4. Solvability of elliptic PDEs

4.1 Weak formulation

Let us first consider the Dirichlet problem for the Laplacian with homogeneous boundary conditions in a bounded open domain $\Omega \subset \mathbb{R}^n$ with C^1 boundary,

$$-\Delta u = f \quad \text{in } \Omega, \tag{4.1}$$

$$u = 0 \quad \text{on } \partial\Omega. \tag{4.2}$$

Assuming that $u, f : \bar{\Omega} \to \mathbb{R}$ are smooth functions, and multiplying the equation by a test function $\phi \in C_c^\infty(\Omega)$, we obtain

$$-\int_\Omega \Delta u \phi \, dx = \int_\Omega f \phi \, dx.$$

Integrating by parts, and discarding the boundary terms due to the compact inclusion of the support of ϕ in Ω, we arrive at

$$\int_\Omega Du \cdot D\phi \, dx = \int_\Omega f \phi \, dx \qquad \text{for all } \phi \in C_c^\infty(\Omega). \tag{4.3}$$

Conversely, if f and Ω are smooth, then any smooth u satisfying (4.3) is necessarily a solution of (4.1). However, notice that (4.3) makes sense under much weaker assumptions on Ω, u and f. There is a flexibility in how to weaken the assumptions on u, so for the added structure of Hilbert spaces, we will consider the case of the derivatives of u in the weak sense being in $L^2(\Omega)$. Then, if Du is in L^2, which will be the case if we assume that $u \in H^1(\Omega)$, then the right hand side of (4.3) is well defined by the Cauchy-Schwartz inequality for all $\phi \in C_c^\infty(\Omega)$, and by extension, for all $\phi \in H_0^1(\Omega)$, which is the closure of $C_c^\infty(\Omega)$ under the H^1 norm. The right hand side of (4.3) will be well defined for all $\phi \in H_0^1$, if $f \in L^2$, or more generally, for all $f \in H^{-1} = (H_0^1(\Omega))^*$, in which case we understand the right hand side of (4.3) as a dual pairing of f and ϕ.

Notice that if we understand the solution u in the weaker sense that it only belongs to $H^1(\Omega)$, then the Dirichlet condition (4.2) is not suited for such functions, since they are defined up to almost everywhere, and $\partial\Omega$ has n-dimensional Lebesgue measure zero. We thus weaken the boundary condition to hold in the trace sense, i.e., by requiring that the solution $u \in H^1(\Omega)$ have trace zero on the boundary $\partial\Omega$. But these are exactly $H_0^1(\Omega)$ functions.

All of the above motivates the following definition.

Definition 4.1. Let $\Omega \subset \mathbb{R}^n$ be open, $f \in H^{-1}(\Omega)$. A function $u : \Omega \to \mathbb{R}$ is called a weak solution of (4.1)-(4.2), if

(i) $u \in H_0^1(\Omega)$, and

(ii)
$$\int_\Omega Du \cdot D\phi \, dx = \langle f, v \rangle \qquad \text{for all } \phi \in H_0^1(\Omega). \tag{4.4}$$

The right hand side of (4.4) is the dual pairing, and the weak solution is understood to be a 'function' in L^2 sense, i.e. it's an equivalence class with respect to the almost everywhere pointwise equality.

Remark 4.2. The boundary conditions (4.2) were assumed to be homogeneous for simplicity, however the general case of non-homogeneous boundary conditions can be reduced to this case as follows. Assume that $g : \partial\Omega \to \mathbb{R}$ is in the range of the trace operator $T : H^1(\Omega) \to L^2(\partial\Omega)$, say $g = Tw$, then the weak formulation for the Dirichlet problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= g \quad \text{on } \partial\Omega, \end{aligned}$$

is obtained by replacing u in Definition 4.1 by $(u - w) \in H_0^1(\Omega)$, and f by $-\Delta(u - w) = f + \Delta w = \widetilde{f} \in H^{-1}$. Otherwise the definition is the same.

Remark 4.3. The roots of the weak formulation (4.4) lie in the variational approach to Dircihlet's problem, in which one looks for a solution to the Dirichlet problem as the minimizer to the energy functional

$$J(u) = \frac{1}{2}\int_\Omega |Du|^2\,dx - \langle f, u\rangle.$$

By properly defining the (Fréchet) derivative of this functional, and considering the minimization problem of $J : H_0^1(\Omega) \to \mathbb{R}$, one can show that the minimizer must be a weak solution of (4.1) in the sense of the Definition 4.1.

4.2 Existence of weak solutions of the Dirichlet problem

Using the weak formulation given by definition 4.1, the existence of weak solutions becomes an immediate consequence of the Riesz representation theorem for a suitably defined Hilbert inner product over $H_0^1(\Omega)$, which induces an equivalent norm to the standard $\|\cdot\|_{H_0^1(\Omega)}$ norm.

