Orthogonal systems in Hilbert space and applications

SML course - Introduction to functional analysis

NGUYEN Tue Tai / 062201848

July 13, 2023

1 Orthogonal systems

In this report, we will consider separable Hilbert space $\mathcal{H}$, which means that there exists a countable dense subset $\mathcal{D}$ of $\mathcal{H}$. We will refer to separable Hilbert spaces as simply Hilbert spaces unless specified. The following discussion is mostly inspired by [1] (some of proofs are also from [2], [3], [7]).

Definition 1. Let $\mathcal{H}$ be a Hilbert space. A sequence (e_n) in $\mathcal{H}$ is an *orthonormal system* if, for any pair of indices m and n,

$$\langle e_m, e_n \rangle = \delta_{mn} = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{if } m \neq n \end{cases} . \tag{1}$$

If $f \in \mathcal{H}$, then $\langle e_n, f \rangle$ is called *the Fourier coefficient* or *the* n^{th} *coefficient* of f relative to the system (e_n).

In the above definition, we just define orthonormal systems as countable (finite or countably infinite) sets of orthonormal vectors because of the following statement.

Theorem 2. *In a separable Hilbert space $\mathcal{H}$, every orthonormal set is finite or countably infinite.*

Proof. Let $\mathcal{D}$ be any dense subset in $\mathcal{H}$ and $\mathcal{N}$ be any set of orthonomal vectors. Then, any two distinct vectors $f,\ g \in \mathcal{N}$ have the distance $\sqrt{2}$ since

$$\|f - g\|^2 = \langle f - g, f - g \rangle = \langle f, f \rangle + \langle g, g \rangle - \langle f, g \rangle - \langle g, f \rangle = 1 + 1 + 0 + 0 = 2. \tag{2}$$

Hence, there exist the two disjoint open balls

$$\mathcal{B}_r(f) = \{h \in \mathcal{H} \mid \|f - h\| < r\}, \qquad \mathcal{B}_r(g) = \{h \in \mathcal{H} \mid \|g - h\| < r\} \tag{3}$$

for $r = \sqrt{2}/3$. Since $\mathcal{D}$ is dense in $\mathcal{H}$, then every open ball in $\mathcal{H}$ contains at least one element of $\mathcal{D}$. Hence, the exist $f_1 \in \mathcal{B}_r(f) \cap \mathcal{D}$ and $g_1 \in \mathcal{B}_r(g) \cap \mathcal{D}$. Since $\mathcal{B}_r(f) \cap \mathcal{B}_r(g) = \varnothing$, we have $f_1 \neq g_1$. If $\mathcal{N}$ is uncountable, then we would have uncountably many such pairs f and g along with f_1 and g_1 in $\mathcal{D}$. Thus, $\mathcal{D}$ would be uncountable in that case. Since $\mathcal{D}$ can be any dense subset, this means $\mathcal{H}$ would not contain any countable dense subsets, which is a contradiction to the separability of $\mathcal{H}$. Therefore, $\mathcal{N}$ is countable. □

From the system (e_n), we have a sequence of scalars (real or complex) $(\langle e_n, f \rangle)$ for every $f \in \mathcal{H}$. This sequence is square-summable as stated in the following theorem.

Theorem 3 (Bessel's inequality)**.** *In a Hilbert space $\mathcal{H}$, let (e_n) be an orthonormal system, then for any $f \in \mathcal{H}$,*

$$\sum_n |\langle e_n, f \rangle|^2 \leq \|f\|^2. \tag{4}$$

Proof. Consider $I \subset \mathbb{N}$ be a finite set of indices. Then, $\mathcal{M} = \text{Vect}\left((e_i)_{i \in I}\right)$ is a closed subspace of $\mathcal{H}$. Let f_1 be a vector of $\mathcal{M}$ such that

$$f_1 = \sum_{i \in I} \langle e_i, f \rangle e_i. \tag{5}$$

Then, let $f_2 = f - f_1$ and we have for any $j \in I$,

$$\begin{aligned}
\langle e_j, f_2 \rangle = \langle e_j, f - f_1 \rangle &= \left\langle e_j, f - \sum_{i \in I} \langle e_i, f \rangle e_i \right\rangle && (6)\\
&= \langle e_j, f \rangle - \sum_{i \in I} \langle e_i, f \rangle \langle e_j, e_i \rangle && (7)\\
&= \langle e_j, f \rangle - \langle e_j, f \rangle = 0. && (8)
\end{aligned}$$

