Function Spaces

§1. Inner products and norms. From linear algebra, we recall that an *inner product* for a complex vector space $\mathcal{V}$ is a function $< \cdot, \cdot >: \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ that satisfies the following properties.

I1. *Positivity:* $< v, v > \geq 0$ for every $v \in \mathcal{V}$, with $< v, v >= 0$ if and only if $v = 0$.

I2. *Conjugate symmetry:* $\overline{< v, w >} =< w, v >$ for every $v \in \mathcal{V}$ and $w \in \mathcal{V}$.

I3. *Homogeneity:* $< cv, w >= c < v, w >$ for all v, $w \in \mathcal{V}$ and $c \in \mathbb{C}$.

I4. *Additivity:* $< u + v, w >=< u, w > + < v, w >$ for all u, v, $w \in \mathcal{V}$ and $c \in \mathbb{C}$.

For real vector spaces, conjugate symmetry is replaced by symmetry, $< v, w >=< w, v >$. In the complex case, there are two simple, immediate consequences of these properties. First, from conjugate symmetry and homogeneity, we have $< v, cw >= \bar{c} < v, w >$; second, from conjugate symmetry and additivity, we get $< v, u + w >=< v, u > + < v, w >$. The *norm* $\| \cdot \|$ associated with an inner product $< \cdot, \cdot >$ on $\mathcal{V}$ is defined to be the quantity $\|v\| := \sqrt{< v, v >}$.

A vector space together with an inner product is called an *inner product space.* Some simple examples of such spaces are: (1) $\mathbb{C}^n$, the space of n-component column vectors with complex entries, with the inner product being $< v, w >= \bar{w}^T v = \sum_{j=1}^n \bar{w}_j v_j$; (2) $\mathcal{P}_n$, the space of polynomials in x having degree n or less and having complex coefficients, with the inner product being $< p, q >= \int_{-1}^1 p(x)\overline{q(x)}dx$; (3) $L^2[0, 2\pi]$, the space of complex-valued, Lebesgue[1] square-integrable functions defined on $[0, 2\pi]$, with the inner product being $< f, g >= \int_0^{2\pi} f(x)\overline{g(x)}dx$. There may be more than one inner product associated with a vector space. For example, for $\mathbb{C}^2$ one may use

$$< v, w >= \begin{pmatrix} \bar{w}_1 & \bar{w}_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

and for $\mathcal{P}_n$ one can also use the inner product $< p, q >= \int_{-\infty}^{\infty} p(x)\overline{q(x)}e^{-x^2} dx$. We will leave as exercises the task of showing that the expressions defined above are really inner products. This is quite easy in the case of $\mathbb{C}^n$ and somewhat more difficult in the other cases.

The inner product space that we will be most concerned with is the one that will be used in connection with signal processing, the space of signals with finite energy, $L^2(\mathbb{R})$. This comprises all Lebesgue square-integrable functions defined on $\mathbb{R}$; its inner product is

$$< f, g >= \int_{-\infty}^{\infty} f(t)\overline{g(t)}dt.$$

An inner product $< \cdot, \cdot >$ provides a vector space $\mathcal{V}$ with a "geometry". The distance between two vectors v and w in $calv$ is $\|v - w\|$, and the two vectors are said to be *orthogonal*

[1]The integral that gets used *theoretically* is the Lebesgue integral. In practice, one thinks in terms of the Riemann integral.

if $< v, w >= 0$. The length of a vector v is $\|v\|$, its distance from 0. If $\|v\| = 1$, then v is said to be a *unit* vector. When $\mathcal{V}$ is a real vector space, then one can also define an angle between two nonzero vectors. Seeing this requires discussing Schwarz's inequality.

PROPOSITION 1.1 (**Schwarz's inequality**): $| < v, w > | \leq \|v\| \, \|w\|$.

