MATH 172: INNER PRODUCT SPACES, SYMMETRIC OPERATORS, ORTHOGONALITY

ANDRÁS VASY

When discussing separation of variables, we noted that at the last step we need to express the inhomogeneous initial or boundary data as a superposition of functions arising in the process of separation of variables. For instance, for the Dirichlet problem $\Delta u = 0$, $u|_{\partial D} = h$, we had to express h as

$$h(\theta) = A_0 + \sum_{n=1}^{\infty} R^n (A_n \cos(n\theta) + B_n \sin(n\theta)), \tag{1}$$

i.e. we had to find constants A_n and B_n so that this expression holds. The basic framework for this is inner product spaces, which we now discuss.

Definition 1. An inner product on a complex vector space V is a map

$$\langle .,. \rangle : V \times V \to \mathbb{C}$$

such that

(i) $\langle .,. \rangle$ is linear in the first slot:

$$\langle c_1 v_1 + c_2 v_2, w \rangle = c_1 \langle v_1, w \rangle + c_2 \langle v_2, w \rangle,\ c_1, c_2 \in \mathbb{C},\ v_1, v_2, w \in V,$$

(ii) $\langle .,. \rangle$ is Hermitian symmetric:

$$\langle v, w \rangle = \overline{\langle w, v \rangle},$$

with the bar denoting complex conjugate,

(iii) $\langle .,. \rangle$ is positive definite:

$$v \in V \Rightarrow \langle v, v \rangle \geq 0, \text{ and } \langle v, v \rangle = 0 \Leftrightarrow v = 0.$$

A vector space with an inner product is also called an inner product space.

While one should write $(V, \langle .., . \rangle)$ to specify the inner product space, one typically says merely that V is an inner product space when the inner product is understood.

For real vector spaces, one makes essentially the same definition, except that, as the complex conjugate does not make sense, one simply has symmetry:

$$V \text{ real vector space } \Rightarrow \langle v, w \rangle = \langle w, v \rangle,\ v, w \in V.$$

We also introduce the notation for the *norm* associated to this inner product:

$$\|v\| = \langle v, v \rangle^{1/2},$$

where the square root is the unique non-negative square root of a non-negative number (see (iii)). Thus,

$$\langle v, v \rangle = \|v\|^2.$$

Recall that in general a norm is defined by:

Definition 2. Suppose V is a vector space. A norm on V is a map

$$\|.\| : V \to \mathbb{R}$$

such that

(i) (positive definiteness) $\|v\| \geq 0$ for all $v \in V$, and $v = 0$ if and only if $\|v\| = 0$.
(ii) (absolute homogeneity) $\|cv\| = |c|\,\|v\|$, $v \in V$, and c a scalar (so $c \in \mathbb{R}$ or $c \in \mathbb{C}$, depending on whether V is real or complex),
(iii) (triangle inequality) $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.

Thus, at this point, we do not yet know that our 'norm' for the inner product is a norm in this vector space sense, though it satisfies (i) by property (iii) of the inner product, and (ii) by properties (i) and (ii) (see (2) below), so the question is whether the triangle inequality holds. We show this shortly, but first we note some examples of inner product spaces:

(i) $V = \mathbb{R}^n$, with inner product
$$\langle x, y\rangle = \sum_{j=1}^n x_j y_j,$$
where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. Thus $\|x\|^2 = \sum_{j=1}^n x_j^2$.
(ii) $V = \mathbb{C}^n$, with inner product
$$\langle x, y\rangle = \sum_{j=1}^n x_j \overline{y_j},$$
where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. Thus $\|x\|^2 = \sum_{j=1}^n |x_j|^2$, which explains why we need Hermitian symmetry for complex vector spaces.
(iii) $V = \mathbb{R}^n$, with inner product
$$\langle x, y\rangle = \sum_{j=1}^n a_j x_j y_j,$$
where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, $a_j > 0$ for all j. Thus $\|x\|^2 = \sum_{j=1}^n a_j x_j^2$.
(iv) $V = C^0(\overline{\Omega})$ (complex valued continuous functions on the closure of Ω), where Ω is a bounded domain in $\mathbb{R}^n$, with inner product
$$\langle f, g\rangle = \int_\Omega f(x)\,\overline{g(x)}\,dx.$$
Thus,
$$\|f\|^2 = \int_\Omega |f(x)|^2\,dx.$$
We often write $\|f\|_{L^2} = \|f\|_{L^2(\Omega)}$ for this norm.
(v) $V = C^0(\overline{\Omega})$, Ω as above, with inner product
$$\langle f, g\rangle = \int_\Omega f(x)\,\overline{g(x)}\,a(x)\,dx,$$
where $a \in C^0(\overline{\Omega})$, $a > 0$ is fixed. Thus,
$$\|f\|^2 = \int_\Omega |f(x)|^2\,a(x)\,dx.$$
We may write $\|f\|_{L^2(\Omega, a(x)\,dx)}$ for this norm.
(vi) Let $N > n/2$, let $V = \{f \in C^0(\mathbb{R}^n) :\ (1 + |x|)^N f \text{ is bounded}\}$, with the inner product
$$\langle f, g\rangle = \int_{\mathbb{R}^n} f(x)\,\overline{g(x)}\,dx.$$

Note that the integral converges under this decay assumption. Thus,

$$\|f\|^2 = \int_{\mathbb{R}^n} |f(x)|^2 \, dx.$$

We often write $\|f\|_{L^2} = \|f\|_{L^2(\mathbb{R}^n)}$ for this norm, in accordance with (vii) below.

(vii) $V = L^2(\mathbb{R}^n) = \{f$ measurable on $\mathbb{R}^n : \int |f|^2 < \infty\}/\sim$ where $\sim$ is the equivalence relation by which two measurable functions are equivalent if they differ on a set of measure 0, with inner product

$$\langle f, g\rangle = \int_{\mathbb{R}^n} f\overline{g},$$

so $\|f\|^2 = \int |f|^2$. Notice that (vi) gives a subspace of this, on which the inner product is the restriction of the L^2-inner product.

