Chapter 2

Results from Linear Analysis

In the last chapter we began to see the need for certain mathematical tools in order to rigorously analyze the finite element method. In an effort to have this book as self-contained as possible, we provide here a short summary of many of the commonly used results from functional analysis.

The main goal of this chapter is to introduce the mathematical tools necessary to precisely formulate and analyze a *general* weak problem and its discrete analogue. The advantage to this abstraction is that we are able to treat a wide class of problems within the same general framework. In later chapters when we investigate particular differential equations, we see that many of the weak formulations fit into this general framework. Thus, if we determine conditions which guarantee existence, uniqueness, and continuous dependence on the data of our general weak problem and derive an error estimate, then we can easily analyze a particular weak formulation by showing that it satisfies the hypotheses for the general problem.

In formulating weak problems we need to determine appropriate classes of function spaces to use for our test and trial spaces and to examine some of their basic properties. The particular types of spaces that are needed are certain Hilbert spaces which are named after the German mathematician David Hilbert (1862-1943). These spaces offer a natural setting for weak problems and can be considered as generalizations of Euclidean space. In general, these Hilbert spaces are infinite dimensional. When attempting to understand various concepts and results on infinite dimensional spaces, it is always helpful to ask oneself what this corresponds to in a finite dimensional setting such as $\mathbb{R}^n$. In many situations we attempt to point out the analogous results in $\mathbb{R}^n$.

We remark that this chapter is by no means a complete exposition of the topic; rather, it is merely intended to prepare the reader for subsequent chapters. For a more detailed exposition of the topics in Section 2.1-2.3, one may consult any functional analysis text; e.g., see [Schechter], [Kreysig], [Yoshida].

2.1 Linear spaces

The goal of this section is to recall some basic definitions for *linear* or *vector spaces*,[1] inner products, and norms and specify some of the notation we use throughout the book. For simplicity of exposition, we only consider real linear spaces.

Definition 2.1. *A* **linear space** *(or* **vector space***) V is a set of objects on which two operations are defined;*[2] *the first determines the sum of two elements belonging to V and the second determines the product of any scalar (a real number) α and any element of V. These sum and product operations must satisfy the following properties:*

(i) $u + v \in V$ *for all* $u, v \in V$*;*

(ii) $u + v = v + u$ *for all* $u, v \in V$*;*

(iii) $u + (v + w) = (u + v) + w$ *for all* $u, v, w \in V$*;*

(iv) *there is an element* $0 \in V$ *such that* $u + 0 = u$ *for all* $u \in V$*;*

(v) *for each* $u \in V$ *there exists an element* $(-u) \in V$ *such that* $u + (-u) = 0$*;*

(vi) $\alpha u \in V$ *for each scalar* α *and all* $u \in V$*;*

(vii) $1u = u$ *for all* $u \in V$*;*

(viii) $\alpha(u + v) = \alpha u + \alpha v$ *for all scalars* α*, and for all* $u, v \in V$*;*

(ix) $(\alpha + \beta)u = \alpha u + \beta u$ *for all scalars* α, β *and for all* $u \in V$*;*

(x) $\alpha(\beta u) = (\alpha\beta)u$ *for all scalars* α, β *and for all* $u \in V$*.*

These axioms are simply the well-known properties satisfied by the set of all vectors in $\mathbb{R}^n$ with the usual definitions for the sum and scalar product operations. However, more general collections of objects such as the set of all continuous functions defined on the interval $[a, b]$ with the usual definitions of sum and product are also linear spaces.

The elements of a linear space V are called *vectors*. An expression of the form

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n ,$$

where $\alpha_i \in \mathbb{R}$ and $u_i \in V$, $i = 1, \ldots, n$, is called a *linear combination* of the vectors u_i. In the simple example of the previous chapter, we saw that our approximate solution was chosen to be a linear combination of functions which formed a basis for the approximating space. The two underlying properties of a basis are linear independence and spanning. Clearly we can always take a linear combination of m vectors and get the zero vector by choosing all the coefficients to be zero. The

[1] The terms are used interchangeably.

[2] There is also an associated field which we always choose to be the real numbers.

concept of linear independence/dependence characterizes whether this is the only way to get the zero vector. Recall that for m vectors in $\mathbb{R}^n$, this reduces to the question of whether the linear system $A\vec{x} = \vec{0}$ has only the trivial solution; here the m columns of A are the vectors. The question of whether a set of m vectors in $\mathbb{R}^n$ span $\mathbb{R}^n$ reduces to the question of whether $A\vec{x} = \vec{b}$ has a unique solution for any $\vec{b} \in \mathbb{R}^n$.

Definition 2.2. *The set of vectors* $\{u_i\}_{i=1}^n$ *is called* **linearly dependent** *if there exist real numbers* α_i, $i = 1, \ldots, n$, *not all of which are zero, such that*

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0\,. \tag{2.1}$$

Otherwise, the set is called **linearly independent***; i.e., the set is linearly independent if the only solution to (2.1) is* $\alpha_i = 0$, $i = 1, \ldots, n$.

Definition 2.3. *A subset of vectors of a finite dimensional vector space* V *is called a spanning set if every vector belonging to* V *can be written as a linear combination of the elements of the subset.*

To define a basis for a linear space, we need enough vectors to span the space but not too many so that they are linearly dependent.

Definition 2.4. *If* V *is a linear space and* $S = \{v_1, v_2, \ldots, v_r\}$ *is a finite set of vectors in* V, *then* S *is called a* **basis** *for* V *if it is a linearly independent spanning set of* V.

To clarify the difference between a finite dimensional and an infinite dimensional linear space, we make the following definition.

Definition 2.5. *A linear space* V *is called finite dimensional of dimension* n *if* V *contains* n *linearly independent elements and if any set of* $(n+1)$ *vectors belonging to* V *is linearly dependent.*

When posing a discrete weak problem, we use a finite dimensional space so we can generate a basis and hence write our approximating solution as a linear combination of the basis elements. In fact, we usually choose our approximating spaces as finite dimensional *subspaces* of the underlying infinite dimensional space on which the weak problem is posed.

Definition 2.6. *A subset* S *of a vector space* V *is called a* **subspace** *of* V *if* $u \in S$ *and* $v \in S$ *implies that* $\alpha u + \beta v \in S$ *for every* $\alpha, \beta \in \mathbb{R}$.