PROOF: We will first look at the case in which $w = 0$. If $w = 0$, then by homogeneity $< v, 0 >= 0 < v, 0 >= 0$, and by positivity $\|w\| = 0$. The inequality is thus satisfied. If $w \neq 0$, then let α be any complex number, and consider

$$F(\alpha) := \|v + \alpha w\|^2 \tag{1.1}$$

Note that $F(\alpha) \geq 0$. Using the properties of a complex inner product, it is easy to express $F(\alpha)$ this way:

$$\begin{aligned} \text{(1.2)}\ F(\alpha) =& \|v\|^2 + \bar{\alpha} < v, w > + \alpha < w, v > + |\alpha|^2 \|w\|^2 \\ =& \|w\|^2 \left(\frac{\|v\|^2}{\|w\|^2} + \bar{\alpha} \frac{< v, w >}{\|w\|^2} + \alpha \frac{\overline{< v, w >}}{\|w\|^2} + |\alpha|^2 \right) \ \text{(Property 2)} \\ =& \|w\|^2 \left(\frac{\|v\|^2}{\|w\|^2} - \frac{| < v, w > |^2}{\|w\|^4} + \left| \alpha + \frac{< v, w >}{\|w\|^2} \right|^2 \right) \ \text{(Complete the square.)} \end{aligned}$$

Choose $\alpha = - < v, w > / \|w\|^2$. Since $F(\alpha)$ is always nonngeative, we see that

$$\frac{\|v\|^2}{\|w\|^2} - \frac{| < v, w > |^2}{\|w\|^4} \geq 0.$$

Multiply both sides by $\|w\|^4$ and rearrange terms. Taking the square root of both sides of the result yields Schwarz's inequality. ∎

When $\mathcal{V}$ is a real vector space, Schwarz's inequality allows us to define the angle between two nonzero vectors. To see this, remove absolute value signs in Schwarz's inequality (real case!),

$$-\|v\| \, \|w\| \leq < v, w > \leq \|v\| \, \|w\|.$$

Dividing both sides by $\|v\| \, \|w\|$, we get that

$$-1 \leq \frac{< v, w >}{\|v\| \, \|w\|} \leq 1.$$

The term in the middle is thus in the range of the cosine. We therefore define the angle θ between v and w to be

$$\theta := \text{Arccos} \left(\frac{< v, w >}{\|v\| \, \|w\|} \right). \tag{1.3}$$

Schwarz's inequality also has an important consequence for lengths of vectors. In ordinary Euclidean geometry, the sum of the lengths of any two sides of a triangle exceeds the length of the third side. The same inequality holds for a general inner product space.

Corollary 1.2 (**The Triangle Inequality**): $\big|\|v\| - \|w\|\big| \le \|v+w\| \le \|v\| + \|w\|$.

Proof: We need a special case of a formula that we derived when we proved Schwarz's inequality; namely,

$$\|v+w\|^2 = \|v\|^2 + < v, w > + < w, v > + \|w\|^2,$$

which follows from (1.1) and (1.2) with $\alpha = 1$. From conjugate symmetry we have that $< w, v > = \overline{< v, w >}$. This and the fact that $\Re\{z\} = \frac{1}{2}(z + \bar{z})$ allow us to write the previous equation as

$$\|v+w\|^2 = \|v\|^2 + 2\Re\{< v, w >\} + \|w\|^2, \tag{1.4}$$

Now, $-| < v, w > | \le \Re\{< v, w >\} \le | < v, w > |$. Applying Schwarz's inequality to this gives

$$-\|v\|\,\|w\| \le \Re\{< v, w >\} \le \|v\|\,\|w\|.$$

Combining this with (1.4) results in

$$\big|\|v\| - \|w\|\big|^2 = \|v\|^2 - 2\|v\|\,\|w\| + \|w\|^2 \le \|v+w\|^2 \le \|v\|^2 + 2\|v\|\,\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

Taking square roots then yields the triangle inequality. ∎

In addition to the triangle inequality, the norm $\|\cdot\|$ coming from the inner product $< \cdot, \cdot >$ satisfies two other basic properties: *positivity*, which means that $\|v\| \ge 0$, with $\|v\| = 0$ if and only if $v = 0$; and *positive homegeneity*—$\|cv\| = |c|\,\|v\|$. As we mentioned earlier, an inner product gives a vector space a geometry that has distances, lengths, and angles. The part of the geometry that concerns distances and lengths comes directly from the norm. For many applications, this is the *only* part that is needed or even possible to obtain.

