

A Short Theory of the Rayleigh–Ritz Method

Harry Yserentant

Dedicated to Rolf Dieter Grigorieff on occasion of his 75th birthday

Abstract — We present some new error estimates for the eigenvalues and eigenfunctions obtained by the Rayleigh–Ritz method, the common variational method to solve eigenproblems. The errors are bounded in terms of the error of the best approximation of the eigenfunction under consideration by functions in the ansatz space. In contrast to the classical theory, the approximation error of eigenfunctions other than the given one does not enter into these estimates. The estimates are based on a bound for the norm of a certain projection operator, e.g., in finite element methods for second order eigenvalue problems, the H^1 -norm of the L_2 -projection onto the finite element space.

2010 Mathematical subject classification: 65N25, 65N15, 65N30.

Keywords: Eigenvalues, Eigenfunctions, Error Estimates.

1. Introduction

The Rayleigh–Ritz method is a variational method to solve the eigenvalue problem for elliptic differential operators, that is, to compute their eigenvalues and the corresponding eigenfunctions. It is the direct counterpart of the Ritz method for the solution of the assigned boundary value problems. The Rayleigh–Ritz method has the advantage of being based on minimal, very general assumptions and produces optimal solutions in terms of the approximation properties of the underlying trial spaces. The theory of the Rayleigh–Ritz method has to a large extent been developed in the context of finite element methods, see [1, 2, 14]. In a more recent paper [12], Knyazev and Osborn derived error estimates into which, in contrast to the older theory, only the best approximation error of the eigenfunction under consideration enters. Here we show that such estimates can be derived in a very simple way utilizing the stability of certain projection operators, in the case of second order problems, the H^1 -stability of the L_2 -projection onto the given ansatz space.

Error estimates of this type are important for adaptive methods since it is obviously much simpler to adapt a mesh to a single eigenfunction than to a whole invariant subspace. A second, more subtle reason is that the regularity of the eigenfunctions often considerably differs and that, for example, approximating the eigenfunction for the minimum eigenvalue can be much more difficult than for the second one. The most prominent example of this type is the eigenfunctions of the Hamilton operator of the hydrogen atom whose knowledge is basic for our understanding of chemistry. These eigenfunctions are classified by three numbers, the principal quantum number $n = 1, 2, \dots$, the angular momentum quantum number $\ell = 0, \dots, n - 1$, and the magnetic quantum number $m = -\ell, \dots, \ell$. The eigenvalues

Harry Yserentant

Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany

E-mail: yserentant@math.tu-berlin.de.

$\lambda = -1/(2n^2)$ depend only on the principal quantum number. They are highly degenerate for larger n ; their multiplicity is n^2 . The regularity of the eigenfunctions is determined by the angular momentum quantum number ℓ . They are contained in the Sobolev spaces H^s for $s < \ell + 5/2$, but not for $s = \ell + 5/2$, see [15]. Similar phenomena occur near reentrant edges and corners. Consider, for example, the eigenvalue problem for the negative Laplace operator on the unit disk, centered at the origin, with a slit along the positive x -axis and Dirichlet boundary conditions on the circular line and the top of the slit and Neumann boundary conditions on its bottom. The eigenvalues and eigenfunctions are in this case characterized by integers $k = 0, 1, \dots$ and $\ell = 1, 2, \dots$ and read

$$\lambda_{k,\ell} = (\omega_{\alpha,\ell})^2, \quad u_{k,\ell}(r, \varphi) = \sin(\alpha\varphi)J_\alpha(\omega_{\alpha,\ell}r)$$

in polar coordinate representation. The functions J_α are the Bessel functions of first kind and fractional order $\alpha = (2k + 1)/4$. They behave near the origin like $\sim r^\alpha$. The constants $\omega_{\alpha,\ell}$ are the zeroes of these Bessel functions in increasing order. This problem has been used in [13] as test example for the adaptive solution of eigenproblems.

