Vector Spaces

Vector space, ν, over the field of complex numbers, $\mathcal{C}$, is a set of elements $|a\rangle$, $|b\rangle, \ldots$, satisfying the following axioms.

$\diamond$ For each two vectors $|a\rangle, |b\rangle \in \nu$ there exists a summation procedure: $|a\rangle + |b\rangle = |c\rangle$, where $|c\rangle \in \nu$. The summation obeys the following laws.

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle \quad \text{(commutative)} \; , \tag{1}$$

$$|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle \quad \text{(associative)} \; . \tag{2}$$

$\diamond$ There exists a zero vector $|0\rangle$, such that $\forall\, |a\rangle$:

$$|a\rangle + |0\rangle = |a\rangle \; . \tag{3}$$

$\diamond$ $\forall\, |a\rangle \quad \exists\, |-a\rangle$ (additive inverse) such that

$$|a\rangle + |-a\rangle = |0\rangle \; . \tag{4}$$

[Here we start using the symbol $\forall$ that means 'for all' ('for each', 'for any') and the symbol $\exists$ that means 'there exists'.]

$\diamond$ There exists a procedure of multiplication by a scalar $\alpha \in \mathcal{C}$. That is $\forall\, |a\rangle \in \nu$, $\forall\, \alpha \in \mathcal{C}$: $\exists\; \alpha\, |a\rangle \in \nu$. Multiplication by a scalar obeys the following laws.

$$\alpha\, (\beta\, |a\rangle\,) = (\alpha\beta)\, |a\rangle \; , \tag{5}$$

$$1 \cdot |a\rangle = |a\rangle \; , \tag{6}$$

$$\alpha(|a\rangle + |b\rangle) = \alpha|a\rangle + \alpha|b\rangle \; , \tag{7}$$

$$(\alpha + \beta)|a\rangle = \alpha|a\rangle + \beta|a\rangle \; . \tag{8}$$

From the above axioms it follows that $\forall\, |a\rangle$

$$0 \cdot |a\rangle = |0\rangle \; , \tag{9}$$

$$(-1) \cdot |\, a\rangle = |-a\rangle \; . \tag{10}$$

Problem 8. On the basis of the axioms:
(a) Show that the zero element is unique.
(b) Show that for any vector $|a\rangle$ there exists only one additive inverse.
(c) Show that for any vector $|a\rangle$ the relation $|a\rangle + |x\rangle = |a\rangle$ implies that $|x\rangle = |0\rangle$.
(d) Derive (9)-(10).
Important (!): Here you are not allowed to use the subtraction procedure, which is not defined yet, and *cannot* be defined *prior* to establishing the uniqueness of additive inverse.

Once the uniqueness of additive inverse is established (Problem 8), it is convenient to define subtraction as simply adding additive inverse:

$$|\,a\rangle - |\,b\rangle \;\equiv\; |\,a\rangle + |-b\rangle \; . \tag{11}$$

Example. As a particular example of a complex vector space consider the space $\mathcal{C}^n$ of the n-rows of complex numbers:

$$|\,x\rangle \;=\; (x_1,\, x_2,\, x_3,\, \ldots,\, x_n) \; . \tag{12}$$

In the space $\mathcal{C}^n$, the addition and multiplication by a complex number are defined componentwise. For $|\,y\rangle \;=\; (y_1,\, y_2,\, y_3,\, \ldots,\, y_n)$ we define

$$|\,x\rangle \,+\, |\,y\rangle \;=\; (x_1 + y_1,\, x_2 + y_2,\, x_3 + y_3, \ldots,\, x_n + y_n) \; . \tag{13}$$

For $\alpha \in \mathcal{C}$ we define

$$\alpha\,|\,x\rangle \;=\; (\alpha\, x_1,\, \alpha\, x_2,\, \alpha\, x_3,\, \ldots,\, \alpha\, x_n) \; . \tag{14}$$

It is easy to make sure that all the axioms of the vector space are satisfied, and thus $\mathcal{C}^n$ is indeed a vector space.

Inner product. Now let us formally introduce the inner-product vector space as a vector space in which for any two vectors $|a\rangle$ and $|b\rangle$ there exists the inner product, $\langle\, b \,|\, a \,\rangle$, which is a complex-valued function of the two vectors satisfying the following properties.

$$\langle\, b \,|\, a \,\rangle = \overline{\langle\, a \,|\, b \,\rangle}\ . \tag{15}$$

Here the bar denotes complex conjugation. From Eq. (15) it directly follows that the inner product of a vector with itself is always real.

