1. Vector spaces

The notion of a vector space should be known from linear algebra:

**Definition 1.** Let $V$ be a set endowed with two operations, the operation "addition", denoted by "+", which maps each pair $(x,y)$ in $V \times V$ into $V$, and the operation "scalar multiplication", denoted by a dot ($\cdot$), which maps each pair $(c,x)$ in $\mathbb{R} \times V$ [or $\mathbb{C} \times V$] into $V$. Thus, a scalar is a real or complex number. The set $V$ is called a real [complex] **vector space** if the addition and multiplication operations involved satisfy the following rules, for all $x$, $y$ and $z$ in $V$, and all scalars $c$, $c_1$ and $c_2$ in $\mathbb{R}$ [$\mathbb{C}$]:

(a) $x + y = y + x$;
(b) $x + (y + z) = (x + y) + z$;
(c) There is a unique zero vector $0$ in $V$ such that $x + 0 = x$;
(d) For each $x$ there exists a unique vector $-x$ in $V$ such that $x + (-x) = 0$;
(e) $1 \cdot x = x$;
(f) $(c_1c_2)x = c_1(c_2x)$;
(g) $c(x + y) = c \cdot x + c \cdot y$;
(h) $(c_1 + c_2)x = c_1x + c_2x$.

It is trivial to verify that the Euclidean space $\mathbb{R}^n$ is a real vector space. However, the notion of a vector space is much more general. For example, let $V$ be the space of all continuous functions on $\mathbb{R}$, with pointwise addition and scalar multiplication defined the same way as for real numbers. Then it is easy to verify that this space is a real vector space.

---

1 In the sequel, $x + (-y)$ will be denoted by $x - y$.  

---
Another (but weird) example of a vector space is the space $V$ of positive real numbers endowed with the "addition" operation $x + y = x \cdot y$ and the "scalar multiplication" $c \cdot x = x^c$. In this case the null vector 0 is the number 1, and $-x = 1/x$.

**Definition 2.** A subspace $V_0$ of a vector space $V$ is a non-empty subset of $V$ which satisfies the following two requirements:

(a) For any pair $x, y$ in $V_0$, $x + y$ is in $V_0$;
(b) For any $x$ in $V_0$ and any scalar $c$, $c \cdot x$ is in $V_0$.

Thus, a subspace $V_0$ of a vector space is closed under linear combinations: any linear combination of elements in $V_0$ is an element of $V_0$.

It is not hard to verify that a subspace of a vector space is a vector space itself, because the rules (a) through (h) in Definition 1 are inherited from the "host" vector space $V$. In particular, any subspace contains the null vector 0, as follows from part (b) of Definition 2 with $c = 0$.

**Definition 3.** An inner product on a real vector space $V$ is a real function $\langle x, y \rangle: V \times V \to \mathbb{R}$ such that for all $x, y, z$ in $V$ and all $c$ in $\mathbb{R}$,

(1) $\langle x, y \rangle = \langle y, x \rangle$
(2) $\langle cx, y \rangle = c \langle x, y \rangle$
(3) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
(4) $\langle x, x \rangle > 0$ when $x \neq 0$.

An inner product on a complex vector space is defined similarly. The inner product is then complex-valued, $\langle x, y \rangle: V \times V \to \mathbb{C}$. Condition (1) then becomes

(1*) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

and (2) now holds for all complex numbers $c$. Note that in both cases, $\langle x, x \rangle$ is real valued.

Finally, the norm of $x$ in $V$ is defined as $\|x\| = \sqrt{\langle x, x \rangle}$.

---

2 The bar denotes the complex conjugate: for $z = a + i \cdot b$, $\overline{z} = a - i \cdot b$. 
For example, in the space $C[0,1]$ of continuous real functions on $[0,1]$, the integral $\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt$ is an inner product. Moreover, in the vector space of zero-mean random variables with finite second moments the expectation $\langle X, Y \rangle = E[XY]$ is an inner product.

As is well-known from linear algebra, for vectors $x, y \in \mathbb{R}^n$, $|x^Ty| \leq ||x|| \cdot ||y||$, which is known as the Cauchy-Schwarz inequality. This inequality carries over to general inner products:

**Theorem 1. (Cauchy-Schwarz inequality)** $|\langle x, y \rangle| \leq ||x|| \cdot ||y||$.

**Proof:** Let the vector space involved be real. Then for any real $\lambda$,

$$0 \leq \langle x+\lambda y, x+\lambda y \rangle = \langle x, x+\lambda y \rangle + \lambda \langle y, x+\lambda y \rangle = \langle x, x \rangle + 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle$$

$$= ||x||^2 + 2\lambda \langle x, y \rangle + \lambda^2 ||y||^2.$$  

Minimizing the latter to $\lambda$ yields the result. The complex case is similar. Q.E.D.

Given the norm $||x|| = \sqrt{\langle x, x \rangle}$, the following properties hold:

1. $||x|| > 0$ if $x \neq 0$;  
2. $||c \cdot x|| = |c| \cdot ||x||$;  
3. $||x+y|| \leq ||x|| + ||y||$. [Triangular inequality]

The properties (1) and (2) follow trivially from Definition 3. In the case of a real vector space the triangular inequality (3) follows from

$$||x+y||^2 = \langle x+y, x+y \rangle = \langle x, x+y \rangle + \langle y, x+y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$$

$$= ||x||^2 + 2\langle x, y \rangle + ||y||^2 \leq ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2$$

$$= (||x|| + ||y||)^2$$

where the last inequality is due to Theorem 1.