Theorem 4.4. *Let $\Omega \in \mathbb{R}^n$ be open, bounded in some direction, and $f \in H^{-1}(\Omega)$. Then there exists a unique weak solution $u \in H_0^1(\Omega)$ of the Dirichlet problem* (4.1)-(4.2) *in the sense of Definition 4.1.*

Proof. We define a binary operation on $H_0^1(\Omega)$ as follows

$$(u, v)_0 = \int_\Omega Du \cdot Dv\,dx. \tag{4.5}$$

It's easy to see that this binary operation is an inner product over $H_0^1(\Omega)$, provided Ω is bounded in some direction, and that the induced norm, $\|\cdot\|_0$, defined by $\|u\|_0 = (u, u)_0$, is equivalent to the standard norm $\|\cdot\|_{H_0^1(\Omega)}$ by Poincaré's inequality. This then implies that the space $H_0^1(\Omega)$ equipped with the $(,)_0$ inner product is a Hilbert space, and $f \in H^{-1}(\Omega)$ is a bounded linear functional on $(H_0^1(\Omega), (,)_0)$. But then by Reisz representation theorem, there exists a unique function $u \in H_0^1(\Omega)$, such that

$$(u, \phi)_0 = \langle f, \phi\rangle \qquad \text{for all } \phi \in H_0^1(\Omega), \tag{4.6}$$

which is equivalent to u being a weak solution. □

This approach to weak solvability of the Dirichlet problem for Laplace's equation can be generalized to other elliptic operators. We consider several examples of such (symmetric) operators, and will consider non-symmetric elliptic operators in the next section.

Example 4.5. Consider the Dirichlet problem for the operator $L = -\Delta + I$,

$$\begin{aligned} -\Delta u + u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

We call $u \in H_0^1(\Omega)$ a weak solution of this problem, if

$$\int_\Omega (Du \cdot D\phi + u\phi)\,dx = \langle f, v\rangle \qquad \text{for all } \phi \in H_0^1(\Omega).$$

In analogy to (4.6) This is equivalent to the condition that

$$(u, \phi)_1 = \langle f, \phi \rangle \qquad \text{for all } \phi \in H_0^1(\Omega), \tag{4.7}$$

where $(,)_1$ is the standard inner product on $H_0^1(\Omega)$. Hence, the Riesz representation theorem will again imply the existence of a unique weak solution.

Remark 4.6. In this example $\Omega \subset \mathbb{R}^n$ is a general open set, and doesn't have to be bounded in some direction, since we used the standard inner product, and thus do not rely on Poincaré's inequality to prove equivalence of induced norms. Moreover, (4.7) implies that $\|u\|_{H_0^1} = \|f\|_{H^{-1}}$, and hence, the operator $L = -\Delta + I$ is an isometry of $H_0^1(\Omega)$ onto $H^{-1}(\Omega)$.

Example 4.7. We can slightly generalize the previous example, by considering the operator $L = -\Delta + \mu I$, where $\mu > 0$ is a real number. Given an open domain $\Omega \in \mathbb{R}^n$, and $f \in H^{-1}(\Omega)$, a function $u \in H_0^1(\Omega)$ is a weak solution of the Dirichlet problem

$$\begin{aligned} -\Delta u + \mu u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

if $(u, \phi)_\mu = \langle f, \phi \rangle$ for all $\phi \in H_0^1(\Omega)$, where

$$(u, v)_\mu = \int_\Omega (Du \cdot Dv + \mu uv)\, dx.$$

It's easy to see that the norm $\|\cdot\|_\mu$ induced by this inner product is again equivalent to the standard norm, precisely because $\mu > 0$. Hence, Riesz representation theorem again will imply the existence of a unique weak solution.

Example 4.8. To generalize the result of the previous example to the case of $\mu < 0$, we have to guarantee that the resulting binary operation still defines an inner product, with the associated norm being equivalent to the standard norm. If Poincaré's inequality holds for the domain Ω with some constant C, i.e. $\|u\|_{L^2(\Omega)}^2 \leqslant C\|Du\|_{L^2(\Omega)}^2$, then we will have

$$\int_\Omega \mu u^2\, dx \geqslant -C|\mu| \int_\Omega |Du|^2\, dx,$$

and hence, also

$$(u, u)_\mu = \int_\Omega (|Du|^2 + \mu u^2)\, dx \geqslant (1 - C|\mu|) \int_\Omega |Du|^2\, dx \geqslant \frac{(1 - C|\mu|)}{(1 + C)} \|u\|_{H_0^1(\Omega)}.$$

Thus, if $-1/C < \mu < 0$, then $\|u\|_\mu$ defines a norm on $H_0^1(\Omega)$ equivalent to the standard norm (the other inequality is trivial). The existence of unique solution will then again follow from Riesz representation theorem.

Remark 4.9. For bounded domains the Dirichlet Laplacian has an infinite sequence of real eigenvalues $\{\lambda_n : n \in \mathbb{N}\}$, and it can be shown that the best constant (smallest constant, giving the sharpest inequality) in the Poincaré inequality is exactly the principal eigenvalue λ_1. Then the above method won't work for $\mu < -1/\lambda_1$. Notice that when $\mu = -\lambda_n$, not only the solution may not exist for an arbitrary $f \in H^{-1}(\Omega)$, but even if a weak solution exists, it will not be unique, since adding an eigenfunction to a solution will still be a solution. Hence, we do not expect existence of a unique weak solution when $\mu < -1/C$, where C is the best constant in the Poincaré inequality.

Example 4.10. As the last example before embarking on the study of solvability for general elliptic operators, let as consider the operator

$$Lu = -\sum_{i,j=1}^n \partial_i (a_{ij} \partial_j u), \tag{4.8}$$

where the coefficients are assumed to be bounded, symmetric ($a_{ij} = a_{ji}$), and satisfy the uniform ellipticity condition. That is, for some $\theta > 0$,

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geqslant \theta|\xi|^2 \qquad \text{for all } x \in \Omega, \text{ and all } \xi \in \mathbb{R}^n.$$

The function $u \in H_0^1(\Omega)$ will be a weak solution of the Dirichlet problem for this operator,

$$\begin{aligned} Lu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

if

$$a(u, \phi) = \langle f, \phi \rangle \qquad \text{for all } \phi \in H_0^1(\Omega),$$

where $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is the symmetric bilinear form associated with the operator, and is given by

$$a(u, v) = \sum_{i,j=1}^{n} \int_\Omega a_{ij}\partial_j u \partial_i v \, dx.$$

Now, if Ω is bounded in some direction, then boundedness of a_{ij}, uniform ellipticiy, and the Poincaré inequality will imply that the symmetric bilinear form a defines an inner product on $H_0^1(\Omega)$, with the induced norm being equivalent to the standard norm of $H_0^1(\Omega)$. This will again imply that $f \in H^{-1}$ is a bounded linear functional on the Hilbert space $(H_0^1(\Omega), a)$, and hence the Riesz representation theorem will once again imply the existence of a unique weak solution of the Dirichlet problem for this operator.