The next axiom of inner product requires also that $\langle a|a\rangle$ be *positive* if $|a\rangle \neq |0\rangle$.

Finally, we require that for any three vectors $|a\rangle$, $|u\rangle$, $|v\rangle$, and any two numbers α and β, the following relation—termed *linearity* of the inner product—holds true:

$$\langle\, a \,|\, \alpha\, u + \beta\, v \,\rangle \;=\; \alpha \,\langle\, a \,|\, u\rangle \;+\; \beta\,\langle\, a \,|\, v\rangle\ . \tag{16}$$

Here we use a convenient notation $|\, u +\, v\,\rangle \equiv |u\rangle + |v\rangle$ and $|\,\alpha\, u\,\rangle \equiv \alpha\,|u\rangle$. From Eqs. (15) and (16) it *follows* that

$$\langle\, \alpha\, u + \beta\, v \,|\, a \,\rangle \;=\; \alpha^* \,\langle\, u \,|\, a\rangle \;+\; \beta^*\,\langle\, v \,|\, a\rangle\ . \tag{17}$$

From the axioms of the inner product it is also easy to deduce that

$$\langle\, a \,|-a\rangle \;<\; 0\ , \qquad \text{if} \quad |a\rangle \neq |0\rangle\ , \tag{18}$$

and that

$$\langle\, a \,|\, 0\rangle \;=\; 0\ . \tag{19}$$

Problem 9. Derive Eqs. (17)-(19) and from the axioms of the inner product.

Example 1. In the above-discussed vector space $\mathcal{C}^n$, the inner product of the two vectors, $|a\rangle \;=\; (a_1, a_2, \,\ldots\,, a_n)$ and $|b\rangle \;=\; (b_1, b_2, \,\ldots\,, b_n)$, can be defined as

$$\langle\, a \,|\, b\,\rangle \;=\; \sum_{j=1}^{n} a_j^*\, b_j\ . \tag{20}$$

More generally, one can define

$$\langle\, a \,|\, b\,\rangle \;=\; \sum_{j=1}^{n} a_j^*\, b_j\, w_j\ , \tag{21}$$

where w_j's are some fixed real positive coefficients.

Example 2. In the vector space of integrable complex-valued functions $f(x)$ defined in the interval $x \in [a, b]$ the inner product can be defined as

$$\langle f \,|\, g \rangle \;=\; \int_a^b f^*(x)\, g(x)\, dx \; . \tag{22}$$

More generally, one can define

$$\langle f \,|\, g \rangle \;=\; \int_a^b f^*(x)\, g(x)\, w(x)\, dx \; , \tag{23}$$

where $w(x)$ is some fixed positive real function.

Finite-dimensional inner-product vector space

Linear combination. Span. Suppose we have a set of n vectors $\{|\phi_j\rangle\}$, $(j = 1, 2, \ldots, n)$ of a vector space ν. The following vector,

$$|a\rangle \;=\; \sum_{j=1}^{n} c_j \,|\phi_j\rangle \quad , \tag{24}$$

where $\{c_j\}$ are some complex numbers, is called a *linear combination* of the vectors $\{|\phi_j\rangle\}$. As is directly seen from the axioms of the vector space, all the linear combinations of the given set $\{|\phi_j\rangle\}$ form some vector space $\tilde{\nu}$—for such a space we will be using the term *sub-space* to stress the fact that $\tilde{\nu} \subseteq \nu$. The vector space $\tilde{\nu}$ is referred to as the *span* of the vectors $\{|\phi_j\rangle\}$. The same idea is also expressed by saying that the vectors $\{|\phi_j\rangle\}$ span the subspace $\tilde{\nu}$.

Basis. If $\tilde{\nu} = \nu$, that is if *any* vector of the space ν can be represented as a linear combination of the vectors $\{|\phi_j\rangle\}$, then the set $\{|\phi_j\rangle\}$ is called *basis* and the space ν is a *finite-dimensional* space. [We can also put this definition as follows: A finite-dimensional space is the space spanned by a finite number of vectors.] For any finite-dimensional vector space there is a minimal possible number n for a vector set to form a basis. This number is called *dimensionality* of the space, and, correspondingly, the space is called n-dimensional space.

Linear independence of a system of vectors is a very important notion, directly relevant to the notion of the dimensionality of the vector space. There are two equivalent definitions of linear independence of a system of vectors, $\{|\phi_j\rangle\}$. The first definition says that the system is linear independent if none of its vectors can be expressed as a linear combination of the others. The second definition says that the system is linear independent if the equality

$$\sum_{j=1}^{n} c_j|\phi_j\rangle \;=\; |0\rangle \tag{25}$$

is only possible if all numbers c_j are equal to zero.