A norm can also be defined directly:

**Definition 4.** A norm on a vector space $V$ is a mapping $||.||: V \to [0,\infty)$ such that for all $x$ and $y$ in $V$ and all scalars $c$, (1), (2) and (3) hold. A vector space endowed with a norm is called a normed
A norm \( \| \cdot \| \) defines a metric \( d(x,y) = \| x-y \| \) on \( V \), i.e., a function that measures the “distance” between two elements \( x \) and \( y \) of \( V \), for which (trivially) the following four properties hold. For all \( x, y \) and \( z \) in \( V \),

\[
\begin{align*}
    d(x,y) &= d(y,x); \\ 
    d(x,y) &> 0 \text{ if } x \neq y; \\ 
    d(x,x) &= 0; \\ 
    d(x,z) &\leq d(x,y) + d(y,z) \text{ [Triangular inequality].}
\end{align*}
\]

More generally,

**Definition 5.** A metric on a space \( V \) is a mapping \( d(.,.): V\times V \rightarrow [0,\infty) \) satisfying the properties (4) through (7) for all \( x, y \) and \( z \) in \( V \). A space endowed with a metric is called a metric space.

In this definition the space \( V \) is not necessarily a vector space: Any space endowed with a metric is a metric space. Moreover, inner products and norms may not be defined on metric spaces.

**Definition 6.** A sequence of elements \( x_n \) of a metric space with metric \( d(.,.) \) is called a Cauchy sequence if for every \( \varepsilon > 0 \) there exists an \( n_0 \) such that for all \( k,m \geq n_0 \), \( d(x_k,x_m) < \varepsilon \).

The notion of a Cauchy sequence plays a critical role in defining Hilbert spaces. See the next section.

**Theorem 2.** Every Cauchy sequence in \( \mathbb{R}^\ell \) or \( \mathbb{C}^\ell \), \( \ell < \infty \), has a limit in the space involved.

**Proof:** Consider first the case \( \mathbb{R} \). Let \( \overline{x} = \limsup_{n \to \infty} x_n \), where \( x_n \) is a Cauchy sequence. I will show first that \( \overline{x} < \infty \). There exists a subsequence \( n_k \) such that \( \overline{x} = \lim_{k \to \infty} x_{n_k} \). Note that \( x_{n_k} \) is also a Cauchy sequence. For arbitrary \( \varepsilon > 0 \) there exists an index \( k_0 \) such that \( |x_{n_k} - x_{n_m}| < \varepsilon \) if
Keeping $k$ fixed and letting $m \to \infty$ it follows that $|x_{n_k} - \bar{x}| < \varepsilon$, hence $\bar{x} \in \mathbb{D}$. Similarly, 

$\bar{x} = \liminf_{n \to \infty} x_n > -\infty$. Now we can find an index $k_0$ and subsequences $n_k$ and $n_m$ such that for $k,m \geq k_0$, $|x_{n_k} - \bar{x}| < \varepsilon$, $|x_{n_k} - \bar{x}| < \varepsilon$, and $|x_{n_k} - x_{n_m}| < \varepsilon$, hence $|x - \bar{x}| < 3\varepsilon$. Since $\varepsilon$ is arbitrary, we must have $\bar{x} = \bar{x} = \lim_{n \to \infty} x_n$. Applying this argument to the real and imaginary parts of a complex Cauchy sequence the result for the case $\mathbb{C}$ follows, and applying the argument to each component of a (complex) vector valued Cauchy sequence the result for the cases $\mathbb{R}^l$ and $\mathbb{C}^l$ follow. Q.E.D.

2. Hilbert spaces

A Euclidean space $\mathbb{R}^n$ is a vector space endowed with the inner product $\langle x,y \rangle = x^T y$, norm $||x|| = \sqrt{x^T x} = \sqrt{\langle x,x \rangle}$ and associated metric $||x - y||$, such that every Cauchy sequence takes a limit in $\mathbb{R}^n$. This makes $\mathbb{R}^n$ a Hilbert space:

Definition 7. A Hilbert space $H$ is a vector space endowed with an inner product and associated norm and metric, such that every Cauchy sequence in $H$ has a limit in $H$.

A Hilbert space is also a Banach space:

Definition 8. A Banach space $B$ is a normed space with associated metric $d(x,y) = ||x - y||$ such that every Cauchy sequence in $B$ has a limit in $B$.

The difference between a Banach space and a Hilbert space is the source of the norm. In the Hilbert space case the norm is defined via the inner product, $||x|| = \sqrt{\langle x,x \rangle}$, whereas in the Banach space case the norm is defined directly, by Definition 4. Thus, a Hilbert space is a Banach space, but the other way around may not be true, because in some cases the norm cannot be associated with an inner product.

An example of a Hilbert space is the space $L^2(a,b)$:
Definition 9. The space $L^2(a,b)$ is the collection of Borel measurable real or complex valued square integrable functions $f$ on $(a,b)$, i.e., $\int_a^b |f(t)|^2\,dt < \infty$, endowed with inner product 

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)}\,dt,$$

and associated norm and metric

$$\|f\| = \sqrt{\int_a^b |f(t)|^2\,dt}, \quad d(f,g) = \|f-g\| = \sqrt{\int_a^b |f(t)-g(t)|^2\,dt},$$

respectively, where the integrals involved are Lebesgue integrals.

Note that in this case $f$ and $g$ are interpreted as being equal if they differ on $(a,b)$ only on a set with Lebesgue measure zero. The proof that $L^2(a,b)$ is a Hilbert space will be given in Section 6 below.

Recall that vectors $x$ and $y$ in $\mathbb{R}^n$ are orthogonal if $x^T y = 0$. More generally,

Definition 10. Elements $x, y$ of a Hilbert space are orthogonal if $\langle x, y \rangle = 0$, also denoted by $x \perp y$, and orthonormal if in addition $\|x\| = 1$ and $\|y\| = 1$.

Note that in a Banach space orthogonality is a non-existing property, because an inner product is not defined on a Banach space. This is the main reason for working with Hilbert spaces.

Definition 11. An orthonormal sequence $e_n, n =1,2,3,...$ in a Hilbert space $H$ is complete if the only member of $H$ which is orthogonal to all $e_n$ is the zero vector.

