6 Eigenvalues of the Laplacian

In this section, we consider the following general eigenvalue problem for the Laplacian,

$$\begin{cases} -\Delta v = \lambda v & x \in \Omega \\ v \text{ satisfies symmetric BCs} & x \in \partial\Omega. \end{cases}$$

To say that the boundary conditions are **symmetric** for an open, bounded set Ω in $\mathbb{R}^n$ means that

$$\langle u, \Delta v\rangle = \langle \Delta u, v\rangle$$

for all functions u and v which satisfy the boundary conditions, where $\langle \cdot, \cdot\rangle$ denotes the L^2 inner product on Ω; that is, for any real-valued functions f and g on Ω,

$$\langle f, g\rangle = \int_\Omega f(x)g(x)\,dx$$

We note that this definition is equivalent to the definition given earlier for the case when Ω is an interval in $\mathbb{R}$.

The most common symmetric boundary conditions are the following:

1. Dirichlet: $v = 0$
2. Neumann: $\frac{\partial v}{\partial \nu} = 0$
3. Robin: $\frac{\partial v}{\partial \nu} + a(x)v = 0$.

6.1 Application to the Heat Equation

Example 1. Heat Equation on a bounded domain $\Omega \subset \mathbb{R}^n$,

$$\begin{cases} u_t = k\Delta u & x \in \Omega, t > 0 \\ u(x,0) = \phi(x) & \\ u(0,t) = 0 & x \in \partial\Omega, t \geq 0. \end{cases}$$

Using separation of variables, we look for a solution of the form $u(x,t) = v(x)T(t)$, which leads to the following eigenvalue problem,

$$\begin{cases} -\Delta v = \lambda v & x \in \Omega \\ v = 0 & x \in \partial\Omega \end{cases}$$

◇

6.2 Facts on Eigenvalues

Theorem 2. *For any of the boundary conditions listed above,*

1. *All eigenvalues are real.*
2. *All eigenfunctions can be chosen to be real-valued.*

3. *Eigenfunctions corresponding to distinct eigenvalues are orthogonal.*

4. *All eigenfunctions may be chosen to be orthogonal by using a Gram-Schmidt process.*

Proof. Proofs of properties (3) and (4) are similar to the 1-dimensional case, discussed earlier. For proofs of (1) and (2), see Strauss. $\square$

Theorem 3. *For the eigenvalue problem above,*

1. *All eigenvalues are positive in the Dirichlet case.*

2. *All eigenvalues are zero or positive in the Neumann case and the Robin case if $a \geq 0$.*

Proof. We prove this result for the Dirichlet case. The other proofs can be handled similarly.

Let v be an eigenfunction with corresponding eigenvalue λ. Then

$$\begin{aligned}
\lambda \int_\Omega v^2\,dx &= -\int_\Omega (\Delta v) v\,dx \\
&= \int_\Omega |\nabla v|^2\,dx - \int_{\partial\Omega} v\frac{\partial v}{\partial \nu}\,dS(x) \\
&= \int_\Omega |\nabla v|^2\,dx
\end{aligned}$$

Therefore,

$$\lambda \int_\Omega v^2\,dx = \int_\Omega |\nabla v|^2\,dx \geq 0.$$

Further, we claim that

$$\int_\Omega |\nabla v|^2\,dx > 0.$$

We prove this claim as follows. Suppose $\int_\Omega |\nabla v|^2\,dx = 0$, then $|\nabla v| = 0$ which implies v is constant on Ω. But, by assumption $v = 0$ on $\partial\Omega$. Therefore, if v is constant on Ω and $v = 0$ on $\partial\Omega$, then $v \equiv 0$. However, the zero function is not an eigenfunction. Therefore,

Therefore,

$$\lambda \int_\Omega v^2\,dx = \int_\Omega |\nabla v|^2\,dx > 0,$$

which implies $\lambda > 0$.

$\square$

6.3 Eigenvalues as Minima of the Potential Energy

In general, it is difficult to explicitly calculate eigenvalues for a given domain $\Omega \subset \mathbb{R}^n$. In this section, we prove that eigenvalues are minimizers of a certain functional. This fact will allow us to approximate eigenvalues for given regions $\Omega \subset \mathbb{R}^n$.

Consider the eigenvalue problem with Dirichlet boundary conditions,

$$\begin{cases} -\Delta u = \lambda u & x \in \Omega \\ u = 0 & x \in \partial\Omega. \end{cases} \tag{6.1}$$

Let $0 < \lambda_1 \leq \lambda_2 \leq \ldots$ be the eigenvalues of (6.1).

For a given function w defined on a set $\Omega \subset \mathbb{R}^n$, we define the **Rayleigh Quotient** of w on Ω as

$$\frac{||\nabla w||^2_{L^2(\Omega)}}{||w||^2_{L^2(\Omega)}} = \frac{\int_\Omega |\nabla w|^2\, dx}{\int_\Omega w^2\, dx}.$$

Theorem 4. (Minimum Principle for the First Eigenvalue) *Let*

$$Y \equiv \{w : w \in C^2(\Omega), w \not\equiv 0, w = 0 \text{ for } x \in \partial\Omega\}.$$

We call this the set of **trial functions** *for* (6.1). *Suppose there exists a function $u \in Y$ such that u minimizes the Rayleigh quotient over all trial functions $w \in Y$. That is,*

$$m \equiv \frac{||\nabla u||^2}{||u||^2} = \min_{w \in Y} \left\{ \frac{||\nabla w||^2}{||w||^2} \right\}.$$

Then m is the first eigenvalue of (6.1). *That is, $m = \lambda_1$ and u is a corresponding eigenfunction.*

Proof. Suppose u is the minimizer of the Rayleigh quotient and m is the Rayleigh quotient of u. That is,

$$m = \frac{\int_\Omega |\nabla u|^2\, dx}{\int_\Omega u^2\, dx}. \tag{6.2}$$

Pick a function $v \in Y$. Let

$$f(\epsilon) \equiv \frac{\int_\Omega |\nabla(u + \epsilon v)|^2\, dx}{\int_\Omega (u + \epsilon v)^2\, dx}.$$

