Brouwer and Galerkin

S.Kesavan
Adjunct Faculty, Indian Institute of Technology, Madras
email: kesh@imsc.res.in

Abstract

We look at an equivalent form of the Brouwer fixed point theorem and explain how it can be used in the Galerkin method for solving equations in Hilbert spaces.

1 The Brouwer fixed point theorem

The two well-known fixed point theorems are the contraction mapping theorem and the Brouwer fixed point theorem. While the former is fairly easy to prove, the latter needs a sophisticated topological tool, *viz.* the topological degree.

Let $\Omega \subset \mathbb{R}^N$ be a bounded domain. We denote its closure by $\overline{\Omega}$ and its boundary by $\partial\Omega$. If $f : \overline{\Omega} \to \mathbb{R}^N$ is a given mapping, a vector $b \in \mathbb{R}^N$ is said to be a boundary value if there exists $x \in \partial\Omega$ such that $f(x) = b$.

Let Ω be as stated above and let $f : \overline{\Omega} \to \mathbb{R}^N$ be a continuous mapping. Let $b \in \mathbb{R}^N$. If b is **not** a boundary value, then the (Brouwer) degree, denoted $d(f, \Omega, b)$ can be defined (cf. Kesavan [2]. Amongst the properties of the degree, the following are most useful.

- If $d(f, \Omega, b) \neq 0$, then there exists $x \in \Omega$ such that $f(x) = b$. Thus, this helps in proving the existence of a solution to the given equation.
- (Homotopy invariance) Let $H : \overline{\Omega} \times [0, 1] \to \mathbb{R}^N$ be a continuous mapping such that for every $t \in [0, 1]$, the point $b \in \mathbb{R}^N$ is not a boundary value of the mapping $H(\cdot, t)$, then
$$d(H(\cdot, 0), \Omega, b) \ = \ d(H(\cdot, 1), \Omega, b).$$
- If I stands for the identity mapping in $\mathbb{R}^N$, then
$$d(I, \Omega, b) \ = , \begin{cases} 1, & \text{if } b \in \Omega, \\ 0, & \text{if } b \in \mathbb{R}^N \backslash \overline{\Omega}. \end{cases}$$

Using these properties, the Brouwer fixed point theorem is proved in two steps, as indicated below.
Let B denote the open unit ball, centered at the origin, in $\mathbb{R}^N$ and let S^{N-1} be its boundary, *i.e.* the unit sphere.
Step 1. First we prove the 'no retraction theorem': there is no continuous map from from $\overline{B}$ onto S^{N-1} which is the identity map when resticted to S^{N-1}.
Step 2: (Brouwer fixed point theorem): Let f be a continuous map of $\overline{B}$ into itself. Then f has a fixed point.

To see this, we assume the contrary. Then, for each $x \in \overline{B}$, the points x and $f(x)$ are distinct and so the line segment joining them is well defined. We produce the line segment starting from $f(x)$ and ending at x to meet the boundary S^{N-1} and call this point, say, $P(x)$. Then one shows that the mapping $x \mapsto P(x)$ is a continuous map (this needs checking) which is, obviously, the identity map when restricted to S^{N-1}, thereby arriving at a contradiction.

In one-dimension, the intermediate value theorem states that if $f : [-1, 1] \to \mathbb{R}$ is a continuous function such that $f(-1)$ and $f(1)$ have opposite signs, then f has to vanish in the interval $(0, 1)$. If, now, $f : [-1, 1] \to [-1, 1]$ is continuous, we can apply the intermediate value theorem to the function $f(x) - x$ and immediately deduce Brouwer's theorem in this case.

We now extend this idea to higher dimensions and deduce Brouwer's theorem. We denote by $(\cdot, \cdot)$, the usual euclidean inner-product, and by $|\cdot|$, the euclidean norm in $\mathbb{R}^N$.

