Orthonormal Bases in Hilbert Space
APPM 5440 Fall 2017 Applied Analysis

Stephen Becker

November 18, 2017(v4); v1 Nov 17 2014 and v2 Jan 21 2016

Supplementary notes to our textbook (Hunter and Nachtergaele). These notes generally follow Kreyszig's book. The reason for these notes is that this is a simpler treatment that is easier to follow; the simplicity is because we generally do not consider uncountable nets, but rather only sequences (which are countable). I have tried to identify the theorems from both books; theorems/lemmas with numbers like 3.3-7 are from Kreyszig.

Note: there are two versions of these notes, with and without proofs. The version without proofs will be distributed to the class initially, and the version with proofs will be distributed later (after any potential homework assignments, since some of the proofs maybe assigned as homework). **This version does not have proofs**.

1 Basic Definitions

NOTE: Kreyszig defines inner products to be linear in the first term, and conjugate linear in the second term, so $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle = \langle x, \bar{\alpha} y\rangle$. In contrast, Hunter and Nachtergaele define the inner product to be linear in the second term, and conjugate linear in the first term, so $\langle \bar{\alpha} x, y\rangle = \alpha\langle x, y\rangle = \langle x, \alpha y\rangle$. We will follow the convention of Hunter and Nachtergaele, so I have re-written the theorems from Kreyszig accordingly whenever the order is important. Of course when working with the real field, order is completely unimportant.

Let X be an inner product space. Then we can define a norm on X by

$$\|x\| = \sqrt{\langle x, x\rangle}, \quad x \in X.$$

Thus, X is also a vector space (or normed linear space), and we can discuss completeness on X.

Definition 1. *A **Hilbert space**, typically denoted $\mathcal{H}$, is a complete inner product space. One must specify the field F, and we will always assume it is either $\mathbb{R}$ or $\mathbb{C}$ (note: the field cannot be arbitrary in a Hilbert space — e.g., finite fields do* not *work).*

Definition 2. *Let X and Y be inner product spaces over the field F. A mapping $T : X \to Y$ is an **isomorphism** if it is an **invertible** (hence **one-to-one**) linear transformation from X **onto** Y such that*

$$\langle Tx, Ty\rangle = \langle x, y\rangle, \quad x, y \in X.$$

*In this case, X and Y are said to be **isomorphic**.*

The above isomorphism definition is in Hunter and Nachtergaele, but not until §8 (cf. Def. 8.28).

Lemma 3 (Lemma 3.3-7)**.** *Let M be a non-empty subset of a Hilbert space $\mathcal{H}$. Then the span(M) is dense in $\mathcal{H}$ if and only if $M^\perp = \{0\}$.*

2 Orthonormal sets

A first result is exercise 6.6 in Hunter/Nachtergaele (or lemma 3.4-2 in Kreyszig): vectors in an orthogonal set are linearly independent.

Theorem 4 (Thm. 3.4-6)**.** *(Bessel inequality) Let (e_k) be an orthonormal sequence in an inner product space X. For every $x \in X$, we have*

$$\sum_{k=1}^{\infty} |\langle e_k, x\rangle|^2 \le \|x\|^2.$$

Definition 5. *Let (e_k) be an orthonormal sequence in an inner product space X. Let $x \in X$. The quantities $\langle e_k, x\rangle$ are called the* ***Fourier coefficients*** *of x with respect to the orthonormal sequence (e_k).*

Now we will discuss the convergence of the following **Fourier series**:

$$\sum_{k=1}^{\infty} \langle e_k, x\rangle e_k.$$

Theorem 6 (Thm. 3.5-2)**.** *Let (e_k) be an orthonormal sequence in a Hilbert space $\mathcal{H}$. Then:*

(a) The series $\sum_{k=1}^{\infty} \alpha_k e_k$ converges if and only if the series $\sum_{k=1}^{\infty} |\alpha_k|^2$ converges.

(b) If the series $\sum_{k=1}^{\infty} \alpha_k e_k$ converges, we write $x = \sum_{k=1}^{\infty} \alpha_k e_k$, then we have $\alpha_k = \langle e_k, x\rangle$, and consequently,

$$x = \sum_{k=1}^{\infty} \langle e_k, x\rangle e_k.$$

(c) For any $x \in \mathcal{H}$, the series $\sum_{k=1}^{\infty} \langle e_k, x\rangle e_k$ converges. (Note that the sum may not equal to x.)