If u minimizes the Rayleigh quotient, then f must satisfy $f'(0) = 0$. Taking the derivative of f, we see that

$$f'(\epsilon) = \frac{(\int_\Omega (u+\epsilon v)^2\, dx)(2\int_\Omega \nabla u \cdot \nabla v + \epsilon |\nabla v|^2\, dx) - (\int_\Omega 2(u+\epsilon v)v\, dx)(\int_\Omega |\nabla(u+\epsilon v)|^2\, dx)}{(\int_\Omega (u+\epsilon v)^2\, dx)^2}.$$

Therefore,

$$f'(0) = \frac{(\int_\Omega u^2\, dx)(2\int_\Omega \nabla u \cdot \nabla v\, dx) - (2\int_\Omega uv\, dx)(\int_\Omega |\nabla u|^2\, dx)}{(\int_\Omega u^2\, dx)^2}.$$

Now, $f'(0) = 0$ implies

$$\left(\int_\Omega u^2\, dx \right) \left(\int_\Omega \nabla u \cdot \nabla v\, dx \right) = \left(\int_\Omega uv\, dx \right) \left(\int_\Omega |\nabla u|^2\, dx \right),$$

which implies

$$\begin{aligned} \int_\Omega \nabla u \cdot \nabla v\, dx &= \frac{\int_\Omega |\nabla u|^2\, dx}{\int_\Omega u^2\, dx} \int_\Omega uv\, dx \\ &= m \int_\Omega uv\, dx, \end{aligned}$$

by (6.2). Using the Divergence theorem, we have

$$-\int_{\Omega}(\Delta u)v\,dx + \int_{\partial\Omega}\frac{\partial u}{\partial\nu}v\,dS(x) = m\int_{\Omega}uv\,dx.$$

By assumption, $v = 0$ on $\partial\Omega$. Therefore, the boundary term vanishes. Therefore,

$$-\int_{\Omega}(\Delta u)v\,dx = m\int_{\Omega}uv\,dx$$

for all $v \in Y$. Now, as this is true for all trial functions v, we conclude that

$$-\Delta u = mu,$$

which means that u is an eigenfunction of (6.1) with corresponding eigenvalue m.

It only remains to show that m is the smallest eigenvalue. Suppose v is another eigenfunction of (6.1) with corresponding eigenvalue λ_i. We just need to show that $\lambda_i \geq m$. Using the Divergence theorem and the fact that v vanishes on the boundary, we have

$$m = \frac{||\nabla u||^2}{||u||^2} \leq \frac{||\nabla v||^2}{||v||^2} = \frac{\int_{\Omega}|\nabla v|^2\,dx}{\int_{\Omega}v^2\,dx} = \frac{-\int_{\Omega}(\Delta v)v\,dx}{\int_{\Omega}v^2\,dx} = \frac{\lambda_i\int_{\Omega}v^2\,dx}{\int_{\Omega}v^2\,dx} = \lambda_i.$$

Therefore, the theorem is proved.

□

Theorem 5. (Minimum Principle for the nth Eigenvalue) *Fix an integer $n \geq 1$. Let v_i, $i = 1, \ldots, n-1$ be the first $n-1$ eigenfunctions of* (6.1)*. Without loss of generality, these eigenfunctions may be chosen to be orthogonal. Let*

$$Y_n \equiv \{w : w \in C^2(\Omega), w \not\equiv 0, w = 0 \text{ for } x \in \partial\Omega, \langle w, v_i\rangle = 0 \text{ for } i = 1, \ldots, n-1\}.$$

Suppose there exists a function $u_n \in Y_n$ which minimizes the Rayleigh quotient over all functions $w \in Y_n$. That is, suppose

$$m_n \equiv \frac{||\nabla u_n||^2}{||u_n||^2} = \min_{w\in Y_n}\frac{||\nabla w||^2}{||w||^2}.$$

Then m_n is the nth eigenvalue of (6.1)*. That is, $\lambda_n = m_n$ and u_n is an eigenfunction of* (6.1) *with eigenvalue m_n.*

Proof. Suppose $u_n \in Y_n$ is the minimizer of the Rayleigh quotient over all functions $w \in Y_n$. That is,

$$m_n \equiv \frac{||\nabla u_n||^2}{||u_n||^2} = \min_{w\in Y_n}\left\{\frac{||\nabla w||^2}{||w||^2}\right\}.$$

Fixing $v \in Y_n$, defining $f(\epsilon)$ as before and using the fact that $f'(0) = 0$, we see that

$$\int_{\Omega}(\Delta u_n + m_n u_n)v\,dx = 0.$$

This is true for *any* $v \in Y_n$. Therefore, we conclude that

$$\int_\Omega (\Delta u_n + m_n u_n) v \, dx = 0 \tag{6.3}$$

for all trial functions v which satisfy $\langle v, v_i \rangle = 0$ for $i = 1, \ldots, n-1$.

To conclude that

$$\Delta u_n + m_n u_n = 0,$$

we need to show that (6.3) is true for *all* trial functions (not just those trial functions which are orthogonal to the first $n-1$ eigenvalues).

