2. Signal Space Concepts

R.G. Gallager

The signal-space viewpoint is one of the foundations of modern digital communications. Credit for popularizing this viewpoint is often given to the classic text of Wozencraft and Jacobs (1965).

The basic idea is to view waveforms, *i.e.*, finite-energy functions, as vectors in a certain vector space that we call *signal-space*. The set of waveforms in this space is the set of finite-energy complex functions $u(t)$ mapping each t on the real time axis into a complex value. We will refer to such a function $u(t)$ as $\boldsymbol{u} : \mathbb{R} \to \mathbb{C}$. Recall that the energy in $\boldsymbol{u}$ is defined as

$$||\boldsymbol{u}||^2 = \int_{-\infty}^{\infty} |u(t)|^2 dt. \tag{1}$$

The set of functions that we are considering is the set for which this integral exists and is finite. This set of waveforms is called the *space of finite-energy complex functions*, or, more briefly, the space of complex waveforms, or, more briefly yet, $\mathcal{L}_2$.

Most of the waveforms that we deal with will be real, and pictures are certainly easier to draw for the real case. However, most of the results are the same for the real and complex case, so long as complex conjugates are included; since the complex conjugate of a real number is just the number itself, the complex conjugates simply provide a minor irritant for real functions. The Fouier transform and Fourier series for a waveform is usually complex, whether the waveform is real or complex, and this provides an additional reason for treating the complex case. Finally, when we treat band-pass functions and their base-band equivalents, it will be necessary to consider the complex case.

CAUTION: Vector spaces are defined over a set of scalars. The appropriate set of scalars for a vector space of real functions is $\mathbb{R}$, whereas the appropriate set of scalars for complex functions is $\mathbb{C}$. In a number of situations, this difference is important. In vector space notation, the set of real functions is a *subset* of the complex functions, but definitely not a *vector subspace*. We will try to flag the cases where this is important, but 'caveat emptor.'

1 Vector spaces

A *vector space* is a set $\mathcal{V}$ of elements defined on a scalar field and satisfying the axioms below. The scalar field of primary interest here is $\mathbb{C}$, in which case the vector space is called a *complex vector space*. Alternatively, if the scalar field is $\mathbb{R}$, the vector space is called a *real vector space*. In either case, the scalars can be added, multiplied, etc. according to the well known rules of $\mathbb{C}$ or $\mathbb{R}$. Neither $\mathbb{C}$ nor $\mathbb{R}$ include ∞. The axioms below do not depend on whether the scalar field is $\mathbb{R}$, $\mathbb{C}$, or some other field. In reading the axioms below, try to keep the set of pairs of real numbers,viewed as points on a plane, in mind and observe that the axioms are saying very natural and obvious statements about addition and scaling of such points. Then reread the axioms thinking about the set of pairs of complex numbers. If you read the axioms very carefully, you will observe that they say nothing about the important geometric ideas of length or angle. We remedy that shortly. The axioms of a vector space are listed below:

(a) Addition: For each $\boldsymbol{v} \in \mathcal{V}$ and $\boldsymbol{u} \in V$, there is a vector $\boldsymbol{v} + \boldsymbol{u} \in \mathcal{V}$ called the sum of $\boldsymbol{v}$ and $\boldsymbol{u}$ satisfying

  (i) Commutativity: $\boldsymbol{v} + \boldsymbol{u} = \boldsymbol{u} + \boldsymbol{v}$,

  (ii) Associativity: $\boldsymbol{v} + (\boldsymbol{u} + \boldsymbol{w}) = (\boldsymbol{v} + \boldsymbol{u}) + \boldsymbol{w}$,

  (iii) There is a unique $\mathbf{0} \in \mathcal{V}$ such that $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$,

  (iv) For each $\boldsymbol{v} \in \mathcal{V}$, there is a unique $-\boldsymbol{v}$ such that $\boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0}$.

(b) Scalar multiplication: For each scalar α and each $\boldsymbol{v} \in \mathcal{V}$ there is a vector $\alpha\boldsymbol{v} \in \mathcal{V}$ called the product of α and $\boldsymbol{v}$ satisfying

  (i) Scalar associativity: $\alpha(\beta\boldsymbol{v}) = (\alpha\beta)\boldsymbol{v}$ for all scalars, α, β, and all $\boldsymbol{v} \in \mathcal{V}$,

  (ii) Unit multiplication: for the unit scalar 1, $1\boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.

(c) Distributive laws:

  (i) For all scalars α and all $\boldsymbol{v}, \boldsymbol{u} \in \mathcal{V}$, $\alpha(\boldsymbol{v} + \boldsymbol{u}) = \alpha\boldsymbol{v} + \alpha\boldsymbol{u}$.

  (ii) For all scalars α, β and all $\boldsymbol{v} \in \mathcal{V}$, $(\alpha + \beta)\boldsymbol{v} = \alpha\boldsymbol{v} + \beta\boldsymbol{v}$.

The space of finite-energy complex functions is a complex vector space. When we view a waveform $v(t)$ as a vector, we denote it by $\boldsymbol{v}$. There are two reasons for doing this: first, it reminds us that we are viewing the waveform as a vector; second, $v(t)$ refers both to the function as a totality and also to the value corresponding to some particular t. Denoting the function as $\boldsymbol{v}$ avoids any possible confusion.

