1}{2p-1} \left(\prod_{i \in \mathrm{dyad}(\{2x\},n)} \mathbf{M}_i \right) 1'^{\top}.$$

We have used the same notation because of the history. ∎

Finally, since our environment is $\mathbb{S}^1$ we need to derive a method to evaluate $\psi^{\mathrm{PER}}(x)$. This method is based on Daubechies - Lagarias algorithm, Theorem 2.1.6 and Section 1.1.3 but, we want to emphasize that our method in $\mathbb{S}^1$ can be understood as a Corollary from such theorems and is stated below but, we will reformulate it a little bit forward as a Proposition in order to derive the method from the proof.

Corollary 2.1.8. *Let x be an arbitrary point of $\mathbb{S}^1$. Then,*

$$\psi^{\mathrm{PER}}(x) = \sum_{\ell\in\mathbb{Z}} \lim_{k\to\infty} u(x+\ell)' v(x+\ell, k),$$

where $u(\cdot)$ and $v(\cdot,k)$ are given by Theorem 2.1.6.

Now we are ready to present one of the main results of this chapter which is the method to evaluate in an effective manner the Daubechies wavelet. Before the statement we want to point out that the following method is also valid for all wavelets which verifies the assumptions of Theorem 2.1.4.

Proposition 2.1.9. *Let $J > 0$ be an integer, $\psi(x)$ be an $\mathbb{R}$-Daubechies wavelet with $p > 1$ vanishing moments and θ be an arbitrary point of $\mathbb{S}^1$. Set $N = 2^J$ and define a N-dimensional row vector $\Psi^{\mathrm{PER}}(\theta)$ such that its entries are ${\psi_{-j,n}(\theta)}^{\mathrm{PER}}$, where $j = 0, 1, \ldots, J-1$ and $n = 0, \ldots 2^j - 1$. Then, for each $j = 0, 1, \ldots, J-1$*

$${\Psi_i}^{\mathrm{PER}}(\theta) = \begin{cases} 1 & \text{if } i = 0, \\ 2^{j/2} \sum_{\ell\in\Lambda_\theta} \lim_{k\to\infty} u(2^j(\theta+\ell) - n)' v(2^j\theta - n, k) & \text{if } i = 1, \ldots, N-1, \end{cases}$$

where the vectors $u(\cdot)$ and $v(\cdot,k)$ are given by Theorem 2.1.6 and the range of the translations is given by

$$n \in \aleph_\theta = \{n \in \mathbb{Z} : \ \max\left(\lceil 2^j\theta \rceil + 2^j\ell - p, 0\right) \le n \le \min\left(\lfloor 2^j\theta \rfloor + 2^j\ell - Ni, 2^j - 1\right)\}.$$

Proof. Notice that the principal difference between the Corollary and Proposition, besides of j and n, is $v(\theta+\ell, k)$ and $v(\theta, k)$. This is because $\{2^j(\theta+\ell)\} = \{2^j\theta\}$ which implies that, for a fixed θ, the products of the matrices it must be done one time and not for all ℓ's. Therefore, since the limit does not depend on ℓ it follows that

$$\lim_{k\to\infty} v(2^j(\theta+\ell) - n, k) = \lim_{k\to\infty} v(2^j\theta - n, k) = v(2^j\theta - n).$$

On the contrary, the integer ℓ affects on the calculation of $u(2^j(\theta+\ell) - n)$ because of the properties of $\lfloor\cdot\rfloor$. Once the convergence it is shown, we continue with the following expression

$$\psi^{\mathrm{PER}}_{-j,n}(\theta) = \sum_{\ell\in\Lambda_\theta} \psi_{-j,n}(\theta+\ell).$$

In order to make more readable the proof, let us define the integer number, related to the number of vanishing moments, $Ni = (1-p)$ and hence the wavelet support is $[Ni, p]$. Now, by definition of support, $Ni \le 2^j(\ell+\theta) - n \le p$. That is, recalling that $0 \le n \le 2^j - 1$

$$\frac{Ni}{2^j} \le \frac{(Ni+n)}{2^j} \le \ell + \theta \le \frac{(p+n)}{2^j} \le \frac{(p+2^j-1)}{2^j}.$$

Therefore,

$$\frac{Ni}{2^j} - \theta \leq \ell \leq \frac{p}{2^j} + 1 - \frac{1}{2^j} - \theta = \frac{1 - Ni}{2^j} + 1 - \frac{1}{2^j} - \theta = 1 - \frac{Ni}{2^j} - \theta.$$

Which leads us to $\lceil \frac{Ni}{2^j} - \theta \rceil \leq \ell \leq \lfloor 1 - \frac{Ni}{2^j} - \theta \rfloor = \lfloor -\frac{Ni}{2^j} - \theta \rfloor + 1$ and, hence,

$$\Lambda_\theta = \left\{ \ell \in \mathbb{Z} : \ \left\lceil \frac{Ni}{2^j} - \theta \right\rceil \leq \ell \leq \left\lfloor -\frac{Ni}{2^j} - \theta \right\rfloor + 1 \right\}.$$

Now, being θ and ℓ fixed one has:

$$Ni \leq 2^j(\ell + \theta) - n \leq p \Leftrightarrow 2^j(\ell + \theta) - Ni \geq n \geq 2^j(\ell + \theta) - p.$$

Therefore, since $0 \leq n \leq 2^j - 1$ it follows that

$$\lceil 2^j(\ell + \theta) - p \rceil \leq n \leq \lfloor 2^j(\ell + \theta) - Ni \rfloor \Leftrightarrow \lceil 2^j\theta \rceil + 2^j\ell - p \leq n \leq \lfloor 2^j\theta \rfloor + 2^j\ell - Ni.$$

Hence, it can be defined

$$\aleph_\theta = \{ n \in \mathbb{Z} : \ \max\left(\lceil 2^j\theta \rceil + 2^j\ell - p, 0\right) \leq n \leq \min\left(\lfloor 2^j\theta \rfloor + 2^j\ell - Ni, 2^j - 1\right) \}.$$

In order to prove the consistency of the last inequality that defines $\aleph_\theta$, notice that, since $j > 0$, $0 \leq 2^j - 1$. Then $0 \leq \lfloor 2^j\theta \rfloor + 2^j\ell - Ni$ which, therefore, implies that $0 \leq \min\left(\lfloor 2^j\theta \rfloor + 2^j\ell - Ni, 2^j - 1\right)$. On the other side, from ℓ's range we have that $2^j(\ell + \theta) \leq p + 2^j - 1$; therefore, $\lceil 2^j\theta \rceil + 2^j\ell - p \leq 2^j - 1$ hence, by definition of Ni (i.e we have $Ni < -1 - Ni$)

$$\lceil 2^j\theta \rceil + 2^j\ell - p \leq 2^j(\theta + \ell) + 1 - p < 2^j(\theta + \ell) - 1 - Ni \leq \lfloor 2^j\theta \rfloor + 2^j\ell - Ni.$$

In other words, $\max\left(\lceil 2^j\theta \rceil + 2^j\ell - p, 0\right) \leq \min\left(\lfloor 2^j\theta \rfloor + 2^j\ell - Ni, 2^j - 1\right)$ which, we claim that has a solution for all ℓ. □

To conclude the proof, it only remains the proof of the above claim. Before we state it, let us make emphasize that for sake of simplicity we have make the proof for our particular case of Daubechies. The proof can be done in general.

Lemma 2.1.10. *Let $\psi(x)$ be a Daubechies wavelet with 10 vanishing moments. Then, $\max\left(\lceil 2^j\theta \rceil + 2^j\ell - p, 0\right) \leq \min\left(\lfloor 2^j\theta \rfloor + 2^j\ell - Ni, 2^j - 1\right)$ has a solution for all ℓ.*

Proof. Let us start the proof by saying that, in order to prove the Lemma, it is enough to show that it works for $2^j \leq 16$. That is, we need to show that for $2^j \leq 16$ and for all $0 \leq n \leq 2^j - 1$ there exists ℓ such that $(1 - 10) = Ni \leq 2^j(\theta + \ell) - n \leq p$. Thus,

$$-9 \leq 2^j(\theta + \ell) - n \leq 10 \Leftrightarrow -9 + n \leq 2^j(\theta + \ell) \leq 10 + n.$$

Now, observe that if $2^j \leq 8$ then $2^j - 1 = 7$. Therefore, $n \leq 7$ and, hence taking $\ell = 0$ we have $-9 + n \leq -2 < 0 \leq 2^j\theta \leq 8 < 10 + n$.

To conclude the proof, it remains the case $2^j = 16$ then $2^j - 1 = 15$. Therefore, $n \leq 15$. Let us proceed by cases.

- $n \leq 9$. In this situation, we have that $-9+n \leq 0 \leq 2^j\theta$. Hence, if $2^j\theta \leq 10+n$ we are done just by taking $\ell = 0$. On the contrary, if $2^j\theta > 10+n$ we must take $\ell = -1$. Indeed, it is true that $2^j(\theta+\ell) \leq 0 < 10+n$. On the other side, $2^j(\theta+\ell) = 2^j\theta - 2^j > 10+n-2^j = -6+n > -9+n$.
- $n > 9$. In this situation we have that $10+n \geq 20 > 2^j\theta$. Hence, if $2^j\theta \geq -9+n$ we are done just by taking $\ell = 0$. On the contrary, if $2^j\theta < -9+n$ we must take $\ell = 1$. Indeed, as above, in this situation we have $2^j(\theta+\ell) \geq 2^j > 15-9 \geq -9+n$. On the other side, $2^j(\theta+\ell) = 2^j\theta + 2^j < -9+n+2^j = 7+n < 10+n$.

□

From such constructive proof of Proposition 2.1.9 we can derive an algorithm to evaluate, and store, a massive set of points of $\mathbb{S}^1$ which is the one of the main tools of the Chapter 5. However, in order to make more clear the exposition we will include those key results in the present chapter just after the following paragraph.

2.1.1 The Haar value

As we have said before, Haar wavelet is very simple and has an explicit formula. This means that the above expressions can be simplified in order to use the Haar basis in an efficient way. In the forthcoming chapters we will use a uniformly distributed mesh of points in $\mathbb{S}^1$. Hence, we will need to derive the explicit formulas for the Haar basis of $\mathbb{S}^1$ (recall that $\psi^{\mathrm{PER}}_{-j,n}(\theta_i) = \psi_{-j,n}(\theta_i)$). To this end, fix $N = 2^{-J+1}$ and $\theta_i = i/N$ with $i = 0, \ldots, 2^N - 1$. Notice that

$$0 \leq \frac{i}{N/2^j} \leq \frac{2^{j-1}N-1}{N/2^j} = \frac{1}{2} - \frac{2^j}{N} \text{ and } \frac{2^{j-1}N-1}{N/2^j} \leq \frac{i}{N/2^j} \leq \frac{2^jN-1}{N/2^j} = 1 - \frac{2^j}{N}.$$

for $i \in \{0, \ldots, 2^{j-1}N-1\}$ and $i \in \{2^{j-1}N, \ldots, 2^jN-1\}$ respectively. Therefore, by Equation (1.7), we have that

$$\psi_{-j,0}(i/N) = \begin{cases} 2^{-j/2} & \text{if } i = 0, \ldots, 2^{j-1}N-1, \\ -2^{-j/2} & \text{if } i = 2^{j-1}N, \ldots, 2^jN-1. \end{cases}$$

Recall that, $\psi_{-j,n}(x)$ is supported over $I_{-j,n} = [2^jn, 2^j(n+1))$ which verifies $I_{-j,n} \cap I_{-j,m} = \emptyset$ if and only if $n \neq m$. That is, fixed a j then the rest of the functions are the same. Indeed, if we define $s = 2^{J-j}$ then we have that $\psi_{-j,0}(x) = 0$ if and only if $x \in [2^j, 1)$. For our particular setting it is $x \in [2^{-j}, 1) = [s/N, N/N)$. That is

$$\psi_{-j,0}(i/N) = \begin{cases} 2^{-j/2} & \text{for } 0 \leq i < s/2, \\ -2^{-j/2} & \text{for } s/2 \leq i < s, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, recall that $n = 0, \dots, 2^j - 1$. Then it follows that $\psi_{-j,n}(l/N) = \psi_{-j,0}((l-ns)/N)$ taking into account that $\psi_{-j,n}((i+ns)/N) = \psi_{-j,0}(i/N)$. In other words, we have the following formula for the rest of the elements:

$$\psi_{-j,n}(l/N) = \begin{cases} 2^{-j/2} & \text{for } 0 \leq l - ns < s/2, \\ -2^{-j/2} & \text{for } s/2 \leq l - ns < s, \\ 0 & \text{if } l - ns \geq 0. \end{cases} \tag{2.6}$$

The above expression is consistent because $l - ns < 0$ if and only if $l < ns$ and, also, $l - ns \geq s$ if and only if $l \geq ns + s = (n+1)s \leq N$. Therefore, we have derived an expression of the Haar basis over a uniform grid of $\mathbb{S}^1$. This simple formulation will be very useful in Chapter 5. Also, can be considered as a clue in order to understand the philosophy behind the following section.

2.2 Massive computation

This current section is devoted to explain how we can calculate a certain matrix Ψ_N^{PER} when Daubechies or Haar wavelets are used. To this end, recall that our framework will be $\mathbb{S}^1$ and $\{\psi_l^{\mathrm{PER}}\}_{l \in \mathbb{Z}^+}$ the orthonormal basis given either Haar or Daubechies with $p > 1$ vanishing moments wavelets. In Section 2.1 we have already made some comments concerning on the calculation of ψ_l^{PER}. Besides of this, the explanation is constructive and tailored for our needs so that at the end we can stitch these ideas to generate a method to evaluate and store the basis $\{\psi_l^{\mathrm{PER}}\}_{l \in \mathbb{Z}^+}$. In view of that, we can trigger an efficient way to calculate such matrix Ψ. But, who is Ψ_N^{PER}? Before answering the question, let us point out that first we will start with some results concerning on the particular structure of Ψ_N^{PER}. After that, we will join Section 2.1 and 2.2 to explain how we have performed a massive calculation of points in an efficient way.

Our framework, in this section will be an equidistantly mesh of points of the unit circle. This section is divided in two items which corresponds to the first two rows and the other ones. First of all, recall that from Theorem 1.1.13 for each level $j > 0$ there are $n = 2^j - 1$ functions that form a bi indexed orthonormal basis which can be compacted as follows $l = 2^j + n$. Therefore, one must take $j = -\lfloor \log_2(l) \rfloor$ and $n = l - 2^j$ to reverse the change of variables. On the other side, the equally spaced mesh of points of $\mathbb{S}^1$ is given by $\theta_i = i/N$ where $N = 2^J$ and the integers i are between 0 and $2^J - 1$. Hence, we define a square matrix Ψ_N^{PER} to be the $N \times N$ matrix whose entries (i, l) are $\psi_l^{\mathrm{PER}}(\theta_i)$ being $N > 0$ fixed.

The first two columns. The first column of the storage matrix it is simply, by Theorem 1.1.13, a column of 1's which, if one wants to economize memory on a computer it

can not be stored. The second column is a little bit tricky. Indeed, by definition of Ψ_N^{PER} it corresponds to $\psi_{0,0}^{\mathrm{PER}}$ and we claim that we only must evaluate the first half of θ_i's.

Lemma 2.2.1. *Let $\psi(x)$ be an $\mathbb{R}$-Daubechies wavelet with $p > 1$ vanishing moments then*

$$\psi_{0,0}^{\mathrm{PER}}(\theta) = -\psi_{0,0}^{\mathrm{PER}}(\theta + 1/2) \ \forall\ \theta \in [0, 1/2].$$

Proof. By Proposition 2.1.9 we need to show that $\forall\ \theta \in [0, 1/2]$

$$\begin{aligned} \psi_{0,0}^{\mathrm{PER}}(\theta) &= \sum_{\ell \in \Lambda_\theta} \lim_{k\to\infty} u(\theta+\ell)' v(\theta, k) = -\psi_{0,0}^{\mathrm{PER}}(\theta + 1/2) \\ &= -\sum_{\ell \in \Lambda_\theta} \lim_{k\to\infty} u(\theta+1/2+\ell)' v(\theta+1/2, k). \end{aligned}$$

To this end,

$$v(\theta+1/2,k) = \frac{1}{2p-1} 1' \prod_{i \in \mathrm{dyad}(\{2\theta+1\},k)} \mathbf{M}_i = \frac{1}{2p-1} 1' \prod_{i \in \mathrm{dyad}(\{2\theta\},k)} \mathbf{M}_i = v(\theta, k),$$

where $1'$ is the row-vector of ones. Now, notice that

$$\lfloor 2\theta + 2\ell \rfloor = \begin{cases} 2\ell & \text{if } 0 \le \theta < 1/2, \\ 2\ell + 1 & \text{if } 1/2 \le \theta < 1. \end{cases} \Rightarrow 1 - \lfloor 2\theta + 2\ell \rfloor = \begin{cases} 1 - 2\ell & \text{if } 0 \le \theta < 1/2, \\ -2\ell & \text{if } 1/2 \le \theta < 1. \end{cases}$$

That is,

$$(-1)^{1-\lfloor 2\theta+2\ell \rfloor} = \begin{cases} -1 & \text{if } 0 \le \theta < 1/2, \\ 1 & \text{if } 1/2 \le \theta < 1. \end{cases} \tag{2.7}$$

On the other side, for $i = 0, \ldots, 2p-2$ define

$$\iota_i = i + 1 - \lfloor 2\theta + 2\ell \rfloor = \begin{cases} i + 1 - 2\ell & \text{if } 0 \le \theta < 1/2, \\ i - 2\ell & \text{if } 1/2 \le \theta < 1. \end{cases}$$

In view of that and in order to proof the Lemma, we must show that the vector $u(\theta)$ has the same components, perhaps in a different order, when θ is in $[0, 1/2)$ and $[1/2, 1)$. Once this kind of equality is proven, by Equation (2.7) the proof follows because of the "*complementary*" signs of each entry. To this end, recall that $u(\theta)$, in our case, is a vector defined as

$$u_i(\theta) = (-1)^{1-\lfloor 2\theta \rfloor} h[\iota_i],\ i = 0,\ \ldots,\ 2p-2,$$

where if for some i the index ι_i is negative or larger than $2p-1$, then the corresponding component of u_i is set to be 0 because the support of the filter h is $0, \ldots, 2p-1$. For our

particular instance, and from the expressions above, it follows that for $i = 0, \ldots, 2p-2$ the vector u has the following formulation

$$u_i(\theta+\ell) = \begin{cases} 0 & i+1-2\ell < 0,\ i-2\ell < 0, \\ -h[i+1-2\ell] & 0 \le \theta < 1/2, \\ h[i-2\ell] & 1/2 \le \theta < 1, \\ 0 & i+1-2\ell > 2p-1,\ i-2\ell > 2p-1. \end{cases}$$

In order to prove this kind of '*complementary*" property we split the argument in two cases focusing the reasoning on the nonzero components of u. To this end, recall that from Proposition 2.1.9 we have a constraint on ℓ given by $\lceil 1-p-\theta \rceil \le \ell \le \lfloor 1-(1-p)-\theta \rfloor$

(a) $\theta \in (0, 1/2)$ (therefore $\iota_i = i+1-2\ell$). As we have said, $0 \le \iota_i \le 2p-1$. Thus, $0 \le i+1-2\ell \le 2p-1$ and hence $i+1 \ge 2\ell \ge 1-2p+i+1$ which gives the constraints on ℓ for $\theta \in (0,1/2)$: $1-p+\frac{i}{2} \le \ell \le \frac{i+1}{2}$. That is, the good range of ℓ's is given by $\max(1-p, 1-p+i/2)$ and $\min(\lfloor \frac{i+1}{2} \rfloor, p-1)$. Which, we claim, is equivalent to

$$1-p+\left\lceil \frac{i}{2} \right\rceil \le \ell \le \left\lfloor \frac{i+1}{2} \right\rfloor. \tag{2.8}$$

Hence, the component becomes

$$u_i(\theta+\ell) = \sum_{\ell=1-p+\lceil \frac{i}{2} \rceil}^{\lfloor \frac{i+1}{2} \rfloor} h[\iota_i] = \sum_{\iota=i+1-2\left(1-p+\lceil \frac{i}{2} \rceil\right)}^{i+1-2\lfloor \frac{i+1}{2} \rfloor} h[\iota] = \sum_{\substack{\iota=\mu(i) \\ 2\iota}}^{2(p-1)+\mu(i)} h[\iota],$$

where $\mu(i) = i+1-2\left\lceil \frac{i}{2} \right\rceil$ the last equality follows from Lemma 2.2.2. Now, we are almost done because it only remains to prove that on components of the vector u there are the sum of the odd components of the filter h and the sum of the even components in terms of $\mu(i)$. The proof, again taking into account Lemma 2.2.2, goes as follows

$$u_i(\theta+\ell) = \sum_{\substack{\iota=\mu(i) \\ 2\iota}}^{2(p-1)+\mu(i)} h[\iota] = \sum_{\substack{\iota=0 \\ 2\iota}}^{2(p-1)} h[\mu(i)+\iota] = \sum_{\iota=0}^{2(p-1)} h[\mu(i)+2\iota].$$

It only remains to show that $\max(1-p, 1-p+i/2) = 1-p+i/2$ and $\min(\lfloor \frac{i+1}{2} \rfloor, p-1) = \lfloor \frac{i+1}{2} \rfloor$. The first one it is obvious and the second one follows from recalling that at most i is equal to $2p-2$. Thus

$$\frac{i+1}{2} \le \frac{2p-2+1}{2} = p - \frac{1}{2}.$$

Hence $\min(\lfloor \frac{i+1}{2} \rfloor, p-1) = \lfloor \frac{i+1}{2} \rfloor$ and, consequently, Equation (2.8) is proven.

(b) $\theta \in (1/2, 1)$ (therefore $\iota_i = i - 2\ell$). Doing similar arguments as before the constraint on ℓ for $\theta \in [1/2, 1)$ is $\max\left(1-p, \frac{i+1}{2} - p\right) \leq \ell \leq \min\left(\lfloor \frac{i}{2} \rfloor, p-1\right)$ which, since for the range of i's $i/2 \leq (2p-2)/2 = p-1$, is equivalent to

$$\left\lceil \frac{i+1}{2} \right\rceil - p \leq \ell \leq \left\lfloor \frac{i}{2} \right\rfloor.$$

That is,

$$u_i(\theta + \ell) = \sum_{\ell = \left\lceil \frac{i+1}{2} \right\rceil - p}^{\left\lfloor \frac{i}{2} \right\rfloor} h[\iota_i] = \sum_{\substack{\iota = i - 2\left(-p + \left\lceil \frac{i+1}{2} \right\rceil\right) \\ 2\iota}}^{i - 2\left\lfloor \frac{i}{2} \right\rfloor} h[\iota] = \sum_{\substack{\iota = i - 2\left(-p + \left\lceil \frac{i-1}{2} \right\rceil\right) \\ 2\iota}}^{2(p-1) + i - 2\left\lfloor \frac{i}{2} \right\rfloor} h[\iota],$$

where the last equality comes from the fact that $2 - 2 = 0$. Indeed, $i - 2\left(-p + \left\lceil \frac{i+1}{2} \right\rceil\right) = 2p - 2 + i + 2 - 2\left\lceil \frac{i+1}{2} \right\rceil = 2(p-1) + i - 2\left(-p + \left\lceil \frac{i-1}{2} \right\rceil\right)$. Now, again by Lemma 2.2.2

$$u_i(\theta + \ell) = \sum_{\substack{\iota = \widetilde{\mu}(i) \\ 2\iota}}^{2(p-1) + \widetilde{\mu}(i)} h[\iota] = \sum_{\substack{\iota = 0 \\ 2\iota}}^{2(p-1)} h[\widetilde{\mu}(i) + \iota] = \sum_{\iota = 0}^{2(p-1)} h[\widetilde{\mu}(i) + 2\iota].$$

where $\widetilde{\mu}(i) = i - 2\left\lfloor \frac{i}{2} \right\rfloor$ and as in the case of $\theta \in [0, 1/2)$ the components of u are the sum of the odd (or even) components of the filter h in terms of $\widetilde{\mu}(i)$.

Now we have all the tools to conclude the proof. Indeed, by Proposition 1.1.4 it follows that $\sum_{n \in \mathbb{Z}} h[2n] = \sum_{n \in \mathbb{Z}} h[2n+1]$ which are precisely, by the above computations, the components of the vectors u' and $\widetilde{u}$ where they are vectors for $\theta \in (0, 1/2]$ and $\theta \in (1/2, 1]$ respectively. Finally, from Equation (2.7), fixed i it follows that $u'_i = -\widetilde{u}_i$. In other words, we have shown that

$$\psi^{\mathrm{PER}}_{0,0}(\theta) = -\psi^{\mathrm{PER}}_{0,0}(\theta + 1/2) \;\forall\; \theta \in (0, 1/2]$$

because of u. Finally, by the continuity of $\psi^{\mathrm{PER}}_{0,0}(\theta)$ it follows that, just by taking a sequence θ_n such that its limit is zero, the above expression is true for all $\theta \in [0, 1/2)$. Indeed,

$$\psi^{\mathrm{PER}}_{0,0}(0) = \lim_{n \to \infty} \psi^{\mathrm{PER}}_{0,0}(\theta_n) = -\lim_{n \to \infty} \psi^{\mathrm{PER}}_{0,0}(\theta_n + 1/2) = -\psi^{\mathrm{PER}}_{0,0}(1/2).$$

□

The following Lemma is the proof of some of the properties used in the above Lemma.

Lemma 2.2.2. *For $0 \leq i \leq 2p-2$ consider the following binary functions given by $\mu(i) = i + 1 - 2\left\lceil \frac{i}{2} \right\rceil$ and $\widetilde{\mu} = i - 2\left\lfloor \frac{i}{2} \right\rfloor$. Then*

(a) for all $0 \leq i \leq 2p-2$ it is true that $\mu(i) = i + 1 - 2\lfloor \frac{i+1}{2} \rfloor$,

(b) for all $0 \leq i \leq 2p-2$ it is verified that $\widetilde{\mu}(i) = i + 1 - 2\left\lceil \frac{i-1}{2} \right\rceil$,

(c) for all $0 \leq i \leq 2p-2$ the functions μ and $\widetilde{\mu}$ are binary and complementary. That is,

$$\mu(i) = \begin{cases} 1 & \text{if } i \text{ is even} \\ 0 & \text{if } i \text{ is odd} \end{cases} \quad \text{and} \quad \widetilde{\mu}(i) = \begin{cases} 0 & \text{if } i \text{ is even} \\ 1 & \text{if } i \text{ is odd} \end{cases}.$$

Proof. We need to prove the three properties.

(a) In order to prove the statement, it suffice to show that $2\left\lceil \frac{i}{2} \right\rceil = 2\left\lfloor \frac{i+1}{2} \right\rfloor$. But if i is even then $\left\lceil \frac{i}{2} \right\rceil = \frac{i}{2}$ and $\left\lfloor \frac{i+1}{2} \right\rfloor = \left\lfloor \frac{i}{2} + \frac{1}{2} \right\rfloor = \frac{i}{2}$. On the other hand, when i is odd, the equality is true by the properties of $\lceil \cdot \rceil$ and taking into account that $i+1$ is even.

(b) The proof of the statement follows similar argument as the previous one. Indeed, it suffice to prove that $2\left\lceil \frac{i-1}{2} \right\rceil = 2\left\lfloor \frac{i}{2} \right\rfloor$. When i is even, then $\left\lceil \frac{i-1}{2} \right\rceil = \left\lceil \frac{i}{2} - \frac{1}{2} \right\rceil = \left\lfloor \frac{i}{2} \right\rfloor$. On the contrary, if i is odd $i-1$ is even and $\left\lceil \frac{i-1}{2} \right\rceil = \frac{i-1}{2} = \left\lfloor \frac{i}{2} \right\rfloor$.

(c) By the above item, $\mu(i) = i + 1 - 2\left\lceil \frac{i}{2} \right\rceil = i + 1 - 2\lfloor \frac{i+1}{2} \rfloor = i + 1 - 2\left\lfloor \frac{i}{2} \right\rfloor$. That is, $\mu(i) = \widetilde{\mu}(i) + 1$. Now, it only remains to show the values of $\mu(i)$ and this is done splitting i in the even and odd case. Actually, if i is even then $\mu(i) = i + 1 - 2\frac{i}{2} = i + 1 - i = 1$ and $\mu(i) = i + 1 - 2\frac{i+1}{2} = i + 1 - i - 1 = 0$ if i is odd.

□

And, after such tedious Lemmae, it is almost clear what we have claimed at the beginning of the present section: only the half of the work must be done. Now, it remains to show how are the rest of the columns. We will see that, as we have pointed out when using Haar's wavelet, Ψ_N^{PER} has a kind of auto-similarity when one uses a Daubechies wavelet.

The rest of columns. For the rest of columns one can save time computation (and/or memory) in a similar manner as the previous Lemma. It turns out that the $l > 1$ columns of Ψ_N^{PER} follows some kind of circulant, by blocks, matrix. That is, for each block of size $\lfloor \log_2(l) \rfloor$ (which corresponds to the level $-j$) one can establish a rule to construct, from some "*privileged values*", the complete block. For instance, if $l = 1$ one of "*privileged*" set of points are $\theta \in [0, 1/2]$. Indeed,

$$\psi_{-1,1}^{\mathrm{PER}}(\theta + 1/2) = \sum_{\ell \in \mathbb{Z}} \psi(2\theta + 1 + 2\ell - 1) = \sum_{\ell \in \mathbb{Z}} \psi(2\theta + 2\ell) = \psi_{-1,0}^{\mathrm{PER}}(\theta).$$

The above chain of equalities guides us to the following Lemma.

Lemma 2.2.3. *Let $\psi(x)$ be an $\mathbb{R}$-Daubechies wavelet with $p > 1$ vanishing moments and consider and fix a block of size $\lfloor \log_2(l) \rfloor$ which generates a sequence of consecutive naturals k between 0 and $2^j - 1$. Then for all $x \in [0, 1/2^j]$ and $0 \leq n \leq 2^j - 1$*

$$\psi^{\mathrm{PER}}_{-j,k}(x) = \psi^{\mathrm{PER}}_{-j,n}\left(x + \frac{n-k}{2^j}\right),$$

where $n - k = n - k \pmod{2^j}$. Also this fills in all the block of size $\lfloor \log_2(l) \rfloor$ of the matrix Ψ^{PER}_N.

Proof. Let l fix a block of size $\lfloor \log_2(l) \rfloor$ and consider a fixed k which is, by construction, less than 2^j (notice that n also). We must distinguish two cases:

$n - k \geq 0$: then $n - k \pmod{2^j} = n - k$ and therefore,

$$\begin{aligned}
\psi^{\mathrm{PER}}_{-j,n}\left(x + \frac{n-k}{2^j}\right) &= \sum_{\ell \in \mathbb{Z}} \psi\left(2^j x + n - k + 2^j \ell - n\right) \\
&= \sum_{\ell \in \mathbb{Z}} \psi\left(2^j (x + \ell) - k\right) \\
&= \psi^{\mathrm{PER}}_{-j,k}(x).
\end{aligned}$$

$n - k < 0$: then $n - k \pmod{2^j} = n - k + 2^j$ and therefore,

$$\begin{aligned}
\psi^{\mathrm{PER}}_{-j,n}\left(x + \frac{n-k}{2^j}\right) &= \sum_{\ell \in \mathbb{Z}} \psi\left(2^j x + n - k + 2^j + 2^j \ell - n\right) \\
&= \sum_{\ell \in \mathbb{Z}} \psi\left(2^j (x + \ell) - k\right) \\
&= \psi^{\mathrm{PER}}_{-j,k}(x),
\end{aligned}$$

where the last equalities comes from Theorem 1.1.13 because for each j, there are only 2^j elements in the basis which is almost the end of the proof.

In order to conclude it, it only remains to point out that moving k for all the integers between 0 and 2^j, and also n, we recover all the entries (i, l) of the block of size $\lfloor \log_2(l) \rfloor$. □

Remark 2.2.4. In view of the above Lemma, it is important to stress that $\ell \in \mathbb{Z}$. We omit $\ell \in \Lambda_\theta$ because the assertion follows from the techniques used in Lemma 2.2.1 and the proof of Lemma 2.2.3. However, this gives us a method to fill in the matrix Ψ^{PER}_N. Indeed, for a block of size $\lfloor \log_2(l) \rfloor$ we fix $\theta_i = i/N$ and n. ∎

To conclude this "*storage*" part we recover the comments which concerns to the "*goodness*" of the matrices generated above. The matrix Ψ_N^{PER}, if $\psi(x)$ is an $\mathbb{R}$-Daubechies wavelet with $p \geq 1$ vanishing moments, is sparse if $N = 2^J$ is big enough. That is, Ψ_N^{PER} has the most of the elements equal to zero due to the compact support of $\psi(x)$. Indeed, by Theorem 1.1.9 and Proposition 1.1.15, we have that each ψ_l^{PER} has compact support. Therefore, unfolding $l = 2^j + n$, we have that at each j the support is divided by 2. This allows us to ensure that there exist a j_0 such that $\psi_l^{\mathrm{PER}} = \psi_l$ if $j > j_0$. Hence, because of the support, we will have a large set of columns such that they are populated by zero except a range which is inversely proportional to 2^j. In other words, typically the matrix Ψ_N^{PER} will have the entries of the j_0 first columns with nonzero $\psi_l^{\mathrm{PER}}(\theta_i)$'s. After that j_0, the columns will become zero except for a small range of entries. This, at a first approach, god property will produce bad consequences as we will see later on in Section 5.2.2. Finally, recall that the support of Haar is $[0, 1)$ and it is the smallest one of the Daubechies family wavelets. Hence the above arguments are also valid for the Haar wavelet. And, moreover, the number of non-zero elements of the $N \times N$ matrix Ψ_N^{PER} is $N \log_2(N) + N = N(J + 1)$.

It only remains what we have sketched at the beginning of the current section: the algorithms to calculate Ψ_N^{PER} (and $\widetilde{\Psi}_N^{\mathrm{PER}}$ also). This will be the main topic of the following lines.

2.2.1 The algorithms

Thereupon we will join the results of the previous section and those ones in Section 1.1.3, 2.1. Hence, we will provide two algorithms to work and compute Daubechies wavelet with $p > 1$ vanishing moments. To this end and in order to fix ideas, let us start by recalling that for the double-precision binary floating-point the `IEEE 754` standard specifies a `binary64` as having:

(a) sign bit: 1 bit,

(b) exponent width: 11 bits,

(c) significant precision: 53 bits which 52 are explicitly stored as d_i's.

Therefore, a number x is represented as follows

$$(-1)^{\mathsf{sign}} \left(1 + \sum_{i=1}^{52} d_{52-i} 2^{-i}\right) \times 2^{\mathbf{e}-1023}.$$

Keeping this in mind, we sketch how one can compute $v(\cdot)$ from Theorem 2.1.6. Indeed, when $0 \leq x < 1$ then $\mathrm{sign} = 0$ and the number x is:

- If exponent is zero then $x = 0.\text{Mantisa} \times 2^{-1022}$ (called *subnormal numbers*). A word of caution must be done. Indeed, the maximum of the subnormal `double` is $2.2250738585072009 \times 10^{-308}$ which is very smaller than the number $7.8886090522101180541 \times 10^{-31} = 2^{-100}$. In view of that, all subnormal numbers can be mapped to zero when doing numerics. In this case, exponent will be strictly positive as in the next case.

- If exponent is non-zero then $x = 1\text{Mantisa} \times 2^{\mathbf{e}}$, where according to the `IEEE 754` convention $\mathbf{e} = \text{exponent} - 1023$. The fact that $x < 1$ implies that $\mathbf{e} \leq -1$; that is, $0 < \text{exponent} \leq 1022$. Therefore we have that $x = 0.1\text{Mantisa} \times 2^{\mathbf{e}+1}$. Hence we can conclude that $x = 0.00\ldots001\text{Mantisa}$ where the number of initial zeros is $-(\mathbf{e}+1) = 1022 - \text{exponent}$.

Summarizing, since we are in $\mathbb{S}^1$ we take numbers $\theta \in [0,1)$. Therefore, if we eliminate the subnormal numbers, then we have $\theta = 0.00\ldots001\text{Mantisa}$ where the number of initial zeros is $1022 - \text{exponent}$. Clearly, $0 \leq 1022 - \text{exponent} < 1022$ must be verified.

On the other side, in order to get $\psi_l^{\text{PER}}(\theta_i)$ we need to use Proposition 2.1.9. Nevertheless, it should be noted that first of all it is necessary to calculate the Daubechies - Lagarias products $v(\theta)$ from Theorem 2.1.6.

Remark 2.2.5. Recall that $v(x)$ from Theorem 2.1.6 is given by the following infinite product

$$v(x) = \lim_{n\to\infty} \frac{1}{2p-1} 1' \prod_{i \in \text{dyad}(\{2x\},n)} \mathbf{M}_i,$$

where the matrices $\mathbf{M}_0$ and $\mathbf{M}_1$ are explicitly given by Equation (2.1) and (2.2). For the particular instance of 0 and using the idea behind of the well known Power Method, a procedure which produces the eigenvalue of the biggest modulus using powers of matrix, we can derive $v(0)$. Indeed, the above expression is equivalent to

$$v(0) = \lim_{n\to\infty} \frac{1}{2p-1} 1' \prod_{i \in \text{dyad}(\{0\},n)} \mathbf{M}_i = \lim_{n\to\infty} \frac{1}{2p-1} 1' \prod_{i=0}^{n} \mathbf{M}_0.$$

Therefore, $v(0)$ is the eigenvector of biggest eigenvalue of $\mathbf{M}_0$ because the last infinite product is only composed by the matrix $\mathbf{M}_0$. ∎

Thereafter, we are ready to derive the implementation of the calculation of the Daubechies - Lagarias products $v(x)$ from Theorem 2.1.6 and using Proposition 2.1.9.

Algorithm 1: The calculation of the Daubechies - Lagarias products $v(x)$. Recall that & is the Binary `AND` operator and copies a bit to the result if it exists in both operands. Also, $\gg$ is the Binary Right Shift operator. The left operands value is moved right by the number of bits specified by the right operand. Finally, $v^{k)}$ stands for the k-th multiplication to perform.

`DaubechiesLagariasProducts(`θ, $\mathbf{M}_0$, $\mathbf{M}_1$, m_0, v`){`

Data: The point $\theta \in \mathbb{S}^1$, the matrices $\mathbf{M}_0,\ \mathbf{M}_1$ and the eigenvector m_0 of $\mathbf{M}_0$.

Result: The value of the vector $v(\theta)$.

$\tilde{\theta} = \{2\theta\},\ \mathfrak{m} = \mathrm{Mantisa}(\theta),\ v = 1'$

if $\tilde{\theta} = 0$ **then**

  $v = m_0$

else

  $\mathrm{pb} = \mathfrak{m}\ \&\ 1,\ \mathfrak{m} \gg= 1,\ \mathrm{nppb} = 1$

  for $i = 0$ **to** 51 **do**

    $b = \mathfrak{m}\ \&\ 1$

    if $b = \mathrm{pb}$ **then** $\mathrm{nppb}++$

    else $\mathbf{M}_{\mathrm{pb}} v^{\mathrm{nppb})},\ \mathrm{pb} = b,\ \mathrm{nppb} = 1$

    $\mathfrak{m} \gg= 1$

  end for

  if $\mathrm{pb} = 1$ **then** $\mathrm{nppb}++,\ \mathbf{M}_{\mathrm{pb}} v^{\mathrm{nppb})}$

  else $\mathbf{M}_0 v^{\mathrm{nppb})},\ \mathbf{M}_1 v^{\mathrm{nppb})}$

  $-(\mathbf{e}+1) = 1022 - \mathrm{exponent}(\tilde{\theta})$

  if $-(\mathbf{e}+1) < 53$ **then**

    $\mathbf{M}_0 v^{\mathrm{nppb})}$

  else

    $\mathbf{M}_0 v^{-(\mathbf{e}+1)-\lfloor -(\mathbf{e}+1)/52 \rfloor 52)}$

    for $i = 0$ **to** $-(\mathbf{e}+1)/52$ **do**

      $\mathbf{M}_0 v^{52)}$

    end for

  end if

end if

}

The above algorithm has several comments which are outlined in the following Remark.

Remark 2.2.6. We want to emphasize several things from the Algorithm 1.

(a) The first thing to notice is at the first line. Indeed, we need to multiply many times the matrices $\mathbf{M}_0$ and $\mathbf{M}_1$. In general, the matrix multiplication if carried out naively spends $\mathcal{O}(N^3)$ running time. However, this time can be improved in our case. Indeed, setting v to be the column vector of ones and starting from the right hand side, doing $\mathbf{M}_{i_\mathsf{f}} v$, then the time decreases to $\mathcal{O}(N^2)$, where i_f is the last digit of the finite binary expansion of θ. Even doing it naively. Such naively way, since the matrices $\mathbf{M}_0$ and $\mathbf{M}_1$ are relatively small, seems to be the best choice to proceed.

(b) Related with the above point, since the matrix product of $\mathbf{M}_0$ and $\mathbf{M}_1$ is not commutative, we need the binary expansion in the right order. And the right order is the backwards order because we start the multiplication with $\mathbf{M}_{i_\mathsf{f}}$. Of course, taking N to be a power of two, such binary expansion for the points i/N, with $i = 0, \ldots, N-1$ is easy to derive. Therefore, in such situation the Algorithm 1 can be better performed.

(c) It is important to emphasize that the procedure `DaubechiesLagariasProducts`, as it can be seen in the proof of Proposition 2.1.9, will be called only once to calculate $\psi_l^{\mathrm{PER}}(\theta_i)$.

■

Now we can give the implementation, which is a direct consequence of the proof of Proposition 2.1.9, for the evaluation of ψ^{PER}. Let us emphasize that the main ingredient $v(x)$ is given by Algorithm 1 and the filter $h[n]$ is pre-stored.

Algorithm 2: The calculation of $\psi_l^{\mathrm{PER}}(\theta)$ using the Daubechies - Lagarias products. Recall that ? : is the Ternary Conditional Operator. Also, can be commonly referred to as the conditional operator. Notice that is the verbatim application of Proposition 2.1.9.

```
PERpsijntheta(θ, 2^j, Ni, h[n]){
    Data: The point θ ∈ S^1, the dimension 2^j, Ni = 1 − p and the filter h[n].
    Result: The value of ψ_l^PER(θ).
    ⋮
    DaubechiesLagariasProducts(θ, v);
    for ℓ = ⌈Ni/2^j − θ⌉ to ⌊−Ni/2^j − θ⌋ + 1 do
        for n = max(⌈2^j θ⌉ + 2^j ℓ − p, 0) to min(⌊2^j θ⌋ + 2^j ℓ − Ni, 2^j − 1) do
            θ̃ = θ + 2^j ℓ − n;
            if θ̃ ≤ Ni or θ̃ ≥ 1 − Ni then
                return 0;
            else
                for i = 0 to 2(1 − Ni) do
                    if 0 ≤ i + (1 − ⌊2θ̃⌋) ≤ 2(1 − Ni) then ψ_{j,n}^PER(θ) += (i + (1 − ⌊2θ̃⌋)
                    mod (2) ? h[i + (1 − ⌊2θ̃⌋)] :  − h[i + (1 − ⌊2θ̃⌋)])v(i) ;
                end for
            end if
        end for
    end for
}
```

After this computational aspects of the above theory, we are ready to work with wavelets $\psi_l^{\mathrm{PER}}(\theta)$ and compute the desired wavelet coefficients $d_{-j}^{\mathrm{PER}}[n]$. Of course, there are several ways to get such coefficients. In particular, we will explain three of them in the following sections.

2.3 Two ways to get the coefficients

In the present section, we present and develop two ways to get $\langle f, \psi_{-j,n}\rangle$, namely the Fast Wavelet Transform and the solution of a (non)-linear system of equations. Also, some ideas of an ad-hoc Gaussian quadrature rule will be given. The first two ones will be strongly used in Chapter 4 and 5 where our natural environment will be $\mathbb{S}^1$. To this end, and as usual along this memory, firstly we will present the methods in $\mathbb{R}$ and then, if it is needed, we will modify the methods in order to be in $\mathbb{S}^1$.

