This is a postprint version of the following published document:

Ibort, A. & Pérez-Pardo, J. M. (2013). Numerical Solutions of the Spectral Problem for Arbitrary Self-Adjoint Extensions of the One-Dimensional Schrödinger Equation. *SIAM Journal on Numerical Analysis*, 51(2), pp. 1254–1279.

DOI: <https://doi.org/10.1137/110856800>

© 2013, Society for Industrial and Applied Mathematics

NUMERICAL SOLUTIONS OF THE SPECTRAL PROBLEM FOR ARBITRARY SELF-ADJOINT EXTENSIONS OF THE 1D SCHRÖDINGER EQUATION*

A. IBORT[†] AND J. M. PÉREZ-PARDO^{†‡}

Abstract. A numerical algorithm to solve the spectral problem for arbitrary self-adjoint extensions of 1D regular Schrödinger operators is presented. It is shown that the set of all self-adjoint extensions of 1D regular Schrödinger operators is in one-to-one correspondence with the group of unitary operators on the finite dimensional Hilbert space of boundary data, and they are characterized by a generalized class of boundary conditions that include the well-known Dirichlet, Neumann, Robin, (quasi-)periodic boundary conditions, etc. The corresponding numerical algorithms to solve the eigenvalue problem for each generalized boundary condition are constructed and their convergence is proved. An appropriate basis of boundary functions must be introduced to deal with arbitrary boundary conditions as described by the general theory and the conditioning of its computation is analyzed. Some significant numerical experiments are also discussed as well as the comparison with some standard algorithms. In particular it is shown that an appropriate perturbation of well behaved boundary conditions, both for the free particle and the harmonic oscillator, leads to the theoretically predicted result of the unbounded below behaviour of the groundlevel and the localization of the corresponding eigenvectors at the boundary (edge states).

1. Introduction. The study of the self-adjointness of Schrödinger operators has been a fundamental mathematical problem since the beginning of Quantum Mechanics and there is a vast literature on the subject (see for instance [Re75], the review [Si00] and references therein). In spite of this, there is a continuous flow of new results and even surprises (see for instance the recent papers where some apparent paradoxical aspects of the spectrum of certain self-adjoint extensions of the Schrödinger operator in 2D are analyzed [Be08], [Ma09], [Be09]).

Consider the evolution of a quantum system on a D -dimensional riemannian manifold M with boundary ∂M under the influence of a potential V which is given by the Schrödinger equation $i\hbar \frac{\partial \Psi}{\partial t} = H\Psi$, with H the Hamiltonian operator of the system given by

$$H = -\frac{\hbar^2}{2m} \Delta_\eta + V(x) = -\frac{\hbar^2}{2m} \frac{1}{\sqrt{|\eta|}} \frac{\partial}{\partial x^j} \sqrt{|\eta|} \eta^{jk} \frac{\partial}{\partial x^k} + V(x), \quad (1.1)$$

where Δ_η stands for the Laplace-Beltrami operator on M defined by the metric tensor η given by $\eta = \eta_{jk}(x) dx^k dx^j$, $|\eta| = |\det \eta_{jk}(x)|$ and $\eta^{ij} \eta_{jk} = \delta_k^i$. The second order differential operator H is formally self-adjoint, however in order to define a unitary evolution of the quantum state Ψ , a self-adjoint extension of it must be specified. If H denotes one of such self-adjoint extensions, because of Stone's theorem, a one-parameter group U_t of unitary operators exists such that $U_t = \exp(-itH/\hbar)$ and the quantum evolution of a given initial state Ψ_0 will be uniquely determined as $\Psi_t = U_t \Psi_0$. Self-adjoint extensions of the operator H are usually chosen by fixing the boundary values of the functions where the operator H acts, typically Dirichlet or Neumann boundary conditions. However they are by no means the most general choice of boundary conditions determining self-adjoint extensions of the operator H and a variety of other possibilities exist. The development of quantum information technologies makes relevant the discussion of more general classes of boundary conditions, i.e., of general self-adjoint extensions for the operator H , and the numerical computation of their spectrum in order to integrate Schrödinger's equation.

In this paper we present an algorithm based in the finite element method to solve numerically the spectral problem for *all* possible self-adjoint extensions of 1D regular Schrödinger operators. In this particular instance, the Schrödinger operator takes the form of a Sturm-Liouville operator. There exist many methods for approaching this problem (see [Ac09], [Ch99], [Ch09], [Fa57], [Pr73] just to mention a few of them), however they cannot be used to solve the problem of general self-adjoint extensions of it, as explained below. Most of the previous algorithms solve the wide class

*This work has been partially supported by the Spanish MICINN grants MTM2010-21186-C02-02 and QUITEMAD programme P2009 ESP-1594.

[†]Depto. de Matemáticas, Univ. Carlos III de Madrid, Avda. de la Universidad 30, 28911 Leganés, Madrid, Spain.

[‡]Supported by the UCIIM University through the PhD Program Grant M02-0910.

of self-adjoint extensions given by the equations of the form:

$$\begin{aligned}\alpha \cdot u(a) + \beta \cdot u'(a) &= 0 \\ \gamma \cdot u(b) + \delta \cdot u'(b) &= 0,\end{aligned}\tag{1.2}$$

where a and b are the endpoints of the interval, or for slightly more general boundary conditions like in [Ch09]. For instance, (quasi-)periodic boundary conditions can not be implemented with the above setting.

In this work we will take the approach developed in [As05] to describe the set of self-adjoint extensions of Laplace-Beltrami operators. We will show in section 2 that in 1D the set of self-adjoint extensions of the Schrödinger operator H is in one-to-one correspondence with the group of unitary operators on the space of boundary data for the problem and we will provide explicit constructions of such correspondence so that it can be actually used in the implementation of the corresponding numerical algorithms. Similar results are valid in higher dimensions, however there are additional technical difficulties related to the functional spaces where the domains of the operators are defined, that make the presentation and the discussion of the results more difficult. Because of this and in order to make the presentation of the results as simple and as short as possible we will consider in this paper just the 1D situation leaving the discussion in higher dimensions for later works. Thus the manifold M will be assumed to be one-dimensional and compact and the potential function V will be assumed to be regular and bounded below.

The numerical algorithm which is developed to solve the eigenvalue problem for all self-adjoint extensions of the Hamiltonian operator H described in terms of boundary data is based on the finite element method, even though the treatment of the boundary basis functions is novel. The determination of the boundary basis function is analyzed as well as its conditioning. The stability and convergence of the method is proved by describing the corresponding self-adjoint extensions in weak form. A number of experiments are presented that are consistent with the obtained results and shows the efficiency of the corresponding computer algorithms. These results are compared with the standard methods in some significant examples showing again the good behavior of the proposed algorithm. In particular it is shown that for a family of self-adjoint extensions the groundlevel of the system must go to $-\infty$, a behaviour that is described nicely by the algorithm presented here as well as the presence of *edge states*, i.e., eigenfunctions which are localized at the boundary.

The paper is organized as follows. Section 2 is devoted to prove the general theorems that will allow for a convenient description of the self-adjoint extensions of the Schrödinger operator H in 1D in terms of boundary data. The main results of the finite element method for the eigenvalue problem for completely general self-adjoint extensions are discussed in section 3. An appropriate basis of boundary functions is introduced there in order to implement the boundary conditions as described by the general theory. Moreover, in this section, results concerning the convergence and the stability of the numerical scheme are also proved. In section 4 we present some numerical experiments displaying relevant features of the algorithm, as well as a comparison with other methods.

2. Self-adjoint extensions of Schrödinger operators. As it was stated in the introduction, we will restrict our attention to the case of Schrödinger operators on 1D compact manifolds and regular potentials bounded below. The Schrödinger operator for a particle moving on a smooth manifold M with boundary ∂M and riemannian metric η is given by the Hamiltonian operator H defined in eq. (1.1). Notice first that a compact 1D manifold M consists on a finite number of closed intervals I_α , $\alpha = 1, \dots, n$. Each interval will have the form $I_\alpha = [a_\alpha, b_\alpha] \subset \mathbb{R}$ and the boundary of the manifold $M = \coprod_{\alpha=1}^n [a_\alpha, b_\alpha]$ is given by the family of points $\{a_1, b_1, \dots, a_n, b_n\}$. Functions Ψ on M are determined by vectors $(\Psi_1, \dots, \Psi_n)$ of complex valued functions $\Psi_\alpha: I_\alpha \rightarrow \mathbb{C}$. A riemannian metric η on M is given by specifying a riemannian metric η_α on each interval I_α , this is, by a positive smooth function $\eta_\alpha(x) > 0$ on the interval I_α , i.e., $\eta|_{I_\alpha} = \eta_\alpha(x)dx^2$. Then the L^2 inner product on I_α takes the form $\langle \Psi_\alpha, \Phi_\alpha \rangle = \int_{a_\alpha}^{b_\alpha} \bar{\Psi}_\alpha(x) \Phi_\alpha(x) \sqrt{\eta_\alpha(x)} dx$ and the Hilbert space of square integrable functions on M is given by $L^2(M) = \bigoplus_{\alpha=1}^n L^2(I_\alpha, \eta_\alpha)$.