Hence, $\langle g, f_2 \rangle = 0$ for any $g \in \mathcal{M}$, which implies that $f_2 \in \mathcal{M}^\perp$. Then, we have

$$\begin{aligned}
\|f\|^2 = \|f_1 + f_2\|^2 &= \langle f_1 + f_2, f_1 + f_2 \rangle && (9)\\
&= \langle f_1, f_1 \rangle + \langle f_2, f_2 \rangle + \langle f_1, f_2 \rangle + \langle f_2, f_1 \rangle && (10)\\
&= \|f_1\|^2 + \|f_2\|^2. && (11)
\end{aligned}$$

Because $\|f_2\|^2 \geq 0$, we have

$$\begin{aligned}
\|f\|^2 \geq \|f_1\|^2 &= \left\langle \sum_{i \in I} \langle e_i, f \rangle e_i, \sum_{j \in I} \langle e_j, f \rangle e_j \right\rangle && (12)\\
&= \sum_{i,j \in I} \overline{\langle e_i, f \rangle} \langle e_j, f \rangle \langle e_i, e_j \rangle && (13)\\
&= \sum_{i \in I} \overline{\langle e_i, f \rangle} \langle e_i, f \rangle && (14)\\
&= \sum_{i \in I} |\langle e_i, f \rangle|^2. && (15)
\end{aligned}$$

Because the sum on the right-hand side is bounded for any I, the series $\sum_n |\langle e_n, f \rangle|^2$ converges and we get the inequality (4). □

A natural question: for which orthonormal system (e_n) the equality occurs in the Bessel's inequality. To answer that, we come to the following definition.

Definition 4. An orthonormal system (e_n) in a Hilbert space $\mathcal{H}$ is said to be *complete* or *total* if the set of all finite linear combinations of vectors of (e_n) is dense in $\mathcal{H}$. A complete orthonormal system is also called a *Hilbert basis*.

If (e_n) is a complete orthonormal system, this means that for any $f \in \mathcal{H}$ and any $\varepsilon > 0$, there exists a finite linear combination $\sum_{i \in I} \lambda_i e_i$ such that

$$\left\| f - \sum_{i \in I} \lambda_i e_i \right\| \leq \varepsilon. \tag{16}$$

Theorem 5 (Parseval's identity)**.** *Let (e_n) be an orthonormal system in a Hilbert space $\mathcal{H}$. This system is complete if and only if for any $f \in \mathcal{H}$, we have*

$$\|f\|^2 = \sum_n |\langle e_n, f \rangle|^2. \tag{17}$$

Proof. If (e_n) is complete, then for any $f \in \mathcal{H}$ and any $\varepsilon > 0$, there exists a finite linear combination $\sum_{i \in I} \lambda_i e_i$ such that

$$\left\| f - \sum_{i \in I} \lambda_i e_i \right\|^2 \leq \varepsilon. \tag{18}$$

From the proof of Theorem 3, we had $\mathcal{M} = \text{Vect}\left((e_i)_{i\in I}\right)$ as a closed subspace of $\mathcal{H}$ and found $f_1 \in \mathcal{M}$, $f_2 \in \mathcal{M}^\perp$ such that $f = f_1 + f_2$. From the Orthogonal Projection Theorem (Theorem 10), we have $f_1 = P_{\mathcal{M}}(f)$ and

$$\|f_2\| = \|f - f_1\| = \inf_{g\in\mathcal{M}} \|f - g\|, \tag{19}$$

where

$$\|f_2\|^2 = \|f\|^2 - \|f_1\|^2 = \|f\|^2 - \sum_{i\in I} |\langle e_i, f\rangle|^2. \tag{20}$$

Because $\sum_{i\in I} \lambda_i e_i \in \mathcal{M}$, we have

$$\|f\|^2 - \sum_{i\in I} |\langle e_i, f\rangle|^2 = \|f_2\|^2 = \inf_{g\in\mathcal{M}} \|f - g\|^2 \leq \left\| f - \sum_{i\in I} \lambda_i e_i \right\|^2 \leq \varepsilon. \tag{21}$$

Hence,

$$\|f\|^2 \leq \sum_{i\in I} |\langle e_i, f\rangle|^2 + \varepsilon \leq \sum_{n} |\langle e_n, f\rangle|^2 + \varepsilon. \tag{22}$$

Since this inequality is true for any $\varepsilon > 0$, we have

$$\|f\|^2 \leq \sum_{n} |\langle e_n, f\rangle|^2 \tag{23}$$

and taking into account Bessel's inequality, we get Parseval's identity.