2.3.1 Fast Wavelet Transform

Recall that it is possible to approximate maps $f \in \mathscr{L}^2(\mathbb{R})$ by linear combinations of wavelets. Indeed, by using the expression (1.2) and Theorem 1.1.5, the projection of f to

$\mathcal{V}_{-J} \subset \mathscr{L}^2(\mathbb{R})$:

$$\sum_{n\in\mathbb{Z}} \langle f, \phi_{-J,n}\rangle\, \phi_{-J,n},$$

is a good approximation of f provided that $J > 0$ is large enough. We want to rewrite such an approximation as linear combination of wavelets of the form

$$f \sim \sum_{n\in\mathbb{Z}} \langle f, \phi_{0,n}\rangle\, \phi_{0,n} + \sum_{j=0}^{J-1}\sum_{n\in\mathbb{Z}} \langle f, \psi_{-j,n}\rangle\, \psi_{-j,n} \in \mathcal{V}_0 \oplus \bigoplus_{j=0}^{J-1} \mathcal{W}_j.$$

To do it, as usual, we define the coefficients

$$a_j[n] := \langle f, \phi_{j,n}\rangle \quad \text{and} \quad d_j[n] := \langle f, \psi_{j,n}\rangle$$

for $j, n \in \mathbb{Z}$. With this notation, the initial approximation of f becomes

$$\sum_{n\in\mathbb{Z}} a_{-J}[n]\phi_{-J,n}.$$

By (1.1), for every $j \in \mathbb{Z}$ we have

$$\sum_{n\in\mathbb{Z}} a_{-j}[n]\phi_{-j,n} = \sum_{n\in\mathbb{Z}} a_{-j+1}[n]\phi_{-j+1,n} + \sum_{n\in\mathbb{Z}} d_{-j+1}[n]\psi_{-j+1,n}. \tag{2.9}$$

Now, to obtain the coefficients $a_{-j+1}[n]$ and $d_{-j+1}[n]$ from $a_{-j}[n]$, we use the *Fast Wavelet Transform (FWT)* given by (see [Mal98, Theorem 7.7]):

$$\begin{cases} a_{j+1}[p] := \displaystyle\sum_{n\in\mathbb{Z}} h[n-2p]a_j[n] \quad \text{and} \quad d_{j+1}[p] := \sum_{n\in\mathbb{Z}} g[n-2p]a_j[n]; \\ \text{with } g[p] = (-1)^{1-p}h[1-p] \end{cases} \tag{2.10}$$

for every $j, p \in \mathbb{Z}$. Hence, from the iterative use of (2.9) and (2.10) starting with the approximation $\sum_{n\in\mathbb{Z}} a_{-J}[n]\phi_{-J,n}$ we obtain the approximation of f that we are looking for:

$$f \sim \sum_{n\in\mathbb{Z}} a_0[n]\phi_{0,n} + \sum_{j=0}^{J-1}\sum_{n\in\mathbb{Z}} d_{-j}[n]\psi_{-j,n} \in \mathcal{V}_0 \oplus \bigoplus_{j=0}^{J-1} \mathcal{W}_j.$$

For (numerical) applications such infinite approximations are usually not available since we often work with finite information about our function. For this we need a similar theory for subspaces of $\mathcal{V}_j$ and $\mathcal{W}_j$ of finite dimension. For $j \geq 0$ we define

$$\mathcal{V}^*_{-j} := \left\langle \phi_{-j,0}, \phi_{-j,1}, \ldots, \phi_{-j,2^j-1} \right\rangle \subset \mathcal{V}_{-j}, \text{ and}$$
$$\mathcal{W}^*_{-j} := \left\langle \psi_{-j,0}, \psi_{-j,1}, \ldots, \psi_{-j,2^j-1} \right\rangle \subset \mathcal{W}_{-j},$$

where $\langle f_1, f_2, \dots, f_n \rangle$ denotes the subspace of $\mathscr{L}^2(\mathbb{R})$ generated by the linear combinations of $f_1, f_2, \dots, f_n$. From the comment at the end of Section 7.3.1 of [Mal98] (see also [Fra99, Lemma 3.26] for a more detailed account), it follows that

$$\mathcal{V}^*_{-j+1} = \mathcal{W}^*_{-j} \oplus \mathcal{V}^*_{-j} \tag{2.11}$$

for every $j > 0$.

Now, given a function $f \in \mathscr{L}^2(\mathbb{R})$ we can take a good finite approximation of the map given by its projection to $\mathcal{V}^*_{-J}$:

$$f \sim \sum_{n=0}^{2^J-1} a_{-J}[n]\phi_{-J,n}, \tag{2.12}$$

provided that J is large enough. Again, we are interested in writing such an approximation as linear combination of wavelets, but in this case this expansion must be finite:

$$f \sim a_0[0]\phi_{0,0} + \sum_{j=0}^{J-1}\sum_{n=0}^{2^j-1} d_{-j}[n]\psi_{-j,n} \in \mathcal{V}^*_0 \oplus \bigoplus_{j=0}^{J-1} \mathcal{W}^*_{-j}.$$

To obtain this expression observe that (2.11) implies

$$\sum_{n=0}^{2^j-1} a_{-j}[n]\phi_{-j,n} = \sum_{n=0}^{2^{j-1}-1} a_{-j+1}[n]\phi_{-j+1,n} + \sum_{n=0}^{2^{j-1}-1} d_{-j+1}[n]\psi_{-j+1,n} \tag{2.13}$$

for $j > 0$. Now, to obtain the coefficients $a_{-j+1}[n]$ and $d_{-j+1}[n]$ from $a_{-j}[n]$, instead of using formulae (2.10), we use the following circular convolution version of them (see [Mal98, Section 7.5.1] or the proof of [Fra99, Lemma 3.26]):

$$\begin{cases} a_{-j+1}[p] := \displaystyle\sum_{n=0}^{2^j-1} h[n-2p]a_{-j}[n] \quad \text{and} \quad d_{-j+1}[p] := \sum_{n=0}^{2^j-1} g[n-2p]a_{-j}[n]; \\ \text{with } g[p] = (-1)^{1-p}h[p] \end{cases} \tag{2.14}$$

for every $j > 0$ and $p \in \{0, 1, \dots, 2^{j-1} - 1\}$. Hence, with the iterative use of (2.13) and (2.14) starting with the approximation (2.12) we obtain

$$f \sim a_0 + \sum_{j=0}^{J-1}\sum_{n=0}^{2^j-1} d_{-j}[n]\psi_{-j,n} \in \mathcal{V}^*_0 \oplus \bigoplus_{j=0}^{J-1} \mathcal{W}^*_{-j}. \tag{2.15}$$

as we wanted.

Remark 2.3.1. It is important to point out that the "*finiteness*" of the FWT turns the orthonormal basis of $\mathcal{W}^*_{-j}$ into an orthonormal basis of $\mathbb{S}^1$ for $j > 0$. Therefore we do not need anything more when we will have to deal with maps which naturally live in $\mathbb{S}^1$ in the forthcoming. ∎

To effectively compute such an approximation one remaining problem is left : namely, to find a good estimate of the initial coefficients $a_{-J}[n] = \langle f, \phi_{-J,n}\rangle$. In the literature there is a lot of discussion on how to compute these coefficients, but a simple customary approach is to use the following estimate (see, for instance, [Fra99, Lemma 5.54] and its proof):

Lemma 2.3.2. *Assume that f verifies $|\langle f, \phi_{j,n}\rangle| < \infty$ for every $j, n \in \mathbb{Z} \times \mathbb{Z}$ and*

$$|f(x) - f(y)| \leq C_1 \, |x - y|^{\alpha} \text{ with } \alpha \in (0, 1]$$

for all real numbers x, y and a constant $C_1 < \infty$. Suppose that the scaling function ϕ from an MRA $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ is such that

$$\phi \in \mathscr{L}^1(\mathbb{R}), \ \widehat{\phi}(0) = \int_{\mathbb{R}} \phi(x) \, dx = 1 \text{ and } \int_{\mathbb{R}} |x|^{\alpha} \, \phi(x) \, dx < C_2.$$

Then, for every $j, n \in \mathbb{Z} \times \mathbb{Z}$,

$$\left| \langle f, \phi_{j,n}\rangle - 2^{j/2} f(2^j n) \right| < C_1 C_2 2^{j\left(\alpha + \frac{1}{2}\right)}.$$

As a corollary of this lemma we see that if f is Lipschitz, then

$$a_{-J}[n] \approx 2^{-J/2} f(2^{-J} n).$$

Let us explain the problem of such approximation. To this end, let us define

$$\widetilde{a}_{-J}[n] = a_{-J}[n] \pm \varepsilon,$$

where $a_{-J}[n]$ are the *real* $|\langle f, \phi_{j,n}\rangle|$ coefficients. Since, by Proposition (1.1.4), the scaling filter verifies $\sum\limits_{n\in\mathbb{N}} h[n] = \sqrt{2}$ then, by Equation (2.10), one has that

$$\widetilde{a}_{j+N}[p] = \widetilde{a}_{j+N}[p] \pm 2^{N/2}\varepsilon.$$

Therefore, at each step the coefficients $a_j[n]$ are worse inducing some error, in the finite case, to the $d_j[n]$ ones. But moreover, in Lemma 2.3.2 one only demands the existence of $0 < C_1 < \infty$ such that

$$|f(x) - f(y)| \leq C_1 \, |x - y|^{\alpha} \text{ with } \alpha \in (0, 1].$$

The problem is that such a C_1 exists but it can highly complicate the election of a good J in order to minimize the error that gives a "*sufficiently*" good approximation.

2.3.2 A (non)-linear system of equations

Recall that, given a function f we want to calculate

$$\langle f, \psi_{-j,n} \rangle = \int_{\mathrm{supp}(\psi_{-j,n}(x))} f(x)\psi_{-j,n}(x)\, dx$$

in order to get an expression like

$$a_0 + \sum_{j=0}^{J} \sum_{n=0}^{2^j-1} \langle f, \psi_{-j,n} \rangle \, \psi_{-j,n}.$$

A good way to attack the above problem is take a similar approach of the well Collocation Method. That is, perform a linear system of equations, being $\langle f, \psi_{-j,n} \rangle$ the unknowns, since one can create such system using the following identity

$$f(x_i) = a_0 + \sum_{j=0}^{J} \sum_{n=0}^{2^j-1} \langle f, \psi_{-j,n} \rangle \, \psi_{-j,n}(x_i),$$

where $x_i \in \mathbb{R}$ are given. Now we know how to evaluate $\psi_{j,n}(x_i)$ and hence we can create a system of equations in order to calculate the desired wavelet coefficients. This will be, precisely, the main topic of the following lines. But let us make a final comment related to the value of the wavelet at a single point and the system to solve.

Recall that we are specially interested in Daubechies wavelet, a kind of wavelets which are compactly supported. That is, depending on whether x_i is placed then the wavelet value is zero. Therefore, the system to solve is a large sparse system (specially when one considers the Haar wavelet).

Let us resume us explanation by emphasizing that the above system of equations gives rise a non linear system of equations. To solve it, we will use Newton's method. Following its one dimensional formulation, one has to multiply by the inverse of the Jacobian matrix $\mathbf{JF}(x_n)$

$$x_{n+1} = x_n - \mathbf{JF}^{-1}(x_n)\mathbf{F}(x_n).$$

But, one can reformulate the above expression by solving the system of linear equations

$$\mathbf{JF}(x_n)(\mathrm{X}) = -\mathbf{F}(x_n)$$

for the unknown $\mathrm{X} = x_{n+1} - x_n$ and updating correctly the seed x_0.

Remark 2.3.3. As we will see in Chapter 3 and 5 we will work with a Newton's method tailored for our particular needs. In addition, $\mathbf{JF}$ and $\mathbf{F}$ will be very specific. Hence, we will stop here the explanation and we will be focused on how we will solve the aforesaid linear system iteratively. ∎

Solving the system. Let A be a $n \times n$ nonsingular matrix and suppose that we want to solve $Ax = b$, where b is a vector of dimension n. To this end, consider the polynomial $m(x)$ of degree k such that $m(A)b = 0$ and given by

$$m(x) = x^k - \sum_{j=0}^{k-1} \alpha_j x^j.$$

This polynomial $m(x)$ is, precisely, the minimum polynomial for the matrix A which, by the Cayley-Hamilton theorem, its degree verifies $k \leq n$ and, also, notice that $m(A)b = 0$ gives us the solution of the system. Indeed,

$$m(A)b = \left(A^k - \sum_{j=0}^{k-1} \alpha_j A^j \right) b = 0 \Rightarrow A\overbrace{\left(\frac{A^{k-1}b - \alpha_{k-1}A^{k-2}b - \cdots - \alpha_1 b}{\alpha_0} \right)}^{\text{the unknown } x} = b.$$

On the other side, consider the following k-dimensional vectorial space given by

$$\mathcal{K}_k(A, b) = \left\langle b, Ab, A^2b, \ldots, A^k b \right\rangle$$

where k is, also, the degree of the minimal polynomial $m(x)$. In general, the linear subspaces spanned by the images of b under the powers of A are is called the *k-Krylov subspace*. The method used to solve $Ax = b$, roughly speaking, is based on "*good approximations*" of x by a *sufficiently good* polynomial. All of the theoretical tools for understanding the Krylov solvers are borrowed from [Gut07, Mey00, Saa03, vdV09] and this will be the topic of the following paragraphs and this kind of techniques are justified because of the sparsity of the system that we want to solve.

Definition 2.3.4. A *Krylov space solver* is an iterative method starting from some initial approximation x_0 and the corresponding residual $r_0 := b - Ax_0$ and generating for all, or at least most n, until it possibly finds the exact solution, iterates x_n, such that

$$x_n - x_0 = p_{n-1}(A)r_0 \in \mathcal{K}_n(A, r_0)$$

with a polynomial p_{n-1} of degree $n - 1$. For some n, x_n may not exist or p_{n-1} may have lower degree. ∎

Remark 2.3.5. Observe that, by definition, $r_n := b - Ax_n$ and hence if the sequence of the *residuals*, r_n, converges to zero then x_n is *the* solution. ∎

One idea to derive the Krylov space solvers is through the projection methods to solve linear systems. Let $Ax = b$ be the linear system that we want to solve and suppose that $\mathcal{K}$ and $\mathcal{L}$ are two orthogonal linear subspaces of dimension m: $\mathcal{K} \perp \mathcal{L}$. A *projection method* is a process to find $\tilde{x}$ such that, under some conditions $\tilde{x} \in \mathcal{K}$ and the

residual, $r = b - A\tilde{x}$, is orthogonal to $\mathcal{L}$. In other words, the idea behind the projection methods is find $\tilde{x} \in \mathcal{K}$ such that $r \perp \mathcal{L}$ in the following way:

$$A^{-1}b \approx x_m = x_0 + p_{n-1}(A)r_0$$

and, of course, if the initial guess is zero then $A^{-1}b \approx p_{n-1}(A)b$. That is, Krylov space solvers must be understood as iterative methods to solve $Ax = b$ from an initial guess x_0.

Definition 2.3.6. The orthogonality condition over the residual, $r \perp \mathcal{L}$, is called the *Petrov-Glaerkin condition*. ∎

Let us make a final informal comment related to the "*shape*" and the "*size*" of the matrix A which will be the motivation for the following paragraph. There are several questions refereed to the matrix like is A symmetric? How are the eigenvalues? It is well conditioned? One can get easily its transpose? Depending on the sign of the answers, obviously, a concrete Krylov space solver method will be better than another one. For our purposes, since we are considering Daubechies wavelets, our matrix A will be very sparse (due to the compact support, nonsingular, non symmetric, with difficulties to calculate its transpose and with storage problems. It seems that the best adapted Krylov method to solve our (non)-linear system is the Generalized Minimum Residual Method. Since it is not very common its use and, also, for sake of completeness we have decided to include a quick overview of the GMRES algorithm.

The GMRES algorithm. As it has been already said, we will need to solve a linear system of equations because we are using the Newton's Method. The way to solve such linear system will be the Generalized Minimum Residual Method. In the following lines we will make a quick overview of such method. To this end, let $Ax = b$ be the linear system of equations that we want to solve, where A is $n \times n$ nonsingular matrix. The Generalized Minimum Residual Method, GMRES from now on, is a projection method where the subspaces $\mathcal{K}$ and $\mathcal{L}$ are taken to be $\mathcal{K}_m$ and $A\mathcal{K}_m$ respectively being

$$\mathcal{K}_m := \left\langle v, Av, A^2v, \ldots, A^{m-1}v \right\rangle, \; v = \frac{r_0}{\|r_0\|_2}.$$

In aim to create an orthonormal basis, one can perform the Arnoldi iteration given by the following algorithm.

The algorithm breaks down when w is the zero vector. This happens when the minimal polynomial of A is of degree m. As a sub-product, the Arnoldi iteration generates a $(m+1) \times m$ Hessenberg matrix, $\widetilde{H}_m$, whose entries different from zero are the coefficients $h_{i,j}$ of Algorithm 3.

Algorithm 3: The Arnoldi process, expressed in a vectorial way for convenience.

`Arnoldi(`A`,`V`)`**{**

Data: Matrix A and the previous m vectors, v_m, from the mk-th iterate of `gmres` stored in the matrix V.

Result: The new vector v_{m+1} from the Arnoldi process related to the m-th iterate of `gmres`.

for $j = 1$ **to** m **do**

$u = Av_j$;

$h_{1,m} = v_1^\top u, \ldots, h_{m,m} = v_m^\top u_m$;

$w = u - (h_{1,m}v_1 + \cdots + h_{m,m}v_m)$;

$h_{m+1,m} = \|w\|_2$;

$v_{m+1} = v_{m+1}/\|w\|_2$

end for

}

Proposition 2.3.7 (Arnoldi's Algorithm, Proposition 6.6 [Saa03]). *Let us consider a vector v_1 such that $\|v_1\|_2 = 1$ and it is n-dimensional. Assume that the Algorithm 3 does not stop before the m-th step. Then the vectors $v_1, v_2, \ldots, v_m$ form an orthonormal basis of the Krylov subspace*

$$\mathcal{K}_m := \left\langle v_1, Av_1, A^2v_1, \ldots, A^{m-1}v_1 \right\rangle.$$

Now, consider V_m to be the $n \times m$ matrix with column vectors $v_1, \ldots, v_m$ it can be shown that $AV_m = V_{m+1}\widetilde{H}_m$ since for $j = 1, 2, \ldots, m$

$$Av_j = \sum_{i=1}^{j+1} h_{i,j}v_i.$$

With this trick in mind let us derive, as it is done in Section 6.5 from [Saa03], the GMRES algorithm by considering $x \in x_0 + \mathcal{K}_m$ rewritten as $x = x_0 + V_m y$, where y is an m-vector. Now define

$$J(y) = \|b - A(x_0 + V_m y)\|_2$$

and, by construction of V_{m+1}, one has that

$$b - A(x_0 + V_m y) = r_0 - AV_m y = \beta v_1 + V_{m+1}\widetilde{H}_m y = V_{m+1}(\beta e_1 - \widetilde{H}_m y)$$

where $\beta = \|r_0\|_2$. Therefore, since V_{m+1} has its column vectors orthonormal it is true that $J(y) = \|b - A(x_0 + V_m y)\|_2 = \|\beta e_1 - \widetilde{H}_m y\|_2$. The GMRES gives the unique vector of $x_0 + \mathcal{K}_m$ such that $J(y)$ is minimum. That is, the iterative construction is the following

$$x_m = x_0 + V_m y_m,$$

where $y_m = \text{argmin}\|\beta e_1 - \widetilde{H}_m y\|_2$. We are ready to write down the GMRES algorithm.

Algorithm 4: GMRES algorithm. The matrix V are the vectors v_k stored in columns.

```
gmres(A, x, b, maxit){
    Data: Matrix A, x, b ∈ R^n and maxit ≤ n is the maximum numbers of iterates
          allowed.
    Result: x ∈ R^n such that minimizes ||Ax − b||_2.
    nb = ||b||_2, r = b − Ax_0, ρ = ||r||_2, v_1 = r/ρ, g = ρ · e_1;
    while k < maxit do
        Arnoldi(A,V);
        Compute y_m = argmin||ρe_1 − H̃_m y||_2;
    end while
    x = x_0 + V y_m;
}
```

Proposition 2.3.8 (GMRES Algorithm, Proposition 6.10 [Saa03]). *Let A be a nonsingular matrix. Then, the GMRES algorithm breaks down at step j, that is $h_{j+1,j} = 0$, if and only if the approximate solution x_j is exact.*

Remark 2.3.9. It is clear that Algorithm 4 is a very simplified version of the GMRES algorithm. In fact, with such presentation it is masked the tricky way to derive y_m. Actually, to do such calculation in an efficient way is by solving an upper triangular linear system of equations (see [Saa03, Section 6.5.3]). And, moreover, from such triangular linear system, derived from Givens transformations of the matrix $\widetilde{H}_m$, it can be controlled, at each iteration, how far away we are from the solution. ∎

As we will see in Chapter 5 we will need another solver: Transpose-Free Quasi-Minimal Residual method, TFQMR from now on. For sake of completeness, in Algorithm 5 it is displayed its complete implementation. However, we only describe some informal comments related to TFQMR because the underlying ideas of such method are similar as those ones of GMRES. First of all, let us state the following Corollary.

Corollary 2.3.10 (Corollary 3.6.1 [Kel95]). *Let A a $n \times n$ matrix and let $x_0, b \in \mathbb{R}^n$ given. Then within $(n+1)/2$ iterations the TFQMR iteration will either break down or terminate with the solution.*

That is, TFQMR goes faster to the solution. In fact, it can be seen that the residual of GMRES can control the quasi-residual of TFQMR. The name of quasi arises from the fact that the basis of Krylov subspace is constructed to be biorthogonal. Also, the biorthogonalization minimizes some breakdowns in the underlying Arnoldi procedure. However, if the memory it is not a problem then GMRES is often a good choice. In other words, is better to perform a good precondition strategy than try to accelerate the convergence (see e.g [Saa03]). In view of that and in order to conclude this kind of

Algorithm 5: TFQMR algorithm. Notice that `tol` is set to be $\varepsilon \|b\|_2$ for a given ε. The multiplications A against y_i, u_i (with $i = 1, 2$) must be interpreted as the construction of the Krylov subspace.

`tfqmr(`A, x, b, `maxit, tol)`**{**
  Data: Matrix $A, x, b \in \mathbb{R}^n$ and `maxit` $\leq n$ is the maximum numbers of iterates allowed.
  Result: $x \in \mathbb{R}^n$ such that quasi-minimizes $\|Ax - b\|_2$.
  $w = y_1 = r_0 = b - Ax_0$, $u_1 = v = Ay_1$;
  $\rho_0 = r_0^\top r_0,\ \tau = \|r_0\|_2,\ \theta = 0,\ \eta = 0$;
  while $k <$ `maxit` **and** $\tau\sqrt{m+1} <$ `tol` **do**
    $\sigma_{k-1} = r_0^\top v,\ \alpha = \rho_{k-1}/\sigma_{k-1},\ y_2 = y_1 - \alpha v,\ u_2 = Ay_2$;
    for $j = 1$ **to** 2 **do**
      $w = w - \alpha u_j,\ d = y_j + (\theta^2\eta/\alpha)d$;
      $\theta = \|w\|_2/\tau,\ c = 1/1 + \theta^2,\ \tau = \tau\theta c,\ \eta = c^2\alpha$;
      $x_k = x_{k-1} + \eta d$;
      Compute $\tau\sqrt{m+1}$;
    end for
    $\rho_k = r_0^\top w,\ \beta = \rho_k/\rho_{k-1},\ y_1 = w + \beta y_2,\ u_1 = Ay_1$;
    $v = u_1 + \beta(u_2 + \beta v)$;
  end while
}

paragraph, let us recall that a right precondition strategy which is, instead of solving the original linear system $Ax = b$ solve firstly $A\mathbf{P}^{-1}y = b$ and $\mathbf{P}x = y$ to get the desired solution x. Of course, a left precondition strategy is solve $\mathbf{P}Ax = \mathbf{P}b$. In both strategies, what is wanted is a better matrix or a simpler linear system than the original ones. Clearly, the above techniques can be reconfigured to use the precondition strategies. We will come back to such topic in Chapter 5.

2.3.3 Quadrature rules using wavelets

In what follows, we will explain how one can derive a quadrature rule using a Daubechies wavelet as a weight function. However, since our experiments do not yield to an efficient and fast method (see the final comments of the Chapter 4) it will be presented in a rough way for two reasons. In view of that, the reader may omit this part. The first reason is to underscore one way of the aforesaid inefficiency. The second one, is to show another usage of the filter $h[n]$. Concretely, the quadrature rule is constructed using $h[n]$ and, hence, the values of $\psi(x)$ are no longer needed.

Recall that there are several ways to create a quadrature rule and one of them is to allow that the distance between interpolation points not be constant, such as the Gaussian quadrature formulas which are, usually, more accurate. Roughly speaking, an

n-point Gaussian quadrature rule is constructed to give exact results for polynomials of degree at most $2n-1$. This is done by a choice of some special points x_i and create weights ℓ_i for $i=1,\ldots,n$. It is customary to use, as a domain of integration, for a Gaussian quadrature rule the interval $[-1,1]$ and therefore the method has the following form

$$\int_{-1}^{1} f(x)\,dx \approx \sum_{i=1}^{n} \ell_i f(x_i).$$

The problem is to find a correct $w(x)$ that provides us a "*good*" weight function in our case.

The first step to construct a quadrature rule is to give a sequence of orthogonal polynomials respect the inner product defined by the *positive* weight $w(x)$. But if one looks a picture of a Daubechies wavelet, $\psi(x)$, clearly it is not necessarily nonnegative and hence, in general, it cannot be a *canonical weight* function. Following [BBDK01, BBD+02, HV05], for example, one can fix the problem with the wavelets: if it is required that $w(x)\geq 0$, just define

$$w(x) = \psi(x) + c\chi_{\mathrm{supp}(\psi(x))}(x), \tag{2.16}$$

with $c>0$ sufficiently big to become $w(x)\geq 0$. This method is known as *lifting device* an was introduced in [BBDK01]. Therefore, if we pick a "*good*" c, we have a good expression for a weight given by Equation (2.16) and hence we are able to construct a sequence of orthogonal polynomials using a Daubechies wavelet as a weight defining the inner product by

$$\langle f,g\rangle_w = \int (f(x)\cdot g(x))\cdot w(x)\,dx. \tag{2.17}$$

Remark 2.3.11. Before continue, let us make a comment about the techniques that we will develop in the following. We have performed some quadrature formulas for refinable functions with filters $h[n]$ from the tables. But another step beyond can be done following the parametrization of $h[n]$ (see [Reg07, RS05]) instead of the "*explicit*" $h[n]$ for the Daubechies wavelets in the sense of Remark 1.1.12. The reason is simple, with this parametrization one expects the increase of the degree of accuracy on the computer calculations and also allows us the use of a symbolic calculator. ∎

Recall that an important tool on the topic of orthogonal polynomials on the Real Line is the management of the moments under the action of a weight. Now we are ready to write a rule to evaluate the moments for our "*special*" weight by a recursion. To this end, we will state some technical results to forge such recursion which leads us to derive the quadrature formulas.

Lemma 2.3.12. *Let us define* $I_k = \int_{\mathbb{R}} x^k \phi(x)\ dx$ *then the following relation on the* Continuous Moments *for the scaling function*

$$I_k = \frac{\sqrt{2}}{2^{k+1}-2} \sum_{n\in\mathbb{Z}} h[n] \sum_{i=0}^{k-1} \binom{k-1}{i} n^{k-1-i} I_i$$

is verified.

Proof. Let us notice that with this notation and by the admissibility condition we have that $I_0 = 1$. Now, by using Equation (1.5)

$$\begin{aligned}
\int_{\mathbb{R}} x^k \phi(x)\ dx &= \frac{1}{2^{k+1}} \int_{\mathbb{R}} t^k \phi\left(\frac{t}{2}\right)\ dt \\
&= \frac{1}{2^{k+1}} \int_{\mathbb{R}} x^k \phi\left(\frac{x}{2}\right)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \int_{\mathbb{R}} \sum_{n\in\mathbb{Z}} h[n] x^k \phi(x-n)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n] \int_{\mathbb{R}} (x+n)^k \phi(x)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n] \int_{\mathbb{R}} \sum_{i=0}^{k} \binom{k}{i} x^i n^{k-i} \phi(x)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n] \sum_{i=0}^{k} \binom{k}{i} n^{k-i} \int_{\mathbb{R}} x^i \phi(x)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n] \int_{\mathbb{R}} x^k \phi(x)\ dx + \\
&\quad \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n] \sum_{i=0}^{k-1} \binom{k-1}{i} n^{k-1-i} \int_{\mathbb{R}} x^i \phi(x)\ dx.
\end{aligned}$$

Then

$$\left(1 - \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n]\right) \int_{\mathbb{R}} x^k \phi(x)\ dx = \\ = \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} h[n] \sum_{i=0}^{k-1} \binom{k-1}{i} n^{k-1-i} \int_{\mathbb{R}} x^i \phi(x)\ dx.$$

But $\sum h[n] = \sqrt{2}$:

$$\int_{\mathbb{R}} x^k \phi(x)\ dx = \frac{\sqrt{2}}{2^{k+1}-2} \left(\sum_{n\in\mathbb{Z}} h[n] \sum_{i=0}^{k-1} \binom{k-1}{i} n^{k-1-i} \int_{\mathbb{R}} x^i \phi(x)\ dx \right).$$

Hence

$$I_k = \frac{\sqrt{2}}{2^{k+1}-2} \sum_{n\in\mathbb{Z}} h[n] \sum_{i=0}^{k-1} \binom{k-1}{i} n^{k-1-i} I_i$$

as we want to show. □

And, in the same way of the previous Lemma we have the following result.

Lemma 2.3.13. *Let us define* $\mu_k = \int_{\mathbb{R}} x^k \psi(x)\ dx$ *then the following relation on the* Continuous Moments *for the wavelet*

$$\mu_k = \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} g[n] \sum_{i=0}^{k} \binom{k}{i} n^{k-i} I_i,$$

where

$$I_m = \frac{\sqrt{2}}{2^{m+1}-2} \sum_{n\in\mathbb{Z}} h[n] \sum_{j=0}^{m-1} \binom{m-1}{j} n^{m-1-j} I_j$$

and $I_0 = 1$ *is verified.*

Proof. Let us notice that by using Equation (1.5) one has

$$\begin{aligned}
\int_{\mathbb{R}} x^k \psi(x)\ dx &= \frac{1}{2^{k+1}} \int_{\mathbb{R}} t^k \psi\left(\frac{t}{2}\right)\ dt \\
&= \frac{1}{2^{k+1}} \int_{\mathbb{R}} x^k \psi\left(\frac{x}{2}\right)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \int_{\mathbb{R}} x^k \sum_{n\in\mathbb{Z}} g[n]\phi(x-n)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} g[n] \int_{\mathbb{R}} x^k \phi(x-n)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} g[n] \int_{\mathbb{R}} (x+n)^k \phi(x)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} g[n] \int_{\mathbb{R}} \sum_{i=0}^{k} \binom{k}{i} x^i n^{k-i} \phi(x)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} g[n] \sum_{i=0}^{k} \binom{k}{i} n^{k-i} \int_{\mathbb{R}} x^i \phi(x)\ dx \\
&= \frac{\sqrt{2}}{2^{k+1}} \sum_{n\in\mathbb{Z}} g[n] \sum_{i=0}^{k} \binom{k}{i} n^{k-i} I_i
\end{aligned}$$

and I_i is done by Lemma 2.3.12. □

Now, since the weight has two parts, we need a final step to have a rule to evaluate the moments of $w(x)$ taking into account that c is sufficiently "*good enough*".

Lemma 2.3.14. *It is verified that*

$$m_k = c\int_a^b x^k\,dx = c\frac{b^{k+1}-a^{k+1}}{k+1}.$$

Proof. It is a straightforward computation. □

Summing up all the previous results, we can state the following proposition which will be one of the main tools to get the sequence of orthogonal polynomials respect the weight $w(x)$. It will be stated in terms of Daubechies wavelets but it can be applied to all kind of wavelets.

Proposition 2.3.15. *Let $\psi(x)$ a Daubechies wavelet of $p \geq 1$ vanishing moments and define a nonnegative weight function $w(x) = \psi(x) + c\chi_{\mathrm{supp}(\psi(x))}(x)$. Then the corresponding moments, for $k \geq 0$, are given by*

$$M_k := \int_{\mathbb{R}} w(x)x^k\,dx = \begin{cases} m_k & \text{if } 0 \leq k < p, \\ \mu_k + m_k & \text{otherwise.} \end{cases} \tag{2.18}$$

Proof. The proposition follows by a direct consequence of the addition property of the integral, Lemma 2.3.14 and Lemma 2.3.12. □

Now, we are ready to give a recurrence to forge orthogonal polynomials but let us start by mentioning that we have forced that $w(x)$ to be nonnegative on and it is measurable in Lebesgue's sense for which all moments exists since $\psi(x)$ and $\chi_{\mathrm{supp}(\psi(x))}(x)$ are continuous. In other words, we can apply the classic theory of orthogonal polynomials on the Real Line to construct a sequence of orthogonal polynomials $\{p_n(x)\}_{n\in\mathbb{N}}$ by the well known *three-term recurrence relation*.

Given $p_{-1}(x) \equiv 0$, $p_0(x) \equiv 1$ and $\beta_0 = 0$ the following three recurrence relation for $n \geq 1$:

$$p_{n+1}(x) = (x - \alpha_n)\,p_n(x) - \beta_n p_{n-1}(x), \tag{2.19}$$

where

$$\alpha_n := \frac{\langle x\cdot p_n(x), p_n(x)\rangle_w}{\langle p_n(x), p_n(x)\rangle_w}, \qquad \beta_n := \frac{\langle p_n(x), p_n(x)\rangle_w}{\langle p_{n-1}(x), p_{n-1}(x)\rangle_w}$$

and $\langle\cdot,\cdot\rangle_w$ denotes the scalar product respect $w(x)$ given by Equation (2.17), gives us a sequence of monic orthogonal polynomials of degree n.

Remark 2.3.16. Notice that if the polynomial $p_n(x)$ has degree n then the polynomial $x \cdot p_n(x) \cdot p_n(x)$ has degree $2n+1$. Hence, if $n < (p-1)/2$ and we are using a wavelet with p vanishing moments, α_n is computed by using *only* Lemma 2.3.14. In the same way, if $n < p/2$ then β_n is computed by using Lemma 2.3.14. Both quantities do not depend on c. Moreover, taking into account the above paragraph one has that if $n < (p-1)/2$ then the n first orthogonal polynomials of the three-term recurrence relation for $w(x) = \psi(x) + c\chi_{\text{supp}(\psi(x))}(x)$ coincide with the Legendre polynomials (which are a special case of Jacobi polynomials $(1-x)^\alpha(1+x)^\beta$). Perhaps modified, since the Legendre polynomials, as Jacobi polynomials, are constructed over the interval $[-1,1]$. ∎

Now, we are almost done with the orthogonal polynomials.

Lemma 2.3.17. *Let $p_n(x) = \sum_{i=0}^{n} a_{i,n}x^{n-i}$ a polynomial of degree n. Then*

$$P_{2n}(x) = p_n(x) \cdot p_n(x) = \sum_{k=0}^{2n} \sum_{i=\max(0,k-n)}^{\min(k,n)} a_{i,n} \cdot a_{k-i,n} x^{2n-k} \tag{2.20}$$

is a polynomial of degree $2n$ and

$$\int P_{2n}(x)w(x)\,dx = \sum_{k=0}^{2n} \sum_{i=\max(0,k-n)}^{\min(k,n)} a_{i,n} \cdot a_{k-i,n} M_{2n-k}.$$

Proof. The degree of $P_{2n}(x)$ is clearly the sum of the degree of $p_n(x)$ and $p_n(x)$ i.e $2n$. The values of the coefficients are given by the Cauchy product (essentially a discrete and finite convolution) where the lower and upper summation limits can be checked by inspection. On the other side, the integral is clearly done by definition of M_k and applying the well known properties of the integral. □

Notice that

$$\int x \cdot P_{2n}(x)w(x)\,dx = \sum_{k=0}^{2n} \sum_{i=\max(0,k-n)}^{\min(k,n)} a_{i,n} \cdot a_{k-i,n} M_{2n-k+1}$$

it is verified. Using the above Lemma, one expects that it is possible to compute the coefficients α_n and β_n "*directly*". But one more step can be done taking into account that

$$\begin{aligned} p_{n+1}(x) &= \sum_{i=0}^{n+1} a_{i,n+1}x^{n+1-i} = (x-\alpha_n)\,p_n(x) - \beta_n p_{n-1}(x) \\ &= (x-\alpha_n)\sum_{i=0}^{n} a_{i,n}x^{n-i} - \beta_n \sum_{i=0}^{n-1} a_{i,n-1}x^{n-1-i} \\ &= \sum_{i=0}^{n} a_{i,n}x^{n+1-i} - \alpha_n \sum_{i=0}^{n} a_{i,n}x^{n-i} - \beta_n \sum_{i=0}^{n-1} a_{i,n-1}x^{n-1-i}, \end{aligned}$$

we get the following proposition which gives us an "*explicit*" expression for the coefficients of the orthogonal polynomial of degree n, $p_n(x)$. Again, it will be stated in terms of Daubechies wavelets but it can be applied to all kind of wavelets.

Proposition 2.3.18. *Let us consider $w(x) = \psi(x) + c\chi_{\mathrm{supp}(\psi(x))}(x)$ as a weight function, where $c > 0$ is "good enough" and $\psi(x)$ is a Daubechies wavelet of $p \geq 1$ vanishing moments. Let*

$$p_{n+1}(x) = \sum_{i=0}^{n+1} a_{i,n+1} x^{n+1-i}$$

be a monic orthogonal polynomial given by Equation (2.19) of degree $n+1$, with $n \geq 2$. Then, the coefficients $a_{i,n+1}$ are recursively given by:

$$\begin{cases} a_{0,n+1} = 1, \\ a_{1,n+1} = a_{1,n} - \alpha_n, \\ a_{i,n+1} = a_{i,n} - \alpha_n a_{i-1,n} - \beta_n a_{i-2,n-1} \quad \textit{for } 2 \leq i \leq n, \\ a_{n+1,n+1} = -\alpha_n a_{n,n} - \beta_n a_{n-1,n-1} \end{cases} \tag{2.21}$$

where $\alpha_n = \frac{\lambda_n}{\varrho_n}$, $\beta_n = \frac{\varrho_n}{\varrho_{n-1}}$ with

$$\begin{aligned} \lambda_n &:= \sum_{k=0}^{2n} \sum_{i=\max(0,k-n)}^{\min(k,n)} a_{i,n} \cdot a_{k-i,n} M_{2n-k+1}, \\ \varrho_n &:= \sum_{k=0}^{2n} \sum_{i=\max(0,k-n)}^{\min(k,n)} a_{i,n} \cdot a_{k-i,n} M_{2n-k}. \end{aligned}$$

Proof. It is a direct consequence of the previous results, where the seeds for $a_{i,0}$ are given by the polynomials $p_{-1}(x)$, $p_0(x)$ and $p_1(x)$. □

Now, once we have a kind of explicit expression for the orthogonal polynomials we need some results concerning the zeros of each $p_n(x)$ to give the quadrature rule and this is the main topic of the following paragraphs. To this end, let $p_n(x)$ be one of the orthogonal polynomials given by Equation (2.21) and consider its zeros $x_{1,n}^{k_1}, \ldots, x_{i,n}^{k_i}, \ldots, x_{m,n}^{k_m}$ and perhaps with some multiplicity k_i. Since we have constructed the weight to be positive and using classical results on the theory of orthogonal polynomials on the Real Line (see [MM08, Theorem 2.2.5 and 2.2.6]) it follows that all of the zeros of $p_n(x)$ are real, distinct and they are located in the interior of the interval $[-p+1, p]$. Moreover, if $p_n(x), p_{n+1}(x)$ are two orthogonal polynomial of degree n and $n+1$, respectively, given by Equation (2.21) then the zeros of $p_n(x)$ and $p_{n+1}(x)$ interlace. That is,

$$x_{k,n+1} < x_{k,n} < x_{k+1,n+1} \quad \text{for } k = 1, \ldots, n \text{ and } n \in \mathbb{N}.$$

Precisely such zeros are the main ingredient for the quadrature rule. Indeed, following Chapter 5 of [MM08], let us introduce a n-point quadrature rule, respect the weight $w(x)$, given by

$$\int_{\text{supp}\,(w(x))} f(x)w(x)\,dx = \sum_{i=1}^{n} \ell_i f(x_i) + \mathcal{R}(f)_w$$

where x_i are the zeros of the orthogonal polynomial of degree n, the quantities ℓ_i are given by

$$\ell_i = \frac{1}{q_n'(x_i)} \int \frac{q_n(x)}{(x - x_i)} w(x)\,dx, \tag{2.22}$$

$q_n(x)$ is the *node polynomial* given by $q_n(x) = \prod_{i=1}^{n} (x - x_i)$ and $\mathcal{R}(f)_w$ is the remainder. This kind of construction is similar, or based on, to the *Barycentric Interpolation* which is a faster way to evaluate the Lagrange polynomial just by noticing that, in Equation (2.22), we are calculating the integral of a polynomial respect the weight $w(x)$ since

$$\frac{q_n(x)}{(x - x_i)} = \prod_{\substack{j=1 \\ j \neq i}}^{n} (x - x_j).$$

Let us explain a way to construct $q_n(x)$ in our setting. It is a straightforward (and tedious) check that each coefficient of $q_n(x)$, let us denote them by $a_{k,n}$, has $\binom{n}{n-k}$ terms to add and each term is the multiplication of $(n - k)$ quantities: the nodes x_i. In order to avoid the problem of the operations and the errors induced by the operations, let us recover the Newton's identities (see [Mea92]) for our special case just by considering the nodes x_i as parameters instead of variables. That is, one can consider the monic polynomial in z with roots $x_1, \ldots, x_n$:

$$\prod_{i=1}^{n} (z - x_i) = \sum_{k=0}^{n} (-1)^k a_{k,n} z^{n-k},$$

where the coefficients $a_{k,n}$ are given by the elementary symmetric polynomials in the roots just by setting $a_{k,n} = e_k(x_1, \ldots, x_n)$. After that, define the power sums of the roots by $s_k = p_k(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^k$. Hence, by Newton's identities the roots can be expressed recursively in terms of the coefficients of the polynomial, $a_{k,n}$, just by using

$$\begin{aligned}
s_1 &= a_{1,n}, \\
s_2 &= a_{1,n}s_1 - 2a_{2,n}, \\
s_3 &= a_{1,n}s_2 - a_{2,n}s_1 + 3a_{3,n}, \\
s_4 &= a_{1,n}s_3 - a_{2,n}s_2 + a_{3,n}s_1 - 4a_{4,n}, \\
&\vdots
\end{aligned}$$

Conversely, one can easily get the coefficients $a_{k,n}$ from the above recursion:

$$a_{k,n} = \frac{s_k - \left(\sum_{i=1}^{k-1} (-1)^{i+1} a_{i,n} s_{k-i}\right)}{(-k)^{k+1}} \text{ and } a_{0,n} = 1. \tag{2.23}$$

And that is what we are looking for: a kind of "closed" expression for the coefficients that allows us to take care of the propagation of the error and, consequently, the error must be under control. We can summarize all the above calculations with the following conjecture which is the main goal claimed at the beginning of this section: the rough quadrature rule with a Daubechies wavelet. Rough in the sense that, for example, $\mathcal{R}(f)_w$ or the constant c are not given. Moreover, some of the recursions presented becomes unreachable as n increases. This is the reason to state it as a Conjecture.