The second order differential operator Δ_η takes in 1D the form of a Sturm-Liouville operator on each subinterval and is formally self-adjoint in the sense that

$$\int_M \bar{\Psi} \Delta_\eta \Phi \sqrt{\eta} dx = - \int_M \frac{d\bar{\Psi}}{dx} \frac{d\Phi}{dx} \sqrt{\eta} dx = \int_M (\overline{\Delta_\eta \Psi}) \Phi \sqrt{\eta} dx, \quad (2.1)$$

for $\Psi, \Phi \in C_c^\infty(M \setminus \partial M)$ complex-valued smooth functions with compact support in the interior of M . In fact, the differential expression (2.1) defines a symmetric operator on $L^2(M)$, the space of square integrable functions on M with respect to the measure defined by the metric η , with dense domain $C_c^\infty(M \setminus \partial M)$. The operator Δ_η is symmetric but not self-adjoint on $L^2(M)$ because the adjoint operator $\Delta_\eta^\dagger$ has a dense domain $D^\dagger$ on $L^2(M)$ that in the particular instance of 1D is exactly given by $\mathcal{H}^2(M)$, the Sobolev space of order 2 on M , i.e., the space of functions on $L^2(M)$ possessing first and second weak derivatives that are square integrable. In higher dimensions this is not true because the domain of the adjoint could contain functions which are not in $\mathcal{H}^2(M)$ (see for instance [Gr68]). Clearly $C_c^\infty(M \setminus \partial M) \subset \mathcal{H}^2(M)$. The operator Δ_η is closable on $L^2(M)$ and its closed extension, that we will keep denoting by Δ_η , has domain $D_0 := \mathcal{H}_0^2(M)$, the Sobolev space of order 2 on M with functions and first derivatives vanishing at the boundary. Self-adjoint extensions of the operator Δ_η are given by operators $\Delta_D := \Delta_\eta^\dagger|_D$ with domain D such that $D_0 \subset D \subset D_0^\dagger$, $\Delta_D|_{D_0} = \Delta_\eta$ and $\Delta_D = \Delta_D^\dagger$.

Von Neumann's theorem establishes (see for instance [We80], Thm. 8.12) that there is a one-to-one correspondence between self-adjoint extensions Δ_D of the Laplace-Beltrami operator Δ_η and unitary operators $K: \mathcal{N}_+ \rightarrow \mathcal{N}_-$, where the deficiency spaces $\mathcal{N}_\pm$ are defined as:

$$\mathcal{N}_\pm = \{\Psi \in \mathcal{H}^2(M) \mid \Delta_\eta^\dagger \Psi = \pm i\Psi\}.$$

Thus, given a unitary operator $K: \mathcal{N}_+ \rightarrow \mathcal{N}_-$, the domain D of the operator Δ_D is given by

$$D = D_0 \oplus (I + K)\mathcal{N}_+, \quad (2.2)$$

and the extended operator Δ_D takes the explicit form: $\Delta_D(\Psi_0 \oplus (I + K)\xi_+) = \Delta_\eta \Psi_0 \oplus i(I - K)\xi_+$, for all $\Psi_0 \in D_0$ and $\xi_+ \in \mathcal{N}_+$.

Unfortunately von Neumann's theorem is not always suitable for the explicit construction of the most general self-adjoint extensions of the Laplace-Beltrami operator because we need to determine first the deficiency spaces $\mathcal{N}_\pm$. Moreover, the explicit construction of the self-adjoint domain as given in (2.2), is not well suited for numerical purposes. We can take however a different route inspired in the classical treatment of formally self-adjoint differential operators. If we rewrite eq. (2.1) for functions $\Psi, \Phi \in D^\dagger = \mathcal{H}^2(M)$ instead of $C_c^\infty(M \setminus \partial M)$, a simple computation shows:

$$\int_M \bar{\Psi} \Delta_\eta \Phi \sqrt{\eta} dx = \int_M (\overline{\Delta_\eta \Psi}) \Phi \sqrt{\eta} dx + \int_{\partial M} (\bar{\psi} \dot{\varphi} - \bar{\dot{\psi}} \varphi) \text{vol}_{\partial\eta}, \quad (2.3)$$

where $\text{vol}_{\partial\eta}$ stands for the induced riemannian metric at the boundary, $\varphi := \Phi|_{\partial M}$, $\dot{\varphi} := \frac{d\Phi}{d\nu}|_{\partial M}$, with ν is the exterior normal vector to ∂M , and respectively for ψ and $\dot{\psi}$. We thus obtain the Lagrange boundary form Σ for the Laplace-Beltrami operator:

$$\Sigma((\psi, \dot{\psi}), (\varphi, \dot{\varphi})) = \int_{\partial M} \bar{\psi} \dot{\varphi} \text{vol}_{\partial\eta} - \int_{\partial M} \bar{\dot{\psi}} \varphi \text{vol}_{\partial\eta} = \langle \psi, \dot{\varphi} \rangle_{L^2(\partial M)} - \langle \dot{\psi}, \varphi \rangle_{L^2(\partial M)}. \quad (2.4)$$

In what follows, if there is no risk of confusion, we will omit the subscript $L^2(\partial M)$ that denotes the L^2 inner product on the boundary manifold ∂M with respect to the measure defined by the volume form $\text{vol}_{\partial\eta}$ and we will simply write $\langle \psi, \varphi \rangle = \int_{\partial M} \bar{\psi} \varphi \text{vol}_{\partial\eta}$. The Lagrange boundary bilinear form Σ defines a continuous bilinear form on the Hilbert space $L^2(\partial M) \oplus L^2(\partial M)$,

$$\Sigma((\psi_1, \psi_2), (\varphi_1, \varphi_2)) = \langle \psi_1, \varphi_2 \rangle - \langle \psi_2, \varphi_1 \rangle, \quad \forall (\psi_1, \psi_2), (\varphi_1, \varphi_2) \in L^2(\partial M) \oplus L^2(\partial M). \quad (2.5)$$

The trace map $\gamma: \mathcal{H}^2(M) \rightarrow L^2(\partial M) \oplus L^2(\partial M)$ given by $\gamma(\Psi) = (\psi, \dot{\psi})$, is continuous and induces a homeomorphism from $\mathcal{H}^2(M)/\ker \gamma$ into $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M) \subset L^2(\partial M) \oplus L^2(\partial M)$ (see for

instance [Ad75], Thm. 7.20), where $\mathcal{H}^{3/2}$ and $\mathcal{H}^{1/2}$ are Sobolev spaces of fractional order. The previous observations provide a simple characterization of self-adjoint extensions of the operator Δ_η in 1D.

THEOREM 2.1. *There is a one-to-one correspondence between self-adjoint extensions Δ_D of the Laplace–Beltrami operator Δ_η and non-trivial maximal closed isotropic subspaces W of Σ contained in $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$. The correspondence being explicitly given by $D \mapsto \gamma(D)$.*

Proof. Let $D \subset \mathcal{H}^2(M)$ be the domain of a self-adjoint extension Δ_D of the operator Δ_η . Consider the subspace $W := \gamma(D) \subset \mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M) \subset L^2(\partial M) \oplus L^2(\partial M)$ consisting on the set of pairs of functions $(\varphi, \dot{\varphi})$ that are respectively the restriction to ∂M of a function $\Phi \in D$ and its normal derivative. Notice that the subspace D is closed in $\mathcal{H}^2(M)$. Hence, because γ is an homeomorphism from $\mathcal{H}^2(M)/\ker \gamma$ to $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$, $\gamma(D)$ is a closed subspace of $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$. Because of eq. (2.3), it is clear that $\Sigma|_W = 0$ and finally the subspace W is maximally isotropic in $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$, because if this were not the case, there will be a closed isotropic subspace $W' \subset \mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$ containing W . Then $\gamma^{-1}(W')$ defines a domain D' containing D such that the operator $\Delta_\eta|_{D'}$ would be symmetric on it, in contradiction with the self-adjointness assumption.

Let $W \subset L^2(\partial M) \oplus L^2(\partial M)$ be a maximal closed Σ -isotropic subspace in $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$. Then consider the subspace $D_W := \gamma^{-1}(W) \subset \mathcal{H}^2(M)$ of functions Ψ such that $(\psi, \dot{\psi}) \in W$. It is clear that for any pair of functions Ψ, Φ on D_W , because W is isotropic with respect to Σ , eq. (2.3) gives $\langle \Psi, \Delta_\eta \Phi \rangle = \langle \Delta_\eta \Psi, \Phi \rangle$ and the operator Δ_η is therefore symmetric in D_W . Moreover, because of the maximality of W in $\mathcal{H}^{3/2}(\partial M) \oplus \mathcal{H}^{1/2}(\partial M)$ it is easy to see that $D_W = D_W^\dagger$, hence it is self-adjoint. $\square$

The key observation now is that, because $\dim M = 1$, $\mathcal{H}^q(\partial M) = L^2(\partial M) = \mathbb{C}^{2n} \quad \forall q$, where n is the number of subintervals. Hence, the self-adjoint extensions of the Laplace–Beltrami operator are in one-to-one correspondence with the maximal isotropic subspaces of the Lagrange boundary form Σ . It is a well known result (see [As05], [Br08] and references therein) that the maximal isotropic subspaces of the bilinear form (2.5) are in one-to-one correspondence with the graphs of unitary operators $U : L^2(\partial M) \rightarrow L^2(\partial M)$ through the relationship

$$\varphi_1 - i\varphi_2 = U(\varphi_1 + i\varphi_2).$$

The last step in discussing all self-adjoint extensions of the Hamiltonian H consists in realizing that the operator multiplication by a regular function is essentially self-adjoint and its unique self-adjoint extension has domain $L^2(M)$. Hence, the self-adjoint extensions of H coincide with the self-adjoint extensions of Δ_η .