Remark 4.11. The bilinear form a of course arises from integration by parts of the left hand side of the equation after multiplying by the function v. Thus, having the derivative in front of the entire term $a_{ij}\partial_j u$ is crucial, since we are not assuming that the coefficients a_{ij} are weakly differentiable. In such cases we will say that the elliptic operator is in the *divergence form*.

4.3 Generel linear elliptic PDEs

As in the previous section, we are interested in solving the PDE

$$Lu = f \quad \text{in } \Omega,$$

subject to homogeneous Dirichlet boundary conditions on $\partial\Omega$. Here we generalize the linear operator L, and consider an operator of the form

$$Lu = -\sum_{i,j=1}^{n} \partial_i(a_{ij}\partial_j u) + \sum_{i=1}^{n} \partial_i(b_i u) + cu. \tag{4.9}$$

Notice that the leading order terms, as well as the first order terms are in the divergence form, which will be useful when studying the weak formulation of the problem. If $a_{ij}, b_i \in C^1(\Omega)$, then an operator in non-divergence form can be always written in the divergence form, by possibly modifying the coefficient of first and zeroth order terms. However, we will assume only boundedness of the coefficients, thus, for the weak formulation the divergence form (of the highest order terms) is necessary.

Definition 4.12. The operator L given by (4.9) is called elliptic at the point $x_0 \in \Omega$, if the matrix $(a_{ij}(x_0))$ is positive definite. And the operator will be elliptic in all of Ω, if it is elliptic at every point.

We will assume the stronger notion of ellipticity, that of uniform ellipticity, given by the next definition.

Definition 4.13. The operator L given by (4.9) is called uniformly elliptic in Ω, if there exists a constant $\theta > 0$, sich that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geqslant \theta|\xi|^2 \tag{4.10}$$

for x almost everywhere in Ω and every $\xi \in \mathbb{R}^n$.

Remark 4.14. Uniform ellipticity means that the eigenvalues of the matrix $(a_{ij}(x))$ are bounded from below by θ uniformly in x almost everywhere in Ω. We will use the uniform ellipticity with the vector $\xi = Du$, which will in turn allow us to control the integral of $|Du|^2$ in terms of the integral of $\sum_{i,j=1}^{n} a_{ij}\partial_i u \partial_j u$.

Example 4.15. The Laplacian, $L = -\Delta$ is uniformly elliptic, since the matrix of coefficients of the leading order terms is the unit matrix, and thus the uniform ellipticity condition (4.10) holds with $\theta = 1$.

Let $\mu \in \mathbb{R}$, and consider the Dirichlet problem for the operator $L + \mu I$,

$$\begin{aligned} Lu + \mu u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{4.11}$$

In the sequel we will always make the following assumptions on the operator L given by (4.9):

(i) (boundedness) the coefficient functions $a_{ij}, b_i, c : \Omega \to \mathbb{R}$ satisfy

$$a_{ij}, b_i, c \in L^\infty(\Omega) \tag{4.12}$$

(ii) (symmetry in the leading terms) the coefficients of the leading terms are symmetric: $a_{ij} = a_{ji}$

(iii) (uniform ellipticity) the operator is uniformly elliptic, i.e. (4.10) holds.

To obtain the weak formulation for the problem (4.11), we proceed as before: multiply the equation by a test function $\phi \in C_c^\infty(\Omega)$, integrate over Ω, and integrate by parts, assuming all the functions as well as the domain are smooth. This leads to the condition that $u \in H_0^1(\Omega)$ is a weak solution of (4.11), if

$$\int_\Omega \left[\sum_{i,j=1}^{n} a_{ij}\partial_i u \partial_j \phi - \sum_{i=1}^{n} b_i u \partial_i \phi + cu\phi \right] dx + \mu \int_\Omega u\phi \, dx = \langle f, \phi \rangle \tag{4.13}$$

for all $\phi \in H_0^1(\Omega)$.

We define the bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ associated with the operator L as

$$a(u, v) = \int_\Omega \left[\sum_{i,j=1}^{n} a_{ij}\partial_i u \partial_j v - \sum_{i=1}^{n} b_i u \partial_i v + cuv \right] dx. \tag{4.14}$$

This form is well-defined on $H_0^1(\Omega)$, and is bounded as we will see later. Notice, however, that it is not symmetric, unless $b_i = 0$. Using this bilinear form, we can write the weak formulation (4.13) in a more concise form.

Definition 4.16. Let $\Omega \in \mathbb{R}^n$ be open, $f \in H^{-1}(\Omega)$, and L is given by (4.9), whose coefficients are bounded, symmetric in the leading terms, and satisfy uniform ellipticity. Then $u : \Omega \to \mathbb{R}$ is a weak solution of (4.11), if:

(i) $u \in H_0^1(\Omega)$, and

(ii)
$$a(u, \phi) + \mu(u, \phi)_{L^2} = \langle f, \phi \rangle \qquad \text{for all } \phi \in H_0^1(\Omega), \tag{4.15}$$

where $(,)_{L^2}$ is the standard inner product of $L^2(\Omega)$.