In such applications, one has a norm, but no inner product. A *norm* for a vector space $\mathcal{V}$ is a function $\|\cdot\| : \mathcal{V} \to [0, \infty)$ that satisfies these three properties:

N1. *Positivity:* $\|v\| \ge 0$ for all $v \in \mathcal{V}$, with $\|v\| = 0$ if and only if $v = 0$.

N2. *Positive Homogeneity:* $\|cv\| = |c|\,\|v\|$ for all $v \in \mathcal{V}$ and $c \in \mathbb{C}$.

N3. *The Triangle Inequality:* $\big|\|v\| - \|w\|\big| \le \|v+w\| \le \|v\| + \|w\|$ for all $v, w \in \mathcal{V}$.

A vector space that has a norm $\|\cdot\|$ is called a *normed linear space.* An example of a vector space that has a norm that does *not* come from an inner product is $C[a,b]$, the space of all complex valued functions defined and continuous on the interval $[a,b]$; its norm is defined by

$$\|f\| = \max_{a \le x \le b} |f(x)|.$$

Even the space $\mathbb{C}^n$ can be given a norm that does not come from an inner product. Indeed, it has a whole family of them. Let $v = (\, v_1 \;\; \cdots \;\; v_n \,)^T$ be a column vector in $\mathbb{C}^n$. For $1 \le p < \infty$, define

$$\|v\|_p := \left(\sum_{j=1}^{n} |v_j|^p \right)^{1/p}.$$

This "norms" $\mathbb{C}^n$ for all allowed values of p. In particular, the only norm in this family that can be associated with an inner product is the one for $p = 2$.

§**2. Convergence and completeness.** The spaces that we will be using are infinite dimensional, and so the approximations that we must deal with involve taking limits of sequences. Thus, we will be confronted with the same to questions that we addressed in discussing the Fourier series expansion for the solution to the initial value problem for heat flow in a ring. After we obtained a *formal* series solution, we had to ask whether the series actually defined a function and if it did whether the function solved the heat flow problem.

The property that sets the infinite dimensional spaces that we use here apart from other spaces is that they are *complete*. If $\{v_k\}_{k=1}^{\infty}$ is a sequence of vectors in a normed space $\mathcal{V}$, we say that it is *convergent* if there is a vector v such that for every $\epsilon > 0$ we can find an integer K such that the distance from v_k to v is less than ϵ for all $k > K$; that is,

$$\|v - v_k\| < \epsilon \quad \text{for all } k > K.$$

The difficulty with applying such a definition is that one needs to know in advance that there is such a v. Usually, one only knows enough information to be able to tell that for large k's, the individual members of the sequence are getting close together. For example, look at the sequence of rational numbers defined recursively via

$$q_{k+1} = \frac{1}{2}q_k + \frac{1}{q_k}, \quad q_1 := 2.$$

We do not know—at least at the start—whether this sequence converges or, if it does, to what. We can however show that given $\epsilon > 0$ there is an integer K such $|q_k - q_\ell| < \epsilon$ whenever j and k are greater than K. In words, the q_k's are getting close together for large k's. A sequence with this property is called a *Cauchy sequence*. Is the sequence convergent to a rational number? The answer is *no*. If $q_k \to s$ as $k \to \infty$, then it is easy to show that $s^2 = 2$, so $s = \sqrt{2}$, which is an irrational number. On the other hand, if we ask whether the sequence converges to a real number, the answer is *yes*, and the number is $\sqrt{2}$. The fundamental difference between the reals and the rationals is that in the reals all Cauchy sequences are convergent whereas in the rationals, some are and some are not. The reals are said to be *complete*. In general, a normed linear space $\mathcal{V}$ in which every Cauchy sequence is convergent is also said to be complete. A complete normed linear space is a *Banach space*, and a complete inner product space is called a *Hilbert space*. All of the finite dimesional spaces that we have worked with are complete. The space of signals with finite energy, $L^2(\mathbb{R})$, is a Hilbert space if one uses the inner product defined earlier in §1, and $C[a, b]$, the space of complex-valued functions defined on the closed interval $[a, b]$ is a Banach space, again provided that the norm defined in §1 is used.