We can summarize the preceding analysis by stating that, under the conditions above, the domain of a self-adjoint extension of the Schrödinger operator H in 1D is defined as a closed subspace of functions Ψ on $\mathcal{H}^2(M)$ satisfying:

$$\psi - i\dot{\psi} = U(\psi + i\dot{\psi}) \tag{2.6}$$

for a given unitary operator $U : \mathbb{C}^{2n} \rightarrow \mathbb{C}^{2n}$. Notice for instance that $U = I$ corresponds to Neumann's boundary conditions and $U = -I$ determines Dirichlet's boundary conditions. The formula above, eq. (2.6), provides a powerful and effective computational tool to deal with general self-adjoint extensions of Schrödinger operators, (even in dimensions higher than 1), and will be used extensively in the rest of the paper.

3. Finite element method for the eigenvalue problem in the interval. Our aim now is to find an algorithm to solve the spectral problem for a self-adjoint extension of the 1D Schrödinger operator determined by a unitary matrix U explicitly given by the differential system:

$$\begin{cases} -\Delta_\eta \Phi + V\Phi = \lambda\Phi, & \Phi \in \mathcal{H}^2(M) \\ \varphi - i\dot{\varphi} = U(\varphi + i\dot{\varphi}), & \varphi = \Phi|_{\partial M}, \quad \dot{\varphi} = \frac{d\Phi}{d\nu}|_{\partial M}, \quad U \in U(2n). \end{cases} \tag{3.1}$$

As before, we suppose that the potential V is regular and bounded below. We will develop in the following paragraphs a numerical algorithm to solve such system.

From now on we will keep the *dot* notation for the outwards normal derivative at the boundary $\dot{\varphi} = \frac{d\Phi}{d\nu}\big|_{\partial M}$ and use *primes* to denote the standard derivative $\Phi' = \frac{d\Phi}{dx}$. Because of the particularly simple form that the Lagrange boundary form takes in 1D, it will be also convenient to introduce the following notation:

$$[\bar{\Psi}\Phi']\big|_{\partial M} = \sum_{\alpha=1}^n \bar{\Psi}(b_\alpha)\Phi'(b_\alpha) - \bar{\Psi}(a_\alpha)\Phi'(a_\alpha).$$

It is a well known fact that it is convenient to look for weak solutions of (3.1). Taking the inner product of (3.1) with a vector Ψ on the dense domain $C^\infty(M)$ of smooth functions on M , and integrating by parts we obtain easily the corresponding weak problem:

$$\begin{cases} \langle \Psi', \Phi' \rangle - [\bar{\Psi}\Phi']_{\partial M} + \langle \Psi, V\Phi \rangle = \lambda \langle \Psi, \Phi \rangle, & \forall \Psi \in C^\infty(M) \\ \varphi - i\dot{\varphi} = U(\varphi + i\dot{\varphi}), \quad \varphi = \Phi\big|_{\partial M}, \quad \dot{\varphi} = \frac{d\Phi}{d\nu}\big|_{\partial M}, & U \in U(2n). \end{cases} \quad (3.2)$$

It is easy to prove that (Φ, λ) is a solution of (3.1) if and only if it solves (3.2). In the next subsections we develop an algorithm that approximates such solutions and, moreover, we prove its stability and convergence.

3.1. Finite elements for general self-adjoint boundary conditions. To solve (3.2) we construct a family of finite dimensional subspaces of functions S^N of $L^2(M)$ satisfying the boundary conditions (2.6). Such finite dimensional subspaces will be constructed using finite elements. The finite element model $(K, \mathcal{P}, \mathcal{N})$ that we use is given by $K = [0, 1]$ the unit interval, $\mathcal{P}$ the space of linear polynomials on K and $\mathcal{N}$ the vertex set $\{0, 1\}$.

The domain of our problem is the manifold M which consists on the union of the intervals $I_\alpha = [a_\alpha, b_\alpha]$, $\alpha = 1, \dots, n$. For each N we will construct a non-degenerate subdivision M^N as follows. Let r_α be the integer defined as $r_\alpha = [L_\alpha N / L] + 1$, where $[x]$ denotes the integer part of x . We will assume that each $r_\alpha \geq 2$, and $N \geq 2n$. Let us denote by r the multi-index $(r_1, \dots, r_n)$. Then $|r| = r_1 + \dots + r_n$ satisfies:

$$N \leq |r| \leq N + n.$$

Now we will subdivide each interval I_α into $r_\alpha + 1$ subintervals of length:

$$h_\alpha = \frac{L_\alpha}{r_\alpha + 1}.$$

(It could be possible to use a set of independent steps h_α , one for each interval, however this would create some difficulties later on that we prevent in this way). Each subinterval I_α contains $r_\alpha + 2$ nodes that will be denoted as:

$$x_k^{(\alpha)} = a_\alpha + h_\alpha k, \quad k = 0, \dots, r_\alpha + 1.$$

3.1.1. Bulk functions. Consider now the family $\mathcal{F}_r$ of $|r| - 2n$ piecewise linear functions $\{f_k^{(\alpha)}(x)\}_{k=2}^{r_\alpha-1}$, $\alpha = 1, \dots, n$, $2 \leq k \leq r_\alpha - 1$, that are zero at all nodes except at the k -th node of the interval I_α , where it has value 1, or more explicitly :

$$f_k^{(\alpha)}(x) = \begin{cases} s & x = x_{k-1}^{(\alpha)} + sh_\alpha, \quad 0 \leq s \leq 1, \\ 1-s & x = x_k^{(\alpha)} + sh_\alpha, \quad 0 \leq s \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that these functions are differentiable on each subinterval. All these functions satisfy trivially the boundary conditions (2.6) because they and their normal derivatives vanish at the endpoints of each interval. We will call these functions *bulk functions*.

Fig. 3.1: Boundary function $\beta^{(w)}$.

3.1.2. Boundary functions. We will add to the set of bulk functions a family of functions that implement non-trivially the boundary conditions determining the self-adjoint extension. These functions will be called *boundary functions* and the collection of all of them will be denoted by $\mathcal{B}_r$. Contrary to bulk functions, boundary functions need to be “delocalized” so that they can fulfill any possible self-adjoint extension’s boundary condition.

Because the endpoints $x_0^{(\alpha)} = a_\alpha$, $x_{r_\alpha+1}^{(\alpha)} = b_\alpha$ of the intervals I_α and the adjacent nodes, $x_1^{(\alpha)}$ and $x_{r_\alpha}^{(\alpha)}$, are going to play a prominent role in what follows, we introduce some notation that takes care of them. We will consider an index $l = 1, \dots, 2n$ that labels the endpoints of the intervals. Now for each vector $w = (w_l) \in \mathbb{C}^{2n}$ consider the following functions (see fig. 3.1):

$$\beta^{(w)}(x) = \begin{cases} v_{2\alpha-1} + s(w_{2\alpha-1} - v_{2\alpha-1}) & x = x_0^{(\alpha)} + sh_\alpha, \\ w_{2\alpha-1}(1-s) & x = x_1^{(\alpha)} + sh_\alpha, \\ s w_{2\alpha} & x = x_{r_\alpha-1}^{(\alpha)} + sh_\alpha, \\ w_{2\alpha} + s(v_{2\alpha} - w_{2\alpha}) & x = x_{r_\alpha}^{(\alpha)} + sh_\alpha, \\ 0 & x_2^{(\alpha)} \leq x \leq x_{r_\alpha-1}^{(\alpha)} \end{cases} \quad \begin{matrix} 0 \leq s \leq 1, \\ 1 \leq \alpha \leq n \end{matrix}.$$

Each function on the previous family is determined (apart from the vector w) by the vector $v = (v_l) \in \mathbb{C}^{2n}$ that collects the values of $\beta^{(w)}$ at the endpoints of the subintervals. We select now a family of vectors w sufficient to span a linear space S^N that contains a solution to the spectral problem with boundary conditions (2.6). Thus we denote by $w^{(i)}$ the vector such that $w_l^{(i)} = \delta_{li}$, $i = 1, \dots, 2n$, i.e., the $2n$ vectors $w^{(i)}$ are just the standard basis for $\mathbb{C}^{2n}$. The corresponding functions $\beta^{(w)}$ will now be denoted simply by $\beta^{(i)}$. Notice that each boundary function $\beta^{(i)}$ is completely characterized by the unique non extremal node where it doesn’t vanish¹ and the values at the endpoints. We denote by $v_l^{(i)}$, $l = 1, \dots, 2n$, the boundary values of the functions $\beta^{(i)}$ above.