Now let h be an arbitrary trial function. Let

$$v(x) \equiv h(x) - \sum_{k=1}^{n-1} c_k v_k(x) \qquad \text{where } c_k \equiv \frac{\langle h, v_k \rangle}{\langle v_k, v_k \rangle}$$

and the v_i are the first $n-1$ eigenfunctions. We claim that

$$\int_\Omega (\Delta u_n + m_n u_n) h \, dx = 0.$$

We note that

$$\begin{aligned}
\int_\Omega (\Delta u_n + m_n u_n) h \, dx &= \int_\Omega (\Delta u_n + m_n u_n) \left[v + \sum_{k=1}^{n-1} c_k v_k \right] dx \\
&= \int_\Omega (\Delta u_n + m_n u_n) v \, dx + \sum_{k=1}^{n-1} c_k \int_\Omega (\Delta u_n + m_n u_n) v_k \, dx.
\end{aligned}$$

Now, first, we claim that v is orthogonal to v_i for $i = 1, \ldots, n-1$, and, therefore, the first term on the right-hand side above vanishes. We prove this claim as follows. Let v_i be an arbitrary eigenfunction for $i = 1, \ldots, n-1$. Then

$$\begin{aligned}
\langle v, v_i \rangle = \int_\Omega v v_i \, dx &= \int_\Omega \left(h - \sum_{k=1}^{n-1} c_k v_k \right) v_i \, dx \\
&= \int_\Omega h v_i \, dx - \sum_{k=1}^{n-1} c_k \int_\Omega v_k v_i \, dx \\
&= \int_\Omega h v_i \, dx - c_i \int_\Omega v_i v_i \, dx \\
&= \int_\Omega h v_i \, dx - \int_\Omega h v_i \, dx = 0,
\end{aligned}$$

using the definition of c_i and the fact that eigenfunctions are orthogonal. Therefore,

$$\int_\Omega (\Delta u_n + m_n u_n) v \, dx = 0.$$

Next, we claim that for all eigenfunctions v_i, $i = 1, \ldots, n-1$,

$$\int_\Omega (\Delta u_n + m_n u_n) v_i \, dx = 0.$$

We prove this claim as follows. Fix an eigenfunction v_i. Let λ_i be its corresponding eigenvalue. Then

$$\begin{aligned}
\int_\Omega (\Delta u_n + m_n u_n) v_i \, dx &= -\int_\Omega (\nabla u_n \cdot \nabla v_i) \, dx + \int_{\partial\Omega} \frac{\partial u_n}{\partial \nu} v_i \, dS(x) + \int_\Omega m_n u_n v_i \, dx \\
&= \int_\Omega u_n \Delta v_i \, dx - \int_{\partial\Omega} u_n \frac{\partial v_i}{\partial \nu} \, dS(x) + \int_{\partial\Omega} \frac{\partial u_n}{\partial \nu} v_i \, dS(x) \\
&\quad + \int_\Omega m_n u_n v_i \, dx \\
&= (-\lambda_i + m_n) \int_\Omega u_n v_i \, dx.
\end{aligned}$$

By assumption, $u_n \in Y_n$ which implies u_n is orthogonal to the first $n-1$ eigenvalues. Therefore, $\int_\Omega u_n v_i \, dx = 0$ for $i = 1, \ldots, n-1$. Therefore,

$$\int_\Omega (\Delta u_n + m_n u_n) v_i \, dx = 0.$$

Consequently, we conclude that

$$\int_\Omega (\Delta u_n + m_n u_n) h \, dx = 0,$$

where h is an arbitrary trial function. Consequently, we conclude that

$$\Delta u_n + m_n u_n = 0,$$

and, therefore, u_n is an eigenfunction with eigenvalue m_n.

Now, clearly, $m_n \geq \lambda_{n-1} \geq \lambda_{n-2} \geq \ldots$ because $Y_n \subset Y_{n-1} \subset \ldots$. We can prove that the other eigenvalues $\lambda_{n+1}, \lambda_{n+2}, \ldots$ are larger than m_n using the same technique as in the previous theorem, and the fact that the eigenfunctions v_k for $k \geq n+1$ satisfy $\langle v_k, v_i \rangle = 0$ for $i = 1, \ldots, n-1$.

□

We can now use the above minimization principles to approximate eigenvalues for given regions $\Omega \subset \mathbb{R}^n$.

Example 6. Let $\Omega = [0, 1]$. Use the trial function $v(x) = x(1-x)$ to approximate the first eigenvalue of (6.1) for this region Ω.

(Note: Of course, we already know the eigenvalues for $\Omega = [0,1]$ are given by $\lambda_n = (n\pi)^2$. We use this example just to demonstrate how the above technique works.)

We calculate the Rayleigh quotient of v,

$$\frac{||\nabla v||^2}{||v||^2} = \frac{\int_0^1 |v'(x)|^2 \, dx}{\int_0^1 v^2(x) \, dx} = \frac{\int_0^1 (1-2x)^2 \, dx}{\int_0^1 (x - x^2)^2 \, dx} = \frac{\int_0^1 (1 - 4x + 4x^2) \, dx}{\int_0^1 (x^2 - 2x^3 + x^4) \, dx} = \frac{(1/3)}{(1/30)} = 10.$$

Of course, the first eigenvalue is actually $\pi^2 \approx 9.8696$, but with a fairly simple choice of trial function, we get a fairly good approximation.

◇

6.4 Minimax Principle

In this section, we present another theorem regarding the eigenvalues of (6.1). This theorem is known as the minimax principle. It will allow us to prove a relationship between eigenvalues of sets contained within larger sets. In particular, we will show that if $\Omega_1 \subset \Omega_2$, then $\lambda_n(\Omega_1) \geq \lambda_n(\Omega_2)$, where λ_n is the nth eigenvalue of (6.1). This fact will give us another means of approximating eigenvalues of arbitrary domains Ω. In addition, it will allow us to prove the completeness of eigenfunctions of (6.1) in the L^2-sense. Before we get to these results, however, we present some motivation for the minimax principle. This motivation will also provide us another means of approximating eigenvalues.

Rayleigh-Ritz Approximation.

Let $w_1, \ldots, w_n$ be n arbitrary trial functions. (Recall: w is a trial function if it is $C^2(\Omega)$ and vanishes on $\partial\Omega$, but is not identically zero.) Let

$$w \equiv \sum_{k=1}^{n} c_k w_k,$$

be a linear combination of these trial functions. Suppose we made a really good choice of trial functions, and, in particular, chose w in such a way that w was an eigenfunction of (6.1) with eigenvalue λ. Of course, this is not likely by randomly guessing, but we will use this idea to find a way of *approximating* eigenvalues.