Since the form a given by (4.14) is not symmetric unless $b_i = 0$, we have

$$a(v,u) = a^*(u,v),$$

where

$$a^*(u,v) = \int_\Omega \left[\sum_{i,j=1}^n a_{ij} \partial_i u \partial_j v - \sum_{i=1}^n b_i (\partial_i u) v + cuv \right] dx. \tag{4.16}$$

This is the bilinear form associated with the formal adjoint L^* of L,

$$L^* u = - \sum_{i,j=1}^n \partial_i (a_{ij} \partial_j u) - \sum_{i=1}^n b_i \partial_i u + cu. \tag{4.17}$$

Using the weak formulation (4.15) via the bilinear form associated with the uniformly elliptic operator L, we would like to prove the existence of a unique weak solution by a method similar to the analogous proof for the Dirichlet Laplacian. In this case, however, the bilinear form a is not symmetric, and cannot be used to define an inner product. Fortunately, a similar result to the Riesz representation theorem holds for non-symmetric bilinear forms as well, which is due to Lax and Milgram.

4.4 Lax-Milgram theorem, solvability of general elliptic PDEs

We will state the Lax-Milgram theorem in the general setting of an abstract Hilbert space $\mathcal{H}$, and will subsequently apply it to the bilinear form associated with the uniformly elliptic operator in the Hilbert space $H_0^1(\Omega)$.

Theorem 4.17. *(Lax-Milgram) Let $\mathcal{H}$ be a Hilbert space with the inner product $(\cdot,\cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$, and let $b : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be a bilinear form on $\mathcal{H}$. Further assume that there exist constants $C_1, C_2 > 0$, such that*

(i) $C_1 \|u\|_{\mathcal{H}}^2 \leqslant b(u,u)$ for all $u \in \mathcal{H}$

(ii) $|b(u,v)| \leqslant C_2 \|u\|_{\mathcal{H}} \|v\|_{\mathcal{H}}$ for all $u, v \in \mathcal{H}$.

Then for every bounded linear functional $f : \mathcal{H} \to \mathbb{R}$, there exists a unique element $u \in \mathcal{H}$, such that

$$\langle f, v \rangle = b(u,v) \qquad \text{for all } v \in \mathcal{H}.$$

Remark 4.18. By using $v = u$ in the second condition in the above theorem, we can understand the two conditions as the two inequalities of the equivalence $b(u,u) \sim \|u\|_{\mathcal{H}}^2$. The first condition is the positive definiteness of the bilinear form, while the second condition is the boundedness. Thus the Lax-Milgram theorem states that every bounded functional on a Hilbert space can be represented by the functional $b(u,\cdot)$, provided the bilinear form b is bounded and positive definite. Since the inner product is bounded and positive definite, we can see that the Lax-Milgram theorem generalized the Riesz representation theorem, and in general no symmetry is assumed for the bilinear form b.

Proof. Notice that for every fixed $u \in \mathcal{H}$, the mapping $v \mapsto b(u,v)$ is a bounded linear functional on $\mathcal{H}$. By the Riesz representation theorem there exists a unique element $w \in \mathcal{H}$, such that

$$b(u,v) = (w,v) \quad \text{for all } v \in \mathcal{H}.$$

Denote the operator mapping u to w by B, i.e. $w = Bu$, and $b(u,v) = (Bu,v)$ for all $v \in \mathcal{H}$.

Using the hypothesis of the theorem, one can show that the operator B is linear, one to one, and that the range of B, $\mathrm{ran}(B)$, is closed in $\mathcal{H}$. These would imply that $\mathrm{ran}(B) = \mathcal{H}$. But then every element of $\mathcal{H}$ has a preimage under B, and from the Riesz representation theorem for f, we have

$$\langle f, v \rangle = (w,v) = (Bu,v) = b(u,v) \quad \text{for all } v \in \mathcal{H},$$

where u is the preimage of the element $w \in \mathcal{H}$ under the operator B.

Uniqueness follows from linearity of b, and condition (i). □

To use the Lax-Milgram theorem to prove the existence of a unique weak solution of (4.11), we need to show that the associated bilinear form satisfies the hypothesis of the theorem. This will depend on the following energy estimates.

Theorem 4.19. *Let a be the bilinear form on $H_0^1(\Omega)$ given by* (4.14)*, and the coefficients are bounded, symmetric in the higher order terms, and satisfy the uniform ellipticity condition* (4.10)*. Then there exist constants $C_1, C_2 > 0$ and $\gamma \in \mathbb{R}$, such that for all $u, v \in H_0^1(\Omega)$, the following estimates hold:*

$$C_1 \|u\|^2_{H_0^1(\Omega)} \leqslant a(u,u) + \gamma \|u\|^2_{L^2(\Omega)} \tag{4.18}$$

$$|a(u,v)| \leqslant C_2 \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}. \tag{4.19}$$

Remark 4.20. The constant γ in inequality (4.18) can be taken to be $\gamma = \theta - c_0$, if $b_i = 0$, and $\gamma = \frac{1}{2\theta}\sum_{i=1}^n \|b_i\|^2_{L^\infty} + \frac{\theta}{2} - c_0$, if $b_i \neq 0$. Here θ is the constant in the uniform ellipticity condition, and $c_0 = \text{ess inf}_\Omega c$.

Proof. The second inequality, (4.19), is the boundedness of the bilinear form a, and follows directly from the boundedness of the coefficients.