3.1.3. The boundary matrix. The $2n$ extremal values $v_l^{(i)}$ of the boundary functions $\beta^{(i)}$ are undefined, but we are going to show that the $2n$ conditions (2.6) imposed on the boundary functions constitute for them a determinate system of linear equations. Because the boundary functions are constructed to be piecewise linear, the normal derivatives of these functions at the boundary can be expressed as follows. For the left boundaries of the intervals, i.e., at the points a_α , we have:

$$\left. \frac{d\beta^{(i)}}{dn} \right|_{x=a_\alpha} = -\frac{1}{h_\alpha} (w_{2\alpha-1}^{(i)} - v_{2\alpha-1}^{(i)}), \quad (3.3)$$

and respectively for the right boundaries:

¹If $i = 2\alpha - 1$, it is the node $x_1^{(\alpha)}$ and if $i = 2\alpha$, it corresponds to the node $x_{r_\alpha}^{(\alpha)}$.

$$\left. \frac{d\beta^{(i)}}{dn} \right|_{x=b_\alpha} = \frac{1}{h_\alpha} (v_{2\alpha}^{(i)} - w_{2\alpha}^{(i)}) = -\frac{1}{h_\alpha} (w_{2\alpha}^{(i)} - v_{2\alpha}^{(i)}). \quad (3.4)$$

Thus the vector containing the normal derivatives of the function $\beta^{(i)}$, consistently denoted by $\dot{\beta}^{(i)}$, is given by:

$$\dot{\beta}_l^{(i)} = -\frac{1}{h_l} (w_l^{(i)} - v_l^{(i)}) = -\frac{1}{h_l} (\delta_l^i - v_l^{(i)}), \quad (3.5)$$

where we use again the consistent notation $h_l = h_\alpha$, if $l = 2\alpha - 1$, or $l = 2\alpha$. For each boundary function $\beta^{(i)}$, the boundary conditions (2.6) read simply as the system of $2n$ equations on the components of the vector $v^{(i)}$:

$$v^{(i)} - i\dot{\beta}^{(i)} = U(v^{(i)} + i\dot{\beta}^{(i)}),$$

or, componentwise:

$$v_l^{(i)} \left(1 - \frac{i}{h_l}\right) + \frac{i}{h_l} w_l^{(i)} = \sum_{j=1}^{2n} U_{lj} \left[v_j^{(i)} \left(1 + \frac{i}{h_j}\right) - \frac{i}{h_j} w_j^{(i)} \right], \quad l = 1, \dots, 2n.$$

Collecting coefficients and substituting the expressions $w_l^{(i)} = \delta_{il}$ we get

$$\sum_{j=1}^{2n} \left[\left(1 - \frac{i}{h_j}\right) \delta_{lj} - U_{lj} \left(1 + \frac{i}{h_j}\right) \right] v_j^{(i)} = \left[-\frac{i}{h_i} \delta_{li} - U_{li} \frac{i}{h_i} \right]. \quad (3.6)$$

This last equation can be written as the set of linear systems

$$FV = C, \quad (3.7)$$

with V a $2n \times 2n$ matrix whose entries are given by $V_{ji} = v_j^{(i)}$, $i, j = 1, \dots, 2n$. The i -th column of V contains the boundary values of the boundary function $\beta^{(i)}$. The $2n \times 2n$ matrix F with entries

$$F_{lj} = \left(1 - \frac{i}{h_j}\right) \delta_{lj} - U_{lj} \left(1 + \frac{i}{h_j}\right),$$

will be called the boundary matrix of the subdivision of the domain M determined by the integer N , and

$$C_{li} = -\frac{i}{h_i} (\delta_{li} + U_{li}),$$

are the inhomogeneous terms of the linear systems (3.7), one for each boundary function. Using a compact notation we get:

$$F = \text{diag}(\mathbf{1} - i/\mathbf{h}) - U \text{diag}(\mathbf{1} + i/\mathbf{h}), \quad C = -i(I + U) \text{diag}(1/\mathbf{h}),$$

where $1/\mathbf{h}$ denotes the vector whose components are $1/h_l$. Notice that F depends just on U and the integer N defining the size of discretization of the approximate weak spectral problem defined below.

3.2. Conditioning of the boundary matrix. We study now the behavior of the system (3.7) under perturbations, in other words, we compute the condition number of the boundary matrix F and show that it is low enough to assure the accuracy of the numerical determination of our family of boundary functions $\beta^{(i)}$. The relative condition number we want to compute is

$$\kappa(F) = \|F\| \|F^{-1}\|.$$

In our case, the boundary matrix F can be expressed as

$$F = \bar{D} - UD = (I - UD\bar{D}^{-1})\bar{D},$$

with $D_{ij}(\mathbf{h}) = D_{ij} = (1 + \frac{i}{h_i}\delta_{ij})$. Notice that the product $UD\bar{D}^{-1}$ is a unitary matrix which we will denote as $U_0(\mathbf{h})$ or simply U_0 if we don't want to emphasize the $\mathbf{h}$ dependence of U_0 . Thus, $F = (I - U_0)\bar{D}$ and

$$\|F\| = \|(I - U_0)\bar{D}\| \leq \|I - U_0\|\|\bar{D}\| \leq 2\|D\|.$$

On the other hand

$$\|F^{-1}\| = \|\bar{D}^{-1}(I - U_0)^{-1}\| \leq \|\bar{D}^{-1}\|\|(I - U_0)^{-1}\| = \frac{\|D^{-1}\|}{\min_{\lambda \in \text{spec}(U_0)}\{|1 - \lambda|\}}$$

and thus we obtain:

$$\kappa(F) \leq \kappa(D) \frac{2}{\min_{\lambda \in \text{spec}(U_0)}\{|1 - \lambda|\}}.$$

As D is a diagonal matrix its condition number is given by

$$\kappa(D) = \frac{\sqrt{\frac{1}{h_{\min}^2} + 1}}{\sqrt{\frac{1}{h_{\max}^2} + 1}} \leq \frac{h_{\max}}{h_{\min}},$$

with $h_{\max}$ ($h_{\min}$) the biggest (smallest) step of the discretization determined by N . We get finally,

$$\kappa(F) \leq \frac{h_{\max}}{h_{\min}} \frac{2}{|1 - \lambda|}, \quad (3.8)$$

with λ the closest element of the spectrum of U_0 to 1. Of course, because U_0 is unitary, it may happen that 1 is in its spectrum, so that the condition number is not bounded (and the system (3.7) is incompatible). Because the matrix U_0 depends on $\mathbf{h}$, its eigenvalues will depend on $\mathbf{h}$ too. We want to study the dependence of the closest eigenvalue to 1, or 1 for that matter, with respect to perturbations of the vector $\mathbf{h}$.

LEMMA 3.1. *Suppose that X_0 is an eigenvector with eigenvalue 1 of U_0 and that the perturbed matrix $\hat{U} = U_0 + \delta U$, for $\|\delta U\|$ small enough, is such that $1 \in \sigma(\hat{U})$. Then $\bar{X}_0^T \delta U X_0 = 0$ to first order in δU .*

Proof. Clearly, if $1 \in \sigma(\hat{U})$ and δU is small enough, it exists a vector $\hat{X} = X_0 + \delta X$, with $\|\delta X\| \leq C\|\delta U\|$, such that $\hat{U}\hat{X} = 1\hat{X}$. Then,

$$\begin{aligned} (U_0 + \delta U)(X_0 + \delta X) &= X_0 + \delta X \\ \Leftrightarrow U_0 \delta X + \delta U X_0 + \delta U \delta X &= \delta X. \end{aligned}$$

Because $U_0 X_0 = X_0$ and U_0 unitary, $\bar{X}_0^T U_0 = \bar{X}_0^T$, then multiplying on the left by $\bar{X}_0^T$ and keeping only first order terms, we get the desired condition:

$$\bar{X}_0^T \delta U X_0 = 0.$$

□

Because of the previous lemma, if $\hat{U} = U_0 + \delta U$ is a unitary perturbation of U_0 such that $\bar{X}_0^T \delta U X_0 \neq 0$, for any eigenvector X_0 with eigenvalue 1, then $1 \notin \sigma(\hat{U})$. Now if we consider a unitary perturbation $\hat{U}$ of U_0 such that $1 \notin \sigma(\hat{U})$ we want to estimate how far is 1 from the spectrum of $\hat{U}$. Consider the eigenvalue equation for the perturbed matrix. The perturbed eigenvalue $\hat{\lambda} = 1 + \delta\lambda$ will satisfy:

$$(U_0 + \delta U)(X_0 + \delta X) = (1 + \delta\lambda)(X_0 + \delta X) \quad (3.9)$$

Multiplying on the left by $\bar{X}_0^T$ and solving for $|\delta\lambda|$ it follows that,

$$|\delta\lambda| \geq \frac{|X_0^H \delta U X_0 + X_0^H \delta U \delta X|}{|1 + X_0^H \delta X|} \geq \frac{|X_0^H \delta U X_0| - |X_0^H \delta U \delta X|}{1 + \|\delta X\|}, \quad (3.10)$$

for $\|\delta U\|$ small enough. Considering the particular form of the matrix $U_0 = UD\bar{D}^{-1}$ we have that $\delta U = U\delta(D\bar{D}^{-1})$ and therefore $\|\delta U\| = \|\delta(D\bar{D}^{-1})\|$. Having into account that $\delta(D\bar{D}^{-1})$ is a diagonal matrix, in fact $\delta(D\bar{D}^{-1})_{kk} = \frac{-2i}{(h_k - i)^2} \delta h_k$, its singular values are the modulus of its eigenvalues, and we have the following proposition.

