Linear vector space

Read : Shankar ch 1
Sakurai 1.2, 1.3, 1.5
Cohen-Tanoudji, chapter 2A-2E

Definition

A linear vector space V is a collection of objects ψ_a, ψ_b ---, called vectors, which satisfy the following postulates:

1. If ψ_a and ψ_b are vectors in V, there is a unique vector $\psi_a + \psi_b$ in V, called the sum of ψ_a and ψ_b. In other words, an operation called addition is defined in the vector space such that the space is closed under addition.

2. The vector addition is commutative and associative, i.e.,

$$\psi_a + \psi_b = \psi_b + \psi_a$$

$$\psi_a + (\psi_b + \psi_c) = (\psi_a + \psi_b) + \psi_c$$

3. There is a vector in V called the null vector and denoted by ϕ satisfying

$$\psi_a + \phi = \phi + \psi_a$$

for every ψ_a in V.

4. For every vector ψ_a in V there is another vector ψ_a' in V such that

$$\psi_a + \psi_a' = \phi .$$

We denote ψ_a' as $-\psi_a$.

(Note: We use the notation $\psi_a - \psi_b$ to mean $\psi_a + (-\psi_b)$).

5. If ψ_a is a vector and λ is an arbitrary number (real or complex), called a scalar, there is a uniquely defined vector $\lambda\psi_a$ in V satisfying

(a) $\lambda(\psi_a + \psi_b) = \lambda\psi_a + \lambda\psi_b$ i.e., multiplication is distributive with respect to ~~vecto~~ vector addition.

(b) $(\lambda\mu)\psi_a = \lambda(\mu\psi_a)$ i.e., multiplication by a scalar is associative

(c) $(\lambda+\mu)\psi_a = \lambda\psi_a + \mu\psi_a$, i.e., multiplication is distributive with respect to addition scalars.

(d) Multiplication by scalars 0 and 1 are defined by

$$0\,\psi_a = \phi$$
$$1\,\psi_a = \psi_a$$

for any ψ_a in V.

~~Examples of vector spaces~~

Examples of linear vector space

1. Consider all real numbers ~~x~~ x in the range $-\infty$ to ∞, i.e.,

$$-\infty < x < \infty$$

i.e., $x \in \mathbb{R}$. ($\mathbb{R}$ is the set of all real numbers)

Take any two ~~re~~ real numbers x_1 and x_2. If we add two real numbers we get another real number in $\mathbb{R}$. Thus

$$x_1 + x_2 \in \mathbb{R}.$$

Next take any real number x. If we ~~ra~~ multiply x by another real number λ, we get a ~~another~~ real number in $\mathbb{R}$, i.e.,

$$\lambda x \in \mathbb{R}.$$

If we take a real number x, then there exists another real number $-x$ such that

$$x + (-x) = 0$$

So the real numbers form a vector space with the real number themselves as vectors in the space. The number 0 is the null vector ϕ of the space. The scalars λ by which the vectors are multiplied are also real numbers.

Thus the real numbers form a real linear vector space over a field which are also real numbers. The addition and multiplication are just the normal addition and multiplication of real numbers.

2. The set of n-tuples of numbers $(x_1, x_2, \cdots x_n)$ when the addition of vectors and multiplication by a scalar are defined by

$$(x_1\, x_2 \cdots x_n) + (y_1\, y_2 \cdots y_n) = (x_1+y_1,\, x_2+y_2, \cdots x_n+y_n)$$

and

$$\lambda(x_1\, x_2 \cdots x_n) = (\lambda x_1,\, \lambda x_2, \cdots \lambda x_n).$$

3. The collection of all square-integrable complex valued functions of a real variable form a vector space. Consider all functions

$$f : \mathbb{R} \to \mathbb{C}$$

| $\mathbb{R}$ = set of real numbers
| $\mathbb{C}$ = " " complex "

such that

$$\int_{-\infty}^{\infty} f^*(x) f(x)\, dx \equiv \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \quad (\text{i.e., finite}).$$

The sum of two functions and the product of a function by a complex scalar are defined in the usual way.

The reason the square-integrable functions form a (complex) vector space is that the space is closed under addition. The vectors ~~are~~ of the space are the square-integrable functions. In other words, it can be shown that if f and g are two vectors, in this case two functions $f(x)$ and $g(x)$ both of which are square integrable, then

$f(x)+g(x)$ is also square integrable and hence the sum belongs to the vector space.