Now, if w was an eigenfunction of (6.1), then we know that

$$\lambda \langle w_j, w \rangle = \lambda \int_\Omega w_j w \, dx = -\int_\Omega w_j \Delta w \, dx = \int_\Omega \nabla w_j \cdot \nabla w \, dx = \langle \nabla w_j, \nabla w \rangle ,$$

using the fact that w, w_j are trial functions, and, therefore, vanish on the boundary of Ω. Further, using the definition of w, we have

$$\lambda \left\langle w_j, \sum_{k=1}^{n} c_k w_k \right\rangle = \left\langle \nabla w_j, \nabla \left(\sum_{k=1}^{n} c_k w_k \right) \right\rangle ,$$

which implies

$$\lambda \sum_{k=1}^{n} c_k \langle w_j, w_k \rangle = \sum_{k=1}^{n} c_k \langle \nabla w_j, \nabla w_k \rangle .$$

Define

$$\begin{aligned} a_{jk} &= \langle \nabla w_j, \nabla w_k \rangle \\ b_{jk} &= \langle w_j, w_k \rangle . \end{aligned} \tag{6.4}$$

Therefore, *if* w was actually an eigenfunction, we would have

$$\lambda \sum_{k=1}^{n} c_k b_{jk} = \sum_{k=1}^{n} c_k a_{jk} \qquad \text{for } j = 1, \ldots, n.$$

In other words, defining the $n \times n$ symmetric matrices

$$A = (a_{jk}) \qquad B = (b_{jk}),$$

and letting

$$c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

we would have

$$\lambda Bc = Ac \implies [A - \lambda B]c = 0.$$

In particular, this would mean $A - \lambda B$ is a singular matrix, and, therefore, $\det[A - \lambda B] = 0$.

Again, it is not likely that we will be so lucky in randomly choosing w, but we use this technique to approximate our first n eigenvalues. We now state our approximation technique. Let $w_1, \ldots, w_n$ be any n trial functions. For this choice of trial functions, define a_{jk}, b_{jk} as in (6.4) and let A, B be the corresponding $n \times n$ symmetric matrices with entries a_{jk}, b_{jk}, respectively. Then the n roots of the polynomial equation

$$\det(A - \lambda B) = 0$$

are approximations to the first n eigenvalues $\lambda_1, \ldots, \lambda_n$.

We now turn to the *minimax principle* for eigenvalues of (6.1). We use the Rayleigh-Ritz approximation method to motivate the minimax principle. First, we prove the following lemma.

Lemma 7. *Let A, B be $n \times n$ symmetric matrices. In addition, let B be positive definite. For $i = 1, \ldots, n$ let λ_i^* be the n roots of the characteristic equation* $\det(A - \lambda B) = 0$.

The n roots λ_i^* are all real, and the largest root λ_n^* satisfies

$$\lambda_n^* = \max_{\substack{c \in \mathbb{R}^n \\ c \neq 0}} \frac{Ac \cdot c}{Bc \cdot c}. \tag{6.5}$$

Remark. For $A = (a_{jk})$ and $B = (b_{jk})$ where a_{jk} and b_{jk} are defined as in (6.4) for some choice of trial functions w_i, A and B will satisfy the hypotheses of this lemma. We will use this lemma to motivate the minimax principle.

In order to prove Lemma 7, we first prove the following claim.

Claim 8. *Let A, B be $n \times n$ real, symmetric matrices. In addition, assume B is positive definite. Let $\lambda_1^*, \ldots, \lambda_n^*$ be the n roots of the characteristic equation*

$$\det(A - \lambda B) = 0.$$

Then the roots λ_i^* are all real. In addition, there exists a set of vectors $\{v_i\}$ which forms a basis for $\mathbb{R}^n$ and such that each v_i sastisfies the equation

$$Av_i = \lambda_i^* Bv_i$$

for some λ_i^*. Further,

$$Bv_i \cdot v_j = 0 \qquad \textit{for } i \neq j.$$

Remark. In the proof below, we will use the following fact regarding positive definite, real symmetric matrices. If B is a positive definite, real symmetric matrix, then there exists a lower triangular matrix L whose diagonal entries are positive and such that $B = LL^T$. This is called the Cholesky decomposition. Ref: *Linear Algebra with Applications*, S. Leon.

Proof. First, we will show that the roots of the characteristic equation

$$\det(A - \lambda B) = 0$$

are all real. Assume λ_i^* is a root of this equation. Then, using the fact that $B = LL^T$ for some lower triangular matrix L whose diagonal entries are positive, we have

$$\det(A - \lambda_i^* LL^T) = 0 \iff \det(L^{-1}A(L^T)^{-1} - \lambda_i^* I) = 0.$$

Therefore, λ_i^* is a root of the characteristic equation $\det(A - \lambda B) = 0$ if and only if λ_i^* is an eigenvalue of the matrix

$$M = L^{-1}A(L^T)^{-1}.$$

By a quick calculation, we see that M is real symmetric, and, therefore, all its eigenvalues are real. In addition, M has an orthonormal eigenbasis $\{u_i\}$.

We will now use this orthonormal eigenbasis $\{u_i\}$ to construct a basis for $\mathbb{R}^n$ consisting of solutions of $Av = \lambda Bv$ for some $\lambda \in \mathbb{R}$. Let

$$v_i \equiv (L^T)^{-1}u_i.$$

As the $\{u_i\}$ form a basis for $\mathbb{R}^n$ and $(L^T)^{-1}$ has rank n, we see that the set $\{v_i\}$ forms a basis for $\mathbb{R}^n$.