PROPOSITION 3.2. *Given the matrix $U_0 = UD\bar{D}^{-1}$ with eigenvalue 1, then 1 is not an eigenvalue of the perturbed matrix $U + \delta U = U_0(D\bar{D}^{-1} + \delta(D\bar{D}^{-1}))$ with $\|\delta(D\bar{D}^{-1})\|$ small enough, and the perturbation $\delta\lambda$ of such eigenvalue satisfies the following bound,*

$$|\delta\lambda| \geq \frac{\sigma_{\min}(\delta(D\bar{D}^{-1})) - C\sigma_{\max}^2(\delta(D\bar{D}^{-1}))}{1 + C\sigma_{\max}(\delta(D\bar{D}^{-1}))}.$$

Proof. Because of Lemma 3.1 it is sufficient to show that $\bar{X}_0^T \delta(D\bar{D}^{-1}) X_0 \neq 0$. But this is an easy consequence of the fact that $|\sum_{i=1}^{2n} (|X_{0,i}|^2 \delta(D\bar{D}^{-1}))_{ii}| \geq 2\delta h_{\min}$. Furthermore $\|\delta X\| \leq C\|\delta U\|$ and the bound follows from equation (3.10). $\square$

If we neglect terms $|h_i^2| \ll 1$ and $|\delta h_i| \ll |h_i|$ we finally get the desired bound for the condition number:

$$\kappa(F) \leq \frac{h_{\max}}{h_{\min}} \frac{1}{\delta h}, \quad (3.11)$$

where $\delta h = \min\{\delta h_\alpha\}$. Then, if for a given N we obtain a boundary matrix F which is bad conditioned, it suffices to change the size of the discretization, i.e., to increase N , to improve the condition number. Of course, if N is already quite large, then the bound (3.11) could be useless. For typical values $h \approx 10^{-2} \sim 10^{-3}$, it can be taken $\delta h \approx 10^{-4} \sim 10^{-5}$, to provide condition numbers $\kappa(F) \approx 10^4 \sim 10^5$.

3.3. The spectral pencil. For any $N > 2n$ we define the finite-dimensional approximation space S^N as the linear span of the of bulk and boundary functions, i.e., $S^N = \text{span}\{f_k^\alpha, \beta^l \mid \alpha = 1, \dots, n, k = 2, \dots, r_\alpha - 1, l = 1, \dots, 2n\}$. All functions $f_k^{(\alpha)}$ and $\beta^{(l)}$ are linearly independent, thus the dimension of S^N will be $|r| = r_1 + \dots + r_n$. It is convenient to arrange the basis above as follows:

$$\beta^{(1)}, f_1^{(1)}, \dots, f_{r_1}^{(1)}, \beta^{(2)}, \beta^{(3)}, f_1^{(2)}, \dots, f_{r_2}^{(2)}, \beta^{(4)}, \dots, \beta^{(2n-1)}, f_1^{(n)}, \dots, f_{r_n}^{(n)}, \beta^{(2n)}.$$

Using this ordering, we will rename the elements of this basis as f_a , with $a = 1, \dots, |r|$, and an arbitrary element $\Phi_N \in S^N$ will be written as $\Phi_N(x) = \sum_{a=1}^{|r|} \Phi_a f_a(x)$. We consider now the approximate eigenvalue problem:

$$\langle \Psi'_N, \Phi'_N \rangle - [\bar{\Psi}_N \Phi'_N]_{\partial M} + \langle \Psi_N, V \Phi_N \rangle - \lambda_N \langle \Psi_N, \Phi_N \rangle = 0, \quad \forall \Psi_N \in S^N. \quad (3.12)$$

Introducing the expansion above in the approximate problem (3.12) we get:

$$\sum_{a,b} \Psi_a \left[\langle f'_a(x), f'_b(x) \rangle + [\bar{f}_a(x) f'_b(x)]_{\partial M} + \langle f_a(x), V f_b(x) \rangle - \lambda \langle f_a(x), f_b(x) \rangle \right] \Phi_b = 0. \quad (3.13)$$

As (3.13) holds for every $\Psi_N \in S^N$, this equation is equivalent to the matrix pencil

$$A = \lambda B, \quad (3.14)$$

where $A_{ab} = \langle f'_a(x), f'_b(x) \rangle + [\bar{f}_a(x) f'_b(x)]_{\partial M} + \langle f_a(x), V f_b(x) \rangle$ and $B_{ab} = \langle f_a(x), f_b(x) \rangle$. Notice that A and B are hermitian matrices, what improves the algorithms for finding the pencil's solutions.

In fact, when solving numerically the spectral pencil (3.14), it is relevant to keep the pencil's hermitian character. Notice that the boundary functions $\beta^{(l)}(x)$ satisfy

$$[\bar{\beta}^{(l)}\beta^{(m)'}]_{\partial M} - [\beta^{(m)}\bar{\beta}^{(l)'}]_{\partial M} = 0, \quad (3.15)$$

because their boundary values are elements of a maximally isotropic subspace of the Lagrange boundary form. Using (3.5) and the definition of the boundary values of the boundary functions in terms of V we have that:

$$[\bar{\beta}^{(l)}\beta^{(m)'}]_{\partial M} = \sum_{k=1}^{2n} \frac{1}{h_k} \bar{V}_{kl} (V_{km} - \delta_{km}) = \sum_{k=1}^{2n} \frac{1}{h_k} \bar{V}_{kl} V_{km} - \frac{1}{h_m} \bar{V}_{ml}.$$

This identity together with equation (3.15) leads to

$$\frac{1}{h_j} \bar{V}_{ji} = \frac{1}{h_i} V_{ij}. \quad (3.16)$$

The hermiticity relation (3.16) is satisfied by the numerical solutions of (3.7) only up to the computing error and hence the pencil (3.14) is hermitian only up to this order. We will force the numerical solution of matrix V to satisfy (3.16) so that the hermiticity of the pencil is preserved exactly. This is convenient because the algorithms for solving the general eigenvalue problem are much better behaved in the hermitian case [De97].

To end this discussion we must realize that, with the basis f_a for S^N that we have just constructed, the matrices A and B are almost 3-diagonal and the unique elements different from zero, besides the 3-diagonal ones, are those related to the matrix elements of the boundary functions. In fact, we can consider a number of cases. If the function f_a is an interior bulk function, i.e., not corresponding to the nodes x_2^α nor $x_{r_\alpha-1}^{(\alpha)}$, it is obvious that the only nontrivial inner products $\langle f_a, f_b \rangle$ and $\langle f'_a, f'_b \rangle$ will correspond to $b = a - 1, a, a + 1$. If the function f_a is an extreme bulk function, for instance $f_2^{(\alpha)}$, then it has nontrivial inner products only with $\beta^{(2\alpha-1)}$ and $f_3^{(\alpha)}$. If f_a is now a boundary function $\beta^{(l)}$, then the only non-vanishing inner products will be with the other boundary functions and an extreme bulk function, namely $f_2^{(\alpha)}$ if $l = 2\alpha - 1$ or $f_{r_\alpha-1}^{(\alpha)}$ if $l = 2\alpha$.

3.4. Convergence of the numerical scheme. Now we discuss the convergence of the proposed numerical scheme. Consider first the following space of functions

$$\mathcal{D}_U = \left\{ \Phi \in \mathcal{C}^\infty(M) \mid \varphi - i\dot{\varphi} = U(\varphi + i\dot{\varphi}), \varphi = \Phi|_{\partial M}, \dot{\varphi} = \frac{d\Phi}{d\nu}|_{\partial M} \right\}.$$

THEOREM 3.3. *The union of all the finite dimensional spaces S^N is the closure of $\mathcal{D}_U$ in the Sobolev norm of order 1.*

$$\cup_{N>0} S^N = \overline{\mathcal{D}_U}^{\|\cdot\|_{\mathcal{H}^1}}.$$

Proof. The proof uses standard arguments in finite element theory. The family of subdivisions M^N of the manifold M defined by the integer N consists on collections of $N+n \leq |r|+n \leq N+2n$ closed subintervals $I_{\alpha,a}$, $a = 1, \dots, r_\alpha + 1$, of the real line. The domain K of the reference element is star-shaped (an interval again) and the family of functions $\mathcal{P}$ contains linear polynomials. Thus we are in the conditions of Thm. 4.4.4 of [Bre08] and we have the local estimate:

$$\|\Psi - \Psi_N\|_{\mathcal{H}^k(K)} \leq C(\text{diam}(K))^{m-k} \|\Psi\|_{\mathcal{H}^m(K)},$$

with $m > 1/2$. In our case we consider $m = 2$. The family of subdivisions M^N , consisting on the family of $r_\alpha + 1$ subintervals $I_{\alpha,a}$ of length h_α of the intervals I_α , satisfies that $\max\{\text{diam}(I_{\alpha,a} \mid I_{\alpha,a} \in M^N)\} \leq L/N =: h$. This family is also quasi-uniform, i.e., $\min\{\text{diam}(I_{\alpha,a} \mid I_{\alpha,a} \in M^N)\} \geq$