Example 2.7 Consider the infinite dimensional linear space of all continuous functions defined on $\Omega = [0,1]$ with the usual definition of addition and scalar multiplication; we denote this space as $C^0(\Omega)$. Define the following two subsets of $C^0(\Omega)$

$$S_1 = \{v \in C^0(\Omega) \;:\; v(0) = 0\}$$

and

$$S_2 = \{v \in C^0(\Omega) \ : \ v(0) = 1\}\,.$$

The set S_1 is a subspace of $C^0(\Omega)$ since if we take a linear combination of any two continuous functions that are zero at $x = 0$ then the result is a continuous function that is zero at $x = 0$. However, the set S_2 is not a subspace because if we add two functions which are one at $x = 0$ then the resulting function has the value two at $x = 0$. This will be important to us when we are satisfying inhomogeneous boundary conditions. ∎

Mappings or operators on linear spaces play an important role, especially linear mappings.

Definition 2.8. *A mapping f of a linear space V onto a linear space W , denoted $f : V \to W$, is called a linear mapping or equivalently a linear operator provided*

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) \quad \forall\, u, v \in V,\ \alpha, \beta \in \mathbb{R}\,. \tag{2.2}$$

The kernel of a mapping $f : V \to W$ is defined to be the set $\{v \in V \ : \ f(v) = 0\}$ and the range is defined to be the set of all $w \in W$ such that there exists a $u \in V$ where $f(u) = w$.

For example, matrix multiplication using an $m \times n$ matrix is a linear map from $\mathbb{R}^n \to \mathbb{R}^m$. The kernel of the mapping is just the null space of the matrix and the range is just the span of the columns of the matrix.

The structure of a general linear space is not rich enough to be of use in analyzing the finite element method. In this section the goal is to build a particular class of linear spaces which have the properties that we need to state and analyze our problems. In particular, we need a distance function or metric to measure the "size" of a vector, such as an error vector. However, to be useful the metric must be defined in such a way that there is a relationship between the algebraic structure of the vector space and the metric. To guarantee this relationship we first introduce the concept of a *norm* which uses the algebraic properties of the space and then we use the norm to define a metric.

2.1.1 Norms

One familiar distance or metric function is the Euclidean distance formula for measuring the length of a given vector in $\mathbb{R}^n$ or equivalently the distance between two points in $\mathbb{R}^n$. This concept of length of a vector in $\mathbb{R}^n$ can be generalized to include other measures such as the maximum component of a vector in $\mathbb{R}^n$. This generalization is accomplished by introducing the notion of a *norm* on $\mathbb{R}^n$ which is a real-valued function from $\mathbb{R}^n$ to $\mathbb{R}$ satisfying important properties that the Euclidean distance possesses. This concept of a norm can be extended to general linear spaces. A norm on a linear space V can be used to measure the "size" of an element of V, such as the size of an error.

Definition 2.9. *A norm on a linear space V is a real-valued function from V to $\mathbb{R}$, denoted by $\|\cdot\|$, such that*

(i) $\|u\| \geq 0$ *for all* $u \in V$ *and* $\|u\| = 0$ *if and only if* $u = 0$*;*

(ii) $\|\alpha u\| = |\alpha|\ \|u\|$ *for all* $u \in V$ *and all* $\alpha \in \mathbb{R}$*;*

(iii) $\|u + v\| \leq \|u\| + \|v\|$ *for all* $u, v \in V$*.*

The last property is known as the *triangle inequality* due to its interpretation in $\mathbb{R}^n$ using the standard Euclidean norm. In the exercises we consider the three most common norms on $\mathbb{R}^n$.

If we relax the first property of a norm to allow $\|u\| = 0$ for $u \neq 0$, but still require properties (ii) and (iii), then we call the resulting function a *semi-norm.* A linear space V equipped with a norm as defined above is called a *normed linear space* so that we think of a normed linear space as a pair $(V, \|\cdot\|)$.

Example 2.10 If we return to our linear space $C^0(\Omega)$ where $\Omega = [0,1]$ we can define a norm as

$$\|f\| \equiv \max_{x \in [0,1]} |f(x)|\,.$$

Clearly, all three properties of the norm are satisfied. To measure the difference between two vectors, $f, g \in C^0(\Omega)$ we determine

$$\|f - g\| = \max_{x \in [0,1]} |f(x) - g(x)|\,.$$

∎

Since there is always a choice of norms to use on a given vector space, we would like to know if these different measures are somehow comparable.

Definition 2.11. *Two norms, $\|\cdot\|_a$, $\|\cdot\|_b$ defined on a linear space V are said to be equivalent if there are constants C_1, C_2 such that*

$$C_1\|u\|_a \leq \|u\|_b \leq C_2\|u\|_a\,. \tag{2.3}$$

Of course, if (2.3) holds then we also have

$$\frac{1}{C_2}\|u\|_b \leq \|u\|_a \leq \frac{1}{C_1}\|u\|_b\,.$$

In a course in linear algebra, it is usually proved that all norms on $\mathbb{R}^n$ are equivalent. In the exercises, the actual constants in the equivalence relations for the three standard norms on $\mathbb{R}^n$ are investigated; of course, these constants can depend upon n. In functional analysis, one can show a more general, *i.e.,* that in a finite dimensional vector space all norms are equivalent. For the proof of the result, see [Schechter].

Lemma 2.12. *If V is a finite dimensional normed linear space, then all norms are equivalent.*

2.1.2 Inner products

Recall that in dealing with vectors in $\mathbb{R}^n$, one defines a scalar product of two vectors $\vec{a} = (a_1, a_2, \cdots, a_n)$ and $\vec{b} = (b_1, b_2, \cdots, b_n) \in \mathbb{R}^n$ as

$$(\vec{a}, \vec{b}) = \vec{a}^T \vec{b} = \sum_{i=1}^{n} a_i b_i \,.$$

The result of the scalar product is just a number so it can be viewed as a function from $\mathbb{R}^n \to \mathbb{R}$. The scalar product is useful in many applications such as determining if two vectors are perpendicular or equivalently, orthogonal. This concept can be generalized to elements of a linear space in the following manner.

Definition 2.13. *An* **inner product** *or* **scalar product** *on a (real) linear space V is a real-valued function from V to $\mathbb{R}$, denoted by $(\cdot, \cdot)$, satisfying*

(i) $(u, u) \geq 0$ *for all* $u \in V$ *and* $(u, u) = 0$ *if and only if* $u = 0$*;*

(ii) $(u, v) = (v, u)$ *for all* $u, v \in V$*;*

(iii) $(\alpha u + \beta v, w) = \alpha\,(u, w) + \beta\,(v, w)$ *for all* $u, v, w \in V$ *and all* $\alpha, \beta \in \mathbb{R}$*.*

A vector space V equipped with an inner product is aptly called an *inner product space.*

Example 2.14 Returning to our example $C^0(\Omega)$ of a linear space we can define an inner product as

$$(f, g) = \int_0^1 f(x) g(x)\, dx \,.$$

The three properties of the inner product are easily shown to be satisfied by using the properties of integrals. See exercises. ∎

Analogous to the case of $\mathbb{R}^n$, we say that two vectors in an inner product space are orthogonal if their scalar product is zero.