Problem 10. Prove the equivalence of the two definitions.

In an n-dimensional vector space, all the vectors of an n-vector basis $\{|\phi_j\rangle\}$ are *linear independent*. Indeed, if the vectors $\{|\phi_j\rangle\}$ were *linear dependent* (= not linear independent), then—in accordance with one of the definitions of linear independence—some vector vector $|\phi_{j_0}\rangle$ could be expressed as a linear combination of the other vectors of the basis, and we would get a basis of $(n-1)$ vectors—all the vectors of the original basis, but the vector $|\phi_{j_0}\rangle$. But this is impossible by definition of the dimensionality saying that n is the *minimal* possible number for the basis elements!

Without loss of generality we may deal only with a linear-independent basis, since we can always eliminate linear dependent vectors. In what follows, we assume that our basis is linear independent.

Problem 11. Show that if the basis $\{|\phi_j\rangle\}$ is linear independent, then for any vector $|x\rangle$ the coefficients c_j of the expansion

$$|x\rangle \;=\; \sum_{j=1}^{n} c_j|\phi_j\rangle \tag{26}$$

are unique.

Orthonormal basis. With the inner-product structure, we can introduce a very convenient for applications type of basis. The two vectors are called *orthogonal* if their inner product equals zero. Correspondingly, a basis is called orthogonal if all its vectors are orthogonal to each other. A basis is called normal if for each its vector the product of this vector with itself equals unity. An orthonormal basis (ONB) is a basis which is both orthogonal and normal. Summarizing, the basis $\{|e_j\rangle\}$ is orthonormal if

$$\langle\, e_i \,|\, e_j \,\rangle = \delta_{ij} \ , \tag{27}$$

where δ_{ij} is the Kronecker delta symbol, equal to one for $i = j$, and zero for $i \neq j$.

An extremely important feature of ONB's is that $\forall\,|x\rangle \in \nu$, the coefficients x_j—referred to as components of the vector $|x\rangle$ with respect to the given basis—in the expansion

$$|x\rangle \;=\; \sum_{j=1}^{n} x_j\,|e_j\rangle \ , \tag{28}$$

are given by the inner products of the vector $|x\rangle$ with the basis vectors

$$x_j \;=\; \langle e_j|x\rangle \ . \tag{29}$$

This relation is readily seen by constructing inner products of the r.h.s. of (28) with the basis vectors.

Gram-Schmidt orthonormalization procedure. Staring from an arbitrary basis$\{|\phi_j\rangle\}$, one can always construct an orthonormal basis $\{|e_j\rangle\}$. This is done by the following procedure.

$$\begin{aligned}
|\tilde{e}_1\rangle &= |\phi_1\rangle \ , \qquad & |e_1\rangle &= |\tilde{e}_1\rangle/\sqrt{\langle\tilde{e}_1|\tilde{e}_1\rangle} \ , \\
|\tilde{e}_2\rangle &= |\phi_2\rangle - \langle e_1|\phi_2\rangle\,|e_1\rangle \ , \qquad & |e_2\rangle &= |\tilde{e}_2\rangle/\sqrt{\langle\tilde{e}_2|\tilde{e}_2\rangle} \ , \\
|\tilde{e}_3\rangle &= |\phi_3\rangle - \langle e_1|\phi_3\rangle\,|e_1\rangle - \langle e_2|\phi_3\rangle\,|e_2\rangle \ , \qquad & |e_3\rangle &= |\tilde{e}_3\rangle/\sqrt{\langle\tilde{e}_3|\tilde{e}_3\rangle} \ , \\
&\cdots\cdots\cdots\cdots\cdots\cdots && \\
|\tilde{e}_n\rangle &= |\phi_n\rangle - \langle e_1|\phi_n\rangle\,|e_1\rangle - \ldots - \langle e_{n-1}|\phi_n\rangle\,|e_{n-1}\rangle \ , \qquad & |e_n\rangle &= |\tilde{e}_n\rangle/\sqrt{\langle\tilde{e}_n|\tilde{e}_n\rangle} \ .
\end{aligned}$$

By construction, each successive vector $|\tilde{e}_{j_0}\rangle$ is orthogonal to all previous vectors $|e_j\rangle$, and then it is just properly normalized to yield $|e_{j_0}\rangle$.