We now need to show that each v_i satisfies the equation

$$Av_i = \lambda_i^* Bv_i$$

for some λ_i^*. By assumption, u_i is an eigenvector of M with corresponding eigenvalue λ_i^*. Therefore,

$$\begin{aligned}
Mu_i = \lambda_i^* u_i &\implies L^{-1}A(L^T)^{-1}u_i = \lambda_i^* u_i \\
&\implies (L^T)^{-1}L^{-1}A(L^T)^{-1}u_i = \lambda_i^*(L^T)^{-1}u_i \\
&\implies (LL^T)^{-1}Av_i = \lambda_i^* v_i \\
&\implies B^{-1}Av_i = \lambda_i^* v_i \\
&\implies Av_i = \lambda_i^* Bv_i.
\end{aligned}$$

Therefore, we have found a set of vectors $\{v_i\}$ which forms a basis for $\mathbb{R}^n$ and such that each v_i satisfies the desired equation for some $\lambda_i^* \in \mathbb{R}$.

It remains only to show that

$$Bv_i \cdot v_j = 0 \qquad \text{for } i \neq j.$$

Using the definition of v_j and the fact that $B = LL^T$ for some lower triangular matrix L whose diagonal entries are positive, we have

$$Bv_i \cdot v_j = LL^T(L^T)^{-1}u_i \cdot (L^T)^{-1}u_j = Lu_i \cdot (L^T)^{-1}u_j = u_i^T L^T (L^T)^{-1}u_j = u_i^T u_j = u_i \cdot u_j = 0,$$

using the fact that the set $\{u_i\}$ is orthonormal. □

We now have the necessary ingredients to prove Lemma 7.

Proof of Lemma 7. By definition, it is easy to see that B is positive definite. Therefore, applying Claim 8, we see that all roots of the characteristic equation

$$\det(A - \lambda B) = 0$$

are real. Let λ_n^* be the largest of these roots. We now need to prove (6.5).

We will begin by showing that

$$\max_{\substack{c \in \mathbb{R}^n \\ c \neq 0}} \frac{Ac \cdot c}{Bc \cdot c} \leq \lambda_n^*. \tag{6.6}$$

Let $c \in \mathbb{R}^n$. By Claim 8, we can write c as a linear combination of the v_i where each v_i is a solution of

$$Av_i = \lambda_i^* B v_i$$

for one of the λ_i^*. Therefore, writing

$$c = a_1 v_1 + \ldots + a_n v_n,$$

and using the fact that B is positive definite, we have

$$\begin{aligned}
Ac \cdot c &= A(a_1 v_1 + \ldots + a_n v_n) \cdot c \\
&= (a_1 \lambda_1^* B v_1 + \ldots + a_n \lambda_n^* B v_n) \cdot (a_1 v_1 + \ldots + a_n v_n) \\
&= a_1^2 \lambda_1^* B v_1 \cdot v_1 + \ldots + a_n^2 \lambda_n^* B v_n \cdot v_n \\
&\leq a_1^2 \lambda_n^* B v_1 \cdot v_1 + \ldots + a_n^2 \lambda_n^* B v_n \cdot v_n \\
&= \lambda_n^* (B(a_1 v_1 + \ldots + a_n v_n)) \cdot (a_1 v_1 + \ldots + a_n v_n) \\
&= \lambda_n^* (Bc \cdot c).
\end{aligned}$$

Consequently, we have shown that

$$\frac{Ac \cdot c}{Bc \cdot c} \leq \lambda_n^*.$$

Taking the maximum of both sides over all $c \in \mathbb{R}^n$, we have proven (6.6).

It remains only to show that

$$\lambda_n^* \leq \max_{\substack{c \in \mathbb{R}^n \\ c \neq 0}} \frac{Ac \cdot c}{Bc \cdot c}. \tag{6.7}$$

We do so by finding a specific $c \in \mathbb{R}^n$ such that

$$\lambda_n^* \leq \frac{Ac \cdot c}{Bc \cdot c}.$$

We know that λ_n^* is a root of the characteristic equation $\det(A - \lambda B) = 0$. Therefore, there exists a vector $v_n \not\equiv 0 \in \mathbb{R}^n$ such that $(A - \lambda_n^* B)v_n = 0$. Let $c = v_n$. Therefore,

$$\frac{Av_n \cdot v_n}{Bv_n \cdot v_n} = \frac{\lambda_n^* B v_n \cdot v_n}{Bv_n \cdot v_n} = \lambda_n^*,$$

and we have proved (6.7).

$\square$

We now return to motivating the minimax principle. Recall from our Rayleigh-Ritz approximation that for a given set $\Omega \subset \mathbb{R}^n$, we can approximate the first n eigenvalues of (6.1) by looking for the roots of the characteristic equation

$$\det(A - \lambda B) = 0$$

where $A = (a_{jk})$ and $B = (b_{jk})$ for a_{jk}, b_{jk} defined in (6.4) for some choice of trial functions $\{w_i\}$ for Ω. From Lemma 7, we have shown that the largest root of $\det(A - \lambda B)$ is given by

$$\lambda_n^* = \max_{\substack{c \in \mathbb{R}^n \\ c \neq 0}} \frac{Ac \cdot c}{Bc \cdot c}.$$

Therefore, for a fixed set of trial functions $w_1, \ldots, w_n$ for a given set $\Omega \subset \mathbb{R}^n$, we have the following formula for the approximation of the nth eigenvalue of (6.1). Let $A = (a_{jk}) = (\langle \nabla w_j, \nabla w_k \rangle)$, $B = (b_{jk}) = (\langle w_j, w_k \rangle)$. We see that

$$\begin{aligned} Ac \cdot c &= \left\langle \sum_{j=1}^n c_j \nabla w_j, \sum_{k=1}^n c_k \nabla w_k \right\rangle \\ Bc \cdot c &= \left\langle \sum_{j=1}^n c_j w_j, \sum_{k=1}^n c_k w_k \right\rangle. \end{aligned}$$

Therefore, for A, B defined in terms of the trial functions $w_1, \ldots, w_n$, we see that the largest root of the characteristic equation $\det(A - \lambda B)$ is given by

$$\lambda_n^*(w_1, \ldots, w_n) = \max_{c \in \mathbb{R}^n} \left\{ \frac{||\nabla w||^2}{||w||^2} : w = \sum_{i=1}^n c_i w_i \right\}. \tag{6.8}$$

This value $\lambda_n^*(w_1, \ldots, w_n)$ will give us an approximation to the nth eigenvalue of (6.1). We will show below that if we take the *minimum* of $\lambda_n^*(w_1, \ldots, w_n)$ over all possible sets of trial functions, then we will get the *exact* value of the nth eigenvalue of (6.1).