Definition 2.15. *Let V be an inner product space. Then $u, v \in V$ are orthogonal if and only if*

$$(u, v) = 0 \,. \tag{2.4}$$

One can use the inner product to define a norm for a vector space. Indeed, if we let $\|v\| = (v, v)^{1/2}$ for all $v \in V$, one can readily show that this defines a norm on V; see the exercises for details. We refer to a norm defined in this manner on an inner product space as the *induced norm.*

To complete this section we present an inequality for inner product spaces which is extremely useful. Recall that the scalar product of two vectors in $\mathbb{R}^n$ can also be written as $(\vec{a}, \vec{b}) = \|\vec{a}\|\,\|\vec{b}\| \cos\theta$ where $\|\cdot\|$ denotes the standard Euclidean norm and θ is the angle between the two vectors. The Cauchy-Schwarz inequality generalizes this result to an inner product space.

Lemma 2.16. *Let V be an inner product space. The Cauchy-Schwarz inequality is given by*

$$(u, v) \leq (u, u)^{\frac{1}{2}} (v, v)^{\frac{1}{2}} \quad \forall\, u, v, \in V\,. \tag{2.5}$$

If $\|\cdot\|$ denotes the induced norm on V then this inequality can also be written as

$$(u, v) \leq \|u\|^{\frac{1}{2}} \|v\|^{\frac{1}{2}} \quad \forall\, u, v, \in V\,. \tag{2.6}$$

Proof. To verify (2.5), we first note that it is trivially satisfied if $u = 0$ or $v = 0$ so we consider the case where $u, v \neq 0$. By the first property of inner products, we know that $(u - \alpha v, u - \alpha v) \geq 0$ for any $\alpha \in \mathbb{R}$. Using the linearity property of the inner product we rewrite this as

$$\begin{aligned} 0 \leq (u - \alpha v, u - \alpha v) &= (u, u) - 2\alpha(u, v) + \alpha^2(v, v) \\ &= (u, u) - \alpha(u, v) - \alpha\Big[(u, v) - \alpha(v, v)\Big]\,. \end{aligned}$$

Now the term in brackets is zero if we choose $\alpha = (u, v)/(v, v)$. Note that this is possible since we are considering the case $v \neq 0$. Thus

$$(u, u) - \frac{(u, v)^2}{(v, v)} \geq 0$$

and simplification yields the Cauchy-Schwarz inequality (2.5). The second form of the inequality given in (2.6) follows directly from the definition of the norm on V induced by the scalar product.

2.1.3 Topological concepts

One of our goals in analyzing the finite element method is to determine the error between the solution of the discrete weak problem and the solution of the continuous weak problem. We can use the concept of norm introduced in the last section to measure the distance between these two solutions. For a normed vector space V, we define the *distance* ρ between two vectors u and v as $\rho(u, v) = \|u - v\|$.

In discretizing a problem, we expect to have a sequence of solutions which are generated by using successively finer meshes. We expect that these solutions converge, in some sense, to the solution of the continuous problem. We now make precise what this means.

Definition 2.17. *A sequence of vectors $u_1, u_2, u_3, \ldots$ belonging to a normed linear space V is called* **convergent** *if there exists a vector $u \in V$ such that given any $\epsilon > 0$, there exists a postive integer $N = N(\epsilon)$ such that*

$$\|u_n - u\| < \epsilon \quad \forall\, n \geq N\,.$$

We call u the limit of the sequence $\{u_i\}_{i\geq 1}$ and write

$$\lim_{n\to\infty} u_n = u \quad \textit{or} \quad u_n \to u \textit{ in } V \textit{ as } n \to \infty\,.$$

It can be shown that a convergent sequence has only one limit and $u_n \to u$ in V if and only if $\|u_n - u\| \to 0$ as $n \to \infty$.

An important tool in analysis is the Cauchy sequence. If we use Definition 2.17 to show that a sequence is convergent, then we need to know its limit. However, sometimes we don't know the actual limit or it may not even be in our linear space. Oftentimes the important issue is that a sequence converges rather than what its limit is. A Cauchy sequence is one in which its terms ultimately become arbitrarily close. In fact, we can discount a finite number of terms at the beginning of the sequence and then guarantee that any two of the remaining terms are closer than some prescribed value.

Definition 2.18. *A sequence of vectors $\{u_i\}_{i\geq 1}$, $\{u_i\} \in V$, is called a* **Cauchy sequence** *if, given any $\epsilon > 0$, there exists an integer $N = N(\epsilon)$ such that*

$$\|u_n - u_m\| \leq \epsilon \quad \forall\, m, n \geq N\,.$$

Here $\|\cdot\|$ defines a norm on a normed linear space V.

Every convergent sequence is clearly a Cauchy sequence since

$$\|u_m - u_n\| = \|(u_m - u) + (u - u_n)\|\, le\, \|u - u_m\| + \|u - u_n\|$$

and we can make the right-hand side arbitrarily small as $m, n \to \infty$. However, the converse is not always true as the following example illustrates.

Example 2.19 The Weierstrass Approximation Theorem states that a continuous function defined on $[a, b]$ can be uniformly approximated as closely as desired by a polynomial defined on $[a, b]$. More precisely, suppose f is a continuous function defined on $[a, b]$. For every $\epsilon > 0$, there exists a polynomial function $p(x)$ such that $\max_{x\in[a,b]} |f(x)?p(x)| < \epsilon$. Thus we can construct a sequence of polynomials in the linear space of polynomials defined on $[a, b]$ with the max-norm which form a Cauchy sequence but its limit is not a polynomial. ■

We would like to avoid this situation by imposing on our space of functions the property that every Cauchy sequence in V has a limit in V. In addition to properties (i)–(x) which characterize linear spaces, we would like to add the property of *completeness*, *i.e.,*

(xi) if $\{v_n\}$ is a sequence of elements in V such that $\|v_n - v_m\| \to 0$ as $m, n \to \infty$, then there exists an element $v \in V$ such that $\|v - v_n\| \to 0$ as $n \to \infty$.

Another way to state property (xi) is to require that every Cauchy sequence in V is convergent.

A *complete normed vector space*, *i.e.,* a collection of objects satisfying properties (i)–(xi) with a norm defined on the space, is of such importance that it is given a special name: a *Banach space*. Euclidean n-dimensional space is the most familiar example of a Banach space. Given any (noncomplete) normed space S it can be proved that by adding new elements, S can be extended to a complete normed space (a Banach space), V. This process is referred to as the *completion* of S or the *closure* of S in V.