The vector sum, $\boldsymbol{v} + \boldsymbol{u}$ is the function mapping each t into $v(t) + u(t)$. The scalar product $\alpha\boldsymbol{v}$ is the function mapping each t into $\alpha v(t)$. The axioms are pretty easily verified for this space of functions with the possible exception of showing that $\boldsymbol{v} + \boldsymbol{u}$ is a finite-energy function if $\boldsymbol{v}$ and $\boldsymbol{u}$ are. Recall that for any complex numbers v and u, we have $|v + u|^2 \leq 2|v|^2 + 2|u|^2$. Thus, if $\boldsymbol{v}$ and $\boldsymbol{u}$ are finite-energy functions,

$$\int_{-\infty}^{\infty} |v(t) + u(t)|^2 \, dt \leq \int_{-\infty}^{\infty} 2|v(t)|^2 \, dt + \int_{-\infty}^{\infty} 2|u(t)|^2 \, dt < \infty \qquad (2)$$

Also, if $\boldsymbol{v}$ is finite-energy, then $\alpha\boldsymbol{v}$ has α^2 times the energy of $\boldsymbol{v}$, which is still finite. We conclude that the space of finite-energy complex functions forms a vector space using $\mathbb{C}$ as the scalars. Similarly, the space of finite-energy real functions forms a vector space using $\mathbb{R}$ as the scalars.

2 Inner product vector spaces

As mentioned above, a vector space does not in itself contain any notion of length or angle, although such notions are clearly present for vectors viewed as two or three dimensional tuples of complex or real numbers. The missing ingredient is an inner product.

An *inner-product* on a complex vector space $\mathcal{V}$ is a complex valued function of two vectors, $\boldsymbol{v}, \boldsymbol{u} \in \mathcal{V}$, denoted by $\langle \boldsymbol{v}, \boldsymbol{u} \rangle$, that has the following properties:

(a) Hermitian symmetry: $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle^*$:

(b) Hermitian bilinearity: $\langle \alpha\boldsymbol{v} + \beta\boldsymbol{u}, \boldsymbol{w} \rangle = \alpha\langle \boldsymbol{v}, \boldsymbol{w} \rangle + \beta\langle \boldsymbol{u}, \boldsymbol{w} \rangle$
(and consequently $\langle \boldsymbol{v}, \alpha\boldsymbol{u} + \beta\boldsymbol{w} \rangle = \alpha^*\langle \boldsymbol{v}, \boldsymbol{u} \rangle + \beta^*\langle \boldsymbol{v}, \boldsymbol{w} \rangle$);

(c) Strict positivity: $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$, with equality if and only if $\boldsymbol{v} = \boldsymbol{0}$.

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

An inner product on a real vector space is the same, except that the complex conjugates above are redundant.

The *norm* $||\boldsymbol{v}||$, or the *length* of a vector $\boldsymbol{v}$ is defined as

$$||\boldsymbol{v}|| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}.$$

Two vectors $\boldsymbol{v}$ and $\boldsymbol{u}$ are defined to be *orthogonal* if $\langle \mathbf{v}, \boldsymbol{u} \rangle = 0$.

Thus we see that the important geometric notions of length and orthogonality are both defined in terms of the inner product.

3 The inner product space $\mathbb{R}^2$

For the familiar vector space $\mathbb{R}^n$ of real n-tuples, the inner product of vectors $\boldsymbol{v} = (v_1, \ldots v_n)$ and $\boldsymbol{u} = (u_1, \ldots, u_n)$ is usually defined as

$$\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \sum_{j=1}^{n} v_j u_j.$$

You should take the time to verify that this definition satisfies the inner product axioms above. In the special case $\mathbb{R}^2$ (see Figure 1), a vector $\boldsymbol{v} = (v_1, v_2)$ is a point in the plane, with v_1 and v_2 representing its horizontal and vertical location. Alternatively, $\boldsymbol{v}$ can be viewed as the directed line from $\boldsymbol{0}$ to the point (v_1, v_2).

The inner product determines geometric quantities such as length and relative orientation of vectors. In particular, by plane geometry,

- the distance from $\boldsymbol{0}$ to $\boldsymbol{u}$ is $||\boldsymbol{u}|| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}$,

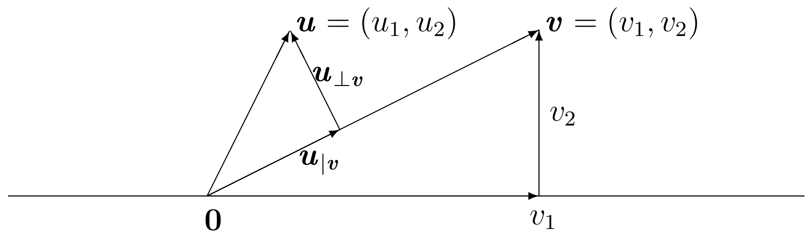

Figure 1: Two vectors, $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$ in $\mathbb{R}^2$. Note that $||\boldsymbol{v}||^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle = v_1^2 + v_2^2$ is the squared length of $\boldsymbol{v}$ (viewed as a directed line).

- the distance from $\mathbf{v}$ to $\mathbf{u}$ (and from $\mathbf{u}$ to $\mathbf{v}$) is $||\boldsymbol{v} - \boldsymbol{u}||$
- $\cos(\angle(\boldsymbol{v} - \boldsymbol{u})) = \frac{\langle \boldsymbol{v}, \boldsymbol{u} \rangle}{||\boldsymbol{v}|| \, ||\boldsymbol{u}||}$

4 One Dimensional Projections

Using the cosine formula above, or using direct calculation, we can view the vector $\boldsymbol{u}$ as being the sum of two vectors, $\boldsymbol{u} = \boldsymbol{u}_{|\boldsymbol{v}} + \boldsymbol{u}_{\perp \boldsymbol{v}}$, where $\boldsymbol{u}_{|\boldsymbol{v}}$ is collinear with $\boldsymbol{v}$ and $\boldsymbol{u}_{\perp \boldsymbol{v}} = \boldsymbol{u} - \boldsymbol{u}_{|\boldsymbol{v}}$ is orthogonal to $\boldsymbol{v}$ (see Figure 1). We can view this as dropping a perpendicular from $\boldsymbol{u}$ to $\boldsymbol{v}$. The word 'perpendicular' is a synonym[1] for orthogonal and is commonly used for $\mathbb{R}^2$ and $\mathbb{R}^3$. In particular,

$$\boldsymbol{u}_{|\boldsymbol{v}} = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{||\boldsymbol{v}||^2} \, \boldsymbol{v} \tag{3}$$

This vector $\boldsymbol{u}_{|\boldsymbol{v}}$ is called the *projection* of $\boldsymbol{u}$ onto $\boldsymbol{v}$. Rather surprisingly, (3) is valid for any inner product space.