<u>Proof</u>: Let $f(x)$ and $g(x)$ be two square integrable function, i.e.,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$$

and
$$\int_{-\infty}^{\infty} |g(x)|^2 dx < \infty .$$

Then using the inequality

$$\int_{-\infty}^{\infty} |f+g|^2 dx \leq \left[\sqrt{\int_{-\infty}^{\infty} |f|^2 dx} + \sqrt{\int_{-\infty}^{\infty} |g|^2 dx} \right]^2$$

it is obvious that

$$\int_{-\infty}^{\infty} |f+g|^2 dx < \infty \quad \text{(i.e., finite)}.$$

4. The set of all $n \times n$ ~~elements~~ matrices with complex elements form a complex linear vector space. For illustration let us take 2×2 complex matrices.

Suppose we have a 2×2 matrix A

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where the elements a, b, c and d can, in general be complex. Then A belongs to a vector (complex) space.

We have

$$\phi = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$-A = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}, \qquad \lambda A = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$$

The set of all these matrices form a vector space.

Inner-product space or a unitary vector space.

For a general linear vector space, product of vectors (i.e., multiplication of two vectors) need not be defined. However, we will restrict ourselves to spaces in which a scalar product or an inner product is defined.

A linear vector space is called unitary if a scalar product is defined in it. To every pair of vectors ψ_a and ψ_b in V there corresponds a unique scalar (in general complex), called the scalar product. The scalar product is defined to have the following properties:

(a) $(\psi_a, \psi_b) = (\psi_b, \psi_a)^*$

(b) $(\psi_a, \lambda\psi_b) = \lambda(\psi_a, \psi_b)$

(c) $(\lambda\psi_a, \psi_b) = \lambda^*(\psi_a, \psi_b)$

(d) $(\psi_a, \psi_b + \psi_c) = (\psi_a, \psi_b) + (\psi_a, \psi_c)$

(e) $(\psi_a, \psi_a) \geqslant 0$; the equality holds only if ψ_a is the null vector.

It follows from the above postulated properties of the scalar product, that the scalar product is linear with respect to post factors, i.e,

$$(\psi_a, \lambda\psi_b + \mu\psi_c) = \lambda(\psi_a, \psi_b) + \mu(\psi_a, \psi_c)$$

and antilinear with respect to the pre-factors, i.e.,

$$(\lambda\psi_a + \mu\psi_b, \psi_c) = \lambda^*(\psi_a, \psi_c) + \mu^*(\psi_b, \psi_c).$$

Examples of scalar product.

Ex 1. Consider the vector space consisting of all square integrable functions of a real variable in the domain $[a,b]$. This space is denoted by $L^2[a,b]$.

Suppose $f \in L^2[a,b]$

i.e $\int_a^b f^*(x) f(x) dx \equiv \int_a^b |f(x)|^2 dx < \infty$

We can define the scalar product of two vectors f and g as

$$(f,g) \overset{def}{\equiv} \int_a^b f^*(x) g(x) dx = \text{complex number}$$

We can show

$$|(f,g)| = \left[\sqrt{\int_a^b |f(x)|^2 dx}\right]\left[\sqrt{\int_a^b |g(x)|^2 dx}\right]$$

Since both f and g are square integrable, $|(f,g)|$ is finite, i.e., the scalar product of f and g exists.

The scalar product defined above satisfies all the properties that a scalar product is postulated to have.

Ex 2. Now consider the vector space consisting of n-tuples of complex numbers. Such a vector space is denoted as $\mathbb{C}^n$.

A vector $\psi_a \in \mathbb{C}^n$ may be expressed as

$$\psi_a = (a_1\ a_2 \cdots a_n)^T.$$

The scalar product may then be defined as

$$(\psi_a, \psi_b) \overset{def}{\equiv} (a_1^*\ a_2^* \cdots a_n^*)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$= \sum_{i=1}^{n} a_i^* b_i$$

This scalar product also satisfies all the properties of a scalar product.