Proposition 1.1 *Let $g : \mathbb{R}^N \to \mathbb{R}^N$ be continuous. Let $R > 0$. Assume that $(g(x), x) \geq 0$, for all $x \in \mathbb{R}^N$ such that $|x| = R$. Then there exists $x_0 \in \mathbb{R}^N$ with $|x_0| \leq R$ such that $g(x_0) = 0$.*

Proof: Assume that $g(x)$ does not vanish for $|x| = R$, for, otherwise, we are done. Define, for $|x| \leq R$ and for $t \in [0, 1]$,

$$H(x, t) \ = \ tg(x) + (1 - t)x.$$

Then, clearly, $H(x, t)$ does not vanish for $|x| = R$ when $t = 0$ and when $t = 1$. Let $t \in (0, 1)$. If $|x| = R$ and if $H(x, t) = 0$, we hve $g(x) = -\frac{1-t}{t}x$ and so

$$0 \ \leq \ (g(x), x) \ = -\frac{1-t}{t}R^2 \ < \ 0,$$

which is a contradiction. Thus $H(\cdot,t)$ does not vanish on the boundary of the ball B_R, centred at the origin and of radius R, for all $t \in [0,1]$. Hence, by the properties of the degree listed earlier, we get

$$d(g, B_R, 0) \;=\; d(I, B_R, 0) \;=\; 1,$$

and so, again by a property of the degree, there exists $x_0 \in B_R$ such that $g(x_0) = 0$. ■

Remark 1.1 We have not only proved the existence of a solution to the equation $g(x) = 0$, but we also have an estimate for the norm of the solution. ■

Remark 1.2 It is clear that the proposition is true if we have the condition $(g(x), x) \leq 0$ for all $x \in \mathbb{R}^N$ such that $|x| = R$. ■

Corollary 1.1 *(Brouwer fixed point theorem) Let B_R be the open ball in $\mathbb{R}^N$, centred at the origin and of radius $R > 0$. Let $f : \overline{B_R} \to \overline{B_R}$ be continuous. Then f has a fixed point.*

Proof: Set $g(x) = x - f(x)$. Then, for $|x| = R$, by the Cauchy-Schwarz inequality, we have

$$(g(x), x) \;=\; R^2 - (f(x), x) \;\geq\; 0,$$

since $|f(x)| \leq R$. The result now follows from the preceding proposition. ■

In fact, all the three results - the Brouwer fixed point theorem the above proposition and the no retraction theorem - are all equivalent, *i.e.* each implies the other two.

Proposition 1.2 *Let $R > 0$ and let B_R denote the open ball in $\mathbb{R}^N$ centred at the origin and of radius R. Let S_R denote its boundary, the sphere with centre at the origin and of radius R. The following statements are equivalent.*
(i) *Let $f : \overline{B_R} \to \overline{B_R}$ be continuous. Then f has a fixed point.*
(ii) *Let $g : \mathbb{R}^N \to \mathbb{R}^N$ be such that $(g(x), x) \geq 0$ for all $x \in S_R$. Then there exists $x_0 \in \overline{B_R}$ such that $g(x_0) = 0$.*
(iii) *There is no continuous mapping from $\overline{B_R}$ onto S_R which is the identity when restricted to S_R.*

Proof: (i) $\Rightarrow$ (ii): Assume that g does not vanish in $\overline{B_R}$. Then the mapping

$$f(x) \;=\; -R\frac{g(x)}{|g(x)|}$$

is well-defined, continuous and maps $\overline{B_R}$ into itself. In fact, the image is contained in S_R. Thus, by (i), there exists a fixed point, say, x_0, of f. Then it follows that $x_0 \in S_R$. But then,

$$0 \;<\; R^2 \;=\; |x_0|^2 \;=\; (f(x_0), x_0) \;=\; -R\frac{(g(x_0), x_0)}{|g(x_0)|} \;\leq\; 0,$$

a contradiction.