Conversely, if any $f \in \mathcal{H}$ satisfies Parseval's identity, then for any $\varepsilon > 0$, there exists a finite set of indices I such that

$$\|f\|^2 - \sum_{i\in I} |\langle e_i, f\rangle|^2 \leq \varepsilon^2. \tag{24}$$

Again, from the proof of Theorem 3, we have

$$\|f\|^2 - \sum_{i\in I} |\langle e_i, f\rangle|^2 = \|f\|^2 - \|f_1\|^2 = \|f_2\|^2 = \|f - f_1\|^2 = \left\| f - \sum_{i\in I} \langle e_i, f\rangle e_i \right\|^2. \tag{25}$$

Hence,

$$\left\| f - \sum_{i\in I} \langle e_i, f\rangle e_i \right\| \leq \varepsilon, \tag{26}$$

which implies that (e_n) is complete. □

Corollary 6. *An orthonormal system (e_n) in a Hilbert space $\mathcal{H}$ is complete if and only if the relation $\langle e_n, f\rangle = 0$ for every n, implies $f = 0$.*

Proof. The forward statement comes directly from Parseval's identity.

Conversely, assume that $\langle e_n, f\rangle = 0$ for every n, implies $f = 0$. Let $\mathcal{M}$ be the set of all linear combinations of (e_n). Then, we have $\mathcal{M}^\perp = \{0\}$, which implies that $\mathcal{M}$ is dense in $\mathcal{H}$. Hence, (e_n) is complete. □

Corollary 7. *Let (e_n) be a complete orthonormal system in a Hilbert space $\mathcal{H}$. For any f and g in $\mathcal{H}$,*

$$\langle f, g\rangle = \sum_{n} \overline{\langle e_n, f\rangle} \langle e_n, g\rangle = \sum_{n} \langle f, e_n\rangle \langle e_n, g\rangle \tag{27}$$

Proof. Using the polarisation identity and Parseval's identity, we have

$$4\langle f,g\rangle = \|f+g\|^2 - \|f-g\|^2 - i\|f+ig\|^2 + i\|f-ig\|^2 \tag{28}$$
$$= \sum_n |\langle e_n, f+g\rangle|^2 - \sum_n |\langle e_n, f-g\rangle|^2 - i\sum_n |\langle e_n, f+ig\rangle|^2 + i\sum_n |\langle e_n, f-ig\rangle|^2 \tag{29}$$
$$= \sum_n \left(|\langle e_n, f+g\rangle|^2 - |\langle e_n, f-g\rangle|^2 - i|\langle e_n, f+ig\rangle|^2 + i|\langle e_n, f-ig\rangle|^2\right), \tag{30}$$

where

$$|\langle e_n, f+g\rangle|^2 - |\langle e_n, f-g\rangle|^2 = \langle f+g, e_n\rangle\langle e_n, f+g\rangle - \langle f-g, e_n\rangle\langle e_n, f-g\rangle \tag{31}$$
$$= 2\langle f, e_n\rangle\langle e_n, g\rangle + 2\langle g, e_n\rangle\langle e_n, f\rangle, \tag{32}$$

$$-i|\langle e_n, f+ig\rangle|^2 + i|\langle e_n, f-ig\rangle|^2 = i\langle f-ig, e_n\rangle\langle e_n, f-ig\rangle - i\langle f+ig, e_n\rangle\langle e_n, f+ig\rangle \tag{33}$$
$$= 2\langle f, e_n\rangle\langle e_n, g\rangle - 2\langle g, e_n\rangle\langle e_n, f\rangle. \tag{34}$$

Hence,

$$4\langle f,g\rangle = \sum_n 4\langle f, e_n\rangle\langle e_n, g\rangle \tag{35}$$
$$\Leftrightarrow \quad \langle f,g\rangle = \sum_n \langle f, e_n\rangle\langle e_n, g\rangle. \tag{36}$$

□

Corollary 8. *A orthonormal system* (e_n) *is complete in a Hilbert space* $\mathcal{H}$ *if and only if for any* $f \in \mathcal{H}$*, the sequence* $\left(\sum_{i=1}^n \langle e_i, f\rangle e_i\right)_n$ *strongly converges to* f*.*

Proof. If (e_n) is a complete orthonormal system, then Parseval's identity is satisfied for any $f \in \mathcal{H}$, which implies that for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\|f\|^2 - \sum_{i=1}^n |\langle e_i, f\rangle|^2 \leq \varepsilon. \tag{37}$$

From the proof of Theorem 5, we have

$$\left\| f - \sum_{i=1}^n \langle e_i, f\rangle e_i \right\|^2 = \|f\|^2 - \sum_{i=1}^n |\langle e_i, f\rangle|^2. \tag{38}$$

Hence, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\left\| f - \sum_{i=1}^n \langle e_i, f\rangle e_i \right\|^2 \leq \varepsilon. \tag{39}$$

As a result, we write

$$f = \mathrm{s}-\lim_{n\to\infty} \sum_{i=1}^n \langle e_i, f\rangle e_i = \sum_{n=1}^\infty \langle e_n, f\rangle e_n. \tag{40}$$

The converse statement is evident by checking the definition of complete orthonormal systems.
□

We see that complete orthonormal systems have many useful properties for us to study separable Hilbert spaces. Furthermore, the existence of a complete orthonormal system and the separability of Hilbert spaces are equivalent.

