B. A FEW WORDS ON SPACES, VECTORS AND FUNCTIONS

1 VECTOR SPACE

A vector space means a set V of elements $x, y, \ldots$, that form an Abelian group and can be "added" together[1] and "multiplied" by numbers $z = \alpha x + \beta y$ thus producing $z \in V$. The multiplication (α, β are, in general, complex numbers) satisfies the usual rules (the group is Abelian, because $x + y = y + x$):

$1 \cdot x = x,$
$\alpha(\beta x) = (\alpha\beta)x,$
$\alpha(x + y) = \alpha x + \alpha y,$
$(\alpha + \beta)x = \alpha x + \beta x.$

Example 1. ***Integers.*** The elements $x, y, \ldots$ are integers, the "addition" means simply the usual addition of integers, the numbers $\alpha, \beta, \ldots$ are also integers, "multiplication" means just usual multiplication. Does the set of integers form a vector space? Let us see. The integers form a group (with the addition as the operation in the group). Checking all the above axioms, we can easily prove that they are satisfied by integers. Thus, the integers (with the operations defined above) form a vector space.

Example 2. ***Integers with real multipliers.*** If, in the previous example, we admitted α, β to be real, the multiplication of integers x, y by real numbers would give real numbers (not necessarily integers). Therefore, in this case $x, y, \ldots$ *do not* represent any vector space.

Example 3. ***Vectors.*** Suppose $x, y, \ldots$ are vectors, each represented by a N-element sequence of real numbers (they are called the vector "components") $x = (a_1, a_2, \ldots, a_N)$, $y = (b_1, b_2, \ldots, b_N)$, etc. Their addition $x + y$ is an operation that produces the vector $z = (a_1 + b_1, a_2 + b_2, \ldots, a_N + b_N)$. The vectors form an Abelian group, because $x + y = y + x$, the unit ("neutral") element is $(0, 0, \ldots, 0)$, the inverse element to $(a_1, a_2, \ldots, a_N)$ is equal to $(-a_1, -a_2, \ldots, -a_N)$. Thus, the vectors form a group. "Multiplication" of a vector by a real number α means $\alpha(a_1, a_2, \ldots, a_N) = (\alpha a_1, \alpha a_2, \ldots, \alpha a_N)$. Please check that the four axioms above are satisfied. Conclusion: the vectors form a vector space.

[1] See Appendix C, to form a group any pair of the elements can be "added" (operation in the group), the addition is associative, there exists a unit element and for each element an inverse exists.

Note that if only the positive vector components were allowed, they would not form an Abelian group (no neutral element), and on top of this their addition (which might mean a subtraction of components, because α, β could be negative) could produce vectors with non-positive components. Thus vectors with all positive components do not form a vector space.

Example 4. ***Functions.*** This example is important in the context of this book. This time the vectors have real components.[2] Their "addition" means the addition of two functions $f(x) = f_1(x) + f_2(x)$. The "multiplication" means multiplication by a real number. The unit ("neutral") function means $f = 0$, the "inverse" function to f is $-f(x)$. Therefore, the functions form an Abelian group. A few seconds are needed to show that the four axioms above are satisfied. Such functions form a vector space.

Linear independence. A set of vectors is called a set of linearly independent vectors if no vector of the set can be expressed as a linear combination of the other vectors of the set. The number of linearly independent vectors in a vector space is called the dimension of the space.

Basis means a set of n linearly independent vectors in n-dimensional space.

2 EUCLIDEAN SPACE

A vector space (with multiplying real numbers α, β) represents the Euclidean space, if for any two vectors x, y of the space we assign a real number called an *inner* (*or scalar*) *product* $\langle x|y\rangle$ with the following properties:

- $\langle x|y\rangle = \langle y|x\rangle$,
- $\langle \alpha x|y\rangle = \alpha\langle x|y\rangle$,
- $\langle x_1 + x_2|y\rangle = \langle x_1|y\rangle + \langle x_2|y\rangle$,
- $\langle x|x\rangle = 0$, only if $x = 0$.