Theorem 3. Let $e_n, n =1,2,3,...$ be a complete orthonormal sequence in a Hilbert space $H$. Then for every $x$ in $H$, $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ and $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2$.

Proof: Observe that

$$0 \leq \|x - \sum_{n=1}^{m} \langle x, e_n \rangle e_n \|^2 = \left\langle x - \sum_{n=1}^{m} \langle x, e_n \rangle e_n, x - \sum_{n=1}^{m} \langle x, e_n \rangle e_n \right\rangle$$

$$= \left\langle x, x - \sum_{n=1}^{m} \langle x, e_n \rangle e_n \right\rangle - \sum_{n=1}^{m} \langle x, e_n \rangle \left\langle e_n, x - \sum_{n=1}^{m} \langle x, e_n \rangle e_n \right\rangle$$

$$= \|x\|^2 - \sum_{n=1}^{m} |\langle x, e_n \rangle|^2$$
hence, letting $m \to \infty$, we have $\sum_{n=1}^{\infty} |<x,e_n>|^2 \leq \|x\|^2 < \infty$. Therefore,

$$\lim_{m \to \infty} \sum_{n=m}^{\infty} |<x,e_n>|^2 = 0$$

(9)

Let $y_m = \sum_{n=1}^{m} <x,e_n>e_n$ and note that

$$\|x - y_m\|^2 = \min_{\beta_1,\ldots,\beta_m} \|x - \sum_{n=1}^{m} \beta_n e_n\|^2.$$  

(10)

Thus, we can write

$$x = y_m + z_m = \sum_{n=1}^{m} <x,e_n>e_n + z_m,$$

where $<z_m,e_n> = 0$ for all $n \leq m$,

(11)

hence $<y_m,z_m> = 0$ and therefore

$$\|x\|^2 = \|y_m\|^2 + \|z_m\|^2.$$  

(12)

Moreover,

$$\|y_m - y_k\|^2 = \sum_{n=\min(k,m)}^{\max(k,m)} |<x,e_n>|^2 \leq \sum_{n=\min(k,m)}^{\infty} |<x,e_n>|^2 - 0$$

(13)

for $\min(k,m) \to \infty$, hence it follows from (9) that $y_m$ is a Cauchy sequence. Therefore, it follows from the properties of a Hilbert space that $y_n$ converges to a limit $y \in H$:

$$\lim_{m \to \infty} \|y - y_m\| = \lim_{m \to \infty} \|y - \sum_{n=1}^{m} <x,e_n>e_n\| = 0.$$  

(14)

Thus, we can write $y$ as

$$y = \sum_{n=1}^{m} <x,e_n>e_n + u_m = y_m + u_m,$$  

where $\lim_{m \to \infty} \|u_m\| = 0$.  

(15)

This result gives rise to the notation

$$y = \sum_{n=m}^{\infty} <x,e_n>e_n.$$  

(16)

However, keep in mind that the actual definition of $y$ is given by (15). On the other hand, without loss of generality we may treat $y$ as the right-hand side of (16), as will be demonstrated below.

Let $z = x - y$. If $<z,e_k> = 0$ for all $k$ then it follows from the completeness of $\{e_k\}$ that $z = 0$, hence $x = y$. To prove this, note that by (15), $z = x - \sum_{n=1}^{m} <x,e_n>e_n - u_m$ for all $m$. Then for $m > k$,

$$<z,e_k> = <x,e_k> - \sum_{n=1}^{m} <x,e_n><e_k,e_n> - <e_k,u_m>$$

$$= <x,e_k> - <e_k><e_k,e_k> - <e_k,u_m>$$

$$= <x,e_k> - <e_k,e_k> - <e_k,u_m>$$

(17)

hence by Theorem 1 and (15),

$$|<z,e_k>| = \lim_{m \to \infty} |<e_k,u_m>| \leq \lim_{m \to \infty} \|u_m\| \|e_k\| = \lim_{m \to \infty} \|u_m\| = 0.$$  

(18)

Thus, $x = y$. It follows now from (11) and (15) that $z_m = u_m$, hence
\[ x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n + u_m, \text{ where } \langle e_n, u_m \rangle = 0 \text{ for all } n \leq m, \text{ and } \lim_{m \to \infty} ||u_m|| = 0. \]  

(19)

Again, this result is denoted by

\[ x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n. \]  

(20)

This proves the first part of the theorem. The second part follows now easily from (8). Q.E.D.

Note that the result (18) could have been obtained more directly by treating \( y \) as the right-hand side of (16), because then \( z = x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \), hence

\[ \langle z, e_k \rangle = \langle x, e_k \rangle - \sum_{n=1}^{\infty} \langle x, e_n \rangle \langle e_n, e_k \rangle = \langle x, e_k \rangle - \langle x, e_k \rangle = \langle x, e_k \rangle - \langle x, e_k \rangle = 0. \]

Similarly, (19) implies that for any \( \nu \in H \),

\[ \langle x, \nu \rangle = \lim_{m \to \infty} \sum_{n=1}^{m} \langle x, e_n \rangle \langle e_n, \nu \rangle + \lim_{m \to \infty} \langle u_m, \nu \rangle = \sum_{n=1}^{\infty} \langle x, e_n \rangle \langle e_n, \nu \rangle + \lim_{m \to \infty} \langle u_m, \nu \rangle = \sum_{n=1}^{\infty} \langle x, e_n \rangle \langle e_n, \nu \rangle \]

(21)

where the last equality follows from the fact that by Theorem 1 and (19),

\[ \lim_{m \to \infty} \langle u_m, \nu \rangle \leq \lim_{m \to \infty} ||u_m|| ||\nu|| = 0, \]

whereas the result (21) would have followed trivially if we have used (20).