2. The Rayleigh–Ritz Method: Framework and Assumptions

We start from the usual abstract framework with two real Hilbert spaces \mathcal{H}_0 and $\mathcal{H}_1 \subseteq \mathcal{H}_0$ and a symmetric, coercive, and bounded bilinear form $a : \mathcal{H}_1 \times \mathcal{H}_1 \rightarrow \mathbb{R}$. The inner product on \mathcal{H}_0 is denoted by (u, v) and the induced norm by $\|u\|_0$. We equip the space \mathcal{H}_1 for simplicity with the energy norm $\|u\|$ induced by the bilinear form $a(u, v)$. For convenience we assume that \mathcal{H}_1 is compactly embedded into \mathcal{H}_0 and that both spaces are infinite dimensional. Then there exists an infinite sequence $0 < \lambda_1 \leq \lambda_2 \leq \dots$ of eigenvalues of finite multiplicity tending to infinity and an assigned sequence of eigenvectors u_1, u_2, \dots in \mathcal{H}_1 for which

$$(u_k, u_\ell) = \delta_{k\ell}, \quad a(u_k, u_\ell) = \lambda_k \delta_{k\ell}.$$

The example that we have in mind are second order elliptic eigenvalue problems over bounded domains Ω . In this case, $\mathcal{H}_0 = L_2(\Omega)$, and \mathcal{H}_1 is a subspace of the Sobolev space $H^1(\Omega)$, depending on the boundary conditions. The central results of this paper are, however, of more general nature and also apply, for example, to eigenfunctions for eigenvalues below the essential spectrum. Such situations arise in quantum chemistry.

The aim is to approximate the eigenvalues λ_k and the vectors in the assigned eigenspaces. For this, one chooses an n -dimensional subspace \mathcal{S} of \mathcal{H}_1 . Then there exist discrete eigenvectors u'_1, u'_2, \dots, u'_n in \mathcal{S} for eigenvalues $0 < \lambda'_1 \leq \lambda'_2 \leq \dots \leq \lambda'_n$, satisfying the relations

$$(u'_k, u'_\ell) = \delta_{k\ell}, \quad a(u'_k, u'_\ell) = \lambda'_k \delta_{k\ell}.$$

As will be shown, the discrete eigenvalues λ'_k approximate then the original eigenvalues λ and the discrete eigenvectors u'_k the corresponding eigenvectors u in a sense explained later. This already fixes the method, which replicates the weak form

$$a(u, v) = \lambda(u, v), \quad v \in \mathcal{H}_1,$$

of the original eigenvalue problem and is determined by the choice of the subspace \mathcal{S} replacing its solution space \mathcal{H}_1 . Typical approximation spaces \mathcal{S} are finite element spaces.

We will measure the approximation properties of the chosen subspace \mathcal{S} in terms of the a -orthogonal projection operator $P : \mathcal{H}_1 \rightarrow \mathcal{S}$ defined by

$$a(Pu, v) = a(u, v), \quad v \in \mathcal{S}.$$

With respect to the energy norm the projection Pu is the best approximation of $u \in \mathcal{H}_1$ by an element of \mathcal{S} , which means that for all $v \in \mathcal{S}$

$$\|u - Pu\| \leq \|u - v\|.$$

Our main assumption is that the correspondingly defined \mathcal{H}_0 -orthogonal projection Q from \mathcal{H}_0 onto \mathcal{S} is stable in the energy norm, that is, that there exists a constant κ with

$$\|Qv\| \leq \kappa \|v\|, \quad v \in \mathcal{H}_1. \quad (2.1)$$

The constant κ must be independent of hidden discretization parameters. This does not follow from the general assumptions made above but holds in many important cases. Examples are spectral methods in which the approximation spaces \mathcal{S} are built up from eigenfunctions of a nearby eigenvalue problem, say, in the case of a second order problem, from eigenfunctions of the Laplace operator. In this case the \mathcal{H}_0 - respectively L_2 -orthogonal projection coincides with the H^1 -orthogonal projection and satisfies the condition by definition.