~~Norm of a vector.~~

<u>Ex 3</u> Euclidean 3-space $\mathbb{R}^3$. The vectors of $\mathbb{R}^3$ are 3-tuples of real numbers which could be represented as column vectors. Thus if ψ_a and ψ_b are in $\mathbb{R}^3$,

$$\psi_a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \qquad a_i = \text{real}$$

$$\psi_b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} . \qquad b_i = \text{real}$$

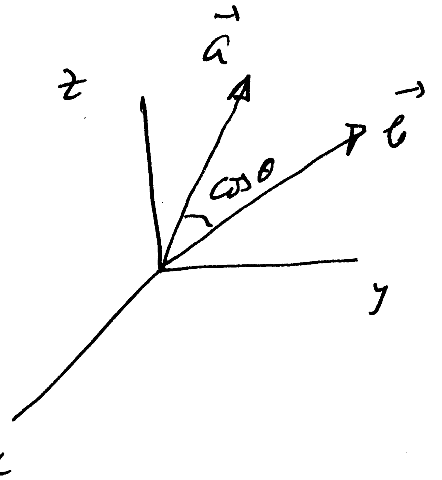
We could define the scalar product of ψ_a and ψ_b as

$$(\psi_a, \psi_b) \overset{\text{def}}{\equiv} (a_1 \; a_2 \; a_3) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \psi_a^T \psi_b$$

$$= \sum_{i=1}^{3} a_i b_i$$

This scalar product also ~~satisfies~~ has all the postulated properties of a scalar product.

In case of the vector space $\mathbb{R}^3$, the vectors Ψ_a and Ψ_b could be represented as directed lines $\vec{a}$ and $\vec{b}$ in a three-dimensional coordinate system.



The scalar product (Ψ_a, Ψ_b) is the usual dot product

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

$$= |\vec{a}|\,|\vec{b}| \cos\theta$$

where $|\vec{a}|$ and $|\vec{b}|$ are the magnitudes of the vectors $\vec{a}$ and $\vec{b}$ defined as

$$|\vec{a}| \equiv \sqrt{(\Psi_a, \Psi_a)} = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

$$|\vec{b}| \equiv \sqrt{(\Psi_b, \Psi_b)} = \sqrt{b_1^2 + b_2^2 + b_3^2}\,.$$

Norm of a vector

If a vector space is endowed with a scalar product, then the scalar product gives us the concept of the 'magnitude' or 'length' of a vector. In a general vector space the 'magnitude' or 'length' of a vector is called the norm of the vector. We simply define the norm of a vector ψ_a as

$$\| \psi_a \| \stackrel{\text{def}}{\equiv} \sqrt{(\psi_a, \psi_a)}$$

The norm has the following properties:

(a) $\| \psi_a \| \geq 0$, the equality holds only if the vector is null.

(b) $\| \psi_a + \psi_b \| \leq \| \psi_a \| + \| \psi_b \|$

This is called the triangle inequality

(c) $\| \psi_a - \psi_b \| = \| \psi_b - \psi_a \|$

Metric induced by the scalar product

The norm induced by the scalar product allows us to develop the concept of 'distance' between vectors in a vector space. We say ~~the~~ two vectors ψ_a and ψ_b are 'close' if

$$\| \psi_a - \psi_b \|$$

is small. The metric in a vector space assigns a real number to ~~two vectors~~ the vector $\psi_a - \psi_b$. This real number is a measure of how close the two vectors are. We simply define the metric $d(\psi_a, \psi_b)$ as

$$d(\psi_a, \psi_b) \overset{\text{def}}{\equiv} \| \psi_a - \psi_b \| .$$

Thus, if there are three vectors ψ_a, ψ_b and ψ_c and if $d(\psi_a, \psi_b) < d(\psi_a, \psi_c)$

then we say ψ_a is closer to ψ_b than to ψ_c.

Schwarz inequality.

We will now prove a very important inequality called Schwartz inequality which states

$$\left|(\psi_a, \psi_b)\right| \leq \sqrt{(\psi_a, \psi_a)(\psi_b, \psi_b)}$$

or

$$\left|(\psi_a, \psi_b)\right| \leq \|\psi_a\| \, \|\psi_b\| .$$

Proof

Let

$$\psi = \psi_a + \lambda \psi_b$$

Then

$$(\psi, \psi) = (\psi_a + \lambda\psi_b, \psi_a + \lambda\psi_b)$$

$$= (\psi_a, \psi_a) + \lambda(\psi_a, \psi_b) + \lambda^*(\psi_b, \psi_a)$$

$$+ |\lambda|^2 (\psi_b, \psi_b) \geqslant 0 .$$

The best inequality is obtained if λ is chosen so as to minimize the left hand side of the above equation. By differentiation, the value of λ which accomplishes this is found to be

$$\lambda = - \frac{(\psi_b, \psi_a)}{(\psi_b, \psi_b)}$$

Substitute this value of λ in the above equation yields the Schwarz inequality.