Conjecture 2.3.19. *Let us consider $w(x) = \psi(x) + c\chi_{\text{supp}(\psi(x))}(x)$ as a weight function, where $c > 0$ is "good enough" and $\psi(x)$ is a Daubechies wavelet of $p \geq 1$ vanishing moments. Let $q_n(x) = \prod_{i=1}^{n}(x - x_i)$ be its* node polynomial *of the n-the orthogonal polynomial generated by Equation (2.21), where x_i are its zeros. Finally, consider $q_{i,n}(x) = \frac{q_n(x)}{(x-x_i)}$. Then $\ell_i = \int_{\mathbb{R}} q_{i,n}(x)w(x)\,dx = \sum_{k=1}^{n-1} a_{k,n-1}M_k$, where the coefficients $a_{k,n}$ are given by Equation (2.23) and the continuous moments M_k are given by Proposition 2.3.15. Then the n-point Gauss quadrature generated by $\psi(x)$ is given by*

$$\sum_{i=1}^{n} \ell_i f(x_i) + \mathcal{R}(f)_w. \tag{2.24}$$

Remark 2.3.20. It is a direct consequence of Equation (2.23), integral properties and Proposition 2.3.15. However the problem appears on determining $\mathcal{R}(f)_w$ for our particular instances as we will see at the end of Chapter 4. ∎

Let us end the present chapter with some comments. Along this chapter we have been focused to give methodology to work and compute with wavelets in a fast way either in $\mathbb{R}$ or $\mathbb{S}^1$. However, some of them are well suited "*only*" for $\mathbb{R}$. What a first glance seems to be a good choice at the end will cause several round off errors as we will see in the Chapter 4. Moreover, focusing, on the aforesaid quadrature rule, which is an inherent idea from those ones in [BBDK01, BBD+02, HV05] but using Daubechies wavelet, it seems to be imperative to extend the quadrature rule to $\mathbb{S}^1$. As a matter of fact, our natural framework will be $\mathbb{S}^1$. Thus, as in the previous section, the above quadrature rule and, hence, all the results above must be translated to the unit circle and correctly stated (treating the circle as the interval $[0, 1]$ with endpoints identified and then making some change of variables) This is a future work topic.

Chapter 3

Statement of the Problem

In this chapter we will give a short compilation of the theoretical framework concerning on the kind of dynamical systems that we will have to deal along this memory. For the convenience of the reader and in order to facilitate access to the individual topics this chapter is rendered as self contained as possible although the explanation is tailored, as much as it is possible, for our particular needs and issues. We refer the reader to [CR09, FR11] for a more comprehensive account.

This chapter is divided in two differentiated blocks. In the first one we will give a survey of results and definitions faced to the tackle the problem by using Ergodic Theory. Also it is done a study of the geometric and regularity properties involved. The second part, already focused in our specific framework, it will be stated a series of definitions, properties and results, from the functional analysis subject. These concepts, besides they are the connection between the two parts of the chapter, will bring us clues for understanding the problems generated at Chapter 5. Also will be helpful for avoiding and solving them.

3.1 Dynamics: a shortcut

The existence of attractive non-continuous invariant graphs for the quasi periodically forced dynamical systems has generated a great interest since the theoretical results are still poorly developed in comparison periodically forced systems. This is in contrast with the numerical results and experiments that tries to ensure and justify properties of the system. To this end, we need to recall some concepts.

3.1.1 Skew products on the cylinder

Let $\mathbb{X}$ be a finite dimensional compact smooth Riemannian manifold and consider a map $f\colon \mathbb{X} \longrightarrow \mathbb{X}$, not necessarily invertible. A *(semi)-discrete Dynamical System* is determined by the iteration of f. Therefore, the k-th iterate of a point $x_0 \in \mathbb{X}$ will be given by $x_k = f(x_{k-1}) = f^k(x_0)$ with the usual convention that f^0 is the identity map and from such point of view, one can define the *forwards orbit* of a point x the set of all positive iterates of f. One can ask for many invariants objects of such f related to $\mathbb{X}$. A set $\Delta \subset \mathbb{X}$ is said to be *invariant* if $f(\Delta) = \Delta$ and $f^{-1}(\Delta) = \Delta$ which means that the set

Δ is *forwards invariant* and *backwards invariant*. Let us introduce another set related with positive orbits and the forward invariance: the ω-*limit set*. The *omega limit set* is the set of all possible accumulation points of the sequence $\{f^k(x)\}_{k\in\mathbb{N}}$. Finally, we will say that two dynamical systems, $f_1\colon \mathbb{X}_1 \longrightarrow \mathbb{X}_1$ and $f_2\colon \mathbb{X}_2 \longrightarrow \mathbb{X}_2$, are *topologically conjugate* if there exists a bijective function $h\colon \mathbb{X}_1 \longrightarrow \mathbb{X}_2$ such that $h \circ f_2 = f_1 \circ h$ and if h is onto one says that f_2 is *semi-conjugate* to f_1.

We are going to work with a special (semi)-discrete Dynamical Systems: *skew products over a rotation* which are bundle maps of the form

$$\begin{array}{rccc} \mathfrak{F}: & \mathbb{S}^1\times\mathbb{R} & \longrightarrow & \mathbb{S}^1\times\mathbb{R} \\ & (\theta,x) & \longmapsto & (R_\omega(\theta),F(\theta,x)), \end{array} \tag{3.1}$$

where $R_\omega(\theta)=\theta+\omega$ being $\omega\in\mathbb{R}$, $F\colon \mathbb{S}^1\times\mathbb{R}\longrightarrow\mathbb{R}$ is continuous and $\mathcal{C}^1$ with respect x. Finally, $\mathbb{S}^1=\mathbb{R}/\mathbb{Z}$ is the base space of the skew product and $\mathbb{R}$ is the fiber space which in many situations can be replaced by a finite interval $[a,b]\subset\mathbb{R}$. When R_ω is an irrational rotation, that is when $\omega\in\mathbb{R}\backslash\mathbb{Q}$ the system is called *quasi periodically forced skew product*. These systems have the property that any fiber, $\{\theta\}\times\mathbb{R}$, is mapped into another fiber, $\{R_\omega(\theta)\}\times\mathbb{R}$ and, moreover, it can be seen that for all subset $\Delta\subset\mathbb{S}^1\times\mathbb{R}$ then it follows that $\pi_\theta(\mathfrak{F}(\Delta)=R_\omega(\pi_\theta(\Delta))$ where $\pi_\theta\colon\mathbb{S}^1\times\mathbb{R}\longrightarrow\mathbb{S}^1$ denotes the projection with respect to the first component.

From such kind of systems we will be interested in the asymptotic behavior of the forward orbits $\mathfrak{F}$ which lands on graphs of a measurable functions, i.e pairs of the form $\mathcal{K}=\{(\theta,\varphi(\theta)) : \theta\in\mathbb{S}^1\}$, and we will say that the graph is a *torus*. If φ satisfies the functional equation, called *invariance equation*, $F(\theta,\varphi(\theta))-\varphi(R_\omega(\theta))=0$, then the torus $\mathcal{K}$ is invariant under $\mathfrak{F}$ and its inner dynamics is the rigid rotation R_ω. We want to emphasize that the above equation will be the main tool (and topic) in Chapter 5.

Before continuing the explanation we will give a definition of Lyapunov exponent which is a kind of measure of how the orbits grow in average. From such definition, thats why we are using it, it is possible to give, as we will see, an explicit formula to compute the maximal Lyapunov exponent for our systems.

Definition 3.1.1. Let $(f,\mathbb{X},\mathfrak{B})$ be a measurable (semi-)dynamical system such that $\mathbb{X}$ is a finite dimensional compact smooth Riemannian manifold, $\mathfrak{B}$ is the Borel σ-algebra defined on $\mathbb{X}$, and is a differentiable map from $\mathbb{X}$ to itself. Finally let T_x denote the tangent space, at point $x\in\mathbb{X}$, of $\mathbb{X}$. The *Lyapunov exponent* in a point $x\in\mathbb{X}$ and to the direction $v\in T_x\mathbb{X}$ is given by

$$\lambda(z,v)=\limsup_{n\to\infty}\frac{1}{n}\log\|Df^n(z)\cdot v\|\,.$$

And the *maximal Lyapunov exponent* of $x \in \mathbb{X}$ is defined as:

$$\begin{aligned}\lambda_{\max}(x) &= \limsup_{n\to\infty} \frac{1}{n} \log |Df^n(x)| \\ &= \max\left\{ \limsup_{n\to\infty} \frac{1}{n} \log \|Df^n(x)v\| : v \in T_x\mathbb{X} \setminus \{0\} \right\}.\end{aligned}$$

where $\|\cdot\|$ a vector norm and $|\cdot|$ a matrix norm compatible with it. ∎

It is important to notice that for all point $x \in \mathbb{X}$ and for all direction $v \in T_x\mathbb{X}$ Lyapunov exponent exists and their values, due the finite dimension of $\mathbb{X}$, do not depend on the choice of the norm. To derive the expression of the maximal Lyapunov exponent, notice that by Oseledec's Theorem (see [BP02, Theorem 2.1.1] for example) there exists a linear filtration, in terms of the m Lyapunov exponents, namely $-\infty \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m < \infty$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \|Df^n(x) \cdot v\|$$

exists for every v in the tangent space to $\mathbb{X}$ at x $v \in T_x\mathbb{X}$, and

$$\lim_{n\to\infty} \frac{1}{n} \log |\det(D_z f^n(x))| = \sum_{i=1}^{m} \lambda_i. \tag{3.2}$$

Thus, in our case, we have two Lyapunov exponents the vertical one, namely λ_v, and the other λ_o. For the first one it is easy to get an analytic expression:

$$\lambda_v = \limsup_{n\to\infty} \frac{1}{n} \log \left\| \begin{pmatrix} 1 & 0 \\ \dfrac{\partial x_n}{\partial \theta} & \dfrac{\partial x_n}{\partial x} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = \limsup_{n\to\infty} \frac{1}{n} \log \left| \frac{\partial x_n}{\partial x} \right|.$$

For the other Lyapunov exponent we can use Equation (3.2), as it is done in [CR09], to assert that $\lambda_o = 0$. Actually,

$$\lambda_v + \lambda_o = \limsup_{n\to\infty} \frac{1}{n} \log \left| \det \begin{pmatrix} 1 & 0 \\ \dfrac{\partial x_n}{\partial \theta} & \dfrac{\partial x_n}{\partial x} \end{pmatrix} \right| = \limsup_{n\to\infty} \frac{1}{n} \log \left| \frac{\partial x_n}{\partial x} \right| = \lambda_v.$$

And hence, we have that the desired maximal Lyapunov exponent is given by $\lambda_{\max} = \max(\lambda_v, \lambda_o)$ which is negative if and only if $\lambda_v \leq 0$.

Certainly, in [Wal82, Theorem 1.14] it is stated the Birkhoff Ergodic Theorem for non-invertible systems with σ-finite measure. Recall that a positive (or signed) measure μ on a σ-algebra $\mathfrak{B}$ of subsets of set $\mathbb{X}$ is called *finite* if $\mu(\mathbb{X})$ is a finite real number (rather than ∞). The measure μ is called *σ-finite* is $\mathbb{X}$ is the countable union of measurable sets with finite measure. Finally, a set in a measure space is said to have *σ-finite measure* if it is a countable union of sets with finite measure. We want to point out that since we will use the Birkhoff Ergodic Theorem several times in the forthcoming chapters we will state it for sake of completeness.

Theorem 3.1.2 (Birkhoff). *Suppose* $\mathbf{T}\colon (\mathbb{X}, \mathfrak{B}, \mu) \longrightarrow (\mathbb{X}, \mathfrak{B}, \mu)$ *to be a measure preserving application (where we allow* $(\mathbb{X}, \mathfrak{B}, \mu)$ *to be* σ*-finite) and* $f \in \mathscr{L}^1(\mu)$*. Then*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} f(\mathbf{T}^i(x)) = f^*(x)$$

converges μ*-almost everywhere, where* $f^* \in \mathscr{L}^1(\mu)$*. Also* $f^* \circ \mathbf{T} = f^*$ *almost everywhere and if* $\mu(\mathbb{X}) < \infty$ *then* $\int_{\mathbb{X}} f^* \, d\mu = \int_{\mathbb{X}} f \, d\mu$*. Moreover, if* $\mathbf{T}$ *is ergodic, then* f^* *is constant* μ*-almost everywhere. If* $(\mathbb{X}, \mathfrak{B}, \mu)$ *is a probability space and* $\mathbf{T}$ *is ergodic we have that for all* $f \in \mathscr{L}^1(\mu)$ *then*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} f(\mathbf{T}^i(x)) = \int_{\mathbb{X}} f \, d\mu$$

μ*-almost everywhere.*

Recall that, as usual, we say that $\mathbf{T}$ is *ergodic respect to* μ (or alternatively that μ is ergodic with respect to $\mathbf{T}$) if for every $B \in \mathfrak{B}$ with $\mathbf{T}^{-1}(B) = B$, either $\mu(B) = 0$ or $\mu(B) = 1$.

Along the main part of this memory, the second component of $\mathfrak{F}$ will be given by a product of two maps each one of them depending only on θ and x. That is, $F(\theta, x) = f(x)g(\theta)$ and we will assume that $f(0) = 0$ in order to make the unit circle $x \equiv 0$ $\mathfrak{F}$-invariant. In view of that, it makes sense the calculation of λ_v along $x \equiv 0$

$$\lambda_v = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \left(\log(|f'(0)|) + \log(|g(\theta_i)|)\right) = \log(f'(0)) + \int_{\mathbb{S}^1} \log|g(\theta)| \, d\mu, \tag{3.3}$$

where the last equality is by using Theorem 3.1.2 which forces the log-integrability of g to be part of the hypothesis. In view of that the circle $x \equiv$ is a repellor when λ_v is positive and imposing that f and g to be bounded $+\infty$ is also a repellor. Therefore, between them we must have an attractor different from $x \equiv 0$. In order to finish this section, since there has been a lot of discussion with the concept of *strangeness*, let us define rigorously what does strange non-chaotic attractor means for us.

Definition 3.1.3 ([AC09]). Let $(f, \mathbb{X}, \mathfrak{B})$ be a measurable (semi-)dynamical system such that $\mathbb{X}$ is a finite dimensional compact smooth Riemannian manifold, $\mathfrak{B}$ is the Borel σ-algebra defined on $\mathbb{X}$ is a differentiable map from $\mathbb{X}$ to itself and the measure μ is equivalent to the Lebesgue measure. A *Strange non-Chaotic Attractor* is a closed set $\mathcal{A} \subset \mathbb{X}$ such that

- *Strangeness*: $\mathcal{A}$ it is not a finite set of points neither a piecewise differentiable manifold.

- *non-Chaoticity*: The set of points in $\rho(\mathcal{A}) := \{x : \omega(x) \subset \mathcal{A}\}$, $\rho(\mathcal{A})$ is called *realm of attraction*, whose maximal upper Lyapunov exponent is positive has zero Lebesgue measure.
- *Attractor*: $\mathcal{A}$ is an attractor in the sense of Milnor. That is, the set $\rho(\mathcal{A})$ has positive Lebesgue measure and there is no strictly smaller closed set $\mathcal{A}' \subset \mathcal{A}$ such that $\mu(\rho(\mathcal{A}'))$ is positive.

■

In order to illustrate with examples, in fact is a theoretical result which can be found at [Har12, Proposition 4.3], and besides the well known Keller's Theorem (Theorem 3.2.1), which is the main topic of Section 3.2, let us give the following Proposition which is what it is meant for *Strange*.

Proposition 3.1.4 ([Har12]). *Define the following set*

$$Z_\varphi = \bigcap_{k>0} \{\theta \in \mathbb{S}^1 : \varphi(\theta) < 1/k\}.$$

Then, for all $\theta \notin Z_\varphi$, there exist $C_{\theta,x} > 0$ such that $|\varphi(\theta_n) - x_n| \leq C_{\theta,x}\alpha^n |\varphi(\theta) - x|$, where α is taken to be the supremum of $xf'(x)/f(x)$ with $x \in (0, M]$.

From the proof, $C_{\theta,x} = M/\min(x, \varphi(\theta))$. When $\varphi(\theta)$ has zeros, this constant is not uniform because of the density of zeros.

Finally, to conclude this generic part we will explain a numerical issue of Theorem 3.1.2. As we have said, λ_v along $x \equiv 0$ is easy to calculate in the sense that an analytical expression can be obtained when $F(\theta, x) = f(x)g(\theta)$. Following the same arguments for the aforesaid Lyapunov exponent on $x \equiv 0$ and using that

$$\frac{\partial x_n}{\partial x} = \prod_{k=0}^{n-1} f'(x_k)g(\theta + k\omega)$$

it can be derived, in the same way as Equation (1) from [Kel96], the vertical Lyapunov exponent along the invariant curve φ. Indeed, the formula is

$$\lambda_\varphi = \int_{\mathbb{S}^1} \log(|f'(\varphi(\theta))|) + \log(|g(\theta)|)\, d\theta$$

which, using again Theorem 3.1.2, is taken to be the limit of

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log(|f'(\varphi(\theta_k))|) + \log(|g(\theta_k)|).$$

But we can go a little bit in-depth with such expressions; if we define

$$\log\left(\frac{\partial x_k}{\partial x}\right) := \log(|f'(\varphi(\theta_k))|) + \log(|g(\theta_k)|)$$

then the n-th approximation of the Lyapunov exponent is given by

$$\lambda_\varphi^{n)} := \frac{1}{n}\sum_{k=0}^{n-1}\log\left(\frac{\partial x_k}{\partial x}\right).$$

Hence, following this thread, one can define the increment of $\lambda_\varphi^{n)}$ by $\Delta_\varphi^{n)} := \lambda_\varphi^{n)} - \lambda_\varphi^{n-1)}$. Therefore $\Delta_\varphi^{n)}$ must be tend to zero when n tends to infinity $\lambda_\varphi^{n)} = \lambda_\varphi^{n-1)} + \Delta_\varphi^{n)}$. On the other side, observe that

$$(n-1)\lambda_\varphi^{n-1)} = \sum_{k=0}^{n-2}\log\left(\frac{\partial x_k}{\partial x}\right).$$

Thus, we have that

$$\lambda_\varphi^{n)} = \frac{(n-1)\lambda_\varphi^{n-1)} + \log\left(\frac{\partial x_{n-1}}{\partial x}\right)}{n} = \lambda_\varphi^{n-1)} + \frac{\log\left(\frac{\partial x_{n-1}}{\partial x}\right) - \lambda_\varphi^{n-1)}}{n}$$

which allows us to redefine the aforementioned increment by

$$\Delta_\varphi^{n)} := \frac{\log\left(\frac{\partial x_{n-1}}{\partial x}\right) - \lambda_\varphi^{n-1)}}{n}. \tag{3.4}$$

Remark 3.1.5. In Chapter 6 we will perform an experiment with a model with an additive forcing. That is $F(\theta, x) = f(x)+g(\theta)$. Notice that the above expression, Equation (3.4), with the convenient changes also holds. ∎

Such way of writing will be helpful for the numerics done in the forthcoming chapters. To conclude this section we want to make a twofold comment. On the one hand we want to emphasize that with such definition the Strange non-Chaotic Attractors found in the literature verifies it (see [CR09] for further information and references). On the other side, one can relate the dynamical properties, and hence the validation of Definition 3.1.3, with a functional approach. This is, precisely the main topic of the Section 3.4.

3.2 Keller's techniques

The previous section contains a brief summary and some technical results of a family of pinched skew products which can generate, for some values of the parameters, a Strange Non-Chaotic Attractor. Now, let us see how the previous techniques are applied to a particular instance of a bi-parametric skew product:

$$\begin{pmatrix}\theta_{n+1}\\ x_{n+1}\end{pmatrix} = \mathfrak{F}_{\sigma,\varepsilon}(\theta_n, x_n) = \begin{pmatrix} R_\omega(\theta_n)\\ f_\sigma(x_n)g_\varepsilon(\theta_n)\end{pmatrix}. \tag{3.5}$$

Furthermore, we will try to have enough tools to face the problems that will arise in the following chapters. Moreover, the understanding of such phenomenons can be useful for the study of the invariant set generated by the System (3.5). Finally a word of caution must be done. Indeed, all this machinery can be seen as useless since in [Sta97, Sta99] some regularity properties are established for systems of such type. However, such tools are interesting to see how the discontinuities are created and also to classify theoretically, in terms of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, the invariant function. Related to such comment, the geometric properties of the invariant function will be a key point to determine such classification.

3.2.1 Keller's Theorem

As we have said in Section 3.1, since by hypothesis $f(0) = 0$, the system (3.5) has the circle $x \equiv 0$ invariant in contrast to $x \equiv \varepsilon$ which it is not invariant for $\varepsilon > 0$. On the other side, from Equation (3.3), the vertical Lyapunov exponent for almost every $\theta \in \mathbb{S}^1$ it is the logarithm of

$$\sigma := \begin{cases} f'(0)\exp(\int_{\mathbb{S}^1} g(\theta)d\theta) & \text{if } \int_{\mathbb{S}^1} g(\theta)d\theta > -\infty,\\ 0 & \text{otherwise.}\end{cases}$$

Therefore if $\sigma > 1$ the circle $x \equiv 0$ is a repellor. Moreover, by using Birkhoff Ergodic Theorem, recall that the Lyapunov Exponent at $x \equiv 0$ is

$$\kappa(f,g) := \int_{\mathbb{S}^1} \log\left|\frac{\partial f(x)g(\theta)}{\partial x}\Big|_{x=0}\right| d\theta = \log(f'(0)) + \int_{\mathbb{S}^1} \log|g(\theta)|\, d\theta.$$

When $\kappa(f,g)$ is positive, $x \equiv 0$ is a repellor for System (3.5). Moreover, since f and g are bounded, infinity is also a repellor and the system must have an attractor different from $x \equiv 0$. These attractors, which are the objects that we want to study, are typically very complicated. In order to fix ideas, we are going to restrict ourselves to the study of a particular subfamily of model (3.5), which is

$$\begin{pmatrix}\theta_{n+1}\\ x_{n+1}\end{pmatrix} = \mathfrak{F}(\theta_n, x_n)_{\sigma,\varepsilon} = \begin{pmatrix} R_\omega(\theta_n)\\ 2\sigma\tanh(x)\cdot(\varepsilon + |\cos(2\pi\theta)|)\end{pmatrix}, \tag{3.6}$$

with $\omega = \frac{\sqrt{5}-1}{2}$, $\sigma > 0$ and $\varepsilon \geq 0$. Apart from the parameter ε, it is the natural restriction to $\mathbb{R}^+$ of the system considered in [GOPY84] (see Figure 3.1, where a graph of the attractor of this system with $\sigma = 1.5$ and $\varepsilon = 0$ is shown).

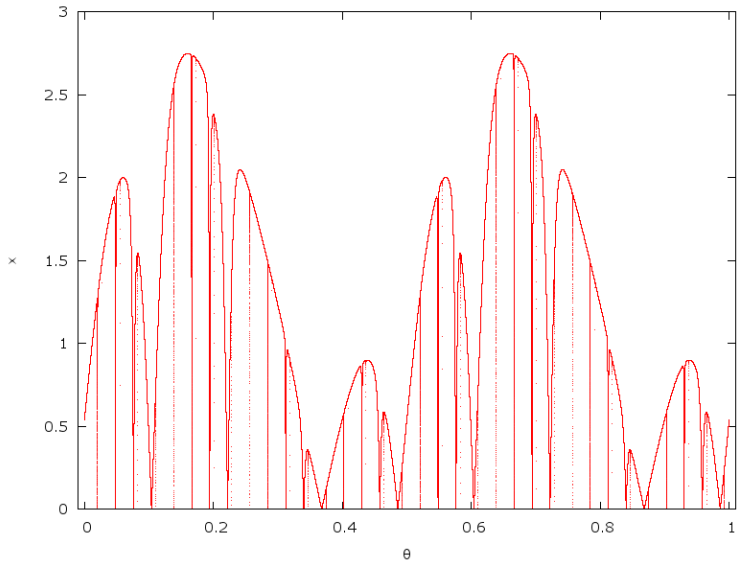


Figure 3.1: The attractor of System (3.6) $\sigma = 1.5$ and $\varepsilon = 0$. Notice the abrupt changes in the graph of the attractor.

Concretely, when $\varepsilon = 0$ the Lyapunov Exponent $\kappa(f, g)$ at $x \equiv 0$ is $\log(\sigma)$. Hence, the interesting case (for us) occurs when $\sigma > 1$. The attractor of System (3.5) and its dynamics is described by the following theorem:

Theorem 3.2.1 (G. Keller [Kel96]). *Under the above assumptions there exists an upper semi continuous function $\varphi\colon \mathbb{S}^1 \longrightarrow \mathbb{R}^+$ whose graph is invariant under System (3.5) and satisfies:*

(a) The Lebesgue measure on the circle, lifted to the graph of φ is a Sinai-Ruelle-Bowen measure (that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\mathfrak{F}^k(\theta, x)) = \int_{\mathbb{S}^1} f(\theta, \varphi(\theta))\, d\theta$$

for every $f \in \mathcal{C}^0(\mathbb{S}^1 \times \mathbb{R}^+, \mathbb{R})$ and Lebesgue almost every $(\theta, x) \in \mathbb{S}^1 \times \mathbb{R}^+$),

(b) if $\kappa(f, g) \leq 0$ then $\varphi \equiv 0$,

(c) if $\kappa(f, g) > 0$ then $\varphi(\theta) > 0$ for almost every θ,

(d) if $\kappa(f, g) > 0$ and g vanishes at some point then the set $\{\theta \in \mathbb{S}^1 : \varphi(\theta) > 0\}$ is meager and φ is almost everywhere discontinuous,

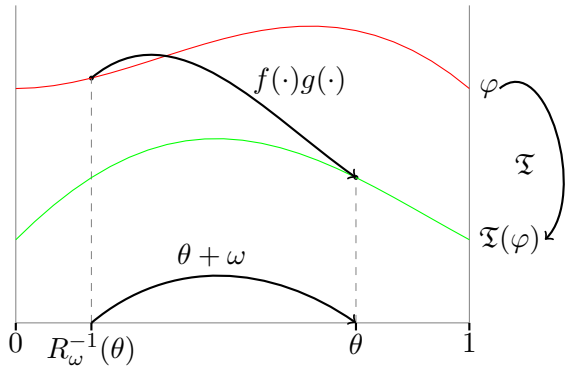


Figure 3.2: The notion of Transfer Operator $\mathfrak{T}$ from Keller's definition

(e) if $\kappa(f,g)>0$ and $g>0$ then φ is positive and continuous; if $g\in\mathcal{C}^1$ then so is φ,

(f) if $\kappa(f,g)\neq 0$ then $|x_n-\varphi(\theta_n)|\to 0$ exponentially fast for almost every θ and every $x>0$.

Observe that, when $\kappa(f,g)>0$ and g vanishes at some point (i.e. when the system is *pinched*), it follows from statements (c,d) that φ is discontinuous almost everywhere. In the particular case of System (3.6), the pinching condition implies that $\varepsilon=0$ and, since $|\cos(2\pi\theta)|$ vanishes for $\theta\in\left\{\frac{1}{4},\frac{3}{4}\right\}$, it follows that the set

$$\left\{(\tfrac{i}{4}+n\omega \pmod 1, 0)\,:\; n\in\mathbb{N},\; i\in\{1,3\}\right\} \tag{3.7}$$

is both a subset of the attractor and is dense (and invariant) in $x\equiv 0$. On the other hand, if $\varepsilon>0$ we can not have a dense set of *pinched* points.

The proof of the above theorem is based on the iteration of the *Transfer Operator* of the system. Since in Section 4.2 from Chapter 4 we will use this construction and it is the key point of the present section let us explain it. To do so, let $\mathcal{P}$ be the space of all functions not necessarily continuous from $\mathbb{S}^1$ to $\mathbb{R}$: $\mathcal{P}(\mathbb{S}^1,\mathbb{R})$. If we look for a functional version of the System (3.5) in the space $\mathcal{P}$ one can define the *Transfer Operator* $\mathfrak{T}\colon\mathcal{P}\longrightarrow\mathcal{P}$ as

$$\mathfrak{T}(\varphi)(\theta)=f(\varphi(R_\omega^{-1}(\theta)))\cdot g(R_\omega^{-1}(\theta)).$$

Remark 3.2.2. In Section 3.4 we will define the notion of Transfer Operator $\mathbf{T}$. The "*Keller*" notion of $\mathfrak{T}$ must be understood as the situation sketched in Figure 3.2. ∎

Remark 3.2.3. From the above definition we obtain

$$\mathfrak{T}(\varphi)(\theta)=\pi_x\left(\mathfrak{F}(R_\omega^{-1}(\theta),\varphi(R_\omega^{-1}(\theta)))\right)$$

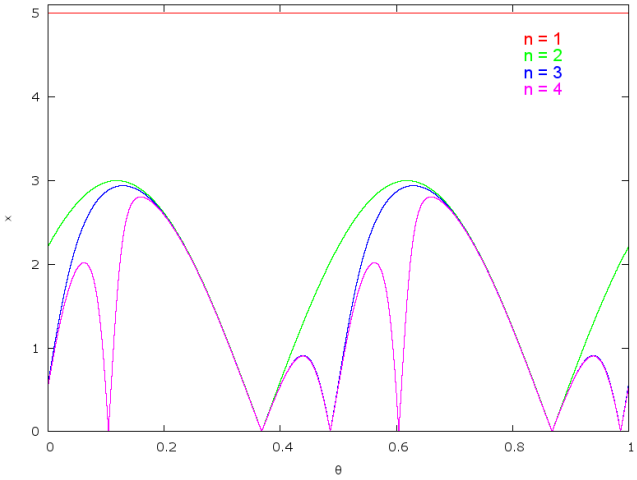


Figure 3.3: The constant function $c = 5$ and three iterations of the Transfer Operator $\mathfrak{T}$ for System 3.6 with $\sigma = 1.5$ and $\varepsilon = 0$. The function c is plotted in red, $\mathfrak{T}(c)$ in green, $\mathfrak{T}^2(c)$ in blue and $\mathfrak{T}^3(c)$ in magenta.

where $\pi_x \colon \mathbb{S}^1 \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ denotes the projection with respect to the second component. ∎

Notice that the graph of a function $\varphi \colon \mathbb{S}^1 \longrightarrow \mathbb{R}$ is invariant for the System (3.5) if and only if $\mathfrak{T}(\varphi) = \varphi$. Precisely, to obtain the map φ from Theorem 3.2.1, Keller takes a sufficiently large constant function $\varphi_0 = c$ (with $c > (\sup_{x\in\mathbb{R}} f(x))\left(\max_{\theta\in[0,1]} g(\theta)\right)$) and iterates it under the transfer operator $\mathfrak{T}$ (see Figure 3.3). In such a way he gets, since the map f is monotone, a non-increasing sequence of continuous maps given by

$$\varphi_k = \mathfrak{T}(\varphi_{k-1}) = \mathfrak{T}^k(c). \tag{3.8}$$

Then, following the Keller's proof, one has that

$$\varphi := \lim_{k\to\infty} \varphi_k = \inf_{k\to\infty} \varphi_k$$

exponentially fast. This idea will be one of the key points in our strategies for the estimation of the regularity of the two parameter family of Strange Non-Chaotic Attractors given by System (3.6) and for the knowledge of the shape of attractor which is, precisely, the main topic of the forthcoming sections.

3.3 Geometry of the non-pinched Transfer Operator

As we have said at the beginning of the current section, the sequence given by Equation (3.8) will be important for us in order to have an idea of how the invariant graph

of the System (3.5) is created although the techniques presented here can not be seen as a panacea. We have to mention that some of the ideas developed in the following are, for us and at the beginning, a tool for study where one has to apply the Gaussian quadrature rule developed at Section 2.3.3.

3.3.1 Stark's Theorem

As we have said, using [Sta97, Sta99], the *regularity* of the graph which contains the attractor for the family that we are considering, in the case *non-pinched*, is a graph such that has the same regularity as the skew product. We will completely justify the use of such techniques and results. This effort is justified because it will help us to understand some of the mechanisms of the regularity loss of φ. To this end, we will start with the general setting of [Sta97, Sta99] in the same way of Equation (3.5) consider the *skew product*:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} F(x_n) \\ G(x_n, y_n) \end{pmatrix}$$

where $x_n \in \mathbb{X}$ represents the state of the driving system and $y_n \in \mathbb{Y}$ is the state of the driven system. Precisely, assuming that there exists a F-invariant set $\Lambda \subset \mathbb{X}$, $\lambda < 0$ and $C > 0$, one can impose conditions on the driven system:

$$\mathtt{d}_{\mathbb{Y}}(G_x^n(y), G_x^n(y')) \leq Ce^{\lambda n}\mathtt{d}_{\mathbb{Y}}(y, y') \tag{3.9}$$

for all $x \in \Lambda \subset \mathbb{X}$ where $G_x^n(y) : \mathbb{Y} \longrightarrow \mathbb{Y}$ is defined

$$G_x^{n+1}(y) = G(f^n(x), G_x^n(y)),$$

where $G_x^1(y) = G_x(y) = G(x, y)$. Now, once a distance $\mathtt{d}_{\mathbb{Y}}$ is given, one can define, as usual, the α-Hölder condition for a function f, namely:

$$\mathtt{d}_{\mathbb{Y}}(f(y), f(y')) \leq K_{\mathbb{Y}}\mathtt{d}_{\mathbb{Y}}(y, y')^{\alpha},$$

where $K_{\mathbb{Y}} > 0$. With all these tools, we can state the following theorem, which is a slight modification of those appearing in [HPS77].

Theorem 3.3.1 (Theorem 1.2, [Sta99]). *Suppose $\mathbb{X}$ is a metric space, $F : \mathbb{X} \longrightarrow \mathbb{X}$ a homeomorphism such that*

(a) for all $x, x' \in \mathbb{X}$

$$\mathtt{d}_{\mathbb{X}}(F^{-n}(x), F^{-n}(x')) \leq Ce^{\mu n}\mathtt{d}_{\mathbb{X}}(x, x'),$$

where $C > 0$, $\mu \geq 0$ and $\mathtt{d}_{\mathbb{X}}(\cdot, \cdot)$ is a distance on $\mathbb{X}$,

$\Lambda \subset \mathbb{X}$ is a closed f-invariant set, $\mathbb{Y}$ is a complete metric space and $G : \mathbb{X} \times \mathbb{Y} \longrightarrow \mathbb{Y}$ is continuous function such that

(a) satisfies Equation (3.9) for all $x \in \Lambda$,

(b) for some $y_0 \in \mathbb{Y}$ *the function* $G(x, y_0) : \mathbb{X} \times \mathbb{Y} \longrightarrow \mathbb{Y}$ *is bounded and*

(c) suppose in addition that there exists $0 < \alpha \leq 1$ *such that for any bounded* $W \subset \mathbb{Y}$ *there exists a constant* $K_W \geq 0$ *satisfying*

$$\mathtt{d}_{\mathbb{Y}}(G(x, y), G(x', y)) \leq K_W \mathtt{d}_{\mathbb{X}}(x, x')^{\alpha}$$

Then there exists a bounded continuous function $\Phi : \Lambda \longrightarrow \mathbb{Y}$ *such that*

(a) the graph of Φ *is* (f, g)*-invariant,*

(b) the graph is attracting for all $(x, y) \in \Lambda \times \mathbb{Y}$ *and*

(c) Φ *is uniformly Hölder on* Λ *with exponent* γ*, for* $0 < \gamma \leq \alpha$*, such that* $\lambda + \gamma\mu < 0$*. In particular if* $\lambda < -\mu$ *and* $\alpha = 1$*, then* Φ *is Lipschitz.*

Let us mention that the above theorem has a bug for our purposes in the sense of the following Remark.

Remark 3.3.2. Recall that we are considering families of skew products that when $\varepsilon = 0$ the system becomes pinched. However, if one looks at the statement of the above Theorem it does not mention the possibility of the pinching condition, that is when $g(\theta)$ vanish. Therefore, one may think that the hypothesis are verified and hence the question of the regularity, when the System (3.5) is pinched, is solved. We claim that it is not true. In particular Equation (3.9) does not hold in the pinched case.

To prove this claim we will strongly use the invariance of the circle $x \equiv 0$ and the Keller's Theorem. Indeed, for all $\theta \in \mathbb{S}^1$, one has that

$$(x_n, \theta_n) \xrightarrow{n \to \infty} (\theta_n, \varphi(\theta_n)) \text{ exponentially fast}$$

and $\varphi(\theta) > 0$ almost everywhere $\theta \in \mathbb{S}^1$. That is, there exists $\delta > 0$ such that

$$\mu(\{\theta \in \mathbb{S}^1 \ : \ \varphi(\theta) > 0\}) > 0$$

where μ is the Lebesgue measure in $\mathbb{S}^1$. On the other hand, let

$$\#((n \in (0, 1, \ldots, n) \mid R^n_\omega(\theta) \in \{\theta \in \mathbb{S}^1 \ : \ \varphi(\theta) > 0\})$$

be the counter measure. By the ergodicity of R_ω, one has that the above quantity is infinite for almost everywhere $\theta \in \mathbb{S}^1$. Now, let θ be such that $G^n_\theta(x)$ converges to $\varphi(\theta_n)$

exponentially fast. Then if $n > n_0$ one has, by the above arguments, that $G_\theta^n(x) > \delta/2$. Moreover, recall that $\equiv 0$ is invariant, hence $G_\theta^n(0) = 0$ for all $n \in \mathbb{Z}$. Therefore

$$\mathrm{d}_{[0,M]}(G_\theta^n(x), G_\theta^n(0)) > \delta/2 \text{ if } n > n_0. \tag{3.10}$$

We wish to arrange that $Ce^{\lambda n}\mathrm{d}_{[0,M]}(x,0)$ be small enough. To this end, since $\lambda < 0$, let n_1 be such that

$$Ce^{\lambda n}x < \delta/2 \text{ if } n > n_1. \tag{3.11}$$

Now taking $n > \max(n_0, n_1)$, by Equation (3.10) and (3.11), the claim easily follows. Hence it is not true, in the pinched case, that Equation (3.9) holds as we wanted to show. ∎

Once all this stuff is introduced we can see that in our case how we can apply the above Theorem. As we have said at the introduction of this section, the Keller model will be our testing ground. Therefore, all the knowledge concerning to such model will be useful to accept or reject the numerical results in Chapters 4 and 5 and, hence, the "*goodness*" of the extrapolation to other models as it will be done in Chapter 6. Now, recall the definition of a distance in $\mathbb{S}^1$

$$\mathrm{d}_{\mathbb{S}^1}(x,x') = \begin{cases} |x-x'| & \text{if } |x-x'| \leq 1/2, \\ 1-|x-x'| & \text{if } |x-x'| \geq 1/2 \end{cases} \tag{3.12}$$

which induces $\mathbb{S}^1$ to be a metric space. We will redefine the sequence given by (3.8) as follows:

$$\varphi^k_{(\sigma,\varepsilon)} := \varphi_k, \text{ where } \varphi_k = \mathfrak{T}(\varphi_{k-1}) = \mathfrak{T}^k(c). \tag{3.13}$$

The above equation stands for the k-th iterate of the Transfer Operator, of the System (3.6), respect the parameters σ and ε. Now, consider the homeomorphism $F = R_\omega$ and set $G = 2\sigma\tanh(x)\cdot(\varepsilon + |\cos(2\pi\theta)|)$ to be the continuous function, $\mathbb{X} = \mathbb{S}^1$ with the distance given by Equation (3.12) and $\mathbb{Y} = [0, M]$ where the constant verifies $M > (\sup_{x\in\mathbb{R}} 2\sigma\tanh(x))\left(\max_{\theta\in[0,1]}(\varepsilon + |\cos(2\pi\theta)|\right)$. Finally, by the ergodicity of R_ω we will take $\Lambda = \mathbb{X} = \mathbb{S}^1$ as R_ω-invariant set.

Lemma 3.3.3. *With R_ω and $2\sigma\tanh(x)\cdot(\varepsilon+|\cos(2\pi\theta)|)$ the conditions of Theorem 3.3.1 are verified when $\varepsilon > 0$. That is, there exists a bounded continuous function $\Phi : \mathbb{S}^1 \longrightarrow [0,M]$ such that is uniformly Lipschitz which is $(R_\omega, 2\sigma\tanh(x)\cdot(\varepsilon+|\cos(2\pi\theta)|))$-invariant.*

Proof. We will prove all the items of the Theorem 3.3.1:

(a) We must check that $\mathtt{d}_{\mathbb{S}^1}(R_\omega^{-n}(\theta), R_\omega^{-n}(\theta')) \leq Ce^{\mu n}\mathtt{d}_{\mathbb{S}^1}(\theta,\theta')$ but observing that one has $|\theta - n\omega - (\theta' - n\omega)| = |\theta - \theta'|$. That is, take $C = 1$ and $\mu = 0$ then

$$\mathtt{d}_{\mathbb{S}^1}(R_\omega^{-n}(\theta), R_\omega^{-n}(\theta')) \leq e^{0\cdot n}\mathtt{d}_{\mathbb{S}^1}(\theta,\theta')$$

as we wanted to check.

(b) We must verify that $2\sigma\tanh(x_0)\cdot(\varepsilon + |\cos(2\pi\theta)|)$ is bounded for some point $x_0 \in [0, M]$ but this is clear since the map $(\varepsilon + |\cos(2\pi\theta)|)$ is continuous and it is considered on $\mathbb{S}^1$.

(c) We must check that there exists $0 < \alpha \leq 1$ and a constant $K_W \geq 0$ which verifies:

$$\mathtt{d}_{\mathbb{S}^1}(2\sigma\tanh(x)\cdot(\varepsilon + |\cos(2\pi\theta)|), 2\sigma\tanh(x)\cdot(\varepsilon + |\cos(2\pi\theta')|)) \leq K_W\mathtt{d}_{\mathbb{S}^1}(\theta,\theta')^\alpha.$$

To this end, by Mean Value Theorem. Moreover, since g is absolutely continuous thus is differentiable almost everywhere and in our case satisfies $|g'(\theta)| \leq K$ for almost all $\theta \in \mathbb{S}^1$, then g is Lipschitz continuous with Lipschitz constant at most K_W and therefore the assertion follows.

□

The reader can consider the above Lemma as a triviality since Keller's Theorem asserts the conclusion. However, since in Theorem 3.2.1 it is only required that $g(\theta)$ must be continuous one can make the following natural question:

It is possible to generalize the Theorem from [Sta99] for a large class of functions, namely the set of continuous functions or, moreover, for a functions such that belongs on a "*other*" class of Hölder?

There are examples with positive answer to the previous question. For example, and in the same way as item (e) from Example 1.2.4, let g (of the System (3.5)) be such that

$$\lim_{x\to 0}\frac{|x|^\alpha}{g(x)} = 0.$$

This means that $g(x)$ can not verify any Hölder condition but it is continuous. By Keller's Theorem using such g the invariant function φ is continuous when $\varepsilon > 0$. However, it can not verify any Hölder condition because of g. Concretely, fix a g with logarithmic cusp on a concrete θ_p. Then, using the Keller's construction such cusp is permanent along the iteration of the Transfer Operator.

In view of that, let us focus a little bit more the situation by considering the sequence given by $\{\varphi^k_{(\sigma,\varepsilon)}\}_{k\in\mathbb{N}}$ from Equation (3.13), once $\varepsilon > 0$ and σ are fixed such that the conditions of Theorem 3.2.1 are satisfied. Just by experimenting with the parameters

σ and ε, one can observe that "*there are not so many pinched regions*". Therefore the question is: how and where the spikes appear? Recall that at the beginning of Section 3.2.1, we have defined a dense set for the System (3.6), see (3.7), where as it can be checked in Figure 3.3 spikes appear over these points and consequently the differentiability problems arise.

In the same way, controlling the set where $g(\theta) = \varepsilon$ (that is when $|\cos(2\pi\theta)|$ vanish) is a good way to understand the mechanism of creation of spikes. As a consequence the loss of differentiability can be controlled. But this it is not almost true for all values of ε and σ. However, one can get more information about the wavelet coefficients of a concrete $\varphi^k_{(\sigma,\varepsilon)}$ (namely $d^k_j[n]$) and, thus, control the difference between $d^k_j[n]$ and $d^l_j[n]$:

$$|d^k_j[n] - d^l_j[n]| = \left| 2^{j/2} \int_{\mathrm{supp}(\psi)} \left(\varphi^k_{(\sigma,\varepsilon)}(2^j(x+n)) - \varphi^l_{(\sigma,\varepsilon)}(2^j(x+n)) \right) \psi(x)\, dx \right|. \tag{3.14}$$

More concretely, consider $\theta_p \in \mathbb{S}^1$, $U_{\theta_p} \subset \mathbb{S}^1$ a neighbourhood of θ and $\{\varphi^k_{(\sigma,\varepsilon)}\}_{k\in\mathbb{N}}$ be the sequence given by (3.13). Define, for $t \in [0,1]$, the following set

$$\mathbf{E}^{(k)}_i = \{\theta \in U_\theta \subset \mathbb{S}^1 \ : \ \varphi^k_{(\sigma,\varepsilon)}(t\theta_0 + (1-t)\theta_1) \leq \max\left(\varphi^k_{(\sigma,\varepsilon)}(\theta_0), \varphi^k_{(\sigma,\varepsilon)}(\theta_1)\right)\}.$$

Therefore the formal derivative $\varphi^k_{(\sigma,\varepsilon)}$ has a change of sign on U_θ (possibly with an asymptote if it is not defined). Notice that, allowing the possibility of the appearance of the asymptotes on the derivative the notion of weak derivative arise. Hence, the Sobolev spaces must be considered. On the contrary, it is natural to expect that the change of sign of the formal derivative φ'_k depends on σ and ε. That is, there are privileged zones where $\varphi_{(\sigma,\varepsilon)}$ is smooth enough. Moreover, in such U_{θ_p} is where it can be placed the logarithmic cusp (or any kind of singularity) and they can be used to classify φ. In other words, the above integral should be placed around the set that contains all neighborhoods U_{θ_p} for a concrete $\varphi^k_{(\sigma,\varepsilon)}$ and $\varphi^l_{(\sigma,\varepsilon)}$. As matter of fact, being $\varepsilon > 0$ this was the idea that we have in mind from when we have tried to apply the quadrature rule from Section 2.3.3 and with the use of Proposition 3.1.4 try to see the evolution of the wavelet coefficients as ε goes to zero. The counterpart is that such plan is "*feasible*" when $\varepsilon > 0$. Indeed, the "*great problem*" is when $\varepsilon = 0$ and was another reason to leave a part the Gaussian quadrature rules given in Section 2.3.3. The reason is the dense set of $\mathbf{E}^{(k)}_i$ with an infinite set of openings which, at the end, produces the upper semi continuity.