Inner product and distance. The concept of the inner product is used to introduce

- the *length of the vector* x defined as $\|x\| \equiv \sqrt{x|x}$, and
- the *distance between two vectors* x and y as a non-negative number $\|x - y\| = \sqrt{\langle x - y|x - y\rangle}$. The distance satisfies some conditions which we treat as obvious from everyday experience:
- the distance from Paris to Paris has to equal zero (just insert $x = y$);
- the distance from Paris to Rome has to be the same as from Rome to Paris (just exchange $x \leftrightarrow y$);
- the Paris–Rome distance is equal to or shorter than the sum of two distances: Paris–X and X–Rome for any town X (a little more difficult to show).

[2] Note the similarity of the present example with the previous one: a function $f(x)$ may be treated as a vector with an infinite number of components. The components are listed in the sequence of increasing $x \in R$, the component $f(x)$ corresponding to x.

Schwarz inequality. For any two vectors belonging to the Euclidean space the Schwarz inequality holds[3]

$$\big|\langle x|y\rangle\big| \leqslant \|x\|\|y\| \tag{B.1}$$

or equivalently $|\langle x|y\rangle|^2 \leqslant \|x\|^2\|y\|^2$.

Orthogonal basis means that all basis vectors x_j, $j = 1, 2, \ldots, N$, are orthogonal to each other: $\langle x_i|x_j\rangle = 0$ for $i \neq j$.

Orthonormal basis is an orthogonal basis set with all the basis vectors of length $\|x_i\| = 1$. Thus for the orthonormal basis set we have $\langle x_i|x_j\rangle = \delta_{ij}$, where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$ (Kronecker delta).

Example 5. ***Dot product.*** Let us take the vector space from Example 3 and introduce the dot product (representing the inner product) defined as

$$\langle x|y\rangle = \sum_{i=1}^{N} a_i b_i. \tag{B.2}$$

Let us check whether this definition satisfies the properties required for a inner product:

- $\langle x|y\rangle = \langle y|x\rangle$, because the order of a and b in the product is irrelevant.
- $\langle \alpha x|y\rangle = \alpha\langle x|y\rangle$, because the sum says that multiplication of each a_i by α is equivalent to multiplying the inner product by α.
- $\langle x_1 + x_2|y\rangle = \langle x_1|y\rangle + \langle x_2|y\rangle$, because if vector x is decomposed into two vectors $x = x_1 + x_2$ in such a way that $a_i = a_{i1} + a_{i2}$ (with a_{i1}, a_{i2} being the components of x_1, x_2, respectively), the summation of $\langle x_1|y\rangle + \langle x_2|y\rangle$ gives $\langle x|y\rangle$.
- $\langle x|x\rangle = \sum_{i=1}^{N}(a_i)^2$, and this equals zero if, and only if, all components $a_i = 0$. Therefore, the proposed formula operates as the inner product definition requires.

3 UNITARY SPACE

If three changes were introduced into the definition of the Euclidean space, we would obtain another space: the unitary space. These changes are as follows:

- the numbers $\alpha, \beta, \ldots$ instead of being real are complex;
- the inner product, instead of $\langle x|y\rangle = \langle y|x\rangle$ has the property $\langle x|y\rangle = \langle y|x\rangle^*$;
- instead of $\langle \alpha x|y\rangle = \alpha\langle x|y\rangle$ we have:[4] $\langle \alpha x|y\rangle = \alpha^*\langle x|y\rangle$.

[3] The Schwarz inequality agrees with what everyone recalls about the dot product of two vectors: $\langle x|y\rangle = \|x\|\|y\|\cos\theta$, where θ is the angle between the two vectors. Taking the absolute value of both sides, we obtain $|\langle x|y\rangle| = \|x\|\|y\||\cos\theta| \leqslant \|x\|\|y\|$.

[4] While we still have $\langle x|\alpha y\rangle = \alpha\langle x|y\rangle$.

After the new inner product definition is introduced, related quantities: the length of a vector and the distance between the vectors are defined in exactly the same way as in the Euclidean space. Also the definitions of orthogonality and of the Schwarz inequality remain unchanged.

4 HILBERT SPACE

This is for us the most important unitary space – its elements are wave functions, which instead of $x, y, \ldots$ will be often denoted as $f, g, \ldots, \phi, \chi, \psi, \ldots$ etc. The wave functions which we are dealing with in quantum mechanics (according to John von Neumann) are the elements (i.e. vectors) of the Hilbert space. The inner product of two functions f and g means $\langle f|g\rangle \equiv \int f^* g \, d\tau$, where the integration is over the whole space of variables, on which both functions depend. The length of vector f is denoted by $\|f\| = \sqrt{\langle f|f\rangle}$. Consequently, the orthogonality of two functions f and g means $\langle f|g\rangle = 0$, i.e. an integral $\int f^* g \, d\tau = 0$ over the whole range of the coordinates on which the function f depends. The Dirac notation (1.9) is in fact the inner product of such functions in a unitary space.