Since we work with finite dimensional subspaces when we discretize, we often have sequences on these subspaces and need to know if their limit is in the subspace.

Definition 2.20. *A subset S of a Banach space V is said to be a* **closed subspace** *of V if it is a subspace of V with the property that whenever $\{u_i\}_{i\geq 1}$ is a convergent sequence in V such that $u_i \in S$, $i = 1, 2, \ldots$, then $u = \lim_{n\to\infty} u_n$ belongs to S also.*

It can be shown that every finite dimensional subspace is closed; this is important for us since our approximating spaces are finite dimensional.

Our search for the appropriate function spaces to use in analyzing the finite element method is almost at an end. In the next section we add a final property to our complete, normed linear space, that of an inner product.

2.1.4 Hilbert spaces

A complete inner product space is called a *Hilbert space*; these spaces extend the ideas of the Euclidean space $\mathbb{R}^n$ to infinite dimensional spaces. For example, the parallelogram law in $\mathbb{R}^2$ states that the the sum of the squares of the lengths of the two diagonals in a parallelogram equals the sum of the squares of the lengths of the four sides. This law can be shown to hold in all Hilbert spaces and is written as

$$\|f+g\|^2 + \|f-g\|^2 = 2\big(\,\|f\|^2 + \|g\|^2\,\big)\,, \tag{2.7}$$

where $\|\cdot\|$ denotes the induced norm and f, g are any elements of the Hilbert space. See the exercises for a proof of this result.

Clearly, every Hilbert space is a Banach space; one simply uses the norm induced by the inner product, *i.e.,* $\|v\| = (v,v)^{1/2}$. However, the converse is not true. A standard counterexample is to consider the Banach space of all bounded linear functions with the uniform or max norm. In this example, one can demonstrate that the parallelogram law fails to hold so it can not also be a Hilbert space; see the exercises for details. The most commonly used spaces of admissible test and trial functions for weak formulations of boundary value problems for partial differential equations are Hilbert spaces.

Example 2.21 An example of a Hilbert space that is central to our discussions is $L^2(\Omega)$ where Ω denotes an open, connected subset of $\mathbb{R}^n$. To construct this space, we consider the set of real-valued, continuous functions $u(\mathbf{x}) = u(x_1, x_2, \ldots, x_n)$ defined on Ω where $(x_1, x_2, \ldots, x_n)$ denotes a point in $\mathbb{R}^n$. Addition and scalar multiplication are defined in the usual manner. We define an inner product as

$$(u, v) = \int_\Omega u(\mathbf{x})v(\mathbf{x})\, d\Omega\,, \tag{2.8}$$

where $d\Omega$ is the volume element in Ω. Clearly this satisifes all the properties of an inner product given in Definition 2.13. In order to guarantee that the integral defining this inner product exists, we restrict our attention to functions $u(\mathbf{x})$ on Ω such that

$$\int_\Omega |u(\mathbf{x})|^2\, d\Omega < \infty\,.$$

We now define S to be the space described above; *i.e.,*

$$S = \{u \mid u = u(\mathbf{x}), u(\mathbf{x}) \text{ is continuous for all } \mathbf{x} \in \Omega \text{ and } \int_\Omega |u(\mathbf{x})|^2\, d\Omega < \infty\}\,.$$

Then S is an inner product space with the inner product defined by (2.8). The norm on S is given by

$$\|u\| = (u, u)^{1/2} = \left(\int_\Omega |u(\mathbf{x})|^2\, d\Omega\right)^{1/2}\,.$$

In general, the space S is not complete. For example in $\mathbb{R}$, let $\Omega = (-1, 1)$ and let S be defined as above. Consider the sequence $u_1(x), u_2(x), \cdots$ where

$$u_j(x) = \begin{cases} -1 & \text{for} \quad -1 < x \le -1/j \\ jx & \text{for} \quad -1/j \le x \le 1/j \\ 1 & \text{for} \quad 1/j < x < 1\,. \end{cases}$$

It is straightforward to show that $\{u_j(x)\}$, $j = 1, 2, \cdots$, is a Cauchy sequence in S. Moreover, the sequence converges to the discontinuous function $f(x)$ where

$$f(x) = \begin{cases} -1 & \text{for} \quad -1 < x < 0 \\ 0 & \text{for} \quad x = 0 \\ 1 & \text{for} \quad 0 < x < 1\,. \end{cases}$$

However, $f(x) \notin S$ and so there is no continuous function $u(x)$ on $(-1, 1)$ for which $\|u_n - u\| \to 0$ as $n \to \infty$. By adding new elements to S we can complete the space to form a Hilbert space V. These additional functions may be piecewise continuous, but, in general, are highly discontinuous. This extended *complete* space V is called $L^2(\Omega)$ which is a complete inner product space, *i.e.,* a Hilbert space. ■

Remark The space $L^2(\Omega)$ is really a special case of the Banach space of functions on Ω which are p-integrable denoted $L^p(\Omega)$, $p \ge 1$, The norm is given by

$$\|u\|_{L^p} = \left(\int_\Omega |u|^p\, d\Omega\right)^{1/p}$$

for $1 \leq p < \infty$ and for $p = \infty$

$$\|u\|_{L^\infty} = \sup_\Omega |u| \,.$$

Remark In a manner analogous to the construction of $L^2(\Omega)$, we can construct a weighted L^2-space. Given a weight $w(\mathbf{x})$, integrable on Ω, we define the inner product as

$$(u, v) = \int_\Omega u(\mathbf{x})v(\mathbf{x})w(\mathbf{x})\, d\Omega \,.$$

We denote this space $L^2(\Omega; w)$.

2.2 Best approximations

In this section we want to investigate some geometric properties of Hilbert spaces. A central idea in approximation theory is to determine an element of a subspace of a given vector space which is closest (with respect to the given metric) to a particular element of the vector space; that is, to find the *best approximation* of the given vector in the subspace. (In fact, this is the basis for least squares methods.) We would like to know when it is possible to assert in advance that a best approximating element exists. Moreover, we want to know whether this best approximating element is unique.