The estimate (4.18) is a consequence of the uniform ellipticity. Indeed, by the uniform ellipticity,

$$\begin{aligned} \theta \|Du\|^2_{L^2} = \theta \int_\Omega |Du|^2\, dx &\leqslant \sum_{i,j=1}^n a_{ij} \partial_i u \partial_j u \, dx \\ &\leqslant a(u,u) + \sum_{i=1}^n \int_\Omega b_i u \partial_i u \, dx - \int_\Omega c u^2 \, dx \\ &\leqslant a(u,u) + \sum_{i=1}^n \|b_i\|_{L^\infty} \|u\|_{L^2} \|\partial_i u\|_{L^2} - c_0 \|u\|^2_{L^2}. \end{aligned}$$

Inequality (4.18) would follow, if one uses Cauchy's inequality with ϵ for the middle term on the right, and hides the $\|\partial_i u\|^2_{L^2}$ term with an ϵ coefficient on the left. □

Remark 4.21. The estimate (4.18) is called Garding's inequality, and it is the crucial *a priori* estimate, that establishes the bound for the H_0^1 norm of the solution in terms of the bilinear form of the elliptic operator.

Using Theorem 4.19, we can now apply the Lax-Milgram theorem to problem (4.11).

Theorem 4.22. *Let $\Omega \in \mathbb{R}^n$ be open, $f \in \mathcal{H}^{-1}(\Omega)$, and L be the differential operator* (4.9)*. Suppose the coefficients are bounded, symmetric in the highest order terms, and satisfy the uniform ellipticity condition, and let $\gamma \in \mathbb{R}$ be the constant for which Theorem 4.19 holds. Then for every $\mu \geqslant \gamma$ there exists a unique weak solution $u \in H_0^1(\Omega)$ of the Dirichlet problem* (4.11)*.*

Proof. For $\mu \in \mathbb{R}$, we define the bilinear form $a_\mu : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ by

$$a_\mu(u,v) = a(u,v) + \mu(u,v)_{L^2}, \tag{4.20}$$

where a is the bilinear form associated with the operator L and is given by (4.14).

It is easy to see that a_μ is bounded. It also satisfies condition (i) of the Lax-Milgram theorem by Garding's inequality (4.18), provided $\mu \geqslant \gamma$. Hence, we can apply the Lax-Milgram theorem to show that for every $f \in H^{-1}(\Omega)$, there exists a unique function $u \in H_0^1(\Omega)$, such that

$$\langle f, v \rangle = a_\mu(u,v) \quad \text{for all } v \in H_0^1,$$

which is equivalent to u being a weak solution. □

Remark 4.23. The above proof of existence of a unique weak solution applies to L^* given by (4.17) as well, with a replaced by a^* from (4.16) in the proof, even though the first order term is not in the divergence form.

4.5 Fredholm operators on a Hilbert spaces

The solvability of the problem (4.11) for μ large enough implies that the operator $K = (L+\mu I)^{-1} : \mathcal{H}^{-1}(\Omega) \to H_0^1(\Omega)$ is well defined and, as we will later see, bounded. If we restrict K to $L^2(\Omega)$, and think of it as a map into $L^2(\Omega)$, then, provided Ω is bounded, the operator K will be compact, since $\operatorname{ran}(K) \subset H_0^1(\Omega)$, which is compactly embedded into $L^2(\Omega)$ for bounded Ω by Relich's theorem. The operator $(L-\lambda I)^{-1}$ is called the resolvent of L, thus the above property states that L has a compact resolvent. As we will see in the next section, this fact leads to characterization of solvability of the equation $Lu - \lambda u = f$ for for arbitrary $\lambda \in \mathbb{R}$, $f \in L^2(\Omega)$. In this section we give the formal definitions of compact and Fredholm operators on a Hilbert space, and state some of the properties of such operators without proof.

Let $\mathcal{H}$ be a Hilbert space equipped with the inner product $(,)$, and the associated norm $\|\cdot\|$. The space of bounded linear operators $T : \mathcal{H} \to \mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$. This space is a Banach space with respect to the operator norm

$$\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \in \mathcal{H}, x \neq 0\right\}.$$

The adjoint of $T \in \mathcal{L}(\mathcal{H})$ is the linear operator $T^* \in \mathcal{L}(\mathcal{H})$, such that

$$(Tx, y) = (x, T^*y) \qquad \text{for all } x, y \in \mathcal{H}.$$

An operator is self-adjoint, if $T = T^*$. The kernel and range of $T \in \mathcal{L}(\mathcal{H})$ are the subspaces

$$\ker(T) = \{x \in \mathcal{H} : Tx = 0\}, \quad \operatorname{ran}(T) = \{y \in \mathcal{H} : y = Tx \text{ for some } x \in \mathcal{H}\}.$$

Definition 4.24. A linear operator $T \in \mathcal{L}(\mathcal{H})$ is called compact, if it maps bounded sets to precompact sets.

This is equivalent to the following: for every bounded sequence $\{x_n\} \subset \mathcal{H}$, there exists a converging subsequence of the sequence $\{Tx_n\}$.

Example 4.25. Any bounded linear map of finite rank, i.e. a linear operator whose range is finite dimensional is compact. As a consequence, every linear operator on a finite-dimensional Hilbert space is compact.

For compact self-adjoint operators the following spectral theorem holds.