Theorem 9. (Minimax Principle) *Let Y denote the set of trial functions associated with* (6.1) *(see Theorem 4). The nth eigenvalue of* (6.1) *is given by*

$$\lambda_n = \min_{(w_1, \ldots, w_n) \in Y} \lambda_n^*(w_1, \ldots, w_n).$$

That is, the minimum is taken over all possible sets of n linearly independent trial functions.

Proof. First, we will show that $\lambda_n \leq \min \lambda_n^*$. Fix n linearly independent trial functions $w_1, \ldots, w_n$. Let

$$w(x) \equiv \sum_{j=1}^n c_j w_j(x)$$

be a linear combination of the n trial functions such that w is orthogonal to the first $n-1$ eigenfunctions $v_1, \ldots, v_{n-1}$ of (6.1). That is, choose c_j such that

$$\langle w, v_k \rangle = \sum_{j=1}^{n} c_j \langle w_j, v_k \rangle = 0 \qquad \text{for } k = 1, \ldots, n-1.$$

We know we can solve this system, because we have only $n-1$ equations for our n unknowns $c_1, \ldots, c_n$. Now from the Minimum Principle for the nth Eigenvalue, we know that

$$\lambda_n \leq \frac{||\nabla w||^2}{||w||^2},$$

because

$$\lambda_n = \min_{v \in Y_n} \left\{ \frac{||\nabla v||^2}{||v||^2} \right\},$$

where Y_n is as defined in Theorem 5, and $w \in Y_n$. Therefore,

$$\lambda_n \leq \frac{||\nabla w||^2}{||w||^2} \leq \max_{c \in \mathbb{R}^n} \frac{\left|\left|\nabla \sum_{j=1}^n c_j w_j\right|\right|^2}{\left|\left|\sum_{j=1}^n c_j w_j\right|\right|^2} = \lambda_n^*(w_1, \ldots, w_n).$$

Now taking the minimum of both sides over all possible sets of n linearly independent trial functions, we see that

$$\boxed{\lambda_n \leq \min_{(w_1, \ldots, w_n) \in Y} \lambda_n^*(w_1, \ldots, w_n).}$$

Now, we need to show that $\lambda_n \geq \min \lambda_n^*$. In particular, we will show there exists a particular choice of trial functions $w_1, \ldots, w_n$ such that $\lambda_n^*(w_1, \ldots, w_n) \leq \lambda_n$. Let $w_1, \ldots, w_n$ be the first n eigenfunctions of (6.1) with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Without loss of generality, we may assume they are orthogonal and normalized.

By definition,

$$\lambda_n^*(w_1, \ldots, w_n) = \max_{c \in \mathbb{R}^n} \left\{ \frac{\left|\left|\nabla \sum_{j=1}^n c_j w_j\right|\right|^2}{\left|\left|\sum_{j=1}^n c_j w_j\right|\right|^2} \right\}.$$

Now, using the fact that the w_j are eigenfunctions, orthogonal, and normalized, we have

$$\begin{aligned}
\left|\left|\nabla \sum_{j=1}^n c_j w_j\right|\right|^2 &= \left\langle \sum_{j=1}^n c_j \nabla w_j, \sum_{j=1}^n c_j \nabla w_j \right\rangle = -\left\langle \sum_{j=1}^n c_j w_j, \sum_{j=1}^n c_j \Delta w_j \right\rangle \\
&= \left\langle \sum_{j=1}^n c_j w_j, \sum_{j=1}^n \lambda_j c_j w_j \right\rangle = \sum_{j=1}^n \langle c_j w_j, \lambda_j c_j w_j \rangle = \sum_{j=1}^n \lambda_j c_j^2 \langle w_j, w_j \rangle = \sum_{j=1}^n \lambda_j c_j^2.
\end{aligned}$$

Again, using the fact that the w_j are orthogonal and normalized, we have

$$\left|\left|\sum_{j=1}^n c_j w_j\right|\right|^2 = \left\langle \sum_{j=1}^n c_j w_j, \sum_{j=1}^n c_j w_j \right\rangle = \sum_{j=1}^n \langle c_j w_j, c_j w_j \rangle = \sum_{j=1}^n c_j^2 \langle w_j, w_j \rangle = \sum_{j=1}^n c_j^2.$$

Therefore, we have

$$\lambda_n^*(w_1,\ldots,w_n) = \max_{c\in\mathbb{R}^n} \frac{\sum_{j=1}^n \lambda_j c_j^2}{\sum_{j=1}^n c_j^2} \le \max_{c\in\mathbb{R}^n} \frac{\sum_{j=1}^n \lambda_n c_j^2}{\sum_{j=1}^n c_j^2} = \lambda_n.$$

Therefore, for this choice of trial functions $w_1,\ldots,w_n$, we have

$$\lambda_n^*(w_1,\ldots,w_n) \le \lambda_n,$$

and, consequently,

$$\boxed{\min_{(w_1,\ldots,w_n)\in Y} \lambda_n^*(w_1,\ldots,w_n) \le \lambda_n.}$$

Therefore, our theorem is proved.