In the above Theorem 6(c), we know that the sum may not be x itself. To ensure the sum is x, we need to assume that (e_k) is a **total** orthonormal set which is defined as follows.

Definition 7. *Let X be an inner product space and let M be a subset. Then M is called* ***total*** *if $\overline{span(M)} = X$.*

Definition 8. *A total orthonormal set in an inner product space is called an* ***orthonormal basis***. ***N.B.*** *Other authors, such as Reed and Simon, define an orthonormal basis as a* maximal *orthonormal set, e.g., an orthonormal set which is not properly contained in any other orthonormal set. The two definitions are equivalent (Hunter and Nachtergaele's theorem).*

Theorem 9 (Thm. 4.1-8)**.** *Every Hilbert space contains a total orthonormal set. (Furthermore, all total orthonormal sets in a Hilbert space $\mathcal{H} \neq \{0\}$ have the same cardinality, which is known as the* ***Hilbert dimension****).*

See Kreyszig, where he states this without proof in §3.6 and proves it in §4.1. The corresponding result in Hunter/Nachtergaele is Theorem 6.29. The proof requires the axiom of choice or Zorn's lemma.

Theorem 10 (Thm. 3.6-2)**.** *Let M be a subset of an inner product space X. Then:*

(a) If M is total in X, then $x \perp M$ implies $x = 0$.

(b) Assume that X is complete. If $x \perp M$ implies $x = 0$, then M is total in X.

Lemma 11 (Lemma 3.5-3)**.** *Let X be an inner product space and let $x \in X$. Let (e_k), $k \in I$, be an orthonormal set in X. Then at most countably many of the Fourier coefficients $\langle e_k, x\rangle$ are non-zero.* (This lemma is the key to Kreyszig's simpler approach to non-separable Hilbert spaces)

Let (e_k), $k \in I$, be an orthonormal set in an inner product space X. Let $x \in X$. The above lemma shows that the sum $\sum_k |\langle e_k, x\rangle|^2$ is a countable sum. Hence, Bessel's inequality can be applied to conclude $\sum_k |\langle e_k, x\rangle|^2 \le \|x\|^2$ in this case.

Now we have the following criterion for totality.

Theorem 12 (Thm. 3.6-3)**.** *Let M be an orthonormal set in a Hilbert space $\mathcal{H}$. Then M is total in $\mathcal{H}$ if and only if for all $x \in \mathcal{H}$, the following **Parseval relation** holds (where we are summing over the non-zero terms only)*

$$\sum_k |\langle e_k, x\rangle|^2 = \|x\|^2.$$

Now we discuss Hilbert spaces that contain countable orthonormal sets. Note: the following theorem is the same as exercise 6.10 in Hunter/Nachtergaele ("A Hilbert space is a separable metric space iff it has a countable orthonormal basis"):

Theorem 13 (Thm. 3.6-4)**.** *Let $\mathcal{H}$ be a Hilbert space.*

1. *If $\mathcal{H}$ is separable, every orthonormal set in H is countable.*
2. *If $\mathcal{H}$ contains an orthonormal sequence which is total in $\mathcal{H}$, then $\mathcal{H}$ is separable.*

And the major result of Hilbert space theory is the following:

Theorem 14 (Thm. 3.6-5)**.** *Two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, over the same field ($\mathbb{R}$ or $\mathbb{C}$), are isomorphic if and only if they have the same (Hilbert) dimension.*

3 Advanced definitions

The following is taken from Combettes and Bauschke's "Convex Analysis and Monotone Theory in Hilbert Spaces". Let M be a nonempty set and let $\preceq$ a binary relation on $M \times M$. Consider the following statements, and for all statements, let a and b be any arbitrary element of M, then

1. $a \preceq a$.
2. $(\forall c \in M), \quad \big(a \preceq b \text{ and } b \preceq c\big) \implies a \preceq c$.
3. $(\exists c \in M), \quad a \preceq c \text{ and } b \preceq c$.
4. $\big(a \preceq b \text{ and } b \preceq a\big) \implies a = b$.
5. $a \preceq b$ or $b \preceq a$.

if (1), (2) and (3) hold, then we call $(M, \preceq)$ a **directed set**. If (1), (2) and (4), we call $(M, \preceq)$ a **partially ordered set**. A partially ordered set with the property (5) is a **totally ordered set** (or a **chain**).

This allows us to define Zorn's lemma, and to talk about nets.