Theorem 4.1 (One Dimensional Projection Theorem) *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in an inner product space (either on $\mathbb{C}$ or $\mathbb{R}$). Then $\boldsymbol{u} = \boldsymbol{u}_{|\boldsymbol{v}} + \boldsymbol{u}_{\perp \boldsymbol{v}}$ where $\boldsymbol{u}_{|\boldsymbol{v}} = \alpha \boldsymbol{v}$ (*i.e.*, $\boldsymbol{u}_{|\boldsymbol{v}}$ is collinear with $\boldsymbol{v}$) and $\langle \boldsymbol{u}_{\perp \boldsymbol{v}}, \boldsymbol{v} \rangle = 0$ (*i.e.*, $\boldsymbol{u}_{\perp \boldsymbol{v}}$ is orthogonal to $\boldsymbol{v}$). The scalar α is uniquely given by $\alpha = \langle \boldsymbol{u}, \boldsymbol{v} \rangle / ||\boldsymbol{v}||^2$.*

The vector $\boldsymbol{u}_{|\boldsymbol{v}}$ is called the *projection* of $\boldsymbol{u}$ on $\boldsymbol{v}$.

Proof: We must show that $\boldsymbol{u}_{\perp \boldsymbol{v}} = \boldsymbol{u} - \boldsymbol{u}_{|\boldsymbol{v}}$ is orthogonal to $\boldsymbol{v}$.

$$\begin{aligned} \langle \boldsymbol{u}_{\perp \boldsymbol{v}}, \boldsymbol{v} \rangle &= \langle \boldsymbol{u} - \boldsymbol{u}_{|\boldsymbol{v}}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle - \langle \boldsymbol{u}_{|\boldsymbol{v}}, \boldsymbol{v} \rangle \\ &= \langle \boldsymbol{u}, \boldsymbol{v} \rangle - \alpha \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle - \alpha ||\boldsymbol{v}||^2 \end{aligned}$$

which is 0 if and only if $\alpha = \langle \boldsymbol{u}, \boldsymbol{v} \rangle / ||\boldsymbol{v}||^2$ $\square$

[1] Two vectors are called perpendicular if the cosine of the angle between them is 0. This is equivalent to the inner product being 0, which is the definition of orthogonal.

Another well known result in $\mathbb{R}^2$ that carries over to any inner product space is the Pythagorean theorem: If $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal, then

$$||\boldsymbol{u}+\boldsymbol{v}||^2 = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 \tag{4}$$

To see this, note that

$$\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

The cross terms disappear by orthogonality, yielding (4).

There is an important corollary to Theorem 4.1 called the *Schwartz inequality.*

Corollary 1 (Schwartz inequality) *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in an inner product space (either on $\mathbb{C}$ or $\mathbb{R}$). Then*

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \leq ||\boldsymbol{u}||\ ||\boldsymbol{v}|| \tag{5}$$

Proof: Since $\boldsymbol{u}_{|\boldsymbol{v}}$ and $\boldsymbol{u}_{\perp\boldsymbol{v}}$ are orthogonal, we use (4) to get

$$||\boldsymbol{u}||^2 = ||\boldsymbol{u}_{|\boldsymbol{v}}||^2 + ||\boldsymbol{u}_{\perp\boldsymbol{v}}||^2.$$

Since each term above is non-negative,

$$\begin{aligned} ||\boldsymbol{u}||^2 &\geq ||\boldsymbol{u}_{|\boldsymbol{v}}||^2 = |\alpha|^2||\boldsymbol{v}||^2 \\ &= \left|\frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{||\boldsymbol{v}||^2}\right|^2 ||\boldsymbol{v}||^2 = \frac{|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|^2}{||\boldsymbol{v}||^2} \end{aligned}$$

This is equivalent to (5). □

5 The inner product space of finite-energy complex functions

The inner product of two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in a vector space $\mathcal{V}$ of finite-energy complex functions is defined as

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \int_{-\infty}^{\infty} u(t)v^*(t)dt \tag{6}$$

Since the vectors are finite energy, $\langle \boldsymbol{v}, \boldsymbol{v}\rangle < \infty$. It is shown in problem 6.3 that this integral must be finite. Note that the Schwartz inequality proven above cannot be used for this, since we have not yet shown that this is an inner product space. However, with this finiteness, the first two axioms of an inner product space follow immediately.

The final axiom of an inner product space is that $\langle \boldsymbol{v}, \boldsymbol{v}\rangle \geq 0$ with equality iff $\boldsymbol{v} = \boldsymbol{0}$. We have already seen that if a function $v(t)$ is zero for all but a finite number of values of t (or even at all but a countable number), then its energy is 0, but the function is not identically 0. This is a nit-picking issue at some level, but it comes up when we look at the difference between functions and expansions of those functions.

In very mathematical treatments, this problem is avoided by defining equivalence classes of functions, where two functions are in the same equivalence class if their difference has zero energy. In this type of very careful treatment, a vector is an equivalence class of functions. However, we don't have the mathematical tools to describe these equivalence classes precisely. As a more engineering point of view, we continue to view functions as vectors, but say that two functions are $\mathcal{L}_2$ equivalent if their difference has zero energy. We then ignore the value of functions at discontinuities, but do not feel the need to specify the class of functions that are $\mathcal{L}_2$ equivalent. Thus, for example, we refer to a function and the inverse Fourier transform of its Fourier transform as $\mathcal{L}_2$ equivalent. We all recognize that it is unimportant what value a function has at a point of discontinuity, and the notion of $\mathcal{L}_2$ equivalence lets us ignore this issue.