(ii) $\Rightarrow$ (iii): If $g : \overline{B_R} \to S_R$ is a continuous mapping which, when restricted to S_R, is the identity, we have, for all $x \in S_R$,

$$(g(x), x) \;=\; |x|^2 \;\geq\; 0.$$

Thus g must vanish somewhere, which is impossible since it takes values only in S_R.

(iii) $\Rightarrow$ (i): This is the standard proof of Brouwer's theorem outlined at the beginning of this section. See Kesavan [2] for details. ■

2 The Galerkin method

The Galerkin method is a useful approximation method to solve linear and nonlinear equations in Hilbert spaces. In this section, we will illustrate the method by using it to prove the famous Lax-Milgram lemma. In the next section, we will illustrate its use in solving a nonlinear equation and we will use the equivalent form of Brouwer's theorem proved in the previous section.

Let A be an $N \times N$ matrix with real entries. It is said to be positive definite if, for every $x \in \mathbb{R}^N,\; x \neq 0$, we have

$$(Ax, x) \;>\; 0.$$

Remark 2.1 If A is a matrix with complex entries and if we still denote the usual inner-product in $\mathbb{C}^N$ by $(\cdot, \cdot)$, then $(Az, z) = 0$ for all $z \in \mathbb{C}^N$ implies that $A = 0$. Similarly if $(Az, z) \geq 0$ for all $z \in \mathbb{C}^N$, it follows that $A = A^*$, *i.e.* A is self-adjoint. Neither of these results is true in the real case as seen from the matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \blacksquare$$

Since the unit ball is compact, we then deduce that, if A is positive definite, there exists $\alpha > 0$ such that

$$(Ax, x) \ \geq \ \alpha |x|^2, \text{ for every } x \in \mathbb{R}^N.$$

It also follows that the linear map associated to A is injective and so A is invertible. The Lax-Milgram lemma is an infinite-dimensional version of this result.

Let H be a real Hilbert space whose norm is denoted by $\|\cdot\|$ and the inner-product by $(\cdot, \cdot)$ and let $a : H \times H \to \mathbb{R}$ be a bilinear form. We say that a is continuous if there exists a constant $M > 0$ such that

$$|a(x, y)| \ \leq \ M\|x\| \ \|y\|,$$

for every x and y in H. The bilinear form is said to be *elliptic* if there exists a constant $\alpha > 0$ such that

$$a(x, x) \ \geq \ \alpha\|x\|^2,$$

for every $x \in H$.

Theorem 2.1 *(Lax-Milgram Lemma) Let H be a real Hilbert space which is separable and let $a : H \times H \to \mathbb{R}$ be a continuous and elliptic bilinear form. Given any $f \in H$, there exists a unique vector $u \in H$ such that*

$$a(u, v) \ = \ (f, v),$$

for every $v \in H$.

Proof: Since H is separable, we can find an orthonormal basis $\{w_i\}_{i=1}^{\infty}$. Let W_m be the span of $\{w_1, \cdots, w_m\}$, where m is a positive integer.

Step 1. Let m be a positive integer. Using the notation established above, we consider the following problem: find $u_m \in W_m$ such that, for every $v \in W_m$,

$$a(u_m, v) \ = \ (f, v). \tag{2.1}$$

By linearity in the second variable of a, it is enough for us to find $u_m \in W_m$ such that

$$a(u_m, w_i) \;=\; (f, w_i), \text{ for } 1 \le i \le m.$$

Since we can write

$$u_m \;=\; \sum_{j=1}^{m} u_j^m w_j,$$

we get the system of m linear equations in the m unknowns $\{u_1^m, \cdots, u_m^m\}$,

$$\sum_{j=1}^{m} a(w_j, w_i) u_j^m \;=\; (f, w_i), \text{ for } 1 \le i \le m.$$

The matrix $A = (a_{ij})$ where $a_{ij} = a(w_j, w_i)$ is positive definite. For,

$$(Ax, x) \;=\; \sum_{i=1}^{m} \sum_{j=1}^{m} a(w_j, w_i) x_i x_j \;= a(w, w) \;\ge\; \alpha \|w\|^2,$$

where $w = \sum_{i=1}^{m} x_i w_i$. Hence, there exists a unique solution u_m to (2.1).