The finite element case is more complicated. For quasi-uniform meshes estimates of this type can be traced back to the seminal paper [3] of Bank and Dupont. Consider a second-order problem, let $\mathcal{H}_0 = L_2(\Omega)$, and assume that $\mathcal{H}_1 \subseteq H^1(\Omega)$. Let the functions in the finite element space satisfy an inverse inequality

$$|v|_{H^1} \lesssim h^{-1} \|v\|_{L_2}, \quad v \in \mathcal{S}. \quad (2.2)$$

Let $\Pi : \mathcal{H}_1 \rightarrow \mathcal{S}$ be a quasi-interpolation operator for which

$$\|u - \Pi u\|_{L_2} \lesssim h |u|_{H^1}, \quad |\Pi u|_{H^1} \lesssim |u|_{H^1} \quad (2.3)$$

holds for all functions u in the solution space \mathcal{H}_1 . Then

$$|Q(u - \Pi u)|_{H^1} \lesssim h^{-1} \|Q(u - \Pi u)\|_{L_2} \leq h^{-1} \|u - \Pi u\|_{L_2} \lesssim |u|_{H^1}$$

for all these functions u . Since $Q\Pi u = \Pi u$, this implies the desired estimate

$$|Qu|_{H^1} \lesssim |Q(u - \Pi u)|_{H^1} + |\Pi u|_{H^1} \lesssim |u|_{H^1}.$$

Unfortunately, the h in the inverse inequality (2.2) behaves like the minimum element diameter and the h in (2.3) like the maximum element diameter. If the ratio of the maximum and the minimum element diameter becomes large, the proof breaks down.

The H^1 -stability of the L_2 -projection can, however, actually be proven under much weaker conditions on the grids. Partly rather technical sufficient conditions for the H^1 -stability of the L_2 -projection are given in [7–10]. In [11], the case of piecewise linear elements in two space dimensions is treated. The most comprehensive result of this type is due to Bank and Yserentant [5]. It is shown in this paper that the L_2 -projection remains H^1 -stable for example for the highly nonuniform grids that are generated by bisection-like refinement schemes like the red-green refinement in two [4] and three [6] space dimensions and Lagrangian type finite elements of polynomial order up to twelve in two and up to seven in three space dimensions.

For realistic meshes and grid sizes changing in the average, on larger scales less abruptly than in an extreme case possible with such refinement schemes, the argumentation from [5] works without any condition to the polynomial order of the finite elements or any further condition to the triangulation up to shape regularity. The results in [5] even cover the case of hp -like finite element methods as long as the maximum order of the finite elements remains limited. There is therefore less doubt that the L_2 -projection remains H^1 -stable in almost all cases of practical interest so that (2.1) is a reasonable assumption at least for finite element discretizations of second order elliptic eigenvalue problems.

3. The Error Estimates

We first turn to the approximation of the eigenfunctions or, in the abstract setting, the eigenvectors. Starting point of our considerations is the following representation of the error between an eigenvector $u \in \mathcal{H}_1$ for an eigenvalue λ and its best \mathcal{H}_0 -approximation by a linear combination of discrete eigenvectors u'_k for eigenvalues λ'_k in a given neighborhood Λ of λ .

Lemma 3.1. *The error between an eigenvector $u \in \mathcal{H}_1$ of the original problem for the eigenvalue λ and its \mathcal{H}_0 -orthogonal projection onto the space spanned by the discrete eigenvectors $u'_k \in \mathcal{S}$ for the eigenvalues λ'_k in a given neighborhood Λ of λ possesses the representation*

$$u - \sum_{\lambda'_k \in \Lambda} (u, u'_k) u'_k = R(u - Pu) + (I - Q)(u - Pu), \quad (3.1)$$

where the mapping $R : \mathcal{H}_0 \rightarrow \mathcal{S}$ is defined by the expression

$$Rf = \sum_{\lambda'_k \notin \Lambda} \frac{\lambda'_k}{\lambda'_k - \lambda} (f, u'_k) u'_k.$$