We note that the equality sign holds if and only if $(\psi, \psi) = 0$, i.e., ψ is the null vector, i.e., $\psi = \phi$, in other words

$$\psi_a + \lambda \psi_b = \phi \text{ (null)}$$

i.e $\quad \psi_a = - \lambda \psi_b + \phi$

or $\quad \psi_a = - \lambda \psi_b$.

Hence, the equality holds if ψ_a and ψ_b are multiples of each other, or if ψ_a and ψ_b are "parallel".

It follows from the Schwarz inequality that the scalar product (ψ_a, ψ_b) is finite if the norms of ψ_a and ψ_b are finite.

Analogy of Schwarz inequality with vectors in a three-dimensional Euclidean space $\mathbb{R}^3$.

In $\mathbb{R}^3$, the vectors can be represented by directed lines (i.e. arrows). We have the scalar product ordinary vectors in the form

$$\vec{A}.\vec{B} = |\vec{A}|\,|\vec{B}| \cos\theta$$

Since cosine of any angle lies between -1 and 1, we have

$$|\vec{A}.\vec{B}| \leq |\vec{A}|\,|\vec{B}|$$

The analogue of this equation for a general vector space is the Schwarz inequality

$$|(\psi_a, \psi_b)| \leq \|\psi_a\|\,\|\psi_b\|.$$

Orthogonality and linear independence

A vector whose norm is unity is called a unit vector. For any given non-null vector, a unit vector can be formed by dividing the vector by its norm. Thus

$$u_a = \frac{\psi_a}{\|\psi_a\|}$$

is normalized.

Two vectors ψ_a and ψ_b are ~~norm~~ orthogonal if their inner product is zero, i.e., if

$$(\psi_a, \psi_b) = 0.$$

The unit vectors $u_1, u_2, \ldots u_N$ form an orthonormal set if they are mutually orthogonal, i.e., if

$$(u_i, u_j) = \delta_{ij}.$$

Linear independence

The set of vectors $\psi_1, \psi_2, \cdots \psi_N$ are linearly independent if none of them can be expressed as a linear combination of the others. Mathematically this means that the equation

$$\sum_{j=1}^{N} c_j \psi_j = 0$$

cannot be satisfied except by $c_j = 0$ for all j.

Orthonormality and linear independence

A set of mutually orthogonal vectors (not necessarily normalized) are necessarily linearly independent. The converse is not true, howerever. That is, a set of linearly independent vectors may not be mutually orthogonal.

It is always possible to orthonormalize a set of linearly independent vectors. By this we mean that from a given set of N linearly independent vectors, it is possible to form a set of N orthonormal vectors. This procedure is called Schmidt orthonormalization method.

<u>Schmidt orthonormalization method.</u>

Suppose $\psi_1, \psi_2, \cdots \psi_N$ is a set of linearly independent vectors. Let

$$u_1 = \frac{\psi_1}{\|\psi_1\|} \quad \text{------} \quad (1)$$

Then $(u_1, u_1) = 1$, i.e., u_1 is normalized.

Next construct the vector ψ_2' as follows:

$$\psi_2' = \psi_2 - u_1 (u_1, \psi_2) \quad \cdots \cdots \quad (2)$$

i.e., to obtain ψ_2' we have subtracted the 'component' of ψ_2 along the u_1 "direction". Then it follows that

$$\begin{aligned} (u_1, \psi_2') &= (u_1, \psi_2) - (u_1, u_1)(u_1, \psi_2) \\ &= (u_1, \psi_2) - (u_1, \psi_2) \\ &= 0 \end{aligned}$$

i.e., ψ_2' is orthogonal to u_1. We then normalize ψ_2', i.e.,

$$u_2 = \frac{\psi_2'}{\| \psi_2' \|} \quad \cdots \cdots \quad (3)$$

We can continue the process until we exhaust all the vectors. For example, in the next step we can write

$$\psi_3' = \psi_3 - u_1 (u_1, \psi_3) - u_2 (u_2, \psi_3) \quad \cdots \text{(4)}$$