Example 2.22 Consider the situation illustrated in Figure 2.1. Here we assume that we have a given plane S in $\mathbb{R}^3$, a vector $\mathbf{u} \notin S$, and we want to find a vector $\mathbf{s}$ in S that is nearest $\mathbf{u}$; *i.e.,* $\|\mathbf{u} - \mathbf{s}\| \leq \|\mathbf{u} - \phi\|$ for all $\phi \in S$ where we are using the standard Euclidean norm. Clearly in this case, there is a a unique $\mathbf{s}$ and it is found by drawing a perpendicular from $\mathbf{u}$ to S; that is, *projecting* the vector $\mathbf{u}$ onto S. Also, we can uniquely write the vector $\mathbf{u}$ as the sum of the vector $\mathbf{s} \in S$ and a vector not in S. ∎

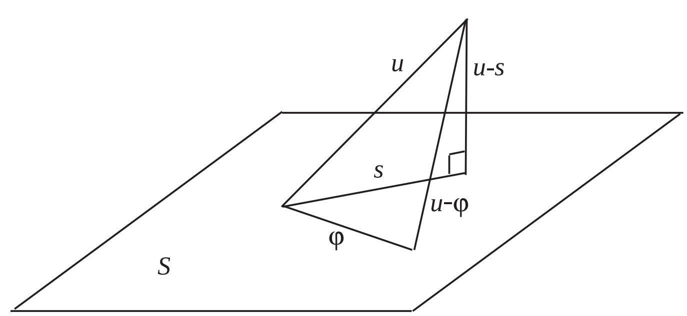

Figure 2.1.

There are analogous results for general Hilbert spaces. However, they do not hold for a general subspace but only a *closed subspace* (see Definition 2.20). In the finite element setting, we are guaranteed that the subspace is closed since it is finite dimensional.

The following result is known as the *Projection Theorem* and it states that given an element in a Hilbert space, its orthogonal projection onto a closed subspace is the element of the subspace that is "nearest" the given vector where the distance is measured using the induced norm. Of course one must keep in mind that this depends upon the choice of the inner product (and thus the norm) on the given Hilbert space.

Theorem 2.23. (The Projection Theorem) *Let S be a closed subspace of a Hilbert space V which is not the whole of V. Then given $u \in V$ there exists a unique element $Pu \in S$ such that*

$$\|u - Pu\| = \inf_{\phi \in S} \|u - \phi\| \tag{2.9}$$

where Pu satisfies

$$(u - Pu, \phi) = 0 \quad \text{for every } \phi \in S\,. \tag{2.10}$$

Proof. Let $u \in V$ such that $u \notin S$. We know that for all $\phi \in S$, $\|u - \phi\| > 0$ so that if we define the distance from u to S as the lower bound

$$\delta = \inf_{\phi \in S} \|u - \phi\| \,,$$

then there exists a sequence $\phi_n \in S$ such that $\|u - \phi_n\| \to \delta$ as $n \to \infty$. Our goal is so show that $\{\phi_n\}$ is a Cauchy sequence in S and thus conclude that the limit of the sequence is also in S because S is a closed subspace of the Hilbert space V; we call Pu this limit.

To demonstrate that $\{\phi_n\}$ is a Cauchy sequence we apply the parallelogram law (2.7)

with $f = u - \phi_m$ and $g = u - \phi_n$. We have

$$\|(u - \phi_m) + (u - \phi_n)\|^2 + \|(u - \phi_m) - (u - \phi_n)\|^2 = 2\,\|u - \phi_m\|^2 + 2\,\|u - \phi_n\|^2$$

and simplifying the left-hand side gives

$$4\left\|\left(u - \frac{\phi_m + \phi_n}{2}\right)\right\|^2 + \|\phi_n - \phi_m\|^2 = 2\,\|u - \phi_m\|^2 + 2\,\|u - \phi_n\|^2 \tag{2.11}$$

because

$$\|(u - \phi_m) + (u - \phi_n)\|^2 = \|2u - \phi_m - \phi_n\|^2 = 4\left\|u - \frac{1}{2}\,(\phi_m + \phi_n)\right\|^2 \,.$$

Now $\phi_n, \phi_m \in S$ and S is a subspace of V, so we have that $(\phi_m + \phi_n)/2 \in S$ and thus by the definition of δ, the first term on the left of (2.11) is nonnegative and at least as large as $4\delta^2$. Thus

$$\|\phi_n - \phi_m\|^2 \le 2\,\|u - \phi_m\|^2 + 2\,\|u - \phi_n\|^2 - 4\delta^2 \,.$$

We conclude that because $\|\phi_m - u\| \to \delta$, $\|\phi_n - u\| \to \delta$ as $m, n \to \infty$, the right-hand side goes to zero as $m, n \to \infty$, and hence $\{\phi_n\}$ is a Cauchy sequence in a closed subspace and therefore convergent.

Uniqueness of the limit $s \in S$ is proved in the standard way by assuming there are two elements in S which satisfy (2.9). Let s_1 and s_2 have the property that

$$\delta = \inf_{\phi \in S} \|u - \phi\| = \|u - s_1\| = \|u - s_2\| \ .$$

Then because $(s_1 + s_2)/2 \in S$, we have that

$$\delta \leq \left\| u - \frac{1}{2}(s_1 + s_2) \right\| \leq \frac{1}{2} \|u - s_1\| + \frac{1}{2} \|u - s_2\| = \delta$$

where we have used the triangle inequality for the last inequality. This implies that for $u - s_1$ and $u - s_2$ the triangle inequality must hold as an equality. However, this can only be true if $u - s_1 = \alpha(u - s_2)$ for some α. If we choose $\alpha = 1$ then this leads to a contradiction because it would imply $s_1 = s_2$; if $\alpha \neq 1$ then we have $(\alpha - 1)u = s_2 - s_1$ and this contradicts the fact that $u \notin S$. Consequently we have uniqueness.

To prove (2.10) we assume that there is some $\hat{\phi} \in S$ such that $(u - Pu, \hat{\phi}) \neq 0$ and show that this assumption leads to the existence of an $s \in S$ such that $\|u - s\| < \inf_{\phi \in S} \|u - \phi\|$; thus we obtain a contradiction. Let $s \in S$ be given by

$$s = Pu + \frac{\left(u - Pu, \hat{\phi}\right)}{\left(\hat{\phi}, \hat{\phi}\right)} \hat{\phi} \, .$$

Then

$$\begin{aligned}
\|u - s\|^2 &= \left(u - Pu - \frac{\left(u - Pu, \hat{\phi}\right)}{\|\hat{\phi}\|^2} \hat{\phi}, u - Pu - \frac{\left(u - Pu, \hat{\phi}\right)}{\|\hat{\phi}\|^2} \hat{\phi} \right) \\
&= \|u - Pu\|^2 - \frac{2}{\|\hat{\phi}\|^2} \left(u - Pu, \hat{\phi}\right) \left(u - Pu, \hat{\phi}\right) + \frac{1}{\|\hat{\phi}\|^4} \left(u - Pu, \hat{\phi}\right)^2 \|\hat{\phi}\|^2 \\
&= \|u - Pu\|^2 - \frac{1}{\|\hat{\phi}\|^2} \left(u - Pu, \hat{\phi}\right)^2 \, .
\end{aligned}$$