First, note that $L^2(\mathbb{R}^n)$ is indeed a vector space, namely (as multiplication by constants certainly gives another element of V) $f, g \in V$ implies $f + g \in V$ (note that $f + g$ is well-defined a.e. since $\int |f|^2 < \infty$ implies that f is a.e. finite, and similarly with g, and changing both f and g on a set of measure 0 changes $f + g$ on a set of measure 0 as well). To see this, one just needs to note that for all x with $f(x), g(x)$ finite,

$$|(f(x)+g(x))|^2 \le (|f(x)|+|g(x)|)^2 = |f(x)|^2+|g(x)|^2+2|f(x)||g(x)| \le 2(|f(x)|^2+|g(x)|^2),$$

so $\int |f + g|^2 \le 2\int |f|^2 + 2\int |g|^2$, giving $f + g \in L^2$.

That $\langle \cdot, \cdot\rangle$ is well defined follows from f, g a.e. finite since $\int |f|^2, \int |g|^2 < \infty$, and $|f(x)\overline{g(x)}| \le \frac{|f(x)|^2+|g(x)|^2}{2}$ for all x for which $f(x)$, $g(x)$ finite, i.e. on the complement of a set of measure 0, so by the monotonicity of the integral $|fg|$ is integrable, so fg is integrable. Further, changing f, g on sets of measure 0 only changes the product on a set of measure 0, thus leaves the integral unchanged, so it is indeed well-defined independent of the choice of the representative of the equivalence class.

In addition, $\|f\|^2 = 0$ means $\int |f|^2 = 0$, thus $f(x) = 0$ a.e. x, i.e. f is in the equivalence class of the zero function. All other properties required in the definition of an inner product follow from the basic properties of $\mathbb{C}$, and the linearity and monotonicity of the integral.

(viii) $\Omega \subset \mathbb{R}^n$ measurable, $V = L^2(\Omega) = \{f$ measurable on $\Omega : \int_\Omega |f|^2 < \infty\}/\sim$, $\sim$ as in (vii), with inner product

$$\langle f, g\rangle = \int_\Omega f\overline{g},$$

so $\|f\|^2 = \int_\Omega |f|^2$. Notice that (iv) is a subset of this, on which the inner product is the restriction of the $L^2(\Omega)$-inner product. (This is the trivial zero dimensional vector space if the measure of Ω is 0!)

A few properties on inner products should be observed immediately. First, the inner product is conjugate-linear in the second variable (which simply means linear if the vector space is real):

$$\begin{aligned}\langle v, c_1w_1 + c_2w_2\rangle &= \overline{\langle c_1w_1 + c_2w_2, v\rangle} = \overline{c_1\langle w_1, v\rangle + c_2\langle w_2, v\rangle} \\ &= \overline{c_1}\,\overline{\langle w_1, v\rangle} + \overline{c_2}\,\overline{\langle w_2, v\rangle} = \overline{c_1}\langle v, w_1\rangle + \overline{c_2}\langle v, w_2\rangle.\end{aligned}$$

A map $V \times V \to \mathbb{C}$ which is linear in the first variable and conjugate-linear in the second variable is called *sesquilinear*. Second, by (ii), if $v \in V$ then $\langle v, v\rangle = \overline{\langle v, v\rangle}$, so $\langle v, v\rangle$ is real. Thus, (iii) is the statement that this real number is non-negative, and it is actually positive if $v \neq 0$. Also, by linearity, denoting the 0 vector in V by 0_V (usually we simply denote this by 0, but here this care clarifies the calculation),

$$\langle 0_V, v\rangle = \langle 0 \cdot 0_V, v\rangle = 0\langle 0_V, v\rangle = 0,\ \ v \in V,$$

and by Hermitian symmetry then

$$\langle v, 0_V\rangle = \overline{\langle 0_V, v\rangle} = 0.$$

We also note a useful properties of $\|.\|$:

$$\|cv\|^2 = \langle cv, cv\rangle = c\langle v, cv\rangle = c\bar{c}\langle v, v\rangle = |c|^2\|v\|^2,\ \ c \in \mathbb{C},\ \ v \in V, \tag{2}$$

so

$$\|cv\| = |c|\,\|v\|,$$

proving the absolute homogeneity of the 'norm'.

One concept that is tremendously useful in inner product spaces is orthogonality:

Definition 3. Suppose V is an inner product space. For $v, w \in V$ we say that v is orthogonal to w if $\langle v, w\rangle = 0$.

Note that $\langle v, w\rangle = 0$ if and only of $\langle w, v\rangle = 0$, so v is orthogonal to w if and only if w is orthogonal to v – so we often say simply that v and w are orthogonal.

As an illustration of its use, let's prove Pythagoras' theorem:

Lemma 0.1. *Suppose V is an inner product space, $v, w \in V$ and v and w are orthogonal. Then*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 = \|v - w\|^2.$$

Proof. Since $v-w = v+(-w)$, the statement about $v-w$ follows from the statement for $v + w$ and $\| - w\| = \|w\|$. Now,

$$\langle v+w, v+w\rangle = \langle v, v+w\rangle+\langle w, v+w\rangle = \langle v, v\rangle+\langle v, w\rangle+\langle w, v\rangle+\langle w, w\rangle = \langle v, v\rangle+\langle w, w\rangle$$

by the orthogonality of v and w, proving the result. □

One use of orthogonality is the following:

Lemma 0.2. *Suppose $v, w \in V$, $w \neq 0$. Then there exist unique $v_\parallel, v_\perp \in V$ such that $v = v_\parallel + v_\perp$, $v_\parallel = cw$ for some $c \in \mathbb{C}$ and $\langle v_\perp, w\rangle = 0$.*

Proof. If $v = v_\parallel + v_\perp$ then taking the inner product with w and using $v_\parallel = cw$ we deduce

$$\langle v, w\rangle = \langle cw, w\rangle + \langle v_\perp, w\rangle = c\|w\|^2,$$

so as $w \neq 0$,

$$c = \frac{\langle v, w\rangle}{\|w\|^2}.$$

Thus, $v_\parallel = cw$ and $v_\perp = v - cw$, giving uniqueness.