□

We will now use the minimax principle to prove the following theorem

Theorem 10. *If $\Omega_1 \subset \Omega_2$, then $\lambda_n(\Omega_1) \ge \lambda_n(\Omega_2)$, where $\lambda_n(\Omega_i)$ is the nth eigenvalue of the Dirichlet problem* (6.1) *on Ω_i.*

Proof. Let $Y(\Omega_i)$ be the set of trial functions for Ω_i, $i = 1, 2$. Recall

$$Y(\Omega_i) = \{w : w \in C^2(\Omega_i) : w \not\equiv 0, w = 0 \text{ for } x \in \partial\Omega_i\}.$$

For emphasis, we let

$$\lambda_n^*(w_1,\ldots,w_n)|_{\Omega_i} \equiv \lambda_n^*(w_1,\ldots,w_n) \quad \text{where the } L^2 \text{ norms are taken over } \Omega_i.$$

By the minimax principle, we know that

$$\begin{aligned}
\lambda_n(\Omega_i) &= \min_{(w_1,\ldots,w_n)\in Y(\Omega_i)} \lambda_n^*(w_1,\ldots,w_n)|_{\Omega_i} \\
&= \min_{(w_1,\ldots,w_n)\in Y(\Omega_i)} \max_{c\in\mathbb{R}^n} \left\{ \frac{||\nabla w||^2_{L^2(\Omega_i)}}{||w||^2_{L^2(\Omega_i)}} : w = \sum_{i=1}^n c_i w_i \right\}.
\end{aligned}$$

For a fixed set of trial functions $w_1,\ldots,w_n$ for Ω_1, let $c^* = c^*(w_1,\ldots,w_n)$ be the vector in $\mathbb{R}^n$ which maximizes the Rayleigh quotient. That is,

$$\lambda_n^*(w_1,\ldots,w_n)|_{\Omega_i} = \frac{||\nabla w^*||^2_{L^2(\Omega_1)}}{||w^*||^2_{L^2(\Omega_1)}} \qquad \text{where } w^* \equiv \sum_{i=1}^n c_i^* w_i.$$

Now, we can extend each of the trial functions w_i to be a trial function for Ω_2 by extending w_i to be zero outside Ω_1. Let $\widetilde{w}_i$ denote w_i extended to Ω_2 in this way. Therefore, it is clear that

$$\lambda_n^*(\widetilde{w}_1,\ldots,\widetilde{w}_n)|_{\Omega_2} = \frac{||\nabla \widetilde{w}^*||^2_{L^2(\Omega_2)}}{||\widetilde{w}^*||^2_{L^2(\Omega_2)}} \qquad \text{where } \widetilde{w}^* \equiv \sum_{i=1}^n c_i^* \widetilde{w}_i.$$

As the functions $\widetilde{w}_i$ are zero outside Ω_1, we see that

$$\lambda_n^*(w_1,\ldots,w_n)|_{\Omega_1} = \lambda_n^*(\widetilde{w}_1,\ldots,\widetilde{w}_n)|_{\Omega_2}. \tag{6.9}$$

Now suppose $v_1, \ldots, v_n$ are the n trial functions for Ω_1 which minimize $\lambda_n^*(w_1, \ldots, w_n)|_{\Omega_1}$. Then using (6.9), we see that

$$\lambda_n^*(\widetilde{v}_1, \ldots, \widetilde{v}_n)\big|_{\Omega_2} = \lambda_n^*(v_1, \ldots, v_n)\big|_{\Omega_1},$$

and, therefore,

$$\lambda_n(\Omega_2) = \min_{\text{trial functions} \in Y(\Omega_2)} \lambda_n^*|_{\Omega_2} \le \lambda_n^*(v_1, \ldots, v_n)\big|_{\Omega_1} = \lambda_n(\Omega_1),$$

as claimed.

□

Remark. In the above theorem, when extending the functions w_i to Ω_2 by defining w_i to be zero outside Ω_1, we overlooked the smoothness issues of the extended functions. This is a technical point which we will not get into here.

Corollary 11. *For Ω a bounded subset of $\mathbb{R}^n$, the eigenvalues of the Dirichlet problem,*

$$\begin{cases} -\Delta u = \lambda u & x \in \Omega \\ u = 0 & x \in \partial\Omega \end{cases}$$

form an infinite sequence $\{\lambda_n\}$ such that $\lambda_n \to +\infty$ as $n \to +\infty$.

Proof. For Ω a bounded subset of $\mathbb{R}^n$, let

$$R \equiv \{(x_1, \ldots, x_n) \in \mathbb{R}^n : |x_i| \le M, i = 1, \ldots, n\}$$

for M sufficiently large such that $\Omega \subset R$. Now we can explicitly calculate the eigenvalues of R. In particular, the eigenvalues are given by

$$\lambda_{m_1,\ldots,m_n}(R) = \sum_{i=1}^{n} \left(\frac{m_i \pi}{2M}\right)^2,$$

where the m_i are positive integers. We see that these eigenvalues form an infinite sequence which goes to infinity. By the above theorem, we know that the corresponding eigenvalues for Ω satisfy $\lambda_m(\Omega) \ge \lambda_m(R)$. Therefore, we conclude that the eigenvalues of Ω form an infinite sequence which goes to infinity. □

We now turn to proving the completeness of the eigenfunctions for the Dirichlet problem (6.1).

Theorem 12. *The eigenfunctions of the Dirichlet problem* (6.1) *are complete in the L^2 sense. That is to say, if $\{v_n\}$ is the set of eigenfunctions of* (6.1) *for a set $\Omega \subset \mathbb{R}^n$, then for any function $f \in L^2(\Omega)$,*

$$\left\| f - \sum_{n=1}^{N} c_n v_n \right\|_{L^2(\Omega)}^2 = \int_\Omega |f - \sum_{n=1}^{N} c_n v_n|^2 \, dx \to 0 \qquad \text{as } N \to +\infty \tag{6.10}$$

where

$$c_n \equiv \frac{\langle f, v_n \rangle}{\langle v_n, v_n \rangle}.$$

Proof. We will prove this in the case when f is a trial function; that is, $f \in C^2(\Omega)$, $f(x) = 0$ for $x \in \partial\Omega$. To prove this theorem in the case of general $f \in L^2(\Omega)$, you can use the fact that any L^2 function can be approximated in L^2 norm by a C^2 function which vanishes on $\partial\Omega$.