David Hilbert (1862–1943), German mathematician, professor at the University of Göttingen. At the II Congress of Mathematicians in Paris Hilbert formulated 23 goals for mathematics he considered to be very important. This had a great impact on mathematics and led to some unexpected results (e.g., Gödel theorem, cf. p. 851). Hilbert's investigations in 1900–1910 on integral equations resulted in the concept of the Hilbert space. Hilbert also worked on the foundations of mathematics, on mathematical physics, number theory, variational calculus, etc. This hard working and extremely prolific mathematician was deeply depressed by Hitler's seizure of power. He regularly came to his office, but did not write a single sheet of paper.

Let us imagine an infinite sequence of functions (i.e. vectors) $f_1, f_2, f_3, \ldots$ in a unitary space, Fig. B.1. The sequence will be called a Cauchy sequence, if for a given $\varepsilon > 0$ a natural number N can be found, such that for $i > N$ we will have $\|f_{i+1} - f_i\| < \varepsilon$. In other words, in a Cauchy sequence the distances between consecutive vectors (functions) decrease when we go to sufficiently large indices, i.e. the functions become more and more similar to each other. *If the converging Cauchy sequences have their limits (functions) which belong to the unitary space, such a space is called a Hilbert space.*

A basis in the Hilbert space is such a set of linearly independent functions (vectors) that any function belonging to the space can be expressed as a linear combination of the basis set functions. Because of the infinite number of dimensions, the number of the basis set functions is infinite. This is difficult to imagine. In a way analogous to a 3D Euclidean space, we may imagine an orthonormal basis as the unit vectors protruding from the origin in an infinite number of directions (like a "hedgehog", Fig. B.2).

Each vector (function) can be represented as a linear combination of the hedgehog functions. We see that we may rotate the "hedgehog" (i.e. the basis set)[5] and the completeness of the basis will be preserved, i.e. any vector of the Hilbert space can be represented as a linear combination of the new basis set vectors.

[5]The new orthonormal basis set is obtained by a unitary transformation of the old one.

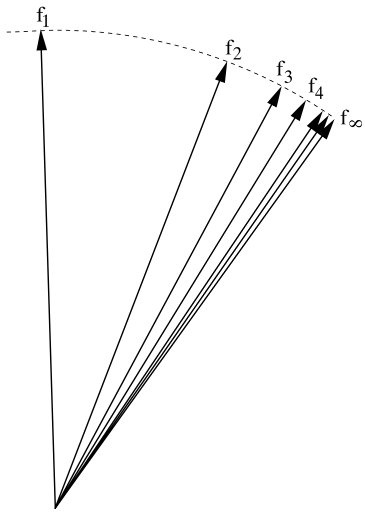

Fig. B.1. A pictorial representation of the Hilbert space. We have a vector space (each vector represents a wave function) and a series of unit vectors f_i that differ less and less (Cauchy series). If any convergent Cauchy series has its limit belonging to the vector space, the space represents the *Hilbert space.*

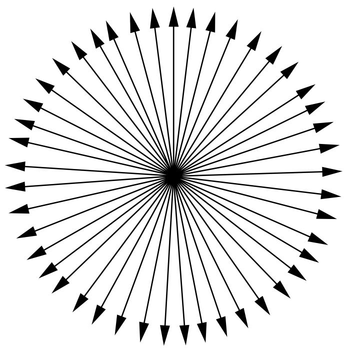

Fig. B.2. A pictorial representation of something that surely cannot be represented: an orthonormal basis in the Hilbert space looks like a hedgehog of the unit vectors (their number equal to ∞), each pair of them orthogonal. This is analogous to a 2D or 3D basis set, where the hedgehog has two or three orthogonal unit vectors.