Isomorphism. With the tool of the orthonormal basis, it is easy to show that each n-dimensional inner-product space is *isomorphic*—the precise meaning of this word will become clear later—to the above-introduced vector space $\mathcal{C}^n$. Fixing some orthonormal basis $\{|e_j\rangle\}$ in our vector space ν, we note that Eqs. (28)-(29) then explicitly yield a one-to-one mapping between the vectors of the spaces ν and $\mathcal{C}^n$:

$$|x\rangle \leftrightarrow (x_1, x_2, \ldots, x_n)\,. \tag{30}$$

By the axioms of the inner product, from (28)-(29) one makes sure that if $|x\rangle \leftrightarrow (x_1, x_2, \ldots, x_n)$ and $|y\rangle \leftrightarrow (y_1, y_2, \ldots, y_n)$, then

$$|x\rangle + |y\rangle \leftrightarrow (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)\,, \tag{31}$$

$$\alpha\,|x\rangle \leftrightarrow \alpha\,(x_1, x_2, \ldots, x_n)\,, \tag{32}$$

$$\langle x|y\rangle = (x_1, x_2, \ldots, x_n)\cdot(y_1, y_2, \ldots, y_n) \equiv \sum_{j=1}^{n} x_j^*\, y_j\,, \tag{33}$$

That is we see a complete equivalence of the two vector spaces, and it is precisely this type of equivalence which we understand by the word isomorphism. The crucial *practical* importance of this isomorphism is that it allows one to reduce all operations with vectors of an arbitrary nature to operations with numbers, by introducing a basis and explicitly evaluating components of vectors in accordance with Eq. (29).

Problem 12. Consider the vector space of real-valued polynomials of the power not larger than 4:

$$P_4(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4\,. \tag{34}$$

This is a 5-dimensional vector space in which the functions $\phi_0(x) = 1$, $\phi_1(x) = x$, $\phi_2(x) = x^2$, $\phi_3(x) = x^3$, and $\phi_4(x) = x^4$ form a basis. Confining the variable x to the interval $x \in [-1, 1]$, we introduce the inner product of two polynomials, $f(x)$ and $g(x)$, by

$$\langle f|g\rangle = \int_{-1}^{1} f(x)\, g(x)\, dx\,. \tag{35}$$

Use Gram-Schmidt procedure to orthonormalize the basis $\{\phi_j(x)\}$, $j = 0, 1, 2, 3, 4$. As a result, you are supposed to obtain five polynomials $\{e_j(x)\}$, $j = 0, 1, 2, 3, 4$, that are orthogonal to each other in the sense of the inner product (35). Check the orthogonality and normalization of the new basis by explicit integration. Expand the function x^4 in terms of the basis $\{e_j(x)\}$.

Normed vector space

A vector space ν is said to be normed if $\forall\, x \in \nu$ there is defined a non-negative real number $\|x\|$ satisfying the following requirements (axioms of norm).

$$\|x+y\| \leq \|x\| + \|y\| \qquad \text{(Triangle inequality)}\ . \tag{36}$$

$\forall \alpha \in \mathcal{C},\, \forall\, x \in \nu$:

$$\|\alpha x\| = |\alpha| \cdot \|x\| \qquad \text{(Linearity of the norm)}\ . \tag{37}$$

Note that the linearity of the norm implies $\|0\| = 0$.

Finally, we require that from $\|x\| = 0$ it should follow that $x = 0$. Hence, the norm is positive for all vectors, but zero.

Problem 13. Consider the vector space R^2, i.e. the set of pairs (x, y) of real numbers.
(a) Show that the function $\|\cdot\|_M$ defined by $\|(x,y)\|_M = \max\{|x|, |y|\}$ is a norm on R^2.
(b) Show that the function $\|\cdot\|_S$ defined by $\|(x,y)\|_S = |x| + |y|$ is a norm on R^2.
(c) In any normed space ν the unit ball $\mathcal{B}_1$ is defined to be $\{u \in \nu \mid \|u\| \leq 1\}$. Draw the unit ball in R^2 for each of the norms, $\|\cdot\|_M$ and $\|\cdot\|_S$.

An inner-product vector space is *automatically* a normed vector space, if one defines the norm as

$$\|x\| = \sqrt{\langle\, x \,|\, x \,\rangle}\ . \tag{38}$$

It can be shown that with the definition (38) all the axioms of norm are satisfied and, moreover, there take place the following *specific to the inner-product norm* properties.