§**3. Orthogonal sets.** Let $\mathcal{V}$ be an inner product space, with the inner product being $< \cdot, \cdot >$. A finite or infinite set of vectors $\{v_1, v_2, \ldots\}$ is said to be *orthogonal* if none of its vectors are 0 and if for every pair of distinct vectors—i.e., $v_j \neq v_k$—we have $< v_j , v_k >= 0$. If in addition, the vectors all are normalized so that $\|v_j\| = 1$, then the set is said to be *orthonormal*. The following results are elementary but useful facts; they will be stated without proof.

PROPOSITION 3.1: *Every orthogonal set is linearly independent.*

PROPOSITION 3.2: *Every orthogonal set* $\{v_1, v_2, v_3, \ldots\}$ *can be transformed into an orthonomal set; namely,* $\{u_1 = v_1/\|v_1\|, u_2 = v_2/\|v_2\|, u_3 = v_3/\|v_3\|, \ldots\}$.

PROPOSITION 3.3: *Every linearly independent set can be transformed into an orthonomal set having the same span via the Gram-Schmidt process.*

For future reference, we will now simply list a number of different sets of orthogonal functions, together with the inner product space they belong to.

1. (Fourier) $\{e^{inx}/\sqrt{2\pi}\}_{n=-\infty}^{\infty}$, $\mathcal{V} = L^2[-\pi,\pi]$ and $< f, g >= \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$.

2. (Fourier) $\{\sqrt{\frac{2}{\pi}}\sin(nx)\}_{n=1}^{\infty}$, $\mathcal{V} = L^2[0,\pi]$ and $< f, g >= \int_0^{\pi} f(x)\overline{g(x)}dx$.

3. (Legendre) $\{P_\ell(x)\}_{\ell=0}^{\infty}$, $\mathcal{V} = L^2[-1,1]$ and $< f, g >= \int_{-1}^{1} f(x)\overline{g(x)}dx$, where

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}(x^2-1)^\ell, \quad \|P_\ell\| = \sqrt{\frac{2\ell+1}{2}}.$$

4. (Hermite) $\{H_k(x)e^{-x^2/2}\}_{k=0}^{\infty}$, $\mathcal{V} = L^2(\mathbb{R})$ and $< f, g >= \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx$, where

$$H_k(x) = (-1)^k e^{x^2}\frac{d^k}{dx^k}e^{-x^2}, \quad \|H_k(x)e^{-x^2/2}\| = \sqrt{\sqrt{\pi}2^k k!}$$

5. (Laguerre) $\{L_k(x)e^{-x/2}\}_{k=0}^{\infty}$, $\mathcal{V} = L^2([0,\infty)]$ and $< f, g >= \int_0^{\infty} f(x)\overline{g(x)}dx$,

$$L_k(x) = e^x\frac{d^k}{dx^k}(x^k e^{-x}), \quad \|L_k(x)e^{-x/2}\| = k!$$

6. (Haar) $\{2^{j/2}\psi(2^j x - k)\}_{j,k=-\infty}^{\infty}$, $\mathcal{V} = L^2(\mathbb{R})$ and $< f, g >= \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx$, where $\psi(x)$ is given by

$$\psi(x) := \begin{cases} 1 & \text{if } 0 \le x < 1/2 \\ -1 & \text{if } 1/2 \le x < 1 \\ 0 & \text{if } x < 0 \text{ or } x \ge 1. \end{cases}$$

The first two examples are of course the familiar sets used for Fourier series and Fourier sine series; the functions in both have been normalized so that they form orthonormal sets. Examples 3-5 all involve orthogonal polynomials that are important in quantum mechanics: the Legendre polynomials come up in connection with angular momentum; the Hermite polynomials arise in computing eigenstates for the harmonic oscillator; and, the Laguerre polynomials are associated with eigenstates of the hyrogen atom. The final example is the set constructed by Haar from scalings and translations of the function ψ. These are the Haar wavelets.