Drawing to close this section, recall that from Keller's proof if follows that

$$\varphi_k(\theta) \xrightarrow{k\to\infty} \varphi(\theta) \text{ exponentially fast}$$

but, moreover, Keller's Theorem proves, also, the pointwise convergence for almost $\theta \in \mathbb{S}^1$ which implies a kind of uniform convergence. Indeed, by Egorov's Theorem,

there is a close relation between the uniform convergence and the convergence almost everywhere of a sequence of Lebesgue measurable functions. We state it for completeness and we remark that such kind of convergence is called *almost uniform convergence*.

Theorem 3.3.4 (Egorov). *Let $(\mathbb{X}, \mathfrak{B}, \mu)$ be a measure space and let $E \subset \mathbb{X}$ be a measurable set with $\mu(E) < \infty$. Let $\{f_n\}_{n\in\mathbb{Z}}$ be a sequence of measurable functions on E such that each f_n is finite almost everywhere in E and $\{f_n\}_{n\in\mathbb{Z}}$ converges almost everywhere in E to a finite limit. Then for every $\varepsilon > 0$, there exists a subset A of E with $\mu(E - A) < \varepsilon$ such that $\{f_n\}_{n\in\mathbb{Z}}$ converges uniformly on A.*

We do not want to find such measurable sets but, again as it can be seen at Keller's proof, all sets $\{\theta \in \mathbb{S}^1 \,:\, \varphi(\theta) < \varepsilon\}$ are open. Therefore, this means that φ is continuous at each point that $\varphi(\theta) = 0$ (whether pinched or non-pinched) and, also, it is in $\mathscr{L}^1(\mathbb{S}^1)$. With such comment we mean that it seems that there are regions where one can apply standard techniques of uniform convergence on a very weird function. Having said that, we are ready to classify the regularity of φ.

Proposition 3.3.5. *The upper semi continuous function $\varphi\colon \mathbb{S}^1 \longrightarrow \mathbb{R}^+$ whose graph is invariant under System (3.5) is in $\mathscr{L}^1(\mathbb{S}^1)$. Moreover, is continuous at*

$$\mathbf{Z}_{(n)} := \{Z_g + kw \pmod 1 \,:\, k = 0,\, \ldots,\, n\} = \bigcup_{k=0}^{n} R_\omega^k(Z_g) \subset \mathbb{S}^1$$

where $Z_g \subset \mathbb{S}^1$ is the finite and discrete set where the continuous function g vanishes. Finally, φ is in $\mathscr{B}^0_{\infty,\infty}(\mathbb{S}^1)$ when $\varepsilon = 0$ and $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, with $s \in (0, 1]$ when $\varepsilon > 0$ being $s = 0$ for a certain subclass of functions.

Proof. The case $\varepsilon > 0$ is done by Theorem 3.3.1. On the other side, the pinched case (i.e $\varepsilon = 0$), it is done as follows. By Example 1.2.4 (a) we know that $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ with $s \leq 0$. But, since $\varphi \in \mathscr{L}^\infty(\mathbb{S}^1)$ and, by construction, it is integrable then one has that $\Phi \in \mathscr{B}^{1+s}_{\infty,\infty}$, where $\Phi' = \varphi$. The latter "*belonging property*" of Φ holds even with $s = 0$. Therefore, since φ is bounded then Φ is Lipschitz. That is, φ must be in $\mathscr{B}^0_{\infty,\infty}(\mathbb{S}^1)$ and moreover, when $s < 0$ $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ are distributional spaces and φ, by Keller's Theorem, is a *usual* function. □

We want to finish this first block with a few words related with the above Proposition and comments. From Keller's Theorem we already know that φ is an upper semi continuous function and, in turn, this implies which kind of singularities could have φ. Even though, from the fact that the limit function has a "*big quantity*" of discontinuities one may think that it has negative regularity but it turns out that φ is, at most worst possible regularity space, in $\mathscr{B}^0_{\infty,\infty}(\mathbb{S}^1)$. As a matter of fact, $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ with $s < 0$ can be understood as the dual of the ones with $s > 0$. Hence, if $s < 0$ then φ's such that

are in $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ must be attractors with area. We will come back to such topic at the end of Chapter 6. Nevertheless, in a far away scope from the regularity topics there are some interesting points for further studies such that as why φ it can not be extended to a graph of a continuous function on $\mathbb{S}^1$, give a characterization of $\theta \in \mathbb{S}^1$ such that $\varphi(\theta) > 0$, how weird are the continuity points of φ?

3.4 Spectral techniques

As we have said at the end most of the contents of the Section 3.1, can be reviewed in functional terms. To this end, let us rewrite the invariance equation:

$$\varphi(R_\omega(\theta)) - F(\theta, \varphi(\theta)) = 0. \tag{3.15}$$

Before continue the explanation, we want to remark the following obvious rewritings of the above equality as the next ones $F(\theta, \varphi(\theta)) - \varphi(R_\omega(\theta)) = 0$ or in the non rotated version $\varphi(\theta) - F((R_\omega^{-1}(\theta), \varphi(R_\omega^{-1}(\theta))) = 0$ and $F((R_\omega^{-1}(\theta), \varphi(R_\omega^{-1}(\theta))) - \varphi(\theta) = 0$. Such expressions of the invariance equation, besides its ink saving usefulness, must be taking into account when doing numerics in a safe and cheap way. Let us resume the explanation considering $\mathcal{C}^0(\mathbb{X}, \mathbb{Y})$ the space of continuous functions (or operators) between the Banach spaces $\mathbb{X}$ and $\mathbb{Y}$ and defining the operator defined as

$$\begin{array}{rccc} \mathbf{T}: & \mathcal{C}^0(\mathbb{S}^1, \mathbb{R}) & \longrightarrow & \mathcal{C}^0(\mathbb{S}^1, \mathbb{R}) \\ & \varphi(\theta) & \longmapsto & \mathbf{T}(\varphi)(\theta), \end{array} \tag{3.16}$$

where $\mathbf{T}(\varphi)(\theta) = \varphi(R_\omega(\theta)) - F(\theta, \varphi(\theta))$. From such point of view φ is an invariant torus if and only if $\mathbf{T}(\varphi)(\theta) = 0$. One can see (see [dlLO99]) that the operator given by Equation (3.16) is Fréchet differentiable with $\mathbf{DT}\colon \mathcal{C}^0(\mathbb{S}^1, \mathbb{R}) \longrightarrow \mathcal{C}^0(\mathbb{S}^1, \mathbb{R})$ defined by

$$\mathbf{DT}(\varphi)h(\theta) = h(R_\omega(\theta)) - \frac{\partial F(\theta, \varphi(\theta))}{\partial x} h(\theta)$$

as its derivative.

Recall that an operator $F\colon \mathbb{X} \longrightarrow \mathbb{Y}$, where $\mathbb{X}$ and $\mathbb{Y}$ are normed linear spaces, is called *Fréchet differentiable* at $x \in \mathbb{X}$ if there exists a bounded linear operator $\mathbf{D}F\colon \mathbb{X} \longrightarrow \mathbb{Y}$ which satisfies the following relation

$$\lim_{h \to 0} \frac{\|F(x+h) - F(x) - \mathbf{D}Fh\|_{\mathbb{Y}}}{\|h\|_{\mathbb{X}}} = 0.$$

The motivations of such definitions are the close relationship between the functional solutions of Equation (3.16) and the dynamical properties of the linearized dynamics

around φ as it is pointed out in [HdlL06a, HdlL06b, HdlL07, Mat68]. To this end, let us define the *transfer operator* $\mathcal{M}$ as the bounded linear operator

$$\begin{array}{rccc} \mathcal{M}: & \mathcal{C}^0(\mathbb{S}^1,\mathbb{R}) & \longrightarrow & \mathcal{C}^0(\mathbb{S}^1,\mathbb{R}) \\ & h(\theta) & \longmapsto & \mathcal{M}(h)(\theta), \end{array} \tag{3.17}$$

where $\mathcal{M}(h)(\theta) = M(R_\omega^{-1}(\theta))h(R_\omega^{-1}(\theta))$ and M is the *transfer matrix* which gives the linearized dynamics as follows:

$$\begin{array}{ccc} \mathbb{S}^1\times\mathbb{R} & \longrightarrow & \mathbb{S}^1\times\mathbb{R} \\ (\theta,x) & \longmapsto & (R_\omega(\theta),Mx) \end{array}$$

being $M = \frac{\partial F(\theta,\varphi(\theta))}{\partial x}$. The last formula is a linear cocycle which for our framework, the one dimensional setting, is close to simple and it is the (in fact one of the) promised relationship. Indeed, if M vanishes at least at one point then the cocycle is not invertible everywhere which implies that the dynamics collapses. On the other hand, M is invertible everywhere and its inverse is well-defined if $M \neq 0$. Finally, we will say that the linear cocycle is *reducible* if there exist a constant matrix Γ and a change of variables $\Psi(\theta)$ such that

$$\Psi^{-1}(R_\omega(\theta))\cdot M(\theta)\cdot\Psi(\theta) = \Gamma.$$

Such application it is known as the *Floquet transform*. Having introduced these concepts one step more can be done. Namely, the relation of functional analysis and the dynamical properties which will leads us to some practical issues in Chapter 5 as for example the continuity with respect perturbations of M of $\mathcal{M}$. But, the most important fact is that the spectrum of $\mathcal{M}$ is *assembled* by an union of annuli centered at zero. Moreover, 0 belongs to one of the annuli if M is not invertible. These ideas are compressed in the following Theorem.

Theorem 3.4.1 ([CL99, Mat68]). *Let $\mathcal{M}$ be a transfer operator over a rotation R_ω. Then, if ω is irrational the spectrum of $\mathcal{M}$ is rotational invariant.*

Now, let us drop the ball on the grass and let play with all these concepts that mixes the functional properties with the dynamical ones.

3.4.1 The one dimensional case

As we have said, if φ is an invariant curve of $\mathfrak{F}$ its linearized behavior is given by the *linear skew product*

$$\begin{array}{rccc} \mathfrak{L}_\mathfrak{F}: & \mathbb{S}^1\times\mathbb{R} & \longrightarrow & \mathbb{S}^1\times\mathbb{R} \\ & (\theta,x) & \longmapsto & (R_\omega(\theta),\Upsilon(\theta)x), \end{array}$$

where $\Upsilon(\theta) = \frac{\partial F(\theta,\varphi(\theta))}{\partial x}$. It can be seen that, being ω Diophantine, $\mathfrak{LF}$ is reducible if and only if Υ does not vanish. But it can be said a little it more despite of the not requirement to be Diophantine of ω. Recall that a number ω verifies a Diophantine condition if there exists $\alpha > 0$ and $\beta \geq 1$ such that $|q\omega - 2\pi p| \geq \alpha/|q|^{\beta}$ for all integers p and q.

Proposition 3.4.2 (Proposition 5 [JT08]). *Let $\mathcal{M}$ be the transfer operator associated to $\mathfrak{LF}$. Then the following statements are equivalent:*

(a) The linear skew product $\mathfrak{LF}$ is reducible.

(b) The spectrum of $\mathcal{M}$ does not contain 0 and coincides with the closure of the set of eigenvalues.

(c) $\mathcal{M}$ has a non zero eigenvalue.

(d) $\mathcal{M}$ has a real non zero eigenvalue.

The above Proposition and the following Theorem, as we have said and it will be seen, will have important implications in Chapter 5. Before the statement of the following Theorem, which relates the growth of the orbits and the loss of reducibility, let us bear in mind that the *spectral radius* of a bounded linear operator $\mathcal{M}$, namely $\varsigma(\mathcal{M})$, is the supremum among the absolute values of the elements in its spectrum, namely $\mathrm{Spec}\,(\mathcal{M})$.

Theorem 3.4.3 (Theorem 3.8 and 3.9 [JT08]). *Let Λ be the Lyapunov exponent of $\mathfrak{LF}$ and $\mathcal{M}$ be its transfer operator associated, where it suffices that $\mathcal{M}\colon \mathcal{C}^0(\mathbb{S}^1,\mathbb{R}) \longrightarrow \mathcal{C}^0(\mathbb{S}^1,\mathbb{R})$.*

(a) Then, it is verified that $\varsigma(\mathcal{M}) = e^{\Lambda}$.

(b) If Υ vanish then $\mathrm{Spec}\,(\mathcal{M}) = \{z \in \mathbb{C} : \ |z| \leq e^{\Lambda}\}$ and $\mathfrak{LF}$ it is not reducible.

It is important to remark that our goal in Chapter 5 will not be to derive results related to reducibility methods. However, we will use a discretization of Equation (3.15), and hence of the transfer operator given by Equation (3.16). Therefore, and as a consequence of that, the typology of problems and properties which we have described up to now will be inherited for our particular setting.

Finally, to conclude the current chapter and in order to motivate the following one we can recover Proposition 3.1.4 which can be sketched as "*when the transfer vanishes all becomes complicated*". Thus, understanding the evolution of the complication can gives use utile ideas for the Chapters 4 and 5 and, also, the ones that we have depicted here.

Part II

Wavelets against Dynamical Systems

Chapter 4

Filtering Strange Non-Chaotic Attractors

Up to now, we have focused in gather together several tools and problems concerning to wavelets and quasi-periodically forced skew products on the cylinder. Now, we can combine them. But, how we can get the wavelet coefficients of the invariant object φ? We will get the coefficients applying verbatim the first idea depicted in Section 2.3. That is, we will use the Fast Wavelet Transform to perform an algorithm to control whether $\varphi \in \mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ even for $s \leq 0$ doing ad-hoc techniques which are valid in a general setting.

However, it is necessary to check where are the limits of such methodology. That is, we will try to relate the quality of the coefficients computed with the Fast Wavelet Transform with the regularity of φ. Roughly speaking we will say that the coefficients are good enough if they can recover the regularity of φ. Precisely, at the end of this chapter we will make a reflection of the viability of such method. Also, and in the same way of the above comments, we will conclude the chapter with some words of caution referred to the use of the quadrature rules. Indeed, along this memory we have made some overtones about the viability of using such techniques. We will completely justify when one can use them at the end of the present chapter.

4.1 An Algorithm to estimate regularities on $\mathscr{L}^\infty$

In [dlLP02], numerical implementations of wavelet analysis to estimate the *positive* regularity of conjugacies between critical circle maps are done. Due to Theorem 1.2.11 and Remark 1.2.9, we can generalize such techniques to any value (positive or not) of the regularity measured in terms of the Besov Spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$. The algorithm described below explains how to implement this generalization.

Assume that a function $f \in \mathscr{L}^\infty(\mathbb{S}^1)$ is given. Hence, one has

$$|\langle f, \psi_{j,n}\rangle| < \infty$$

for all $j, n \in \mathbb{Z} \times \mathbb{Z}$. Then we can perform the following steps:

Step 1. Compute the coefficients

$$d^{\mathrm{PER}}_{-j}[n] = \langle f^{\mathrm{PER}}, \psi_{-j,n}\rangle$$

for $j = 0, \dots, J-1$ and $0 \leq n \leq 2^j - 1$.

Step 2. By using the coefficients $d_{-j}^{\mathrm{PER}}[n]$ from Step 2, calculate

$$s_{-j} = \log_2 \left(\sup_{0 \leq n \leq 2^j - 1} \left| d_{-j}^{\mathrm{PER}}[n] \right| \right)$$

for $j = 0, \dots, J-1$.

Step 3. Make a linear regression to estimate the slope τ of the graph of the pairs $(-j, s_{-j})$ with $j = 0, \dots, J-1$. Then, when there is evidence of linear correlation between the variables $-j$ and s_{-j}, we set $s = \mathrm{R}(\tau)$.

If $k > \max(s, 5/2 - s)$ then, by Theorem 1.2.11, $f \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ and, hence, f has regularity s. Otherwise we need to repeat the algorithm with a Daubechies wavelet having a larger value of k until $k > \max(s, 5/2 - s)$.

Remark 4.1.1. As we have said, there is a lot of methods and literature of how one can get the wavelets coefficients. Here, we will use the Fast Wavelet Transform. In view of that, we can perform the following steps which are the same as Section 4.1 but we have added a previous step:

Step Previous. Compute $a_{-J}^{\mathrm{PER}}[n] := \langle f^{\mathrm{PER}}, \phi_{-J,n} \rangle$ for $0 \leq n \leq 2^J - 1$. This can be done, for instance, by using Lemma 2.3.2. ∎

After that, one has a method to decide the quality of the coefficients based on the quality of the regularity given by the above algorithm, the previous knowledge of the regularity of the function and, what we have called, the reconstruction of the regularity method. Indeed, by Theorem 1.2.8, $f^{\mathrm{PER}} \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ if and only if there exists $C > 0$ such that

$$\sup_{n \in \mathbb{Z}} |\langle f^{\mathrm{PER}}, \psi_{j,n} \rangle| \leq C 2^{\tau j},$$

even for $s = 0$. Therefore, given an s the coefficients $\langle f, \psi_{j,n} \rangle$ must be on a (infinite) straight line of slope τ. On the other side, recall that in the simple regression model of n points $\{x_i, y_i\}$, where $i = 1, 2, \dots n$, the aim is to find the equation of the straight line $y = \alpha + \beta x$ which would give the *best* fit for the data points. Usually, α and β are taken to be

$$\hat{\beta} = \frac{\mathrm{Cov}[x, y]}{\mathrm{Var}[x]} \quad \text{and} \quad \hat{\alpha} = \bar{y} - \hat{\beta}\, \bar{x}, \tag{4.1}$$

where $\mathrm{Cov}[\cdot, \cdot]$ is the *covariance* of the data $\mathrm{Var}[\cdot]$ is the *variance* and $\bar{\cdot}$ is the *standard average*. With all this things in mind, we can perform the following quality method (for the wavelet coefficients).

Method 4.1.2 (Quality test). Given an $f^{\mathrm{PER}} \in \mathscr{B}^s_{\infty,\infty}(\mathbb{R})$, set $\hat{\beta} = s$ in Equation (4.1) and calculate $\hat{\alpha}$ with the coefficients from **Step 2**. ∎

Remark 4.1.3. This criteria/method can be used to see the how the norm of f evolves. Indeed, from Theorem 1.2.8, there exists a $C > 0$ such that

$$\sup_{n\in\mathbb{Z}} |\langle f^{\mathrm{PER}}, \psi_{j,n}\rangle| \leq C2^{\tau j}.$$

Set $\hat{\alpha}$ to be $\log_2(C)$. Therefore, if C goes to ∞ then $\hat{\alpha}$ also. ∎

To test the quality of this algorithm, and hence get the degree of accuracy of the wavelet coefficients, we will try it with the Weierstraß function since we have an explicit expression for it, we have an analytic formula for its regularity in terms of its parameters and it is "*strange*" everywhere. This idea is borrowed from [dlLP02], but since we use more data than [dlLP02] we reproduce the example.

Example 4.1.4. From Section 1.2.1 (in particular Example 1.2.4 (d)) we know that $\mathfrak{W}_{A,B} \in \mathscr{B}^{-\log_B(A)}_{\infty,\infty}(\mathbb{R})$. To test the algorithm we fix the parameter $B = 2$ and we take $A \in [0.56745, 0.86475]$. Hence, $\mathfrak{W}_{A,2} \in \mathscr{B}^{s}_{\infty,\infty}(\mathbb{R})$ with $s = -\log_2(A) \in [0.2051\ldots, 0.8174\ldots]$. Then, observe that

$$1 < \max\left(s, \frac{5}{2} - s\right) = \frac{5}{2} - s < 3.$$

Therefore the above algorithm is valid in this case only for Daubechies Wavelets with $k \geq 3$ vanishing moments.

To perform the above algorithm we take $J = 30$ (that is, we use a sample of the graph of $\mathfrak{W}_{A,2}$ of 2^{30} points). To carry out Step Previous, by Lemma 2.3.2, we can estimate

$$a_{-J}[n] \approx 2^{-J/2}\mathfrak{W}_{A,2}(2^{-J}n).$$

Then, after executing Steps 1-3 of the algorithm we obtain the results depicted in Figure 4.1. We want to remark that the best numerical estimate of the regularity of $\mathfrak{W}_{A,2}(x)$ with $A \in [0.56\cdots, 0.86\cdots]$ computed with a Daubechies wavelet of 10 vanishing moments is obtained for $A = 0.86\cdots$ (that is, when the regularity is closer to zero). The fact that we have to work with the Daubechies wavelet of 10 vanishing moments can be explained as follows. Daubechies Wavelets with higher vanishing moments have bigger domain and regularity (see [Mal98]) and, hence, they are less adapted to approximate the Weierstraß function, which has highly concentrated oscillations. It turns out that the value of 10 vanishing moments is the best adapted (in the sense that minimizes the error) to the Weierstraß function for the range of parameters considered.

We also want to remark that all the computed Pearson correlation coefficients are bigger than 0.999. This agrees with the fact that the Weierstraß function is self-similar. Then, the coefficients $d_j[n]$, (as pointed out in Remark 1.2.10) must be approximately on a straight line. This is indeed the case as Figure 4.2 shows for a particular case. It turns out that the Daubechies wavelet with 10 vanishing moments also maximizes globally the

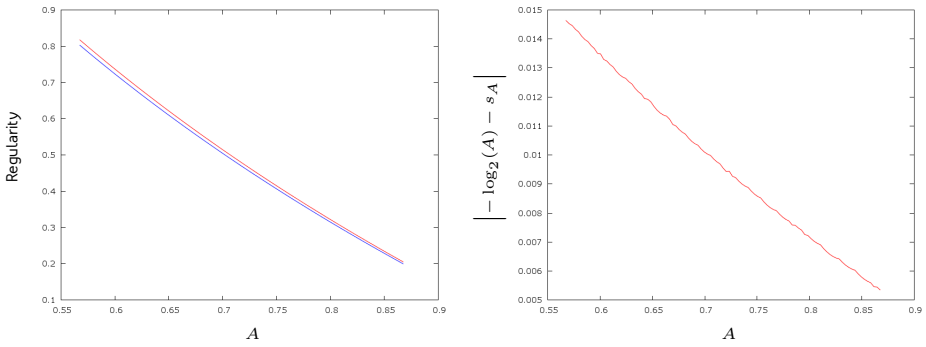


Figure 4.1: On the left picture the theoretical and estimated regularity of $\mathfrak{W}_{A,2}$ with $A \in [0.56\cdots, 0.86\cdots]$ are shown. The theoretical curve is plotted in blue and the *numerical* one in red. The estimated regularity is computed with a Daubechies Wavelet with 10 vanishing moments. On the right picture the *Error* function $|-\log_2(A)-s_A|$ is plotted (here s_A denotes the estimated regularity of $\mathfrak{W}_{A,2}$). Notice that the error is decreasing as the regularity gets closer to zero.

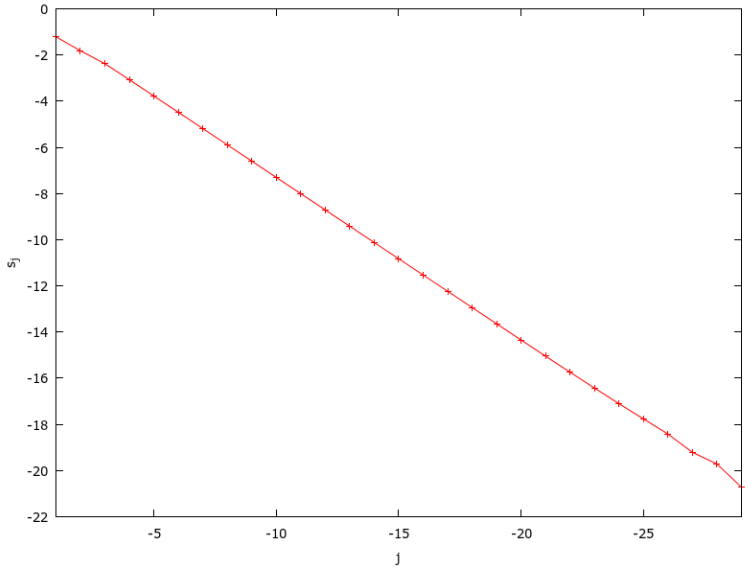


Figure 4.2: The graph of the pairs (j, s_j) with $-29 \leq j \leq 0$ for $\mathfrak{W}_{0.86745,2}$ (the regularity is 0.205147).

computed Pearson correlation coefficients in the range of parameters that we consider. Moreover, since the theoretical regularity is known, one can compare the straight line of the *theoretical* regularity against the *numerical* one (given by the pairs (j, s_j)).

With all these comments in mind one has that the wavelet coefficients (with 10 vanishing moments) of the Weierstraß function are closer enough to the real ones because the difference between the *numerical* regularity and the *theoretical* one is small enough. However, the last sentence is kind of act of faith in the sense that the small enough depends on the quality of the coefficients and, in our setting, they depends on the quality of the Fast Wavelet Transform. In other words, it is important to notice that, after some experiments that we have done there are differences between Fast Wavelet Transform of the `GNU-GSL` library and the "*real wavelet*" value. Therefore, the wavelets coefficients are given with error which are similar to those ones represented in Figure 4.1. ∎

4.2 The *numerical* regularity of the attractors

As it has been already said we aim at estimating the regularity of the attractor of System (3.6) in order to have a degree of accuracy of the wavelet coefficients of an invariant object. Before continue the explanation, let us recall it

$$\begin{pmatrix} \theta_{n+1} \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} R_\omega(\theta_n) \\ 2\sigma \tanh(x) \cdot (\varepsilon + |\cos(2\pi\theta)|) \end{pmatrix}.$$

By Keller Theorem, this attractor is the graph of a map $\varphi\colon \mathbb{S}^1 \longrightarrow \mathbb{R}^+$. Thus we have to use the algorithm described in the previous section applied to the function φ^{PER} (see Lemma 1.2.7).

But φ^{PER} in the pinched case is discontinuous almost everywhere (and the corresponding attractor is called *strange*). Therefore, we are *not* allowed to apply verbatim the algorithm from the previous section. In the rest of this section we will describe how to solve this problem in the implementation of the algorithm from the previous section.

We introduce the following notation for the wavelet coefficients of ϱ^{PER}:

$$a_j^{\mathrm{PER}}[n] := \langle \varrho^{\mathrm{PER}}, \phi_{j,n} \rangle \quad \text{and} \quad d_j^{\mathrm{PER}}[n] := \langle \varrho^{\mathrm{PER}}, \psi_{j,n} \rangle$$

for $j, n \in \mathbb{Z}$.

To compute an approximation of the type (2.15) for φ^{PER}, since we do not have an explicit formula for φ, we will use Theorem 3.2.1(f) and the transfer operator to get a sufficiently good numerical approximation of this function. Indeed, by Theorem 3.2.1(f), for almost every $\theta_0 \in \mathbb{S}^1$, any $x_0 > 0$ and any $\varepsilon > 0$ there exists N_0 such that for every $n \geq N_0$ we have:

$$|x_n - \varphi(\theta_n)| < \varepsilon$$

where $(\theta_n, x_n) = \mathfrak{F}^n(\theta_0, x_0)$. Moreover, the points (θ_n, x_n) with $n \in \{N_0, N_0 + 1, \dots, N_0 + 2^J - 1\}$ approximate exponentially fast the points $(\theta_n, \varphi(\theta_n))$ from $\mathrm{graph}(\varphi)$. Therefore,

$$\left\{(\theta_n, x_n) : n = N_0, N_0 + 1, \dots, N_0 + 2^J - 1\right\} \tag{4.2}$$

is an approximate mesh of $\mathrm{graph}(\varphi)$ provided that J is large enough. Precisely, this large enough must be estimated, given a tolerance, using Equation (3.4). To fix the mesh we choose a random point θ_0 and we fix some x_0 such that

$$x_0 > 2\sigma(\varepsilon + 1) \geq \left(\sup_{x\in\mathbb{R}} 2\sigma \tanh(x)\right)\left(\max_{\theta\in[0,1]}(\varepsilon + |\cos(2\pi\theta)|)\right).$$

However, this approximate mesh has two problems to be used in our computations:

Problem (1) as we will see, we need a mesh of the graph of φ at dyadic points of the form $i2^{-J}$ for $i = 0, 1, \dots, 2^J - 1$,

Problem (2) we can not use Lemma 2.3.2 to estimate the initial coefficients $a_{-J}[n]$ since our map φ is discontinuous almost everywhere (and, hence, not Lipschitz). Moreover, if one considers the closure of φ some topological artifacts can appear.

4.2.1 A solution to Problem (1): a $\mathcal{C}^1$ homeomorphism

As we have said, we need a mesh of the graph of φ at dyadic points of the form $\theta_i = i2^{-J}$ for $i = 0, 1, \dots, 2^J - 1$ but, clearly, if we obtain the points (θ_n, x_n) just as iterates of a single point by $\mathfrak{F}$ this condition is not satisfied. The natural approach which would be to approximately compute the points of the graph of φ based at the dyadic points by interpolating the obtained values is not feasible since, by Theorem 3.2.1, we know that φ is upper semi-continuous and discontinuous everywhere. Then we propose the following solution which consists in moving to a conjugate system with the desired properties. To do this, first we relabel the points $\{(\theta_n, x_n)\}_{n=N_0}^{N_0+2^J-1}$ to a sequence $\{(\widetilde{\theta}_i, z_i)\}_{i=0}^{2^J-1}$ so that

$$0 \leq \widetilde{\theta}_0 < \widetilde{\theta}_1 < \cdots < \widetilde{\theta}_{2^J-1} < 1$$

(we do this simply by sorting the data (4.2) with respect to the first coordinate; see Remark 4.2.3). In particular if $n \in \{N_0, N_0 + 1, \dots, N_0 + 2^J - 1\}$ and $i = i(n) \in \{0, 1, \dots, 2^J - 1\}$ is such that $\widetilde{\theta}_i = \theta_n$, then $z_i = x_n$.

Now we consider a $\mathcal{C}^1$ homeomorphism $h\colon \mathbb{S}^1 \longrightarrow \mathbb{S}^1$ such that $h\left(i2^{-J}\right) = \widetilde{\theta}_i$ for $i = 0, \dots, 2^J - 1$, $h(1) = \widetilde{\theta}_0 + 1$. Such map h can be obtained by taking h to be, for instance, a cubic spline in the intervals $\left[i2^{-J}, (i+1)2^{-J}\right]$ for $i = 0, \dots, 2^J - 1$.

Clearly, $\left\{\left(i2^{-J}, z_i\right)\right\}_{i=0}^{2^J-1}$ is now an approximate mesh of $\mathrm{graph}(\varrho)$ with $\varrho = \varphi \circ h$, based at the dyadic points. Thus, *we will use the list of pairs* $\left\{\left(i2^{-J}, z_i\right)\right\}_{i=0}^{2^J-1}$ *to estimate the regularity of* ϱ.

Remark 4.2.1. The map ϱ has the following dynamical interpretation. Consider the homeomorphism $H\colon \mathbb{S}^1 \times \mathbb{R}^+ \longrightarrow \mathbb{S}^1 \times \mathbb{R}^+$ defined by $H(\theta, x) = (h(\theta), x)$. Then, one can check that, $\mathrm{graph}(\varrho)$ is the attractor of the dynamical system

$$\left(H^{-1} \circ \mathfrak{F} \circ H\right)(\theta, x) = (h^{-1}(R_\omega(h(\theta))), f(x)g(h(\theta))),$$

which is conjugate to System (3.6). ∎

To obtain the regularity of φ we need to relate the regularities of ϱ and φ in terms of the Besov Spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$. To do this we use the fact that h is a $\mathcal{C}^1$ homeomorphism and that a homeomorphism is a bijective mapping of $\mathscr{B}^s_{\infty,\infty}(\mathbb{R})$ onto itself (we refer the reader to Section 4.3 from [Tri06] for a more detailed explanation). More precisely,

Proposition 4.2.2. *Let* $f \in \mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ *with* $s \in \mathbb{R}$ *and let* $h\colon \mathbb{S}^1 \longrightarrow \mathbb{S}^1$ *be a* $\mathcal{C}^m$ *diffeomorphism with* $m \geq s$. *Then* $f \circ h$ *belongs to* $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$.

Thus, by Proposition 4.2.2 and Proposition 3.3.5, the regularity of φ and ϱ coincide and we can estimate the regularity of ϱ by using the mesh $\left\{\left(i2^{-J}, z_i\right)\right\}_{i=0}^{2^J-1}$.

We remark that the exact formula for h is irrelevant for our algorithm. We only use the fact that such a map h exists and the fact that it can be taken $\mathcal{C}^1$. To obtain the data mesh $\left\{\left(i2^{-J}, z_i\right)\right\}_{i=0}^{2^J-1}$ that approximates ϱ (the attractor of the conjugate system) we simply have to sort the obtained mesh $\{(\theta_n, x_n)\}_{n=N_0}^{N_0+2^J-1}$ for φ with respect to the first coordinate θ and replace $\widetilde{\theta}_i$ by $i2^{-J}$. Of course, this does not add any further computational error to the mesh other than the errors coming from the iteration of the system and truncation errors derived from the choice of J. Furthermore, the exponential contraction of the system to the attractor (see Theorem 3.2.1(f)) still holds for the conjugate system, thus assuring that there is no loss of precision when replacing φ by ϱ (see Subsection 4.2.2).

Remark 4.2.3. The process of sorting the data of an array of 2^{30} points from $\mathbb{S}^1 \times \mathbb{R}^+$ (stored as pairs of `double` variables in `C` programming language) turns to be the bottleneck of the whole algorithm (and the most time consuming task of the whole program). Moreover, even the process of computing and filling the array with the initial mesh of the function φ already spends a "visible" amount of CPU time. Indeed, the iteration, storing and sorting process (with a standard sort algorithm like Heapsort) of this data spends about 2200 CPU seconds, with a remarkable variability which depends on the initial sorting of the data, in a computer with a Xeon processor at 3 GHz and 32 Gb

of RAM memory. In order to reduce the time elapsed in the sorting process we use the following trick based on the fact that the dynamical system generating the θ_i data is the irrational rotation R_ω. In this case we know that the Lebesgue measure is the unique ergodic measure of R_ω and, hence, its averaged spatial distribution is uniform and it is controlled approximately by the Birkhoff Ergodic Theorem applied to the Lebesgue measure. Indeed, we have

$$\sharp\left(\left\{\theta, R_\omega(\theta), \dots, R_\omega^{k-1}(\theta)\right\} \cap \left[\tfrac{i}{N}, \tfrac{i+1}{N}\right)\right) \approx \frac{k}{N}$$

for k large enough and for every $i \in \{0, 1, \dots, N-1\}$. The interpretation of this equation is that the statement

$$\sharp\left(\left\{\theta_{N_0}, \theta_{N_0+1}, \dots, \theta_{N_0+2^J-1}\right\} \cap \left[\tfrac{i}{2^J}, \tfrac{i+1}{2^J}\right)\right) = 1 \tag{4.3}$$

holds with high frequency for J large enough (observe that in this case we have $\left\{\theta_{N_0}, \theta_{N_0+1}, \dots, \theta_{N_0+2^J-1}\right\} = \left\{\theta_{N_0}, R_\omega(\theta_{N_0}), \dots, R_\omega^{2^J-1}(\theta_{N_0})\right\}$). Moreover, when (4.3) holds, we have $i = \left\lfloor 2^J \theta_l \right\rfloor$, where θ_l is the unique element from the set $\left\{\theta_{N_0}, \theta_{N_0+1}, \dots, \theta_{N_0+2^J-1}\right\} \cap \left[\tfrac{i}{2^J}, \tfrac{i+1}{2^J}\right)$ and $\lfloor\cdot\rfloor$ denotes the integer part function. This observation gives a good "hash function" and the following efficient algorithm to store and sort the data $\{(\theta_n, x_n)\}_{n=N_0}^{N_0+2^J-1}$. First, for $n = N_0, N_0+1, \dots N_0+2^J-1$ we compute the point $(\theta_n, x_n) = \mathfrak{F}(\theta_{n-1}, x_{n-1})$. Then, we store it in the position $i = \left\lfloor 2^J \theta_n \right\rfloor$ of the array data, if this slot is free. Otherwise, we store the point (θ_n, x_n) in a free position $j = j(i)$ of the array data such that $|j - i|$ is minimal. According to the above observations this will happen with low frequency and the array data will be almost sorted. Moreover, the positions of the array data which are not sorted are close the place where they should be when the array is sorted. This is the situation when the direct insertion sorting algorithm can be used with very good results. This means that we are using a method of order $\mathcal{O}(2^J + d)$ where d is the number of insertions (which are very low due to the way we have stored all data) instead of a method of order $\mathcal{O}(J2^J)$ as the Heapsort algorithm.

With this trick, the iteration, storing and sorting process lasts about 300 CPU seconds, almost without variability, which clearly improves the efficiency of the program. ∎

4.2.2 A solution to Problem (2): calculating the coefficients $a_{-J}^{\mathrm{PER}}[n]$ of ϱ^{PER}

When ϱ is regular enough, Lemma 2.3.2 gives $2^{-J/2}\varrho\left(\frac{n}{2^J}\right)$ as an estimate for the coefficients $a_{-J}^{\mathrm{PER}}[n]$. Recall that Theorem 3.3.1 and the next-to-last item of Theorem 3.2.1 ensures this good behavior. But, as we have pointed out, φ (and hence ϱ^{PER}) is discontinuous almost everywhere and the above estimate of $a_{-J}^{\mathrm{PER}}[n]$ is, a priori, not valid.

However, as we will see, the element $z_n \approx \varrho\left(n2^{-J}\right)$ from our data give indeed a good estimate for $a_{-J}^{\mathrm{PER}}[n]$ because our mesh is based at the dyadic points $n2^{-J}$.

As it has been already said in Section 3.2.1, φ is the pointwise limit of a non-increasing sequence of continuous (and, hence, uniformly continuous) functions $\varphi_k\colon \mathbb{S}^1 \longrightarrow \mathbb{R}^+$ defined by

$$\varphi_0(\theta) = c \qquad \text{and} \qquad \varphi_{k+1}(\theta) = \mathfrak{T}(\varphi_k)(\theta)$$

for every $\theta \in \mathbb{S}^1$ and $c > \sup_{x\in\mathbb{R}} \tanh(x)$. Consequently, $\varrho(\theta) = \lim_{k\to\infty} \varphi_k(h(\theta))$ for every θ.

Remark 4.2.4. If we take $x_0 = c = \varphi_0(\theta_0)$ then $x_k = \varphi_k(\theta_k)$ for every $k \geq 1$. To see it notice that, from the definition of the points (θ_n, x_n) and $\mathfrak{F}$, we get

$$\theta_k = R_\omega(\theta_{k-1}) \qquad \text{and} \qquad x_k = \pi_x(\mathfrak{F}(\theta_{k-1}, x_{k-1}))$$

for every $k \geq 1$. Assume that $x_{k-1} = \varphi_{k-1}(\theta_{k-1})$ fore some $k \geq 0$. Then, by Remark 3.2.3,

$$\begin{aligned} x_k &= \pi_x(\mathfrak{F}(\theta_{k-1}, x_{k-1})) = \pi_x(\mathfrak{F}(\theta_{k-1}, \varphi_{k-1}(\theta_{k-1}))) \\ &= \mathfrak{T}(\varphi_{k-1})(R_\omega(\theta_{k-1})) = \varphi_k(\theta_k). \end{aligned}$$

■

Since the scaling function ϕ of a Daubechies Wavelet is continuous, so is $\phi_{-J,n}$ for each n. Hence, from the definition of the coefficients $a_{-J}^{\mathrm{PER}}[n]$ and the Dominated Convergence Theorem we have:

$$\begin{aligned} a_{-J}^{\mathrm{PER}}[n] &= \int_{\mathrm{supp}(\phi_{-J,n})} (\varphi \circ h)^{\mathrm{PER}}(\theta)\phi_{-J,n}(\theta)\, d\theta \\ &= \lim_{k\to\infty} \int_{\mathrm{supp}(\phi_{-J,n})} (\varphi_k \circ h)^{\mathrm{PER}}(\theta)\phi_{-J,n}(\theta)\, d\theta \\ &= \lim_{k\to\infty} a_{-J}^{k,\mathrm{PER}}[n], \end{aligned}$$

where $a_{-J}^{k,\mathrm{PER}}[n] := \left\langle (\varphi_k \circ h)^{\mathrm{PER}}, \phi_{-J,n} \right\rangle$. From the proof of the Dominated Convergence Theorem, it can be shown that $a_{-J}^{k,\mathrm{PER}}[n]$ converge exponentially fast to $a_{-J}^{\mathrm{PER}}[n]$. Therefore, if k is large enough, by Lemma 2.3.2 we have

$$a_{-J}^{\mathrm{PER}}[n] \sim a_{-J}^{k,\mathrm{PER}}[n] \approx 2^{-J/2}(\varphi_k \circ h)^{\mathrm{PER}}(n2^{-J}) = 2^{-J/2}\varphi_k(h(n2^{-J}))$$

for $n = 0, \ldots, 2^J - 1$ (where $\sim$ means exponentially close).

From the definition of h it follows that, given $n \in \{0, 1, \ldots, 2^J - 1\}$, there exists $k \in \{N_0, N_0 + 1, \ldots, N_0 + 2^J - 1\}$ such that $h\left(n2^{-J}\right) = \widetilde{\theta}_n = \theta_k$. Therefore, by Remark 4.2.4,

$$\varphi_k\left(h\left(n2^{-J}\right)\right) = \varphi_k(\theta_k) = x_k = z_n.$$

Hence, if N_0 is large enough,

$$a_{-J}^{\mathrm{PER}}[n] \approx 2^{-J/2}\varphi_k(h(n2^{-J})) = 2^{-J/2}z_n$$

for $n = 0, \ldots, 2^J - 1$. This gives the necessary approximation of the coefficients $a_{-J}^{\mathrm{PER}}[n]$ to initialize the algorithm.

Remark 4.2.5. Such issue of the application of the Dominated Convergence Theorem is what we meant by control the difference between the wavelet coefficients of the Transfer Operator in Equation (3.14). However, as we will see in the forthcoming sections and at the end of the current chapter, the explicit calculation becomes very difficult. This comment is another way to motivate Chapter 5. ∎

4.2.3 A summary on the implementation of the algorithm

Step Previous 1. Fix $\sigma > 1$, choose a random $\theta_0 \in [0, 1)$ and $x_0 > 1$ and, by using the recurrence $(\theta_n, x_n) = \mathfrak{F}(\theta_{n-1}, x_{n-1})$, generate the data

$$\left\{(\theta_n, x_n) : n = N_0, N_0 + 1, \ldots, N_0 + 2^J - 1\right\},$$

with N_0 given by Equation (3.4) with a tolerance 10^{-14} and $J = 30$ (or apply a generic transient N_0).

Step Previous 2. Sort the above data to obtain a sequence $\{(\widetilde{\theta}_n, z_n)\}_{n=0}^{2^J-1}$ so that

$$0 \leq \widetilde{\theta}_0 < \widetilde{\theta}_1 < \cdots < \widetilde{\theta}_{2^J-1} < 1,$$

and delete the concrete values of the points $\widetilde{\theta}_n$. This defines a map ϱ with the same regularity that the map φ such that $\varrho\left(n2^{-J}\right) \approx z_n$ for $n = 0, \ldots, 2^J - 1$.