**Definition 12.** The coefficients \( \langle x, e_n \rangle \) in the series representation \( x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \) are called the Fourier coefficients of \( x \).

**Definition 13.** Let \( A \) be a set of vectors in a Hilbert space \( H \). Then \( \text{Lin}(A) \) is the intersection of all sub-spaces of \( H \) which contain \( A \), and \( \text{Clin}(A) \) is the closure of \( \text{Lin}(A) \).

In other words, \( \text{Lin}(A) \) is the set of all linear combinations of the elements of \( A \). If \( A \) is a finite set then \( \text{Clin}(A) = \text{Lin}(A) \). The same applies if \( A \subset \mathbb{R}^n \), because there are only a finite number of linear independent vectors in \( A \), which can be made orthonormal.

**Theorem 4.** Let \( e_n, n = 1, 2, 3, \ldots \) be an orthonormal sequence in a Hilbert space \( H \). Then the following statements are equivalent.

(1) \( e_n \) is complete;

(2) \( H = \text{Clin}(\{e_n, n \geq 1\}) \).
(3) For all $x$ in $H$, $||x||^2 = \sum_{n=-1}^{\infty} |\langle x, e_n \rangle|^2$.

Proof: (1) $\Rightarrow$ (2) and (1) $\Rightarrow$ (3) has already been proved in Theorem 3.

Proof of (3) $\Rightarrow$ (1): Suppose that the sequence $e_n$ is not complete. Then there exists a $z \neq 0$ in $H$ such that $\langle z, e_n \rangle = 0$ for all $n$. But then by part (3), $||z||^2 = \sum_{n=-1}^{\infty} |\langle z, e_n \rangle|^2 = 0$, which contradicts $z \neq 0$.

Proof of (2) $\Rightarrow$ (1): Let $\langle z, e_n \rangle = 0$ for all $n$, and define $E = \{ x \in H: \langle x, z \rangle = 0 \}$. Then $E$ contains all $e_n$, and for fixed $z$, $E$ is a subspace, so that $\text{Lin} \{ e_n, n \geq 1 \} \subset E$. Moreover, it is not hard to verify from the continuity of the inner product that $E$ is closed, hence $H = \text{Clin} \{ e_n, n \geq 1 \} \subset E$. But $z \in H$, and thus $z \in E$, and consequently, $\langle z, z \rangle = 0$. This proves (1).

Q.E.D.

3. Fourier analysis

The following theorem implies that for every Borel measurable real or complex valued function $f$ in $L^2(-\pi, \pi)$ has the series expansion

$$f(x) = (2\pi)^{-1} \sum_{n=-\infty}^{\infty} c_n \exp(i.n.x), \text{ where } c_n = \int_{-\pi}^{\pi} f(y) \exp(-i.n.y)dy.$$  \hspace{1cm} (22)

Recall from the proof of Theorem 3 that (22) should be interpreted as:

For all $m > 0$, $f(x) = (2\pi)^{-1} \sum_{n=-m}^{m} c_n \exp(i.n.x) + e_m(x)$, where

$$\lim_{m \to \infty} \int_{-\pi}^{\pi} |e_m(x)|^2dx = 0.$$  \hspace{1cm} (23)

Theorem 5. The complex functions $e_n(x) = (2\pi)^{-\frac{1}{2}} \exp(i.n.x)$, $n = 0, \pm 1, \pm 2, \ldots.$ form a complete orthonormal sequence in $L^2(-\pi, \pi)$.

The orthonormality of the sequence $\{e_n(x)\}$ follows from the fact that
The rest of the proof of Theorem 5 is too long and is therefore given in the Appendix, Section A.1.

Since every function \( g \) in \( L^2(a,b) \) can be converted to a function \( f \) in \( L^2(-\pi,\pi) \), namely \( f(x) = g(a + (b-a)(x+\pi)/(2\pi)) \) and vice versa \( g(x) = f(\pi - 2\pi(b-x)/(b-a)) \), it follows from (22) that

\[
g(x) = (2\pi)^{-1/2}\sum_{n=-\infty}^{\infty} c_n \exp\left[i.n.(\pi - 2\pi(b-x)/(b-a))\right]
\]

where

\[
c_n = (2\pi)^{-1/2} \int_{-\pi}^{\pi} f(y) \exp(-i.n.y) dy = (2\pi)^{-1/2} \int_{-\pi}^{\pi} g(a + (b-a)(y+\pi)/(2\pi)) \exp(-i.n.y) dy
\]

\[
= \frac{\sqrt{2\pi}}{b-a} \int_{a}^{b} g(x) \exp\left[-i.n.(\pi - 2\pi(b-x)/(b-a))\right] dx
\]

Therefore,

\textbf{Theorem 6. Every function} \( g \) \textit{in} \( L^2(a,b) \), \( -\infty < a < b < \infty \), \textit{can be written as}

\[
g(x) = \sum_{n=-\infty}^{\infty} \gamma_n \exp\left[2\pi n.x/(b-a)\right]
\]

\textit{where} \( \gamma_n = (b-a)^{-1} \int_{a}^{b} g(x) \exp\left[-2\pi n.x/(b-a)\right] dx \).