Pythagorean theorem:

$$\langle\, y \,|\, x \,\rangle \;=\; 0 \qquad \Rightarrow \qquad \|x+y\|^2 \;=\; \|x\|^2 + \|y\|^2\ . \tag{39}$$

Parallelogram law:

$$\|x-y\|^2 + \|x+y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2\ . \tag{40}$$

Cauchy-Bunyakovsky-Schwarz inequality:

$$\mathrm{Re}\,\langle\, y \,|\, x \,\rangle \;\leq\; \|x\| \cdot \|y\| \;, \tag{41}$$

and its stronger version,

$$|\langle\, y \,|\, x \,\rangle| \;\leq\; \|x\| \cdot \|y\| \;. \tag{42}$$

The proofs of the above facts go beyond the scope of our course, but they are not that difficult and, if you will, you can try to find them yourselves, given the following hints. The triangle inequality directly follows from the simplest version of Cauchy-Bunyakovsky inequality, Eq. (41). To prove Eq. (41), utilize the fact that the product $\langle\, x+\lambda y \,|\, x+\lambda y \,\rangle$ is a second order polynomial of λ which is non-negative $\forall \lambda$ (by an axiom of the inner product), which implies a certain constraint on its discriminant. To arrive at Eq. (42), use the same approach, but with $\lambda \to \lambda \langle\, y \,|\, x \,\rangle$.

Convergent sequence. Let ν be a normed vector space. The sequence of vectors $\{x_k,\ k = 1, 2, 3, \ldots\} \in \nu$ is said to be convergent to the vector $x \in \nu$, if $\forall\, \varepsilon > 0\ \exists\, k_\varepsilon$, such that if $k > k_\varepsilon$, then $\|x - x_k\| < \varepsilon$. The fact that the sequence $\{x_k\}$ converges to x is symbolically written as $x = \lim_{k\to\infty} x_k$.

Cauchy sequence. The sequence of vectors $\{x_k,\ k = 1, 2, 3, \ldots\} \in \nu$ is said to be a Cauchy sequence if $\forall\, \varepsilon > 0\ \exists\, k_\varepsilon$, such that if $m, n > k_\varepsilon$, then $\|x_m - x_n\| < \varepsilon$.

Any convergent sequence is necessarily a Cauchy sequence.

Problem 14. Show this.

The question now is: Does *any* Cauchy sequence in a given vector space ν converge to some vector $x \in \nu$? The answer is not necessarily positive and essentially depends on the structure of the vector space ν. (Normed spaces where all Cauchy sequences converge are called *complete spaces*, or Banach spaces.) For any inner-product vector space of a finite dimension the answer is positive and is readily proven by utilizing the fact of existence of the orthonormal basis, $\{|e_i\rangle\}$. First we note that if

$$|\, a \,\rangle = \sum_{i=1}^{n} \alpha_i |\, e_i \,\rangle \;, \tag{43}$$

then

$$\|a\| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \ldots + |\alpha_n|^2} \; . \tag{44}$$

If we have a Cauchy sequence of vectors $|a^{(k)}\rangle$, then for any given $i = 1, 2, \cdots, n$ the i-th coordinates of the vectors, $\alpha_i^{(k)}$, form a Cauchy sequence of complex numbers. Any Cauchy sequence of complex numbers converges to some complex number.—This is a consequence of the fact that a complex-number Cauchy sequence is equivalent to two real-number Cauchy sequences (for the real and imaginary parts, respectively) and a well-known fact of the theory of real numbers that any real-number Cauchy sequence is convergent (completeness of the set of real numbers). We thus introduce the numbers

$$\alpha_i = \lim_{k\to\infty} \alpha_i^{(k)} \; , \tag{45}$$

and can readily check that our vector sequence converges to the vector

$$|\, a\,\rangle = \sum_{i=1}^{n} \alpha_i |\, e_i\rangle \; . \tag{46}$$

A *complete inner-product space* is called *Hilbert space*. We have demonstrated that *all* finite-dimensional inner-product vector spaces are Hilbert spaces.

An important example of complete infinite-dimensional vector space is the space $C[a,b]$ of all continuous complex-valued functions $f(x)$, $x \in [a,b]$ with the norm defined as $\|f\|_{\rm sup} = \max\{|f(x)|, \;\; x \in [a,b]\}$. This norm (which is called 'sup' norm, from 'supremum') guaranties that any Cauchy sequence of functions $f_k(x)$ converges *homogeneously* to some continuous function $f(x)$. A problem with the 'sup' norm however is that it does not satisfy the parallelogram law Eq. (40) which means that it cannot be associated with this or that inner product.

Problem 15. Show by an example that the 'sup' norm does not imply the parallelogram law.