We will come back to this question of $\mathcal{L}_2$ equivalence several times, but for now, we simply recognize that the set of finite-energy complex functions is an inner product space subject to $\mathcal{L}_2$ equivalence; this space is called $\mathcal{L}_2$ for brevity.

At this point, we can use all the machinery of inner product spaces, including projection and the Schwartz inequality.

6 Subspaces of inner product spaces

A *subspace* $\mathcal{S}$ of a vector space $\mathcal{V}$ is a subset of the vectors in $\mathcal{V}$ which form a vector space in their own right (over the same set of scalars as used by $\mathcal{V}$).

Example: Consider the vector space $\mathbb{R}^3$, *i.e.*, the vector space of real 3-tuples. Geometrically, we regard this as a volume where there are three orthogonal coordinate directions, and where u_1, u_2, u_3 specify the length of $\boldsymbol{u}$ in each of those directions. We can define three unit vectors, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ in these three coordinate directions respectively, so that $\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + u_3\boldsymbol{e}_3$.

Consider the subspace of $\mathbb{R}^3$ composed of all linear combinations of $\boldsymbol{u}$ and $\boldsymbol{v}$ where, for example $\boldsymbol{u} = (1, 0, 1)$ and $\boldsymbol{v} = (0, 1, 1)$. Geometrically, the subspace is then the set of vectors of the form $\alpha\boldsymbol{u} + \beta\boldsymbol{v}$ for arbitrary scalars α and β. Geometrically, this subspace is a plane going through the points $\boldsymbol{0}, \boldsymbol{u}$, and $\boldsymbol{v}$. In this plane (as in the original vector space) $\boldsymbol{u}$ and $\boldsymbol{v}$ each have length $\sqrt{2}$ and the cosine of the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is $1/2$.

The projection of $\boldsymbol{u}$ on $\boldsymbol{v}$ is $\boldsymbol{u}_{|\boldsymbol{v}} = (0, 1/2, 1/2)$ and $\boldsymbol{u}_{\perp\boldsymbol{v}} = (1, 1/2, -1/2)$. We can view vectors in this subspace, pictorially and geometrically, in just the same way as we viewed vectors in R^2 before.

7 Linear independence and dimension

A set of vectors, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ in a vector space $\mathcal{V}$ is *linearly dependent* if $\sum_{j=1}^{n} \alpha_j \boldsymbol{v}_j = 0$ for some set of scalars that are not all equal to 0. Equivalently, the set is linearly dependent if one of those vectors, say $\boldsymbol{v}_n$, is a linear combination of the others, *i.e.*, if for some set of scalars,

$$\boldsymbol{v}_n = \sum_{j=1}^{n-1} \alpha_j \boldsymbol{v}_j$$

A set of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ is linearly independent if it is not linearly dependent, *i.e.*, if $\sum_{j=1}^{n} \alpha_j \boldsymbol{v}_j = 0$ only if each scalar is 0. An equivalent condition is that no vector in the set is a linear combination of the others. Often, instead of saying that a set $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is linearly dependent or independent, we say that the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are dependent or independent.

For the example above, $\boldsymbol{u}$ and $\boldsymbol{v}$ are independent. In contrast, $\boldsymbol{u}_{|\boldsymbol{v}}, \boldsymbol{v}$ are dependent. As a more peculiar example, $\boldsymbol{v}$ and $\mathbf{0}$ are dependent ($\alpha\mathbf{0} + 0\boldsymbol{v}$ for all α). Also the singleton set $\{\mathbf{0}\}$ is a linearly dependent set.

The *dimension* of a vector space is the number of vectors in the largest linearly independent set in that vector space. Similarly, the *dimension* of a subspace is the number of vectors in the largest linearly independent set in that subspace. For the example above, the dimension of $\mathbb{R}^3$ is 3 and the dimension of the subspace is 2.[2] We will see shortly that the dimension of $\mathcal{L}_2$ is infinity, *i.e.*, there is no largest finite set of linearly independent vectors representing waveforms.

For a finite dimensional vector space of dimension n, we say that any linearly independent set of n vectors in that space is a *basis* for the space. Since a subspace is a vector space in its own right, any linearly independent set of n vectors in a subspace of dimension n is a basis for that subspace.

Theorem 7.1 *Let $\mathcal{V}$ be an n-dimensional vector space and let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ be a basis for $\mathcal{V}$. Then every vector $\boldsymbol{u} \in \mathcal{V}$ can be represented as a unique linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$.*

Proof: Assume that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ is a basis for $\mathcal{V}$. Since $\mathcal{V}$ has dimension n, every set of $n+1$ vectors is linearly dependent. Thus, for any $\boldsymbol{u} \in \mathcal{V}$,

$$\beta u + \sum_j \alpha_j \boldsymbol{v}_j = 0$$

for some choice of scalars not all 0. Since $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$ are independent, β above must be non-zero, and thus

$$\boldsymbol{u} = \sum_{j=1}^{n} \frac{-\alpha_j}{\beta} \boldsymbol{v}_j \tag{7}$$

[2] You should be complimented if you stop at this point and say "yes, but how does that follow from the definitions." We give the answer to that question very shortly.

If this representation is not unique, then $\boldsymbol{u}$ can also be expressed as $\boldsymbol{u} = \sum_j \gamma_j \boldsymbol{v}_j$ where $\gamma_j \neq -\alpha_j/\beta$ for some j. Subtracting one representation from the other, we get

$$\mathbf{0} = \sum_j (\gamma_j + \alpha_j/\beta)\boldsymbol{v}_j$$

which violates the assumption that $\{\boldsymbol{v}_j\}$ is an independent set. $\square$

A set of vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in \mathcal{V}$ is said to *span* $\mathcal{V}$ if every vector $\boldsymbol{u} \in \mathcal{V}$ is a linear combination of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in \mathcal{V}$. Thus the theorem above can be more compactly stated as: if $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in \mathcal{V}$ is a basis for $\mathcal{V}$, then it spans $\mathcal{V}$.