$\rho/(N+2)$, if $0 < \rho \leq \min L_\alpha$. Hence we can apply Thm. 4.4.20 of [Bre08] to our family of subdivisions with $m = 2$ and $s = 1$, to obtain:

$$\sum_{I_{\alpha,a} \in M^N} \|\Psi - \Psi_N\|_{\mathcal{H}^1(I_{\alpha,a})}^2 \leq Ch^2 \|\Psi\|_{\mathcal{H}^2(M)}^2,$$

for any $\Psi \in \mathcal{D}_U$, C a constant independent of N , Ψ_N the orthogonal projection of Ψ to S^N , $\|\cdot\|_{\mathcal{H}^k(M)}^2$ the Sobolev k -norm of Ψ and $h = L/N$ as before. Then we conclude

$$\|\Psi - \Psi_N\|_{\mathcal{H}^1(M)} \leq \frac{C}{N} \|\Psi\|_{\mathcal{H}^2(M)}. \quad (3.17)$$

The inequality above assures that for every $\Psi \in \mathcal{D}_U$ it exists a sequence $\{\Psi_N\}_{N=0}^\infty$ with $\Psi_N \in S_N$ such that $\lim_{N \rightarrow \infty} \|\Psi - \Psi_N\|_{\mathcal{H}^1(M)} = 0$ and therefore

$$\overline{\mathcal{D}_U}^{\|\cdot\|_{1,2}} \subset \cup_{N>0} S^N.$$

On the other hand we have from the Sobolev's imbedding theorems that

$$S^N \subset \overline{\mathcal{D}_U}^{\|\cdot\|_{1,2}}, \quad \forall N.$$

□

Lets now introduce the following notation for the bilinear form on the left hand side of the first equation (3.2):

$$Q(\Psi, \Phi) := \langle \Psi', \Phi' \rangle - [\bar{\Psi}\Phi']|_{\partial M} + \langle \Psi, V\Phi \rangle. \quad (3.18)$$

Notice that this bilinear form is hermitian, $Q(\Psi, \Phi) = \overline{Q(\Phi, \Psi)}$, if the boundary conditions (2.6) are satisfied.

LEMMA 3.4. *The bilinear form Q (3.18) with domain $\overline{\mathcal{D}_U}^{\|\cdot\|_{\mathcal{H}^1}}$ is semibounded with respect to the Sobolev norm of order 1, i.e. it exists $M \geq 0$ such that*

$$Q(\Phi, \Phi) \geq -M \|\Phi\|_{\mathcal{H}^1}^2.$$

Proof. The potential function V is bounded below by hypothesis and obviously $\langle \Phi', \Phi' \rangle \geq 0$, so it suffices to prove that $|\bar{\Psi}\Phi'|_{\partial M} \leq C \|\Phi\|_{\mathcal{H}^1}^2$. Notice that we can diagonalize the unitary matrix U as:

$$U = S^\dagger D S,$$

where $D = \text{diag}(-1, \dots, -1, z_1, \dots, z_r)$, and where z_i are complex numbers of modulus 1 different from -1 . Now notice that the boundary condition (2.6) can be rewritten having into account the diagonalization of U as $(I - D)S\varphi = i(I + D)S\dot{\varphi}$ or equivalently

$$\begin{bmatrix} 2 & \\ & Z_- \end{bmatrix} \begin{bmatrix} \varphi_0 \\ \varphi_1 \end{bmatrix} = i \begin{bmatrix} 0 & \\ & Z_+ \end{bmatrix} \begin{bmatrix} \dot{\varphi}_0 \\ \dot{\varphi}_1 \end{bmatrix},$$

where we have defined $Z_\pm := \text{diag}(1 \pm z_1, \dots, 1 \pm z_r)$ and

$$\begin{bmatrix} \varphi_0 \\ \varphi_1 \end{bmatrix} := S\varphi, \quad \begin{bmatrix} \dot{\varphi}_0 \\ \dot{\varphi}_1 \end{bmatrix} := S\dot{\varphi}$$

are just the boundary data adapted to the new basis of $\mathbb{C}^{2n}$. Note that this system of equations now implies that $\mathbb{C}^{2n-r} \ni \varphi_0 = 0$ and $\mathbb{C}^r \ni \dot{\varphi}_1 = -iZ_+^{-1}Z_- \varphi_1$. Moreover, as S is a unitary operator in the Hilbert space of the boundary, the inner product $\langle \varphi, \dot{\varphi} \rangle$ verifies

$$\langle \varphi, \dot{\varphi} \rangle = -i \langle \varphi_1, Z_+^{-1}Z_- \varphi_1 \rangle,$$

and then

$$|\langle \varphi, \dot{\varphi} \rangle| \leq \|Z_- Z_+^{-1}\| \|\varphi\|^2 \leq C \|\Phi\|_{\mathcal{H}^1}^2, \quad (3.19)$$

where the last inequality is a particular case of the boundary trace inequalities [Ad75] taking $\varphi \in L^2(\partial M)$.

It can happen that $\{-1\} \notin \sigma(U)$, then $(I + U)$ is invertible and equation (3.19) still holds with $Z_{\pm} = I \pm U$. Therefore the quadratic form is semibounded with respect to the Sobolev norm of order 1. $\square$

Now the convergence of the eigenvalues and eigenvectors follows by standard arguments.

THEOREM 3.5. *The solutions (Φ_N, λ_N) of the approximate eigenvalue problem (3.12) converge in the limit $N \rightarrow \infty$ to solutions (Φ, λ) of the weak problem (3.2).*

Proof. Lemma 3.4 assures that the *min-max* principle [Re75] applies for the quadratic form (3.18) and the convergence of the eigenvalues is proved as follows.

Applying the *min-max* principle to the quadratic forms (3.18) of the weak and the approximate problem and subtracting them we get:

$$\sup_{\varphi_1, \dots, \varphi_{n-1}} \left[\inf_{\substack{\Phi_N \in [\varphi_1, \dots, \varphi_{n-1}]^\perp \\ \Phi_N \in S^N}} \frac{Q(\Phi_N, \Phi_N)}{\|\Phi_N\|} \right] - \sup_{\varphi_1, \dots, \varphi_{n-1}} \left[\inf_{\substack{\Phi \in [\varphi_1, \dots, \varphi_{n-1}]^\perp \\ \Phi \in \mathcal{D}_U}} \frac{Q(\Phi, \Phi)}{\|\Phi\|} \right] = \lambda_N - \lambda. \quad (3.20)$$

The left hand side tends to zero in the limit $N \rightarrow \infty$ by theorem 3.3 and therefore

$$\lim_{N \rightarrow \infty} \lambda_N = \lambda. \quad (3.21)$$

Now let $\Phi_N \in \mathcal{H}^1(M) \subset L^2(M)$ be a solution of (3.2), i.e.

$$Q(\Psi_N, \Phi_N) - \lambda_N \langle \Psi_N, \Phi_N \rangle = 0, \quad \forall \Psi_N \in S^N.$$

We can assume that $\|\Phi_N\|_{\mathcal{H}^1} = 1$ for every N and therefore the Banach–Alaoglu Theorem assures that subsequences $\{\Phi_{N_j}\}$ and accumulation points Φ exist, such that for every $\Psi \in \overline{\mathcal{D}_U}^{\|\cdot\|_{\mathcal{H}^1}}$

$$\lim_{N_j \rightarrow \infty} |\langle \Psi, \Phi - \Phi_{N_j} \rangle_{\mathcal{H}^1}| = 0. \quad (3.22)$$

If we denote by $P_N \Psi$ the orthogonal projection of Ψ into S^N , we have that for every $\Psi \in \overline{\mathcal{D}_U}^{\|\cdot\|_{\mathcal{H}^1}}$

$$\begin{aligned} |Q(\Psi, \Phi) - \lambda \langle \Psi, \Phi \rangle| &\leq |Q(\Psi, \Phi - \Phi_{N_j})| + |\lambda| |\langle \Psi, \Phi - \Phi_{N_j} \rangle| + |\lambda_{N_j} - \lambda| |\langle \Psi, \Phi_{N_j} \rangle| + \dots \\ &\quad \dots + |Q(\Psi - P_N \Psi, \Phi_{N_j})| + |\lambda_{N_j}| |\langle \Psi - P_N \Psi, \Phi_{N_j} \rangle|. \end{aligned}$$

Now (3.17), (3.21) and (3.22) assure that the right hand side vanishes in the limit $N_j \rightarrow \infty$ and we get that the accumulation points of the sequence $\{\Phi_N\}$ are eigenvectors of the weak eigenvalue problem. $\square$

4. Numerical experiments and conclusions. By now we have proved that the procedure described in section 3 results in the finite dimensional eigenvalue problem of the matrix pencil (3.14). Furthermore we have a bound (3.17) on the error committed with this discretization. Now, if we are able to solve the pencil for increasing grid size, we will get better and better approximations to the eigenvalue problem. As remarked in the previous section, the form of the pencil is almost 3-diagonal and both matrices A and B are hermitian. Hence, the resulting problem is algebraically well behaved. We will discuss now some numerical experiments for two particular situations, the free particle in $M = [0, 2\pi]$ ($V(x) \equiv 0$) and the harmonic oscillator in $M = [-\pi, \pi]$ ($V(x) = x^2$). We will use these cases to show results concerning the stability of the procedure and also to show that the bound (3.17) is verified. Furthermore, we will compare the results obtained with this procedure with the ones obtained by other two algorithms which are not based in the finite elements method for a particular choice of boundary conditions close to the singular case described in [Be08]. Notice that in this simple case ($n = 1$) there is only one interval and we have

that $r = N + 1$, so we can use r and N almost interchangeably if N is large enough, what indeed is the actual case. The number $h = \frac{2\pi}{r+1}$ is going to be the length of each subinterval.