We note immediately that ψ_3' is orthogonal to both u_1 and u_2, i.e.,

$$(u_1, \psi_3') = (u_2, \psi_3') = 0$$

We normalize ψ_3' to get u_3, i.e.,

$$u_3 = \frac{\psi_3'}{\|\psi_3'\|} \quad \cdots\cdots \text{(5)}$$

Finally, in the Nth step, we write

$$\psi_N' = \psi_N - u_1 (u_1, \psi_N) - u_2 (u_2, \psi_N) - \cdots u_{N-1} (u_{N-1}, \psi_N).$$

ψ_N' is orthogonal to $u_1, u_2, \cdots u_{N-1}$, i.e.,

$$(u_1, \psi_N') = (u_2, \psi_N') = \cdots = (u_{N-1}, \psi_N') = 0.$$

Normaliting ψ_N' we get

$$u_N = \frac{\psi_N'}{\| \psi_N' \|} \quad \ldots \ldots \text{(6)}$$

Thus, the set $\{u_1, u_2, \ldots u_N\}$ is an orthonormal set of vectors.

Dimension of a vector space

The vector space V is said to be n-dimensional if there exists n linearly independent vectors, but if $n+1$ vectors are linearly dependent. The dimension may be finite or infinite.

Complete vector space

Before defining what a complete vector space is we will give some other ~~definitions~~ definitions.

A sequence of vectors $\{\psi_n\}$ in the vector space V is called a Cauchy sequence if for every $\epsilon > 0$ there exists an integer N such that

$$\| \psi_n - \psi_m \| < \epsilon$$

if $n, m > N$. In other words, the vectors in the sequence ~~come~~ come 'closer' if the index increases. In particular

$$\| \psi_n - \psi_m \| \longrightarrow 0 \quad \text{as} \quad n, m \longrightarrow \infty .$$

Convergence of a sequence of vectors in a vector space.

A sequence $\{\psi_n\}$, $n = 1, 2, \ldots$ in a vector space V converges to a vector ψ in V if for every $\epsilon > 0$ ~~however~~ there is exists an integer N such that

$$\|\psi - \psi_n\| < \epsilon$$

if $n > N$. That is if

$$\lim_{n\to\infty} \|\psi - \psi_n\| = 0 \quad \text{then } \{\psi_n\} \to \psi.$$

and the sequence is called a convergent sequence.

Now we can show every ~~Cauchy~~ convergent sequence is a Cauchy sequence.

<u>Proof</u> Let $\{\psi_n\} \to \psi$.

Then

$$\|\psi_n - \psi_m\| = \|\psi_n - \psi + \psi - \psi_m\|$$

$$\leq \|\psi_n - \psi\| + \|\psi - \psi_m\| \quad \text{(triangle inequality)}$$

Since $\{\psi_n\}$ is a convergent sequence, each term on the right hand side tends to zero as n and m tends to infinity. Hence $\|\psi_n - \psi_m\| \to 0$ as $n, m \to \infty$, i.e., the sequence $\{\psi_i\}$, $i = 1, 2, \ldots$ is a Cauchy sequence.

The converse of the above statement is not true in general. In other words, a Cauchy sequence in a vector space may not converge to a vector in the space. It can be shown that for a finite dimensional vector space the converse ~~this~~ is true, i.e, in a finite dimensional vector space a Cauchy sequence is always a convergent sequence. Exceptions ~~arise~~ may arise in infinite-dimensional vector space.

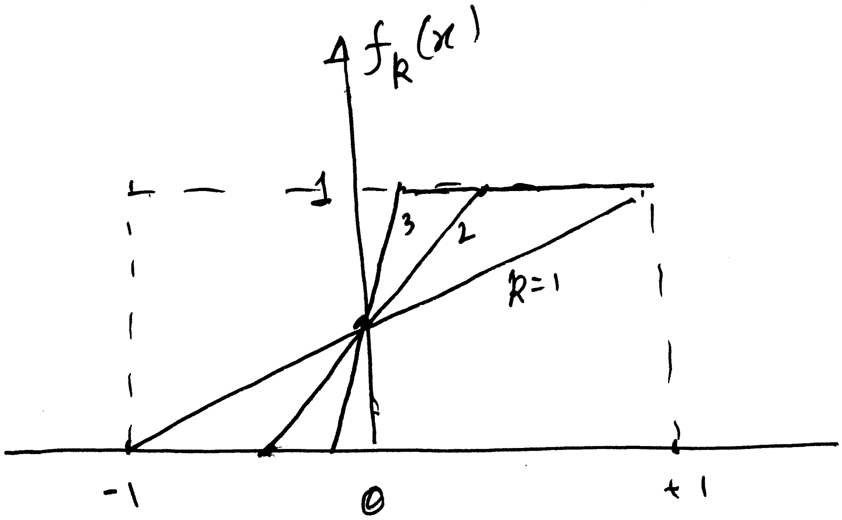
An example of a vector space where Cauchy sequence does not converge to a vector in the vector space.