Let us summarize where we are. We have seen that any n-dimensional space must have a basis consisting of n independent vectors and that any set of n linearly independent vectors is a basis. We have also seen that any basis for $\mathcal{V}$ spans $\mathcal{V}$. We have also seen that the space contains no set of more than n linearly independent vectors. What still remains to be shown is the important fact that no set of fewer than n vectors spans $\mathcal{V}$. Before showing this, we illustrate the wide range of choice in choosing a basis by the following result:

Theorem 7.2 *Let $\boldsymbol{u} \in \mathcal{V}$ be an arbitrary non-zero vector and let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ be a basis for $\mathcal{V}$. Then $\boldsymbol{u}$ combined with $n-1$ of the original basis vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ forms another basis of $\mathcal{V}$.*

Proof: From Theorem (7.1), we can uniquely represent $\boldsymbol{u}$ as $\sum_j \gamma_j \boldsymbol{v}_j$ where not all $\{\gamma_j\}$ are 0. Let i be an index for which $\gamma_i \neq 0$. Then

$$\boldsymbol{v}_i = \frac{1}{\gamma_i}\boldsymbol{u} - \sum_{j \neq i} \frac{\gamma_j}{\gamma_i}\boldsymbol{v}_j \tag{8}$$

Since the $\{\gamma_j\}$ are unique, this representation of $\boldsymbol{v}_i$ is unique. The set of vectors $\boldsymbol{u}, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \dots, \boldsymbol{v}_n$ must then be linearly independent. Otherwise a linear combination of them, with not all zero coefficients, would sum to $\mathbf{0}$; this linear combination could be added to the right hand side of (8), contradicting its uniqueness. Thus this set of vectors forms a basis. $\square$

Theorem 7.3 *For an n-dimensional vector space $\mathcal{V}$, any set of $m < n$ linearly independent vectors do not span $\mathcal{V}$, but can be extended by $n - m$ additional vectors to span $\mathcal{V}$.*

Proof: Let $\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$ be a set of $m < n$ linearly independent vectors and let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ be a basis of $\mathcal{V}$. By Theorem 7.2, we can replace one of the basis vectors $\{\boldsymbol{v}_j\}$ by $\boldsymbol{w}_1$ and still have a basis. Repeating, we can replace one of the elements of this new basis by $\boldsymbol{w}_2$ and have another basis. This process is continued for m steps. At each step, we remove one of the original basis elements, which is always possible because $\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$ are linearly independent. Finally, we have a basis containing $\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$ and $n - m$ vectors of the original basis. This shows that $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ can be extended into a basis. Any vector, say $\boldsymbol{v}_i$, remaining from the original basis is linearly independent of $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$. Thus $\boldsymbol{v}_i \in \mathcal{V}$ is not a linear combination of $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$, which therefore do not span $\mathcal{V}$. □

We can summarize the results of this section about any n-dimensional vector space (or subspace) $\mathcal{V}$ as follows:

- Every set of n linearly independent vectors forms a basis and spans $\mathcal{V}$.
- Every set of n vectors that spans $\mathcal{V}$ is a linearly independent set and is a basis of $\mathcal{V}$.
- Every set of $m < n$ independent vectors can be extended to a basis.
- If a set of n independent vectors in a space $\mathcal{U}$ spans $\mathcal{U}$, then $\mathcal{U}$ is n-dimensional.

These results can now be applied to the example of the previous section. $\mathbb{R}^3$ is a 3-dimensional space according to the definitions here, since the three unit co-ordinate vectors are independent and span the space. Similarly, the subspace $\mathcal{S}$ is a 2-dimenional space, since $\boldsymbol{u}$ and $\boldsymbol{y}$ are independent and, by definition, span the subspace.

For inner product spaces, almost everything we do will be simpler if we consider orthonormal bases, and we turn to this in the next section.

8 Orthonormal bases and the projection theorem

In an inner product space, a set of vectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots$ is *orthonormal* if

$$\langle \boldsymbol{\phi}_j, \boldsymbol{\phi}_k \rangle = \begin{cases} 0 & \text{for} \quad j \neq k \\ 1 & \text{for} \quad j = k \end{cases} \tag{9}$$

In other words, an orthonormal set is a set of orthogonal vectors where each vector is *normalized* in the sense of having unit length. It can be seen that if a set of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$ is orthogonal, then the set

$$\boldsymbol{\phi}_j = \frac{1}{\|\boldsymbol{v}_j\|} \boldsymbol{v}_j$$

is orthonormal. Note that if two vectors are orthogonal, then any scaling (including normalization) of each vector maintains the orthogonality.

If we project a vector $\boldsymbol{u}$ onto a normalized vector $\boldsymbol{\phi}$, then, from Theorem 4.1, the projection has the simplified form

$$\boldsymbol{u}_{|} \;= \langle \boldsymbol{u}, \boldsymbol{\phi} \rangle \boldsymbol{\phi} \tag{10}$$

We now generalize the Projection Theorem to the projection of a vector $\boldsymbol{u} \in \mathcal{V}$ onto a finite dimensional subspace $\mathcal{S}$ of $\mathcal{V}$.

8.1 Finite-dimensional projections

Recall that the one dimensional projection of a vector $\boldsymbol{u}$ on a normalized vector $\boldsymbol{\phi}$ is given by (10). The vector $\boldsymbol{u}_\perp = \boldsymbol{u} - \boldsymbol{u}_|$ is orthogonal to $\boldsymbol{\phi}$.

More generally, if $\mathcal{S}$ is a subspace of an inner product space $\mathcal{V}$, and $\boldsymbol{u} \in \mathcal{V}$, the *projection of* $\boldsymbol{u}$ *on* $\mathcal{S}$ is defined to be a vector $\boldsymbol{u}_{|\mathcal{S}} \in \mathcal{S}$ such that $\boldsymbol{u} - \boldsymbol{v}_{|\mathcal{S}}$ is orthogonal to all vectors in $\mathcal{S}$.

