EIGENVALUES

We fix a domain $D \subset \mathbf{R}^3$. We consider the eigenvalue problem

$$-\Delta u = \lambda u \text{ in } D, \quad u = 0 \text{ on bdy } D.$$

In the following we will denote by $\mathcal{F}(D)$ the set of all differentiable functions u in D which satisfy the Dirichlet boundary condition $u = 0$ on bdy D; we shall also assume the functions to be real valued for simplicity even though most of what we do in this chapter can be easily generalized to complex valued functions. We have shown before that for Dirichlet boundary conditions all the eigenvalues of $-\Delta$ are positive, and we list them as

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ \ldots$$

We have already derived and used Green's first identity, see Chapter 7, page 179,

$$\int\int_{bdy\ D} v\frac{\partial u}{\partial n}\ dS \ = \ \int\int\int_D \nabla v \cdot \nabla u\ d\mathbf{x} + \int\int\int_D v\Delta u\ d\mathbf{x},$$

where $\frac{\partial u}{\partial n} = \nabla u \cdot \mathbf{n}$ is the directional derivative in direction of the outwards pointing normal vector $\mathbf{n}$ on the boundary of D. We also define, as usual, the inner product and norm for functions u, v on D by

$$(u, v) = \int\int\int_D u(\mathbf{x})v(\mathbf{x})\ d\mathbf{x} \quad \text{and} \quad \|u\|^2 = (u, u),$$

where we assume the functions to be real valued for simplicity. Then we have, for functions $u, v \in \mathcal{F}(D)$ that

$$(\nabla u, \nabla v) \ = \ (-\Delta u, v) \ = \ (u, -\Delta v). \tag{1}$$

Indeed this follows directly from the definitions and Greens first identity; our boundary conditions ensure that the left hand side of Greens identity is equal to 0.

We have already constructed an orthogonal set of eigenfunctions of $-\Delta$ for the ball of radius a which can be shown to have the completeness property, or, in other notation, they are an L^2-basis (see previous notes on vibrations of a drumhead).

Theorem 1 Assume that $\mathcal{F}(D)$ has an orthogonal L^2 basis (v_j). Then we have

$$\min_{u \in \mathcal{F}(D)} \frac{\|\nabla u\|^2}{\|u\|^2} \ = \ \lambda_1.$$

Proof. Let $u \in \mathcal{F}(D)$. By our assumptions, we can write

$$u(\mathbf{x}) = \sum_j a_j v_j(\mathbf{x}) \quad \text{and} \quad -\Delta u(\mathbf{x}) = \sum_j \lambda_j a_j v_j(\mathbf{x}).$$

Using this and (1), we get

$$\begin{aligned} (\nabla u, \nabla u) \ &= \ (-\Delta u, u) = (\sum_j \lambda_j a_j v_j, \sum_k a_k v_k) \ = \\ &= \ \sum_{j,k} \lambda_j a_j a_k (v_j, v_k) \ = \ \sum_j \lambda_j a_j^2 \|v_j\|^2 \ = \\ &\geq \ \sum_j \lambda_1 a_j^2 \|v_j\|^2 \ = \ \lambda_1 \|u\|^2. \end{aligned}$$

Hence we have

$$\frac{\|\nabla u\|^2}{\|u\|^2} \geq \lambda_1 \quad \text{for all } u \in \mathcal{F}(D).$$

In particular, λ_1 must be $\leq$ the minimum as in the statement. On the other hand, we do have equality above for $u = v_1$. Indeed, using (1), we have

$$(\nabla v_1, \nabla v_1) \ = \ (-\Delta v_1, v_1) \ = \lambda_1 (v_1, v_1).$$

This finishes the proof.

As outlined in the lecture as well as in the book (for $\ell = 1$), the same proof also works in the 1-dimensional setting, with $D = [0, \ell]$, with Dirichlet boundary condition given by $u(0) = 0 = u(\ell)$. The inner product is given by integration over the interval $[0, \ell]$, and the equalities (1) follow from integration by part. One can then prove

$$\min \frac{\|\nabla u\|^2}{\|u\|^2} \ = \ \min \frac{\|u'\|^2}{\|u\|^2} \ = \ \frac{\pi^2}{\ell^2},$$

using completeness of the set $(\sin(n\pi x/\ell))_n$ along the lines of the proof of Theorem 1.