Because our assumption was that $(u - Pu, \hat{\phi}) \neq 0$, we have

$$\|u - s\| < \|u - Pu\| = \inf_{\phi \in S} \|u - \phi\|$$

which is the contradiction we sought. ∎

In Example 2.22, we wrote the vector which we projected into the subspace as the sum of a vector in the subspace and one orthogonal to the subspace. Theorem 2.23 guarantees that we can do this in a Hilbert space when we are projecting onto a closed subspace. Given u in a Hilbert space V then the unique Pu in a closed subspace S guaranteed by Theorem 2.23 is called the *orthogonal projection* of u onto the closed subspace S. In this case, if we let $r = u - Pu$, then we can write

$$u = Pu + r \quad \text{where } Pu \in S \text{ and } (r, \phi) = 0 \text{ for all } \phi \in S\,. \tag{2.12}$$

Hence the vector $r = u - Pu$ is *orthogonal* to all vectors in S; we call the set of all vectors orthogonal to S the *orthogonal complement* of S and denote it by $S^\perp$. Thus we have written u as the sum of an element in S and one orthogonal to S, i.e., in $S^\perp$. In the exercises, we explore the fact that the analogous result holds for the entire Hilbert; that is, the Hilbert space V can be written as the direct sum of S and $S^\perp$.

Another interpretation of the vector s guaranteed by Theorem 2.23 is given by (2.9). From this equation we call Pu the *best approximation* of u in S. In the case when S is a finite dimensional subspace of a Hilbert space V (and thus automatically closed) then we can explicitly construct the best approximation to a given vector in V by using (2.10). To see this, we let ϕ_i, $i = 1, \ldots, m$ be a basis for S and from (2.10), $(u - Pu, \phi_i) = 0$ for $i = 1, \ldots, m$. Also Pu can be written as a linear combination of the basis elements; *i.e.,* $Pu = \sum_{j=1}^m c_j \phi_j$. Thus

$$\sum_{j=1}^{m} c_j\,(\phi_j, \phi_i) = (u, \phi_i)\,, \quad i = 1, \ldots, m\,, \tag{2.13}$$

which is just a linear system for the unknowns c_j, $j = 1, \ldots, m$. The matrix $\mathcal{G}$ whose entries are given by $\mathcal{G}_{ij} = (\phi_j, \phi_i)$ is known as the *Gram matrix* associated with the basis functions $\{\phi_i\}$ of S and is guaranteed to be nonsingular. This can be easily seen by assuming that if $\mathcal{G}$ is singular, then we could find a vector $d = (d_1, d_2, \ldots, d_m)$ such that $\mathcal{G}d = 0$. This would imply $\left(\sum_{j=1}^m d_j \phi_j, \phi_i\right) = 0$ for all $i = 1, \ldots, m$ and thus the vector $u = \sum_{j=1}^m d_j \phi_j$ would be orthogonal to S which is a contraction.

In the following example we investigate the effect that the choice of the approximating subspace for the best approximation has on the properties of the Gram matrix.

Example 2.24 Consider the problem of determining the best approximation to a function $f(x)$ in two different subspaces of $L^2(0,1)$. We first consider the subspace consisting of all polynomials of degree three or less. In this case an obvious choice of a basis is $\{1, x, x^2, x^3\}$. The specific system we must solve is

$$\begin{pmatrix} \int_0^1 dx & \int_0^1 x\,dx & \int_0^1 x^2\,dx & \int_0^1 x^3\,dx \\ \int_0^1 x\,dx & \int_0^1 x^2\,dx & \int_0^1 x^3\,dx & \int_0^1 x^4\,dx \\ \int_0^1 x^2\,dx & \int_0^1 x^3\,dx & \int_0^1 x^4\,dx & \int_0^1 x^5\,dx \\ \int_0^1 x^3\,dx & \int_0^1 x^4\,dx & \int_0^1 x^5\,dx & \int_0^1 x^6\,dx \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} \int_0^1 f(x)\,dx \\ \int_0^1 x f(x)\,dx \\ \int_0^1 x^2 f(x)\,dx \\ \int_0^1 x^3 f(x)\,dx \end{pmatrix}.$$

Upon performing the integration the system becomes

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} \int_0^1 f(x)\,dx \\ \int_0^1 x f(x)\,dx \\ \int_0^1 x^2 f(x)\,dx \\ \int_0^1 x^3 f(x)\,dx \end{pmatrix}.$$

It is important to note that in this case the Gram matrix is not sparse and is the well-known Hilbert matrix which is notoriously ill-conditioned. Consequently, if we look for the best approximation to a function $f(x)$ out of the space of polynomials of degree n or less, then for even modest values of n our solution is not be reliable using standard matrix solvers.

On the other hand, if we choose as a subspace of $L^2(0,1)$ the space of continuous piecewise linear functions over the uniform partition of $[0,1]$ into N subintervals and choose as a basis the piecewise linear "hat" functions described in the previous chapter, then the resulting matrix is well-conditioned and is tridiagonal. To see the structure of the matrix consider the uniform partition of $[0,1]$ into 3 subintervals with four grid points $x_1 = 0$, $x_2 = \frac{1}{3}$, $x_3 = \frac{2}{3}$, $x_4 = 1$. Let $\phi_i(x)$, $i = 1, \ldots, 4$ denote the basis functions

$$\phi_1 = \begin{pmatrix} 1-3x & 0 \le x \le \frac{1}{3} \\ 0 & \text{elsewhere} \end{pmatrix}, \qquad \phi_2 = \begin{pmatrix} 3x & 0 \le x \le \frac{1}{3} \\ 2-3x & \frac{1}{3} \le x \le \frac{2}{3} \\ 0 & \text{elsewhere} \end{pmatrix},$$

$$\phi_3 = \begin{pmatrix} 3x-1 & \frac{1}{3} \le x \le \frac{2}{3} \\ 3-3x & \frac{2}{3} \le x \le 1 \\ 0 & \text{elsewhere} \end{pmatrix}, \qquad \phi_4 = \begin{pmatrix} 3x-2 & 2/3 \le x \le 1 \\ 0 & \text{elsewhere} \end{pmatrix}.$$

The Gram matrix has the (i,j) entry given by $\int_0^1 \phi_i(x)\phi_j(x)\,dx$. Due to the fact that the basis function $\phi_i(x)$ has local support, we deduce that the Gram matrix is

$$\begin{pmatrix} \int_0^1 \phi_1^2\,dx & \int_0^1 \phi_2\phi_1\,dx & 0 & 0 \\ \int_0^1 \phi_1\phi_2\,dx & \int_0^1 \phi_2\phi_2\,dx & \int_0^1 \phi_3\phi_2\,dx & 0 \\ 0 & \int_0^1 \phi_2\phi_3\,dx & \int_0^1 \phi_3\phi_3\,dx & \int_0^1 \phi_4\phi_3\,dx \\ 0 & 0 & \int_0^1 \phi_3\phi_4\,dx & \int_0^1 \phi_4\phi_4\,dx \end{pmatrix},$$

which is a symmetric tridiagonal matrix. ∎

Being able to find the best approximation to a given function using piecewise polynomials does not directly help us to find our finite element approximation. This is because in order to use (2.13) to find the best approximation to u, we need to know u, which in our case is the unknown solution to the weak problem. However, what we see in the next chapter is that when we measure the error in our finite element approximation, it will be bounded by a constant times the error in the best approximation in the approximating space.