Theorem 9. *A non-trivial Hilbert space* $\mathcal{H}$ $(\mathcal{H} \neq \{0\})$ *is separable if and only if there exists a complete orthonormal system* (e_n) *in* $\mathcal{H}$*.*

Proof. We divide the proof into two parts.

i) *If a non-trivial Hilbert space $\mathcal{H}$ is separable, then there exists a complete orthonormal system (e_n) in $\mathcal{H}$.*

Assume that $\mathcal{H}$ is separable. Then, there exists a dense subset $\mathcal{D} = \{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{H}$ with $\|f_1\| \neq 0$. We will use the *Gram-Schmidt process* to construct inductively an orthonormal system (e_k) from $\mathcal{D}$.

For the base case $n = 1$, we can take $e_1 = f_1/\|f_1\|$ which satisfies that $\mathrm{Vect}(e_1) = \mathrm{Vect}(f_1)$, denoted by $\mathcal{M}_1$. Now for the inductive step, suppose that for $n \geq 1$, there exists $m(n) \in \mathbb{N}$ and an orthonormal set $\{e_1, \ldots, e_{m(n)}\}$ such that $\mathrm{Vect}\left(e_1, \ldots, e_{m(n)}\right) = \mathrm{Vect}\left(f_1, \ldots, f_n\right)$, denoted by $\mathcal{M}_n$. If f_{n+1} is a linear combination of $e_1, \ldots, e_{m(n)}$ and hence, $f_{n+1} \in \mathcal{M}_n$, then we can set $m(n+1) = m(n)$, which means $\mathcal{M}_{n+1} = \mathcal{M}_n$. Otherwise, if f_{n+1} is not in $\mathcal{M}_n$, then we define

$$w_{m(n+1)} = f_{n+1} - P_{\mathcal{M}_n}(f_{n+1}) = f_{n+1} - \sum_{k=1}^{m(n)} \langle e_k, f_{n+1}\rangle e_k, \tag{41}$$

where we got the second equality from the proof of Theorem 5. Hence, $w_{m+1} \in \mathcal{M}_n^{\perp}$ and we define the normalized vector $e_{m(n+1)} = w_{m(n+1)}/\left\|w_{m(n+1)}\right\|$. Then we have

$$\mathrm{Vect}\left(e_1, \ldots, e_{m(n)}, e_{m(n+1)}\right) = \mathrm{Vect}\left(e_1, \ldots, e_{m(n)}, f_{n+1}\right) = \mathrm{Vect}\left(f_1, \ldots, f_n, f_{n+1}\right). \tag{42}$$

Therefore, we have constructed an orthonormal system (e_k) such that for any $n \in \mathbb{N}$, $\mathrm{Vect}\left(e_1, \ldots, e_{m(n)}\right) = \mathrm{Vect}\left(f_1, \ldots, f_n\right)$. Because $\mathcal{D}$ is dense, for any $f \in \mathcal{H}$ and any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|f - f_N\| \leq \varepsilon$. Because $f_N \in \mathrm{Vect}\left(e_1, \ldots, e_{m(N)}\right)$, there exist $\lambda_1, \ldots, \lambda_{m(N)} \in \mathbb{K}$ such that $f_N = \sum_{k=1}^{m(N)} \lambda_k e_k$. As a result,

$$\left\|f - \sum_{k=1}^{m(N)} \lambda_k e_k\right\| \leq \varepsilon. \tag{43}$$

Hence, the set of all linear combinations of (e_k) is also dense in $\mathcal{H}$, which implies that (e_k) is complete.

ii) *If there exists a complete orthonormal system (e_n) in a Hilbert space $\mathcal{H}$, then $\mathcal{H}$ is separable.*

Consider a subset $\mathcal{D}$ of $\mathcal{H}$ defined by

$$\mathcal{D} = \bigcup_n \left\{ \sum_{i=1}^{n} \gamma_i e_i \;\middle|\; \gamma_i \in \mathbb{K} \text{ with } \mathrm{Re}(\gamma_i), \mathrm{Im}(\gamma_i) \in \mathbb{Q} \text{ for } 1 \leq i \leq n \right\} := \bigcup_n \mathcal{D}_n \tag{44}$$

$\big($if $\mathbb{K} = \mathbb{R}$ then $\mathrm{Im}(\gamma_i) = 0\big)$. For each n, there exists a bijection between the subset $\mathcal{D}_n$ and $\mathbb{Q}^{2n}$ when $\mathbb{K} = \mathbb{C}$ or $\mathbb{Q}^n$ when $\mathbb{K} = \mathbb{R}$. Hence, $\mathcal{D}_n$ is countable for all n. Then, $\mathcal{D}$ is a countable union of countable sets, which implies that $\mathcal{D}$ is countable. Next, we will prove that $\mathcal{D}$ is also dense in $\mathcal{H}$.