3.1 Nets

From §1.4 in Combettes/Bauschke, who follow §2 in Kelley's classic topology text "General Topology" (1955). The idea is to generalize the notion of a sequence, so that in this generalized notion, continuous is always the same as sequentially continuous (with the definition of sequentially continuous suitably modified) in all topologies. The theory of nets is due to Moore and Smith (1922); a similar theory that generalizes sequences, using the notion of a **filter**, is due to Cartan in 1937. Neither theory will be relevant for our class.

Let $(\mathcal{A}, \preceq)$ be a directed set. We write $b \succeq a$ to mean $a \preceq b$. Let X be a nonempty set. A **net** or **generalized sequence** in X indexed by $\mathcal{A}$ is denoted by $(x_a)_{a\in\mathcal{A}}$ (or just (x_a) if $\mathcal{A}$ is clear from context).

For example, taking $(\mathbb{N}, \leq)$, we see that every sequence is a net. Not every net is a sequence, since we may have an uncountable index set $\mathcal{A}$.

Let (x_a) be a net in X, and let $Y \subset X$. We say that (x_a) is **eventually** in Y if there is some $c \in \mathcal{A}$ such that for all $a \in \mathcal{A}$, $a \succeq c \implies x_a \in Y$. We say the net is **frequently** in Y if for all $c \in \mathcal{A}$, there is some $a \in \mathcal{A}$ such that $a \succeq c$ and $x_a \in Y$.

3.2 Zorn's lemma

We follow Kreyszig again.

Definition 15 (Chain)**.** *A **chain**, or **totally ordered set**, is a partially ordered set such that every two elements are comparable.*

An **upper bound** of a subset W of a partially ordered set M is an element $u \in M$ such that $x \preceq u$ for every $x \in W$. Such an element need not exist. A **maximal element** of M is some $m \in M$ such that $m \preceq x$ implies $m = x$. Again, such an element need not exist, and if it does, it need not be an upper bound.

Note the funny definition of a maximal element, which is not the same as defining a **greatest element** (an element m such that $x \preceq m$ for all $x \in W$; i.e., an *upper bound* that lives within the set). These notions are not the same, since for a maximal element, we may not be able to compare $x \preceq m$ (or $m \preceq x$) for some x. If the set is totally ordered, then maximal and greatest elements are the same concept.

Similarly we can define a **least element**. A **well-ordered set** is a totally ordered set with the property that every subset has a least element. Zermelo's theorem says that every set can be well-ordered; that is, we can define some $\preceq$ to make it a well-ordered set. This is equivalent to the axiom of choice, and highly counter-intuitive.

Here are four common partial orderings:

1. The usual ordering $\leq$, on subsets of the real numbers (e.g., arbitrary subsets, or integers or $\mathbb{Q}$ or $\mathbb{N}$); this is also a total ordering. Note that $\mathbb{C}$ does not have a canonical total ordering.

2. The ordering on sets of sets, defined by set inclusion $\subset$. E.g., we say $A \preceq B$ if $A \subset B$. For example, if we have $A = \{1, 2\}$ and $B = \{2, 3\}$, then neither $A \subset B$ nor $B \subset A$, so we do not have a total ordering.

   As an example, let $M = \{\{a\}, \{b\}, \{a, b\}\}$, then $\{a, b\}$ is both an upper bound and a maximal element. If $M = \{\{a\}, \{b\}, \{c\}, \{a, b\}\}$, then there is no upper bound, and there are two maximal elements, namely $\{c\}$ and $\{a, b\}$.

3. The ordering induced by the positive semi-definite (PSD) cone. We work on the space of symmetric matrices, and we say $0 \preceq A$ if A is a PSD matrix, and we write $B \preceq A$ if $0 \preceq A - B$, i.e., if $A - B$ is PSD. Consider a matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$, so then neither $A - B$ nor $B - A$ is PSD, so this is not a total ordering. However, it is a *directed* set, since for any two matrices A and B, we can alway find a third matrix C such that $C - A$ and $C - B$ are both PSD.

4. The ordering on $\mathbb{R}^n$ where $(x_1, \ldots, x_n) \preceq (y_1, \ldots, y_n)$ if $x_i \leq y_i$ for all i.

The following "lemma" implies, and is implied by, the axiom of choice, so it is often taken to be synonymous with it.

Lemma 16 (Zorn's lemma, 4.1-6)**.** *Let $M \neq \emptyset$ be a partially ordered set, and suppose every chain $C \subset M$ has an upper bound, then M has at least one maximal element.*

Theorem 17 (Hamel basis, 4.1-7)**.** *Every vector space $X \neq \{0\}$ has a Hamel basis.*