The Schrödinger eigenvalue problems that we are going to address for different self-adjoint extensions are the quantum spinless free particle

$$-\frac{d^2}{dx^2}\Psi = \lambda\Psi$$

with associated pencil, eq. (3.14), $L = \lambda B$, and the quantum harmonic oscillator

$$-\frac{1}{2}\frac{d^2}{dx^2}\Psi + \frac{1}{2}x^2\Psi = \lambda\Psi,$$

with associated pencil $A = \frac{1}{2}L + \frac{1}{2}O = \lambda B$, where the matrices L , B and O are given, after straightforward computations, by:

$$\begin{aligned} l_{kk} &= 2, \quad k = 2, \dots, r-1 \\ l_{kk+1} &= l_{k+1k} = -1, \quad k = 1, \dots, r-1 \\ l_{11} &= 2 - V_{11} \\ l_{rr} &= 2 - V_{22} \\ l_{1r} &= -V_{12} \\ l_{r1} &= \overline{l_{1r}} = -V_{21} \\ l_{ij} &= 0, \quad j \geq i+2, \quad j \leq i-2 \\ L_{ij} &= \frac{1}{h}l_{aij} \end{aligned} \tag{4.1a}$$

$$\begin{aligned} b_{kk} &= 4, \quad k = 2, \dots, r-1 \\ b_{kk+1} &= b_{k+1k} = 1, \quad k = 1, \dots, r-1 \\ b_{11} &= 4 + 2[|V_{11}|^2 + |V_{21}|^2] + 2V_{11} \\ b_{rr} &= 4 + 2[|V_{22}|^2 + |V_{12}|^2] + 2V_{22} \\ b_{1r} &= 2[\bar{V}_{11}V_{12} + \bar{V}_{21}V_{22}] + 2 \\ b_{r1} &= \bar{b}_{1r} \\ b_{ij} &= 0, \quad j \geq i+2, \quad j \leq i-2 \\ B_{ij} &= \frac{h}{6}b_{ij} \end{aligned} \tag{4.1b}$$

$$\begin{aligned} o_{kk} &= 20r^2 - 80rk + 80k^2 - 80k + 40r + 28, \quad k = 2, \dots, r-1 \\ o_{kk+1} &= o_{k+1k} = 5r^2 - 20rk + 20k^2 + 1, \quad k = 1, \dots, r-1 \\ o_{11} &= (\bar{V}_{11}V_{11} + \bar{V}_{21}V_{21})(10r^2 + 10r + 4) + (\bar{V}_{11} + V_{11})(1 + 5r^2) + 20r^2 - 40r + 28 \\ o_{rr} &= (\bar{V}_{12}V_{12} + \bar{V}_{22}V_{22})(10r^2 + 10r + 4) + (\bar{V}_{22} + V_{22})(1 + 5r^2) + 20r^2 - 40r + 28 \\ o_{1r} &= (\bar{V}_{21} + V_{12})(1 + 5r^2) + (\bar{V}_{21}V_{22} + \bar{V}_{11}V_{12})(10r^2 + 10r + 4) \\ o_{r1} &= \overline{o_{1r}} \\ o_{ij} &= 0, \quad j \geq i+2, \quad j \leq i-2 \\ O_{ij} &= \frac{h^3}{120}c_{ij} \end{aligned} \tag{4.1c}$$

One needs, in each case, to obtain the matrix V to solve previously the corresponding system of equations (3.7) for each self-adjoint extension. The solutions of these generalized eigenvalue problems have been obtained using the *Octave* built-in function **eig**. The details of this routine can be found in [An92].

In order to check if the bound (3.17) is verified we have to compare the numerical solution obtained solving the generalized eigenvalue problem with the exact solution. Therefore we will solve the Dirichlet Problem for the free particle, whose solutions are $\Phi(x) = \sin \alpha x$. In this particular case the matrix V solution of (3.6) is the zero matrix. Note that this forces the boundary values to be all zero, as it is supposed to be in the Dirichlet Problem. We will now compute $\|\Phi - \Phi_N\|_{\mathcal{H}^1}$ for increasing values of N . For calculating it we have performed a sum over

all subintervals, knowing that at each subinterval the approximate solution is a linear polynomial. In figure 4.1 we can see the $\mathcal{H}^1$ -norm of the difference between the analytic solution for the groundstate of the Dirichlet Problem, i.e. $\sin x/2$, and the computed solution plotted against the size of the discretization.

Fig. 4.1: Evolution of the error, measured in the $\mathcal{H}^1$ -norm, for the groundstate of the Dirichlet Problem with increasing lattice size.

After a linear fit of the logarithms of this data, one observes that they can be adjusted to a linear polynomial with slope $p = -1.002 \pm 0.001$, what indeed shows that

$$\|\Phi - \Phi_N\|_{\mathcal{H}^1} = K(r+1)^p \leq \frac{K}{r+1} \leq \frac{K}{N},$$

therefore satisfying the theoretical bound (3.17).

Our next step will be to give a test of the stability of the method against variations of the input parameters. The unique parameter that this procedure has is the self-adjoint extension whose eigenvalue problem we want to solve. We will perturb an initial self-adjoint extension, described by a unitary matrix U , and observe the behavior of the relative error in the eigenvalues. In other words, we are interested now in studying the relation

$$\frac{|\delta\lambda|}{|\lambda|} = K(\varepsilon) \frac{\|\delta U\|}{\|U\|} \quad (4.2)$$

and we would expect, if the algorithm were stable, that the condition number $K(\varepsilon)$ grows at most polynomially with the perturbation ε . However one must be very careful in this process since the original eigenvalue problem presents divergences in certain circumstances (explained below) which could lead to the wrong conclusion that the method isn't stable. As a consequence of lemma 3.4, we see that when a self-adjoint extension is parameterized by a unitary matrix U that has eigenvalues close but not equal to -1 , it happens that some eigenvalues of the considered problem take very large negative values. However, matrices with -1 in the spectrum can lead to self-adjoint extensions that are positive definite, for example the Dirichlet self-adjoint extension. Thus, following a path in the space of self-adjoint extensions, it could happen that an infinitesimal change in the arc parameter led to an infinite jump in the eigenvalues. Such self-adjoint extensions

are precisely the ones that lead in higher dimensions to the problem identified by M. Berry as a *Dirichlet singularity*² [Be08],[Be09], [Ma09] and they will be the target of our latter tests. For proving the stability it is therefore necessary to perturb the unitary matrix along a direction of its tangent space such that this situation is avoided. A path in the space of self-adjoint extensions where these jumps don't occur is, for instance, the one described by the so called "quasi-periodic boundary conditions" [As83]. In this case, the self-adjoint domain is described by functions that satisfy the following boundary conditions $u(0) = e^{i\theta}u(2\pi)$ and $u'(0) = e^{i\theta}u'(2\pi)$, that correspond to the unitary matrix

$$U(\theta) = \begin{pmatrix} 0 & e^{i\theta} \\ e^{-i\theta} & 0 \end{pmatrix}.$$

Notice that this particular choice of boundary conditions, that are non-local in the sense that they mix the boundary data at both endpoints, is naturally treatable by the discretization procedure introduced in section 3 and that goes beyond the ones usually addressed, described by equations of the form (1.2).

Let us then consider perturbations of the periodic case in the quasi-periodic direction, i.e. we consider

$$U(\varepsilon) = U + i\varepsilon A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + i\varepsilon \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We have calculated the numerical solutions for values of ε between 10^{-5} and 10^{-3} in steps of 10^{-5} and the discretization size used was $N = 250$. In this case the relative perturbation is $\|i\varepsilon A\|/\|U\| = \varepsilon$, hence the relative error ratio is

$$K(\varepsilon) = \frac{\frac{|\delta\lambda|}{|\lambda|}}{\frac{\|\delta U\|}{\|U\|}} = \frac{1}{\varepsilon} \frac{|\delta\lambda|}{|\lambda|}.$$

The results are plotted in figure 4.2. As can be seen, $K(\varepsilon)$ grows slowly with ε . Fits of the resulting data show that, in every case, the error ratios satisfy power law growths of type $K(\varepsilon) = a \cdot \varepsilon^b + c$ with exponents -0.89 ± 0.02 , -0.42 ± 0.03 , -0.03 ± 0.03 , 0.28 ± 0.03 for the 1st, 2nd, 3rd and 4rd excited states respectively. Although the groundstate error ratio has linear growth, and this certainly isn't a bad behavior in the sense that it is polynomial, the result is not trustable because the quotient $|\delta\lambda|/|\lambda|$ is bad conditioned. This is due the fact that the groundlevel is $\lambda_0 = 0$ in the periodic case.