Hence, here we have an example of a complete space which is not an inner-product space, and thus is not a Hilbert space.

Countably infinite orthonormal system. Let $\{|e_j\rangle\}$, $j = 1, 2, 3, \ldots$ be an infinite countable orthonormal set of vectors in some infinite-dimensional inner-product vector space. The series

$$\sum_{j=1}^{\infty} \langle e_j|x\rangle \, |e_j\rangle \tag{47}$$

is called Fourier series for the vector $|x\rangle$ with respect to the given orthonormal systems; the numbers $\langle e_j|x\rangle$ are called Fourier coefficients.
Theorem. Partial sums of the Fourier series form a Cauchy sequence.
Proof. We need to show that $\{|x^{(n)}\rangle\}$, where

$$|x^{(n)}\rangle = \sum_{j=1}^{n} \langle e_j|x\rangle \, |e_j\rangle \tag{48}$$

a Cauchy sequence. From (39) we have

$$\|x^{(m)} - x^{(n)}\|^2 = \sum_{j=n+1}^{m} |\langle e_j|x\rangle|^2 \qquad (m > n) \; , \tag{49}$$

which means that it is sufficient to show that the real-number series

$$\sum_{j=1}^{\infty} |\langle e_j|x\rangle|^2 \tag{50}$$

converges. The series (50) is non-negative, and to prove its convergence it is enough to demonstrate that it is bounded from above. This is done by utilizing the inequality

$$\langle\, x - x^{(n)} |\, x - x^{(n)} \rangle \;\geq\; 0 \; . \tag{51}$$

A straightforward algebra shows that

$$\langle\, x - x^{(n)} |\, x - x^{(n)} \rangle \;=\; \|x\|^2 - \sum_{j=1}^{n} |\langle e_j|x\rangle|^2 \; . \tag{52}$$

Hence, we obtain

$$\sum_{j=1}^{n} |\langle e_j|x\rangle|^2 \;\leq\; \|x\|^2 \qquad \text{(Bessel inequality)} \; , \tag{53}$$

and prove the theorem.

Moreover, rewriting Eq. (52) as

$$\|x - x^{(n)}\| \;=\; \sqrt{\|x\|^2 - \sum_{j=1}^{n} \,|\langle e_j|x\rangle|^2} \;. \tag{54}$$

we see that for the series (48) to *converge* to the vector $|x\rangle$ it is necessary and sufficient to satisfy the condition

$$\sum_{j=1}^{\infty} \,|\langle e_j|x\rangle|^2 = \|x\|^2 \qquad \text{(Parseval relation)} \;. \tag{55}$$

Problem 16. Derive Eqs. (49) and (52).

Previously, we introduced the notion of span for a system of finite number of vectors. In a Hilbert space, the notion of span can be generalized to a countably infinite number of vectors in a system. Consider all the Cauchy sequences that can be constructed out of *finite-number* linear combinations of the vectors of our system. Without loss of generality, we can consider only orthonormal systems (ONS), and only special Cauchy sequences. Namely, the ones that correspond to Fourier series for elements of the original vector space in terms of the given ONS. The vector properties of the new set follow from the definition of the Fourier series. Denoting by $|x'\rangle$ the Fourier series corresponding to the vector $|x\rangle$, we see that $|x'\rangle + |y'\rangle = |(x+y)'\rangle$ and $\alpha|x'\rangle = |(\alpha x)'\rangle$. The vector $|x'\rangle$ is called *projection* of the vector $|x\rangle$ on the subspace spanned by the orthonormal system.

By the definition of the span $\tilde{\nu}$, corresponding ONS forms an orthonormal basis (ONB) in it, that is any vector $x \in \tilde{\nu}$ can be represented as the Fourier series over this ONS, that converges to x. Correspondingly, if ONS spans the whole Hilbert space, then it forms ONB in this Hilbert space, and for *any* vector of the Hilbert space the Fourier series converges to it.

Isomorphism of Hilbert spaces with a countably infinite basis. Given a Hilbert space with a countably infinite basis $\{|e_j\rangle\}$, one can construct an isomorphism of this space with a particular space, l_2, which consists of *infinite* complex-number rows, $(x_1, x_2, x_3, \ldots)$, subject to the condition

$$\sum_{j=1}^{\infty} |x_j|^2 < \infty \;. \tag{56}$$

The construction of the isomorphism is absolutely the same as in the previously considered case of the finite-dimensional space, and we do not need to repeat it. The only necessary remark is that the convergence of corresponding series in the space l_2 is guaranteed by Eq. (56). This isomorphism means that all the Hilbert spaces with a countably infinite basis have a similar structure which is very close to that of a finite-dimensional Hilbert space.