Just as we can expand a function in a Fourier series, we can expand a vector in a series of orthogonal vectors, provided the set contains "enough" vectors. In finite dimensional spaces, "enough" just means that the orthogonal set has as many vectors as the dimension of the space. With infinite dimensional spaces, the situation is a little more complicated. The approach that we take here will give us not only conditions for there to be "enough" vectors in an orthogonal set, but it will also give us the concept of a least-squares projection.

§**4. Least-squares approximation and orthogonal expansions.** The term "least-squares" comes from the kind of minimization done in statistics. It is also used in a function space setting. Here, we will do the following. Let $S = \{v_1, v_2, \ldots, v_n, \ldots\}$ be an orthogonal set of vectors in an inner product space $\mathcal{V}$. The problem that we want to look at is the following:

> Among all linear combinations from the finite set $S_N = \{v_1, v_2, \ldots, v_N\}$, which will be closest to a given vector $v \in \mathcal{V}$—where "close" is measured in terms of the norm for $\mathcal{V}$?

What we must do to solve this problem is to find the coefficients $c_1, \ldots, c_N$ that minimize the quantity, $\|v - \sum_{j=1}^N c_j v_j\|^2$, which is the square of the distance between v and $\sum_{j=1}^N c_j v_j$. First, use the properties of the inner product and the orthogonalty of the v_j's to manipulate this quantity into the form:

$$\begin{aligned}\|v - \sum_{j=1}^N c_j v_j\|^2 &= \|v\|^2 - \sum_{j=1}^N (c_j < v_j, v > + \bar{c}_j < v, v_j >) + \sum_{j=1}^N \sum_{k=1}^N c_j \bar{c}_k < v_j, v_k > \\ &= \|v\|^2 - \sum_{j=1}^N (c_j < v_j, v > + \bar{c}_j < v, v_j >) + \sum_{j=1}^N |c_j|^2 \|v_j\|^2 \\ &= \|v\|^2 - \sum_{j=1}^N (c_j \overline{< v, v_j >} + \bar{c}_j < v, v_j >) + \sum_{j=1}^N |c_j|^2 \|v_j\|^2 \\ &= \|v\|^2 - \sum_{j=1}^N \frac{| < v, v_j > |^2}{\|v_j\|^2} + \sum_{j=1}^N \|v_j\|^2 \left| c_j - \frac{< v, v_j >}{\|v_j\|^2} \right|^2,\end{aligned}$$

where the last equation follows from the previous one by completing the square. Second, obverve that on the right hand side of the last equation, the only term that depends on the c_j's is the last one, which is nonnegative. Choosing $c_j = \frac{<v,v_j>}{\|v_j\|^2}$ thus minimizes $\|v - \sum_{j=1}^N c_j v_j\|^2$. We have proved the following.

THEOREM 4.1: *The distance between v and an arbitrary vector $\sum_{j=1}^N c_j v_j$ in the span of $\{v_1, v_2, \ldots, v_N\}$ is minimized when $c_j = \frac{<v,v_j>}{\|v_j\|^2}$. Indeed, one always has*

$$\|v - \sum_{j=1}^N c_j v_j\|^2 \geq \|v - \sum_{j=1}^N \left(\frac{< v, v_j >}{\|v_j\|^2} \right) v_j\|^2 = \|v\|^2 - \sum_{j=1}^N \frac{| < v, v_j > |^2}{\|v_j\|^2}. \tag{4.1}$$

The vector $\sum_{j=1}^N \left(\frac{<v,v_j>}{\|v_j\|^2} \right) v_j$ is the least squares approximant to v; it is also called the *orthogonal projection* of v onto $\mathrm{Span}\{v_1, \ldots, v_N\}$. It is worth noting that the vector v

can be written as the sum

$$v = \left(v - \sum_{j=1}^{N} \left(\frac{< v, v_j >}{\|v_j\|^2} \right) v_j) \right) + \left(\sum_{j=1}^{N} \left(\frac{< v, v_j >}{\|v_j\|^2} \right) v_j \right),$$

where the vectors in parentheses are orthogonal to one another.

The theorem proved above gives us a general version of an inequality that we met earlier in connection with mean convergence of Fourier series.