Step Previous 3. Set $a_{-J}^{\mathrm{PER}}[n] \approx 2^{-J/2}z_n$ for $n = 0, \ldots, 2^J - 1$.

Now, Steps 1-3 of the algorithm from Section 4.1 remain unaltered. As a result we get an estimate of the regularity of the (strange) attractor of System (3.6) for the chosen value of σ and ε.

4.3 Conclusions and results

We have performed two kind of exercises. The first one it has been the application verbatim of the above algorithm. The second one, once we have seen the results, we have tested the quality of the wavelet coefficients by using Method 4.1.2.

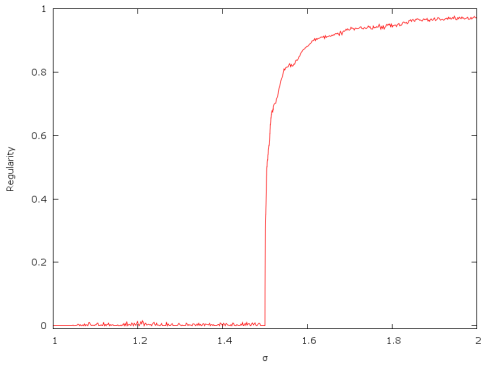


Figure 4.3: The estimate of the regularity $\mathrm{R}(\widetilde{s})$ of the (strange) attractor of System (3.6) for $\sigma \in [1, 2]$ and ε given by the parametrization $\varepsilon(\sigma)$. The results are obtained by using a sample of 2^{30} points (that is, $J = 30$), a transient $N_0 = 10^5$ and the Daubechies Wavelet with 16 vanishing moments. For this number of vanishing moments we obtain the minimum variance of Pearson correlation coefficient. It can be seen, by direct inspection of the picture, that when the system becomes pinched Method 4.1.2 does not give the seal of approval to the wavelet coefficients. However the pinching value $\sigma = 1.5$ is detected.

4.3.1 A verbatim application

We have used the above algorithm with the System (3.6) for $\sigma \in [1, 2]$ and ε given by the function

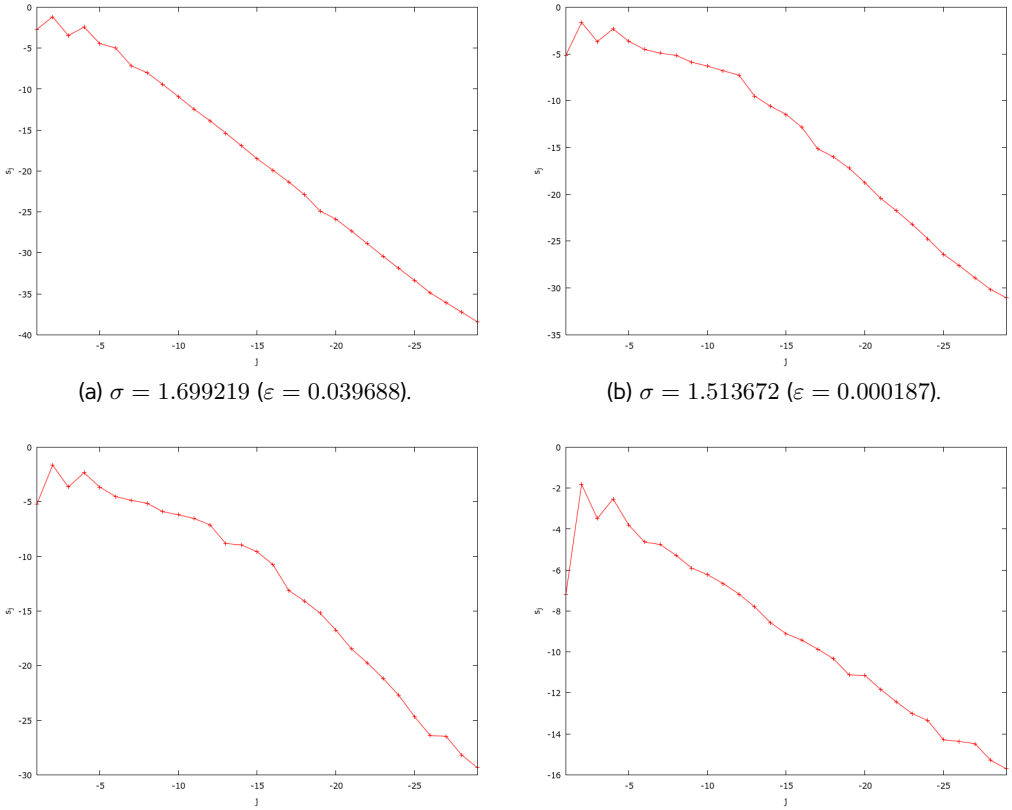
$$\varepsilon(\sigma) = \begin{cases} (\sigma - 1.5)^2 & \text{when } 1.5 \leq \sigma \leq 2, \\ 0 & \text{when } 1 \leq \sigma \leq 1.5. \end{cases}$$

With this parametrization the system is pinched if and only if $\sigma \in [1, 1.5]$. In Figure 4.3 we plot the estimated regularities of System (3.6) as a function of σ with the above parametrization.

In Figure 4.3 one can clearly appreciate three regions with different qualitative behavior. One of them corresponds to the pinched case (i.e. $\sigma \in [1, 1.5]$) and the other two to the non-pinched one: $\sigma \in (1.5, \widetilde{\sigma})$ and $\sigma \in [\widetilde{\sigma}, 2]$ with $\widetilde{\sigma} \approx 1.527$. In what follows we discuss in detail these three regions. The problems can be guessed by inspection of Figure 4.4.

Non pinched case: $\sigma \in [\widetilde{\sigma}, 2]$

In this region we have $\varepsilon = (\sigma - 1.5)^2 \gtrapprox 7.29 \times 10^{-4}$ and hence we are far from the pinched case. The function φ whose graph is the attractor is continuous but not differentiable (see [Kel96, Sta99, Sta97]). Moreover, since we are far from the pinched case, φ is rather well



(a) $\sigma = 1.699219$ ($\varepsilon = 0.039688$).

(b) $\sigma = 1.513672$ ($\varepsilon = 0.000187$).

(c) $\sigma = 1.507812$ ($\varepsilon = 0.000061$).

(d) $\sigma = 1.425781$ ($\varepsilon = 0$).

Figure 4.4: Graphs of the pairs (j, s_j) with $-29 \leq j \leq 0$ for the (strange) attractor of System (3.6) different values of σ and ε given by the parametrization $\varepsilon(\sigma)$. In agreement with the computed Pearson correlation coefficient they are approximately linear. Also observe that the first values of j will be the worst fitted for the straight line of the regression in concordance with the comment below Lemma 2.3.2.

behaved since still we have lack of differentiability in few points (see Figure 4.5). This is confirmed by the estimated regularities that, not surprisingly (see Remark 3.3.5), are in the interval $(0, 1)$ and "far" from zero: $\mathrm{R}(\widetilde{s}) \in [0.6822, 0.9669]$. But these regularities must be closer to one by Theorem 3.3.1.

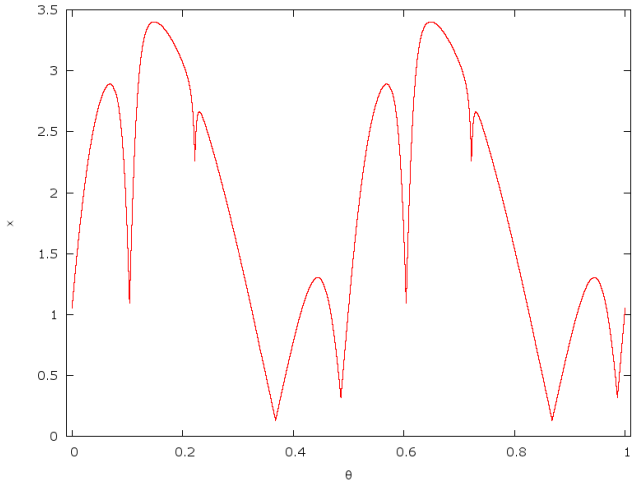


Figure 4.5: The attractor of System (3.6) for $\sigma = 1.699219$ (and $\varepsilon = 0.039688$). In this case $\mathrm{R}(\widetilde{s}) = 0.91431$.

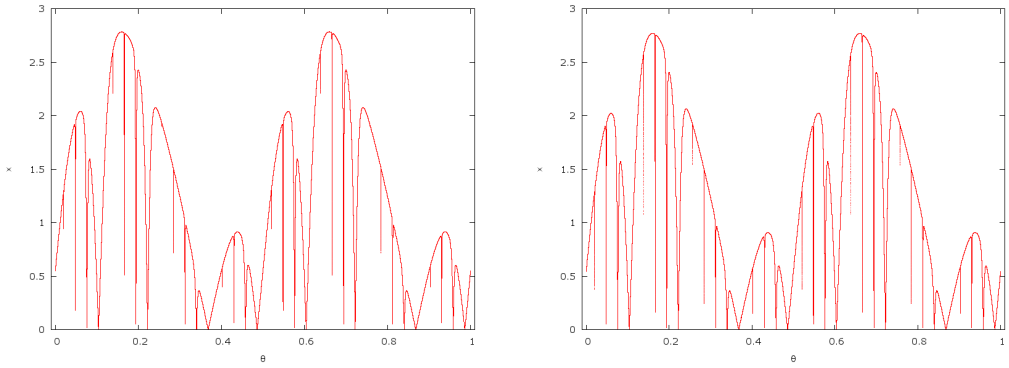


Figure 4.6: On the left picture it is shown the attractor of System (3.6) for $\sigma = 1.513672$ (and $\varepsilon = 0.000187$). The regularity of this attractor is $\mathrm{R}(\widetilde{s}) = 0.6266$. On the right picture the parameter are $\sigma = 1.507812$ ($\varepsilon = 0.000061$) and $\mathrm{R}(\widetilde{s}) = 0.4951$. Notice the *big* difference between the regularities of the two *parametrically close attractors*.

Approaching the pinched case: $\sigma \in (1.5, \widetilde{\sigma})$

In this region, since for $\sigma = 1.5$ we are already in the pinched case, the function becomes more irregular (see Figure (4.6)). Therefore, the regularity falls to zero abruptly and we have big differences in estimated regularity between parametrically close attractors. But moreover, these estimated regularities must be, again by Theorem 3.3.1, closer to one.

The pinched case: $\sigma \in [1, 1.5]$

In this case, according to 3.2.1 and Proposition 3.3.5, the attractor is pinched and, hence, discontinuous almost everywhere (see Figure 3.1). Then it is not surprising that the regularity is equal to zero for the whole range of parameters.

In view of the above three regions and these kind of results some questions arise: are the wavelet coefficients "*sufficiently*" good enough besides the inherent error induced from the initialization of Fast Wavelet Transform? The wavelet used, to make the above pictures, is the best for all range of parameters?

4.3.2 Testing the quality of coefficients

In contrast with the Fourier setting, one of the advantages of the wavelet context usage is that one has infinitely many orthonormal bases. Recall that, in order to be under hypothesis of Theorem 1.2.11, we must use Daubechies family with more than $\max(s, 5/2 - s)$ vanishing moments. However, since the number of vanishing moments, k, cause an increment of the support size of the wavelet the first wavelet coefficients are even worse because the support of φ remains unaltered (in fact the ratio of growth between φ and ψ is 1 to $2k-1$). With this comment we mean that a good strategy seems to choose the wavelet in terms of the parameters σ and specially ε and k (besides the assumptions on the vanishing moments mentioned above).

In order to make more clear the previous comments and ideas, we have performed the following exercise. Since the region *Approaching the pinched case* is the worst one (in therms of predicting the regularity) we have fixed $\sigma = 1.5$ and we have calculated the wavelet coefficients using the **Step Previous** and **Step 2** of the algorithm given in Section 4.1. After that, we have reconstructed the regularity with Method 4.1.2 in order to check if such wavelet coefficients are able to reconstruct the regularity of the function φ which is given by Proposition 3.3.5. The result of such experiment can be seen at Figure 4.7.

The first thing to notice, as we have pointed out at the beginning of this section, is that there are better regions, in terms of ε and k, where some wavelets reconstruct the regularity better than other. Also a we have a remarkable fact: the Pearson correlation coefficient is always big enough whether in the good and the bad regions. On the other

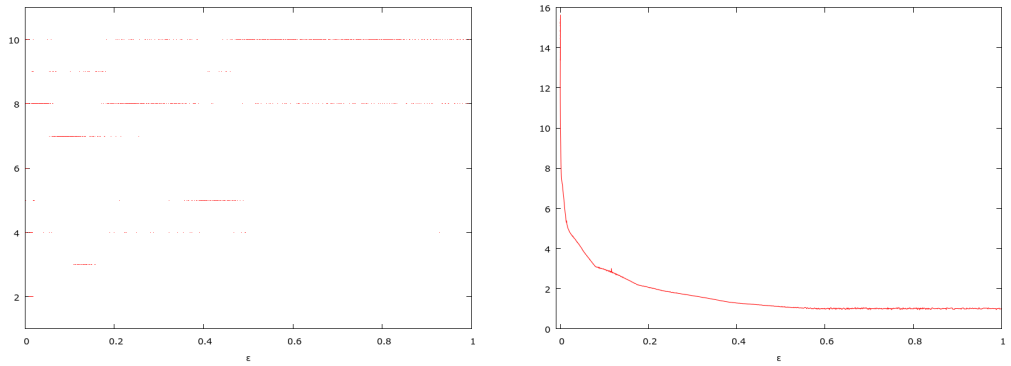


Figure 4.7: On the left picture it is shown, in terms of the Vanishing Moments which wavelets reach the maximum Pearson coefficient (always greater than 0.99). On the right hand side, it can be suspected the functional space jump in terms of the norm of φ. In both cases, $\varepsilon \in (0, 1]$ and $\sigma = 1.5$ for the attractor of System (3.6).

side, the wavelet coefficients, given by the Fast Wavelet Transform, show the functional space jump predicted at Proposition 3.3.5 (see Figure 4.7). Also they define a behavior on the regularity problem but they seem to be not able to assess *the* value.

Drawing to a close this section we want to highlight that, in view of the experiments that we have performed with the Weierstraß function or moving the parameter of the invariant function φ, the implementation of the Fast Wavelet Transform that we have used, the one from `GNU-GSL` library, it is not sufficiently fine. But, moreover, it seems that Fast Wavelet Transform has some drawbacks, as the initialization specially when one has to deal with highly irregular functions. This leads us to think that the "*practical*" Fast Wavelet Transform is only valid to define tendencies when the problem of the regularity is considered. Hence, other methods must be employed such as quadrature rules. However, this tendencies can be taken to be numerical evidence of the regularity space jumps especially in view of Proposition 3.3.5. That is, the Fast Wavelet Transform is about to lay the foundation stone of another way to see how φ becomes strange. Although other methods more accurate must be performed.

The naive reason of take the quadrature rules way is simple. Indeed, since we have used Fast Wavelet Transform, and hence a_{-J}^{PER}, we did not take into account the special shape of the Transfer Operator which is described by in the previous chapter. However, if we use a quadrature rule to calculate $d_j^{\mathrm{PER}}[n] := \left\langle {\varphi_{(\sigma,\varepsilon)}^k}^{\mathrm{PER}}, \psi_{j,n} \right\rangle$, for a sufficiently big $k > 0$, we can use the information given by Theorem 3.2.1 and the Dominated Convergence Theorem. Hence, increase the degree of accuracy of the coefficients.

Recall that in Section 2.3.3 we have defined

$$\int_{\operatorname{supp}(w(x))} f(x)w(x)\,dx = \sum_{i=1}^{n} \ell_i f(x_i) + \mathcal{R}(f)_w$$

to be a n-point Gaussian quadrature rule. The problem of such method, to be perfectly honest, is the viability and reliability of the computations performed as well as the problem of how to convert $\varphi \in \mathbb{S}^1$ to a map of $\mathbb{R}$. Let us be more precise. The quadrature given in Conjecture 2.3.19, due to its computational advantages (in terms of time spent and memory) as well as its simplicity, made us feel hopeful at a first glance. That said, one of the first problematic milestones is the degree of accuracy of the coefficients $h[n]$. Indeed, it is known that the roots of a polynomial are a continuous function of the coefficients. However, a not sufficiently good approximation of $h[n]$'s gives us a sequence of polynomials, given by Equation (2.21), which they loose its orthogonality as well as they become ill-conditioned polynomials. The theoretical solution of such problems can be done following the thread depicted in Remark 1.1.12. That is, use tools of computational algebraic geometry such as the Gröbner basis (see [Reg07, RS05] an the references there in to get a more comprehensive explanation) to get the best parameterization of the coefficients $h[n]$. Nevertheless, taking into account our poor knowledge of a computational algebraic software this theoretical tool becomes unworkable for us when the degree of the polynomial grows. Also, algebraic geometry is a topic very far away of our scope.

On the other side, when one tries to bound the remainder $\mathcal{R}(f)_w$ in order to have a good quadrature rule, for the pinched case, it turns out the problem of the zeros. In other words, the wavelet coefficients for $\varepsilon \neq 0$ can be computed in a reasonable time as well as can be considered good enough in terms of reconstruction of φ. Instead, consider a naive application of the Birkhoff Ergodic Theorem such as

$$d_j^{\mathrm{PER}}[n] := \int_{\Omega_{j,n}^{\mathrm{PER}}} \varphi(\theta)\psi_{j,n}^{\mathrm{PER}}(\theta)\,d\theta = \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} \overbrace{\varphi \cdot \psi_{j,n}^{\mathrm{PER}} \cdot \chi_{\Omega_{j,n}^{\mathrm{PER}}}}^{f \text{ of Birk. Erg. Thm.}}(R_\omega^i(\theta)),$$

where $\chi_{\Omega_{j,n}^{\mathrm{PER}}}$ is the characteristic function of the domain of $\psi_{j,n}^{\mathrm{PER}}$. Then, even though its poor speed of convergence it seems to be better even the quadrature rule aforementioned. The origin of this, at a first glance, surprising artifact can be explained as follows. Indeed, recall that the base space of the skew product is $\mathbb{S}^1$ whereas the quadrature rule is defined over $\mathbb{R}$. We tried to use the same trick as the one used before Lemma 1.2.7 (the definition of f^{PER}). Unfortunately, it seems to be not enough powerful for our purposes and problems. In view of that the extension of the Conjecture 2.3.19 on $\mathbb{S}^1$ will be, for us, an open problem to solve in a future work.

In view of these comments, another method must be applied such as the resolution of the (discretized) invariance equation (3.15). This will be the main topic of the next chapter.

Chapter 5

The Invariance Equation Through Wavelets

Now we have a simple tool to work with and compute the wavelet coefficients: the Fast Wavelet Transform. It allows us, for example, determine some tendencies of φ in terms of the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$. However, being not satisfied with this approach and in view of we have pointed out at the ending of the previous chapter, we will go further.

This current chapter must be seen as a melting-pot of all the techniques, subjects and concepts that we have explained and presented up to now. Actually, we aim to get an expression like

$$\varphi \sim a_0 + \sum_{j=0}^{J} \sum_{n=0}^{2^j-1} d_{-j}^{\mathrm{PER}}[n] \psi_{-j,n}^{\mathrm{PER}}.$$

To get it, the central role will be played by the discretization of the Transfer Operator, $\mathbf{T}$ which is given by Equation (3.16), in order to solve the invariance equation, given by (3.15), using Newton's method. Following this thread, it will leads us to use the method to compute and store either Haar or Daubechies wavelet matrices given in Section 2.2 of Chapter 2. Moreover, using Haar wavelet we give the "*exact*" solution of the linear system derived using a simple recursion. Such solution can be obtained in $\mathcal{O}(N)$ time. Recall that, roughly speaking, this means that for big input size the computer running time increases linearly respect the size of the data entry.

Moving to the final part of the chapter, we will perform a battery of experiments with the Skew Product given by Equation (3.6). Such performance should enable us two things. The first one is the check and comparison between the results of Chapter 4 and the new ones. The second item is to take advantage of the short running time (and memory!) of the calculation when Haar wavelet is considered. Combining them we will derive more properties of φ besides of the regularity pointed out in Chapter 4.

5.1 The (non)-linear system

We aim to obtain the coefficients $d_{-j}^{\mathrm{PER}}[n]$, which recall that they are given by

$$d_{-j}^{\mathrm{PER}}[n] := \langle \varphi, \psi_{-j,n} \rangle = \int_{\mathrm{supp}(\psi_{-j,n}^{\mathrm{PER}})} \varphi(\theta)\psi_{-j,n}^{\mathrm{PER}}(\theta)\, d\theta.$$

To this end, recall that we are considering skew products of the following form $(\theta_{n+1}, x_{n+1}) = (R_\omega(\theta_n), F_{\sigma,\varepsilon}(\theta_n, x_n))$. On the other side, when $F_{\sigma,\varepsilon}(\theta, x)$ satisfies Theorem 3.2.1, we know that there exists an upper semi continuous function $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{R}$ such that $\varphi(R_\omega(\theta)) = F_{\sigma,\varepsilon}(\theta, \varphi(\theta))$ whose graph is invariant. That is, we can consider to forge the non linear system of equations which will give us the desired coefficients $d_{-j}^{\mathrm{PER}}[n]$ for a large class of skew products. Indeed, recall the definition of the operator $\mathbf{T}$ given by Equation (3.16):

$$\begin{array}{rcl} \mathbf{T} : C^n(\mathbb{S}^1 \times \mathbb{R}, \mathbb{R}) & \longrightarrow & C^n(\mathbb{S}^1 \times \mathbb{R}, \mathbb{R}) \\ \varphi(\theta) & \longmapsto & \varphi(R_\omega(\theta)) - F_{\sigma,\varepsilon}(\theta, \varphi(\theta)). \end{array}$$

Let us explain how we will get the desired coefficients $d_{-j}^{\mathrm{PER}}[n]$ aforesaid. To this end, following mainly [Jor01], we will consider the discretization of the operator $\mathbf{T}$. Indeed, we want to express the function φ as a finite expansion like

$$\varphi \sim a_0 + \sum_{l \in \Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}},$$

where ψ_l^{PER} is an orthonormal wavelet basis of $\mathbb{S}^1$. Recall that $l = 2^j + n$ (therefore one must take $j = -\lfloor \log_2(l) \rfloor$ and $n = l - 2^j$) and Λ is the set of l's. Now, being J large enough and the points $\theta_i = i/2^J \in \mathbb{S}^1$, one can perform *the* equation with $\varphi(\theta_i) = a_0 + \sum_{l \in \Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(\theta_i)$ doing the following:

$$F_{\sigma,\varepsilon}\left(\theta_i, a_0 + \sum_{l \in \Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(\theta_i)\right) = a_0 + \sum_{l \in \Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(R_\omega(\theta_i)) \tag{5.1}$$

which is, clearly, the discretization of $\mathbf{T}$. In view of that, we can translate such problem to a "*root finding*" problem by considering each component of the invariance equation

$$\begin{array}{rcl} I_i : \mathbb{R}^{2^J} \times \{\theta_i\} & \longrightarrow & \mathbb{R} \\ (\mathrm{D}_N^{\mathrm{PER}}, \theta_i) & \longmapsto & I_i(d_l^{\mathrm{PER}})\big|_{\theta_i} \end{array}$$

where we take, by definition,

$$I_i(\mathrm{D}_N^{\mathrm{PER}})\big|_{\theta_i} := a_0 + \sum_{l \in \Lambda} d_l^{\mathrm{PER}} \psi_l(R_\omega(\theta_i)) - F_{\sigma,\varepsilon}\left(\theta_i, a_0 + \sum_{l \in \Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(\theta_i)\right)$$

and hence we can redefine the problem derived from Equation (5.1), in terms of the I_i's, by find a vector $\mathrm{D}_l^{\mathrm{PER}} \in \mathbb{R}^{2^J}$ such that the following function

$$\begin{array}{rcl} \mathbf{F} : \mathbb{R}^{2^J} & \longrightarrow & \mathbb{R}^{2^J} \\ \mathrm{D}_N^{\mathrm{PER}} & \longmapsto & (I_0(\mathrm{D}_N^{\mathrm{PER}})\big|_{\theta_0}, \ldots, I_{2^{J-1}}(\mathrm{D}_N^{\mathrm{PER}})\big|_{\theta_{2^{J-1}}}) \end{array}$$

has a zero at $\mathrm{D}_N^{\mathrm{PER}}$. To do this, as we have pointed out in Section 2.3.2, we will use Newton's method. That is, we will find $\mathrm{D}_{N,\star)}^{\mathrm{PER}}$ such that $\left|\mathrm{D}_{N,\star)}^{\mathrm{PER}} - \mathrm{D}_N^{\mathrm{PER}}\right| < \varepsilon$ using the recurrence

$$\mathbf{JF}(\mathrm{D}_{N,n)}^{\mathrm{PER}})(\mathrm{X}) = -\mathbf{F}(\mathrm{D}_{N,n)}^{\mathrm{PER}}) \tag{5.2}$$

for the unknown $\mathrm{X} = \mathrm{D}_{N,n+1)}^{\mathrm{PER}} - \mathrm{D}_{N,n)}^{\mathrm{PER}}$ and given a seed $\mathrm{D}_{N,0)}^{\mathrm{PER}}$. Fixed a $J > 0$ it can be seen that the (i,l)-th entry of the Jacobian matrix of $\mathbf{JF}$, in our particular setting, is given by

$$\mathbf{JF}_{i,l} = \begin{cases} 1 - \dfrac{\partial F\left(\theta_i, a_0 + \sum\limits_{l\in\Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(\theta_i)\right)}{\partial x} & \text{if } l = 0, \\ \psi_l^{\mathrm{PER}}(R_\omega(\theta_i)) - \dfrac{\partial F_{\sigma,\varepsilon}\left(\theta_i, a_0 + \sum\limits_{l\in\Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(\theta_i))\right)}{\partial x} \psi_l^{\mathrm{PER}}(\theta_i) & \text{otherwise.} \end{cases} \tag{5.3}$$

We want to emphasize that the above equation is, precisely, the discretization of the transfer operator given by Equation (3.17) aforementioned in Chapter 3. In other words, from the point of view of iterative solvers, the convergence towards the solution is closely related with the reducibility properties of the skew product considered. On the other side, we claim that if $\psi(x)$ has compact support then the Jacobian matrix has most of its entries populated by zeros. That is, in our particular setting $\mathbf{JF}$ will be sparse matrix and is "*highly structured*" as it can be seen in Section 2.2.

We end by recalling that ψ_l^{PER} is the l-th element of an orthonormal basis of $\mathbb{S}^1$. Hence, being $N = 2^J$ fixed and taking into account Equation (5.3), the Jacobian matrix can be compactly written as

$$\mathbf{JF} = \left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right),$$

where Ψ_N^{PER} is set to be the $N \times N$ matrix whose columns are $\psi_l^{\mathrm{PER}}(\theta_i)$ and, similarly, let $\tilde{\Psi}_N^{\mathrm{PER}}$ be the $N \times N$ matrix whose columns are $\psi_l^{\mathrm{PER}}(R_\omega(\theta_i))$. Finally, the diagonal matrix given by

$$\frac{\partial F_{\sigma,\varepsilon}\left(\theta_i, a_0 + \sum\limits_{l\in\Lambda} d_l^{\mathrm{PER}} \psi_l^{\mathrm{PER}}(\theta_i)\right)}{\partial x}$$

is stored in Δ_N. To conclude this section devoted to state and perform the (non)-linear system of equations we make the following remark.

Remark 5.1.1. The tools to calculate and store $\mathbf{JF}$ are a direct consequence of those techniques explained in Section 2.2 of Chapter 2. However, it must be pointed out that for the mesh $\tilde{\theta}_i = i/N + \omega \in \mathbb{S}^1$, where $\omega \in \mathbb{R}\backslash\mathbb{Q}$, the above formulation of how one can store $\tilde{\Psi}_N^{\mathrm{PER}}$ remains valid. This is true because, as it can be seen in the proof of Lemma 2.2.3, the block structure of the matrix is induced by $l = 2^j + n$. Notice also that for $\tilde{\theta}_i$, the Daubechies Lagarias products are more complicated to calculate and, thus, it slows down the storage process. ∎

Now, having tailored $\mathbf{JF}$ for our particular needs, we are ready to explain carefully how the Newton's method is applied in the forthcoming section.

5.2 Newton's route

Recall that we are looking for the wavelet coefficients using the Newton's method. In this section we will explain carefully how such technique is applied for our particular setting. That is, fixed an integer $J > 0$ big enough (and hence $N = 2^J$) we need to find, up to a given tolerance, $\mathrm{D}_{N,\star)}^{\mathrm{PER}}$ such that it solves the Equation (5.1). To find it, we will use the recurrence given by Equation (5.2). The vector $\mathrm{D}_{N,\star)}^{\mathrm{PER}} \in \mathbb{R}^N$ will be the vector of wavelet coefficients $d_{-j,\star)}^{\mathrm{PER}}[n]$ related to the Daubechies wavelet ψ_l^{PER} with $p \geq 1$ vanishing moments. Recall that $p = 1$ corresponds to the case of Haar wavelet.

5.2.1 The initial guess

As we have already said, Section 2.3.2, we need an initial seed $\mathrm{D}_{N,0)}^{\mathrm{PER}}$ and, in order to get such first N-dimensional vector, we will apply a practical issue of the Birkhoff Ergodic Theorem. Indeed, let us assume that our dynamical system verifies that $\varphi(\cdot)\psi^{\mathrm{PER}}(\cdot) \in \mathscr{L}^1(\mathbb{S}^1)$. Therefore by the Birkhoff Ergodic Theorem, it follows that

$$d_{-j,0)}^{\mathrm{PER}}[n] := \int_{\Omega_{-j,n}^{\mathrm{PER}}} \varphi(\theta)\psi_{-j,n}^{\mathrm{PER}}(\theta)\, d\theta = \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} \overbrace{\varphi \cdot \psi_{-j,n}^{\mathrm{PER}} \cdot \chi_{\Omega_{-j,n}^{\mathrm{PER}}}}^{f \text{ of Birk. Erg. Thm.}}(R_\omega^i(\theta)), \tag{5.4}$$

where $\chi_{\Omega_{-j,n}^{\mathrm{PER}}}$ is the characteristic function of the domain of $\psi_{-j,n}^{\mathrm{PER}}$. Of course, the aforementioned assumption when one uses the system given by (3.1) makes sense. However, doing numerics the limit is not applicable directly but using the periodicity and taking $N = k$ one may think that Equation (5.4) is the trapezoidal rule.

On the other side, recall that what we have a sample of points $\theta_i = i/N$. Then, using a similar trick as the one depicted in Equation (3.8) we can construct a sample of φ over the points $\theta_i = i/N$. Indeed, define for $i = 0, \ldots, 2^J - 1$ the points $\tilde{\theta}_i = R_\omega^{-k}(\theta_i)$ where $k > k_0$. As we have done before the integer k_0 is defined by Equation (3.4) (up to a given tolerance). That is, we have gone backwards under the rotation in order to

have the preimage of each θ_i. Now, it only remains to go forwards. Indeed, consider for $i = 0, \dots, 2^J - 1$ the points $\mathfrak{F}^k(\tilde{\theta}_i, x_0)$ being, in the same way as in Section 4.2, $x_0 > (2\sigma(\varepsilon + 1))$. Then by, Theorem 3.2.1, the points constructed in such way are an exponentially close sample of $\varphi(\theta_i)$ for $i = 0, \dots, 2^J - 1$. Hence, we are allowed to define Φ_N to be a column vector whose N-components are $\varphi(\theta_i)$ with $\theta_i = i/N$. Being ${\Psi_N^{\mathrm{PER}}}^\top$ the transpose of Ψ_N^{PER} then we have that

$$ {\Psi_N^{\mathrm{PER}}}^\top \Phi_N = \mathrm{D}_{N,0)}^{\mathrm{PER}}, \tag{5.5} $$

where the vector $\mathrm{D}_{N,0)}^{\mathrm{PER}}$ has $d_{-j,0)}^{\mathrm{PER}}[n]$ as entries. Indeed, since the columns of Ψ_N^{PER} are $\psi_l^{\mathrm{PER}}(\theta_i)$ then its transpose has $\psi_l^{\mathrm{PER}}(\theta_i)$ as rows. Therefore, each entry of $\mathrm{D}_{N,0)}^{\mathrm{PER}}$ of Equation (5.5) has a multiplication like

$$ \sum_{i=0}^{N-1} \psi_l^{\mathrm{PER}}(\theta_i)\varphi(\theta_i). $$

Which is, exactly, the Equation (5.4) if one divides the result by N. That is, our initial guess for the Newton's method will be the vector $\mathrm{D}_{N,0)}^{\mathrm{PER}}$ given by Equation (5.5) divided by N.

Remark 5.2.1. One may think that another fast way to obtain the initial seed is applying the FWT. But due to the final comments in Chapter 4 we have discarded it. However, in the case of Haar the matrix ${\Psi^{\mathrm{PER}}}_N^\top$ coincides with the one of the Discrete Haar Transform. In other words, the resultant of ${\Psi_N^{\mathrm{PER}}}^\top \Phi_N$ is equivalent to apply the Discrete Haar Transform to Φ_N. ∎

To conclude this explanation of how we choose the initial seed, let us make a comment concerning why we use the Trapezoidal rule to pick up our initial guess. The reason is quite simple: the Trapezoidal rule tends to become extremely accurate when periodic functions are integrated over their periods. Indeed, an asymptotic error estimate is given by

$$ -\frac{(b-a)^2}{12N^2}[f'(b) - f'(a)] + \mathcal{O}(N^{-3}). $$

That is, under mild conditions $f'(a) = f'(b)$, and only the $\mathcal{O}(N^{-3})$ term remains when periodic functions are considered. Of course, perhaps in our case $\varphi'(0)$ does not exists in the usual sense. But the above error formula is still decreasing when N goes to ∞.

5.2.2 Solving the system

In this following lines we will explain the strategies that we have performed to solve the linear system derived in Equation (5.2). Our goal will be find fast ways to compute solutions of

$$ \mathbf{JF}(\mathrm{X}) = \left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right)(\mathrm{X}) = -\mathrm{Inv}, \tag{5.6} $$

where $\mathbf{JF}$ is given by Equation (5.3). Also, the vector $\mathrm{Inv} \in \mathbb{R}^N$ is the invariant equation evaluated at θ_i and it is computed by means of Equation (5.2). To solve the above linear system of equations, we will refer the reader to some of the concepts introduced Section 2.3.2.

Remark 5.2.2. In order to make the formulas more readable we have omitted the parameters σ and ε in Equation (5.6). In other words, we need to solve

$$\mathbf{JF}_{\sigma,\varepsilon}(\mathrm{X}) = \left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right)(\mathrm{X}) = -\mathrm{Inv}_{\sigma,\varepsilon},$$

where Δ_N also depends on σ and ε. ∎

Let us start by saying that it is difficult to find references that try to analyze the convergence of the GMRES method without making assumptions on the condition number of the Arnoldi matrix. Although, in [Mor97] the author studies linear systems of the form

$$(\lambda \mathrm{Id} + \mathbf{T})u = f$$

for a certain compact operator $\mathbf{T}$ in a Hilbert space and $\lambda \neq 0$. Such study allows to get bounds for the super linear convergence in terms of singular values of the operator $\mathbf{T}$. Furthermore, there are several instances where the linear system must be preconditioned using either right or left preconditioning technique. Such procedure looks for a twofold objective. On the one hand, a good precondition technique must be little costly from the computational point of view. On the other side, the resultant matrix, $\widetilde{M} = \lambda \mathrm{Id} + \Xi$, must have "*better properties*". That is the matrix Ξ must have good properties such as small norm or its eigenvalues are well placed.

Without wishing to be ambitious, we will try to combine those ideas. That is, using a right preconditioning technique we will try to transform the linear system given by Equation (5.6) to a more simpler one. In order to have an intuition of the convergence of the GMRES method, the resultant system will be as much similar as we can as the one presented by [Mor97]. To this end, let us say that the spectral radius of $\mathbf{JF}$ is bounded when an $\mathbb{R}$-Daubechies wavelet with $p \geq 1$ vanishing moments that generates an orthonormal basis of $\mathbb{S}^1$ is considered.

We want to remark that the proof is something silly. But it is a heuristic tool to use when one tries to make the pass to the limit in order to bound the spectral radius of $\mathbf{JF}$. Also, for the same reason, the verbatim application of the Gershgorin circle theorem does not add many information of how and where the eigenvalues of $\mathbf{JF}$ are placed. Despite of this, we already know that the spectral radius is bounded. Indeed, as we have pointed out in Section 3.4, since $\mathbf{JF}$ is a discretization of the Mather spectrum then its spectral radius must be in an annuli. This is a very bad notice for us because GMRES works well in the case of clustered eigenvectors. On the other side, iterative methods, as Jacobi or Gauss-Seidel, are discarded because the diagonal can be zero many times

due the sparsity of $\mathbf{JF}$. In view of that, we need to transform $\mathbf{JF}$ to a better matrix in order to apply safely iterative solvers. This is the main topic of the following two sections.

5.2.2.1 Haar's strategy

Let us zoom in a little bit more the structure of how we can convert the matrix $\mathbf{JF}$ to a better one. Take a Daubechies wavelet with $p \geq 1$ vanishing moments. We already know that the columns of Ψ_N^{PER} are the first N elements of the basis $\{\psi_l^{\mathrm{PER}}\}_{l\in\mathbb{Z}^+}$ evaluated at $\theta_i = i/N$. For a moment, let us assume that using a similar argument as the one in Section 5.2.1 the following chain of equalities, dividing by N, is true

$$\mathbf{JF}\Psi_N^{\mathrm{PER}\top} = (\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N\Psi_N^{\mathrm{PER}})\tilde{\Psi}_N^{\mathrm{PER}\top} = (\mathsf{Id} - \Delta_N\Xi).$$

If the matrix $\Delta_N\Xi$ is well conditioned, or its spectral radius is less than one or $\Xi = \mathsf{Id}$ then we can consider ourselves lucky. We will come back to this last item later on.

Continuing the thread of $(\mathsf{Id}-\Delta_N\Xi)$ is important to recall that if $\mathsf{Id}-\Delta_N\Xi$ is invertible then

$$(\mathsf{Id} - \Delta_N\Xi)^{-1} = \sum_{k=0}^{\infty}(\Delta_N\Xi)^k,$$

where he series above converges whenever $\varsigma(\Delta_N\Xi) < 1$. Hence, the matrix $\Delta_N\Xi$ can be used to derive an iterative method given by

$$\mathrm{X}^{k+1)} = (\mathrm{Inv} - \Delta_N\Xi\mathrm{X}^{k)}),$$

with effective error bounds, to find the solution of $(\mathsf{Id} - \Delta_N\Xi)\mathrm{X} = -\mathrm{Inv}$.

We recover here the idea of the right preconditioning technique but now, using the Haar basis ($p = 1$). Recall, as we have said in Section 1.1.3, in concrete in Corollary 1.1.17, it is verified that $\psi_{j,n}^{\mathrm{PER}}(\theta) = \psi_{j,n}(x)$. Also with the matrix Ψ^{PER} we can forget the superscript $^{\mathrm{PER}}$. That is, when we are considering the Haar matrix we will denote it Ψ and $\tilde{\Psi}$ for its rotated version.

Lemma 5.2.3. *The matrix $\frac{1}{\sqrt{N}}\Psi_N$ is orthogonal. That is $\frac{1}{N}\Psi_N^\top\Psi_N = \frac{1}{N}\Psi_N\Psi_N^\top = \mathsf{Id}$.*

Proof. In order to prove the statement, we must keep in mind that the Haar basis is an orthonormal basis of $\mathbb{S}^1$. Therefore, the matrix $\frac{1}{N}\Psi_N^\top\Psi_N$ has, in soul,

$$\int_{\mathbb{S}^1}\psi_l(\theta)\psi_m(\theta)\,d\theta = \delta_{l,m}$$

at each of its entries. Let us be more precise. We need to show that $\frac{1}{N}\Psi^\top\Psi = \mathsf{Id}$. To this end recall that, by definition of Ψ each its j-th column, with $j > 0$, is given by Equation (2.6) and 1 if $j = 0$. In view of that, we will split the proof in to two cases.

First row of $\Psi_N^\top$. From the definition of matrix multiplication and the above comments it is clear that

$$\Psi_N^\top \Psi_N = \sum_{k=0}^{N-1} \Psi_{i,k}^\top \Psi_{k,j} = \begin{cases} N & \text{if } i = j = 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $\Psi_{i,j}$ stands for a particular entry of the matrix Ψ_N.

The other rows of $\Psi_N^\top$. As we have already said in Section 2.1.1 $\psi_{-j,n}(x)$ is supported over the Dyadic Intervals. This implies that the range where $\Psi_{i,j}$ and $\Psi_{i,j'}$ are different from zero are mutually disjoint if $j \neq j'$. Therefore and, again, by Equation (2.6) we have that

$$\Psi_N^\top \Psi_N = \sum_{k=0}^{N-1} \Psi_{i,k}^\top \Psi_{k,j} = \begin{cases} N & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

To conclude the proof, recall that if two matrices A, B are such that $AB = \mathrm{Id}$ then BA also (see e.g [HJ85]). Therefore, the proof follows dividing $\Psi_N^\top \Psi_N$ by N. □

The above Lemma will be one of the crucial ones of this section. Indeed, since Ψ is orthogonal, we claim that

$$\Psi \cdot \tilde{\Psi}^\top = \Psi \cdot (P \cdot \Psi)^\top = \Psi \cdot \Psi^\top \cdot P^\top = \mathrm{Id} \cdot P^\top = \widetilde{\mathrm{Id}} \tag{5.7}$$

for a certain permutation matrix P. This, at the first glance, naive approach will become very useful for us because our goal is to find fast solutions of Equation (5.6). This will be the main topic of the following lines. But, before we continue the explanation, we need to show the the aforementioned claim.

Lemma 5.2.4. *Let $\{\psi_l\}_{l \in \mathbb{Z}_+}$ the Haar basis in $\mathbb{S}^1$ and fix $\omega \in \mathbb{R} \setminus \mathbb{Q}$. Then, fixed a J and setting $N = 2^J$, for all $i = 0, \ldots, N-1$*

$$\psi_l(i/N) = \psi_{-j,n}(i/N) = \psi_{-j,n}\left(\frac{i - \mathfrak{p} \mod N}{N} + \omega\right),$$

with $\mathfrak{p} = \lfloor \omega N \rfloor$.

Proof. In order to prove the Lemma, we claim that

$$\psi_{-j,n}\left(\frac{i - \mathfrak{p} \mod N}{N} + \omega\right) = \psi_{-j,n}\left(\frac{i}{N} + \alpha\right).$$

Indeed, recall that $i \in \{0, \ldots, N-1\}$ and p is one of such i's. Hence we have two cases.

$i \geq p$ First of all, notice that $i - r \mod N = i - r$. Therefore

$$\begin{aligned}\psi_{-j,n}\left(\frac{i-\mathfrak{p} \mod N}{N}+\omega\right) &= \psi_{-j,n}\left(\frac{i-\mathfrak{p}}{N}+\omega\right)\\ &= \psi_{-j,n}\left(\frac{i}{N}-\frac{\mathfrak{p}}{N}+\frac{\mathfrak{p}}{N}+\alpha\right)\\ &= \psi_{-j,n}\left(\frac{i}{N}+\alpha\right).\end{aligned}$$

$i < \mathfrak{p}$ Now, $i - r \mod N = i - r + N$. Therefore,

$$\begin{aligned}\psi_{-j,n}\left(\frac{i-\mathfrak{p} \mod N}{N}+\omega\right) &= \psi_{-j,n}\left(\frac{i-\mathfrak{p}+N}{N}+\omega\right)\\ &= \psi_{-j,n}\left(\frac{i}{N}+1-\frac{\mathfrak{p}}{N}+\frac{\mathfrak{p}}{N}+\alpha\right)\\ &= \psi_{-j,n}\left(\frac{i}{N}+1+\alpha\right)\\ &= \psi_{-j,n}\left(\frac{i}{N}+\alpha\right),\end{aligned}$$

where the last equality follows from the fact that we are in $\mathbb{S}^1$.