2.3 Bounded linear functionals

Functional is just the name given to a special type of function which assigns a number to each element of a linear space. For example, for functions in $L^2((0,1))$ the integral over the domain is a functional. If V is a given Hilbert space with inner product $(\cdot,\cdot)$ and induced norm $\|\cdot\|$ then $\|v\|$ assigns a number to each element v in V and is thus a functional. If we fix an element u in V, then (v,u) assigns a value (*i.e.,* a scalar) to each element. Such mappings are called *functionals*. Functionals which are linear and bounded are of particular interest.

Definition 2.25. *F is a functional on a Hilbert space V if it assigns to every $v \in V$ a unique number $F(v)$ and we write $F : V \to \mathbb{R}$. A functional is called linear if for every $u, v \in V$ and scalars α, β we have*

$$F(\alpha u + \beta v) = \alpha F(u) + \beta F(v)\,. \tag{2.14}$$

In addition, we say that a functional is bounded if

$$\sup_{v \in V} \frac{|F(v)|}{\|v\|} < \infty,\ v \neq 0\,, \tag{2.15}$$

where $\|\cdot\|$ is the induced norm on V. We call this finite number $\|F\|$.

We note that if F is a bounded linear functional on V then this is equivalent to saying that F is a linear functional which is a continuous function of its arguments.

Example 2.26 Let V be a Hilbert space; for a fixed $u \in V$ the inner product $F(v) = (v,u)$ denotes a bounded linear functional on V. Clearly it defines a functional and is linear because of the linearity of the inner product. Specifically, we have

$$F(\alpha v+\beta w) = (\alpha v+\beta w, u) = (\alpha v, u)+(\beta w, u) = \alpha(v,u)+\beta(w,u) = \alpha F(v)+\beta F(w)\,.$$

Boundedness follows from using the Cauchy-Schwarz inequality

$$F(v) = (v,u) \leq \|v\|\,\|u\|\ .$$

to obtain

$$\frac{|F(v)|}{\|v\|} \leq \|u\| < \infty \quad \forall\, v \in V\,,$$

for $v \neq 0$. ∎

In a similar manner, the linearity of the norm on a Hilbert space can be used to demonstrate that the norm is a bounded linear functional. Clearly, we can think of many other examples of bounded linear functionals, but what is surprising is that the inner product is really the only one on a Hilbert space; i.e., every bounded linear functional can be written as an inner product. This result is known as the Riesz Representation Theorem and is named after the Hungarian mathematician Frigyes Riesz (1880-1956).

Theorem 2.27. (Riesz Representation Theorem) *For every bounded linear functional F on a Hilbert space V there is a unique element $f \in V$ such that*

$$F(v) = (v, f) \quad \text{for all } v \in V\,. \tag{2.16}$$

Moreover, $\|F\| = \|f\|$.

Proof. We first note that if F assigns to each $v \in V$ the value zero, then the proof is immediate by taking $f = 0$. In the sequel we assume that this is not the case. However, we do know that for each $v \in V$ which F assigns to zero we must have that f is orthogonal to it; *i.e.,* $(v, f) = 0$. We call the set of all vectors v such that $F(v) = 0$ the *kernel* of F and denote it by $\mathcal{K}(F)$. Hence we must construct an f that is orthogonal to the kernel of F.

We first demonstrate that $\mathcal{K}(F)$ is a closed subspace of V. To show that it is a subspace we use the linearity of F; *i.e.,* if $u, v \in \mathcal{K}(F)$ then

$$F(\alpha u + \beta v) = \alpha F(u) + \beta F(v) = 0\,.$$

To show that it is a closed subspace of V, we let $\{u_n\}$ be a sequence in $\mathcal{K}(F)$ such that $u_n \to u \in V$ as $n \to \infty$ and show that $u \in \mathcal{K}(F)$. We have

$$|F(u)| = |F(u) - F(u_n)| = |F(u - u_n)| \le \|F\|\,\|u - u_n\|\ ,$$

where we have used the fact that $u_n \in \mathcal{K}(F)$, the linearity of F, and the definition of the norm of a bounded linear functional. The right-hand side of this inequality goes to zero as $n \to \infty$ so that $F(u) = 0$.

We now proceed to construct an f that is orthogonal to $\mathcal{K}(F)$. From the comments following the projection theorem we know that we can write V as the direct sum of $\mathcal{K}(F)$ and $\mathcal{K}(F)^\perp$ since $\mathcal{K}(F)$ is a closed subspace of V. Our strategy is to take an arbitrary $\hat{f} \in \mathcal{K}(F)^\perp$, $\hat{f} \neq 0$ and construct an $f \in \mathcal{K}(F)$ using $\hat{f}$. Consider the vector $F(v)\hat{f} - F(\hat{f})v$. This vector is in $\mathcal{K}(F)$ since

$$F\Big(F(v)\hat{f} - F(\hat{f})v\Big) = F(v)F(\hat{f}) - F(\hat{f})F(v) = 0$$

and thus $\Big(F(v)\hat{f} - F(\hat{f})v, \hat{f}\Big) = 0$ for all $\hat{f} \in \mathcal{K}(F)^\perp$, $v \in V$. Therefore we have $F(v)\|\hat{f}\|^2 = F(\hat{f})(v, \hat{f})$ so that

$$F(v) = \left(v, \frac{F(\hat{f})}{\|\hat{f}\|^2}\hat{f}\right)\,.$$

Hence if we set $f = (F(\hat{f})/\|\hat{f}\|^2)\hat{f}$, we see that for each $v \in \mathcal{K}(F)$, this choice of f is orthogonal to v and we have the desired result.

To show uniqueness of f we assume that there are two vectors f_1 and f_2 such that

$$F(v) = (v, f_1) = (v, f_2) \qquad \forall\, v \in V\,.$$

But this implies that $(v, f_1 - f_2) = 0$ for all $v \in V$; specifically set $v = f_1 - f_2$ from which it follows that $f_1 - f_2 = 0$.