COROLLARY 4.2 (**Bessel's Inequality**):

$$\|v\|^2 \geq \sum_{j=1}^{\infty} \frac{|< v, v_j >|^2}{\|v_j\|^2}. \tag{4.2}$$

PROOF: From (4.1), we see that

$$\|v - \sum_{j=1}^{N} \left(\frac{< v, v_j >}{\|v_j\|^2} \right) v_j\|^2 = \|v\|^2 - \sum_{j=1}^{N} \frac{|< v, v_j >|^2}{\|v_j\|^2} \geq 0. \tag{4.3}$$

Consequently, we have that for all N

$$\|v\|^2 \geq \sum_{j=1}^{N} \frac{|< v, v_j >|^2}{\|v_j\|^2}$$

Thus the partial sums of the series on the right are bounded from above. Since these partial sums form a nondecreasing sequnece, they converge, and of course so does the infinite series. Taking limits in the last equation then yields Bessel's inequality. ∎

We can now answer the question raised at the end of the last section: When does an orthogonal set have "enough" vectors in it to be a basis?

DEFINITION 4.3: An orthogonal set $\{v_j\}_{j=1}^{\infty}$ is said to be *complete*[2] if every $v \in \mathcal{V}$ can be expressed as

$$v = \sum_{j=1}^{\infty} \left(\frac{< v, v_j >}{\|v_j\|^2} \right) v_j ,$$

where the series converges in the sense that $\|v - \sum_{j=1}^{N} \left(\frac{<v, v_j>}{\|v_j\|^2} \right) v_j\|^2 \to 0$ as $N \to \infty$.

COROLLARY 4.4: *An orthogonal set $\{v_j\}_{j=1}^{\infty}$ is complete if and only if for every $v \in \mathcal{V}$ Bessel's inequality is actually an equation; that is,*

$$\|v\|^2 = \sum_{j=1}^{\infty} \frac{|< v, v_j >|^2}{\|v_j\|^2}. \tag{4.4}$$

[2]This is *not* related to an inner product space being complete. There are two distinct concepts here, both with the same name.

Proof: Take the limit as $N \to \infty$ on both sides in equation (4.3). The result is

$$\begin{aligned}\lim_{N\to\infty} \|v - \sum_{j=1}^{N} \left(\frac{< v, v_j >}{\|v_j\|^2} \right) v_j\|^2 &= \|v\|^2 - \lim_{N\to\infty} \sum_{j=1}^{N} \frac{|< v, v_j >|^2}{\|v_j\|^2} \\ &= \|v\|^2 - \sum_{j=1}^{\infty} \frac{|< v, v_j >|^2}{\|v_j\|^2} \end{aligned} \tag{4.5}$$

In view of (4.5), $\lim_{N\to\infty} \|v - \sum_{j=1}^{N} \left(\frac{<v,v_j>}{\|v_j\|^2} \right) v_j\|^2 = 0$ for all v if and only if (4.4) holds for all v; thus, the orthogonal set is complete if and only if (4.4) holds for all v. ∎

Proving that an orthogonal set of functions is complete is generally a task that is beyond the scope of this work. All of the sets of functions listed in §3 are complete.

5. Subspaces. A *subspace* of vector space $\mathcal{V}$ is a subset of $\mathcal{V}$ that is nonempty and that is a vector space in its own right under the operations + and · from $\mathcal{V}$. The following is a standard result from linear algebra; it is quite useful in detemining whether a sub*set* is a sub*space*.

Propositon 5.1: *A subset $\mathcal{U} \subset \mathcal{V}$ is a subspace of $\mathcal{V}$ if and only if*

(a) $0 \in \mathcal{U}$ *($\mathcal{U}$ is nonempty).*
(b) *$\mathcal{U}$ is closed under vector addition, "+". ($u_1 \in \mathcal{U}, u_2 \in \mathcal{U} \Rightarrow u_1 + u_2 \in \mathcal{U}$).*
(c) *$\mathcal{U}$ is closed under scalar mutiplication, "·". ($c \in \mathbb{C}, u \in \mathcal{U} \Rightarrow c \cdot u \in \mathcal{U}$.)*

It is important to be able to combine subspaces to get new subspaces. Here are three ways of doing this; we will give a fourth later on in the section. Let $\mathcal{U}_1, \mathcal{U}_2$ be subspaces of a vector space $\mathcal{V}$.