Theorem 4.26. *Let $T \in \mathcal{L}(\mathcal{H})$ be a compact self-adjoint operator. T has at most countably many distinct real eigenvalues. If there are infinitely many eigenvalues $\{\lambda_n \in \mathbb{R}, n \in \mathbb{N}\}$, then necessarily $\lambda_n \to 0$ as $n \to \infty$. The eigenspace corresponding to each nonzero eigenvalue is finite dimensional, and the eigenvectors associated with distinct eigenvalues are orthogonal. Moreover, $\mathcal{H}$ has an orthonormal basis consisting of eigenvectors of T, including those, if any, for eigenvalue zero.*

Remark 4.27. The fact that the eigenvalues are real and the corresponding eigenspaces are mutually orthogonal follows from self-adjointness of the operator. The compactness of the operator, on the other hand, implies that the spectrum cannot have an accumulation point other than zero, since otherwise we can always choose a bounded set consisting of unit pairwise orthogonal eigenvectors corresponding to the eigenvalues converging to the nonzero accumulation point, which will not have a precompact image. Hence, there can be at most countably many eigenvalues, and they must converge to zero.

This spectral theorem will be used to characterize the spectrum of a uniformly elliptic self-adjoint operator on a bounded domain via the spectrum of its compact resolvent.

We next turn to Fredholm operators.

Definition 4.28. A linear operator $T \in \mathcal{L}(\mathcal{H})$ is called a Fredholm operator, if

(i) $\ker(T)$ has finite dimension

(ii) $\operatorname{ran}(T)$ is closed, and has finite codimension.

The projection theorem for Hilbert spaces, coupled with property (*ii*) in the definition implies that $\mathcal{H} = \operatorname{ran}(T) \oplus \operatorname{ran}(T)^{\perp}$, and $\dim \operatorname{ran}(T)^{\perp} = \operatorname{codim} \operatorname{ran}(T) < \infty$.

Definition 4.29. If $T \in \mathcal{L}(\mathcal{H})$ is Fredholm, then the index of T is the integer

$$\operatorname{ind}(T) = \dim \ker(T) - \dim \operatorname{ran}(T)^{\perp}.$$

Example 4.30. Every linear operator $T : \mathcal{H} \to \mathcal{H}$ on a finite dimensional Hilbert space is Fredholm with zero index, since any finite dimensional linear subspace is closed. The index will be zero due to the formula

$$\dim \mathcal{H} = \dim \ker(T) + \dim \operatorname{ran}(T).$$

Example 4.31. The identity map I on any Hilbert space is Fredholm. Moreover, $\dim \ker(I) = \operatorname{codim} \operatorname{ran}(I) = 0$, and hence, $\operatorname{ind}(I) = 0$.

Using the definition of the Fredholm operator, it's not hard to see the following property.

Theorem 4.32. *If $T \in \mathcal{L}(\mathcal{H})$ is Fredholm, then so is T^*, and*

$$\dim \ker(T^*) = \operatorname{codim} \operatorname{ran}(T), \quad \operatorname{codim} \operatorname{ran}(T^*) = \dim \ker(T), \quad \operatorname{ind}(T^*) = -\operatorname{ind}(T).$$

One can show that the set of compact operators is open as a subset of $\mathcal{L}(\mathcal{H})$ in the topology of the operator norm. Moreover, the set of Fredholm operators is closed under addition of compact operators.

Theorem 4.33. *Suppose $T \in \mathcal{L}(\mathcal{H})$ is Fredholm, and $K \in \mathcal{L}(\mathcal{H})$ is compact, then:*

(i) *There exists an $\epsilon > 0$, such that for any $H \in \mathcal{H}$ with $\|H\| < \epsilon$, $T + H$ is Fredholm. Moreover, for every such operator,* $\operatorname{ind}(T + H) = \operatorname{ind}(T)$.

(ii) *$T + K$ is Fredholm and* $\operatorname{ind}(T + K) = \operatorname{ind}(T)$.

Remark 4.34. The first statement of the theorem implies that not only the set of Fredholm operators is open in the operator norm topology, but that it is the union of connected components characterized by the index.

For Fredholm operators with zero index the following result holds, know as the Fredholm alternative, which characterizes the solvability of the linear equation corresponding to a Fredholm operator.

Theorem 4.35. *Let $T \in \mathcal{L}(\mathcal{H})$ be a Fredholm operator with* $\operatorname{ind}(T) = 0$*, then one of the following alternatives holds:*

(1) $\ker(T^*) = \ker(T) = 0; \operatorname{ran}(T) = \operatorname{ran}(T^*) = \mathcal{H}$.

(2) $\ker(T^*) \neq 0; \dim \ker(T) = \dim \ker(T^*) < \infty; \operatorname{ran}(T) = \ker(T^*)^{\perp}, \operatorname{ran}(T^*) = \ker(T)^{\perp}$.

Remark 4.36. The Fredholm alternative for the Fredholm operator T with zero index can be interpreted as the solvability of the linear equation $Tx = y$. Indeed, the two alternatives are equivalent to the following:

(1) $T^*z = 0$ has the only solution $z = 0$; and $Tx = y$ has a unique solution $x \in \mathcal{H}$ for every $y \in \mathcal{H} = \operatorname{ran}(T)$.

(2) $T^*z = 0$ has a nonzero solution, in which case the dimension of the solution space is equal to the dimension of the solution space of the equation $Tx = 0$; the equation $Tx = y$ is solvable, iff $(y, z) = 0$ for every z solving $T^*z = 0$.

Remark 4.37. The Fredholm alternative is a consequence of the fact that, if $T \in \mathcal{L}(\mathcal{H})$, then

$$\mathcal{H} = \overline{\operatorname{ran}(T)} \oplus \ker(T^*), \quad \text{and } \overline{\operatorname{ran}(T)} = \ker(T^*)^{\perp}.$$

In the case of a Fredholm operator, there are finitely many solvability conditions expressed by the orthogonality in the second alternative.