Proof. We first represent the expression on the left-hand side of (3.1) in the form

$$u - \sum_{\lambda'_k \in \Lambda} (u, u'_k) u'_k = \sum_{\lambda'_k \notin \Lambda} (u, u'_k) u'_k + u - \sum_{k=1}^n (u, u'_k) u'_k,$$

and replace the inner products in the first sum on the right-hand side by

$$(u, u'_k) = \frac{\lambda'_k}{\lambda'_k - \lambda} (u - Pu, u'_k).$$

This is possible as u is an eigenvector and the u'_k are discrete eigenvectors; therefore

$$\lambda(u, u'_k) = a(u, u'_k) = a(Pu, u'_k) = \lambda'_k (Pu, u'_k).$$

Since the \mathcal{H}_0 -orthogonal projection can be written as eigenvector expansion

$$Qf = \sum_{k=1}^n (f, u'_k) u'_k,$$

the error representation

$$u - \sum_{\lambda'_k \in \Lambda} (u, u'_k) u'_k = R(u - Pu) + (u - Qu)$$

follows. As $QPu = Pu$, this yields (3.1). □

A first consequence of the error representation (3.1) is the following error estimate in the \mathcal{H}_0 -norm that is still independent of the \mathcal{H}_1 -stability (2.1) of the \mathcal{H}_0 -orthogonal projection Q .

Theorem 3.2. *Let $u \in \mathcal{H}_1$ be an eigenvector for the eigenvalue λ and let Λ be an arbitrarily given neighborhood of this eigenvalue. Then*

$$\left\| u - \sum_{\lambda'_k \in \Lambda} (u, u'_k) u'_k \right\|_0 \leq \max(1, \gamma) \|u - Pu\|_0, \tag{3.2}$$

where $\gamma = 0$ if there is no discrete eigenvalue λ'_k outside Λ and γ is otherwise given by

$$\gamma = \max_{\lambda'_k \notin \Lambda} \left| \frac{\lambda'_k}{\lambda'_k - \lambda} \right|.$$

Proof. As $(u'_k, u'_\ell) = \delta_{k\ell}$, for all $f \in \mathcal{H}_0$

$$\|Rf\|_0^2 = \sum_{\lambda'_k \notin \Lambda} \left| \frac{\lambda'_k}{\lambda'_k - \lambda} (f, u'_k) \right|^2 \leq \gamma^2 \|Qf\|_0^2.$$

The proposition thus follows from the error representation (3.1) using the \mathcal{H}_0 -orthogonality of the terms $R(u - Pu)$ respectively $Q(u - Pu)$ and $(I - Q)(u - Pu)$. □

Taking into account the energy norm or \mathcal{H}_1 -stability of the \mathcal{H}_0 -orthogonal projection Q , a very similar estimate of the primarily interesting energy norm error can be derived:

Theorem 3.3. *Let $u \in \mathcal{H}_1$ be an eigenvector for the eigenvalue λ and let Λ be an arbitrarily given neighborhood of this eigenvalue. Then*

$$\left\| u - \sum_{\lambda'_k \in \Lambda} (u, u'_k) u'_k \right\| \leq (1 + (\gamma + 1)\kappa) \|u - Pu\|, \tag{3.3}$$

where γ is the same constant as in the previous theorem and κ the constant from (2.1).

Proof. The proof is again based on the error representation (3.1) and transfers almost verbatim from that of the previous theorem. As $a(u'_k, u'_\ell) = \lambda'_k \delta_{k\ell}$, for all $f \in \mathcal{H}_0$

$$\|Rf\|^2 = \sum_{\lambda'_k \notin \Lambda} \lambda'_k \left| \frac{\lambda'_k}{\lambda'_k - \lambda} (f, u'_k) \right|^2 \leq \gamma^2 \|Qf\|^2.$$

The only difference is that one can no longer argue using the orthogonality of the single terms but has to switch to the triangle inequality. The bound (2.1) for the energy norm of the projection operator Q enters in form of the estimate

$$\|Q(u - Pu)\| \leq \kappa \|u - Pu\|$$

for the projection of the approximation error. □

The larger the neighborhood Λ of the eigenvalue λ under consideration is chosen, the more discrete eigenvectors u'_k are used to approximate the assigned eigenvector u and the smaller the error is, but the less specific the relation between the original and the discrete eigenvectors becomes. To establish a correspondence between the continuous and discrete eigenvalues, we need a crude error estimate as the following one:

Lemma 3.4. *If \mathcal{E}_k , $k \leq n$, is the subspace of \mathcal{H}_1 spanned by the eigenvectors u_1, \dots, u_k of the original problem for the eigenvalues $\lambda_1 \leq \dots \leq \lambda_k$ and if there is an $\varepsilon < 1$ such that*

$$\|u - Pu\|_0 \leq \varepsilon \|u\|_0 \quad (3.4)$$

for all $u \in \mathcal{E}_k$, the following error estimate holds for the k -th eigenvalue:

$$\lambda_k \leq \lambda'_k \leq \frac{1}{(1 - \varepsilon)^2} \lambda_k.$$

Proof. Let \mathcal{E}'_k be the k -dimensional subspace of \mathcal{S} that is spanned by the first k discrete eigenvectors u'_1, \dots, u'_k . By the min-max characterization of the eigenvalue λ_k then

$$\lambda_k \leq \max_{u \in \mathcal{E}'_k} \frac{a(u, u)}{(u, u)} = \lambda'_k.$$

The functions Pu_1, \dots, Pu_k are linearly independent of each other. If namely a linear combination of these functions vanishes, the corresponding linear combination of the functions u_1, \dots, u_k vanishes because of (3.4) too. By the min-max characterization of λ'_k therefore

$$\lambda'_k \leq \max_{u \in \mathcal{E}_k} \frac{a(Pu, Pu)}{(Pu, Pu)}.$$

As $a(Pu, Pu) \leq a(u, u)$ and $\|Pu\|_0 \geq (1 - \varepsilon)\|u\|_0$ for $u \in \mathcal{E}_k$, the upper estimate

$$\lambda'_k \leq \frac{1}{(1 - \varepsilon)^2} \max_{u \in \mathcal{E}_k} \frac{a(u, u)}{(u, u)} = \frac{1}{(1 - \varepsilon)^2} \lambda_k$$

for the discrete eigenvalue λ'_k follows. □

The lemma ensures that the discrete eigenvalues λ'_k tend in the limit to their continuous counterparts λ_k . If the distance Δ of the eigenvalue λ under consideration to the neighboring eigenvalues of the original problem is sufficiently large, an adequate choice for the neighborhood Λ in (3.2), (3.3) is therefore the interval of length Δ with midpoint λ . Then

$$\gamma \leq \frac{2\lambda}{\Delta} + 1.$$

Asymptotically then only approximate eigenvectors u'_k for eigenvalues λ'_k tending to λ are taken into account. If the eigenvalue λ belongs to a cluster of closely neighboring eigenvalues, the neighborhood Λ should be chosen accordingly and (3.2) and (3.3) be interpreted as a result on the approximation by an element in the corresponding discrete invariant subspace. In any case, the approximation order in the given norms defined in a proper sense completely corresponds to that for a corresponding boundary value problem with solution u .

A similar result as for the eigenvectors holds for the eigenvalues:

Theorem 3.5. *Let $u \in \mathcal{H}_1$, $\|u\|_0 = 1$, be an eigenvector for the eigenvalue λ and assume that already a discrete eigenvalue $\lambda'_k \geq \lambda$ exists for which $\lambda'_k - \lambda \leq \lambda$. Then*

$$\min_{\lambda'_k \geq \lambda} (\lambda'_k - \lambda) \leq (1 + \alpha\kappa)^2 \|u - Pu\|^2,$$

where $\alpha = 1$ if there is no discrete eigenvalue $\lambda'_k < \lambda$ and α is otherwise given by

$$\alpha = \max_{\lambda'_k < \lambda} \frac{\lambda}{\lambda - \lambda'_k}.$$

Proof. Let u' be the best \mathcal{H}_0 -norm approximation

$$u' = \sum_{\lambda'_k \geq \lambda} (u, u'_k) u'_k$$

of u by a linear combination of the discrete eigenvectors u'_k for eigenvalues $\lambda'_k \geq \lambda$. Then

$$\|u - u'\|^2 = \lambda \|u - u'\|_0^2 + \sum_{\lambda'_k \geq \lambda} (\lambda'_k - \lambda) (u, u'_k)^2.$$