Step 2. (*a priori* estimate) Using the ellipticity of the bilinear form a, we get

$$\alpha \|u_m\|^2 \;\le\; a(u_m, u_m) \;=\; (f, u_m) \;\le\; \|f\| \|u_m\|,$$

whence we get that for all positive integers m,

$$\|u_m\| \;\le\; \frac{\|f\|}{\alpha}.$$

Since $\{u_m\}$ is uniformly bounded in the Hilbert space H, we can extract a weakly convergent subsequence, say, $\{u_{n_k}\}$. Let the weak limit be u. Given $v \in H$, let

$$v^{(m)} \;=\; \sum_{i=1}^{m} v_i w_i,$$

where, $v_i = (v, w_i)$. Then $v^{(m)} \to v$ in H, *i.e.* $\|v^{(m)} - v\| \to 0$. Since $v^{(m_k)} \in W_{m_k}$, we get from (2.1) that

$$a(u_{m_k}, v^{(m_k)}) \;=\; (f, v^{(m_k)}).$$

The right-hand side of this equation obviously converges to (f, v). By the continuity of the bilinear form, for any fixed $z \in H$, the mapping $x \mapsto a(x, z)$ defines a continuous linear functional on H. Hence $a(u_{m_k}, z) \to a(u, z)$, by the definition of weak convergence. Now we have

$$a(u_{m_k}, v^{(m_k)}) \;=\; a(u_{m_k}, v) + a(u_{m_k}, v^{(m_k)} - v).$$

The first term on the right-hand side converges to $a(u, v)$. Since

$$|a(u_{m_k}, v^{(m_k)} - v)| \;\leq\; M\|u_{m_k}\|\|v^{(m_k)} - v\| \;\leq\; \frac{M}{\alpha}\|f\|\|v^{(m_k)} - v\|,$$

it follows that the second term converges to zero. Thus we get that, for all $v \in H$,

$$a(u, v) \;=\; (f, v), \tag{2.2}$$

whch proves the existence of a solution.

Step 3. (Uniqueness) If we have two solutions u_1 and u_2, then, for all $v \in H$, we get that $a(u_1 - u_2, v) = 0$. Set $v = u_1 - u_2$. Thus,

$$\alpha\|u_1 - u_2\|^2 \;\leq\; a(u_1 - u_2, u_1 - u_2) \;=\; 0,$$

which shows that $u_1 = u_2$. This completes the proof. ■

Remark 2.1 In fact $\{u_{m_k}\}$ converges to u in norm. For,

$$\begin{aligned}\alpha\|u - u_{m_k}\|^2 &\leq a(u - u_{m_k}, u - u_{m_k}) \\ &= a(u - u_{m_k}, u - u^{(m_k)}) + a(u - u_{m_k}, u^{(m_k)} - u_{m_k}).\end{aligned}$$

The second term on the right-hand side vanishes since $u^{(m_k)} - u_{m_k} \in W_{m_k}$, in view of (2.1) and (2.2). Thus,

$$\|u - u_{m_k}\| \;\leq\; \frac{M}{\alpha}\|u - u^{(m_k)}\|,$$

and the conclusion follows since the right-hand side of the above inequality tends to zero. ■

Remark 2.2 Given any subsequence of $\{u_m\}$, there is a further subsequence which is weakly convergent. The weak limit satisfies (2.2) which has a unique solution. Therefore, it follows that, in fact, the entire sequence $\{u_m\}$ converges to the unique solution u of (2.2) in norm. ■

For a proof of this theorem, through the context of variational inequalities, see Kesavan [1]. Incidentally, the proof of existence to a variational inequality uses the contraction mapping theorem. Here, we have proved the Lax-Milgram lemma, *directly*, in the case of a *separable* Hilbert space. This is not a serious restriction, since all Hilbert spaces which occur in applications are generally separable.