Note that the earlier definition of projection is a special case of that here if we take the subspace $\mathcal{S}$ to be the one dimensional subspace spanned by $\boldsymbol{\phi}$. In other words, saying that $\boldsymbol{u}_|$ is collinear with $\boldsymbol{\phi}$ is the same as saying it is in the one dimensional subspace spanned by $\boldsymbol{\phi}$.

Theorem 8.1 (Projection theorem) *Let $\mathcal{S}$ be an n-dimensional subspace of an inner product space $\mathcal{V}$ and assume that $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_n$ is an orthonormal basis for $\mathcal{S}$. Then any $\boldsymbol{u} \in V$ may be decomposed as $\boldsymbol{u} = \boldsymbol{u}_{|\mathcal{S}} + \boldsymbol{u}_{\perp\mathcal{S}}$ where $\boldsymbol{u}_{|\mathcal{S}} \in \mathcal{S}$ and $\langle \boldsymbol{u}_{\perp\mathcal{S}}, \boldsymbol{s} \rangle = 0$ for all $\boldsymbol{s} \in \mathcal{S}$. Furthermore, $\boldsymbol{u}_{|\mathcal{S}}$ is uniquely determined by*

$$\boldsymbol{u}_{|\mathcal{S}} = \sum_{j=1}^{n} \langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle \boldsymbol{\phi}_j \tag{11}$$

Note that the theorem above assumes that $\mathcal{S}$ has a set of orthonormal vectors as a basis. We will show later that any finite dimensional inner product space has such an orthonomal basis, so that the assumption does not restrict the generality of the theorem.

Proof: Let $\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{\phi}_i$ be an arbitrary vector in $\mathcal{S}$. We first find the conditions on $\boldsymbol{v}$ under which $\boldsymbol{u} - \boldsymbol{v}$ is orthogonal to all vectors $\boldsymbol{s} \in \mathcal{S}$. We know that $\boldsymbol{u} - \boldsymbol{v}$ is orthogonal to all $\boldsymbol{s} \in \mathcal{S}$ if and only if

$$\langle \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{\phi}_j \rangle = 0, \quad \text{for all } j,$$

or equivalently if and only if

$$\langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle = \langle \boldsymbol{v}, \boldsymbol{\phi}_j \rangle, \quad \text{for all } j. \tag{12}$$

Now since $\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{\phi}_i$, we have

$$\langle \boldsymbol{v}, \boldsymbol{\phi}_j \rangle = \sum_{i=1}^{n} \alpha_i \langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \rangle = \alpha_j, \quad \text{for all } j, \tag{13}$$

This along with (12) uniquely determines the coefficients $\{\alpha_j\}$. Thus $\boldsymbol{u} - \boldsymbol{v}$ is orthogonal to all vectors in $\mathcal{S}$ if and only if $\boldsymbol{v} = \sum_j \langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle \boldsymbol{\phi}_j$. Thus $\boldsymbol{u}_{|\mathcal{S}}$ in (11) is the unique vector in $\mathcal{S}$ for which $\boldsymbol{u}_{\perp\mathcal{S}} = \boldsymbol{u} - \boldsymbol{u}_{|\mathcal{S}}$ is orthogonal to all $\boldsymbol{s} \in \mathcal{S}$. □

As we have seen a number of times, if $\boldsymbol{v} = \sum_j \alpha_j \boldsymbol{\phi}_j$ for orthonormal functions $\{\boldsymbol{\phi}_j\}$, then

$$\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \sum_{j=1}^{n} \alpha_j \boldsymbol{\phi}_j \rangle = \sum_{j=1}^{n} \alpha_j^* \langle \boldsymbol{v}, \boldsymbol{\phi}_j \rangle = \sum_{j=1}^{n} |\alpha_j|^2$$

where we have used (13) in the last step. Thus, for the projection $\boldsymbol{u}_{|\mathcal{S}}$, we have the important relation

$$\|\boldsymbol{u}_{|\mathcal{S}}\|^2 = \sum_{j=1}^{n} |\langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle|^2 \tag{14}$$

8.2 Corollaries of the projection theorem

The projection theorem in the previous subsection assumes the existence of an orthonormal basis for the subspace. We will show that such a basis always exists in the next subsection, and simply assume that here. Thus, for any n-dimensional subspace $\mathcal{S}$ of an inner product space $\mathcal{V}$ and for any $\boldsymbol{u} \in \mathcal{V}$, there is a unique decomposition of $\boldsymbol{u}$ into two orthogonal components, one, $\boldsymbol{u}_{|\mathcal{S}}$ in $\mathcal{S}$, and the other, $\boldsymbol{u}_{\perp\mathcal{S}}$, orthogonal to all vectors in $\mathcal{S}$ (including, of course, $\boldsymbol{u}_{|\mathcal{S}}$). Letting $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_n$ be an orthonormal basis for $\mathcal{S}$, $\boldsymbol{u}_{|\mathcal{S}} = \sum_j \langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle \boldsymbol{\phi}_j$.

From the Pythagorean Theorem,

$$\|\boldsymbol{u}\|^2 = \|\boldsymbol{u}_{|\mathcal{S}}\|^2 + \|\boldsymbol{u}_{\perp\mathcal{S}}\|^2. \tag{15}$$

This gives us the corollary

Corollary 2 (norm bounds)

$$0 \le \|\boldsymbol{u}_{|\mathcal{S}}\|^2 \le \|\boldsymbol{u}\|^2. \tag{16}$$

with equality on the right if and only if $\boldsymbol{u} \in \mathcal{S}$ and equality on the left if and only if $\boldsymbol{u}$ is orthogonal to all vectors in $\mathcal{S}$.

Recall from (14) that $\|\boldsymbol{u}_{|\mathcal{S}}\|^2 = \sum_j |\langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle|^2$. Substituting this in (16), we get Bessel's inequality. Bessel's inequality is very useful in understanding the convergence of orthonormal expansions.