Let (e_n) be a complete orthonormal system in $\mathcal{H}$. From Corollary 8, for every $f \in \mathcal{H}$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\left\|f - \sum_{i=1}^{n} \langle e_i, f\rangle e_i\right\| \leq \frac{\varepsilon}{2}. \tag{45}$$

Because $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $\gamma_i^{(n)} \in \mathbb{K}$ such that $\mathrm{Re}(\gamma_i^{(n)}), \mathrm{Im}(\gamma_i^{(n)}) \in \mathbb{Q}$ and $\left|\langle e_i, f\rangle - \gamma_i^{(n)}\right| \leq \varepsilon/2n$ for each $1 \leq i \leq n$. Then, we have

$$\left\|\sum_{i=1}^{n} \left[\langle e_i, f\rangle - \gamma_i^{(n)}\right] e_i\right\| \leq \sum_{i=1}^{n} \left|\langle e_i, f\rangle - \gamma_i^{(n)}\right| \|e_i\| \leq \sum_{i=1}^{n} \frac{\varepsilon}{2n} = \frac{\varepsilon}{2}. \tag{46}$$

Therefore, for every $f \in \mathcal{H}$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\left\| f - \sum_{i=1}^{n} \gamma_i^{(n)} e_i \right\| \leq \left\| f - \sum_{i=1}^{n} \langle e_i, f \rangle e_i \right\| + \left\| \sum_{i=1}^{n} \left[\langle e_i, f \rangle - \gamma_i^{(n)} \right] e_i \right\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{47}$$

As a result, for every $f \in \mathcal{H}$ and $\varepsilon > 0$, we can find an element $f_n = \sum_{i=1}^{n} \gamma_i^{(n)} e_i \in \mathcal{D}$ such that $\|f - f_n\| \leq \varepsilon$. Therefore, $\mathcal{D}$ is dense in $\mathcal{H}$, which implies that $\mathcal{H}$ is separable.

□

2 Some examples

In this section, given $\Omega \subset \mathbb{R}$ and an appropriate bounded measurable function $w : \Omega \to \mathbb{R}_+$, we consider the Hilbert space

$$\mathcal{H} = L_w^2(\Omega) := \left\{ f : \Omega \to \mathbb{K} \;\middle|\; \int_\Omega |f(x)|^2 w(x)\, \mathrm{d}x < \infty \right\} = \left\{ f : \Omega \to \mathbb{K} \;\middle|\; \sqrt{w} f \in L^2(\Omega) \right\} \tag{48}$$

with an inner product $\langle \cdot, \cdot \rangle$ defined for $f, g \in \mathcal{H}$ by

$$\langle f, g \rangle = \int_\Omega \overline{f(x)} g(x) w(x)\, \mathrm{d}x\,. \tag{49}$$

The function w is called the *weight function*.

2.1 Fourier series

Consider $\Omega = (-\pi, \pi)$ and $w(x) = 1$. Define the orthonormal system $(e_k)_{k \in \mathbb{Z}}$ by

$$e_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}. \tag{50}$$

The Riesz-Fischer theorem states that *Fourier series* of any $f \in L^2(-\pi, \pi)$,

$$\sum_{k \in \mathbb{Z}} \langle e_k, f \rangle e_k = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} c_k(f) e^{ikx} \tag{51}$$

with

$$c_k(f) = \langle e_k, f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-ikx}\, \mathrm{d}x\,, \tag{52}$$

strongly converges or converges in the norm $\|\cdot\|_{L^2}$ to f, i.e.

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} \left| f(x) - \frac{1}{\sqrt{2\pi}} \sum_{|k| \leq n} c_k(f) e^{ikx} \right|^2 \mathrm{d}x = 0. \tag{53}$$

Hence, the sequence (e_k) is a complete orthonormal system in $L^2(-\pi, \pi)$. Furthermore, the Carleson's theorem states that Fourier series of any function $f \in L^2(-\pi, \pi)$ converges to f almost everywhere, i.e.

$$f(x) = \lim_{n \to \infty} \frac{1}{\sqrt{2\pi}} \sum_{|k| \leq n} c_k(f) e^{ikx} \tag{54}$$

for almost every $x \in (-\pi, \pi)$. The Fourier series has greatly many applications in physics and engineering.