4.6 The Fredholm alternative for elliptic equations

As we mentioned in the beginning of the previous section, Theorem 4.22 for the weak solvability implies that the operator $L + \mu I$ for $\mu \geqslant \gamma$ is invertible, and we may define the inverse operator $K = (L + \mu I)^{-1}$, which maps $H^{-1}(\Omega)$ onto $H_0^1(\Omega)$. That is, $Kf = u$, iff $a_\mu(u, v) = \langle f, v\rangle$ for all $v \in H_0^1(\Omega)$, where a_μ is the bilinear form (4.20).

Clearly K is linear, and it is bounded due to the Garding inequality (4.18). Let us now assume that Ω is bounded. If we restrict K to $L^2(\Omega)$, and use the compact embedding of $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ for bounded domains, then the map

$$K : L^2(\Omega) \to H_0^1(\Omega) \hookrightarrow L^2(\Omega)$$

maps bounded sets in $L^2(\Omega)$ to bounded sets in $H_0^1(\Omega)$, which are precompact in $L^2(\Omega)$. Hence, as a map from $L^2(\Omega)$ to $L^2(\Omega)$, K is compact.

If $f \in L^2(\Omega)$, then for the dual pairing of $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ we have $\langle f, v\rangle = (f, v)_{L^2}$. Hence,

$$Kf = u \quad \text{iff} \quad a_\mu(u, v) = (f, v)_{L^2} \qquad \text{for all } v \in H_0^1(\Omega). \tag{4.21}$$

Using the bilinear form $a_\mu^*(u, v) = a^*(u, v) + (u, v)_{L^2}$, where a^* is the bilinear form associated with L^*, the formal dual of L, given by (4.16), we can define the operator K^* in a similar way:

$$K^*g = v \quad \text{iff} \quad a_\mu^*(v, u) = (g, u)_{L^2} \qquad \text{for all } u \in H_0^1(\Omega). \tag{4.22}$$

That is,

$$K^* = (L^* + \mu I)^{-1}\big|_{L^2(\Omega)}.$$

It is not hard to see that K^* is the adjoint of the operator K.

Theorem 4.38. *The operator K defined by* (4.21) *is a linear bounded operator $K : L^2(\Omega) \to L^2(\Omega)$. Its adjoint is the operator K^* given by* (4.22)*. If Ω is bounded, then K is a compact operator.*

Proof. Boundedness follows directly from Garding's inequality (4.18). To show that K^* is the adjoint of K, take $f, g \in L^2(\Omega)$, for which $Kf = u$, $K^*g = v$. Then using (4.21) and (4.22), we have

$$(Kf, g)_{L^2} = (u, g)_{L^2} = (g, u)_{L^2} = a_\mu^*(v, u) = a_\mu(u, v) = (f, v)_{L^2} = (f, K^*g).$$

Compactness follows from Relich's theorem, as explained above. ◻

Observe that, if K is compact on $L^2(\Omega)$, then so is the operator σK for every $\sigma \in \mathbb{R}$. But then the operator $(I + \sigma K)$ will be Fredholm by Theorem 4.33, and $\operatorname{ind}(I + \sigma K) = \operatorname{ind}(I) = 0$. Hence, the Fredholm alternative, Theorem 4.35 holds for this operator. We then have a Fredholm alternative for the elliptic operator $(L - \lambda I)$ as well, as encapsulated in the following theorem.

Theorem 4.39. *Let $\Omega \subset \mathbb{R}$ be open and bounded, and L is a uniformly elliptic operator* (4.9)*, for which Theorem 4.22 holds. Let L^* be the formal adjoint of L, given by* (4.17)*, and $\lambda \in \mathbb{R}$. Then one of the following alternatives holds:*

(1) *The only weak solution of the equation $L^*v - \lambda v = 0$ is $v = 0$. For every $f \in L^2(\Omega)$ there exists a unique weak solution $u \in H_0^1(\Omega)$ of the equation $Lu - \lambda u = f$. In particular, the only solution of $Lu - \lambda u = 0$ is $u = 0$.*

(2) *The equation $L^*v - \lambda v = 0$ has a nonzero weak solution v. The solution space of the equations $Lu - \lambda u = 0$ and $L^*v - \lambda v = 0$ are finite dimensional and have the same dimension. For $f \in L^2(\Omega)$ the equation $Lu - \lambda u = f$ has a weak solution $u \in H_0^1(\Omega)$, iff $(f, v)_{L^2} = 0$ for every weak solution $v \in H_0^1(\Omega)$ of $L^*v - \lambda v = 0$, and if the solution u exists, it is not unique.*

Remark 4.40. Notice that for smooth u, v, for which we can perform integration by parts, $(Lu, v) = (u, L^*v)$, and we can see that if $Lu = f$, then necessarily

$$(f, v) = (Lu, v) = (u, L^*v) = 0$$

for all v solving $L^*v = 0$. The Fredholm alternative implies that this condition is also sufficient for the solvability of the elliptic equations, which is a consequence of the index of the operator I being equal to zero.

Proof. Since $K = (L + \mu I)^{-1}$ is compact, and hence $(I + \sigma K)$ is Fredholm on $L^2(\Omega)$, the Fredholm alternative, Theorem 4.35, holds for the equation

$$u + \sigma K u = g \qquad u, g \in L^2(\Omega), \tag{4.23}$$

for any $\sigma \in R$. We consider the two alternatives separately.