Inner product space of functions. As we already know, in the space $C[a,b]$ (of continuous complex-valued functions) the inner product can be defined as

$$\langle\, f \,|\, g \,\rangle = \int_a^b f^* g \, w \, \mathrm{d}x \ , \tag{57}$$

where $w \equiv w(x)$ is some real positive-definite function [for the sake of briefness, below we set $w \equiv 1$, since we can always restore it in all the integrals by $\mathrm{d}x \to w(x)\,\mathrm{d}x$.] All the axioms of the inner product are satisfied. What about completeness? The norm now is given by

$$\|f\| = \sqrt{\int_a^b |f|^2 \, \mathrm{d}x} \ , \tag{58}$$

and it is easy to show that the space $C[a,b]$ is not complete with respect to this norm. Indeed, consider a Cauchy sequence of functions $f_k(x)$, $x \in [-1,1]$, where $f_k(x) = 0$ at $x \in [-1,0]$, $f_k(x) = kx$ at $x \in [0,1/k]$, and $f_k(x) = 1$ at $x \in [1/k,1]$. It is easily seen that in the limit of $k \to \infty$ the sequence $\{f_k\}$ converges to the function $f(x)$ such that $f(x) = 0$ at $x \in [-1,0]$, $f(x) = 1$ at $x \in (0,1]$. But the function $f(x)$ is discontinuous at $x = 0$, that is $f \notin C[a,b]$.

Is it possible to construct a Hilbert space of functions? The answer is "yes", but the theory becomes quite non-trivial. Here we confine ourselves with outlining some most important results. First, one has to understand the integral in the definition of the inner product (57) as a *Lebesgue integral*, which is a more powerful integration scheme than the usual one. Then, one introduces the set of functions, $\mathcal{L}_2[a,b]$, by the following rule: $f \in \mathcal{L}_2[a,b]$ if the Lebesgue integral

$$\int_a^b |f|^2 \, \mathrm{d}x \tag{59}$$

does exist. For the set $\mathcal{L}_2[a,b]$ to be a vector space, one has to introduce a weaker than usual notion of equality. Namely, two functions, $f(x)$ and $g(x)$, if considered as vectors of the space $\mathcal{L}_2[a,b]$, are declared "equal" if

$$\int_a^b |f-g|^2 \, \mathrm{d}x = 0 \; . \tag{60}$$

Fore example, the function $f(x) = 0$ at $x \in [a,b)$ and $f(b) = 1$ is "equal" to the function $g(x) \equiv 0$. It can be proven that this definition of equality implies that the two "equal" functions coincide *almost everywhere*. The term 'almost everywhere' means that the set of points where the functions do not coincide is of the *measure zero*. [A set of points is of measure zero if it can be covered by a countable set of intervals of arbitrarily small total length.] Correspondingly, the convergence by the inner product norm is actually a convergence *almost everywhere*.

According to the *Riesz-Fisher theorem*, the vector space $\mathcal{L}_2[a,b]$ is a Hilbert space with countably infinite basis.

Also important is the *Stone-Weierstrass theorem* which states that the set of polynomials $\{x^k, k = 0, 1, 2, \ldots\}$ forms a basis in $\mathcal{L}_2[a,b]$. That is any $f \in \mathcal{L}_2[a,b]$ can be represented as a series

$$|f\rangle = \sum_{k=0}^{\infty} a_k \, x^k \; , \tag{61}$$

which converges to $|f\rangle$ by the inner product norm, that is the series (61) converges to $f(x)$ almost everywhere. The basis of polynomials $\{x^k, k = 0, 1, 2, \ldots\}$ is not orthogonal. However, one can easily orthonormalize it by the Gram-Schmidt process. The orthonormalized polynomials are called *Legendre polynomials*. Clearly, the particular form of Legendre polynomials depends on the interval $[a,b]$.

It is worth noting that while the space $\mathcal{L}_2[a,b]$ contains pretty weird functions—for example, featuring an infinite number of jumps—the polynomial basis consists of analytic functions. This is because any function in $\mathcal{L}_2[a,b]$ can be approximated to any given accuracy—in the integral sense of Eq. (57)—by some polynomial. In this connection they say that the set of all polynomials is *dense* in $\mathcal{L}_2[a,b]$.