On the other hand, if we *let*

$$c = \frac{\langle v, w\rangle}{\|w\|^2},\ \ v_\parallel = cw,\ \ v_\perp = v - cw,$$

then $v_\perp + v_\parallel = v$ and $v_\parallel = cw$ are satisfied, so we merely need to check $\langle v_\perp, w\rangle = 0$. But

$$\langle v_\perp, w\rangle = \langle v, w\rangle - c\langle w, w\rangle = \langle v, w\rangle - \frac{\langle v, w\rangle}{\|w\|^2}\|w\|^2 = 0,$$

so the desired vectors $v_\perp$ and $v_\parallel$ indeed exist. □

One calls $v_\parallel$ the *orthogonal projection* of v to the span of w.

The final piece we need to show that our 'norm' is indeed a norm is the triangle inequality. This is achieved by the Cauchy-Schwarz inequality:

Lemma 0.3. *In an inner product space V,*

$$|\langle v, w\rangle| \leq \|v\|\|w\|,\ \ v, w \in V.$$

Proof. If $w = 0$, then both sides vanish, so we may assume $w \neq 0$. Write $v = v_{\parallel} + v_{\perp}$ as in Lemma 0.2, so

$$v_{\parallel} = cw,\ \ c = \frac{\langle v, w\rangle}{\|w\|^2}.$$

Then by Pythagoras' theorem, using $\langle v_{\parallel}, v_{\perp}\rangle = c\langle w, v_{\perp}\rangle = 0$,

$$\|v\|^2 = \|v_{\parallel}\|^2 + \|v_{\perp}\|^2 \geq \|v_{\parallel}\|^2 = |c|^2\|w\|^2 = \frac{|\langle v, w\rangle|^2}{\|w\|^2}.$$

Multiplying through by $\|w\|^2$ and taking the non-negative square root completes the proof of the lemma. □

A useful consequence of the Cauchy-Schwarz inequality is the triangle inequality for the norm:

Lemma 0.4. *In an inner product space V,*

$$\|v + w\| \leq \|v\| + \|w\|.$$

Proof. One only needs to prove the equivalent estimate where one takes the square of both sides:

$$\|v + w\|^2 \leq \|v\|^2 + 2\|v\|\,\|w\| + \|w\|^2.$$

But

$$\|v+w\|^2 = \langle v+w, v+w\rangle = \|v\|^2 + \langle v, w\rangle + \langle w, v\rangle + \|w\|^2 \leq \|v\|^2 + 2\|v\|\,\|w\| + \|w\|^2,$$

where the last inequality is the Cauchy-Schwarz inequality. □

As a consequence:

Corollary 0.5. *If V is an inner product space, $\|v\| = \langle v, v\rangle^{1/2}$ defines a norm on V. Thus $d(v, w) = \|v - w\|$ is a metric on V.*

In general, when talking about an inner product space, one always takes the norm and the metric to be those in these corollary, unless otherwise specified. In particular, convergence, etc., are with respect to this metric.

Some examples of a normed vector space where the norm does not come from an inner product are:

(i) $V = C^0(\overline{\Omega})$ (complex valued continuous functions on the closure of Ω), where Ω is a bounded domain in $\mathbb{R}^n$, and

$$\|f\|_{C^0} = \sup_{x\in\overline{\Omega}} |f(x)|;$$

(ii) $V = L^1(\mathbb{R}^n)$ with norm

$$\|f\|_{L^1} = \int |f|.$$

Note that in a normed vector space, with convergence defined as for inner product spaces, the vector space operations are (jointly) continuous:

Lemma 0.6. *If V is a normed vector space, $v_j \to v$, $w_j \to w$ in V and $c_j \to c$ in the scalars, then $c_j v_j \to cv$ and $v_j + w_j \to v + w$ in V.*

This lemma states that the vector space operations

$$+ : V \times V \to V, \cdot : \mathbb{C} \times V \to V \text{ or } \cdot : \mathbb{R} \times V \to V$$

are continuous, when one equips the product spaces with the product metric, so e.g. on $V \times V$ one has $d_{V\times V}((v, v'), (w, w')) = d_V(v, w) + d_V(v', w')$.

Proof. Consider the sequence $\{v_j + w_j\}_{j=1}^{\infty}$ first. Then

$$\|(v_j + w_j) - (v + w)\| = \|(v_j - v) + (w_j - w)\| \le \|v_j - v\| + \|w_j - w\| \to 0$$

as $v_j \to v$, $w_j \to w$. Similarly,

$$\begin{aligned}\|c_j v_j - cv\| &= \|(c_j v_j - c_j v) + (c_j v - cv)\| \le \|c_j v_j - c_j v\| + \|c_j v - cv\| \\ &= \|c_j(v_j - v)\| + \|(c_j - c)v\| = |c_j|\,\|v_j - v\| + |c_j - c|\,\|v\| \to 0\end{aligned}$$

since convergent sequences in $\mathbb{R}$ or $\mathbb{C}$ are bounded. □

We used in the proof that convergent sequences in the scalars are bounded; this is also true for convergent sequences in a normed vector space V (indeed, in any metric space!). Namely, if $v_j \to v$ in V, then for j sufficiently large, $\|v_j - v\| \le 1$, hence

$$\|v_j\| = \|(v_j - v) + v\| \le \|v_j - v\| + \|v\| \le \|v\| + 1.$$

Thus, for all but finitely many j, $\|v_j\| \le \|v\| + 1$, hence the sequence $\{\|v_j\|\}_{j=1}^{\infty}$ is bounded.