The Lax-Milgram lemma is the corner stone of the exitence theory for weak solutions of elliptic boundary value problems. For several examples of its use in this direction, see Kesavan [1].

Another example of the Galerkin method to solve a linear problem, without the use of the Lax-Milgram lemma, is the solution of the Schrödinger equation. See Kesavan[1] for details.

3 A nonlinear equation

We will now illustrate the application of the Galerkin method to solve a nonlinear equation (cf. Kesavan [2]). Let H be a separable Hilbert space whose norm and inner-product are denoted by $\|\cdot\|$ and $(\cdot,\cdot)$, respectively. A mapping $A : H \to H$ is said to be *monotone* if, for every $x, y \in H$, we have

$$(Ax - Ay, x - y) \;\geq\; 0.$$

We now establish some notation. Let $\{w_i\}_{i=1}^{\infty}$ be an orthonormal basis of H. We denote by W_m, the span of $\{w_1, \cdots, w_m\}$, which is an m-dimensional subspace of H. Let $v \in W_m$. Then

$$v \;=\; \sum_{i=1}^{m} (v, w_i) w_i.$$

Set $v_i = (v, w_i)$. Consider the vector $\mathbf{v} \in \mathbb{R}^m$ given by $\mathbf{v} = (v_1, \cdots, v_m)$. Thus we have a linear bijection between W_m and $\mathbb{R}^m$. In fact, since we are working with an orthonormal set, it is immediate to see that this bijection is an isometry, *i.e.*

$$\|v\|^2 \;=\; |\mathbf{v}|^2.$$

Theorem 3.1 *Let H be separable Hilbert space. Let $A : H \to H$ be a continuous map which is monotone and which maps bounded sets onto bounded sets. Then, given any $f \in H$, there exists a unique solution $u \in H$ of the equation*

$$u + Au \;=\; f. \tag{3.3}$$

Further,

$$\|u\| \;\leq\; \|A(0) - f\|. \tag{3.4}$$

Proof: Step 1. (Uniqueness) If u_1 and u_2 were two solutions of (3.3), we have

$$u_1 - u_2 + Au_1 - Au_2 \;=\; 0.$$

Thus,

$$\|u_1 - u_2\|^2 + (Au_1 - Au_2, u_1 - u_2) \;=\; 0,$$

and hence, by virtue of the monotonicity of A, we have $u_1 - u_2 = 0$.

Step 2. (*a priori* estimate) If $u \in H$ is a solution of (3.3), then

$$\|u\|^2 + (Au - A0, u) \;=\; (f - A0, u),$$

and the estimate (3.4) follows, again thanks to the monotonicity of A.

Step 3. Let $\{w_i\}_{i=1}^\infty$ be an orthonormal basis of H. Let W_m be as defined above and for $v \in W_m$ we associate the vector $\mathbf{v} \in \mathbb{R}^m$, as explained above. Define $T : \mathbb{R}^m \to \mathbb{R}^m$ by

$$(T\mathbf{v})_i = (v, w_i) + (Av, w_i) - (f, w_i), 1 \leq i \leq m.$$

Then, T is continuous and (denoting the usual inner product in $\mathbb{R}^m$ by $(\cdot, \cdot)_m$) we have

$$\begin{aligned} (T\mathbf{v}, \mathbf{v})_m &= \|v\|^2 + (Av, v) - (f, v) \\ &= \|v\|^2 + (Av - A0, v) - (f - A0, v) \\ &\geq \|v\|^2 - \|A0 - f\|\|v\|. \end{aligned}$$

Setting $R = \|A0 - f\|$, we have that $(T\mathbf{v}, \mathbf{v})_m \geq 0$ for all $|\mathbf{v}| = \|v\| = R$. Hence, by Proposition 1.1, there exists a $\mathbf{u}_m \in \mathbb{R}^m$ such that $|\mathbf{u}_m| \leq R$ and $T\mathbf{u}_m = 0$. Thus, $u_m \in W_m$ satisfies

$$(u_m, w_i) + (Au_m, w_i) \;=\; (f, w_i),\; 1 \leq i \leq m,$$

i.e.

$$(u_m, v) + (Au_m, v) \;=\; (f, v) \;\text{ for all }\; v \in W_m$$

and $\|u_m\| \leq \|A0 - f\|$.