(i) **Intersection.** $\mathcal{U}_1 \cap \mathcal{U}_2$ is a subspace. (But $\mathcal{U}_1 \cup \mathcal{U}_2$ is *not* a subspace, at least not in general.)

(ii) **Sum.** $\mathcal{U}_1 + \mathcal{U}_2 := \{u_1 + u_2 \in \mathcal{V} \,|\, u_1 \in \mathcal{U}_1, \text{ and } u_2 \in \mathcal{U}_2\}$

(iii) **Direct Sum.** We will say the sum $\mathcal{U}_1 + \mathcal{U}_2$ is *direct* if for every $v \in \mathcal{U}_1 + \mathcal{U}_2$, the vectors u_1, u_2 for which $v = u_1 + u_2$ are *unique*. We write $\mathcal{U}_1 \oplus \mathcal{U}_2$.

Before we can discuss the fourth way of combining subspaces to get another subspace, we need to talk about *orthogonal subspaces*. This concept plays a crucial role in the theory and application of wavelets. Again, let $\mathcal{U}_1$ and $\mathcal{U}_2$ be subspaces of a vector space $\mathcal{V}$, and also let $\mathcal{V}$ have the inner product $< \cdot, \cdot >$. $\mathcal{U}_1$ and $\mathcal{U}_2$ are said to be *orthogonal* if for every pair $u_1 \in \mathcal{U}_1$ and $u_2 \in \mathcal{U}_2$,

$$\langle u_1, u_2 \rangle = 0.$$

When this happens, we write $\mathcal{U}_1 \perp \mathcal{U}_2$.

Examples:

1. In $\mathcal{V} = \mathbb{R}^3$, let $\mathcal{U}_1 = \text{Span } \{\vec{i} + \vec{j} - \vec{k}, \vec{i} + \vec{j}\}$ and $\mathcal{U}_2 = \text{Span } \{\vec{i} - \vec{j}\}$. Clearly, $\mathcal{U}_1 \perp \mathcal{U}_2$.

2. For Fourier series, if we let

$$\mathcal{U}_+ = \text{Span}\,\{e^{ix}, e^{2ix}, \ldots\} \quad \text{and} \quad \mathcal{U}_- = \text{Span}\,\{e^{-ix}, e^{-2xi}, \ldots\},$$

we see that $\mathcal{U}_+$ and $\mathcal{U}_-$ are orthogonal; that is, $\mathcal{U}_+ \perp \mathcal{U}_-$.

PROPOSITION 5.2: *If $\mathcal{U}_1$ and $\mathcal{U}_2$ are orthogonal subspaces, then the sum $\mathcal{U}_1 + \mathcal{U}_2$ is direct. That is,*

$$\mathcal{U}_1 + \mathcal{U}_2 = \mathcal{U}_1 \oplus \mathcal{U}_2.$$

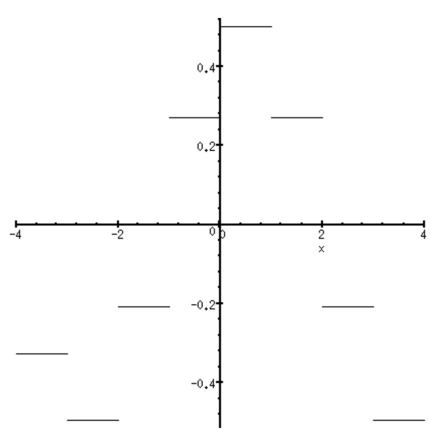
PROOF: If v is in $\mathcal{U}_1 + \mathcal{U}_2$, then we can decompose v into a sum, $v = u_1 + u_2$, with $u_1 \in \mathcal{U}_1$ and $u_2 \in \mathcal{U}_2$. Suppose that we also have $v = u_1' + u_2'$, again with $u_1' \in \mathcal{U}_1$ and $u_2' \in \mathcal{U}_2$. Equating the two expressions for v gives $u_1 + u_2 = u_1' + u_2'$, from which we see that $\mathcal{U}_1 \ni u_1 - u_1' = u_2' - u_2 \in \mathcal{U}_2$. Thus, $u_1 - u_1'$ is in both $\mathcal{U}_1$ and $\mathcal{U}_2$. Since these spaces are orthogonal, we have that $0 =< u_1 - u_1', u_1 - u_1' >= \|u_1 - u_1'\|^2$. From the properties of inner products, we get $u_1 - u_1' = 0$, so $u_1 = u_1'$. Similar reasoning gives us that $u_2 = u_2'$. The decomposition of v into u_1 and u_2 is therefore unique, and the sum is direct. ∎