We can now show that the inner product is jointly continuous with respect to the notion of convergence we discussed:

Lemma 0.7. *Suppose that V is an inner product space. If $v_j, w_j \in V$, and $v_j \to v$, $w_j \to w$ in V then $\langle v_j, w_j\rangle \to \langle v, w\rangle$. Thus,*

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$$

is continuous.

Proof. We have

$$\begin{aligned}\langle v_j, w_j\rangle - \langle v, w\rangle &= \langle v_j, w_j\rangle - \langle v_j, w\rangle + \langle v_j, w\rangle - \langle v, w\rangle \\ &= \langle v_j, w_j - w\rangle + \langle v_j - v, w\rangle.\end{aligned}$$

Thus,

$$|\langle v_j, w_j\rangle - \langle v, w\rangle| \le |\langle v_j, w_j - w\rangle| + |\langle v_j - v, w\rangle| \le \|v_j\|\,\|w_j - w\| + \|v_j - v\|\,\|w\| \to 0$$

since $\{\|v_j\|\}_{j=1}^{\infty}$ is bounded. □

Suppose now that we have a sequence of orthogonal non-zero vectors $\{x_j\}_{j=1}^{\infty}$ in an inner product space V, and suppose that

$$v = \sum_{j=1}^{\infty} c_j x_j,$$

i.e. this sum converges to v, which, recall, means that with $v_N = \sum_{j=1}^{N} c_j x_j$, $v_N \to v$ as $N \to \infty$. Then, by the continuity of the inner product, for any w,

$$\langle v, w\rangle = \sum_{j=1}^{\infty} c_j \langle x_j, w\rangle.$$

Applying this with $w = x_k$, all inner products on the right hand side but the one with $j = k$ vanish, and we deduce that

$$\langle v, x_k\rangle = c_k \|x_k\|^2,$$

hence

$$c_k = \frac{\langle v, x_k \rangle}{\|x_k\|^2}. \tag{3}$$

As these orthogonal collections of vectors are so useful, we make a definition.

Definition 4. An *orthogonal set* S of vectors in an inner product space V is a subset of V consisting of non-zero vectors such that if $x, x' \in S$ and $x \neq x'$ then x is orthogonal to x'.

An orthogonal set is called *orthonormal* if all of its elements have norm 1.

Now, one could check easily by an explicit computation that the functions we have considered, in terms of which we would like to express our initial or boundary data, are orthogonal to each other. Thus, *if* we can write our data as an infinite linear combination of these functions, then the coefficients can be determined easily. For instance, if we want to write

$$h(\theta) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta),$$

where we consider the inner product on functions on $\mathbb{S}^1$ to be given by

$$\langle f, g \rangle = \int_0^{2\pi} f(\theta)\,\overline{g(\theta)}\,d\theta,$$

then we must have

$$\begin{aligned}
a_0 &= \frac{\langle h, 1 \rangle}{\|1\|^2} = \frac{\int_0^{2\pi} h(\theta)\,d\theta}{\int_0^{2\pi} 1\,d\theta} = \frac{1}{2\pi}\int_0^{2\pi} h(\theta)\,d\theta,\\
a_n &= \frac{\int_0^{2\pi} h(\theta)\,\cos(n\theta)\,d\theta}{\int_0^{2\pi} \cos^2(n\theta)\,d\theta} = \frac{1}{\pi}\int_0^{2\pi} h(\theta)\,\cos(n\theta)\,d\theta,\\
b_n &= \frac{\int_0^{2\pi} h(\theta)\,\sin(n\theta)\,d\theta}{\int_0^{2\pi} \sin^2(n\theta)\,d\theta} = \frac{1}{\pi}\int_0^{2\pi} h(\theta)\,\sin(n\theta)\,d\theta.
\end{aligned}$$

Here we used the computation $\cos(2n\theta) = 2\cos^2(n\theta) - 1$, so $\cos^2(n\theta) = \frac{1}{2}(\cos(2n\theta) + 1)$, and thus

$$\int_0^{2\pi} \cos^2(n\theta)\,d\theta = \frac{1}{2}\int_0^{2\pi} (\cos(2n\theta) + 1)\,d\theta = \frac{1}{4n}\sin(2n\theta)|_0^{2\pi} + \frac{1}{2}\theta|_0^{2\pi} = \pi,$$

with a similar result with cosine replaced by sine. In particular, returning to (1), we deduce that

$$\begin{aligned}
A_0 &= \frac{1}{2\pi}\int_0^{2\pi} h(\theta)\,d\theta,\\
A_n &= \frac{1}{\pi R^n}\int_0^{2\pi} h(\theta)\,\cos(n\theta)\,d\theta,\\
B_n &= \frac{1}{\pi R^n}\int_0^{2\pi} h(\theta)\,\sin(n\theta)\,d\theta.
\end{aligned}$$

We of course need to discuss whether we can indeed write h in this form, but before this we should give a conceptual reason why the functions we considered are automatically orthogonal to each other. For this purpose we need to consider symmetric operators on V. Indeed, the operators A we want to consider are not defined on all of V (or if we define A on V, it will not map V to itself), so we have to enlarge our framework.

So we consider operators defined on a linear subspace D of V. In order to make this definition well-behaved, we need to assume that D is *dense* in V, i.e. if $v \in V$ then there exists $v_j \in D$ such that $v_j \to v$.

Definition 5. A linear operator $A : D \to V$ is called *symmetric* if

$$\langle Av, w\rangle = \langle v, Aw\rangle$$

for all $v, w \in D$.

Recall that an eigenvector of A is an element $v \neq 0$ of D such that $Av = \lambda v$ for some $\lambda \in \mathbb{C}$; λ is then an eigenvalue of A.

Lemma 0.8. *Suppose $A : D \to V$ is symmetric. Then all eigenvalues of A are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal to each other.*

Proof. Suppose that λ is an eigenvalue of A, so for some $v \in D$, $v \neq 0$, $Av = \lambda v$. Then

$$\lambda\|v\|^2 = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\|v\|^2,$$

so dividing through by $\|v\|^2$ we deduce that $\lambda = \overline{\lambda}$, so λ is real.