If \( g \) \textit{is real valued}, then

\[
\gamma_n = (b-a)^{-1} \int_{a}^{b} g(x) \cos\left[2\pi n.x/(b-a)\right] dx - i(b-a)^{-1} \int_{a}^{b} g(x) \sin\left[2\pi n.x/(b-a)\right] dx
\]

\[
= \alpha_n - i\beta_n,
\]

\textit{say, so that the result of Theorem 6 reads}

\[
g(x) = \gamma_0 + \sum_{n=1}^{\infty} \gamma_n \exp\left[2\pi n.x/(b-a)\right] + \sum_{n=1}^{\infty} \overline{\gamma}_n \exp\left[-2\pi n.x/(b-a)\right]
\]

\[
= \alpha_0 + 2 \sum_{n=1}^{\infty} \alpha_n \cos\left[2\pi n.x/(b-a)\right] + 2 \sum_{n=1}^{\infty} \beta_n \sin\left[2\pi n.x/(b-a)\right] \textit{ a.e. on } (a,b).
\]
4. Functions of two or more variables

Let \( g(x_1, x_2) \) be a function in \( L^2((a_1, b_1) \times (a_2, b_2)), \, -\infty < a_j < b_j < \infty, \, j = 1, 2 \). Since
\[
\int_{a_1}^{b_1} \int_{a_2}^{b_2} |g(x_1, x_2)|^2 \, dx_1 \, dx_2 < \infty,
\]
the set \( N_2 = \{x_2 \in (a_2, b_2) : \int_{a_1}^{b_1} |g(x_1, x_2)|^2 \, dx_1 = \infty\} \) has Lebesgue measure zero. Consequently, for every fixed \( x_2 \in (a_2, b_2) \setminus N_2 \) the function \( g(x_1, x_2) \) is a member of \( L^2(a_1, b_1) \). Applying Theorem 6 then yields
\[
g(x_1, x_2) = \sum_{n=\infty}^{\infty} \gamma_n(x_2) \exp\left[2\pi n i x_1/(b_1 - a_1)\right] \text{ a.e. on } (a_1, b_1),
\]
where
\[
\gamma_n(x_2) = (b_1 - a_1)^{-1} \int_{a_1}^{b_1} g(x_1, x_2) \exp\left[-2\pi n i x_1/(b_1 - a_1)\right] \, dx_1.
\]
But (29) is a member of \( L^2(a_2, b_2) \), so that by Theorem 6,
\[
\gamma_n(x_2) = \sum_{n=\infty}^{\infty} \gamma_{n,m} \exp\left[2\pi n i x_2/(b_2 - a_2)\right] \text{ a.e. on } (a_2, b_2),
\]
where
\[
\gamma_{n,m} = (b_1 - a_1)^{-1}(b_2 - a_2)^{-1} \int_{a_1}^{b_1} \int_{a_2}^{b_2} g(x_1, x_2) \exp\left[-2\pi n i x_1/(b_1 - a_1)\right] \exp\left[-2\pi m i x_2/(b_2 - a_2)\right] \, dx_1 \, dx_2.
\]
Substituting (30) in (28) yields
\[
g(x_1, x_2) = \sum_{n=\infty}^{\infty} \sum_{m=\infty}^{\infty} \gamma_{n,m} \exp\left[2\pi n i x_1/(b_1 - a_1)\right] \exp\left[2\pi m i x_2/(b_2 - a_2)\right].
\]
This result can be extended to functions of more than two variables:

**Theorem 7.** Let \( C = \times_{j=1}^{k} (a_j, b_j), \, -\infty < a_j < b_j < \infty \), let \( D \) be the diagonal matrix with diagonal elements \( b_1 - a_1, \ldots, b_k - a_k \), and denote by \( \mathbb{I}^k \) the set of \( k \)-dimensional vectors with integer components. Then for every function \( g(x) \) in \( L^2(C) \), \( g(x) = \sum_{n \in \mathbb{I}^k} \gamma(n) \exp[2\pi i n^T D^{-1} x] \)
where for \( n \in \mathbb{I}^k \), \( \gamma(n) = \det[D^{-1}] \int_C g(y) \exp[-2\pi i n^T D^{-1} y] \, dy \).

Again, we can write \( \gamma(n) = \alpha(n) - i \beta(n) \), where
\[
\alpha(n) = \det[D^{-1}] \int_C g(y) \cos[2\pi n^T D^{-1} y] \, dy,
\]
\[
\beta(n) = \det[D^{-1}] \int_C g(y) \sin[2\pi n^T D^{-1} y] \, dy,
\]
so that if \( g \) is real valued,
$$g(x) = \sum_{n \in \mathbb{Z}} (\alpha(n) - i \beta(n)) [\cos(2\pi n T D^{-1} x) + i \sin(2\pi n T D^{-1} x)]$$

$$= \alpha(0) + \sum_{n \in \mathbb{N}, \{0\}} \alpha(n) \cos(2\pi n T D^{-1} x) + \sum_{n \in \mathbb{N}, \{0\}} \beta(n) \sin(2\pi n T D^{-1} x).$$

(34)

5. Hilbert spaces of random variables

As mentioned before, for zero-mean random variables $X$ and $Y$ with finite second moments the expectation $E[XY]$ can be interpreted as an inner product:

**Theorem 8.** Let $H$ be the space of zero-mean random variables with finite second moments defined on a common probability space $\{\Omega, \mathcal{F}, P\}$, endowed with the inner product $\langle X, Y \rangle = E[XY]$, norm $\|X\| = \sqrt{E[X^2]}$ and metric $\|X-Y\|$. Then $H$ is a Hilbert space.

**Proof:** It is trivial to verify that the space of these random variables is a vector space. Therefore, we only need to show that every Cauchy sequence $X_n, n \geq 1$, has a limit in $H$, as follows. Since by Chebishev’s inequality,

$$P(|X_n - X_m| > \varepsilon) \leq E[(X_n - X_m)^2] / \varepsilon^2 = \|X_n - X_m\|^2 / \varepsilon^2 \to 0 \text{ as } n,m \to \infty$$

for every $\varepsilon > 0$, it follows that $|X_n - X_m| \to 0$ in probability as $n,m \to \infty$. Therefore, there exists a subsequence $n_k$ such that $|X_{n_k} - X_{m_k}| \to 0$ a.s. as $n,m \to \infty$. The latter implies that there exists a null set $N$ such that for every $\omega \in \Omega \setminus N$, $X_{n_k}(\omega)$ is a Cauchy sequence in $\mathbb{R}$, hence

$$\lim_{k \to \infty} X_{n_k}(\omega) = X(\omega) \text{ exists for every } \omega \in \Omega \setminus N.$$  

Now for every fixed $m$,

$$(X_{n_k} - X_m)^2 \to (X-X_m)^2 \text{ a.s. as } k \to \infty.$$  

By Fatou’s lemma (see the Appendix, Section A.2) and the Cauchy property the latter implies that $\|X-X_m\|^2 = E[(X-X_m)^2] \leq \liminf_{k \to \infty} E[(X_{n_k} - X_m)^2] \to 0$ as $m \to \infty$. Moreover, it is easy to verify that $E[X] = 0$ and $E[X^2] < \infty$. Q.E.D.