(1) Suppose the only solution of $v + \sigma K^* v = 0$ is $v = 0$. Then, applying $(L^* + \mu I)$ to this equation, we see that the only solution of $L^*v + (\mu + \sigma)v = 0$ is $v = 0$. The Fredholm alternative then implies that for every $g \in L^2(\Omega)$ there is a unique solution of (4.23). Now take an arbitrary function $f \in L^2(\Omega)$, and let $g = Kf$, then the unique solution of (4.23) for this g will be in the range of K. We may then apply $(L + \mu I)$ to (4.23) to conclude that there is a weak solution $u \in \operatorname{ran}(K) \subset H_0^1(\Omega)$ of the equation

$$Lu + (\mu + \sigma)u = f. \tag{4.24}$$

This solution must be unique, since otherwise (4.23) would have multiple solutions. Taking $\sigma = -(\lambda + \mu)$ leads to the first alternative in the theorem.

(2) Suppose $v + \sigma K^* v = 0$ has a finite dimensional subspace of solutions $v \in L^2(\Omega)$. Then $v = -\sigma K^* v \in \operatorname{ran}(K^*)$, and applying $(L^* + \mu I)$ to this equation leads to

$$L^*v + (\mu + \sigma)v = 0.$$

By the Fredholm alternative, equation $u + \sigma K u = 0$ has a finite dimensional solution space of the same dimension, and by applying $(L + \mu I)$ to this, so does the equation

$$Lu + (\mu + \sigma)u = 0.$$

Also, for any $f \in L^2(\Omega)$, (4.23) is solvable for $g = Kf$, if and only if $(g, v) = 0$ for all solutions v of $v + \sigma K^* v = 0$. But

$$(v, g)_{L^2} = (v, Kf)_{L^2} = (K^* v, f)_{L^2} = -\frac{1}{\sigma}(v, f)_{L^2},$$

and hence $Lu - \lambda u = f$ will have a weak solution iff (4.23) does for $\sigma = -(\lambda + \mu)$ and $g = Kf$, which by Fredholm alternative will happen iff $(v, g) = 0$, or equivalently $(f, v) = 0$ for every solution v of $v + \sigma K^* v = 0$. But the solutions of the last equation are exactly the solutions of $L^*v - \lambda v = 0$. □

4.7 The spectrum of a self-adjoint elliptic operator

Suppose L is a symmetric uniformly elliptic operator in some domain $\Omega \subset \mathbb{R}^n$ of the form

$$Lu = -\sum_{i,j=1}^{n} \partial_i (a_{ij} \partial_j u) + cu, \tag{4.25}$$

where $a_{ij} = a_{ji}$, and $a_{ij}, c \in L^\infty(\Omega)$. The associated symmetric bilinear form will be

$$a(u, v) = \int_\Omega \left(\sum_{i,j=1}^{n} a_{ij} \partial_i u \partial_j v + cuv \right) dx.$$

If the domain Ω is bounded, then the resolvent $K = (L+\mu I)^{-1}$ is a compact self-adjoint operator on $L^2(\Omega)$ for μ large enough. Hence, Theorem 4.26 holds for this K. Since, as we saw in the last section, L has the same eigenfunctions as K, we have a corresponding spectral theorem for the elliptic operator L.

Theorem 4.41. *Let $\Omega \subset \mathbb{R}^n$ be open, bounded, then the operator L given by* (4.25) *has an increasing sequence of real eigenvalues of finite multiplicity*

$$\lambda_1 < \lambda_2 \leqslant \lambda_3 \leqslant \cdots \leqslant \lambda_n \leqslant \ldots,$$

such that $\lambda_n \to \infty$. Moreover, there is an orthonormal basis $\{\phi_n : n \in \mathbb{N}\}$ of $L^2(\Omega)$ consisting of eigenfunctions $\phi_n \in H_0^1(\Omega)$, which are weak solutions of

$$L\phi_n = \lambda_n \phi_n.$$

Proof. First, notice that if $K\phi = 0$ for some $\phi \in L^2(\Omega)$, then applying $(L+\mu I)$ to this equation will yield $\phi = 0$, so K doesn't have zero as one of its eigenvalues. This in particular will imply that K must necessarily have infinitely many eigenvalues, since otherwise K could have only finitely many linearly independent eigenfunctions, which could not span $L^2(\Omega)$.

Now if $K\phi = \kappa\phi$, for $\phi \in L^2(\Omega)$, then $\phi \in \operatorname{ran}(K) \subset H_0^1(\Omega)$, and applying $(L + \mu I)$ to this equation gives

$$L\phi = \left(\frac{1}{\kappa} - \mu\right)\phi.$$

So ϕ is an eigenfunction of L corresponding to the eigenvalue $\lambda = 1/\kappa - \mu$. This means that $a(\phi,\phi) = \lambda\|\phi\|_{L^2}^2$, hence by Garding's inequality (4.18),

$$C_1\|\phi\|_{H_0^1} \leqslant a(\phi,\phi) + \gamma\|\phi\|_{L^2}^2 = (\lambda+\gamma)\|\phi\|_{L^2}^2.$$

for some $\gamma \in R$. It follows that $\lambda > -\gamma$, and so the eigenvalues of L are bounded from below. The limiting property follows from the spectral theorem for the compact operator K. □

Remark 4.42. The boundedness of the domain Ω is crucial, since otherwise the operator K may not be compact, and the spectrum of L then may not be discrete. As an example, consider the Laplacian $L = -\Delta$ on $\mathbb{R}^n$, which has the purely continuous spectrum $[0,\infty)$.