Consider the vector space consisting of all continuous functions of a single real variable x in the range $[-1, 1]$. In this vector space consider a sequence $\{f_k(x)\}$, $k = 1, 2, \cdots$ of the following form:

$$f_k(x) = \begin{cases} 1 & \text{for } \frac{1}{k} \le x \le 1 \\ \frac{kx+1}{2} & \text{for } -\frac{1}{k} < x < \frac{1}{k} \\ 0 & \text{for } -1 \le x \le -\frac{1}{k} \end{cases}.$$

$k = 1, 2, 3, \ldots$

The graph of the sequence of functions is shown below:



~~Let us define the scalars.~~

Note that, in this example each $f_k(x)$ is continuous, but their first derivatives are discontinuous.

Let us define the scalar product in this space as

$$(f,g) = \int_{-1}^{+1} f^*(x) g(x)\, dx$$

so that the metric $d(f,g)$, i.e., the "distance" between vectors f and g can be defined as

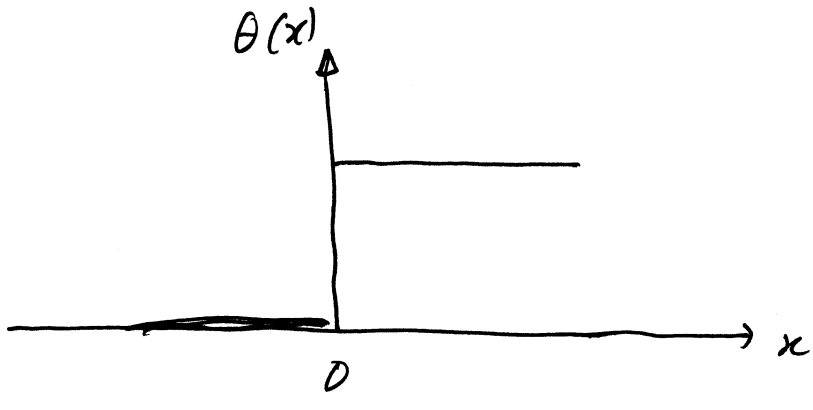
$$d(f,g) \equiv \| f-g \|$$
$$= \sqrt{\int_{-1}^{+1} \left(f^*(x) - g^*(x)\right)\left(f(x) - g(x)\right) dx} \, .$$

With this metric we can show that the sequence $\{f_k\}$ defined above is ~~indee~~ indeed a Cauchy sequence. However, looking at the graph above, we see that as k becomes large, f_k approaches the θ function

$$\theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \end{cases}$$

which is a discontinuous function at $x = 0$.

We show the graph of $\theta(x)$ below



Thus the Cauchy sequence $\{f_k(x)\}$ of continuous functions is converging to a discontinuous function which lies outside the vector space V.

If, instead of choosing all continuous functions, we had chosen all square integrable functions as ~~dd~~ as defining the vector space, then any Cauchy sequence in the vector space would converge to a vector in the space.

<u>Def</u>: A linear vector space is said to be <u>complete</u> if any Cauchy sequence converges to a vector in the space.

Hilbert space

A complete linear vector space, finite or infinite dimensional, endowed with a scalar product between vectors (and hence endowed with a norm and metric induced by the scalar product), is called a Hilbert space.

A finite-dimensional vector space is always complete. So, a finite-dimensional linear vector space in which a scalar product is defined is a Hilbert space.

An infinite dimensional vector space with a scalar product may or may not be complete. Whether or not an infinite-dimensional vector space is complete depends upon how exactly the vector space is defined and on the metric.

Basis vectors in a Hilbert space.

Finite dimensional space.