Corollary 3 (Bessel's inequality)

$$0 \le \sum_{j=1}^{n} |\langle \boldsymbol{u}, \boldsymbol{\phi}_j \rangle|^2 \le \|\boldsymbol{u}\|^2.$$

with equality on the right if and only if $\boldsymbol{u} \in \mathcal{S}$ and equality on the left if and only if $\boldsymbol{u}$ is orthogonal to all vectors in $\mathcal{S}$.

Another useful characterization of the projection $\boldsymbol{u}_{|\mathcal{S}}$ is that it is the vector in $\mathcal{S}$ that is closest to $\boldsymbol{u}$. In other words, it is the vector $\boldsymbol{s} \in \mathcal{S}$ that yields the least squared error (LSE) $\|\boldsymbol{u} - \boldsymbol{s}\|^2$:

Corollary 4 (LSE property) *The projection* $\boldsymbol{u}_{|\mathcal{S}}$ *is the* unique *closest vector* $\boldsymbol{s} \in \mathcal{S}$ *to* $\boldsymbol{u}$*; i.e., for all* $\boldsymbol{s} \in \mathcal{S}$,

$$\|\boldsymbol{u} - \boldsymbol{u}_{|\mathcal{S}}\| \leq \|\boldsymbol{u} - \boldsymbol{s}\|$$

with equality if and only if $\boldsymbol{s} = \boldsymbol{u}_{|\mathcal{S}}$.

Proof: By the projection theorem, $\boldsymbol{u} - \boldsymbol{s} = \boldsymbol{u}_{|\mathcal{S}} + \boldsymbol{u}_{\perp\mathcal{S}} - \boldsymbol{s}$. Since both $\boldsymbol{u}_{|\mathcal{S}}$ and $\boldsymbol{s}$ are in $\mathcal{S}$, $\boldsymbol{u}_{|\mathcal{S}} - \boldsymbol{s}$ is in $\mathcal{S}$, so by Pythagoras,

$$\|\boldsymbol{u} - \boldsymbol{s}\|^2 = \|\boldsymbol{u}_{|\mathcal{S}} - \boldsymbol{s}\|^2 + \|\boldsymbol{u}_{\perp\mathcal{S}}\|^2.$$

By strict positivity, $\|\boldsymbol{u}_{|\mathcal{S}} - \boldsymbol{s}\|^2 \geq 0$, with equality if and only if $\boldsymbol{s} = \boldsymbol{u}_{|\mathcal{S}}$. $\square$

8.3 Gram-Schmidt orthonormalization

In our development of the Projection Theorem, we assumed an orthonormal basis $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n$ for any given n-dimensional subspace of $\mathcal{V}$. The use of orthnormal bases simplifies almost everything we do, and when we get to infinite dimensional expansions, orthonormal bases are even more useful. In this section, we give the Gram-Schmidt procedure, which, starting from an arbitrary basis $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ for an n-dimensional subspace, generates an orthonormal basis.

The procedure is not only useful in actually finding orthonormal bases, but is also useful theoretically, since it shows that such bases always exist. In particular, this shows that the Projection Theorem did not lack generality. The procedure is almost obvious given what we have already done, since, with a given basis, $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$, we start with a subspace generated by $\boldsymbol{s}_1$, use the projection theorem on that space to find a linear combination of $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ that is orthogonal to $\boldsymbol{s}_1$, and work ourselves up gradually to the entire subspace.

More specifically, let $\boldsymbol{\phi}_1 = \boldsymbol{s}_1/\|\boldsymbol{s}_1\|$. Thus $\boldsymbol{\phi}_1$ is an orthonormal basis for the subspace $\mathcal{S}_1$ of vectors generated by $\boldsymbol{s}_1$. Applying the projection theorem, $\boldsymbol{s}_2 = \langle \boldsymbol{s}_1, \boldsymbol{\phi}_1 \rangle \boldsymbol{\phi}_1 + (\boldsymbol{s}_2)_{\perp\mathcal{S}_1}$, where $(\boldsymbol{s}_2)_{\perp\mathcal{S}_1}$ is orthogonal to $\boldsymbol{s}_1$ and thus to $\boldsymbol{\phi}_1$.

$$(\boldsymbol{s}_2)_{\perp\mathcal{S}_1} = \boldsymbol{s}_2 - \langle \boldsymbol{s}_1, \boldsymbol{\phi}_1 \rangle \boldsymbol{\phi}_1$$

Normalizing this,

$$\boldsymbol{\phi}_2 = (\boldsymbol{s}_2)_{\perp\mathcal{S}_1}/\|\boldsymbol{u}_{\perp\mathcal{S}_1}\|$$

Since $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are both linear combinations of $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$, the subspace $\mathcal{S}_2$ generated by $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ is also generated by the orthonormal basis $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$.

Now suppose, using induction, we have generated an orthonormal set $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_k$ to generate the subspace $\mathcal{S}_k$ generated by $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k$. We then have

$$(\boldsymbol{s}_{k+1})_{\perp\mathcal{S}_k} = \boldsymbol{s}_{k+1} - \sum_{j=1}^{k} \langle \boldsymbol{s}_{k+1}, \boldsymbol{\phi}_j \rangle \boldsymbol{\phi}_j$$

Normalizing this,

$$\boldsymbol{\phi}_{k+1} = \frac{(\boldsymbol{s}_{k+1})_{\perp \mathcal{S}_k}}{\|(s_{k+1})_{\perp \mathcal{S}_k}\|}$$

Note that $\boldsymbol{s}_{k+1}$ is a linear combination of $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{k+1}$. By induction, $\boldsymbol{s}_1, \dots, \boldsymbol{s}_k$ are all linear combinations of $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_k$. Thus the set of orthonormal vectors $\{\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{k+1}\}$ generates the subspace $\mathcal{S}_{k+1}$ generated by $\boldsymbol{s}_1, \dots, \boldsymbol{s}_{k+1}$.