Now suppose that λ and μ are eigenvalues of A (so they are both real), and $Av = \lambda v$, $Aw = \mu w$. Then

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Av, w\rangle = \langle v, Aw\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle.$$

If $\lambda \neq \mu$ then rearranging and dividing by $\lambda - \mu$ gives $\langle v, w\rangle = 0$, as claimed. □

Now, the operator $A = -\frac{d^2}{d\theta^2}$ defined for functions $f \in C^2(\mathbb{S}^1)$ (which, recall, can be identified with 2π-periodic functions on $\mathbb{R}$), is symmetric, since

$$\begin{aligned}\langle Af, g\rangle &= \int_0^{2\pi} -f''(\theta)\overline{g(\theta)}\, d\theta = -f'(\theta)\overline{g(\theta)}|_0^{2\pi} + \int_0^{2\pi} f'(\theta)\overline{g'(\theta)}\, d\theta \\ &= f(\theta)\overline{g'(\theta)}|_0^{2\pi} - \int_0^{2\pi} f(\theta)\overline{g''(\theta)}\, d\theta = \langle f, Ag\rangle,\end{aligned}$$

where the boundary terms vanishes due to periodicity. A similar computation for the operator $A = -\frac{d^2}{dx^2}$ defined for functions $f \in C^2([0, \ell])$ gives

$$\begin{aligned}\langle Af, g\rangle &= \int_0^{\ell} -f''(x)\overline{g(x)}\, dx = -f'(x)\overline{g(x)}|_0^{\ell} + \int_0^{\ell} f'(x)\overline{g'(x)}\, dx \\ &= -f'(x)\overline{g(x)}|_0^{\ell} + f(x)\overline{g'(x)}|_0^{\ell} - \int_0^{\ell} f(x)\overline{g''(x)}\, dx \\ &= \Big(f(x)\overline{g'(x)} - f'(x)\overline{g(x)}\Big)|_0^{\ell} + \langle f, Ag\rangle\end{aligned}$$

Thus, the operator A is symmetric provided we restrict it to a domain D so that the boundary terms vanish. For instance,

(i) $A = -\frac{d^2}{dx^2}$ defined for functions $f \in C^2([0, \ell])$ such that $f(0) = 0 = f(\ell)$ (Dirichlet boundary condition) is symmetric,

(ii) $A = -\frac{d^2}{dx^2}$ defined for functions $f \in C^2([0, \ell])$ such that $f'(0) = 0 = f'(\ell)$ (Neumann boundary condition) is symmetric,

(iii) $A = -\frac{d^2}{dx^2}$ defined for functions $f \in C^2([0, \ell])$ such that $f'(0) - af(0) = 0$ and $f'(\ell) - bf(\ell) = 0$ for some given $a, b \in \mathbb{R}$ (Robin boundary condition) is symmetric.

As another example, the operator $A = -i\frac{d}{dx}$ defined on C^1 functions which are 2ℓ-periodic, mapping into $C^0([0, 2\ell])$ with the standard inner product, is symmetric since

$$\langle Af, g\rangle = \int_0^{2\ell} -if'(x)\overline{g(x)}\,dx = -if(x)\overline{g(x)}|_0^{2\ell} + i\int_0^{2\ell} f(x)\overline{g'(x)}\,dx = \langle f, Ag\rangle.$$

As a consequence, we deduce:

Corollary 0.9. *The following sets of functions are all orthogonal to each other in the respective vector spaces.*

(i) *In $V = C^0([0, \ell])$, with the standard inner product,*

$$f_n(x) = \sin\left(\frac{n\pi x}{\ell}\right),\ n \geq 1,\ n \in \mathbb{Z}.$$

(ii) *In $V = C^0([0, \ell])$, with the standard inner product,*

$$f_0(x) = 1,\ f_n(x) = \cos\left(\frac{n\pi x}{\ell}\right),\ n \geq 1,\ n \in \mathbb{Z}.$$

(iii) *In $V = C^0([0, 2\ell])$, with the standard inner product,*

$$f_n(x) = e^{in\pi x/\ell},\ n \in \mathbb{Z}.$$

(iv) *In $V = C^0([0, 2\ell])$, with the standard inner product,*

$$f_0(x) = 1,\ f_n(x) = \cos\left(\frac{n\pi x}{\ell}\right),\ n \geq 1,\ n \in \mathbb{Z},\ g_n(x) = \sin\left(\frac{n\pi x}{\ell}\right),\ n \geq 1,\ n \in \mathbb{Z}.$$

Proof. These are all consequences of being eigenfunctions of symmetric operators, namely:

(i) $A = -\frac{d^2}{dx^2}$, with Dirichlet boundary condition, the eigenvalue of f_n is $\lambda_n = \frac{n^2\pi^2}{\ell^2}$,

(ii) $A = -\frac{d^2}{dx^2}$, with Neumann boundary condition, the eigenvalue of f_n is $\lambda_n = \frac{n^2\pi^2}{\ell^2}$,

(iii) $A = -i\frac{d}{dx}$, on 2ℓ-periodic C^1 functions, the eigenvalue of f_n is $\lambda_n = \frac{n\pi}{\ell}$,

(iv) $A = -\frac{d^2}{dx^2}$, on 2ℓ-periodic C^2 functions, the eigenvalue of f_n is $\lambda_n = \frac{n^2\pi^2}{\ell^2}$ which is also the eigenvalue of g_n.