For more complicated regions D in $\mathbf{R}^3$, one usually does not have available an explicit description of the eigenfunctions and eigenvalues of $-\Delta$. One now uses the minimizing property studied so far to find the eigenvalues of $-\Delta$. So in the following theorem we do *not* assume existence of a complete set of eigenfunctions.

Theorem 2 We have

$$\min_{u \in \mathcal{F}(D)} \frac{\|\nabla u\|^2}{\|u\|^2} \ = \ \lambda_1,$$

the smallest eigenvalue of $-\Delta$, provided there exists a function $v_1 \in \mathcal{F}(D)$ such that

$$\|\nabla v_1\|^2 = \lambda_1 \|v_1\|^2.$$

In this case v_1 is an eigenfunction of $-\Delta$ with eigenvalue $-\lambda_1$.

Proof. We first observe that the claim will follow as soon as we can prove

$$(\Delta v_1 + \lambda_1 v_1, v) = 0 \quad \text{for all } v \in \mathcal{F}(D). \tag{$*$}$$

Indeed, if we set $v = \Delta v_1 + \lambda_1 v_1$, we have

$$0 = (\Delta v_1 + \lambda_1 v_1, \Delta v_1 + \lambda_1 v_1) = \int\int\int_D (\Delta v_1 + \lambda_1 v_1)^2(\mathbf{x}) \ d\mathbf{x}.$$

Hence $\Delta v_1 + \lambda_1 v_1 = 0$, which proves the claim.

To prove statement $(*)$, we consider the function

$$f(\epsilon) = \frac{\|\nabla(v_1 + \epsilon v)\|^2}{\|v_1 + \epsilon v\|^2},$$

where $v \in \mathcal{F}(D)$ arbitrary. Observe that

$$\frac{d}{d\epsilon}\|v_1 + \epsilon v\|^2 \ = \ \frac{d}{d\epsilon}(v_1 + \epsilon v, v_1 + \epsilon v) \ = \ 2((v_1, v) + \epsilon(v, v)).$$

Observe that the term with (v, v) disappears for $\epsilon = 0$. We get a similar expression for $\frac{d}{d\epsilon}\|\nabla(v_1 + \epsilon v)\|^2$. By assumption on v_1, the function $f(\epsilon)$ has a minimum at $\epsilon = 0$. Using this and the quotient rule, we get

$$0 = f'(0) = \frac{(v_1, v_1)2(\nabla v_1, \nabla v) - 2(v_1, v)(\nabla v_1, \nabla v_1)}{(v_1, v_1)^2}.$$

Hence the numerator above has to be equal to 0, from which we deduce

$$(\nabla v_1, \nabla v) = \frac{(\nabla v_1, \nabla v_1)}{(v_1, v_1)}(v_1, v).$$

Applying (1) to the left hand side, and the assumption to the right hand side, we get

$$(-\Delta v_1, v) = \lambda_1(v_1, v).$$

Claim $(*)$ follows from this.

Corollary Assume we have an orthogonal set $v_1, v_2, \ldots v_{m-1}$ of eigenfunctions of $-\Delta$ belonging to the $m-1$ smallest eigenvalues. Let $\mathcal{F}^{(m-1)}(D)$ be the set of all functions $u \in \mathcal{F}(D)$ satisfying $(u, v_j) = 0$, $1 \leq j \leq m-1$. Then we have

$$\min_{u \in \mathcal{F}^{(m)}(D)} \frac{\|\nabla u\|^2}{\|u\|^2} = \lambda_m,$$

the m-th smallest eigenvalue of $-\Delta$, provided there exists a function $v_m \in \mathcal{F}(D)^{(m)}$ such that

$$\|\nabla v_m\|^2 = \lambda_m \|v_m\|^2.$$

In this case v_m is an eigenfunction of $-\Delta$ with eigenvalue $-\lambda_m$.

Proof. Goes the same way as the proof of Theorem 2. You only replace $\mathcal{F}(D)$ by $\mathcal{F}^{(m)}(D)$ and λ_1 by λ_m.

Remark The assumption about the existence of minimizing functions v_1 (for Theorem 2) and v_2, v_3 ... (for its corollary) usually holds for reasonable domains D. But it is not easy to prove, so we will not do this here. So, at least in principle, this theorem and its corollary could be used to find eigenvalues and eigenvalues for $-\Delta$. We will next show how our results so far motivate some practical approaches to get good approximations of the eigenvalues.