Orthogonal complements. We can now discuss the fourth way of combining subspaces. Again, let $\mathcal{V}$ be an inner product space, with $< \cdot, \cdot >$ being the inner product. Keep in mind that $\mathcal{V}$ may be a subpsce of some larger space. If we have a subspace $\mathcal{U} \subset \mathcal{V}$, there are many subspaces for which $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. From the point of view of multiresolution analysis, the most important is the orthogonal complement of $\mathcal{U}$ relative to $\mathcal{V}$.

$$\mathcal{V} \overset{\perp}{\ominus} \mathcal{U} := \{w \in \mathcal{V} : \langle w, u \rangle = 0 \text{ for all } u \in \mathcal{U}\}.$$

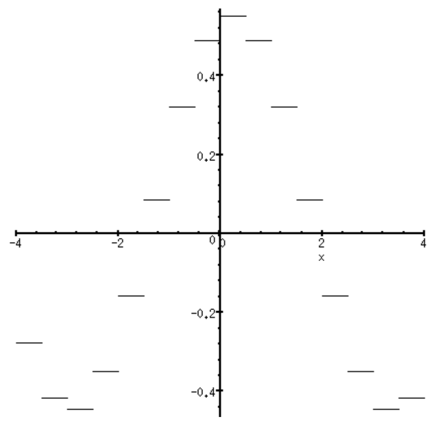
That is, $\mathcal{V} \overset{\perp}{\ominus} \mathcal{U}$ comprises all the vectors in $\mathcal{V}$ that are perpendicular to *every* vector in $\mathcal{U}$; it amounts to a "subtraction" of vector spaces. As we mentioned above, there are many spaces for which $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. One of them is $\mathcal{W} = \mathcal{V} \overset{\perp}{\ominus} \mathcal{U}$. Indeed, since this space is very nearly the only one that we will be concerned with, with will use the slightly simpler $\mathcal{V} \ominus \mathcal{U}$ to denote it.

An example. Let $\mathcal{V}_0$ be all square integrable functions that are constant in the intervals $[k/, k+1), k = 0, \pm 1, \pm 2, \cdots$.



Typical function in $\mathcal{V}_0$.

Also, let $\mathcal{V}_1$ be all square integrable functions that are constant in the intervals $[k/2, (k+1)/2)$, $k = 0, \pm 1, \pm 1, \cdots$.



Typical function in $\mathcal{V}_1$.

We leave it as an exercise to show that the orhogonal complement of $\mathcal{V}_0$ relative to $\mathcal{V}_1$ is the vector space

$$\mathcal{W}_0 = \mathcal{V}_1 \ominus \mathcal{V}_0 = \{g \in \mathcal{V}_1 \,|\, g(k) + g((k+1)/2) = 0 \text{ for all } k = 0, \pm 1, \pm 2, \ldots\}$$

This of course is the 0-scale wavelet space. The function ψ that generates the orthonormal wavelet basis $\{\psi(x-k)\}_{k\in\mathbb{Z}}$ is the function

$$\psi(x) := \begin{cases} 1 & \text{if } 0 \leq x < 1/2 \\ -1 & \text{if } 1/2 \leq x < 1 \\ 0 & \text{if } x < 0 \text{ or } x \geq 1, \end{cases}$$

which is the Haar wavelet mentioned earlier.

1