Note that the result of Theorem 8 can be translated in the following way:

---

3 This follows from the fact that if $X_n \to 0$ in probability then every subsequence $n_k$ contains a further subsequence $n_k(m)$ such that $X_{n_k(m)} \to 0$ a.s. as $m \to \infty$.  

---
Corollary 1. Let \( \{\Omega, \mathcal{F}, P\} \) be a probability space, and let \( L_0^2(P) \) be the space of measurable functions \( X(\omega): \Omega \to \mathbb{R} \) satisfying \( \int X(\omega)^2 dP(\omega) < \infty \), \( \int X(\omega) dP(\omega) = 0 \), endowed with the inner product \( \langle X, Y \rangle = \int X(\omega)Y(\omega)dP(\omega) \) and associated norm \( \|X\| = \sqrt{\langle X, X \rangle} \) and metric \( \|X - Y\| \). Then \( L_0^2(P) \) is a Hilbert space.

Moreover, if we take \( \Omega = \mathbb{R} \), \( \mathcal{F} = \mathcal{B} \) and \( P = \mu \), where \( \mathcal{B} \) is the Euclidean Borel field and \( \mu \) is a probability measure on \( \mathcal{B} \), then Corollary 1 reads:

Corollary 2. The space \( L_0^2(\mu) \) of Borel measurable real functions on \( \mathbb{R} \) satisfying \( \int f(x)^2 d\mu(x) < \infty \) and \( \int f(x)d\mu(x) = 0 \), endowed with the inner product \( \langle f, g \rangle = \int f(x)g(x)d\mu(x) \) and associated norm \( \|f\| = \sqrt{\langle f, f \rangle} \) and metric \( \|f - g\| \), is a Hilbert space.

6. The space \( L^2(a,b) \)

The results in Section 5 can be used to prove that the space \( L^2(0,1) \) is a Hilbert space, which then implies that \( L^2(a,b) \) is a Hilbert space. First, consider the subspace \( L_0^2(0,1) \) of Borel measurable functions \( f \) in \( L^2(0,1) \) satisfying \( \int_0^1 f(x)dx = 0 \). It follows straightforwardly from Corollary 2 that \( L_0^2(0,1) \) is a Hilbert space. Now let \( f_n \) be a Cauchy sequence in \( L_0^2(0,1) \). Then \( g_n(x) = f_n(x) - \int_0^1 f_n(u)du \) is a Cauchy sequence in \( L_0^2(0,1) \): there exists a \( g \) in \( L_0^2(0,1) \) such that \( \lim_{n \to \infty} \|f_n - g_n\| = 0 \) because

\[
\left| \int_0^1 f_k(u)du - \int_0^1 f_m(u)du \right| \leq \int_0^1 |f_k(u) - f_m(u)|du \leq \sqrt{\int_0^1 |f_k(u) - f_m(u)|^2du} = \|f_k - f_m\|, \quad (35)
\]

hence \( \|g_k - g_m\| \leq 2\|f_k - f_m\| \). But inequality (35) also implies that \( \int_0^1 f_n(u)du \) is a Cauchy sequence in the Hilbert space \( \mathbb{R} \), and therefore

\( \mu = \lim_{n \to \infty} \int_0^1 f_n(u)du \in \mathbb{R} \).

Next, let \( f = g + \mu \). Then \( f \in L^2(0,1) \) because \( g \in L_0^2(0,1) \subset L^2(0,1) \) and the constant function \( \mu \) is a member of \( L^2(0,1) \). Moreover,

---

Recall that \( X(\omega): \Omega \to \mathbb{R} \) is measurable if for all Borel sets \( B \) the sets \( \{\omega \in \Omega: X(\omega) \in B\} \) are members of \( \mathcal{F} \).
Thus, $L^2(0,1)$ is a Hilbert space.

Along the same lines it is easy to show that:

**Theorem 9.** The space $L^2(a,b)$ is a Hilbert space,

and more generally that:

**Theorem 10.** The space $L^2(\mu)$ of Borel measurable real functions on $\mathbb{R}$ satisfying $\int f(x)^2 \, d\mu(x) < \infty$, endowed with the inner product $\langle f, g \rangle = \int f(x)g(x) \, d\mu(x)$ and associated norm $\|f\| = \sqrt{\langle f, f \rangle}$ and metric $\|f - g\|$, is a Hilbert space.

**Appendix**

**A.1 Proof of Theorem 5**

The proof of Theorem 5 involves the following two steps:

1. Show that any continuous $2\pi$-periodic real function $f$ on $[-\pi, \pi]$, confined to $(-\pi,\pi)$, is contained in $\text{Clin}\{e_n, \ n = 0, \pm 1, \pm 2, \ldots \}$. Then the same applies to complex continuous $2\pi$-periodic functions on $[-\pi,\pi]$, by applying the argument to the real and imaginary parts of $f$.

2. Show that the subset $C_p [-\pi,\pi]$ of continuous $2\pi$-periodic functions on $[-\pi,\pi]$, confined to $(-\pi,\pi)$, is a dense subset of $L^2(-\pi,\pi)$, i.e., $L^2(-\pi,\pi)$ is the closure of $L^2(-\pi,\pi) \cap C_p [-\pi,\pi]$. Then the result follows from Theorem 4.