Step 4. Then, there exists a weakly convergent subsequence of $\{u_m\}$. As $\{Au_n\}$ is bounded (since A maps bounded sets into bounded sets), we can also assume (after taking a further subsequence if necessary) that $u_{n_k} \rightharpoonup u$, $Au_{n_k} \rightharpoonup \chi$ weakly in H, for a subsequence indexed by n_k.

Step 5. Given $v \in H$, the sequence $\{v^{(n)}\}$ defined, as before, by

$$v^{(n)} \;=\; \sum_{i=1}^{n} (v, w_i) w_i,$$

is such that $v^{(n)} \in W_n$ for each n and $\|v^{(n)} - v\| \to 0$. Now,

$$(u_{n_k}, v^{(n_k)}) + (Au_{n_k}, v^{(n_k)}) \;=\; (f, v^{(n_k)}). \tag{3.5}$$

Passing to the limit in (3.5), we get

$$(u, v) + (\chi, v) \;=\; (f, v) \;\text{ for all }\; v \in H. \tag{3.6}$$

Step 6. By the monotonicity of A, we have for any $v \in H$,

$$\begin{aligned} 0 \leq X_{n_k} &= (Au_{n_k} - Av, u_{n_k} - v) \\ &= (Au_{n_k}, u_{n_k}) - (Au_{n_k}, v) - (Av, u_{n_k} - v) \\ &= (f, u_{n_k}) - \|u_{n_k}\|^2 - (Au_{n_k}, v) - (Av, u_{n_k} - v), \end{aligned}$$

using (3.5). Thus, $X = \limsup_{k\to\infty} X_{n_k} \geq 0$ and

$$\begin{aligned} X &= (f, u) - \liminf_{k\to\infty} \|u_{n_k}\|^2 - (\chi, v) - (Av, u - v) \\ &\leq (f, u) - \|u\|^2 - (f, v) + (u, v) - (Av, u - v), \end{aligned}$$

using (3.6). Thus,

$$(f - u - Au, u - v) + (Au - Av, u - v) \;\geq\; 0. \tag{3.7}$$

Let $\lambda > 0$ and $w \in H$. Set $v = u - \lambda w$ in (3.7) to get

$$(f - u - Au, w) + (Au - A(u - \lambda w), w) \;\geq\; 0.$$

As $\lambda \to 0$, by the continuity of A, the second term on the left-hand side tends to zero. Thus $(f - u - Au, w) \geq 0$ for all $w \in H$ and, by considering $-w$ in place of w, we conclude that u satisfies (3.3). ■

For an example of the Galerkin method applied to a semilinear elliptic boundary value problem, see Kesavan [1]. The method is always the same. We study the problem in a finite dimensional subspace. The existence of a solution in that space and the uniform bound for that solution follow from Brouwer's theorem (Proposition 1.1). The sequence of approximate solutions will have a weakly convergent subsequence and the weak limit will turn out to be a solution of the original problem. See also Lions [3] for more examples of this technique.

Thus, while the only serious application of the no retraction theorem seems to be the proof of Brouwer's theorem, Proposition 1.1 has many applications, via the Galerkin method, in the theory of existence of solutions to linear and nonlinear boundary value problems.

References

[1] Kesavan, S. Topics in Functional Analysis and Applicatios, Third Edition, New Age International Publishers, 2019.

[2] Kesavan, S. Nonlinear Functional Analysis: A First Course, TRIM Series **28**, Hindustan Book Agency, 2004.

[3] Lions, J. L. Quelques Ḿthodes de Résolution des Problèmes aux Limites Nonlinéaires, Dunod Gauthier-Villars, 1969.