In summary, we have shown that the Gram-Schmidt orthogonalization procedure produces an orthonormal basis $\{\boldsymbol{\phi}_j\}$ for an arbitrary n dimensional subspace $\mathcal{S}$ with the original basis $\boldsymbol{s}_1, \dots, \boldsymbol{s}_n$.

If we start with a set of vectors that are not independent, then the algorithm will find any vector $\boldsymbol{s}_i$ that is a linear combination of previous vectors, will eliminate it, and proceed to the next vector.

9 Orthonormal expansions in $\mathcal{L}_2$

We now have the background to understand the orthogonal expansions such as the Sampling Theorem and the Fourier series that we have been using to convert waveforms into sequences of real or complex numbers. The Projection Theorem of the last section was restricted to finite dimensional subspaces, but the underlying inner product space was arbitrary. Thus it applies to vectors in $\mathcal{L}_2$, *i.e.*, to waveforms.

As an example, consider the Fourier series again, where for each integer k, $-\infty < k < \infty$, we defined $\theta_k(t)$ as $e^{2\pi i k t/T}$ for $\tau \le t \le \tau + T$ and zero elsewhere. Here it is convenient to normalize these functions, *i.e.*,

$$\phi_k(t) = \begin{cases} \sqrt{1/T}\, e^{2\pi i k t/T} & \text{for} \quad \tau \le t \le \tau + T \\ 0 & \text{otherwise} \end{cases}$$

The Fourier series expansion then becomes

$$u(t) = \sum_{k=-\infty}^{\infty} \alpha_k \phi_k(t) \qquad \text{where} \tag{17}$$

$$\alpha_k = \int_{-\infty}^{\infty} u(t)\phi_k^*(t)\, dt \tag{18}$$

As vectors, $\boldsymbol{u} = \sum_k \alpha_k \boldsymbol{\phi}_k$ where $\alpha_k = \langle \boldsymbol{u}, \boldsymbol{\phi}_k \rangle$. If the infinite sum here means anything, we can interpret it as some kind of limit of the partial sums,

$$\boldsymbol{u}^{(n)} = \sum_{k=-n}^{n} \langle \boldsymbol{u}, \boldsymbol{\phi}_k \rangle \boldsymbol{\phi}_k$$

We recognize this as the projection of $\boldsymbol{u}$ on the subspace $\mathcal{S}_n$ spanned by $\boldsymbol{\phi}_k$ for $-n \le k \le n$. In other words, $\boldsymbol{u}^{(n)} = \boldsymbol{u}_{|\mathcal{S}_n}$. From Corollary 3, $\boldsymbol{u}_{|\mathcal{S}_n}$ is the closest point in $\mathcal{S}_n$ to $\boldsymbol{u}$. This approximation minimizes the energy between $\boldsymbol{u}$ and any vector in $\mathcal{S}_n$.

We also have an expression for the energy in the projection $\boldsymbol{u}_{|\mathcal{S}_n}$. From (14), allowing for the different indexing here, we have

$$\|\boldsymbol{u}_{|\mathcal{S}_n}\|^2 = \sum_{k=-n}^{n} |\langle \boldsymbol{u}, \boldsymbol{\phi}_k \rangle|^2 \tag{19}$$

This quantity is non-decreasing with n, but from Bessel's inequality, it is bounded by $\|\boldsymbol{u}\|^2$, which is finite. Thus, as n increases, the terms $\|\langle \boldsymbol{u}, \boldsymbol{\phi}_n \rangle\|^2 + \|\langle \boldsymbol{u}, \boldsymbol{\phi}_{-n} \rangle\|^2$ are approaching 0. In particular, if we look at the difference between $\boldsymbol{u}_{|\mathcal{S}_m}$ and $\boldsymbol{u}_{|\mathcal{S}_n}$, we see that for $n < m$,

$$\lim_{n,m\to\infty} \sum_{j:n<|k|\le m} \|\boldsymbol{u}_{|\mathcal{S}_m} - \boldsymbol{u}_{|\mathcal{S}_n}\|^2 = \lim_{n,m\to\infty} \sum_{k:n<|k|\le m} |\langle \boldsymbol{u}, \boldsymbol{\phi}_k \rangle|^2 = 0$$

This says that the sequence of projections $\{\boldsymbol{u}_{|\mathcal{S}_n}\}$ are approaching each other in terms of energy difference. A sequence for which the terms approach each other is called a *Cauchy sequence*. The problem is that when funny things like functions and vectors get close to each other, it doesn't necessarily mean that they have any limit. Fortunately, the Riesz-Fischer theorem says that indeed these vectors do have a limit (*i.e.*, that Cauchy sequences for $\mathcal{L}_2$ converge to actual vectors). This property that Cauchy sequences converge to something is called completeness. A complete inner product space is called a Hilbert space.

So what have we learned from this rather long sojourn into abstract mathematics? We already knew the Fourier series was a relatively friendly way to turn waveforms into sequences, and we have the intuition that truncating a fourier series simply cuts off the high frequency components of the waveform.

Probably the most important outcome is that *all* orthonormal expansions may be looked at in a common framework and are all subject to the same types of results. The Fourier series is simply one example.

An equally important outcome is that as additional terms are added to a truncated orthonormal expansion, the result is changing (in an energy sense) by increasingly small amounts. One reason for restricting ourselves to finite energy waveforms (aside from physical reality) is that as their energy gets used up in different degrees of freedom (*i.e.*, components of the expansion), there is less to be used in other degrees of freedom, so that in an energy sense, we expect some sort of convergence. The exploration of these mathematical results simply attempts to make this intuition a little more precise.

Another outcome is the realization that these truncated sums can be treated simply with no sophisticated mathematical issues such as equivalence etc. How the truncated sums converge, of course, is tricky mathematically, but the truncated sums themselves are very simple and friendly.