In a finite dimensional vector space, ~~any set~~ of dimension n, any set of linearly independent vectors $\psi_1, \psi_2, \cdots \psi_n$ spans the entire space. In other words any vector ψ in the space can be expressed as linear combination of $\psi_1, \psi_2 \cdots \psi_n$, i.e

$$\psi = \sum_{i=1}^{n} a_i \psi_i . \quad - - - - - - - (1)$$

The vectors $\psi_1, \psi_2 \cdots \psi_n$ form a complete basis for the vector space. The vectors $\psi_1, \psi_2, \cdots \psi_n$, even if linearly independent, may not be orthogonal to each other. It is more convenient to a use a set of n orthonormal vectors ~~as~~ ~~the basis~~ $\phi_1, \phi_2, \cdots \phi_n$ as the basis. Being orthogonal, the vectors $\phi_1, \phi_2, \cdots \phi_n$ are automatically linearly independent. The orthonormal set of basis vectors $\{\phi_i\}$, $i = 1, 2, \cdots n$, can be constructed from the set $\{\psi_i\}$, $i = 1, 2, \cdots n$ by using the Schmidt orthonormalization procedure.

Choosing the orthonormal set as the basis, any vector ψ in the vector space can be written as

$$\psi = \sum_{i=1}^{n} a_i \phi_i \quad \cdots\cdots (2)$$

where

$$(\phi_i, \phi_j) = \delta_{ij} \quad \cdots\cdots (3)$$

Using eq.(3) we have

$$a_i = (\phi_i, \psi) \quad \cdots\cdots (4)$$

Infinite-dimensional vector space

In an infinite-dimensional vector space the number of basis vectors is infinity. Let $\{\phi_1, \phi_2, \phi_3 \cdots\}$ be an infinite set of orthonormal basis vectors spanning the infinite-dimensional Hilbert space. This set of basis vectors is said to be complete if any vector ψ in the Hilbert space can be expanded as a linear combination of the basis vectors, i.e.,

$$\psi = \sum_{i=1}^{\infty} a_i \phi_i . \quad \text{- - - - - - -} \quad (5)$$

In an infinite-dimensional vector space, choosing an infinite number of basis vectors may not ensure that the basis set is complete. It may so happen that there are other linearly independent vectors, may be infinite in number, which have been missed in the first choice of the basis vectors.

Whenever we ~~have~~ have an infinite sum, as in Eq. (5), the issue of convergence arises. Eq. (5) is to be understood in the sense that the sequences consisting of the partial sums

$$f_n = \sum_{i=1}^{n} a_i \phi_i \;;\quad n = 1, 2, 3, \ldots$$

converges to ψ, i.e.,

$$\lim_{n \to} \| \psi - f_n \| \to 0. \qquad (6)$$

Since the vector ψ must have a finite norm, we must have

$$\|\psi\|^2 = (\psi, \psi) = \sum_{i=1}^{\infty} |a_i|^2 < \infty \text{ (finite)}. \qquad (7)$$

If the basis vectors $\{\phi_i\}$ are orthonormal, we have

$$a_i = (\phi_i, \psi),$$

so that Eq. (7) can be written as

$$\sum_{i=1}^{\infty} |(\phi_i, \psi)|^2 < \infty. \quad \ldots\ldots \quad (7)$$

The scalars a_i can be regarded as the components of ψ in the 'directions' ϕ_i.

Ex. Show that the set of all square integrable functions, i.e., set of all functions f such that

$$\int_{-\infty}^{\infty} f^*(x) f(x)\, dx < \infty \quad (\text{i.e., finite})$$

belong to a Hilbert space. This Hilbert space is denoted as $L^2(-\infty, \infty)$.

To show this, consider the following:

1. If f and g are square integrable functions, so is $f+g$, and hence $f+g$ also belongs to the Hilbert space.

$$(\; \|f+g\| \leq \|f\| + \|g\| \;),$$

2. We can define the scalar product between f and g as follows:

$$(f, g) \overset{def}{\equiv} \int_{-\infty}^{\infty} f^*(x) g(x)\, dx .$$

That the scalar product exists follows from the Schwarz inequality

$$|(f,g)| \leq \|f\| . \|g\| < \infty .$$

3. It can also be shown that any Cauchy sequence of square integrable functions converges to a limit which is also square integrable. In other words, the space of all square integrable functions is complete.

Hence the linear vector space consisting of all square integrable functions is indeed a Hilbert space.

— x —

Dirac notation. (Cohen-Tannoudji, page 109)

* "Ket" vectors and "bra" vectors.

Notation

Any element, or vector of a vector space V is called a <u>ket vector</u>, or, more simply, a <u>ket</u>. It is represented by the symbol $|\ \rangle$, inside which is placed a distinctive sign which enables us to distinguish between different kets, for example $|\psi\rangle$.