So, below, we assume that f is a trial function. Let

$$r_N(x) = f(x) - \sum_{n=1}^{N} c_n v_n(x),$$

where $c_n = \frac{\langle f, v_n \rangle}{\langle v_n, v_n \rangle}$. Without loss of generality, we may assume the v_n are mutually orthogonal. We claim that $||r_N||_{L^2} \to 0$ as $N \to +\infty$.

First, we will show that r_N is a trial function which is orthogonal to the first $N-1$ eigenfunctions of (6.1). By assumption, f is a trial function. In addition, the eigenfunctions are trial functions. Therefore, r_N is a trial function. We just need to show that it's orthogonal to the first $N-1$ eigenfunctions. Let v_i be one of the first $N-1$ eigenfunctions. Then using the fact that the eigenfunctions v_i are mutually orthogonal, we have

$$\begin{aligned}
\langle r_N, v_i \rangle &= \left\langle f - \sum_{n=1}^{N} c_n v_n, v_i \right\rangle \\
&= \langle f, v_i \rangle - \sum_{n=1}^{N} c_n \langle v_n, v_i \rangle \\
&= \langle f, v_i \rangle - c_i \langle v_i, v_i \rangle \\
&= \langle f, v_i \rangle - \frac{\langle f, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_i \rangle = 0.
\end{aligned}$$

Consequently, we have shown that r_N is a trial function which is orthogonal to the first $N-1$ eigenfunctions. Therefore, by the minimum principle for the Nth eigenvalue, letting Y denote the set of trial functions, we know that

$$\lambda_N = \min\left\{ \frac{||\nabla w||^2}{||w||^2} : w \in Y, \langle w, v_i \rangle = 0 \text{ for } i = 1, \ldots, N-1 \right\} \leq \frac{||\nabla r_N||^2}{||r_N||^2}. \tag{6.11}$$

Next, we will show that $||\nabla r_N||^2 \leq ||\nabla f||^2$. By computation, we see that

$$\begin{aligned}
||\nabla r_N||^2_{L^2(\Omega)} &= \left|\left| \nabla f - \sum_{n=1}^{N} c_n \nabla v_n \right|\right|^2_{L^2(\Omega)} \\
&= \int_\Omega \left| \nabla f - \sum_{n=1}^{N} c_n \nabla v_n \right|^2 \, dx \qquad (6.12) \\
&= \int_\Omega |\nabla f|^2 - 2\nabla f \cdot \sum_{n=1}^{N} c_n \nabla v_n + \left| \sum_{n=1}^{N} c_n \nabla v_n \right|^2 \, dx
\end{aligned}$$

Now, by the Divergence theorem,

$$\begin{aligned}\int_\Omega \nabla f\cdot\nabla v_n\,dx &= -\int_\Omega f\Delta v_n\,dx + \int_{\partial\Omega} f\frac{\partial v_n}{\partial\nu}\,dS(x)\\ &= \lambda_n\int_\Omega f v_n\,dx,\end{aligned}\tag{6.13}$$

using the fact that v_n is an eigenfunction and f vanishes on the boundary.

Next, in a similar manner, we see that

$$\begin{aligned}\int_\Omega \nabla v_m\cdot\nabla v_n\,dx &= -\int_\Omega v_m\Delta v_n\,dx + \int_{\partial\Omega} v_m\frac{\partial v_n}{\partial\nu}\,dS(x)\\ &= \lambda_n\int_\Omega v_m v_n\,dx.\end{aligned}$$

Now using the fact that the eigenfunctions v_n are mutually orthogonal, we have

$$\int_\Omega \nabla v_m\cdot\nabla v_n\,dx = \begin{cases}\lambda_n\int_\Omega v_n^2\,dx & m=n\\ 0 & m\neq n.\end{cases}\tag{6.14}$$

Now putting (6.13) and (6.14) into (6.12), we have

$$||\nabla r_N||^2_{L^2(\Omega)} = \int_\Omega |\nabla f|^2 - 2\sum_{n=1}^N c_n\lambda_n f v_n + \sum_{n=1}^N c_n^2\lambda_n v_n^2\,dx.$$

Now substituting in for c_n, we have

$$\begin{aligned}\int_\Omega c_n^2\lambda_n v_n^2\,dx &= c_n^2\lambda_n\int_\Omega v_n^2\,dx\\ &= c_n\frac{\langle f, v_n\rangle}{\langle v_n, v_n\rangle}\lambda_n\int_\Omega v_n^2\,dx\\ &= c_n\lambda_n\langle f, v_n\rangle\\ &= c_n\lambda_n\int_\Omega f v_n\,dx.\end{aligned}$$

Therefore, we have

$$\begin{aligned}||\nabla r_N||^2_{L^2(\Omega)} &= \int_\Omega |\nabla f|^2 - 2\sum_{n=1}^N c_n\lambda_n f v_n + \sum_{n=1}^N c_n\lambda_n f v_n\,dx\\ &= \int_\Omega |\nabla f|^2 - \sum_{n=1}^N c_n\lambda_n f v_n\,dx\\ &= \int_\Omega |\nabla f|^2 - \sum_{n=1}^N c_n^2\lambda_n v_n^2\,dx.\end{aligned}$$

Now using the fact that all eigenvalues of the Dirichlet problem are positive, we see that

$$||\nabla r_N||^2_{L^2(\Omega)} \leq \int_\Omega |\nabla f|^2 \, dx = ||\nabla f||^2_{L^2(\Omega)}. \tag{6.15}$$

Now combining (6.11) and (6.15), we see that

$$\lambda_N \leq \frac{||\nabla r_N||^2}{||r_N||^2} \leq \frac{||\nabla f||^2}{||r_N||^2}.$$

Therefore,

$$||r_N||^2 \leq \frac{||\nabla f||^2}{\lambda_N}.$$

Now, by assumption, $f \in C^2(\Omega)$. Therefore, $||\nabla f||^2$ is bounded. In addition, as we showed earlier, $\lambda_N \to +\infty$ as $N \to +\infty$. Therefore, we conclude that

$$||r_N||^2 \to 0 \quad \text{as } N \to +\infty.$$

$\square$