Thus, in all cases but the last the functions are eigenfunctions of a symmetric operator with distinct eigenvalues, hence are orthogonal. In the last case, we have all of the claimed orthogonality by the same argument, except the orthogonality of f_n to g_n. This can be seen easily, however, as the cosines are even around $x = \ell$, while the sines are odd around $x = \ell$, so the product is odd, hence its integral over the interval $[0, 2\ell]$, which is symmetric around ℓ, vanishes. □

We also have the following general symmetry result for the Laplacian:

Proposition 0.10. *Let Ω be a bounded domain with smooth boundary, $V = C^0(\bar{\Omega})$. The Laplacian Δ, defined on any one of the following domains:*

(i) *(Dirichlet) $D = \{f \in C^2(\overline{\Omega}) :\ f|_{\partial\Omega} = 0\}$,*

(ii) *(Neumann) $D = \{f \in C^2(\overline{\Omega}) :\ \frac{\partial f}{\partial n}|_{\partial\Omega} = 0\}$,*

(iii) *(Robin) $D = \{f \in C^2(\overline{\Omega}) :\ \left(\frac{\partial f}{\partial n} - af\right)|_{\partial\Omega} = 0\}$,*

is symmetric. Here, in case (iii), a is a given continuous function on $\partial\Omega$.

Proof. We recall Green's identity,

$$\int_\Omega (f\Delta g - (\Delta f)g)\,dx = \int_{\partial\Omega} \left(f\frac{\partial g}{\partial n} - \frac{\partial f}{\partial n} g \right) dS(x).$$

Replacing g by $\overline{g}$, we have

$$\int_\Omega (f\overline{\Delta g} - (\Delta f)\overline{g})\,dx = \int_{\partial\Omega} \left(f\frac{\partial g}{\partial n} - \frac{\partial f}{\partial n} g \right) dS(x).$$

Under each of the conditions listed above, the right hand side vanishes. Thus,

$$\int_\Omega f\overline{\Delta g}\,dx = \int_\Omega (\Delta f)\overline{g}\,dx,$$

so Δ is symmetric.

We recall how Green's identity is proved: consider the vector field $f\nabla g$. By the divergence theorem,

$$\int_\Omega \operatorname{div}(f\nabla g)\,dx = \int_{\partial\Omega} f\hat{n}\cdot\nabla g\,dS(x) = \int_{\partial\Omega} f\frac{\partial g}{\partial n}\,dS(x),$$

and similarly

$$\int_\Omega \operatorname{div}(g\nabla f)\,dx = \int_{\partial\Omega} g\frac{\partial f}{\partial n}\,dS(x).$$

Subtracting these two yields

$$\begin{aligned}\int_{\partial\Omega} \left(f\frac{\partial g}{\partial n} - \frac{\partial f}{\partial n} g \right) dS(x) &= \int_\Omega (\operatorname{div}(f\nabla g) - \operatorname{div}(g\nabla f))\,dx \\ &= \int_\Omega (f\Delta g + \nabla f\cdot\nabla g - g\Delta f - \nabla g\cdot\nabla f)\,dx = \int_\Omega (f\Delta g - (\Delta f)g)\,dx,\end{aligned}$$

as claimed. □

We still need to discuss whether 'any' v in V can be written as a stated linear combination. This is in general not the case: for instance, if $V = \mathbb{R}^2$, and $x_1 = (1,0)$, then $\{x_1\}$ is an orthogonal set of non-zero vectors, but not every element of V can be written as a linear combination of $\{x_1\}$ (i.e. is not, in general, a multiple of x_1): we need another vector, such as $x_2 = (0,1)$.

We thus make the following definition.

Definition 6. Suppose V is an inner product space. We say that an *orthogonal set* $\{x_n\}_{n=1}^\infty$ is *complete* if for every $v \in V$ there exist scalars c_n such that $v = \sum_{n=1}^\infty c_n x_n$.

A complete orthogonal set is also called an *orthogonal basis.*

If all of its elements have norm 1, an orthogonal basis is called an *orthonormal basis.*

Note that, as discussed before, if the c_n exist, they are automatically unique: they are determined by (3). In fact, we can always *define* a series by the formula (3): given $v \in V$, let

$$v_n = \sum_{k=1}^n c_k x_k,\ c_k = \frac{\langle v, x_k\rangle}{\|x_k\|^2}; \tag{4}$$

the question is whether $v_n \to v$. The coefficients c_k in (4) may be called the *generalized Fourier coefficients* of v. Note that v_n has the useful property that $v - v_n$ is orthogonal to x_j for $1 \le j \le n$. Indeed,

$$\langle v - v_n, x_j\rangle = \langle v, x_j\rangle - \sum_{k=1}^n c_k\langle x_k, x_j\rangle = c_j - c_j = 0,\ j \le n,$$

hence $v-v_n$ is also orthogonal to any vector of the form $\sum_{k=1}^n b_k x_k$, i.e. to the linear span of $\{x_1,\dots,x_n\}$, in particular to v_n itself. Thus, by Pythagoras' theorem, we have

$$\|v\|^2 = \|v-v_n\|^2 + \|v_n\|^2 \geq \|v_n\|^2 = \sum_{j,k=1}^n c_j\overline{c_k}\langle x_j,x_k\rangle = \sum_{k=1}^n |c_k|^2\|x_k\|^2.$$

This is *Bessel's inequality.* Thus, we deduce that for any $v\in V$, the generalized Fourier coefficients satisfy

$$\sum_{k=1}^n |c_k|^2\,\|x_k\|^2 \leq \|v\|^2,$$

hence, as these partial sums form a bounded monotone increasing (which, recall, means non-decreasing) sequence, $\sum_{k=1}^\infty |c_k|^2\,\|x_k\|^2$ converges, and as $\|v\|^2$ is an upper bound for the partial sums, the limit is $\leq \|v\|^2$. If this limit is $\|v\|^2$, then we deduce that $\|v-v_n\|\to 0$ as $n\to\infty$, so with the converse being clear we conclude the following:

Lemma 0.11. *Suppose V is an inner product space. An orthogonal set $\{x_n\}_{n=1}^\infty$ is complete if and only if for every $v\in V$ the generalized Fourier coefficients satisfy*

$$\sum_{k=1}^\infty |c_k|^2\|x_k\|^2 = \|v\|^2.$$

This may be called *Bessel's equality*, but is usually called *Parseval's identity.*