Linear self-adjoint operator

An operator is a function that takes one vector from our vector space and converts it into another one. An operator is said to be *linear* if it satisfies the following requirement:

$$L\,|\,\alpha\,u+\beta\,v\,\rangle \;=\; \alpha L\,|u\rangle+\beta L\,|v\rangle\;. \tag{62}$$

Problem 17. Which of the following operators, acting in the vector space of functions and transforming a function f into a function g, are linear, and which are not?—Explain why.

(a) $g(x) \;=\; \cos(f(x))$

(b) $g(x) \;=\; \int_0^1 \sin(xy)\,f(y)\,dy$

(c) $g(x) \;=\; \int_0^x [f(y)]^2\,dy$.

The operator $L^\dagger$ is said to be adjoint (or Hermitian conjugate) to L, if $\forall\,|a\rangle,|b\rangle\in\nu$:

$$\langle\,b\,|\,L\,|\,a\,\rangle=\overline{\langle\,a\,|\,L^\dagger\,|\,b\,\rangle}\;. \tag{63}$$

With a convenient notation $|\,La\,\rangle\equiv\;L\,|\,a\,\rangle$ and taking into account the axiom (15) we can rewrite (63) as

$$\langle\,b\,|\,La\,\rangle=\langle\,L^\dagger b\,|\,a\,\rangle\;. \tag{64}$$

Clearly, $(L^\dagger)^\dagger\equiv L$.

An operator L is said to be *self-adjoint* (or Hermitian) if $L^\dagger=L$, that is if $\forall\,|a\rangle,|b\rangle$:

$$\langle\,b\,|\,La\,\rangle=\langle\,Lb\,|\,a\,\rangle\;\;. \tag{65}$$

Problem 18. An operator $P_a\;\equiv\;|a\rangle\langle a|$, where $|a\rangle$ is some fixed vector, is defined as follows.

$$P_a|x\rangle\;=\;|a\rangle\langle a|x\rangle\;\equiv\langle a|x\rangle|a\rangle\;. \tag{66}$$

[For your information, such an operator is called *projector* on a given vector $|a\rangle$. Note that $P_a|x\rangle$ is just the vector $|a\rangle$ times a number.] Show that any projector is:

(i) linear operator, and (ii) Hermitian operator.

A vector $|u\rangle$ is called eigenvector of the operator L, if

$$L|u\rangle = \lambda_u |u\rangle \ , \tag{67}$$

where λ_u is some number which is called eigenvalue. For a self-adjoint operator all the eigenvalues are real. To make it sure we first observe that from (65) it follows that for *any* vector $|a\rangle$ the quantity $\langle\, a \,|\, L \,|\, a \,\rangle$ is real provided L is self-adjoint. And we just need to note that $\lambda_u = \langle\, u \,|\, L \,|\, u \,\rangle / \langle\, u \,|\, u \,\rangle$.

Theorem. Any two eigenvectors, $|u_1\rangle$ and $|u_2\rangle$, of the linear self-adjoint operator L are orthogonal if their eigenvalues are different.
Proof. From $L|u_1\rangle = \lambda_{u_1}|u_1\rangle$ and $L|u_2\rangle = \lambda_{u_1}|u_2\rangle$ we get

$$\langle u_2|L|u_1\rangle = \lambda_{u_1}\langle u_2|u_1\rangle \ , \tag{68}$$

$$\langle u_1|L|u_2\rangle = \lambda_{u_2}\langle u_1|u_2\rangle \ . \tag{69}$$

Complex-conjugating the latter equality and taking into account (15), (65), and the fact that eigenvalues of a Hermitian operator are real, we obtain

$$\langle u_2|L|u_1\rangle = \lambda_{u_2}\langle u_2|u_1\rangle \ . \tag{70}$$

Comparing this to (68) we have

$$\lambda_{u_1}\langle u_2|u_1\rangle = \lambda_{u_2}\langle u_2|u_1\rangle \ , \tag{71}$$

which for $\lambda_{u_1} \neq \lambda_{u_2}$ implies

$$\langle u_2|u_1\rangle = 0 \ . \tag{72}$$

In the case $\lambda_{u_1} = \lambda_{u_2}$, the vectors $|u_1\rangle$ and $|u_2\rangle$ are not necessarily orthogonal. What is important, however, is that all the eigenvectors of the same eigenvalue λ form a *vector sub-space* of the original vector space, because any linear combination of two eigenvectors with one and the same eigenvalue λ makes another eigenvector with the same eigenvalue λ. Within this subspace one can choose an orthogonal basis. This way we arrive at an orthogonal basis of the eigenvectors of a self-adjoint operator, which is very useful for many practical applications.