Lastly, we must demonstrate that $\|F\| = \|f\|$. This follows immediately from the definition of F and the Cauchy-Schwarz inequality. We have

$$|F(v)| = |\,(v, f)\,| \le \|v\|\,\|f\|$$

so that if $v \neq 0$,

$$\frac{|F(v)|}{\|v\|} \le \|f\|$$

and thus $\|F\| \le \|f\|$. On the other hand, since $f \in V$, $F(f) = (f, f) = \|f\|^2$. Thus the supremum is attained at $v = f$ and we have equality. ∎

The main goal of this chapter was to introduce the mathematical tools necessary to formulate and analyze a general weak problem. The last tool we need is a bilinear form. We define a *bilinear form* on a Hilbert space V to be a map from $V \times V$ into $\mathbb{R}^1$, denoted by $B(\cdot,\cdot)$, such that

$$B(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 B(u_1, v) + \alpha_2 B(u_2, v)$$
$$B(u, \beta_1 v_1 + \beta_2 v_2) = \beta_1 B(u, v_1) + \beta_2 B(u, v_2)$$

for all $u_i, v_i \in V$ and $\alpha_i, \beta_i \in \mathbb{R}^1$, $i = 1, 2$. That is, $B(\cdot,\cdot)$ is linear in each of its components. An example of a bilinear form on $L^2(0,1)$ is $\int_0^1 u(x)v(x)\,dx$. In fact, any inner product on a Hilbert space defines a bilinear form; this can easily be seen from the linearity of the inner product (see Definition 2.13). Another example of a bilinear form on a Hilbert space V is $(\mathcal{B}u, v)_V$ where $\mathcal{B}$ is a linear operator from V to V.

We say that a bilinear form $B(\cdot,\cdot)$ on V is bounded if there exists a positive constant C such that

$$|B(u, v)| \le C\,\|u\|_V\,\|v\|_V\ .$$

If we fix an element $u \in V$ then the bilinear form $B(u, v)$ represents a linear functional on V; if $B(\cdot,\cdot)$ is bounded, then for a fixed $u \in V$, $B(u, v)$ represents a bounded linear functional $F(v)$ on V. The Riesz Representation Theorem 2.27 then guarantees that there exists a unique element $\hat{u} \in V$ such that $B(u, v)$ can be written as the inner product $(v, \hat{u})$. The ability to associate to each $u \in V$ a unique element $\hat{u}$ is central to our analysis of an abstract weak problem.

Exercises

2.1. The most common examples of norms on $\mathbb{R}^n$ are the Euclidean norm defined by

$$\|\mathbf{x}\|_{\ell_2} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} ,$$

the sum norm defined by

$$\|\mathbf{x}\|_{\ell_1} = \left(\sum_{i=1}^{n} |x_i| \right) ,$$

and the maximum norm defined by

$$\|\mathbf{x}\|_{\ell_\infty} = \max_{1 \le i \le n} |x_i| .$$

Here $\mathbf{x} = (x_1, x_2, \cdots, x_n)$. For any norm, the set $\{\mathbf{x} : \|\mathbf{x}\| \le 1\}$ is called the unit ball.

a. Sketch the unit balls for each of the norms defined above.

b. Show that the norms are equivalent by explicitly determining the comparability constants.

2.2. Let V be a complete inner product space; define $\|u\|$ to be the non-negative number $(u, u)^{1/2}$. Show that this defines a norm on V.

2.3. Let V be a Hilbert space and let $f, g \in V$. Verify the parallelogram law

$$\|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2 .$$

Note that the name comes from the special case of $\mathbb{R}^2$ where we know that the sum of the squares of the sides of a parallelogram is equal to the sum of the squares of the diagonals.

2.4. In the previous exercise we saw that any two elements of a Hilbert space V satisfies the parallelogram law; in fact, one can show that if B is a Banach space which satisfies the parallelogram law then it is also a Hilbert space (see, e.g., [Schechter]). Consider the Banach space of all bounded real functions on the interval $[0, 1]$ with the norm

$$\|u\| = \sup_{0 \le x \le 1} |u(x)| .$$

Find functions $f, g \in B$ which violate the parallelogram law and thus conclude that B is a Banach space but not a Hilbert space. (Hint: for example, find functions f, g such that $\|f\| = \|g\| = \|f - g\| = \|f + g\|$.)

2.5. Let S be a closed subspace of a Hilbert space V. Let $S^\perp$ be defined by

$$S^\perp = \{u \in V : (u, \phi) = 0 \text{ for all } \phi \in S\} .$$

Show that $S^\perp$ is a closed subspace of V. Also show that $S \cap S^\perp = \{0\}$ and thus V can be written as the direct sum of S and $S^\perp$ so that every element $u \in V$ can be written as the sum of an element in S and in $S^\perp$.

2.6. Let P be the projection operator from a Hilbert space V to a *closed* subspace $S \subset V$; *i.e.,* P is an operator $P : V \to S$ such that

$$Pu = \begin{cases} u & \text{if } u \in S \\ u_0 & \text{otherwise,} \end{cases}$$

where $u = u_0 + u_1$ uniquely with $u_0 \in S$ and $u_1 \in S^\perp$.

a. Show that P is linear.

b. Clearly, the range of P is S. What is the kernel of P? Why?

c. Show that $P^2 = P$.

d. Show that $\|P\| = 1$ where

$$\|P\| = \sup_{\phi \in V} \frac{\|P\phi\|}{\|\phi\|} \qquad \text{for } \phi \neq 0\,.$$

e. Show that $I - P$ is the projection operator onto the orthogonal complement of S.

2.7. Prove that if $\{u_i\}$ is a convergence sequence in a normed linear space then the limit is unique.

2.8. (Computational) Consider the function $u(x) = x^3 \sin \pi x$ on $[0, 1]$. We want to determine the *best approximation* in the L^2-norm, $\tilde{u}(x)$, to $u(x)$ out of the space of continuous piecewise linear functions which are zero at $x = 0$ and $x = 1$.

a. Choose a uniform partition of $[0, 1]$ with $h = 0.25$. Write a code to determine the best approximation $\tilde{u}$ to $u(x)$ using the standard "hat" basis functions for continuous piecewise linears. For the integration, use a two-point Gauss quadrature rule. Write your code so that you have a separate function or subroutine which evaluates a basis function at any given point.

b. Repeat (a) with $h = 0.125$ and $h = .0625$. For each value of h determine the $L^2(0, 1)$ error in $u(x)$ and $\tilde{u}(x)$; calculate a numerical rate of convergence (*i.e.,* determine k such that the error is $\mathcal{O}(h^k)$) based upon your two calculations. To calculate the error, apply the two-point Gauss quadrature rule over each subinterval.