In view of that, we need to show that

$$\psi_{-j,n}\left(\frac{i}{N}+\alpha\right) = \psi_{-j,n}\left(\frac{i/N}{2^j}+\frac{\alpha}{2^j}-n\right) = \psi_{-j,n}\left(\frac{i/N}{2^j}-n\right).$$

Since the Haar function can be reformulated as

$$\psi(x) = \begin{cases} 1 - 2\lfloor 2x \rfloor & \text{if } x \in [0,1) \\ 0 & \text{otherwise} \end{cases}$$

it suffices to prove that

$$\left\lfloor 2\left(\frac{i/N}{2^j}+\frac{\alpha}{2^j}-n\right)\right\rfloor = \left\lfloor 2\left(\frac{i/N}{2^j}-n\right)\right\rfloor$$

which, after some manipulations, is equivalent to show that the following equality $\lfloor i/(2^{j-1}N)\rfloor = \lfloor i/(2^{j-1}N) + \alpha/(2^{j-1})\rfloor$ is true. To do so, let us define

$$k := \lfloor i/(2^{j-1}N)\rfloor \text{ and } l := 2^{j-1}N.$$

Those numbers verify that $k \leq i/(2^{j-1}N) < k+1$ and $l = 2^{j-J} > 1$ respectively. Also, notice that the last property implies that $2^{j-1} = l/N \geq 1$. That is,

$$\begin{aligned}\lfloor i/(2^{j-1}N)\rfloor = k &\Leftrightarrow k \leq i/(2^{j-1}N) < k+1\\ &\Leftrightarrow 2^{j-1}k \leq i/N < 2^{j-1}(k+1)\\ &\Leftrightarrow \frac{kl}{N} \leq \frac{i}{N} < \frac{(k+1)l}{N}\\ &\Leftrightarrow \frac{kl}{N} \leq \frac{i}{N} < \frac{(k+1)l-1}{N}.\end{aligned}$$

Now, we are almost done since $0 < \alpha < 1/N$ and by definition of l it follows that

$$2^{j-1}k = \frac{kl}{N} < \frac{i}{N} + \alpha < \frac{(k+1)l-1}{N} + \frac{1}{N} = 2^{j-1}(k+1) \Leftrightarrow$$
$$k < \frac{i}{2^{j-1}N} + \frac{\alpha}{2^{j-1}} < k+1 \Leftrightarrow k = \lfloor i/(2^{j-1}N) + \alpha/(2^{j-1}) \rfloor$$

which completes the proof. □

Now the claim becomes clear. Indeed, we only have to calculate the integer $\mathfrak{p}$ and consider the permutation matrix P such that moves down the rows $\mathfrak{p}$ places. Such matrix P, as it is derived from the above proof, is given by

$$P_{i,j} = \begin{cases} 1 & \text{if } (j+\mathfrak{p}) \mod N, \\ 0 & \text{otherwise.} \end{cases} \tag{5.8}$$

With this trick we have $\tilde{\Psi} = P\Psi$ as we claimed and hence $\Xi = \widetilde{\mathrm{Id}}$. Now we can focus on how we can solve the linear system given by Equation (5.6). To this end, recall that a right precondition strategy which is, instead of solving the original linear system $Ax = b$ solve firstly $A\mathbf{P}^{-1}y = b$ and $\mathbf{P}x = y$ to get the desired solution x. Due to Lemma 5.2.4 we are almost done. Indeed, we will take $\mathbf{P} = \tilde{\Psi}_N$. That is, as it is done at the beginning of the current section, we have to solve

$$(\tilde{\Psi}_N - \Delta_N \Psi_N)\mathrm{X} = -\mathrm{Inv} \sim (\mathrm{Id} - \widetilde{\Delta}_N)y = -\mathrm{Inv}, \tag{5.9}$$

where $\widetilde{\Delta}_N$ is defined in the same way of $\widetilde{\mathrm{Id}}$. That is, $\widetilde{\Delta}_N = \Delta_N P^\top$, where

$$P^\top_{i,j} = \begin{cases} 1 & \text{if } j = (i+N-\mathfrak{p}) \mod N, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the resultant matrix $\widetilde{\Delta}_N$ is given by

$$\widetilde{\Delta}_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ \Delta_{i,j} & \text{if } j = (i+N-\mathfrak{p}) \mod N. \end{cases} \tag{5.10}$$

First of all, we want to remark that now the $\sim$ is well done. That is, the matrix Id is the identity matrix. Not, as we have pointed out above, an approximation of it as it occurs when other wavelet it is considered. The same comment is also valid for the matrix $\widetilde{\Delta}_N$. Moving to the appearance of the linear system, notice that it is a extremely simple linear system of equations. Indeed, it is formed by ones in the diagonal and

two external diagonals (which are complementary). Moreover, recall that the $k+1)$-th iterate of Jacobi method has the expression

$$x_i^{k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j<i} a_{ij}x^{k)} - \sum_{j>i} a_{ij}x^{k)}\right).$$

For our particular instance, the Jacobi method has the following explicit formulation:

$$x_i^{k+1)} = \left(\mathrm{Inv}_i - \widetilde{\Delta}_i x_j^{k)}\right).$$

But we can go further. Indeed, the following Proposition gives us the exact solution, in $\mathcal{O}(N)$ time. To this end, let $J > 0$ be an integer fixed and define $\mathfrak{N} = \{0, 1, \ldots, N-1\}$, where $N = 2^J$ as usual. Define the map ν_ι as

$$\begin{array}{rccl} \nu_\iota \colon \mathfrak{N} & \longrightarrow & \mathfrak{N} & \\ \iota_i & \longmapsto & \iota_i + a_{\mathfrak{p}} & \mod (N), \end{array} \tag{5.11}$$

where $\iota, a_{\mathfrak{p}} \in \mathfrak{N}$. The properties of the points of ν_ι^N are described in the following Lemma.

Lemma 5.2.5. *Let $\{\nu_\iota\}_0^N$ be the orbit of $\iota \in \mathfrak{N}$ given by (5.11). Then*

(a) $\nu_\iota^N = \iota$ for all $\iota \in \mathfrak{N}$.

(b) For all $\iota \in \mathfrak{N}$, $\{\nu_\iota\}_0^N$ is exhaustive if and only if $a_{\mathfrak{p}}$ is odd.

(c) Consider $a_{\mathfrak{p}} \in \mathfrak{N}$ and let k be the minimal integer such that $a_{\mathfrak{p}} \mod 2^{k+1} = 2^k$, where $0 \leq k \leq J$. If $\{\nu_\iota\}_0^N$ is not exhaustive then there are 2^k disjoint parts of $\{\nu_\iota\}_0^N$.

Proof. We must prove the three properties. But, before we start with the proof, we must remark that $\mathfrak{N} \cong (\mathbb{Z}/N\mathbb{Z}, +) = (\mathbb{Z}/2^J\mathbb{Z}, +)$ as a groups.

- We need to show that for all $\iota \in \mathfrak{N}$ the equality $\nu_\iota^N = \iota$ holds. To do this, take $\iota \in \mathfrak{N}$ and just observe that $\nu_\iota^N = \iota + Na_{\mathfrak{p}} \mod (N) = \iota$.
- First of all, notice that if the exhaustive property is verified then $\{\nu_\iota\}_0^N$ is a bijection since $\mathfrak{N}$ is a discrete set. On the other side, the divisors of N are the powers of two up to J. Having said that, let us start with the sufficient condition's proof. So, let us assume that $\{\nu_\iota\}_0^N$ is exhaustive: for all $\iota \in \mathfrak{N}$ there exists an integer $0 \leq k < N-1$ such that $\iota + ka_{\mathfrak{p}} = \iota_k$ for a certain $\iota_k \in \mathfrak{N}$ and $\iota_k \neq \iota_{k'}$ if $k \neq k'$. This implies that, viewing $a_{\mathfrak{p}}$ as a generator of a subgroup of $\mathbb{Z}/N\mathbb{Z}$, has order N. Therefore, $a_{\mathfrak{p}}$ must be odd since, in general, the order of an element m of a cyclic group is $n/\gcd(n, m)$.

Conversely, take $\iota \in \mathfrak{N}$ and assume that $a_{\mathfrak{p}}$ is odd. Then, using the same argument as above $a_{\mathfrak{p}}$ as a generator of a subgroup of $\mathbb{Z}/N\mathbb{Z}$ has order N. Therefore, by definition of order, $\iota + ka_{\mathfrak{p}}$ has N different elements. That is, $\{\nu_\iota\}_0^N$ is exhaustive.

- We know, from the above item, that if $\{\nu_\iota\}_0^N$ is not exhaustive then $a_{\mathfrak{p}}$ must be even. We claim that $a_{\mathfrak{p}}$ can be considered as a power of two itself. Hence, we assume that $a_{\mathfrak{p}} = 2^k$. Also, viewing $a_{\mathfrak{p}}$ as a generator of a subgroup of $\mathbb{Z}/N\mathbb{Z}$, it has some order between the powers of two (up to J). Let 2^{J-k} be this even order to conclude define $a_{\mathfrak{p}} = 2^k$ different sets $\Sigma_\iota = \{\iota + a_{\mathfrak{p}} \mod (N)\} \subset \mathfrak{N}$, with $\iota = 0, \ldots, 2^k - 1$.

  To finish the proof we need to prove the aforementioned claim. To this end, let us define Θ to be the set of the even numbers between 0 and $2^J - 1$. Observe that each $a_{\mathfrak{p}} \in \Theta$ have a prefixed order which is a power of two. Now, as above, take $a_{\mathfrak{p}} = 2^k$. Then $a_{\mathfrak{p}}$ generates $\mathbb{Z}/2^{J-k}\mathbb{Z} \subset \mathbb{Z}/N\mathbb{Z}$ and $\mathbb{Z}/2^{J-k}$ has 2^{J-k-1} generators (the odd numbers between 0 and 2^{J-k}). Then it follows that the order of 2^k and those odd numbers times 2^k is the same. Doing such process for each k, the assertion follows because we have classified all the even numbers by its fixed order 2^k. Following this thread, when $a_{\mathfrak{p}}$ is not a power of two, $a_{\mathfrak{p}} \mod 2^{k+1} = 2^k$ as we wanted.

□

Now, we are ready to face the solution of the linear system given by Equation (5.9) by means of the following proposition.

Proposition 5.2.6. *Fix $J > 0$ and let $Ax = b$ a $N \times N$ linear system where $N = 2^J$. Assume that the matrix A, for $i = 0, \ldots, N-1$, is given by*

$$a_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ -\lambda_i & \text{if } j = i - p \mod (N), \\ 0 & \text{otherwise} \end{cases}$$

with $\lambda_i \in \mathbb{R}$ and $p > 0$ is a fixed integer. Take $\alpha_0 = \lambda_0$, $\beta_0 = b_0$, $\nu_0^0 = 0$ as initial seeds and define, for $k = 0, \ldots, N-1$, the following finite recurrences

$$\alpha_{\nu_0^{k+1}} = \lambda_{\nu_0^{k+1}} \alpha_{\nu_0^k} \quad \text{and} \quad \beta_{\nu_0^{k+1}} = \lambda_{\nu_0^{k+1}} \beta_{\nu_0^k} + b_{\nu_0^{k+1}} \tag{5.12}$$

with $\nu_0^{k+1} = \nu_0^k + p \mod (N)$. If $\alpha_{\nu_0^{N-1}} \neq 1$ then the system has a unique solution which is given by

$$x_k = \begin{cases} \alpha_k x_{N-p} + \beta_k, & \text{for } k = 0, \ldots, p-1, \\ \beta_{\nu_0^{N-1}} / (1 - \alpha_{\nu_0^{N-1}}), & \text{for } k = N-p, \\ \alpha_k x_{N-p} + \beta_k, & \text{for } k = p+1, \ldots, N-1. \end{cases} \tag{5.13}$$

Proof. Without loss of generality, we claim that we can suppose that $\{\nu_0\}_0^N$ is exhaustive in terms of Lemma 5.2.5. The proof will be constructive and it is a slight modification of the Thomas algorithm: the tri-diagonal Gaussian elimination. Certainly, we get $x_0 = b_0 + \lambda_0 x_{N-p}$ from the first equation, $x_0 - \lambda_0 x_{N-p} = b_0$, of the linear system given by (5.12). Also we can define the coefficients $\alpha_0 = \lambda_0$ and $\beta_0 = b_0$. In view of that, and following the idea of the Gaussian elimination, we can replace x_0 in the $\nu_0^1 = p$-th equation of the system. That is, $x_{\nu_0^1} - \lambda_{\nu_0^1} x_0 = b_{\nu_0^1}$:

$$\begin{aligned} x_{\nu_0^1} - \lambda_{\nu_0^1} x_0 = b_{\nu_0^1} \quad &\Rightarrow \quad x_{\nu_0^1} - \lambda_{\nu_0^1}(b_0 + \lambda_0 x_{N-p}) \\ &= \quad \lambda_{\nu_0^1} b_1 + \lambda_{\nu_0^1}\lambda_0 x_{N-p} + x_{\nu_0^1} = b_{\nu_0^1} \\ &\Rightarrow \quad x_{\nu_0^1} = \lambda_{\nu_0^1}\lambda_0 x_{N-p} + \lambda_{\nu_0^1} b_0 + b_{\nu_0^1} \\ x_{\nu_0^1} &= \quad \alpha_1 x_{N-p} + \beta_1, \end{aligned}$$

where $\alpha_1 = \lambda_{\nu_0^1}\lambda_0$ and $\beta_1 = \lambda_{\nu_0^1} b_0 + b_{\nu_0^1}$. Moreover, such coefficients can be written as $\alpha_1 = \lambda_{\nu_0^1}\alpha_0$ and $\beta_1 = \lambda_{\nu_0^1}\beta_0 + b_{\nu_0^1}$. Following this replacement idea, we can repeat such process N steps to produce a list of coefficients α_i and β_i such that they are used to recover each unknown x_i in terms of x_{N-p}. Since we are assuming that ν_0 is exhaustive, such recursive construction sweeps all unknowns x_i of the linear system given by (5.12). That is, the index i takes all the values in $\mathfrak{N}$ in a such a way that $N-p$ is the last one. Moreover, for $i \neq N-p$, each unknown x_i is written in terms of x_{N-p} and its corresponding coefficients α_i and β_i. Because of this, notice $N-p$-th equation of the linear system is $x_{N-p} + \lambda_{N-p} x_{\nu N-1} = b_{N-p}$. Therefore, since $x_{\nu_0^{N-1}}$ is given in terms of x_{N-p}, it follows that

$$x_{N-p} = \frac{\beta_{\nu_0^{N-1}}}{1 - \alpha_{\nu_0^{N-1}}} + b_{N-p}$$

Thus, once we have x_{N-p}, we can recover each x_i by means of Equation (5.13) as we wanted to show. To conclude the proof, it only remains to mention that the proof of the aforesaid claim is done in Lemma 5.2.5. □

Summarizing, the linear solution of the linear system given by (5.12) is done in two steps. The first one is devoted to compute the coefficients α_i and β_i. The second part is to solve back for the actual unknowns x_i from x_{N-p}. Of course, when ν is not exhaustive it is enough to repeat the above process as many times as cycles are.

Remark 5.2.7. In view of Proposition 5.2.6 and Lemma 5.2.5 we can derive, from its constructive proof, a fast numerical algorithm to determine the solution of the linear system given by Equation (5.9). Notice that in terms of computer running time, it is clear what we have claimed before: the $\mathcal{O}(N)$ time. Even when ν is not exhaustive. Indeed,

the solution is constructed in forward-backward manner. Therefore, taking into account the operations at each step of the recursion, there are $2N$ multiplications plus N sums going forwards and N multiplications plus N sums going backwards. Of course, when ν is not exhaustive, by the second property of ν described in Lemma 5.2.5, we have to repeat k times the same kind of N/k operations. ∎

Drawing to close this part, we will make two remarks. The first one can be considered the link between the dynamical properties of the skew product considered and the linear system.

Remark 5.2.8. Let A be a matrix as the one given by Proposition 5.2.6. Firstly, notice that if one subtract the identity matrix from A, the resultant, Δ, is the one of the Jacobi method. Now, to decide if the method converges or not we must have a look at its spectral radius, $\varsigma(\Delta)$. To this end, it can be shown that its characteristic polynomial of Δ is

$$p_c(x) = x^N - \prod_{i=0}^{N-1} a_{i,i+p} = x^N - \Pi_N$$

Clearly, the roots of the $p_c(x)$ are the N-th roots of Π_N. Hence, as a first approximation, we have that the determinant of Δ will be crucial to decide if the Jacobi method is convergent or not.

But, moreover, moving to the dynamical system given by Equation (3.1) something more can be done. To fix ideas we will restrict ourselves to the multiplicative case. Indeed, if we consider the infinite case, N goes to ∞, we have that the spectral radius is bounded if the above infinite product is finite. That is, if by definition of Π_N, it holds that $\Pi_\infty < \infty$. Or, considering the N-root and in terms of the logarithms, we have a bounded $\varsigma(\Delta)$ if

$$\frac{1}{N} \sum_{i=0}^{N-1} \log(f'(\varphi(\theta_i))g(\theta_i) < \infty.$$

Now, taking this expression to the limit and using the Birkhoff Ergodic Theorem once more, we have that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \log(f'(\varphi(\theta_i))g(\theta_i) = \int_{\mathbb{S}^1} \log(f'(\varphi(\theta))g(\theta))\, d\theta < \infty.$$

which is exactly the expression of the Lyapunov exponent λ_φ. That is, we can determine if the Jacobi method converges with the Lyapunov exponent. Indeed, since φ is an attractor of the dynamical system given by Equation (3.5) then $\lambda_\varphi < 0$. Therefore, we have that $\varsigma(\Delta) < \Pi_\infty = e^{\lambda_\varphi} < 1$.

Besides of the Jacobi convergence, λ_φ controls if the system is undetermined or not also. Indeed, we can apply similar arguments as above to the determinant of Δ_N which is, precisely Π_N and also the product of all α_i's from the recursion (5.12). Hence, looking

at the part of x_{N-p} at the final lines of the Proposition 5.2.6's proof, it is clear that Π_N must be different from 1 and such difference is controlled by λ_φ. ∎

The second remark points out the stability, in terms of the propagation of the errors, of the strategy performed.

Remark 5.2.9. Of course, one must check what is more cheaper: the iterative way or the exact way. But, in terms of numerical linear algebra it is important to emphasize that we do not *add* any kind of error. Away from the fact that the linear system is explicitly written and its solution also, the matrix used is neutral in terms of the error propagation. The reason is simple, each and every one matrices used when preconditioning is orthogonal. Therefore since its spectral radius is one it can not be an error propagation source. ∎

Now, we can try to apply the same idea of the right precondition technique when $p > 1$ vanishing moments are considered. As we will see in the next section, the matrices Id and Ξ are not as good as possible. However, some solutions will be proposed.

5.2.2.2 Daubechies' strategy

Let us start by saying that with Daubechies wavelet we have also applied GMRES directly to solve Equation (5.6) up to the limit of its usage. Indeed, recall that in the Arnoldi process (Algorithm 3) is necessary to save some vectors to evaluate the minimizer in GMRES. Hence, when the system becomes big GMRES slows down its efficiency. Moreover, as we have pointed out at the beginning of Section 5.2.2, due to the spectrum the GMRES breakdown is reached closer to the iterate number $k = N$ in our situation. That is, we solve the system given by Equation (5.6) exactly but with the penalization of the sweep of all the matrix and the storage of the Krylov subspace.

However, in view of the previous section we have tried to go further using precondition strategies. Having said that, let us move to the final comment of the aforesaid section. The problem is when Id and Ξ are not as good as possible. Indeed, as we have said in Section 2.2, as farther to the right of the matrix $\tilde{\Psi}^{\mathrm{PER}}$ are you located then more zeros you will have. Therefore, the sample of ψ_l^{PER} at the points $\tilde{\theta}_i$ is not enough to apply the trapezoidal rule in a good way. Actually, at each entry of the matrix $\tilde{\Psi}_N^{\mathrm{PER}}\tilde{\Psi}_N^{\mathrm{PER}\top}/N$ we are doing

$$\frac{1}{N}\sum_{i=0}^{N-1}\psi_l^{\mathrm{PER}}(R_\omega(\theta_i))\psi_m^{\mathrm{PER}}(R_\omega(\theta_i)),$$

assuming that $\tilde{\Psi}_N^{\mathrm{PER}}\tilde{\Psi}_N^{\mathrm{PER}\top} = N\mathrm{Id}$. Hence, $\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}} = N\mathrm{Id}$ but the nonzero coefficients are not enough to ensure an accurate approximation of the integral

$$\int_{\mathbb{S}^1}\psi_l^{\mathrm{PER}}(R_\omega(\theta))\psi_m^{\mathrm{PER}}(R_\omega(\theta))\,d\theta = \delta_{l,m}.$$

That is, even with a given tolerance, we do not have any way to get the identity matrix. Furthermore, the trick of setting $\widetilde{\Delta_n \Xi}$ to be a matrix such that

$$\mathbf{JF}\tilde{\Psi}_N^{\mathrm{PER}\top} = (\mathrm{Id} - \widetilde{\Delta_N \Xi})$$

does not brings us good news because of the uncontrollable behavior of such new matrix $\widetilde{\Delta_n \Xi}$. However, if we fix $N = 2^J$, with $J > 0$, define the matrix $\mathrm{Id} = \frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$ where ψ_l^{PER} it is an orthonormal basis of $\mathbb{S}^1$ generated by a Daubechies wavelet with $p > 1$ vanishing moments. Now, recall that we want to solve Equation (5.6) but, in the same way as Haar, we can use a left preconditioning technique:

$$\left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right)(\mathrm{X}) = -\mathrm{Inv} \Leftrightarrow \tilde{\Psi}_N^{\mathrm{PER}\top}\left(\tilde{\Psi}_N^{\mathrm{PER}} - \Delta_N \Psi_N^{\mathrm{PER}}\right)(\mathrm{X}) = -\tilde{\Psi}_N^{\mathrm{PER}\top}\mathrm{Inv}.$$

That is, we will solve (using either GMRES or TFQMR)

$$\left(\mathrm{Id} - \tilde{\Psi}_N^{\mathrm{PER}\top}\Delta_N \Psi_N^{\mathrm{PER}}\right)(\mathrm{X}) = -\tilde{\Psi}_N^{\mathrm{PER}\top}\mathrm{Inv} = \widetilde{\mathrm{Inv}}. \tag{5.14}$$

instead of Equation (5.6). To this end, we must say that the matrix Id is closer to identity matrix in the sense of the following remark.

Remark 5.2.10. Under the above assumptions on $\psi(x)$, the matrix $\mathrm{Id} = \frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$ is symmetric and its eigenvalues are strictly positive. Moreover the eigenvalues of Id are centered around 1. The proof of the first two assertions are done by definition of the matrix Id. Let us focus on the third property. To this end, since $\mathrm{Id} = \frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$, then, applying roughly the Trapezoidal rule, each entry of Id verifies

$$\begin{aligned}
\frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{l,m} &= \frac{1}{N}\sum_{i=0}^{N-1}\psi_l^{\mathrm{PER}}(R_\omega(\theta_i))\psi_m^{\mathrm{PER}}(R_\omega(\theta_i)) \\
&= \int_{\mathbb{S}^1}\psi_l^{\mathrm{PER}}(R_\omega(\theta))\psi_m^{\mathrm{PER}}(R_\omega(\theta))\, d\theta + \varepsilon_{l,m} \\
&= \delta_{l,m} + \varepsilon_{l,m}.
\end{aligned}$$

Now, let

$$\varepsilon_i = \begin{cases} \max\limits_{l=0,\ldots,N-1}\varepsilon_{l,l} & \text{if } i = 1, \\ \max\limits_{\substack{l,m=0,\ldots,N-1 \\ l\neq m}}\dfrac{\varepsilon_{l,m}}{N} & \text{if } i = 2. \end{cases}$$

$\varepsilon_1 = \max\limits_{l=0,\ldots,N-1}\varepsilon_{l,l}$ and $\varepsilon_2 = \max\limits_{l,m=0,\ldots,N-1 l\neq m}\varepsilon_{l,m}$ To conclude the proof apply the Gershgorin circle theorem. Indeed, all the eigenvalues of Id verifies that $|\lambda_i - (1+\varepsilon_1)| \leq \varepsilon_2 \subset \mathbb{R}$. And this is the crux of the matter. Indeed, even with a N big enough there is an unavoidable lack of precision. In other words, the diagonal of Id it is not plenty of ones because, again, of the samples of $\psi_l^{\mathrm{PER}}(R_\omega(\theta))$. This situation is, certainly, a bug that must be solved because it means that the wavelet values are not good enough (even with all the cautions that we have performed). ∎

The reason for introducing the matrix Id is twofold. The first one is to emphasize the need of have good ψ_l^{PER}-values. On the other hand, since Id does not depend on Δ_N one can store it in order to do not perform the multiplication at each step of the Newton's method. The performance of such storage has implications on the total amount of the number of operations at each Newton step. Since we do not use it we only mention that in the same way as in the Haar's case, we try to generate an image of the matrix Id using a particular instance of such matrix depicted below. Since we know that it is symmetric, we have reinforced such idea only displaying the upper diagonal. The general properties of such matrix are displayed in the Appendix.

$$\begin{pmatrix}
1 & 0 & a_1 & a_1 & a_2 & a_2 & a_2 & a_2 & a_3 & a_3 & a_3 & a_3 & a_3 & a_3 & a_3 & a_3 \\
 & \delta_{0,0} & b_1 & -b_1 & b_2 & b_3 & -b_2 & -b_3 & b_4 & b_5 & b_6 & b_7 & -b_4 & -b_5 & -b_6 & -b_7 \\
 & & \delta_{1,0} & \delta_{1,1} & e_0 & e_1 & e_2 & e_3 & f_0 & f_1 & f_2 & f_3 & f_4 & f_5 & f_6 & f_7 \\
 & & & \delta_{1,0} & e_2 & e_3 & e_0 & e_1 & f_4 & f_5 & f_6 & f_7 & f_0 & f_1 & f_2 & f_3 \\
 & & & & \delta_{2,0} & \delta_{2,1} & \delta_{2,2} & \delta_{2,1} & g_0 & g_1 & g_2 & g_3 & g_4 & g_5 & g_6 & g_7 \\
 & & & & & \delta_{2,0} & \delta_{2,1} & \delta_{2,2} & g_6 & g_7 & g_0 & g_1 & g_2 & g_3 & g_4 & g_5 \\
 & & & & & & \delta_{2,0} & \delta_{2,1} & g_4 & g_5 & g_6 & g_7 & g_0 & g_1 & g_2 & g_3 \\
 & & & & & & & \delta_{2,0} & g_2 & g_3 & g_4 & g_5 & g_6 & g_7 & g_0 & g_1 \\
 & & & & & & & & \delta_{3,0} & \delta_{3,1} & \delta_{3,2} & \delta_{3,3} & \delta_{3,4} & \delta_{3,3} & \delta_{3,2} & \delta_{3,1} \\
 & & & & & & & & & \delta_{3,0} & \delta_{3,1} & \delta_{3,2} & \delta_{3,3} & \delta_{3,4} & \delta_{3,3} & \delta_{3,2} \\
 & & & & & & & & & & \delta_{3,0} & \delta_{3,1} & \delta_{3,2} & \delta_{3,3} & \delta_{3,4} & \delta_{3,3} \\
 & & & & & & & & & & & \delta_{3,0} & \delta_{3,1} & \delta_{3,2} & \delta_{3,3} & \delta_{3,4} \\
 & & & & & & & & & & & & \delta_{3,0} & \delta_{3,1} & \delta_{3,2} & \delta_{3,3} \\
 & & & & & & & & & & & & & \delta_{3,0} & \delta_{3,1} & \delta_{3,2} \\
 & & & & & & & & & & & & & & \delta_{3,0} & \delta_{3,1} \\
 & & & & & & & & & & & & & & & \delta_{3,0}
\end{pmatrix}.$$

To conclude the methodology to solve Equation (5.6) in the Daubechies case, let us make some comments to see the core of the Newton's iteration. To do this, let us assume that $\mathrm{Id} = \tilde{\Psi}^{\mathrm{PER}\top}\tilde{\Psi}^{\mathrm{PER}}$ and $\mathrm{Id} = \Psi^{\mathrm{PER}\top}\Psi^{\mathrm{PER}}$ are *the* identity matrix. Under such assumptions we are solving

$$\left(\mathrm{Id} - \tilde{\Psi}^{\mathrm{PER}-1}\Delta\Psi^{\mathrm{PER}}\right)\mathrm{X} = \widetilde{\mathrm{Inv}}$$

because $\tilde{\Psi}^{\mathrm{PER}\top} = (\tilde{\Psi}^{\mathrm{PER}})^{-1}$. Now recall that $B^{-1}AB$ has the same eigenvalues as A whenever B is invertible (see e.g [HJ85]). Then, taking $B = (\tilde{\Psi}^{\mathrm{PER}\top})^{-1}$ we have that $\Delta\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\top}$ has the same eigenvalues as $\tilde{\Psi}^{\mathrm{PER}\top}\Delta\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\top}(\tilde{\Psi}^{\mathrm{PER}\top})^{-1}$. In other words, in terms of the spectral radius it is true that $\varsigma(\tilde{\Psi}^{\mathrm{PER}-1}\Delta\tilde{\Psi}^{\mathrm{PER}}) = \varsigma(\Delta\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\top})$. Moreover, the right hand matrix controls the left hand side matrix which has an expression similar to the reducibility property (see Section 3.4).

On the other side, assume that $\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\top}$ verifies a similar property as Equation (5.7). Then, because our assumptions the characteristic polynomial of $\Delta\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\top}$ is, following the same thread as in the Haar's case, $x^n + \det(\Delta)$. Hence, by definition of the matrix Δ we have that the spectral radius is bounded by the Lyapunov exponent

of φ. Therefore, since φ is an attractor $\varsigma(\tilde{\Psi}_N^{\mathrm{PER}\,-1}\Delta_N\Psi_N^{\mathrm{PER}})<1$. The problem is that the assumption which leads to the Haar's argument is rather false (on the contrary we do $P\Psi^{\mathrm{PER}}=\tilde{\Psi}^{\mathrm{PER}}$ and this is not the case). However the situation, after numerical explorations on the matrix product $\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\,\top}$, seems to be close to such situation. That is, $\Psi^{\mathrm{PER}}\tilde{\Psi}^{\mathrm{PER}\,\top}$ is a band matrix. On the other side the first two assumptions in view of Remark 5.2.10 are also false but they are *under* control. In other words, it seems that a perturbation argument (setting, for example, $\tilde{\Psi}^{\mathrm{PER}\,\top}\tilde{\Psi}^{\mathrm{PER}}\sim \mathrm{Id}+\varepsilon$) arises as a good way to face the problem of provide regions such that the hypothesis of Newton Kantorovich Theorem are satisfied.

Having said that, let this explanation be a justification of the importance of having under control of the Lyapunov exponent of φ in order to let the Newton's method converge (besides the possible numerical errors on the calculation of ψ_l^{PER}).

Recovering of X. One problem is left for the Haar's case. Indeed, since we are doing a right precondition, we need to multiply the solution y by the inverse of the precondition matrix. That is, in our case $\tilde{\Psi}^{\mathrm{PER}}$. But, again by Lemma 5.2.4, we have that our solution X is particularly simple. Indeed, $\mathrm{X}=\tilde{\Psi}^{\mathrm{PER}}y=P\Psi_N^{\mathrm{PER}}y$ is the solution of the linear system given by Equation (5.9). Notice that this is not the case of Daubechies wavelets with $p>1$ vanishing moments because we solve a left preconditioned system.

5.3 Summary: the method

Before we present the results, we wish to summarize all of the three steps performed to find the desired vector $\mathrm{D}_N^{\mathrm{PER}}\in\mathbb{R}^N$ and, hence, the coefficients $d_{-j}^{\mathrm{PER}}[n]$ using the Newton's method. Besides of such objective, we will include along this description of the method some comments and remarks. Such ideas are focused to give an important bootstrap on the efficiency of the implementation. To do this, let us recall some notation to fix ideas.

(a) As usual along this chapter Ψ_N^{PER} will be the matrix whose columns are the first N elements of the basis $\{\psi_l^{\mathrm{PER}}\}_{l\in\mathbb{Z}^+}$ evaluated at the points in $\mathbb{S}^1$ given by $\theta_i=i/N$. That is

$$\Psi_N^{\mathrm{PER}}{}_{i,l}=\frac{1}{\sqrt{N}}\begin{cases}1 & \text{if } l=0,\\ \psi_l^{\mathrm{PER}}(\theta_i) & \text{if } l\neq 0.\end{cases}$$

In the same way, $\tilde{\Psi}_N^{\mathrm{PER}}$ has the same columns as Ψ_N^{PER} but evaluated at rotated points of the unit circle $\tilde{\theta}_i=i/N+\omega$.

(b) As we have already said let $\mathrm{D}_N^{\mathrm{PER}}\in\mathbb{R}^N$ be the following vector of wavelet coefficients

$$\left(a_0, d_0^{\mathrm{PER}}[0], d_{-1}^{\mathrm{PER}}[0],\ldots, d_{-j}^{\mathrm{PER}}[N-1],\right)^\top$$

where $-j = \lfloor \log_2(N) \rfloor$. Also, $\mathrm{D}^{\mathrm{PER}}_{N,k)} \in \mathbb{R}^N$ stands for the k-th approximation given by the Newton's method.

(c) In the same way as the above item, let $\tilde{\mathrm{D}}^{\mathrm{PER}}_N \in \mathbb{R}^N$ be the vector of the *rotated wavelet coefficients* given by

$$\tilde{\mathrm{D}}^{\mathrm{PER}}_N = \tilde{\Psi}^{\mathrm{PER}}_N \mathrm{D}^{\mathrm{PER}}_N. \tag{5.15}$$

The same definition for $\tilde{\mathrm{D}}^{\mathrm{PER}}_{N,k)}$ holds.

(d) Let $\Phi_N \in \mathbb{R}^N$ be the following vector

$$\Phi_N = (\varphi(\theta_0), \varphi(\theta_1), \ldots, \varphi(\theta_{N-1}))^\top .$$

That is, the attractor φ of the system evaluated at the prescribed points $\theta_i \in \mathbb{S}^1$, with $i = 0, 1, \ldots, N-1$.

Now we are ready to summarize the developed methodology along this chapter.

Solving the Invariance Equation using Daubechies wavelets. Consider a skew product given by Equation (3.1). To manage the invariance equation, given by Equation (3.15), associated to such skew product perform the following steps:

Step Previous. Being fixed $J > 0$ big enough set $N = 2^J$ and a tolerance such that Equation (3.4) is verified to get the vector Φ_N up to a given tolerance. Fix a Daubechies wavelet with $p \geq 1$ vanishing moments.

Step 1: Initial guess. Compute, using the results of Section 2.2, the matrices Ψ^{PER}_N and $\tilde{\Psi}^{\mathrm{PER}}_N$. Use them to set $\frac{1}{N}{\Psi^{\mathrm{PER}}_N}^\top \Phi_N = \mathrm{D}^{\mathrm{PER}}_{N,0)}$ as an initial guess for the recurrence (5.2). In order to avoid computational efforts notice that if the Haar wavelet is used, that is $p = 1$, we have that $\tilde{\mathrm{D}}^{\mathrm{PER}}_{N,0)} = P\Phi_N$ because $\tilde{\Psi}^{\mathrm{PER}}_N {\Psi^{\mathrm{PER}}_N}^\top = P$.

Step 2: Solving the system. In order to find solutions of the linear system given by Equation (5.6) take one of the two possibilities:

(a) Apply, verbatim, GMRES to Equation (5.6) (up to your technical computer limitations) or,

(b) take one of the following strategies:

(a) In the case of Haar apply ${\tilde{\Psi}^{\mathrm{PER}}_N}^\top$ to solve a right preconditioned system. Observe that, at the k-th Newton iterate we have that the i-th component of invariance equation written algebraically is

$$\left[\tilde{\Psi}^{\mathrm{PER}}_N \mathrm{D}^{\mathrm{PER}}_{N,k)}\right]_i - \frac{\partial F_{\sigma,\varepsilon}\left(\theta_i, \left[\Psi^{\mathrm{PER}}_N \mathrm{D}^{\mathrm{PER}}_{N,k)}\right]_i\right)}{\partial x} = 0.$$

In this case, using Equation (5.15) and the definition of $\mathrm{D}_N^{\mathrm{PER}}$, using the rotated coefficients we have an easy expression of the invariance equation

$$\left[\tilde{\mathrm{D}}_{N,k)}^{\mathrm{PER}}\right]_i - \frac{\partial F_{\sigma,\varepsilon}\left(\theta_i, \left[P^\top \tilde{\mathrm{D}}_{N,k)}^{\mathrm{PER}}\right]_i\right)}{\partial x} = 0$$

That is, the invariance equation is much easier with the rotated coefficients than with the standard ones. Clearly, with such change of variables has some cosmetic effect to the Jacobian matrix. Hence, in the case of Haar, apply Proposition 5.2.6 to solve the linear system given by Equation (5.9) and (5.10).

(b) In the case of Daubechies wavelet with $p > 1$ vanishing moments apply $\tilde{\Psi}_N^{\mathrm{PER}\top}$ to solve the left preconditioned version of the system. To solve Equation (5.14) use TFQMR from a certain level $J > J_0 > 0$ (up to J_0 use either TFQMR or GMRES).

Step 3: Updating the recurrence. Recall that in the Haar's situation, we are applying a right preconditioning strategy to Equation (5.6). Hence, being p the vanishing moments of the $\mathbb{R}$-Daubechies wavelet $\psi(x)$

$$\begin{cases} \tilde{\mathrm{D}}_{N,k+1)}^{\mathrm{PER}} = \tilde{\mathrm{D}}_{N,k)}^{\mathrm{PER}} + \mathrm{X} & \text{if } p = 1, \\ \mathrm{D}_{N,k+1)}^{\mathrm{PER}} = \mathrm{D}_{N,k)}^{\mathrm{PER}} + \mathrm{X} & \text{otherwise.} \end{cases}$$

Notice that working with $\tilde{\mathrm{D}}^{\mathrm{PER}}$ has consequences when recovering $\mathrm{D}_{N,\star)}^{\mathrm{PER}}$. Indeed, observe that $\mathrm{D}_{N,\star)}^{\mathrm{PER}}$ is the solution of the linear system, according to Equation (5.15), given by $\tilde{\Psi}^{\mathrm{PER}}\mathrm{D}_{N,\star)}^{\mathrm{PER}} = \tilde{\mathrm{D}}_{N,\star)}^{\mathrm{PER}}$. Moreover, in the case of Haar, such solution is explicitly given by $\mathrm{D}_{N,\star)}^{\mathrm{PER}} = \tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\mathrm{D}}_{N,\star)}^{\mathrm{PER}}$.

Step 3: Updating the recurrence (bis). In the case of Haar, if one wants to recover the attractor just do the following

$$P^\top \tilde{\mathrm{D}}_{N,\star)}^{\mathrm{PER}} = \Psi_N^{\mathrm{PER}}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\mathrm{D}}_{N,\star)}^{\mathrm{PER}} = \Psi_N^{\mathrm{PER}}\mathrm{D}_{N,\star)}^{\mathrm{PER}} = (\varphi(\theta_0), \varphi(\theta_1), \ldots, \varphi(\theta_{N-1}))^\top = \Phi_N.$$

Remark 5.3.1. We want to emphasize that the same strategy performed with Haar can be done with other Daubechies wavelet. Hence, the strategy to work and compute with $\mathrm{D}^{\mathrm{PER}}$ must be the same as in the Haar's case. In view of that, when one uses $\tilde{\mathrm{D}}_{N,\star)}^{\mathrm{PER}}$ and $p \neq 1$ it is mandatory to solve, the linear system $\tilde{\Psi}_N^{\mathrm{PER}}\Phi_N = \mathrm{D}_{N,\star)}^{\mathrm{PER}}$ (once $\mathrm{D}_{N,\star)}^{\mathrm{PER}}$ is obtained from the above Step). Such point of view has drawbacks respect the time because you need to solve two (big) linear systems at each Newton's step. ∎

Now, we apply such method to the System (3.6) with such method(s). This is the main topic of the following section.

5.4 Results

Along this section we will perform the same kind of experimentation as the one in Chapter 4 but now using the wavelet coefficients $\mathrm{D}^{\mathrm{PER}}_{N,\star)}$. To this end, let us recall the system that we have called Keller-GOPY attractor.

$$\begin{pmatrix}\theta_{n+1}\\ x_{n+1}\end{pmatrix} = \begin{pmatrix} R_\omega(\theta_n) \\ 2\sigma\tanh(x)\cdot(\varepsilon+|\cos(2\pi\theta)|)\end{pmatrix}.$$

By Keller's Theorem, this attractor is the graph of a map $\varphi\colon \mathbb{S}^1 \longrightarrow \mathbb{R}^+$ and when $\varepsilon = 0$ the corresponding attractor is called *strange*.

Our strategy is try to see whether the system is pinched or not in terms of regularity. As in the previous chapter, we will use the same parameterization of the above system:

$$\begin{cases}\theta_{n+1} &= \theta_n + \omega \pmod 1,\\ x_{n+1} &= 2\sigma\tanh(x_n)(\varepsilon(\sigma)+|\cos(2\pi\theta_n)|)\end{cases} \tag{5.16}$$

with $\omega = \frac{\sqrt{5}-1}{2}$, $\sigma \in [1,2]$ and for practical reasons two parameterizations

$$\varepsilon_1(\sigma) = \begin{cases}(\sigma-1.5)^2 & \text{if } 1.5 \le \sigma \le 2.0,\\ 0 & \text{if } 1 < \sigma < 1.5.\end{cases} \qquad \varepsilon_2(\sigma) = \begin{cases}(\sigma-2.0)^2 & \text{if } 2.0 \le \sigma \le 2.25,\\ 0 & \text{if } 1.9 < \sigma < 2.0.\end{cases}.$$

In this way the system is pinched if and only if $\sigma \le 1.5$ for ε_1 and $\sigma \le 2.0$ for ε_2. Also, recall that when σ tends to 1 then $\varphi \equiv 0$. The aim of such replay is the poor precision of the signal, for $\sigma \le 1.5$, when one uses the Fast Wavelet Transform. As a matter of fact, it can not be used as a initial seed of the Newton's method due to the extremely complicated geometry of φ since it can not capture most of the zeros of φ. We will come back to such topic in Section 6.2.2.

5.4.1 Haar wavelets

In this part we will use Daubechies wavelet with $p = 1$ vanishing moments with the parameterization $\varepsilon_1(\sigma)$. That is, we will perform the calculation of $\mathrm{D}^{\mathrm{PER}}_{N,\star)}$ using the Haar basis. Hence, recall Theorem 1.2.15 and Corollary 1.2.16 we have the following region of allowed s depicted in Figure 5.1. We focus our attention to the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ and $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$.

In the case of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ a word of caution must be done. Indeed, it is necessary to to remark that in the case of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ we will apply blindly Corollary 1.2.16 adapted to such case. This means that, at the beginning, we do not take care in which is the value of s obtained. Of course, when we will classify φ we will check if the value of s is in the allowed range. We will justify such procedure in the following lines.

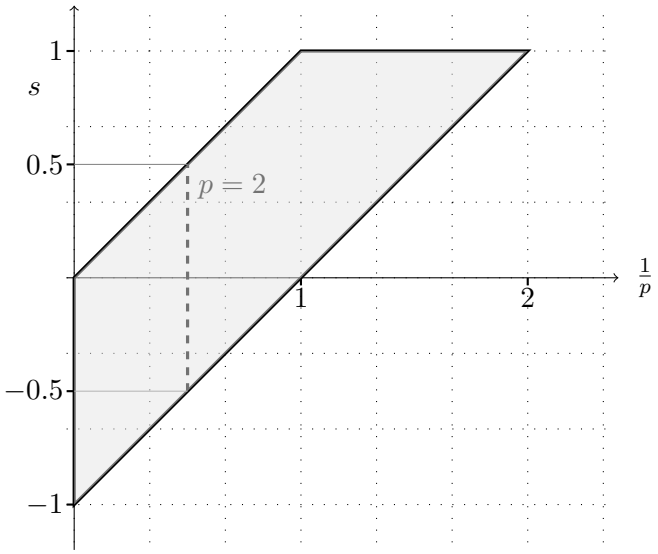


Figure 5.1: The range of allowed s for the spaces $\mathscr{B}^s_{p,q}(\mathbb{S}^1)$ according to Theorem 1.2.15. The dashed vertical line corresponds to $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$.