**Step 1.** Let $f$ be a continuous real function on $[-\pi,\pi]$, and let $f_m = \sum_{n=-m}^{m} <f, e_n> e_n$ and

$$F_m = (m+1)^{-1} \sum_{n=0}^{m} f_n.$$ 

Clearly, $F_m \in \text{Lin}\{e_n, \ n = 0, \pm 1, \pm 2, \ldots \}$. Therefore, $f \in \text{Clin}\{e_n, \ n = 0, \pm 1, \pm 2, \ldots \}$ if $F_m \to f$, so I will show that the latter is true.

Observe that $<f, e_n> = \frac{<e_n, f>}{(2\pi)^{1/2}} = (2\pi)^{-1/2} \int_{-\pi}^{\pi} f(x) \exp(-i.n.x) \, dx$, hence

$$f_m(y) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(x) \sum_{n=-m}^{m} \exp(i.n.(y-x)) \, dx$$

and thus

$$\|f_n - f\| \leq \|g_n - g\| + \left| \int_{0}^{1} f_n(u) \, du - \mu \right| \to 0$$
as $n \to \infty$. Thus, $L^2(0,1)$ is a Hilbert space.
\[ F_m(y) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(x) \left( (m+1)^{-1} \sum_{j=0}^{m} \sum_{n=-j}^{j} \exp(i.n.(y-x)) \right) dx = (2\pi)^{-1} \int_{-\pi}^{\pi} f(x) K_m(y-x) dx, \]

say, where

\[ K_m(t) = (m+1)^{-1} \sum_{j=0}^{m} \sum_{n=-j}^{j} \exp(i.n.t), \quad t \in \mathbb{R}. \]

The latter function is known as the Fejer kernel. Note that for \( t \) such that \( \exp(i.t) = 1 \),

\[ K_m(t) = (m+1)^{-1} \sum_{j=0}^{m} (1 + \sum_{n=-j}^{j} \exp(-i.n.t)) = 1 + m, \]

whereas for all \( t \) such that \( \exp(i.t) \neq 1 \),

\[ K_m(t) = (m+1)^{-1} \sum_{j=0}^{m} \left( 1 + \sum_{n=-j}^{j} \exp(i.n.t) + \sum_{n=-j}^{j} \exp(-i.n.t) \right) \]

\[ = (m+1)^{-1} \sum_{j=0}^{m} \left( 1 - \exp(i.t(1+j)) + \exp(-i.t(1+j)) - 1 \right) \]

\[ = \frac{1}{1 - \exp(i.t)} + \frac{1}{1 - \exp(-i.t)} - 1 \]

\[ - (m+1)^{-1} \frac{\exp(i.t)}{1 - \exp(i.t)} \sum_{j=0}^{m} \exp(i.t.j) - (m+1)^{-1} \frac{\exp(-i.t)}{1 - \exp(-i.t)} \sum_{j=0}^{m} \exp(-i.t.j) \]

\[ = -(m+1)^{-1} \frac{\exp(i.t)}{1 - \exp(i.t)} \frac{1 - \exp(i.t(m+1))}{1 - \exp(i.t)} - (m+1)^{-1} \frac{\exp(-i.t)}{1 - \exp(-i.t)} \frac{1 - \exp(-i.t(m+1))}{1 - \exp(-i.t)} \]

\[ = (m+1)^{-1} \frac{2 - \exp(i.t(m+1)) - \exp(-i.t(m+1))}{(1 - \exp(i.t))(1 - \exp(-i.t))} \]

\[ = (m+1)^{-1} \frac{2 - 2\cos(t(m+1))}{2 - 2\cos(t)} = \frac{1}{m+1} \left( \frac{\sin((m+1)t/2)}{\sin(t/2)} \right)^2. \]

Thus, \( K_m(t) \geq 0 \) and \( K_m(t) = K_m(-t) \). Moreover, it is not hard to show that

\[ \int_{-\pi}^{\pi} K_m(t) dt = 2\pi, \quad (36) \]

and that for any fixed \( \delta \in (0, \pi) \),

\[ \lim_{m \to \infty} \left( \int_{-\pi}^{\delta} K_m(t) dt + \int_{-\delta}^{\pi} K_m(t) dt \right) = 0. \quad (37) \]

Note that (36) implies, replacing \( t \) by \( y - x \), that

\[ \int_{-\pi}^{\pi} K_m(y-x) dx = 2\pi, \]

hence for \( y \in [-\pi, \pi] \),

\[ f(y) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(y) K_m(y-x) dx. \quad (38) \]
Since $f$ and $K_m$ are both $2\pi$-periodic, it follows for each $y, f(x)K_m(y-x)$ is $2\pi$-periodic. Since a $2\pi$-periodic function has the same integral over $[y, -\pi, y+\pi]$ as over $[-\pi, \pi]$, we have

$$F_m(y) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(x)K_m(y-x)dx = (2\pi)^{-1} \int_{-\pi}^{\pi} f(x)K_m(x-y)dx$$

(39)

Hence, it follows from (38) and (39) that for any $\delta \in (0, \pi),$

$$|F_m(y) - f(y)| \leq (2\pi)^{-1} \int_{y-\pi}^{y+\pi} |f(x) - f(y)|K_m(x-y)dx$$

$$= (2\pi)^{-1} \int_{y-\pi}^{y-\delta} |f(x) - f(y)|K_m(x-y)dx + (2\pi)^{-1} \int_{y-\delta}^{y+\delta} |f(x) - f(y)|K_m(x-y)dx$$

$$+ (2\pi)^{-1} \int_{y+\delta}^{y+\pi} |f(x) - f(y)|K_m(x-y)dx$$

(40)