As stated at the beginning of this section, we are going to compare now the results obtained with the method proposed in section 3, from now on referred as FEM, with other two that are not based in the finite element method, namely the Line-based Perturbation Method (LPM) and the Constant reference potential Perturbation Method (CPM) proposed in [Le06] and [Le05] respectively. The software used for the calculations is, in the latter cases, a ready-to-run version provided by the authors³. The objective is now to find the solutions for the two following Schrödinger problems.

$$\text{Free Particle} \quad \begin{cases} -\frac{d^2}{dx^2}\Psi(x) = \lambda\Psi(x), & x \in [0, 2\pi] \\ u'(0) = 0 \\ \tan(-\theta/2) \cdot u(2\pi) + u'(2\pi) = 0 \end{cases} \quad (4.3a)$$

$$\text{Harmonic Oscillator} \quad \begin{cases} -\frac{1}{2}\frac{d^2}{dx^2}\Psi(x) + \frac{1}{2}x^2\Psi(x) = \lambda\Psi(x), & x \in [-\pi, \pi] \\ u'(-\pi) = 0 \\ \tan(-\theta/2) \cdot u(\pi) + u'(\pi) = 0. \end{cases} \quad (4.3b)$$

²Notice again that $U = -1$ is the unitary matrix describing Dirichlet boundary conditions.

³<http://www.twi.ugent.be>

Fig. 4.2: Relative error ratio $K(\varepsilon)$ of the firsts excited levels for the periodic boundary problem plotted against ε .

Both boundary conditions selected for the problems (4.3) are determined by the unitary matrix

$$U = \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

and are of the type described above in the limit $\theta \rightarrow \pi$. When $\theta = \pi$ both operators are positive, but for values of $\theta < \pi$ the groundlevel is negative with increasing absolute value for $\theta \rightarrow \pi$. The LPM routine reported errors for the lowest order eigenvalues for $\theta \geq 2.6$ and was not able to produce any output to compare. The results comparing the CPM and FEM routines for the free particle are shown in the figures 4.3 to 4.5 and the ones for the harmonic oscillator in the figures 4.6 to 4.8. The discretization size used for all the FEM calculations was $r = 5000$. Both methods provide almost the same results for the first excited states (indices 1 to 5) of the free particle, however this is not the case for the groundlevel (fig. 4.3) where it is expected that the eigenvalues take increasingly large negative values for θ approaching π . Note that the lowest value achieved with FEM is $\lambda_0 \simeq -7 \cdot 10^4$. Clearly the CPM routine fails to reproduce the correct eigenvalues in this situation, although it provides good approximations for $\theta \leq 0.989\pi$. In all these cases, the CPM routine used, warns of the low reliability of the results. Despite the numerical eigenvalues are in agreement with the expected ones, except for the groundlevel, the eigenfunctions plotted in figures 4.4 and 4.5 for the case $\theta = 3.1$ show clearly that the solutions obtained with the FEM are more accurate. For instance, the solutions obtained with CPM present discontinuities near the middle of the interval and they are expected to be smooth functions. The eigenfunction obtained for the groundlevel (fig 4.5) is specially remarkable. In this case the FEM solution is a clear example of an *edge state*. These are eigenfunctions that are associated with negative eigenvalues and that are strongly localized at the boundary of the system while they vanish in the interior of the manifold (the interior of the interval in this case). These *edge states* are important in the understanding of certain physical phenomena like the quantum Hall effect (see [Xi92] and references therein). The qualitative results for the quantum harmonic oscillator (fig. 4.6-4.8) comparing both methods are more or less the same that for the free particle with the exception that in this case the CPM routine provides groundlevel energies, for the values of θ explored, that are in agreement with the ones obtained by FEM and that therefore present the correct behavior. However, the CPM routine also warns of the low reliability of this results and, as can be checked

for the eigenvalues plotted in the figures 4.7 and 4.8 for the case $\theta = 3.135$, the approximation of the eigenfunctions is still bad behaved near the center of the interval. As before, the FEM produces in this case an *edge state* for the eigenfunction of the groundlevel (fig. 4.8).

Summarizing we can conclude that the algorithm introduced in this article is able to implement any, even non-local, self-adjoint extension of the 1D Schrödinger problem in a straightforward way. Moreover, due to the convergence and conditioning theorems obtained in sections 3.2 and 3.4 the exact solutions can be achieved even if the original problem is close to being singular. A feature that we leave for subsequent works is that the main ideas and theorems that appear in this article are well suited for the implementation of the algorithm in higher dimensions.

REFERENCES

- [Ac09] L. Aceto, P. Ghelardoni, C. Magherini. *Boundary Value Methods as an extension of Numerov's method for Sturm-Liouville eigenvalue estimates*. Appl. Num. Math. 59, 1644-1656 (2009)
- [Ad75] R.A. Adams. *Sobolev Spaces*. Academic Press NY (1975).
- [An92] E. Anderson, Z. Bai, et. al. *LAPACK users' guide*. SIAM (1992).
- [As83] M. Asorey, J.G. Esteve, A.F. Pacheco. *Planar rotor: The θ -vacuum structure, and some approximate methods in quantum mechanics*. Phys. Rev. D, **27**, 1852-68 (1982).
- [As05] M. Asorey, A. Ibort, G. Marmo. *Global Theory of Quantum Boundary Conditions and Topology Change*. Int. J. Mod. Phys. A, **20** 1001-25 (2005).
- [Be08] M. V. Berry, M.R. Dennis. *Boundary-condition-varying circle billiards and gratings: the Dirichlet singularity*. J. Phys. A: Math. Theor., **41** 135203 (2008).
- [Be09] M.V. Berry. *Hermitian boundary conditions at a Dirichlet singularity: the Marletta-Rozenblum model*. J. Phys., A: Math. Theor., **42** 165208 (2009).
- [Br08] J. Brüning, V. Geyler, K. Pankrashkin. *Spectra of self-adjoint extensions and applications to solvable Schrödinger operators*. Rev. Math. Phys., **20**, 1-70 (2008).
- [Bre08] S.C. Brenner, L. Ridgway Scott *The Mathematical Theory of Finite Element Methods* Second Edition Springer (2008).
- [Ch99] B. Chanane. *Computing eigenvalues of regular Sturm-Liouville Problems*. Appl. Math. Lett. **12**, 119-125 (1999).
- [Ch09] B. Chanane *Computing the eigenvalues of a class of nonlocal Sturm-Liouville problems*. Math. Comput. Modelling **50**, 225-232 (2009).
- [Cl10] S. Clark, F. Gesztesy, M. Mitrea. *Boundary data maps for Schrödinger operators on a compact interval* (2010). [arXiv:1012.0606v1](#)
- [De97] J. W. Demmel, *Applied Numerical Linear Algebra* . SIAM (1997).
- [Fa57] *On the numerical Solution of Sturm-Liouville Differential Equations*. Mathematical Tables and other Aids to Computation **11**, 131-150 (1957).
- [Gr68] G. Grubb. *A characterization of the non local boundary value problems associated with an elliptic operator*. Ann. Scuola Norm. Sup. Pisa, **22**, 425-513 (1968).
- [Ho94] R. A. Horn, C. R. Johnson. *Topics in Matrix Analysis*. CUP (1994).
- [Le05] V. Ledoux, M. Van Daele, G. Vanden Berghe. *MATSLISE, A matlab Package for the numerical solution of Sturm-Liouville and Schrodinger equations*. ACM Trans. on Math. Soft., **31**, 532-554 (2005).
- [Le06] V. Ledoux, M. Rizea, et. al. *Solution of the Schrodinger equation by a high order perturbation method based on a linear reference potential*. Comp. Phys. Com. **175**, 424-439 (2006).
- [Ll08] F. Lledó, O. Post. *Existence of spectral gaps, covering manifolds and residually finite groups*. Rev. Math. Phys., **20**, 199-231 (2008).
- [Ma09] M. Marletta, G. Rozenblum. *A Laplace operator with boundary conditions singular at one point*. J. Phys. A: Math. Theor. **42** 125204 (2009).
- [Pr73] S. Pruess. *Estimating the eigenvalues of Sturm-Liouville Problems by approximating the differential equation*. SIAM J. Numer. Anal. **10**, 55-68 (1973).
- [Re75] M. Reed, B. Simon *Methods of Modern Mathematical Physics, vols. I- IV*. New York. Academic Press (1975).
- [Si00] B. Simon. *Schrödinger operators in the twentieth century*. J. Math. Phys., **41**, 3523-56 (2000).
- [We80] J. Weidman. *Linear Operators in Hilbert spaces*. Springer-Verlag, NY (1980).
- [Xi92] X.G. Wen. *Theory of the edge states in fractional quantum hall effects*. Int. J. Mod. Phys. B, **6**, 1711 (1992).

Fig. 4.3: Groundlevel energies of the free particle problem for increasing values of θ .

Fig. 4.4: First excited eigenfunctions of the free particle problem for $\theta = 3.1$.

Fig. 4.5: Groundlevel eigenfunctions of the free particle problem for $\theta = 3.1$.

Fig. 4.6: Groundlevel energies of the harmonic oscillator problem for increasing values of θ .

Fig. 4.7: First excited eigenfunctions of the harmonic oscillator problem for $\theta = 3.135$.

Fig. 4.8: Groundlevel eigenfunctions of the harmonic oscillator problem for $\theta = 3.135$.