Since u' and $u - u'$ are by definition \mathcal{H}_0 -orthogonal, this implies

$$\|u - u'\|^2 \geq \min_{\lambda'_k \geq \lambda} (\lambda'_k - \lambda) \|u\|_0^2 + \left(\lambda - \min_{\lambda'_k \geq \lambda} (\lambda'_k - \lambda) \right) \|u - u'\|_0^2.$$

As the second term on the right-hand side is by assumption nonnegative and $\|u\|_0 = 1$, thus

$$\min_{\lambda'_k \geq \lambda} (\lambda'_k - \lambda) \leq \|u - u'\|^2.$$

The assertion follows from this estimate analogously to the proof of Theorem 3.3, simply replacing the neighborhood Λ of the eigenvalue λ by the half-open interval $[\lambda, \infty)$. \square

For eigenvalues greater than the minimum eigenvalue, the size of the prefactors depends asymptotically on the separation of the eigenvalue under consideration from the smaller eigenvalues. The speed with which the discrete eigenvalues converge to their continuous counterparts is asymptotically determined by the speed with which the square of the best energy norm approximation error of the assigned eigenfunctions tends to zero. As with the estimates from Theorem 3.2 and Theorem 3.3, pollution effects arising from the approximation error for other eigenfunctions than the one under consideration do not occur.

References

- [1] I. Babuška and J. Osborn, *Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems*, Math. Comput., **52** (1989), pp. 275–297.
- [2] I. Babuška and J. Osborn, *Eigenvalue problems*, in: Handbook of Numerical Analysis, vol. II, Finite Element Methods (Part 1), Elsevier, 1991, pp. 641–792.
- [3] R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comput., **36** (1981), pp. 35–51.
- [4] R. E. Bank, A. H. Sherman, and A. Weiser, *Refinement algorithms and data structures for regular local mesh refinement*, in: Scientific Computing, North Holland, 1983, pp. 3–17.
- [5] R. E. Bank and H. Yserentant, *On the H^1 -stability of the L_2 -projection onto finite element spaces*, Numer. Math. (2013), DOI 10.1007/s00211-013-0562-4.
- [6] J. Bey, *Tetrahedral grid refinement*, Computing, **55** (1995), pp. 355–378.
- [7] J. Bramble, J. Pasciak, and O. Steinbach, *On the stability of the L_2 -projection in $H^1(\Omega)$* , Math. Comp., **71** (2001), pp. 147–156.
- [8] C. Carstensen, *Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomé criterion for H^1 -stability of the L_2 -projection onto finite element spaces*, Math. Comput., **71** (2001), pp. 157–163.
- [9] C. Carstensen, *An adaptive mesh-refining algorithm allowing for an H^1 -stable L_2 -projection onto Courant finite element spaces*, Constr. Approx., **20** (2004), pp. 549–564.

-
- [10] M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces*, Math. Comput., **48** (1987), pp. 521–532.
- [11] M. Karkulik, D. Pavlicek, and D. Praetorius, *On 2D newest vertex bisection: Optimality of mesh closure and H^1 -stability of L_2 -projection*, Constr. Approx. (2013), DOI 10.1007/s00365-013-9192-4.
- [12] A. Knyazev and J. Osborn, *New a priori FEM error estimates for eigenvalues*, SIAM J. Numer. Anal., **43** (2006), pp. 2647–2667.
- [13] K. Neymeyr, *A posteriori error estimation for elliptic eigenproblems*, Numer. Linear Algebra Appl., **9** (2002), pp. 263–279.
- [14] P. Raviart and J. Thomas, *Introduction à l'Analyse Numérique des Équations aux Dérivées Partielles*, Masson, Paris, 1983.
- [15] H. Yserentant, *The hyperbolic cross space approximation of electronic wavefunctions*, Numer. Math., **105** (2007), pp. 659–690.