As we already know, φ must evolve from a continuous function to an upper semi-continuous function when becomes pinched. Concretely, in the case of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ if ε tends to 0 then s also. That is what we have tried to capture in Figure 5.2. Let us explain it carefully in our case. Since $\varphi \in \mathscr{L}^2(\mathbb{S}^1)$ we can approximate it using the Haar wavelet basis. Also, one can ask on how the coefficients vanish using the linear model depicted in Equation (1.9) in the same way as in Chapter 4. But in view of Figure 5.1 we can not classify φ in terms of the regularity spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ with $s > 0$. This is not a surprising fact since $s > 0$ means continuity and $\varepsilon > 0$ implies the continuity of φ. Nevertheless, we have find the same kind of behavior as in Figure 4.3. Concretely, when the system becomes pinched, for $\sigma \leq 1.5$, then $\varphi \in \mathscr{B}^0_{\infty,\infty}(\mathbb{S}^1)$. On the other side when ε is positive, for $\sigma > 1.5$, the estimated s is strictly positive with a high Pearson correlation coefficient (see Figure 5.2). In other words, it seems that the Haar wavelet can detect the pinching value in some sense and it also classifies, in terms of $\varphi \in \mathscr{B}^0_{\infty,\infty}(\mathbb{S}^1)$, with accuracy.

In view of the above results, we have performed a zoom around the pinched case, $\sigma = 1.5$, to see the behavior of how the estimation of the regularity parameter vanishes. In view of Figure 4.6 we have considered $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$ as the regularity space. The results, which are collected in Figure 5.3, describes the same situation as in the previous case: detection of the pinched point and the regularity space.

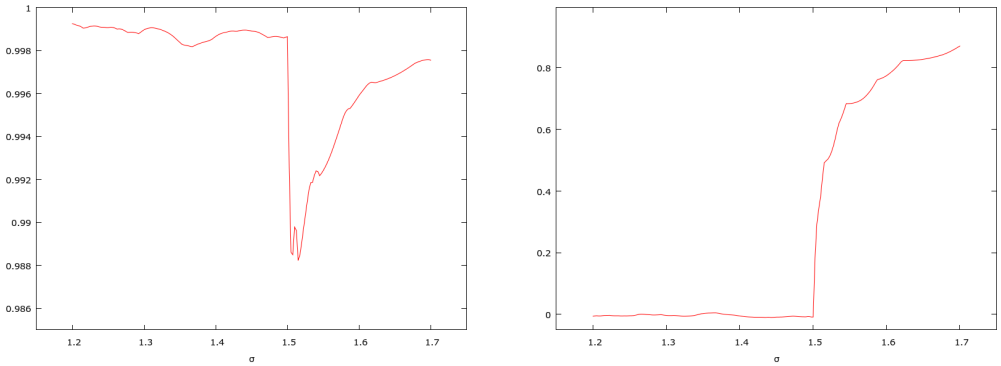


Figure 5.2: On the left hand side, it is plotted the Pearson correlation coefficient along $\varepsilon_1(\sigma)$ with $\sigma \in (1.2, 1.7)$. Notice that the best values are on the pinched zone. On the right hand side the estimation of the regularity parameter along such parameterization. The same comments, in the sense of the detection of the pinched point, as the ones in Figure 4.3 apply.

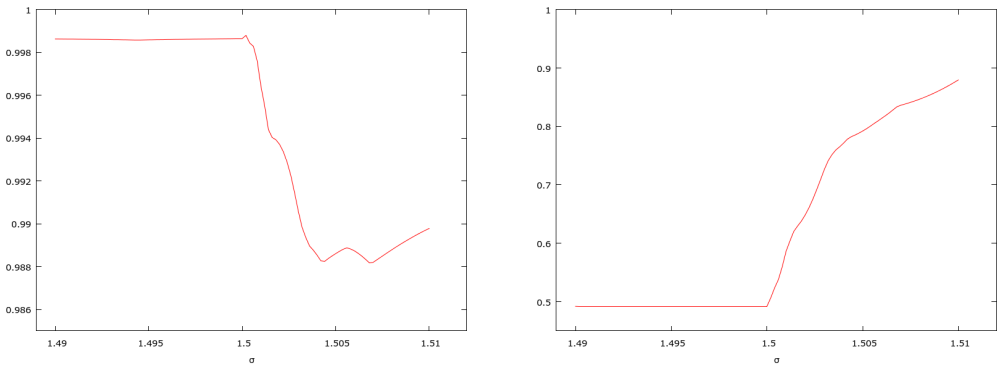


Figure 5.3: On the left hand side, it is plotted the Pearson correlation coefficient along $\varepsilon_1(\sigma)$ with $\sigma \in (1.49, 1.51)$. On the right hand side the estimation, in terms of $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$, of the regularity of φ along such parameterization.

5.4.2 Daubechies wavelets

In this part we will use Daubechies wavelet with $p = 10$ vanishing moments with the parameterization $\varepsilon_2(\sigma)$. The choice of such wavelet is done in view of the results from Chapter 4 concerning to the Pearson correlation coefficient, *efficiency in computing* and the Figure 5.4. Let us explain this three points. To this end, let us start by the

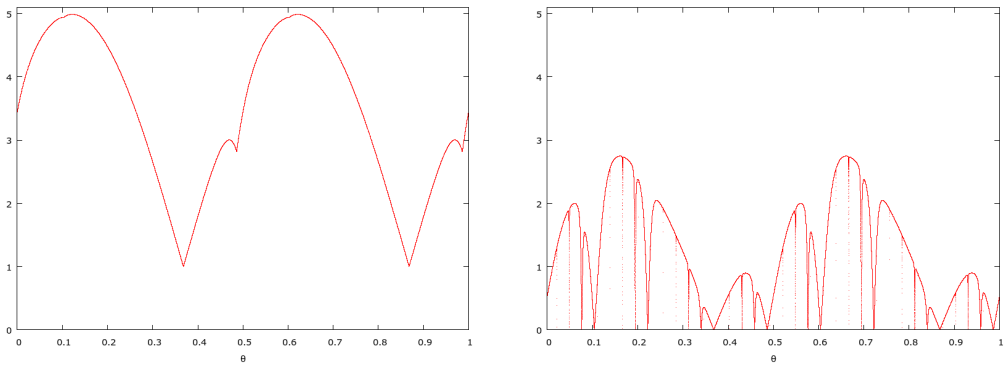


Figure 5.4: On the left hand side, it is plotted the attractor for the System (3.6) with $\sigma = 2.0$ and $\varepsilon = 0.25$ whereas on the right hand side $\varepsilon = 0.0$ and $\sigma = 2.0$.

geometrical situation induced by the dynamics. As we have pointed out in Section 3.3 φ has more visible spikes in terms of ε and σ (compare Figure 3.1 and 5.4). When the system becomes pinched there is a dense set of zeros but only some of them are visible. Hence, the question is how a regular enough wavelet can capture such fact. This links to question of the Person correlation coefficient. Indeed, recall that in Chapter 4 we have dealt with φ from the signal analysis point of view. As a matter of fact, in Section 4.3 Daubechies wavelet with $p = 10$ vanishing moments seems to have the better ratio between Pearson correlation coefficient and regularity assessment. This almost explain our choice. It only remains to say that such wavelet is one of the first wavelets such that has an interlace between regularity of $\psi(x)$ and the sparsity of the matrix Ψ^{PER}. In other words, recall that as p increases the regularity of $\psi(x)$ and its support also. Hence, $\psi^{\mathrm{PER}}(x)$ wraps around $\mathbb{S}^1$ many times (in fact proportional to $\mathrm{supp}(\psi)$). Hence, as p increases the number of columns not purely sparse also increase. For the Daubechies wavelet with $p = 10$ vanishing moments this situation starts at $j = 5$. That is, Ψ^{PER} has 32 columns not completely sparse.

Remark 5.4.1. In view of the above comment let us emphasize that in order to take full advantage of the sparsity properties of Ψ^{PER} the number J, which recall that fix the matrix dimension to 2^J, must be greater than 5. ∎

Let us display the obtained results. However, a word of caution must be done.

Due to the stall induced either by TFQMR or the numerical errors (perhaps) inherited by ψ^{PER} we have diminished our tolerance level to 1×10^{-15} whereas in the Haar's case the it was 1×10^{-16}. Finally, recall that the classification of φ will be done using Theorem 1.2.11. Hence we focus our attention to the spaces $\mathscr{B}^{s}_{\infty,\infty}(\mathbb{S}^1)$.

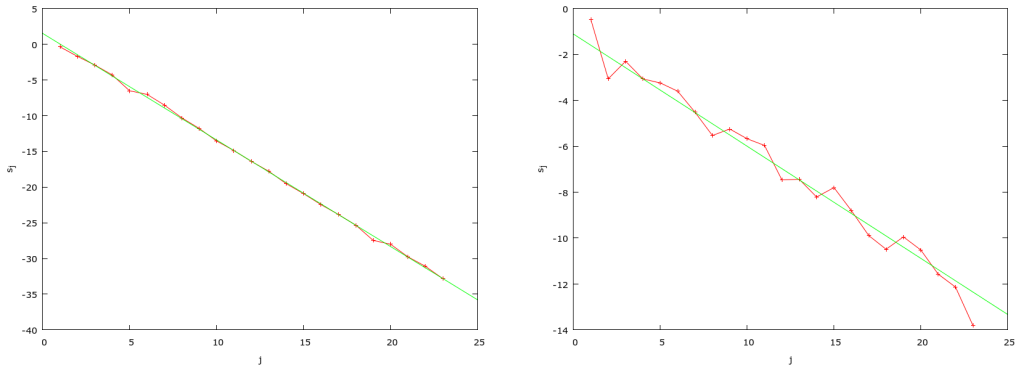


Figure 5.5: Graphs of the pairs (j, s_j) for the φ from System (3.6). On the left hand side, it is plotted the linear model with for the regularity of φ with $\sigma = 2.0$ and $\varepsilon = 0.25$ whereas on the right hand side $\varepsilon = 0.0$ and the same σ with a Person correlation coefficient of 0.999591 and 0.989908 respectively. On both cases, with $N = 2^{24}$, they predict *correctly* the regularity 0.992205 and 0.011645 respectively.

There are two remarkable facts. The first one is that we have fast convergence towards the wavelet coefficients when ε is positive with less iterates than Haar. On the contrary, when the system becomes pinched the number of iterates increases appreciably. However, as it can be seen in Figure 5.5 and 5.6 when one deals with the regularity of φ the Pearson coefficient of the linear models is clearly affected by the pinching condition and it is even worse than the one of the Fast Wavelet Transform. This is not a surprisingly fact since we are dealing with a discontinuous function and we are solving the invariance equation. Hence, the geometric properties of the invariant curve φ are transfered to $\mathrm{D}^{\mathrm{PER}}$ whereas on the Fast Wavelet case such properties are more difficult to capture. In our situation we can trap it as we will see in the following paragraph.

Indeed, as a second remarkable fact, the same comments as the ones in Figure 4.3 and 5.2 apply. That is, the coefficients $\mathrm{D}^{\mathrm{PER}}$ can detect the pinched zone since $\varphi \in \mathscr{B}^{0}_{2,2}(\mathbb{S}^1)$ when $\sigma \leq 2.0$ for $\varepsilon_2(\sigma)$ (see Figure 5.6). However, around the pinched point the Pearson coefficient is irregular. Again, when dealing with regularity, in view of the previous paragraphs and Section 4.3, the coefficients $\mathrm{D}^{\mathrm{PER}}$ can be used to describe tendencies on the evolution of the regularity of φ and to approximate it. This last point, without losing sight of the regularity, is our main goal.

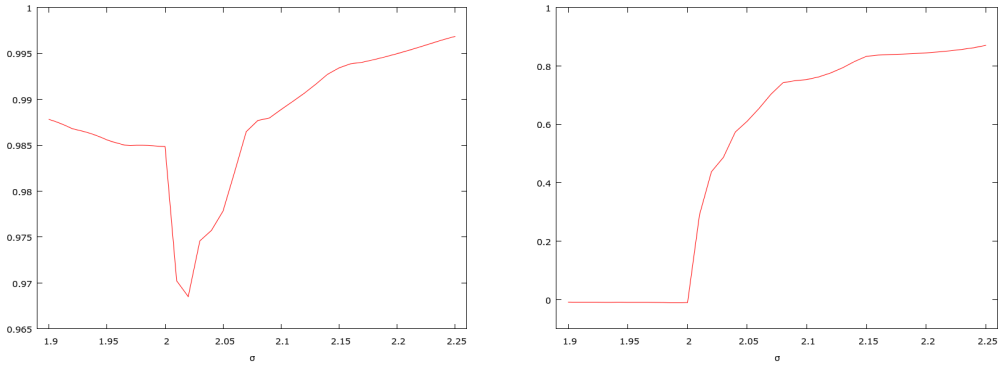


Figure 5.6: On the left hand side, it is plotted the Pearson correlation coefficient along $\varepsilon_2(\sigma)$ On the right hand side the estimation, in terms of $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$, of the regularity of φ along such parameterization.

Finally, in view of the above comments and the ones in Section 5.4.1 as much regular is φ more vanishing moments are required to converge faster towards the wavelet coefficients when ε is positive. However, the price to pay is the lack of precision if ε tends to zero when dealing with regularity topics.

5.5 Conclusions

To conclude and in light of all of the foregoing, we have two different situations and each one of them will be independent. However both can be used in a complementary way. Let us explain it in detail. Clearly, we have two methods that produce the same result: a solution of the invariance Equation, for the System (3.6), which turns out to be, up to a given tolerance, an approximation of φ:

$$\varphi \sim a_0 + \sum_{j=0}^{J-1} \sum_{n=0}^{2^j-1} \left\langle \varphi, \psi^{\mathrm{PER}}_{-j,n} \right\rangle \psi^{\mathrm{PER}}_{-j,n}.$$

However if one is focused on regularity properties, in view of Sections 5.4.1 and 5.4.2, Haar wavelet is not very useful for the positive ones. Having said that, both of them are a good tool to detect the pinching situation.

On the other side, the use of a Daubechies wavelet has inherited the the *Big Matrix* problem. That is, the number of operations to solve the System (5.6) is affordable, respect the user time, up to 2^{16}. Far beyond such quantity the method becomes unreachable and (machine) unstable. In terms of give an approximation of φ this means that we can perform it up to scale [1] $1.52587891 \times 10^{-5}$. After such $J_0 = 16$ we have

[1] In view of this "*big*" scale we do not have displayed the results.

solved System (5.14) using TFQMR. Before continue the explanation let us remark that we have used TFQMR because it is well suited for such linear system. However, this not means that TFQMR *is the unique* iterative solver valid. As a matter of fact, it is better to have a library of solvers such as Kaczmarz, Cimmino, Bi-Conjugate Gradient or even multi-grid methods (see e.g [ESW05, Saa03]).

Now, continuing with TFQMR with $24 \geq J > 16$ we must say that for a certain values of ε, close to zero, such solver can stall the update of the $\mathrm{D}^{\mathrm{PER}}_{N,n)}$ performed by Newton. In other words, TFQMR can provoke a non-convergence situation but close to converge towards $\mathrm{D}^{\mathrm{PER}}_{N,\star)}$. Having said that, we do not have a clear explanation of such fact and, moreover, such problem can be also induced for the numerical errors on the evaluation of ψ^{PER}_l and amplified by Lyapunov exponent.

On the other side, Haar wavelet it allows us to go 2^{10} times far away without any problem (we have stopped at $N = 2^{26}$ for practical reasons). That is, we can solve the aforesaid system in a real $\mathcal{O}(N)$ time. The counterpart is that for the regularity purposes Haar wavelet has "*fewer*" applications than Daubechies wavelet with $p > 2$ vanishing moments (recall that in Remark 1.1.12 $p = 2$ is the 4-tap Daubechies wavelet which it is not regular enough).

However in both situations, when the method converges, we have a semi analytical approximation of φ up to a given tolerance using the Newton's method. That is, given a certain tolerance we have computed numerically an invariant object φ. Clearly, since we can recover φ from $\mathrm{D}^{\mathrm{PER}}_N$ they are a good approximation in the same way as the coefficients given by the Fast Wavelet Transform. In view of that, the control of the regularity (besides its own interest) must be understood as control of the quality on the coefficients. From this point of view, $\mathrm{D}^{\mathrm{PER}}_N$ can be used to detect and describe problematic zones also in the same way as the ones given by the Fast Wavelet Transform.

In view of all this comments, the nature of the problem will decide the strategy to perform. That is, let φ the invariant curve of Equation (3.1). Then, roughly speaking, if an approximation is needed, whatever the cost, use Haar wavelet to get the desired $\mathrm{D}^{\mathrm{PER}}_N$. Surely you will spent many iterates but such effort will be supplied by the fast and exact calculations. On the contrary, make a library of Daubechies wavelets (with $p > 2$) using the Algorithm 2. Then, solve Equation (5.14) for such wavelets. Notice that, such procedure is in contrast with Fourier basis because Fourier is "*unique*". For example, if the question is related to regularity topics along a parameterization, move p along your (pre)-calculated wavelets range to reach the better Pearson coefficient.

Finally, we want to remark that the error at each Newton's step, ϵ, also determines the regularity behavior of φ. Indeed, let $\mathrm{D}^{\mathrm{PER}}_{N,0)}$ be the initial guess for the Newton's method. Then, each correction of the wavelet coefficients can be understood as

$\mathrm{D}^{\mathrm{PER}}_{N,n+1)} = \mathrm{D}^{\mathrm{PER}}_{N,n)} + \epsilon_n$. Hence, setting $\mathcal{E}_n = \sum_{i=1}^{n} \epsilon_i$, we have that

$$\left\|\mathrm{D}^{\mathrm{PER}}_{N,n)}\right\| = \left\|\mathrm{D}^{\mathrm{PER}}_{N,0)} + \mathcal{E}\right\| \leq \left\|\mathrm{D}^{\mathrm{PER}}_{N,0)}\right\| + \|\mathcal{E}\| .$$

In other words, the decay of the wavelet coefficients $\mathrm{D}^{\mathrm{PER}}_{N,n)}$ strongly depends on the one of the initial seed and the ones of the error. And this is precisely the link with the previous chapter. Indeed, the Fas Wavelet Transform plays the role of $\mathrm{D}^{\mathrm{PER}}_{N,0)}$. That is, using the Fast Wavelet Transform we only have a static picture of φ after some iterates. On the other side, the numerical approximation of the invariance equation allows us to capture the zeros of φ in a better way (in the pinched case). Hence, as a conclusion, we remark that the two strategies can be used to save time. Indeed, one can use the Fast Wavelet Transform to explore which Daubechies wavelet better performs the regularity problem. After that, solve the invariance equation with such wavelet up to a given tolerance. That is, seems mandatory to have a good *dictionary* of Daubechies wavelets.

For a future work, the duties must be the solution of the bug of the precision and, also, use the information of the wavelet coefficients to give an estimation of where the attractor is placed. From such information, since the support of the wavelets it is a (dense) cover of $\mathbb{S}^1$ and its images of the ambient space, try to perform other calculations such as the length of φ or its Hausdorff dimension. Also, the Haar wavelet, due its simplicity, can be useful to derive formulas of $\mathcal{E}_\infty$ and, hence, have a "*good choice*" of N. Finally, from the operator point of view, the preconditioning techniques can be understood as an operator. Hence, they can be used to simplify the notion of convergence and perhaps use them to establish Newton Kantorovich hypothesis related to Equations (5.9) and (5.14).

Chapter 6

Testing the Software

The main purpose of this chapter is to put in troubles the software developed from the method derived in Section 5.2 of the previous chapter. Such problems will be faced with the Haar wavelet basis due to the velocity of the computations and the explicit solution of the linear system. That is, along this chapter the vector $\mathrm{D}^{\mathrm{PER}}$ will be constructed in terms of Haar basis. Therefore the functional spaces $\mathscr{B}^s_{2,2}(\mathbb{S}^1)$ and $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ only will allow $s \in (-0.5, 0.5)$ and $s \in (-1, 0)$ respectively.

To do this, firstly we will set our environmental space: the cylinder $[0,1] \times \mathbb{S}^1$. In such cylinder we will apply two kind of skew product whose iterates may produce an overflow problem. One of them will be a multiplicative skew product whereas the other will be in the additive form. Such election is to realize the goodness of the software in both situations.

We want to emphasize that this chapter is mainly experimental. That is, besides of the quick survey of results concerning on the multiplicative skew product systems we do not add more theoretical information. All of such results are from the original paper [AM08]. However, all this numerical experimentation will be useful to refuse (or not) some properties and to test the validity of the software developed up to now. Also, some pictures of the estimation of the Lyapunov exponent on the attractor of these dynamical systems are provided. Note that, since Lyapunov exponent is important to decide either if the matrix is invertible or not, it is important to have such information. Finally, we want to remark that if it is not otherwise said all the numerical computations concerning on the wavelet coefficients are done with $N = 2^{26} = 67108864$ points. That is, we will solve a $2^{26} \times 2^{26} = 2^{52}$ (non) linear system of equations. On the other side, the vector $\mathrm{D}^{\mathrm{PER}}$ is found up to a precision less than 5×10^{-16}. These computations have been done with a single computer with and Intel i7 $-$ 4770K Quad Core processor @ 3.50GHz with 32Gb of RAM DDR3 @ 1867MHz. The figures are obtained by starting from a (pseudo) random pair (θ, x), discarding its first 10^5 iterates and plotting the following 10^5. Finally ω is taken to be the golden mean: $\frac{\sqrt{5}-1}{2}$.

6.1 Two kind of skew products

In view of the results given in Section 5.4 we have performed the same kind of explorations in other types of skew products. That is, we have found the wavelet coefficients

$\mathrm{D}^{\mathrm{PER}}$ of an invariant curve, φ, in order to see how its regularity evolves. The idea of such exercise is try to apply standard continuation techniques on skew products where the nature of the problem causes difficulties on the convergence to $\mathrm{D}^{\mathrm{PER}}$ if the Newton's method is applied verbatim. Indeed, recall that $\lim_{x\to\infty} \tanh(x) = 1$. Therefore, in the case of the skew product given by Equation (3.6) the function $\tanh(x)$ has a crucial role in avoiding overflow problems.

In view of that, our benchmarks will be dynamical systems such that the function $F_{\sigma,\varepsilon}$ will satisfy $\lim_{x\to\infty} F_{\sigma,\varepsilon}(\theta, x) = \pm\infty$. On the other side, the attractor of these skew products is a natural candidate to have suspicious traits of "*fractalization route*" as it can be seen in Figure 6.2 and 6.10.

Unimodal quasiperiodically forced maps

In [AM08] the authors have considered skew products with R_ω as the base map whereas the function in the fiber over θ is taken to be an unimodal map f of the interval $[0, 1]$ onto itself against $g(\theta)$. For us $g(\theta) = (\varepsilon + \sigma_g\theta(1-\theta))$ which is a continuous function from the circle to $[0, 1]$. For all numerical experiments that we have performed the logistic map will be taken as the aforesaid unimodal map (except for one which will be explained later). That is, $F_{\sigma,\varepsilon}(\theta, x) = \sigma_f x(1-x)(\varepsilon + \sigma_g\theta(1-\theta))$.

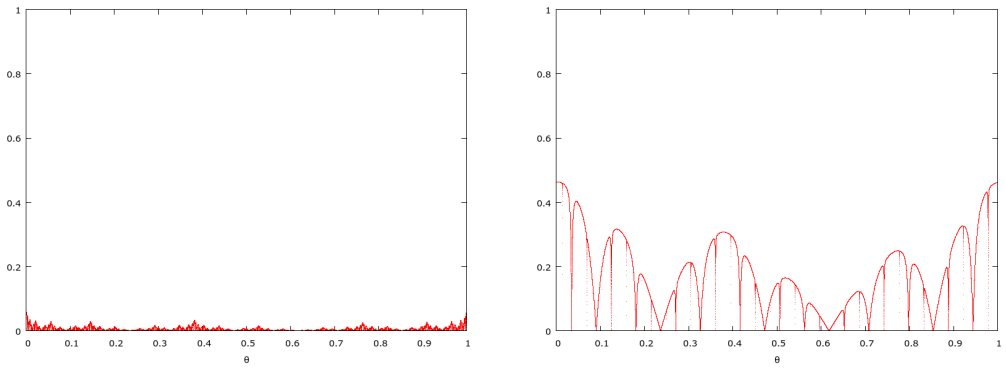


Figure 6.1: The pinched case is considered with $\sigma_f = 4$. On the right hand side $\sigma_g = 1.86$ whereas on the left hand side $\sigma_g = 2.3$.

As in the Keller's situation the circle $x \equiv 0$ is invariant and, therefore, the vertical Lyapunov exponent on it can be explicitly calculated for $\varepsilon = 0$:

$$\kappa(f, g) := \log(f'(0)) + \int_{\mathbb{S}^1} \log|g(\theta)|\, d\theta = \log(f'(0)) + \log(\sigma_g/e^2).$$

Hence, since $\sigma_f = 4$, one of our working situations will start when $\sigma_g > e^2/4$ because then $x \equiv 0$ is a repellor (see [AM08, Corollary 3.3]). Clearly, $\kappa(f, g)$ can be also calcu-

lated for in terms of ε and σ_g. However, following this thread, and being $\sigma_f = 4$, the aforementioned Lyapunov exponent will have a complicated but explicit expression. The knowledge of such $\kappa(f, g)$ will be helpful for us. Finally, it is worth to pointing out that, the unimodality of f implies that some extra assumptions must be added to show that the existence a strange non-chaotic attractor when the system is pinched (i.e $\varepsilon = 0$).

The additive forcing

Up to now, we have been working with models where $x \equiv 0$ was invariant. Such invariance strongly relies on the fact that, by hypothesis, $f(0) = 0$ and $F_{\sigma,\varepsilon}(\theta_n, x_n) = f(x)g(\theta)$. Precisely, using the invariance, for those models we wanted to compute the attractor, φ, so that we can detect the pinching point, $\varepsilon = 0$, in terms of regularity.

Of course, one can consider the case of the additive forcing but we will lose the valuable information of $\kappa(f, g)$. That is, in such situation, we will consider the additive perturbation: $F_{\sigma,\varepsilon}(\theta_n, x_n) = f(x) + g(\theta)$. There is a famous paper [Nis96] where it is described the *fractalization route* seen with support of numerical computations. Currently, there is a strong debate about this route and about the fact that what we get at the end is really a fractal (see Figure6.2). The system under the controversy is given by

$$\begin{pmatrix} \theta_{n+1} \\ x_{n+1} \end{pmatrix} = \mathfrak{F}(\theta_n, x_n) = \begin{pmatrix} R_\omega(\theta_n) \\ F_{\sigma,\varepsilon}(\theta_n, x_n), \end{pmatrix}, \tag{6.1}$$

where $F_{\sigma,\varepsilon}(\theta, x) = \sigma x(1 - x) + \varepsilon \sin(2\pi\theta)$.

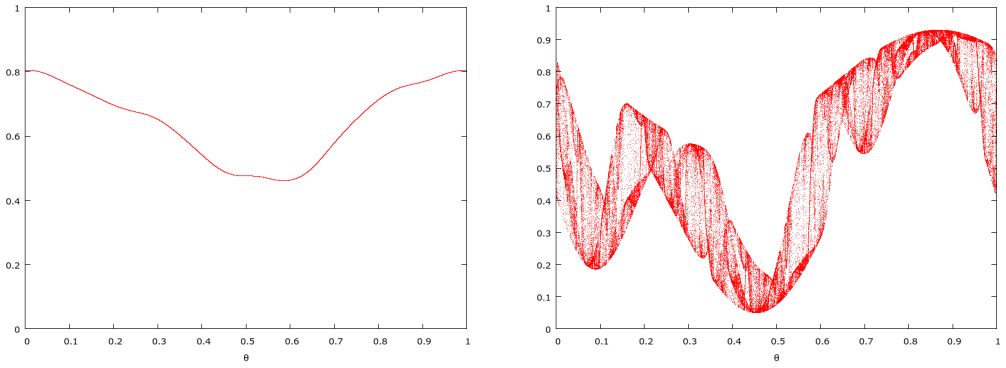


Figure 6.2: The plot of the attractor φ given by the iteration of the System 6.1. On the left hand side $\varepsilon = 0.1$ whereas $\varepsilon = 0.18$ is in the right hand side. Notice the different shapes of both attractors having the same $\sigma = 3.0$.

Using the developed software and the gained experience in similar situations, as

we will see in the following section, we will deal with the Nishikawa-Kaneko model in order to detect changes in the behavior of φ, far beyond the cosmetic ones, if it happens.

6.2 A tour in Alsedà-Misiurewicz System

The authors in [AM08] introduce three operators in order to be able to prove the existence of strange non-chaotic attractors using a contractive property of those operators. In order to fix ideas, let us (re)write the system that we will study

$$\begin{pmatrix}\theta_{n+1}\\ x_{n+1}\end{pmatrix} = \mathfrak{F}(\theta_n, x_n) = \begin{pmatrix} R_\omega(\theta_n)\\ F_{\sigma,\varepsilon}(\theta_n, x_n)\end{pmatrix}, \tag{6.2}$$

where $F_{\sigma,\varepsilon}(\theta, x) = 4x(1-x)(\varepsilon + \sigma_g\theta(1-\theta))$. Now, recall that in Section 3.2 we have introduced the Transfer Operator $\mathfrak{T}(\varphi)(\theta) = f(\varphi(R_\omega^{-1}(\theta)))\cdot g(R_\omega^{-1}(\theta))$. In the same way, the *semi-transfer operator* can be defined as $\mathfrak{S}(\varphi)(\theta) = f(\min\varphi(R_\omega^{-1}(\theta)), c)\cdot g(R_\omega^{-1}(\theta))$, where $c > 0$ is a constant. Using the same ideas as in Equation (3.8) the iteration of $\mathfrak{S}$ converges pointwise towards an upper semi continuous function φ^+. Recall that an upper semi continuous function must have $s = 0$ as a regularity parameter. Finally, the *quarter-transfer operator* is given by $\mathfrak{Q}(\varphi)(\theta) = \min\{\mathfrak{T}(\varphi), \mathfrak{T}(\varphi^+)\}$. Now, we are almost ready to state the equivalent to Keller's Theorem in the present situation. To do this, consider $b := \sup\{x \in (k,1] : \ -f_-(x) < f(x)\}$, where $f_-(x)$ is the left one-sided derivative of f and $\beta = f(\operatorname{ess\,sup}\varphi^+)$.

Theorem 6.2.1 (Theorems 4.3, 4.4, 4.5 and 6.1 [AM08]). *Assume that $0 < \operatorname{ess\,sup}\phi^+ < b$. Then there exists a function $\varphi\colon \mathbb{S}^1 \longrightarrow \mathbb{R}^+$ such that*

(a) $0 \le \varphi \le \phi^+$ and $\varphi \ge \beta'\phi^+$ almost everywhere;

(b) $\mathfrak{T}(\varphi) = \varphi$;

(c) As in the Keller's situation, let $\mathcal{P}$ be the space of all functions not necessarily continuous from $\mathbb{S}^1$ to $\mathbb{R}$: $\mathcal{P}(\mathbb{S}^1, \mathbb{R})$. If $\rho \in \mathcal{P}$ and $\varepsilon\phi^+ \le \rho \le \phi^+$ for some $\varepsilon > 0$ then $\mathfrak{T}^n(\rho)$ converges to φ almost everywhere.

(d) φ is a measurable and positive almost everywhere function;

(e) For almost every $\theta \in \mathbb{S}^1$ and for all $x \in (0,1)$ the trajectory of the pair (θ, x) converges exponentially fast to the graph of φ of falls into $\mathbb{S}^1 \times \{1\}$ and then stays in $\mathbb{S}^1 \times \{0\}$. In particular, for almost every $\theta \in \mathbb{S}^1$ and all but countable number of $x \in (0,1)$ the trajectory of the pair (θ, x) converges exponentially fast to the graph of φ;

(f) If $g \neq 1$ almost everywhere then for almost every $\theta \in \mathbb{S}^1$ and all $x \in (0,1)$ the trajectory of the pair (θ, x) converges exponentially fast to the graph of φ;

(g) If $f(\phi^+)$ then or almost every $\theta \in \mathbb{S}^1$ and all but countable number of $x \in (0,1)$ the trajectory of the pair (θ, x) converges exponentially fast to the graph of φ.

(h) If $\kappa(f,g) > 0$ then $\lambda_\varphi < 0$.

As, in the case of Keller's Theorem, we are interested in the case of $\kappa(f,g) > 0$. However, there are some gaps in the cases that we will consider. Indeed, when $\varepsilon = 0$ the following Lemma is the explanation of some problematic values of σ_g in terms of the hypothesis of the existence of a Strange non-Chaotic Attractor using the above Theorem.

Lemma 6.2.2 (Lemma 7.1 [AM08]). *Consider $g = \sigma_g(\theta(1-\theta)$ to be such that $\sigma_g > e^2/2$ and let f be such its turning point $1/2$. Then $\operatorname{ess\,sup} \varphi^+ > 1/2$.*

Hence, we will move $\sigma_g \in [e^2/4, e^2/2]$. However, to take an idea of where are such places, in the case of $\varepsilon > 0$, we have done the following estimation of the Lyapunov exponents, λ_φ, using the Equation (3.4) (see Figure 6.3).

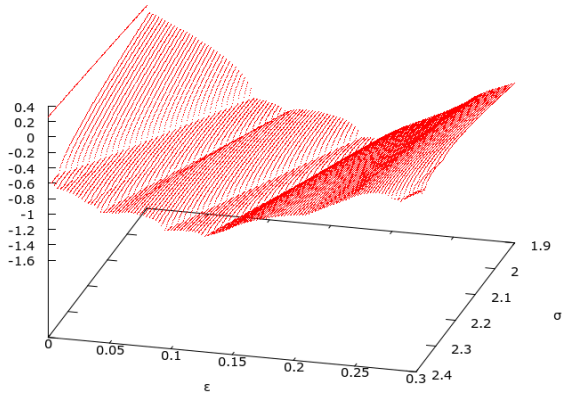


Figure 6.3: Some estimation of λ_φ in terms of σ_g and ε. Because of formatting reasons we have only plotted $\sigma_g \in [1.9, 2.4]$ and $\varepsilon \in [0, 0.3]$. We want to notice that if a piece of the surface which is not plotted is needed, then we will plot it.

In the following sections we will use such estimation, for the whole range, and the above results to make some numerical experiments, conjectures and gain an important experience.

6.2.1 Parameterization

Our first purpose is to see how the wavelet coefficients $\mathrm{D}^{\mathrm{PER}}$ moves along the parameter space for the skew product given by Equation (6.2). Such face is reinforced by the results derived from Equation (5.16). To do this, first of all we have used the following parameterization of ε in terms of σ_g:

(6.3)
$$\varepsilon(\sigma_g) := \begin{cases} -\frac{3323299}{500000}\sigma_g^2 + \frac{60871083}{2500000}\sigma_g - \frac{34284691}{1562500} & \text{if } \sigma_g \in [1.6, 2.0), \\ (\sigma_g - 2.8)^8 + 0.0005 & \text{if } \sigma_g \in [2.0, 3.49), \\ \frac{366298374428641}{861000000000000}\sigma_g^2 - \frac{7013481614144179}{2152500000000000}\sigma_g + \frac{179131268100176321}{287000000000000000} & \text{if } \sigma_g \in [3.49, 3.9], \\ 0 & \text{otherwise.} \end{cases}$$

To fix notation, let us recall that we aim to obtain the coefficients $d_{-j}^{\mathrm{PER}}[n]$ which will be collected in the vector $\mathrm{D}^{\mathrm{PER}}$. Such vector will be found, using the Newton's method, up to a given tolerance (see Chapter 5 for further information). Since now we will move along the parameter space, we will denote $\mathrm{D}_{N,\varepsilon}^{\mathrm{PER}}$ the N-dimensional vector of wavelet coefficients respect the parameter ε. Hence, in the present situation, we are looking for $\mathrm{D}_{N,\varepsilon(\sigma_g)}^{\mathrm{PER}}$ where $\varepsilon(\sigma_g)$ is given by Equation (6.3).

Notice that the values $\sigma_g \in [2.0, 3.49)$ gives us small values of ε. This is not done for free. Indeed, using a symbolic calculator we have calculated $\kappa(f, g)$ over such parameterization to ensure that $\kappa(f, g) > 0$ (see Figure 6.4). Beyond this exercise there is another masked goal: find good regions to start continuation methods for the following sections. Now, the reason for having this flat region becomes clear. All the coefficients of $\mathrm{D}_{N,\varepsilon(\sigma_g)}^{\mathrm{PER}}$ corresponds to a nonpinched region with "*non convergence problems*". Hence, they can be used as a initial seeds for the continuation methods in regions where the attractor is closer to zero or where the Keller's techniques does not apply but there is an invariant curve, even for $\varepsilon = 0$. This will be the main topic of the following section.

In terms of how the coefficients $\mathrm{D}_{N,\varepsilon(\sigma_g)}^{\mathrm{PER}}$ evolves along $\sigma_g \in [1.6, 3.9]$, let us introduce in the same way as in Equation (1.9) and Remark 1.2.10:

(6.4)
$$\tilde{s} = \tau - 1 + \frac{1}{p},$$

where $0 < p \leq \infty$ and τ is obtained by performing a standard linear regression between j and $s_j = \sup_n |\left\langle \varphi, \psi_{-j,n}^{\mathrm{PER}} \right\rangle|$. That is, we are looking at how the Haar wavelet coefficients of φ vanish. In other words, we are taking $\tilde{s}$ as the regularity parameter without the constraints on s inherited by Haar.

Once $\tilde{s}$ has been introduced, first of all let us notice that the values of ε closer to zero must be appreciated for our vanishing parameter $\tilde{s}$ because we are approximating to the pinched case. That is, φ becomes to have a strange behavior since it must converge to an upper semi continuous function when $\varepsilon = 0$. In Figure 6.5 we have displayed how φ evolves along $\varepsilon(\sigma_g)$.

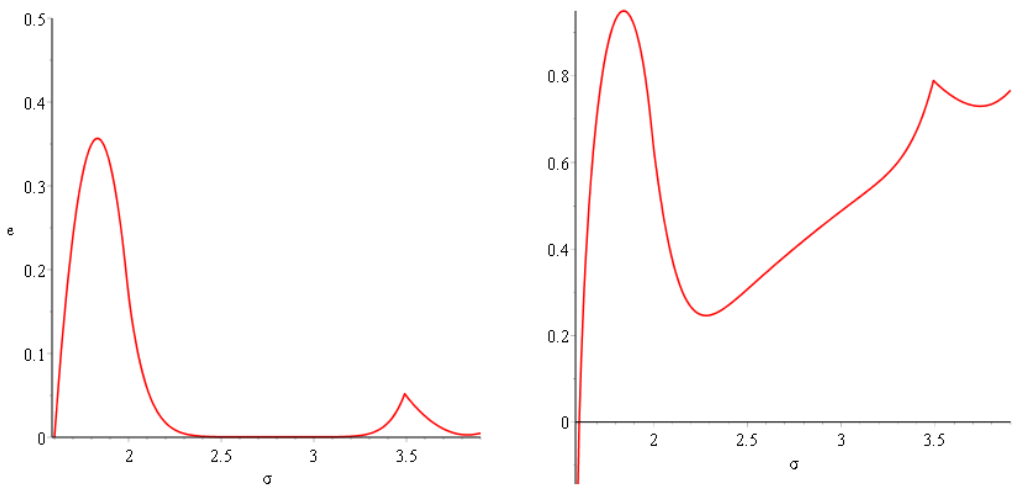


Figure 6.4: The plot the parameterization $\varepsilon(\sigma_g)$ and the vertical Lyapunov exponent of $x \equiv 0$ with such parameterization. Recall that, it is mandatory to have a repellor on $x \equiv 0$ in order to have a finite attractor.

Moving to the regularity evolution we want to make some comments. First of all, there are values of σ_g where the theoretical results of [AM08] does not apply as we will see later. However, the transition to the strangeness is still visible in the same way as in Section 5.4. Moreover, as σ_g gets closer to 4 then φ becomes weird and, again, the graph of regularity parameter still marks such tendency as it can be seen at the last part of Figures 6.5 and 6.6.

As we have already said $\mathrm{D}^{\mathrm{PER}}_{N,\varepsilon(\sigma_g)}$ are in terms of the Haar wavelet basis. Hence, one can also look for $\varphi \in \mathscr{B}^s_{2,2}(\mathbb{S}^1)$ using Corollary 1.2.16 since looking at the pictures in Figure 6.5 it seems that weak derivatives can appear. We only mention that such experiment produces a similar picture as Figure 6.6 (for the allowed ranges of s). However, there are values of σ_g such that $\varphi \notin \mathscr{B}^s_{2,2}(\mathbb{S}^1)$ for $s \in (-0.5, 0.5)$. This is not a surprisingly fact. Indeed, for a big enough values of ε the invariant curve φ is sufficiently smooth as it can be seen from Theorem 3.3.1 and hence peaks can not appear. Moreover, for values of $\sigma_g \leq e^2/4$ when $\varepsilon = 0$ then $\varphi \equiv 0$. Hence, in a naive way, we start from a constant which has infinitely many derivatives and, with $\varepsilon > 0$, the attractor φ is regular enough.

In view of such good convergence zone, we will use such results as a region for initial seeds for continuation methods. However, the loss of regularity seems to be a good indicator to see problematic situations, besides the pinched case, such as the change of sign of the Lyapunov exponent on the invariant curve. In other words, seems that wavelet coefficients also feel the changes on λ_φ. We will return to such topic later

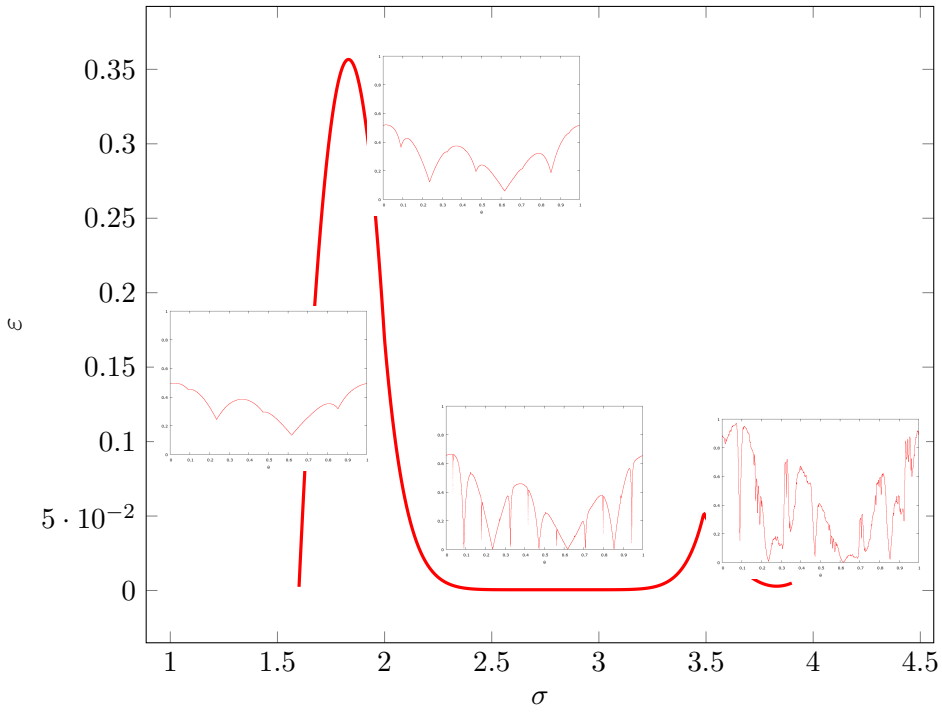


Figure 6.5: The range of ε given by the parameterization $\varepsilon(\sigma_g)$ with $\sigma_g \in [1.6, 3.9]$. there are included several instances of the attractor. Of course, in the flat region φ starts to loose regularity properties.