Now, let's compute $\|v-\sum_{k=1}^n a_kx_k\|^2$, where $a_k\in\mathbb{C}$. This is best done by writing $v = v_\| + v_\perp$, where

$$v_\| = \sum_{k=1}^n c_kx_k,$$

the c_k being the generalized Fourier coefficients, for then $v_\perp$ is orthogonal to all x_k, and hence

$$\|v-\sum_{k=1}^n a_kx_k\|^2 = \|v_\perp + \sum_{k=1}^n (c_k-a_k)x_k\|^2 = \|v_\perp\|^2 + \sum_{k=1}^n |c_k-a_k|^2\|x_k\|^2.$$

In particular,

$$\|v-\sum_{k=1}^n a_kx_k\|^2 \geq \|v_\perp\|^2,$$

and equality holds if and only of $a_k=c_k$ for $k=1,\dots,n$. This is the *least squares approximation*:

Proposition 0.12. *Given $v\in V$, and an orthogonal set $\{x_1,\dots,x_n\}$, the choice of a_k that minimizes the error $\|v-\sum_{k=1}^n a_kx_k\|^2$ of the approximation of v by $\sum_{k=1}^n a_kx_k$ is given by the generalized Fourier coefficients, c_k.*

Thus, even if the generalized Fourier series $\sum_{k=1}^\infty c_kx_k$ does not converge to the function, its partial sums give the best possible approximations of v in the precise sense of this Proposition.

We now actually turn to completeness statements. We prove in the next lecture that all of the orthogonal sets listed in Corollary 0.9 are complete. Moreover, for the Laplacian we have the following result, which is beyond our tools in this class:

Theorem 0.13. *Suppose Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary. For the Laplacian, defined on the subspace D of $C^0(\overline{\Omega})$ given by (i), (ii) or (iii) of Proposition 0.10, the eigenspace for each eigenvalue $\lambda \in \mathbb{R}$ is finite dimensional, the eigenvalues can be arranged in an increasing sequence, tending to infinity, $\lambda_1 \leq \lambda_2 \leq \ldots$, and choosing an orthogonal basis $x_{k,j}$, $j = 1, \ldots, N_k$, of the λ_j-eigenspace, the orthogonal set*

$$\{x_{k,j} : \ k \in \mathbb{Z}, \ k \geq 1, \ j = 1, \ldots, N_k\}$$

is complete.

The analogue of this theorem for variable coefficient symmetric so-called elliptic second order differential operators with appropriate boundary conditions is also valid.

There is one more issue that should be observed. For the generalized Fourier series of v as in (4), and for $n < m$,

$$\|v_n - v_m\|^2 = \sum_{k=n+1}^{m} |c_k|^2 \, \|x_k\|^2. \tag{5}$$

Since $\sum_{k=1}^{\infty} |c_k|^2 \, \|x_k\|^2$ converges, the differences of the partial sums, on the right hand side of (5), go to 0 as $n, m \to \infty$. Thus, by (5), the same holds for the differences of the partial sums of $\sum_{k=1}^{\infty} c_k x_k$, i.e. the partial sums form a *Cauchy sequence*. In an ideal world, this ought to imply that the series $\sum_{k=1}^{\infty} c_k x_k$ converges, but this only holds if our inner product space is *complete*.

Thus, we see that for any inner product space and any orthogonal set inside it, the generalized Fourier series for any $v \in V$ is Cauchy. If V is complete, it will thus converge to some $v' \in V$ (not necessarily to v though: recall the $\mathbb{R}^2$ example).

Now, an example of a complete normed vector space is, for Ω as above, $V = C^0(\overline{\Omega})$ with the C^0-norm:

$$\|f\|_{C^0} = \sup_{x \in \overline{\Omega}} |f(x)|.$$

Unfortunately, an example of an incomplete normed vector space is $V = C^0(\overline{\Omega})$ with the L^2-norm, i.e. our inner product space. On the other hand, $L^2(\mathbb{R}^n)$ and $L^2(\Omega)$ are complete, where Ω is measurable.

Theorem 0.14. *The inner product space $L^2(\mathbb{R}^n)$ is complete, as is $L^2(\Omega)$.*

Proof. The $L^2(\Omega)$ completeness follows from the $L^2(\mathbb{R}^n)$ case by regarding elements of $L^2(\Omega)$ as equivalence classes of functions on $\mathbb{R}^n$ which vanish outside Ω. (Alternatively, the proof given below goes through for $L^2(\Omega)$ directly.)

The proof of $L^2(\mathbb{R}^n)$ being complete is completely analogous to the proof of the completeness of $L^1(\mathbb{R}^n)$. Thus, suppose that $\{f_k\}_{k=1}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. It suffices to show that it has a subsequence converging to some $f \in L^2(\mathbb{R}^n)$, for then the convergence of the full Cauchy sequence follows. As in the case of $L^1(\mathbb{R}^n)$, we may pick a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ such that $\|f_{k_{j+1}} - f_{k_j}\|_{L^2} < 2^{-j}$ for $j \geq 1$. We claim that this subsequence converges pointwise almost everywhere to a function f, $f \in L^2$, and $\|f - f_{k_j}\|_{L^2} \to 0$ as $j \to \infty$, completing the proof of the theorem.