2.2 Classical orthogonal polynomials

In this section, we consider orthogonal systems with polynomials as their elements, which is called *orthogonal polynomials*. Those polynomials (p_n) in each system satisfy the following relation:

$$\langle p_m, p_n \rangle = \int_\Omega p_m(t) p_n(t) w(t)\,\mathrm{d}t = \delta_{mn} h_n. \tag{55}$$

Then, (p_n/h_n) is an orthonormal system. Here, we just consider *classical orthogonal polynomials*, which can be defined by a Rodrigues formula of the form:

$$p_n(t) = \frac{1}{C_n w(t)} \frac{\mathrm{d}^n}{\mathrm{d}t^n} \left(w(t) g(t)^n \right). \tag{56}$$

The information about the three classical orthogonal polynomials taken from [1] is summarized in the two tables as below.

Polynomials	Symbol	C_n	$w(t)$	$g(t)$	Ω
Jacobi	$P_n^{(\alpha,\beta)}$	$(-1)^n 2^n n!$	$(1-t)^\alpha (1+t)^\beta$	$1-t^2$	$(-1,1)$
Laguerre	$L_n^{(\alpha)}$	$n!$	$e^{-t} t^\alpha$	t	$(0,\infty)$
Hermite	H_n	$(-1)^n$	e^{-t^2}	1	$(-\infty,\infty)$

Polynomials	h_n
$P_n^{(\alpha,\beta)}$	$\dfrac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)n!\Gamma(n+\alpha+\beta+1)}$
$L_n^{(\alpha)}$	$\dfrac{\Gamma(n+\alpha+1)}{n!}$
H_n	$\sqrt{\pi} 2^n n!$

Note that Γ is the Gamma function, which is defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\,\mathrm{d}t, \qquad \text{for } x > 0. \tag{57}$$

In addition, the Jacobi polynomials have the following special cases (C_n and h_n can be different, as indicated in the below table from [1]):

(i) Gegenbauer polynomials $G_n^{(p)}$, corresponding to $\alpha = \beta = p - \frac{1}{2}$;

(ii) Chebyshev polynomials of the first kind T_n, corresponding to $\alpha = \beta = -\frac{1}{2}$;

(iii) Chebyshev polynomials of the second kind U_n, corresponding to $\alpha = \beta = \frac{1}{2}$;

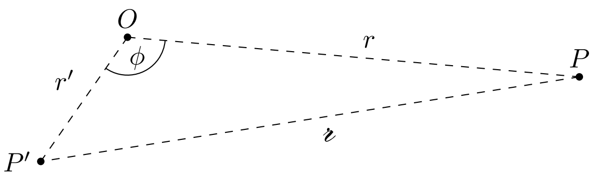
(iv) Legendre polynomials P_n, corresponding to $\alpha = \beta = 0$.

Polynomials	$\alpha = \beta$	C_n	h_n
$G_n^{(p)}$	$p - \frac{1}{2}$	$\frac{(-1)^n 2^n n! \Gamma(2p)\Gamma(n+p+1/2)}{\Gamma(p+1/2)\Gamma(n+2p)}$	$\frac{\pi 2^{1-2p}\Gamma(n+2p)}{n!(n+p)!\Gamma(p)^2}$ if $p \neq 0$
			$\frac{2\pi}{n^2}$ if $p = 0$
T_n	$-\frac{1}{2}$	$\frac{(-1)^n 2^n \Gamma(n+1/2)}{\sqrt{\pi}}$	$\frac{\pi}{2}$ if $n \neq 0$
			π if $n = 0$
U_n	$\frac{1}{2}$	$\frac{(-1)^n 2^{n+1}\Gamma(n+3/2)}{(n+1)\sqrt{\pi}}$	$\frac{\pi}{2}$
P_n	0	$(-1)^n 2^n n!$	$\frac{2}{2n+1}$

All those systems of orthogonal polynomials that we introduced are complete in their corresponding Hilbert space. Most of them come from the theory of differential equations. We shall not go into the detail of their theory but we just list out a few physical examples where they are used.

2.2.1 Multipole expansion

Both the Coulomb and Newtonian potential are proportional to $1/r$. Suppose that we have two points P and P' as in the following diagram.