Since $f$ is bounded, it follows from (37) that the first and second integral at the right hand side of (40) converge to zero as $m \to \infty$, whereas the third term can be bounded by

$$|f(x) - f(y)|K_m(x-y)dx \leq \sup_{|x-y|<\delta} |f(x) - f(y)|(2\pi)^{-1} \int_{-\pi}^{\pi} K_m(t)dt$$

(41)

$$= \sup_{|x-y|<\delta} |f(x) - f(y)|,$$

where the equality follows from (36). Since continuous functions on a compact interval are uniformly continuous, it follows that $f$ is uniformly continuous on $[-\pi, \pi]$, and since $f$ is $2\pi$-periodic it follows that $f$ is uniformly continuous on $\mathbb{R}$. It follows therefore from (40) and (41) that

$$\lim_{m \to \infty} \sup_{-\pi < y < \pi} |F_m(y) - f(y)| = 0.$$  

(42)

Note that (42) implies $\lim_{m \to \infty} \|F_m - f\| = 0$, which completes step 1 of the proof.

Step 2. First, note that every continuous real function on $[-\pi, \pi]$ can be written as a limit of $2\pi$-periodic functions on $[-\pi, \pi]$. In particular, define for a continuous real function $f$ on $[-\pi, \pi], \ldots$
Then $f_n$ is continuous on $[-\pi, \pi]$, and since $f_n(-\pi) = f_n(\pi) = 0$, it can be extended to the real line as a 2\pi-periodic function. Moreover,

$$
\|f_n - f\|^2 = \int_{-\pi}^{\pi - 1/n} (f_n(x) - f(x))^2 \, dx + \int_{\pi - 1/n}^{\pi} (f_n(x) - f(x))^2 \, dx \leq \frac{2}{n} \sup_{-\pi, x \in \pi} f(x)^2 \to 0
$$
as $n \to \infty$. A similar result holds for complex-valued continuous functions. Consequently, the closure of the space $C_p[-\pi,\pi]$ of continuous 2\pi-periodic functions on $[-\pi,\pi]$ contains all the continuous functions on $[-\pi,\pi]$. 

Next, let $B$ be an arbitrary Borel subset of $[-\pi, \pi]$, and let for $n = 1, 2, 3, \ldots$,

$$
f_n(x) = \exp(-n^{-1} \inf_{y \in \overline{B}} |x-y|) - \exp(-n^{-1} \inf_{y \in \overline{B}} |x-y|), \quad x \in [-\pi, \pi],
$$

where $\overline{B}$ is the closure of $B$. Note that $f_n(x)$ is continuous on $[-\pi, \pi]$ because $\inf_{y \in \overline{B}} |x-y|$ is a continuous function on $[-\pi, \pi]$, and so is $\inf_{y \in \overline{B}} |x-y|$. To see this, observe that for any pair $x_1, x_2 \in [-\pi, \pi]$, $\inf_{y \in \overline{B}} |x_1 - y| \leq |x_1 - x_2| + \inf_{y \in \overline{B}} |x_2 - y|$ and $\inf_{y \in \overline{B}} |x_2 - y| \leq |x_1 - x_2| + \inf_{y \in \overline{B}} |x_1 - y|$, hence

$$
\left| \inf_{y \in \overline{B}} |x_1 - y| - \inf_{y \in \overline{B}} |x_2 - y| \right| \leq |x_1 - x_2|,
$$
and similarly,

$$
\left| \inf_{y \in \overline{B}} |x_2 - y| - \inf_{y \in \overline{B}} |x_1 - y| \right| \leq |x_1 - x_2|.
$$

Since $\lim_{n \to \infty} f_n(x) = I(x \in B)$, where $I(.)$ is the indicator function, and therefore by bounded convergence, $\lim_{n \to \infty} \|f_n(x) - I(x \in B)\| = 0$, the function $I(x \in B)$ is contained in the closure of the space $C_p[-\pi,\pi]$, and consequently, all real simple functions on $[-\pi,\pi]$ are included as well. Finally, since a Borel measurable function is a limit of a sequence of simple functions, it is easy to verify from the dominated convergence theorem that all Borel measurable functions $f$ satisfying $\int_{-\pi}^{\pi} |f(t)|^2 \, dt < \infty$ are included in the closure of the space $C_p[-\pi,\pi]$. This completes step 2 of the proof. Q.E.D.
A.2 Fatou’s Lemma

Fatou’s lemma states:

For a sequence $X_n, n \geq 1$, of non-negative random variables, $E[\liminf_{n \to \infty} X_n] \leq \liminf_{n \to \infty} E[X_n]$. 

Proof: Put $X = \liminf_{n \to \infty} X_n$ and let $\varphi(x)$ be a simple function satisfying $0 \leq \varphi(x) \leq x$. Moreover, put $Y_n = \min(\varphi(X), X_n)$. Then $Y_n \to_p \varphi(X)$ because for arbitrary $\varepsilon > 0$, 

$$P[|Y_n - \varphi(X)| > \varepsilon] = P[X_n < \varphi(X) - \varepsilon] \leq P[X_n < X - \varepsilon] \to 0.$$ 

Since $E[\varphi(X)] < \infty$ because $\varphi$ is a simple function, and $Y_n \leq \varphi(X)$, it follows from $Y_n \to_p \varphi(X)$ and the dominated convergence theorem that 

$$E[\varphi(X)] = \lim_{n \to \infty} E[Y_n] = \inf_{n \to \infty} E[Y_n] \leq \liminf_{n \to \infty} E[X_n]. \quad (43)$$ 

Taking the supremum over all simple functions $\varphi$ satisfying $0 \leq \varphi(x) \leq x$ it follows now from (43) and the definition of $E[X]$ that $E[X] \leq \liminf_{n \to \infty} E[X_n]$.

Finally, note that in the case $E[\liminf_{n \to \infty} X_n] = \infty$ Fatou’s lemma states that then also $\liminf_{n \to \infty} E[X_n] = \infty$. Q.E.D.