To see the claim, let

$$s_N(x) = f_{k_1}(x) + \sum_{j=1}^{N-1} (f_{k_{j+1}}(x) - f_{k_j}(x)),$$

and

$$\sigma_N(x) = |f_{k_1}(x)| + \sum_{j=1}^{N-1} |f_{k_{j+1}}(x) - f_{k_j}(x)|.$$

Notice that the sum defining $s_N(x)$ actually telescopes, so $s_N(x) = f_{k_N}(x)$ for all x. Now, as σ_N is monotone increasing in N, the pointwise limits

$$g(x) = \lim_{N\to\infty} \sigma_N(x)$$

exist as extended non-negative real numbers, and $\int \sigma_N^2 \to \int g^2$. Further, by the triangle inequality

$$\|\sigma_N\|_{L^2} \leq \|f_{k_1}\|_{L^2} + \sum_{j=1}^{N-1} \|f_{k_{j+1}} - f_{k_j}\|_{L^2} \leq \|f_{k_1}\|_{L^2} + \sum_{j=1}^{N-1} 2^{-j} \leq \|f_{k_1}\|_{L^2} + 1$$

for all N, so the limit $\|g\|_{L^2}$ satisfies the same inequality:

$$\|g\|_{L^2} \leq \|f_{k_1}\|_{L^2} + 1.$$

In particular, g is finite a.e., so the series $f_{k_1}(x) + \sum_{j=1}^{\infty}(f_{k_{j+1}}(x) - f_{k_j}(x))$, whose partial sums are the s_N, converges absolutely for a.e. x, and thus itself converges for a.e. x. We let

$$f(x) = \lim_{N\to\infty} s_N(x) = f_{k_1}(x) + \sum_{j=1}^{\infty}(f_{k_{j+1}}(x) - f_{k_j}(x)),$$

defined a.e. x. Now

$$\|s_N\|_{L^2} \leq \|f_{k_1}\|_{L^2} + \sum_{j=1}^{N-1} \|f_{k_{j+1}} - f_{k_j}\|_{L^2} \leq \|f_{k_1}\|_{L^2} + 1$$

again by the triangle inequality (applied to the sum defining s_N). Since $|s_N(x)| \leq |\sigma_N(x)| \leq g(x)$ for a.e. x, the dominated convergence theorem implies $\int |s_N|^2 \to \int |f|^2$, and thus $\int |f|^2 \leq (\|f_{k_1}\|_{L^2} + 1)^2$, so $f \in L^2$.

Finally, for $N \geq m$,

$$\|f_{k_N} - f_{k_m}\|_{L^2} \leq \sum_{j=m}^{N-1} \|f_{k_{j+1}} - f_{k_j}\|_{L^2} \leq \sum_{j=m}^{N-1} 2^{-j} \leq 2^{-m+1},$$

and

$$|f_{k_N}(x) - f_{k_m}(x)| \leq \sum_{j=m}^{N-1} |f_{k_{j+1}}(x) - f_{k_j}(x)| \leq \sigma_N(x) \leq g(x)$$

for a.e. x, so by the dominated convergence theorem $\int |f_{k_N} - f_{k_m}|^2 \to \int |f - f_{k_m}|^2$ as $N \to \infty$. Correspondingly, as $\|f_{k_N} - f_{k_m}\|_{L^2} \leq 2^{-m+1}$ for all N, $\|f - f_{k_m}\|_{L^2} \leq 2^{-m+1}$ as well, showing that $f_{k_m} \to f$ in L^2, finishing the proof of the completeness. □

We use this opportunity to make a definition:

Definition 7. A complete inner product space is called a *Hilbert space.*

Moreover, we also have:

Theorem 0.15. *The set of compactly supported continuous functions on* $\mathbb{R}^n$ *is dense in* $L^2(\mathbb{R}^n)$*.*

Proof. This theorem reduces to the density of simple functions in $L^2(\mathbb{R}^n)$, for then one can approximate the characteristic functions of measurable sets by step functions, and then in turn by continuous functions of compact support, exactly as in the case of L^1.

To see the density of simple functions, decompose f into real and imaginary parts, so one may assume f is real, and then further write $f = f_+ - f_-$, $f_+ = \max(f,0)$, $f_- = -\min(f,0)$, and note $f_\pm \in L^2$ since $f_+(x)^2 + f_-(x)^2 = f(x)^2$ for all x. Then take $\phi_{k,\pm} \geq 0$ simple functions such that $\phi_{k,\pm} \nearrow f_\pm$ pointwise; then $\int |f_\pm - \phi_{k,\pm}|^2 \to 0$ since $|f_\pm(x) - \phi_{k,\pm}(x)|^2 \to 0$ for all x, and $|f_\pm - \phi_{k,\pm}|^2 \leq f_\pm^2$, so by the dominated convergence theorem $\int |f_\pm - \phi_{k,\pm}|^2 \to 0$ indeed. □

Now, recall that every incomplete normed vector space V can be completed, i.e. there is a complete normed vector space $\hat{V}$ and an inclusion map $\iota : V \to \hat{V}$ which is linear such that for $v \in V$, $\|\iota(v)\|_{\hat{V}} = \|v\|_V$, and such that for any $v \in \hat{V}$, there is a sequence $\{v_n\}_{n=1}^\infty$ in V such that $\iota(v_n) \to v$ in $\hat{V}$ (i.e. the image of V under ι is dense in $\hat{V}$).

This completion is (essentially) unique. Note that if $\{v_n\}_{n=1}^\infty$ is Cauchy in V, then $\{\iota(v_n)\}_{n=1}^\infty$ is Cauchy in $\hat{V}$, so it converges. Moreover, ι is one-to-one, so one can simply think of elements of V as elements of $\hat{V}$. It is also useful to note that if V is an inner product space, then so is $\hat{V}$.

We thus conclude:

Corollary 0.16. *The completion of $V = C_c(\mathbb{R}^n)$ with the inner product $\langle f, g\rangle = \int f\overline{g}$ is $L^2(\mathbb{R}^n)$.*

The same applies if one takes V as in Example (vi).

We also remark that if V is an inner product space, and $\{x_n\}_{n=1}^\infty$ is a complete orthogonal set in V, then it is also complete in the completion $\hat{V}$ of V: this follows immediately since for $v \in V$ we can take a sequence $v_n \in V$ such that $v_n \to v$, and use that the generalized Fourier series of v_n converges to v_n, as well as the fact that the generalized Fourier series of v converges to some v' to show that $v' = v$.