RAYLEIGH-RITZ APPROXIMATION

We now would like to use the theoretical results above to get good estimates for the lowest eigenvalues of $-\Delta$. A very primitive approach would be to randomly select n functions $w_1, w_2, \ldots w_n$, calculate $\|\nabla w_j\|^2/\|w\|^2$ for $j = 1, 2, \ldots n$ and pick the lowest value. We would need a lot of intuition and luck to get good estimates this way. Instead, we will now study a method which would calculate the minimum of $\|\nabla u\|^2/\|u\|^2$ for all functions u in the linear span of the functions w_1 until w_n. This is just a problem in linear algebra, which can be stated as follows:

Lemma 3 Let W be a finite dimensional vector space (say $W = \mathbf{R}^n$) with inner product (,), and let $w_1, w_2, \ldots w_n$ be linearly independent vectors in W. Moreover, let $D: W \to W$ be a linear map satisfying $(Dw, \tilde{w}) = (w, D\tilde{w})$ for all $w, \tilde{w} \in W$. Then the eigenvalues of D are the roots of $\det(A - \lambda B)$, where A and B are given by $a_{ij} = (Dw_i, w_j)$ and $b_{ij} = (w_i, w_j)$. In particular, the minimum of $(Dw, w)/(w, w)$ is given by the smallest eigenvalue λ_1 of D, and the maximum of $(Dw, w)/(w, w)$ is given by the largest eigenvalue of D.

Proof. Let $v_k = \sum c_i w_i$ be an eigenvector of D with eigenvalue λ_k. Plugging this into $(Dv_k, w_j) = (\lambda_k v_k, w_j)$ gives us

$$\sum_i a_{ij} c_i = \sum_i \lambda_k b_{ij} c_i,$$

for $j = 1, 2, \ldots n$. Denoting by c the row vector $(c_1, c_2, \ldots c_n)$, this can be rewritten in matrix form as

$$c(A - \lambda_k B) = 0.$$

It follows that $\det(A - \lambda_k B) = 0$, i.e. λ_k is a root of the polynomial $\det(A - \lambda B)$. To prove the statement about the minimum, recall that the eigenvectors of D belonging to different eigenvalues are orthogonal. Here is the one-line proof again for eigenvectors v_i and v_j:

$$\lambda_i(v_i, v_j) = (\lambda_i v_i, v_j) = (Dv_i, v_j) = (v_i, Dv_j) = \lambda_j(v_i, v_j).$$

Hence if $\lambda_i \neq \lambda_j$, the equality above can only hold if $(v_i, v_j) = 0$. One can deduce from this that D has an orthogonal basis (v_i) of eigenvectors. The same argument as in the proof of Theorem 1 now shows the last statement, that $(Dw, w) \geq \lambda_1(w, w)$ for all $w \in W$. Similarly, using this line of arguments and the inequality

$$\sum_j \lambda_j a_j^2 \|v_j\|^2 \leq \lambda_n \sum_j a_j^2 \|v_j\|^2,$$

one also proves that $(Dw, w) \leq \lambda_n(w, w)$, with equality reached for $w = v_n$.

Theorem 4 (Rayleigh-Ritz approximation). Let $w_1, w_2, \ ... \ w_n$ be n linearly independent vectors in $\mathcal{F}(D)$, and let A and B be the $n \times n$ matrices defined by

$$a_{ij} = (-\Delta w_i, w_j), \quad b_{ij} = (w_i, w_j).$$

Denote by $\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ ... \ \lambda_n^{(n)}$ the roots of the polynomial $\det(A - \lambda B)$. Then $\lambda_i^{(n)} \to \lambda_i$, the i-th eigenvalue of $-\Delta$ if $n \to \infty$.

Sketch of proof. This follows essentially from Lemma 3, which we apply to the following situation: The vector space W is given by the span of the functions $w_1, \ w_2 \ ... \ w_n$. Let P be the orthogonal projection onto this subspace with respect to our inner product on $\mathcal{F}(D)$, and let $D = P(-\Delta)P : \ W \to W$. Then all the assumptions of Lemma 3 are satisfied, using Eq (1). It is now possible to make the following informal statement mathematically rigorous: The more functions we use to construct the subspace W, the closer the operator $P(-\Delta)P$ approximates the operator $-\Delta$. In particular, the i-th lowest eigenvalue $\lambda_i^{(n)}$ of $P(-\Delta)P$ will approximate the eigenvalue λ_i of $-\Delta$.