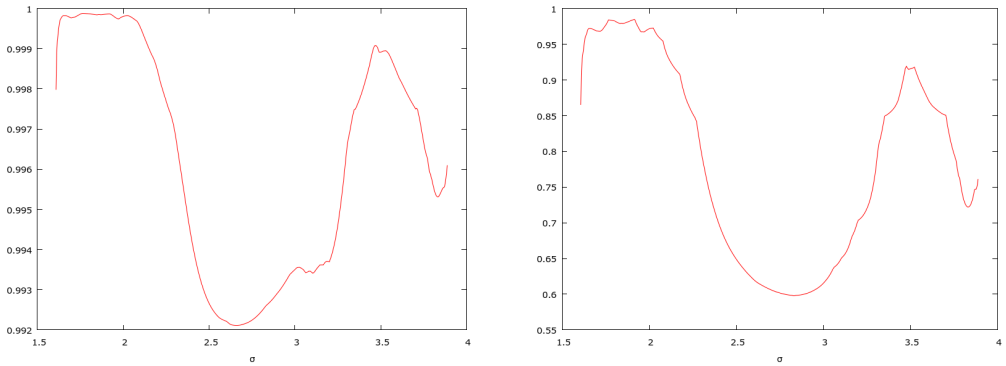


Figure 6.6: On the left hand side, it is plotted the Pearson correlation coefficient along $\varepsilon(\sigma_g)$. On the right hand side there is the regularity parameter $\tilde{s}$ of φ along such parameterization. The flat region of $\varepsilon(\sigma_g)$ can be clearly identified because $\tilde{s}$ must go to 0 because of the pinching condition.

on Section 6.2.3.

6.2.2 Continuation towards the pinched case

In view of the above convergence results along the parameterization of ε by means of $\varepsilon(\sigma)$ we have done a continuation strategy respect the parameter ε. The reason is that the initial seed, using the trapezoidal rule, it is not enough to avoid the basin of attraction of $-\infty$ when $\varepsilon = 0$. In other words, at each reconstruction step, using $\mathrm{D}^{\mathrm{PER}}_{N,0)}$ from the the trapezoidal rule, each $\mathrm{D}^{\mathrm{PER}}_{N,k)}$ is worse than the previous one. Hence the verbatim application of the Newton's method fails to converge and, therefore, we need to perform another strategy. To this end, recall that $\mathrm{D}^{\mathrm{PER}}_{N,k)}$ are obtained, using Newton's method, from Equation (5.1) and (5.6). Let X be such solution. In our situation X strongly depends on σ and ε: $\mathrm{X}_{\sigma,\varepsilon}$. The continuation strategy is based on the fact that, being $h > 0$ small enough, $\mathrm{X}_{\sigma,\varepsilon} \sim \mathrm{X}_{\sigma,\varepsilon+h}$. Hence, one can construct a sequence of $\mathrm{X}_{\sigma,\varepsilon+kh}, k = 0, \ldots, n$ to try to converge to the desired $\mathrm{X}_{\sigma,\widetilde{\varepsilon}}$, where $\widetilde{\varepsilon} = \varepsilon + nh$. Of course, the same strategy can be done respect the parameter σ also.

To do this, we have used heavily Equation (3.4) to have an estimation of the Lyapunov exponent on the invariant curve, λ_φ, in terms of the parameters σ_g and ε. Despite of having a picture of these λ_φ's (see for example Figure 6.7), with such information and the one derived from $\varepsilon(\sigma_g)$ we have several good continuation regions towards the desired ε. Again, we have applied the principle of if the Newton's method converges, it converges.

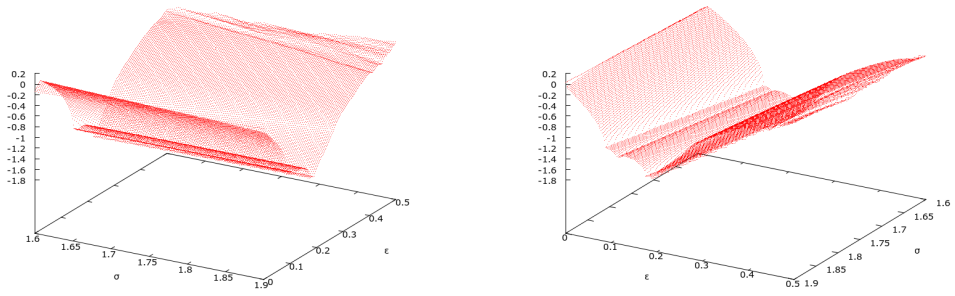


Figure 6.7: Two different perspectives of the estimation of the Lyapunov exponent of the logistic map with $\sigma_f = 4$ forced with a logistic map with $\sigma_g \in [1.6, 1.9]$ and $\varepsilon \in [0, 0.5]$.

We have started our experimentation with two attractors such that they are closer to $\varphi \equiv 0$ but with σ_g slightly larger than $e^2/4$ (see Figure 6.8). Such motivation is twofold. In one hand, to be under the assumptions of Lemma 6.2.2. On the other side, try to bug the convergence towards the wavelet coefficients $\mathrm{D}^{\mathrm{PER}}_N$. It is important to notice that

the values are small than 0.05 which leads to a wavelet coefficients vector $\mathrm{D}_N^{\mathrm{PER}} \equiv 0$ if one starts with $\varepsilon = 0$ directly. This is not strange fact. Indeed, $e^2/4 = 1.847264024732\ldots$ and hence using $\sigma_g = 1.85$ then we are closer to have $x \equiv 0$ as the attractor. Also, the change of sign of λ_φ is near and, hence, the matrix $(\mathrm{Id} - \widetilde{\Delta}_N)$ turns into a non invertible one. On the contrary, using a positive ε the continuation method converges to a $\mathrm{D}_N^{\mathrm{PER}}$ such that $\varphi > 0$. Since we are in the Keller's situation this is a good notice because φ we can recover a strange non-chaotic attractor whose values are small.

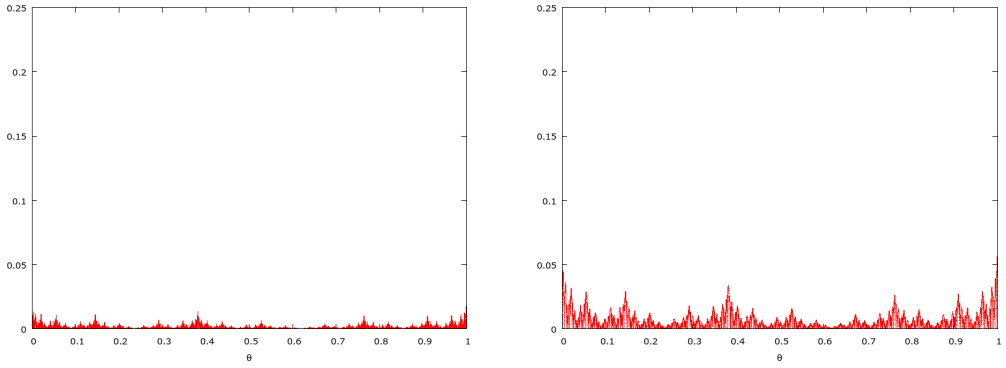


Figure 6.8: Two different values of σ_g and the pinched case is considered. On the right hand side $\sigma_g = 1.85$ whereas on the left hand side $\sigma_g = 1.86$.

In view of such results we have made the same kind of experiment for other values of σ_g (for example for the φ of right hand side Figure 6.1) with similar results: good convergence velocity and the "*real*" regularity parameter s, with a Pearson coefficient greater than 0.99, can be assessed to be equal to zero. Clearly, this zero regularity is what we have expected because of Lemma 6.2.2 and Theorem 6.2.1. That is, φ is a strange non-chaotic attractor and, therefore, it must have zero regularity because of the upper semi continuity of φ.

Of course, we are not saying that we can guarantee the convergence everywhere of the continuation strategies. In fact, there are situations where the convergence to a pinched case is heavy to get. For example, we have started with $\sigma_g = 3.8 > e^2/2$ and $\varepsilon = 0.0001$. The continuation method needs about 7000 adaptive steps to converge towards $\sigma_g = 3.8$ and $\varepsilon = 0$. For adaptive we mean that h can vary, in terms of how convergence evolves, along the continuation performance (see Figure 6.9). Despite of this convergence problems we have promising news respect the regularity. Indeed, for the pinched case and $p = \infty$ we have that $\tilde{s} = 0.006858$. That is, it seems safe to consider that $\varphi \in \mathscr{B}^0_{\infty,\infty}(\mathbb{S}^1)$. On the contrary, for the nonpinched case $\tilde{s} = 1.2368362$ and hence it is out of range. In view of that we can make two comments. We also have a "*strange*" function even in the case of $\sigma_g > e^2/2$. That is we still have a strange

function even in the case where the reduction to the Keller's techniques does not hold. And, also convergence towards the expected regularity space.

The second comment concerns to the shape and the difference between the regularity parameter. As it can be seen at Figure 6.9 the shape of both attractors is almost the same. However this is only a picture. That is, the solution of the invariance equation is able to capture, in a better way, the (dense set of) zeros of φ in the pinched case. Hence, when $\varepsilon = 0$ the coefficients $\mathrm{D}_N^{\mathrm{PER}}$ are more accurate than the ones of the Fast Wavelet Transform. On the other side when $\varepsilon > 0$, the wavelet coefficients are almost the same as the ones of the Fast Wavelet Transform in agree with Lemma 2.3.2.

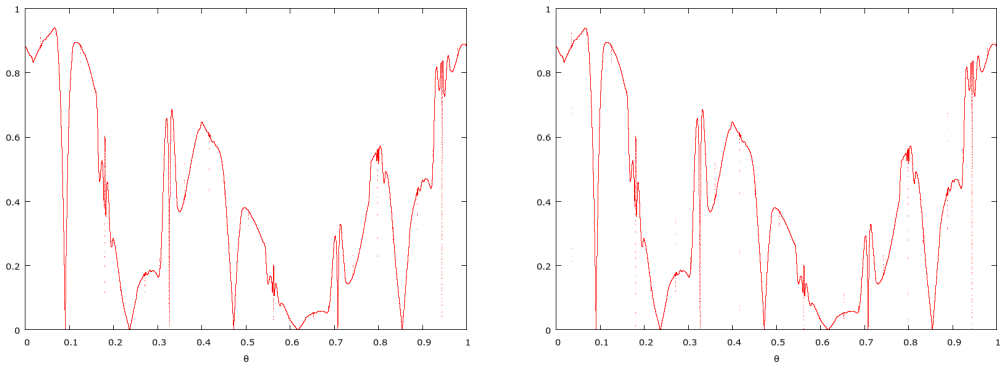


Figure 6.9: Being $\sigma_g = 3.8$ fixed, we have used $\varepsilon = 0.0001$, the left hand side picture, to use $\mathrm{D}_N^{\mathrm{PER}}$ as an initial seed to go to the pinched case. It is worth to pointing out the close shapes of the two pictures. In the pinched case $\tilde{s} = 0.006858$ (with a Pearson coefficient equal to 0.998631).

With those comments we mean that, perhaps, a better strategy in such situation is to consider the invariance equation with the operators $\mathfrak{S}$ and $\mathfrak{Q}$ instead of $\mathfrak{T}$. Moreover, perhaps the rigorous study of the continuation method can be used to derive in a more feasible way an analogous of Theorem 6.2.1 for $\sigma_g > e^2/2$. Or, at least, provide proved safety regions which guarantee a fast convergence to the wavelet coefficients $\mathrm{D}_N^{\mathrm{PER}}$ besides the "*trial-and-error*" way.

6.2.3 Fractalization suspected

As we have already said at the end of the Section 6.2.1 we want to use the idea of continuation methods with situations where the Lyapunov exponent is negative but closer to zero. Dynamically, we will have an attractor which becomes something "*neutral*". In terms of Newton's method this means that we will have problems in the solution of the (non) linear system of equations. In terms of the graph of φ the Lyapunov exponent

seems to mean more freedom to stretch and fold the invariant object (see e.g Figure6.10).

Such situation is related to the one of the right hand side picture of Figure 6.2 and it is the origin of the controversy. In order to gain experience to try to provide some information in the Nishikawa-Kaneko model (and also for the particular instance of the Alsedà-Misiurewicz system) we have made some continuation experiments. Since we are not interested in the regularity[1] results, we have diminished the N to $2^{22} = 4194304$. We remark that the regularity topics are related, for example, with the Hausdorff dimension of the object. Hence, a good estimate of the regularity gives a good estimate of the Hausdorff dimension. In our performance, the regularity assessment it is done by means of a linear model. Hence we need to have as much samples as it is possible. In such situation, $N = 2^{22}$, we have not considered 22 samples a good framework to do the regression.

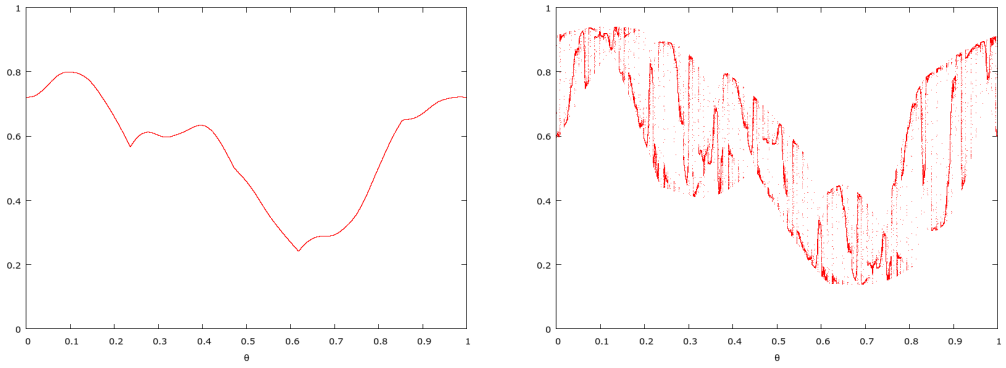


Figure 6.10: Two different values of σ_g and the nonpinched case is considered. On the right hand side $\sigma_g = 2.0$ and $\varepsilon = 0.3$ whereas on the left hand side $\sigma_g = 2.0$ and $\varepsilon = 0.44$. This situation looks like Figure 6.2 or the right hand side picture of Figure 6.11.

Our initial plan was focused on try to be, as much as possible, closer to $\varepsilon = 0.44$. However, our tries were a nightmare because the instability inherited from the system. We can only arrive at $\varepsilon = 0.425042$ (starting on $\varepsilon = 0.20424$) and, looking at the result's files, it is appreciated the bad convergence. Indeed, Newton's method needs more than 300 iterates to reach the solution using a initial seed closer solution. Far beyond the naive strategy performed, this means, that we are closer to non convergence zones. Such regions can arise from the contractive character of the Newton's method. Hence, the problems to reach the desired ε means that either the continuation or Newton's method are close to fail.

[1] However, for sake of completeness let us say that the software returns a value of $\tilde{s}$ such that φ is suspicious of continuity.

Nevertheless, from the data files we have learned the following good region of convergence. Indeed, let ϵ_{Inv} be the error derived from the invariance equation and ϵ_{DPER} how close to zero is the "*current*" Newton iteration. Then, if it is possible to stabilize $1 \times 10^{-6} \leq \epsilon_{\text{Inv}} \leq 1 \times 10^{-8}$ and $1 \times 10^{-3} \leq \epsilon_{\text{DPER}} \leq 1 \times 10^{-4}$ then Newton's method converge (and also the continuation strategy). On the other side, if either $\epsilon_{\text{Inv}} \geq 1 \times 10^{-3}$ and/or $\epsilon_{\text{DPER}} \geq 1 \times 10^{-1}$ the iteration becomes highly unstable and, mainly, does not converge.

Clearly, such regions strongly depends on the values of σ_g and ε (and hence the Lyapunov exponent). This means that such iteration process and the continuation operator induced defines neighborhoods (in fact they are tubes) of convergence $\mathcal{U}_{\sigma_g,\varepsilon}$ towards the desired ε. As a conclusion, these $\mathcal{U}_{\sigma_g,\varepsilon}$'s perhaps can be used to (im)prove the convergence theorems such as Theorem 6.2.1. Also the operators can be helpful to understand such problematic zone.

6.3 Skew Products as bugs

In view of the above results and drawing to close this chapter, let us explain two numerical experiments more. They must be understood as an examples where the developed software does not work (up to now). Both of them, displayed on Figure 6.11 ,may represent two kind of different problems.

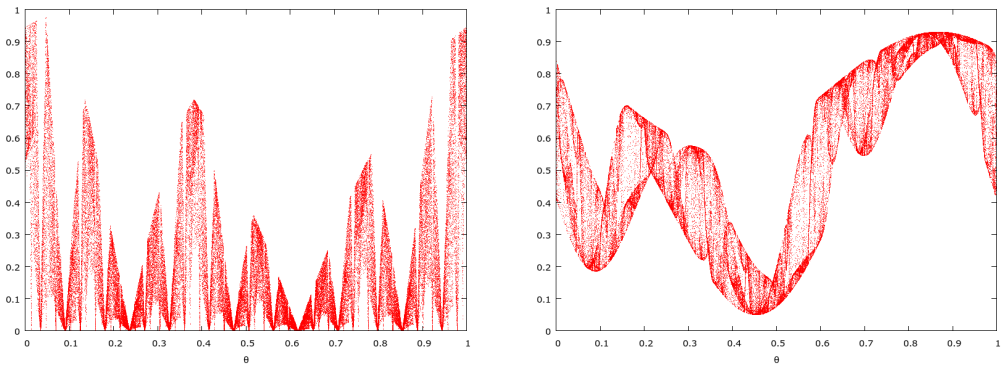


Figure 6.11: On the left picture it is shown the attractor of the tent map forced by a logistic map with $\sigma = 4.0$ and $\varepsilon = 0.0$ [AM08]. On the right hand side it is displayed the Nishikawa-Kaneko model ($\sigma = 3.0$ and $\varepsilon = 0.18$) [Nis96].

As we have already said, there is a controversy of how weird is the attractor of the Nishikawa-Kaneko System. Several techniques can be applied such as wise magnifications among others. In view of that, we have made the following exploration. Being $\sigma = 3.0$ and again, using Equation (3.4), we have taken the coefficients $\mathrm{D}^{\text{PER}}_{N,\varepsilon_0}$ with

$\varepsilon_0 = 0.14$ to be the initial seed of the continuation method. Our goal was $\varepsilon = 0.18$. However, we do not reach our desired value. As a matter of fact, we do not reach the desired $\mathrm{D}^{\mathrm{PER}}$ far beyond the "*good values*" making impossible the use of $\tilde{s}$ to classify it. This in contrast to the previous section. Assuming [2] that $\varphi \in \mathscr{L}^2(\mathbb{S}^1)$ then we can approximate it using the Haar wavelet basis. Hence, this is suspicious that the operator induced on the continuation - iteration in both situations are different which it is not surprisingly because one has multiplicative forcing whereas the other is additive. And in the multiplicative case, the concavity seems to play an important role (see [AM08]) to guarantee the convergence.

On the other side we have taken $F_{\sigma,\varepsilon}(\theta, x) = (1 - |2x - 1|)(\varepsilon + \sigma_g\theta(1 - \theta))$, $\sigma = 4.0$ and $\varepsilon = 0.0$. In such situation we can not get a $\mathrm{D}^{\mathrm{PER}}_{N,0)}$ such that ϵ_{Inv} and $\epsilon_{\mathrm{D}^{\mathrm{PER}}}$ are in the safety region analogous to the one depicted in Section 6.2.3. Even with naive continuation methods. However, the situation is different. Indeed, as it is pointed out in [AM08] φ it is not a usual function and hence the spaces $\mathscr{B}^s_{\infty,\infty}(\mathbb{S}^1)$ with $s \leq 0$ must be considered. Recall that, the negative spaces can be understood as the dual of the positive ones. However, the (numerical) calculations seems to be adapted to such language but we still have theorems of regularity in such cases as, for instance, Theorem 1.2.8 and 1.2.15.

However, in both cases seems that the operator language, specially in the right preconditioned Haar's case, is helpful for understanding the convergence conditions or, perhaps, determine some obstruction problems.

6.4 The method (updated)

We will finish this chapter with some comment concerning on the numerical experiments that we have performed along of it. First of all, in Section 5.3 we have sketched the method to calculate $\mathrm{D}^{\mathrm{PER}}$ with Daubechies wavelet for $p \geq 1$. But, in view of the explanations along this chapter and the previous one we want to remark two things. The first one concerns the initial guess:

Step 1: Initial guess. Compute, using the results of Section 2.2, the matrices Ψ^{PER}_N and $\tilde{\Psi}^{\mathrm{PER}}_N$. Use them to set $\frac{1}{N}{\Psi^{\mathrm{PER}}_N}^{\top}\Phi_N = \mathrm{D}^{\mathrm{PER}}_{N,0)}$ as an initial guess for the recurrence (5.2). In order to get a sufficiently good approximation of φ, up to a given tolerance, use Equation (3.4). Also, moving the parameters use such formula to have a perspective of good continuation zones. Combine the two strategies to get the best initial seed $\mathrm{D}^{\mathrm{PER}}_{N,0)}$ possible.

The second one is about the wavelets. As we have already said, you can use the Daubechies wavelet that you want. Hence, in the same way as it is pointed out at the final lines of Chapter 5 it seems a good idea to make a library of such wavelets. In

[2] Perhaps the assumption is false.

Section 1.1.2 we have justified its use. However, perhaps other wavelets such as Symmlets (see [Mal98, Section 7.2.3]) converts the calculation of $\tilde{\Psi}^{\mathrm{PER}}$ in something more closer to the Haar's case due to its quasi symmetry. Or even relaxing the hypothesis of the orthogonality of wavelets to ones that they are only biorthogonals (see [Mal98, Section 7.4]), then the calculations becomes more easy. In other words, on the Fourier world there is only one basis whereas in the wavelets world there are an infinite varieties of them. The problem is to find the *best* one blindly. To avoid this problem the use of the Fast Wavelet Transform in all of its forms (see e.g [Mal98, Chapter 8] can be used to choose a good candidate of wavelet, such that it verifies Theorem 1.2.8, to solve the invariance equation.

Appendix

In Chapter 5 we have introduced a matrix which, under certain circumstances can be useful. We do not use it. However we display its properties in the following Lemma. We must say that the proof of these properties strongly relies on Lemma 2.2.3: fixed a block j then all the columns corresponding to such block are the same translated by a power of two.

Lemma 6.4.1. *Let $\psi(x)$ be a Daubechies wavelet with $p > 1$ vanishing moments and ψ_l^{PER} its orthonormal basis of $\mathbb{S}^1$. Being fixed $N = 2^J$, with $J > 0$, define the matrix $\mathsf{Id} = \frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$. Then*

(a) The first row of the matrix $\frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$ is the sum of the columns of $\tilde{\Psi}_N^{\mathrm{PER}}$. Moreover, for each level $j = 0, \ldots, J$ we only have to calculate one value.

(b) The second row of the matrix $\frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$ verifies the identity for $k = 1, \ldots, J-1$ $\mathsf{Id}_{1,m} = -\mathsf{Id}_{1,m\widetilde{+}2^l}$, where $m\widetilde{+}2^l = m + 2^l \pmod{2^k}$ with $l = k-1$.

(c) The block of the matrix $\frac{1}{N}\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}}$ j is divided in two differentiated blocks:

(c)-1. The corresponding block of the same j which gives a symmetric circulant matrix. Hence, we only have to calculate 2^j different values.

(c)-2. The rest of the block which corresponds to $k = j+1, \ldots, J-1$. Each of such blocks has different 2^k values. Also, $\mathsf{Id}_{n,m} = \mathsf{Id}_{n+1,m\widetilde{+}2}$ for $n = 0, \ldots, 2^j - 1$ and $m\widetilde{+}2 = m + 2^l \pmod{2^k}$, where $l = 1, \ldots, J-1-j$.

Proof. We need to show the four properties. However, we claim that the item (b) and (c2) are the same. Hence, we will only prove the first property and the corresponding to the level j.

(a) We need to calculate the first row of the matrix:

$$\frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{0,m} = \frac{1}{N}\sum_{i=0}^{N-1}\psi_0^{\mathrm{PER}}(R_\omega(\theta_i))\psi_m^{\mathrm{PER}}(R_\omega(\theta_i))$$

for $m = 0, \ldots, N-1$. But, by definition $\psi_0^{\mathrm{PER}}(R_\omega(\theta)) = 1$ for all $\theta \in \mathbb{S}^1$. Hence the conclusion easily follows since by Lemma 2.2.3 for each $j = \lfloor \log_2(m) \rfloor$ and $n = m - 2^j$ the values of $\psi_{j,n}^{\mathrm{PER}}$ are the same. In view of that, the two first values are very easy to calculate. The first one, concretely, is $\frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{0,0} = 1$. The second one, taking into account Lemma 2.2.1, is $\frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{0,1} = 0$.

(b) Let us fix a level j. The proof will be split in two parts.

(j) Clearly, when $j = 0$, the claim is true. Hence, let $j \geq 1$. We are focused in the block $\frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{\tilde{n},\tilde{m}}$, where the sub-indexes $\tilde{n}, \tilde{m} = 2^j, 2^j+1, \ldots, 2^{j+1}-1$. To simplify the notation, set $n = 2^j - \tilde{n}$ and $m = 2^j - \tilde{m}$. The symmetry is clear because

$$\begin{aligned}\frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{n,m} &= \frac{1}{N}\sum_{i=0}^{N-1}\psi_{-j,n}^{\mathrm{PER}}(R_\omega(\theta_i))\psi_{-j,m}^{\mathrm{PER}}(R_\omega(\theta_i)) \\ &= \frac{1}{N}\sum_{i=0}^{N-1}\psi_{-j,m}^{\mathrm{PER}}(R_\omega(\theta_i))\psi_{-j,n}^{\mathrm{PER}}(R_\omega(\theta_i)) \\ &= \frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})_{m,n}.\end{aligned}$$

Also, following the thread of simplify the notation, we will look at the upper diagonal of the considered block ($n \leq m$), drop the normalization coefficient ($1/N$) and setting $\mathsf{Id} = \frac{1}{N}(\tilde{\Psi}_N^{\mathrm{PER}\top}\tilde{\Psi}_N^{\mathrm{PER}})$. Having introduced the notation, firstly we will show that the first row of such block has a particular structure. To do this, let $k = 0, \ldots, 2^{j-1}$. By Lemma 2.2.3:

$$\psi_{-j,0}^{\mathrm{PER}}(R_\omega(\theta_i)) = \psi_{-j,2^{j-1}\pm k}^{\mathrm{PER}}(R_\omega(\theta_{\widetilde{i+\kappa)}}), \tag{6.5}$$

where $\widetilde{i+\kappa)} = i + \kappa \pmod N$ and $\kappa = (2^{j-1} \pm k)2^{J-j}$. We claim that $\mathsf{Id}_{0,2^{j-1}-k} = \mathsf{Id}_{0,2^{j-1}+k}$ for $k = 0, \ldots, 2^{j-1}$. The assertion, since $\mathsf{Id}_{0,2^j}$ does not exist, is clear for $k = 0$ and $k = 2^{j-1}$. For the rest, we will use the above transformation:

$$\begin{aligned}\mathsf{Id}_{0,2^{j-1}-k} &= \sum_{i=0}^{N-1}\psi_{-j,0}^{\mathrm{PER}}(R_\omega(\theta_i))\psi_{-j,2^{j-1}-k}^{\mathrm{PER}}(R_\omega(\theta_i)) \\ &= \sum_{i=0}^{N-1}\psi_{-j,0}^{\mathrm{PER}}(R_\omega(\theta_i))\psi_{-j,0}^{\mathrm{PER}}(R_\omega(\theta_{\widetilde{i+\kappa)}})) \\ &= \sum_{i=0}^{N-1}\psi_{-j,0}^{\mathrm{PER}}(R_\omega(\theta_i))\psi_{-j,2^{j-1}+k}^{\mathrm{PER}}(R_\omega(\theta_{\widetilde{\iota+\kappa)}})) \\ &= \sum_{i=0}^{N-1}\psi_{-j,0}^{\mathrm{PER}}(R_\omega(\theta_i))\psi_{-j,2^{j-1}+k}^{\mathrm{PER}}(R_\omega(\theta_i)) \\ &= \mathsf{Id}_{0,2^{j-1}+k},\end{aligned}$$

where $\widetilde{\iota+\kappa} = i + \kappa + \kappa \pmod N = i + 2(2^{j-1}+k)2^{J-j} \pmod N = i + 2^J \pmod N = i+N \pmod N$. That is, $\widetilde{\iota+\kappa)} = i$. Now it only remains to prove

that $\mathsf{Id}_{n,m} = \mathsf{Id}_{n+1,m+1}$. Again by Lemma 2.2.3:

$$\begin{aligned}
\mathsf{Id}_{n+1,m+1} &= \sum_{i=0}^{N-1} \psi^{\mathrm{PER}}_{-j,n+1}(R_\omega(\theta_i))\psi^{\mathrm{PER}}_{-j,m+1}(R_\omega(\theta_i)) \\
&= \sum_{i=0}^{N-1} \psi^{\mathrm{PER}}_{-j,n}(R_\omega(\theta_{\widetilde{i+\kappa)}}))\psi^{\mathrm{PER}}_{-j,m}(R_\omega(\theta_{\widetilde{i+\kappa)}})) \\
&= \sum_{\iota=0}^{N-1} \psi^{\mathrm{PER}}_{-j,n}(R_\omega(\theta_\iota))\psi^{\mathrm{PER}}_{-j,m}(R_\omega(\theta_\iota)) \\
&= \mathsf{Id}_{n,m},
\end{aligned}$$

where $\kappa = 2^{J-j}$ and $\iota = \widetilde{i+\kappa)} = i + \kappa \pmod N$.

($k > j$) Let us start by the case $k = j+1$. Hence, we will be interested in a block of 2^j rows and 2^{j+1} columns. Again the indexes must be understood with the corresponding translation by a power of two: $\tilde{n} = 2^j + n$ for the rows and $\tilde{m} = 2^{j+1} + m$ for the columns. As in the previous case the first row of such block will be the "*generator*" for the rest of the rows: let

$$\mathsf{Id}_{0,m} = \sum_{i=0}^{N-1} \psi^{\mathrm{PER}}_{-j,0}(R_\omega(\theta_i))\psi^{\mathrm{PER}}_{-k,m}(R_\omega(\theta_i))$$

be the first row of the block of size $2^j \times 2^{j+1}$. We have, for $m > n$, by Lemma 2.2.3:

$$\psi^{\mathrm{PER}}_{-j,n}(R_\omega(\theta_i)) = \psi^{\mathrm{PER}}_{-j,m}(R_\omega(\theta_{\widetilde{i+\kappa)}}),$$

where $\widetilde{i+\kappa)} = i + \kappa \pmod N$ and $\kappa = (m-n)2^{J-j}$. Hence, we need to show that $\mathsf{Id}_{l,m} = \mathsf{Id}_{l+1,m+2}$ for $l = 0, \ldots, 2^j - 1$. To do this notice that:

$$\begin{aligned}
\mathsf{Id}_{1,m+2} &= \sum_{i=0}^{N-1} \psi^{\mathrm{PER}}_{-j,1}(R_\omega(\theta_i))\psi^{\mathrm{PER}}_{-(j+1),m+2}(R_\omega(\theta_i)) \\
&= \sum_{i=0}^{N-1} \psi^{\mathrm{PER}}_{-j,0}(R_\omega(\theta_{\widetilde{i+\kappa_1)}}))\psi^{\mathrm{PER}}_{-(j+1),m}(R_\omega(\theta_{\widetilde{i+\kappa_2)}})) \\
&= \sum_{\iota=0}^{N-1} \psi^{\mathrm{PER}}_{-j,0}(R_\omega(\theta_\iota))\psi^{\mathrm{PER}}_{-(j+1),m}(R_\omega(\theta_\iota)) \\
&= \mathsf{Id}_{0,m},
\end{aligned}$$

where $\widetilde{\iota+\kappa_1} = (1-0)2^{J-j} \pmod N$ and $\widetilde{\iota+\kappa_2} = (m+2-m)2^{J-(j+1)} \pmod N = 2^{J-j} \pmod N$. In view of that, the assertion follows replacing l and $l+1$ in the formulation of κ_1 for the case $k = j+1$. The proof for a general k it is done with the same trick for κ_1 and setting $\widetilde{\iota+\kappa_2} = 2^l$, where $l = 1, \ldots, J-1-j$.

(c) Finally, to show what we have claimed at the beginning of the proof (the property of the second row), we need to show that $\mathrm{Id}_{1,m} = -\mathrm{Id}_{1,m\widetilde{+}2^l}$. But by Lemma 2.2.1,

$$\psi^{\mathrm{PER}}_{0,0}(R_\omega(\theta_i)) = -\psi^{\mathrm{PER}}_{0,0}(R_\omega(\theta_{\widetilde{i+\kappa)}}),$$

where $\widetilde{i+\kappa)} = i+\kappa \pmod N$ and $\kappa = 2^{J-1}$. Also, for $j \geq 1$, by Lemma 2.2.3 and in the same way as in Equation (6.5):

$$\psi^{\mathrm{PER}}_{-j,0}(R_\omega(\theta_i)) = \psi^{\mathrm{PER}}_{-j,2^{j-1}}(R_\omega(\theta_{\widetilde{i+\kappa)}}),$$

where $\widetilde{i+\kappa)} = i+\kappa \pmod N$ and $\kappa = 2^{J-1}$. Hence the assertion follows.

□

Remark 6.4.2. It is important to notice that the index for the columns, m, are considered in $\mathbb{Z}/2^k\mathbb{Z}$. That is $m\widetilde{+}i = m+i \pmod{2^k}$. In order to do not do the proof unreadable we have not introduced it. Finally, notice that the mysterious range for l corresponds to the one of $k-j$ moving $j = 1, \ldots, J-1$. ∎

Bibliography

[AC09] Lluís Alsedà and Sara Costa. On the definition of strange nonchaotic attractor. *Fund. Math.*, 206:23–39, 2009. (Cited on page 68.)

[AM08] Lluís Alsedà and Michał Misiurewicz. Attractors for unimodal quasiperiodically forced maps. *J. Difference Equ. Appl.*, 14(10-11):1175–1196, 2008. (Cited on pages v, vii, viii, ix, 131, 132, 134, 135, 137, 143, 144, 157 and 159.)

[BBD+02] Arne Barinka, Titus Barsch, Stephan Dahlke, Mario Mommer, and Michael Konik. Quadrature formulas for refinable functions and wavelets. II. Error analysis. *J. Comput. Anal. Appl.*, 4(4):339–361, 2002. (Cited on pages 56 and 63.)

[BBDK01] A. Barinka, T. Barsch, S. Dahlke, and M. Konik. Some remarks on quadrature formulas for refinable functions and wavelets. *ZAMM Z. Angew. Math. Mech.*, 81(12):839–855, 2001. (Cited on pages 56 and 63.)

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223. (Cited on pages 14 and 19.)

[BP02] Luis Barreira and Yakov B. Pesin. *Lyapunov exponents and smooth ergodic theory*, volume 23 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2002. (Cited on page 67.)

[CL99] Carmen Chicone and Yuri Latushkin. *Evolution semigroups in dynamical systems and differential equations*, volume 70 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1999. (Cited on page 82.)

[Coh03] Albert Cohen. *Numerical analysis of wavelet methods*, volume 32 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 2003. (Cited on page 21.)

[CR09] Sara Costa Romero. *Study of a class of skew-products defined on the cylinder*. PhD thesis, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Catalonia, June 2009. (Cited on pages 65, 67 and 70.)

[Dau92] Ingrid Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. (Cited on pages viii, 28 and 29.)

[DL91] Ingrid Daubechies and Jeffrey C. Lagarias. Two-scale difference equations. I. Existence and global regularity of solutions. *SIAM J. Math. Anal.*, 22(5):1388–1410, 1991. (Cited on pages viii, 29 and 30.)

[DL92] Ingrid Daubechies and Jeffrey C. Lagarias. Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals. *SIAM J. Math. Anal.*, 23(4):1031–1079, 1992. (Cited on pages viii, 29 and 31.)

[dlLO99] R. de la Llave and R. Obaya. Regularity of the composition operator in spaces of Hölder functions. *Discrete Contin. Dynam. Systems*, 5(1):157–184, 1999. (Cited on pages v and 81.)

[dlLP02] Rafael de la Llave and Nikola P. Petrov. Regularity of conjugacies between critical circle maps: an experimental study. *Experiment. Math.*, 11(2):219–241, 2002. (Cited on pages viii, 21, 87 and 89.)

[ESW05] Howard C. Elman, David J. Silvester, and Andrew J. Wathen. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2005. (Cited on page 129.)

[FR11] Jordi-Lluís Figueras Romero. *Fiberwise Hyperbolic Invariant Tori in quasiperiodically forced skew product systems*. PhD thesis, Universitat de Barcelona, Barcelona, Catalonia, April 2011. (Cited on pages v and 65.)

[Fra99] Michael W. Frazier. *An introduction to wavelets through linear algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1999. (Cited on pages 48 and 49.)

[GOPY84] Celso Grebogi, Edward Ott, Steven Pelikan, and James A. Yorke. Strange attractors that are not chaotic. *Phys. D*, 13(1-2):261–268, 1984. (Cited on page 71.)

[Gut07] Martin Gutknecht. A brief introduction to krylov space methods for solving linear systems. *Frontiers of Computational Science — Proceedings of the International Symposium on Frontiers of Computational Science 2005 (Y. Kaneda, H. Kawamura, and M. Sasai, eds.)*, pages 53–62, 2007. (Cited on page 51.)

[Har12] Àlex Haro. On strange attractors in a class of pinched skew products. *Discrete Contin. Dyn. Syst.*, 32(2):605–617, 2012. (Cited on pages viii and 69.)

[HdLL06a] À. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: numerical algorithms. *Discrete Contin. Dyn. Syst. Ser. B*, 6(6):1261–1300, 2006. (Cited on pages v, 3 and 82.)

[HdlL06b] A. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: rigorous results. *J. Differential Equations*, 228(2):530–579, 2006. (Cited on pages v, 3 and 82.)

[HdlL07] A. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of hyperbolicity. *SIAM J. Appl. Dyn. Syst.*, 6(1):142–207, 2007. (Cited on pages v, 3 and 82.)

[HJ85] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1985. (Cited on pages 110 and 119.)

[HPS77] M. W. Hirsch, C. C. Pugh, and M. Shub. *Invariant manifolds*. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin, 1977. (Cited on page 75.)

[HV05] Daan Huybrechs and Stefan Vandewalle. Composite quadrature formulae for the approximation of wavelet coefficients of piecewise smooth and singular functions. *J. Comput. Appl. Math.*, 180(1):119–135, 2005. (Cited on pages 56 and 63.)

[HW96] Eugenio Hernández and Guido Weiss. *A first course on wavelets*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1996. With a foreword by Yves Meyer. (Cited on pages viii, 4, 12, 20 and 21.)

[Jor01] Àngel Jorba. Numerical computation of the normal behaviour of invariant curves of n-dimensional maps. *Nonlinearity*, 14(5):943–976, 2001. (Cited on pages v, 3 and 104.)

[Jor14] Àngel Jorba. Dancing with snas. In *New Perspectives in Discrete Dynamical Systems, on the occasion of Lluís Alsedà's 60th birthday*, 2014. (Cited on page vii.)

[JT08] Àngel Jorba and Joan Carles Tatjer. A mechanism for the fractalization of invariant curves in quasi-periodically forced 1-D maps. *Discrete Contin. Dyn. Syst. Ser. B*, 10(2-3):537–567, 2008. (Cited on pages iv, v and 83.)

[Kel95] C. T. Kelley. *Iterative methods for linear and nonlinear equations*, volume 16 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. With separately available software. (Cited on page 54.)

[Kel96] Gerhard Keller. A note on strange nonchaotic attractors. *Fund. Math.*, 151(2):139–148, 1996. (Cited on pages v, 69, 72 and 97.)

[Mal98] Stéphane Mallat. *A wavelet tour of signal processing*. Academic Press Inc., San Diego, CA, 1998. (Cited on pages 4, 7, 8, 9, 10, 20, 21, 47, 48, 89 and 145.)

[Mat68] John N. Mather. Characterization of Anosov diffeomorphisms. *Nederl. Akad. Wetensch. Proc. Ser. A 71 = Indag. Math.*, 30:479–483, 1968. (Cited on page 82.)

[Mea92] D. G. Mead. Newton's identities. *Amer. Math. Monthly*, 99(8):749–751, 1992. (Cited on page 62.)

[Mey00] Carl Meyer. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. With 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual (iv+171 pp.). (Cited on page 51.)

[Mey01] Yves Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures. (Cited on page 21.)

[MM08] Giuseppe Mastroianni and Gradimir V. Milovanović. *Interpolation processes*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2008. Basic theory and applications. (Cited on pages 61 and 62.)

[Mor97] Igor Moret. A note on the superlinear convergence of gmres. *SIAM Journal on Numerical Analysis*, 34, 1997. (Cited on page 108.)

[Nis96] Kunihiko Nishikawa, Takashi; Kaneko. Fractalization of a torus as a strange nonchaotic attractor. *Physical Review E*, 54, 12 1996. (Cited on pages iv, vii, viii, 133, 143, 157 and 159.)

[Reg07] Georg Regensburger. Parametrizing compactly supported orthonormal wavelets by discrete moments. *Appl. Algebra Engrg. Comm. Comput.*, 18(6):583–601, 2007. (Cited on pages 10, 56 and 102.)

[RS96] Thomas Runst and Winfried Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996. (Cited on page 19.)

[RS05] Georg Regensburger and Otmar Scherzer. Symbolic computation for moments and filter coefficients of scaling functions. *Ann. Comb.*, 9(2):223–243, 2005. (Cited on pages 10, 56 and 102.)

[Saa03] Yousef Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003. (Cited on pages 51, 53, 54 and 129.)

[SN96] Gilbert Strang and Truong Nguyen. *Wavelets and filter banks*. Wellesley-Cambridge Press, Wellesley, MA, 1996. (Cited on page 28.)

[Sta97] Jaroslav Stark. Invariant graphs for forced systems. *Phys. D*, 109(1-2):163–179, 1997. Physics and dynamics between chaos, order, and noise (Berlin, 1996). (Cited on pages viii, 71, 75 and 97.)

[Sta99] Jaroslav Stark. Regularity of invariant graphs for forced systems. *Ergodic Theory Dynam. Systems*, 19(1):155–199, 1999. (Cited on pages viii, 71, 75, 78 and 97.)

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970. (Cited on pages 16 and 19.)

[Tri83] Hans Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983. (Cited on pages vii, 14, 15 and 16.)

[Tri06] Hans Triebel. *Theory of function spaces. III*, volume 100 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2006. (Cited on pages vii, viii, 20, 21 and 93.)

[Tri10] Hans Triebel. *Bases in function spaces, sampling, discrepancy, numerical integration*, volume 11 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010. (Cited on pages vii and 23.)

[vdV09] Henk A. van der Vorst. *Iterative Krylov methods for large linear systems*, volume 13 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2009. Reprint of the 2003 original. (Cited on page 51.)

[Vid99] Brani Vidakovic. *Statistical modeling by wavelets*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, 1999. A Wiley-Interscience Publication. (Cited on pages 29, 31 and 32.)

[Wal82] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982. (Cited on page 67.)

List of Figures

Numerical Computation of Invariant Objects with Wavelets

Abstract: In certain classes of dynamical systems invariant sets with a strange geometry appear. For example, under certain conditions, the iteration of two-dimensional quasi-periodically forced skew product $(\theta_{n+1}, x_{n+1}) = (R_\omega(\theta_n), F_{\sigma,\varepsilon}(\theta_n, x_n))$ where $R_\omega(\theta) = \theta + \omega$, $\omega \in \mathbb{R}\backslash\mathbb{Q}$ and $\sigma, \varepsilon \in \mathbb{R}$ gives us Strange Non-Chaotic Attractors, φ.

To obtain analytical approximation of these objects it seems more natural to use wavelets instead of the more usual Fourier approach due to its adaptability. The aim of this thesis is to describe an efficient algorithm for the semi-analitical computation of the invariant object, using both Daubechies and Haar wavelets, by means of the numerical computation of the wavelet coefficients.

The aim for this exercise is twofold. From one side to be able to study possible bifurcations or zoom in the "pinching zone" of the object. From the other side try to get estimates of the regularity of the object. The study of this regularity depending on parameters, for a certain models of skew products, may give another point of view to the fractalization routes described in the literature and that are currently under discussion.

To perform such exercise(s), firstly, we have translated the $\mathbb{R}$-Daubechies wavelets language to $\mathbb{S}^1$. After that, we have carried out two different strategies to get the wavelet coefficients $\mathrm{D}^{\mathrm{PER}}$. The first one based on the Fast Wavelet Transform. The other, solve the Invariance Equation $\varphi(R_\omega(\theta)) = F_{\sigma,\varepsilon}(\theta, \varphi(\theta))$ using the Newton's method. From such coefficients $\mathrm{D}^{\mathrm{PER}}$ we get (numerical) estimations for the aforesaid proposed questions.

Keywords: Wavelets, Fast Wavelet Transform, Besov spaces, quasiperiodically forced system, Lyapunov exponent, Mather spectrum