Scalar product

With each pair of kets $|\phi\rangle$ and $|\psi\rangle$, taken in this order, we associate a complex number, which is their scalar product $(|\phi\rangle, |\psi\rangle)$ and which satisfies various properties discussed earlier (page 11).

Dual vector space.

Linear functional:

A linear functional χ is a linear operation on the kets such that χ operating on a ket $|\psi\rangle$ gives a complex scalar:

$$\chi\,|\psi\rangle \longrightarrow \text{scalar} \quad \text{where } |\psi\rangle \in V.$$

and

$$\chi\left(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle\right) = \lambda_1\,\chi|\psi_1\rangle + \lambda_2\,\chi|\psi_2\rangle.$$

The set of all linear functionals defined on the kets of a vector space V themselves form a linear vector space called the dual space of V and symbolized by V^*.

Bra notation for the vectors of V^*

Any element, or vector, of the space V^* is called a "bra vector", or, more simply, a bra. It is symbolized by $\langle\,|$. For example, the bra $\langle\chi|$

designates the linear functional χ we shall henceforth use the notation $\langle \chi | \psi \rangle$ to denote the number obtained by causing the linear functional $\langle \chi | \in V^*$ to act on the ket $| \psi \rangle \in V$. Thus

$$\chi (| \psi \rangle) = \langle \chi | \psi \rangle .$$

Correspondence between kets and bras

The existence of the scalar product in V will now enable us to show that we can associate with every ket $| \phi \rangle \in V$ an element of V^*, that is a bra, which will be denoted by $\langle \phi |$.

The ket $| \phi \rangle$ does indeed enable us to define a linear functional, the one which associates with each $| \psi \rangle \in V$ a complex number which is equal to the scalar product $(| \phi \rangle , | \psi \rangle)$. Let $\langle \phi |$ be this linear functional. It is thus defined by the relation

$$\langle \phi | \psi \rangle = (| \phi \rangle , | \psi \rangle) .$$

The correspondence is antilinear

Let $\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle$ be a ket. Then

$$\left(\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle, |\psi\rangle\right)$$

$$= \lambda_1^* \left(|\phi_1\rangle, |\psi\rangle\right) + \lambda_2^* \left(|\phi_2\rangle, |\psi\rangle\right)$$

$$= \lambda_1^* \langle\phi_1|\psi\rangle + \lambda_2^* \langle\phi_2|\psi\rangle$$

$$= \left(\lambda_1^* \langle\phi_1| + \lambda_2^* \langle\phi_2|\right)|\psi\rangle$$

Thus

- $$\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle \xrightarrow{dc} \lambda_1^* \langle\phi_1| + \lambda_2^* \langle\phi_2|$$

where "dc" is short for dual correspondence.

Comment

If λ is a complex number and $|\psi\rangle$ is a ket, then $\lambda|\psi\rangle$ is also a ket. We are sometimes led to write $\lambda|\psi\rangle$ as $|\lambda\psi\rangle$:

$$|\lambda\psi\rangle = \lambda|\psi\rangle.$$

One must be careful to remember that $\langle\lambda\psi|$ represents the bra associated with the ket $|\lambda\psi\rangle$. Since the correspondence between a bra and a ket is antilinear, we have

$$\langle\lambda\psi| = \lambda^* \langle\psi| .$$

Dirac notation for the scalar product.

We now have at our disposal two distinct notations for designating the scalar product of $|\psi\rangle$ by $|\varphi\rangle$, namely, $(|\varphi\rangle, |\psi\rangle)$ and $\langle\varphi|\psi\rangle$, $\langle\varphi|$ being the bra associated with the ket $|\varphi\rangle$.

We shall mostly use the Dirac notation $\langle \phi | \psi \rangle$.
In the table below we summarize, in Dirac notation, the properties of the scalar product.

$$\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^*$$

$$\langle \phi | \lambda_1 \psi_1 + \lambda_2 \psi_2 \rangle = \lambda_1 \langle \phi | \psi_1 \rangle + \lambda_2 \langle \phi | \psi_2 \rangle$$

$$\langle \lambda_1 \phi_1 + \lambda_2 \phi_2 | \psi \rangle = \lambda_1^* \langle \phi_1 | \psi \rangle + \lambda_2^* \langle \phi_2 | \psi \rangle$$

$\langle \psi | \psi \rangle$ is real, positive; zero if and only if $| \psi \rangle = \phi$ (null).