We now discuss another way how to characterize the n-th eigenvalue λ_n of $-\Delta$. Let again W be the linear span of n linearly independent functions $w_1, \ w_2, \ ... \ w_n$ in $\mathcal{F}(D)$. Let $\lambda_n^* = \lambda_n^*(W)$ be the largest eigenvalue of $P(-\Delta)P$, where P is the projection onto W as in the last theorem. Then it follows from Lemma 1, by the exactly same reasoning as in the last theorem that

$$\lambda_n^*(W) = \max\{\frac{\|\nabla w\|^2}{\|w\|^2}, \ w \in W\}.$$

Example See Example 2 on p. 306 in the book. There we have for D the disk of radius 1. For simplicity, we only consider radially invariant functions, i.e. functions which are independent of θ in polar coordinates. Then we have

$$-\Delta u = -u_{rr} - \frac{1}{r}u_r = -\frac{1}{r}(ru_r)_r.$$

We then get, using polar coordinates and integration by parts,

$$\int\int |\nabla u|^2 dxdy \ = \ \int_0^1 \int_0^{2\pi} (-\frac{1}{r}(ru_r)_r)u \ rd\theta dr \ = \ 2\pi[(-ru_r u)|_{r=0}^{r=1} + \int_0^1 ru_r^2 \ dr].$$

As $u(1) = 0$ by Dirichlet boundary condition, the first term on the right hand side disappears. So we get

$$\frac{\|\nabla u\|^2}{\|u\|^2} = \frac{\int\int -u\Delta u \ dxdy}{\int\int |u|^2 \ dxdy} = \frac{\int_0^1 ru_r^2 \ dr}{\int_0^1 ru^2 \ dr}.$$

We take as trial functions $w_1 = 1 - r^2$ and $w_2 = (1 - r^2)^2$. Doing the integrations for these functions, see the book, we get the matrices

$$A \;=\; 2\pi \begin{bmatrix} 1 & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} \end{bmatrix} \quad \text{and} \quad B \;=\; 2\pi \begin{bmatrix} \frac{1}{6} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{10} \end{bmatrix}.$$

Then one checks (see p. 206) that $\det(A - \lambda B)$ has the eigenvalues approximately equal to 5.784 and 36.9. The true eigenvalues are given by the squares of the roots of the Bessel function J_0 which would be approximately $5.783 = z_1^2 \sim (2.405)^2$ and 30.5.

Theorem 5 (Minimax Principle) The n-th eigenvalue λ_n of $-\Delta$ is given by

$$\lambda_n = \min \lambda_n^*(W),$$

where W ranges over all possible n dimensional subspaces of $\mathcal{F}(D)$, i.e. over all possible spans of n linearly independent functions in $\mathcal{F}(D)$.

Proof. Pick any n linearly independent functions $w_1, w_2, \; ... \; w_n$. Then we can find a nonzero linear combination $0 \neq w = \sum c_j w_j$ such that w is orthogonal to the first $(n-1)$ eigenfunctions of $-\Delta$. Indeed, this is equivalent to

$$(w, v_k) \;=\; \sum_j (w_j, v_k) c_j \;=\; 0, \quad 1 \leq k \leq n-1.$$

This is a homogeneous linear system of $n-1$ equations with n unknowns; hence it has a nontrivial solution for $c_1, c_2, \; ... \; c_n$ which shows the existence of a vector w as stated above. But then it follows from the corollary of Theorem 2 that $(-\Delta w, w) \geq \lambda_n (w, w)$. Hence also $\lambda_n^*(W) \geq \lambda_n$. As W was the span of an arbitrary linear combination of n linearly independent functions, it also follows that the minimum of all possible $\lambda_n^*(W)$ must be $\geq \lambda_n$.

To show that the minimum actually is equal to λ_n we take for W the span of the first n eigenfunctions $v_1, v_2, \; ... \; v_n$ of $-\Delta$. If $w = \sum c_k v_k$ is a function in W, then we have

$$(-\Delta w, w) = \sum_{k,m} (-\Delta c_k v_k, c_m v_m) \;=\; \sum_{k,m} \lambda_k c_k c_m (v_k, v_m) \;=$$

$$= \; \sum_k \lambda_k c_k^2 (v_k, v_k) \;\leq\; \lambda_n \sum_k c_k^2 (v_k, v_k) = \lambda_n (w, w).$$

Hence the maximum of $(-\Delta w, w)/((w, w)$ is $\leq \lambda_n$. On the other hand, we have $(-\Delta v_n, v_n) = \lambda_n (v_n, v_n)$, i.e. the maximum is indeed equal to λ_n. Hence we have $\lambda_n = \lambda_n^*(W)$ for our choice of W.