We can expand $1/r$ into the following series (from [4])

$$\frac{1}{\mathscr{r}} = \frac{1}{\sqrt{r^2 + (r')^2 - 2rr'\cos\phi}} = \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r'}{r}\right)^n P_n(\cos\phi). \tag{58}$$

This expansion is called *multipole expansion*. When r is significantly large compared to r', we can keep a few of the first terms: the first term is the monopole contribution ($\sim 1/r$); the second is dipole ($\sim 1/r^2$); the third is quadrupole ($\sim 1/r^3$); the forth is octopole ($\sim 1/r^4$) and so on. This expansion is useful when we want to approximate the potential generated by a distribution of charges or masses.

2.2.2 Quantum harmonic oscillation

The time-independent Schrödinger equation for a harmonic oscillator is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2\psi(x) = E\psi(x). \tag{59}$$

Solving this equation gives us normalized eigenfunctions as follows (taken from [5]):

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar(2^n n!)^2}\right)^{1/4}\exp\left(-\frac{m\omega x^2}{2\hbar}\right)H_n\left(\left(\frac{m\omega}{\hbar}\right)^{1/2}x\right) \tag{60}$$

with corresponding energies

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \qquad \text{for } n = 0, 1, 2, 3, \dots \tag{61}$$

2.2.3 Hydrogen atom

The time-independent Schrödinger equation for Hydrogen atom is

$$-\frac{\hbar^2}{2\mu}\left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right)-\frac{e^2}{4\pi\epsilon_0 r}\psi(r,\theta,\varphi)=E\psi(r,\theta,\varphi) \tag{62}$$

where

$$\mu=\frac{m_p m_e}{m_p+m_e} \tag{63}$$

is the reduced mass with m_e, the mass of the electron, and m_p, the mass of the proton. Solving that equation, we get the eigenfunctions (taken from [6]) for $n\in\mathbb{N},\ 0\leq\ell\leq n-1$, and $-\ell\leq m\leq\ell$,

$$\psi_{n\ell m}(r,\theta,\varphi)=R_{n\ell}(r)Y_\ell^m(\theta,\varphi), \tag{64}$$

where

$$R_{n\ell}(r)=\sqrt{\left(\frac{2}{na_0'}\right)^3\frac{(n-\ell-1)!}{2n(n+\ell)!}}e^{-r/na_0'}\left(\frac{2r}{na_0'}\right)^\ell L_{n-\ell-1}^{(2\ell+1)}\left(\frac{2r}{na_0'}\right), \tag{65}$$

with the "reduced" Bohr radius

$$a_0'=\frac{4\pi\epsilon_0\hbar^2}{\mu e^2}, \tag{66}$$

and Y_ℓ^m are spherical harmonics. The associated energies of these wavefunctions are

$$E_n=-\frac{\mathrm{Ry}}{n^2},\qquad \mathrm{Ry}=\frac{\mu e^4}{32\pi^2\epsilon_0^2\hbar^2}. \tag{67}$$

Appendices

A Orthogonal projection theorem

Theorem 10 (Orthogonal projection theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{M}$ a closed subspace of $\mathcal{H}$. For any $f \in \mathcal{H}$, there is a unique vector $g \in \mathcal{M}$ such that*

$$\|f - g\| = \inf_{h \in \mathcal{M}} \|f - h\| =: d(f, \mathcal{M}). \tag{68}$$

In addition, the vector g, called the orthogonal projection of f on $\mathcal{M}$ and denoted $P_{\mathcal{M}}(f)$, is the only vector of $\mathcal{M}$ such that $f - g \in \mathcal{M}^{\perp}$.

Proof. Consider a sequence (g_n) of vectors in $\mathcal{M}$ such that

$$\lim_{n \to \infty} \|f - g_n\| = d(f, \mathcal{M}). \tag{69}$$

We want to prove that (g_n) is a Cauchy sequence. Using the identity

$$\|h_1 - h_2\|^2 + \|h_1 + h_2\|^2 = 2\|h_1\|^2 + 2\|h_2\|^2 \tag{70}$$

and plugging $h_1 = f - g_n$ and $h_2 = f - g_m$ in for any $n, m \in \mathbb{N}$, we have

$$\|g_m - g_n\|^2 = 2\|f - g_n\|^2 + 2\|f - g_m\|^2 - 4\left\|f - \frac{1}{2}(g_n + g_m)\right\|^2. \tag{71}$$

Since $\frac{1}{2}(g_n + g_m) \in \mathcal{M}$,

$$\left\|f - \frac{1}{2}(g_n + g_m)\right\|^2 \geq d(f, \mathcal{M})^2. \tag{72}$$

Furthermore, for any $\varepsilon > 0$, there exists $N(\varepsilon) \in \mathbb{N}$ such that for any $n \geq N(\varepsilon)$,

$$\|f - g_n\|^2 \leq d(f, \mathcal{M})^2 + \frac{\varepsilon^2}{4}. \tag{73}$$

Therefore, for any $n, m \geq N(\varepsilon)$, we have

$$\|g_m - g_n\|^2 \leq \varepsilon^2 \quad \Leftrightarrow \quad \|g_m - g_n\| \leq \varepsilon, \tag{74}$$

which proves that (g_n) is a Cauchy sequence in $\mathcal{M}$. Since $\mathcal{M}$ is closed (or complete), g_n strongly converges to a limit $g \in \mathcal{M}$ such that $\|f - g\| = d(f, \mathcal{M})$.

Suppose there is another $g' \in \mathcal{M}$ such that $\|f - g'\| = d(f, \mathcal{M})$. Then, we have

$$\|g' - g\|^2 = 2\|f - g\|^2 + 2\|f - g'\|^2 - 4\left\|f - \frac{1}{2}(g + g')\right\|^2. \tag{75}$$

Since $\frac{1}{2}(g + g') \in \mathcal{M}$,

$$\left\|f - \frac{1}{2}(g + g')\right\|^2 \geq d(f, \mathcal{M})^2 \tag{76}$$

and we have

$$\|g' - g\|^2 \leq 2\|f - g\|^2 + 2\|f - g'\|^2 - 4d(f, \mathcal{M})^2 = 0, \tag{77}$$

which implies that $g = g'$. Hence, g is the unique vector in $\mathcal{M}$ such that $\|f - g\| = d(f, \mathcal{M})$. Finally, to prove that g is the only vector of $\mathcal{M}$ such that $f - g \in \mathcal{M}^{\perp}$, let $h \neq 0$ be a vector of $\mathcal{M}$ and we have for $\lambda \in \mathbb{R} \setminus \{0\}$,

$$\|f - (g + \lambda h)\|^2 > d(f, \mathcal{M})^2, \tag{78}$$

where

$$\|f-(g+\lambda h)\|^2 = \langle f-g-\lambda h, f-g-\lambda h\rangle \tag{79}$$
$$= \langle f-g, f-g\rangle - \lambda\langle f-g, h\rangle - \lambda\langle h, f-g\rangle + \lambda^2\langle h, h\rangle \tag{80}$$
$$= \|f-g\|^2 - \lambda\left(\langle f-g, h\rangle + \overline{\langle f-g, h\rangle}\right) + \lambda^2\|h\|^2 \tag{81}$$
$$= d(f,\mathcal{M})^2 - 2\lambda\,\mathrm{Re}(\langle f-g, h\rangle) + \lambda^2\|h\|^2. \tag{82}$$

Hence,

$$-2\lambda\,\mathrm{Re}(\langle f-g, h\rangle) + \lambda^2\|h\|^2 > 0, \tag{83}$$

which leads to a contradiction with a suitable value of λ if $\mathrm{Re}(\langle f-g, h\rangle) \neq 0$. Hence, $\mathrm{Re}(\langle f-g, h\rangle) = 0$. Doing the same calculation when replacing h with ih, we get

$$2\lambda\,\mathrm{Im}(\langle f-g, h\rangle) + \lambda^2\|h\|^2 > 0. \tag{84}$$

With the same argument, we get $\mathrm{Im}(\langle f-g, h\rangle) = 0$. Thus, $\langle f-g, h\rangle = 0$. Because h was chosen arbitrarily, $f-g \in \mathcal{M}^\perp$. If there exist $f_1 \in \mathcal{M}$ and $f_2 \in \mathcal{M}^\perp$ such that $f = f_1 + f_2$, then

$$0 = \|f_1 + f_2 - (g + f - g)\|^2 = \|(f_1 - g) + (f_2 - (f-g))\|^2 \tag{85}$$
$$= \|f_1 - g\|^2 + \|f_2 - (f-g)\|^2, \tag{86}$$

because $\langle f_1 - g, f_2 - (f-g)\rangle = 0$, where $f_1 - g \in \mathcal{M}$ and $f_2 - (f-g) \in \mathcal{M}^\perp$. Therefore, $f_1 = g = P_{\mathcal{M}}(f)$ and $f_2 = f - g = f - P_{\mathcal{M}}(f)$. □

References

[1] N. Boccara, Functional Analysis: An Introduction for Physicists

[2] E. Kreyszig, Introductory Functional Analysis with Applications

[3] MIT OpenCourseWare, Introduction to Functional Analysis

[4] D. J. Griffiths, Introduction to Electrodynamics, 4^{th} Edition

[5] R. Shankar, Principles of Quantum Mechanics

[6] D. J. Griffiths, Introduction to Quantum Mechanics

[7] W. Amrein, Hilbert Space Methods in Quantum Mechanics