Linear operator

Operator $\hat{A}$ transforms any vector x from the operator's domain into vector y (both vectors x, y belong to the unitary space): $\hat{A}(x) = y$, which is written as $\hat{A}x = y$. A *linear operator* satisfies $\hat{A}(c_1x_1 + c_2x_2) = c_1\hat{A}x_1 + c_2\hat{A}x_2$, where c_1 and c_2 stand for complex numbers.

We define:

- Sum of operators: $\hat{A} + \hat{B} = \hat{C}$ as $\hat{C}x = \hat{A}x + \hat{B}x$.
- Product of operators: $\hat{A}\hat{B} = \hat{C}$ as $\hat{C}x = \hat{A}(\hat{B}(x))$.

If, for two operators, we have $\hat{A}\hat{B} = \hat{B}\hat{A}$, we say they *commute*, or their *commutator* $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = 0$. In general $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, i.e. the operators do not commute. commutation

- Inverse operator (if it exists): $\hat{A}^{-1}(\hat{A}x) = x$.

Adjoint operator

If, for an operator $\hat{A}$, we can find a new operator $\hat{A}^{\dagger}$, such that for any two vectors x and y of the unitary space[6] we have[7]

$$\langle x|\hat{A}y\rangle = \langle \hat{A}^{\dagger}x|y\rangle \tag{B.3}$$

then we say that $\hat{A}^{\dagger}$ is the *adjoint* operator to $\hat{A}$.

Hermitian operator

If $\hat{A}^{\dagger} = \hat{A}$, we will call operator $\hat{A}$ a *self-adjoint* or *Hermitian* operator:[8]

$$\langle x|\hat{A}y\rangle = \langle \hat{A}x|y\rangle. \tag{B.4}$$

Unitary operator

A unitary operator $\hat{U}$ transforms a vector x into $y = \hat{U}x$ both belonging to the unitary space (the domain is the unitary space) and the *inner product is preserved*:

$$\langle \hat{U}x|\hat{U}y\rangle = \langle x|y\rangle.$$

This means that any unitary transformation preserves the angle between the vectors x and y, i.e. the angle between x and y is the same as the angle between $\hat{U}x$ and $\hat{U}y$. The transformation also preserves the length of the vector, because $\langle \hat{U}x|\hat{U}x\rangle = \langle x|x\rangle$. This is why operator $\hat{U}$ can be thought of as a *transformation related to a motion in the unitary space* (rotation, reflection, etc.). For a unitary operator we have $\hat{U}^{\dagger}\hat{U} = 1$, because $\langle \hat{U}x|\hat{U}y\rangle = \langle x|\hat{U}^{\dagger}\hat{U}y\rangle = \langle x|y\rangle$.

5 EIGENVALUE EQUATION

If, for a particular vector x, we have

$$\hat{A}x = ax, \tag{B.5}$$

where a is a complex number and $x \neq 0$, x is called the eigenvector[9] of operator $\hat{A}$ corresponding to eigenvalue a. Operator $\hat{A}$ may have an infinite number, a finite number including number zero of the eigenvalues, labelled by subscript i:

[6] The formal definition is less restrictive and the domains of the operators $\hat{A}^{\dagger}$ and $\hat{A}$ do not need to extend over the whole unitary space.

[7] Sometimes we make a useful modification in the Dirac notation: $\langle x|\hat{A}y\rangle \equiv \langle x|\hat{A}|y\rangle$.

[8] The self-adjoint and Hermitian operators differ in mathematics (a matter of domains), but we will ignore this difference in the present book.

[9] In quantum mechanics, vector x will correspond to a function (a vector in the Hilbert space) and therefore is called the eigenfunction.

$$\hat{A}x_i = a_i x_i.$$

Hermitian operators have the following important properties:[10]

> If $\hat{A}$ represents a Hermitian operator, its eigenvalues a_i are real numbers, and its eigenvectors x_i, which correspond to different eigenvalues, are orthogonal.

The number of linearly independent eigenvectors which correspond to a given eigenvalue a is called the degree of degeneracy of the eigenvalue. Such vectors form the basis of the *invariant space of operator* $\hat{A}$, i.e. any linear combination of the vectors represents a vector that is also an eigenvector (with the same eigenvalue a). If the eigenvectors corresponded to different eigenvalues, their linear combination *is not* an eigenvector of $\hat{A}$. Both things need a few seconds to show. degeneracy

One can show that the eigenvectors of a Hermitian operator form the complete basis set[11] in Hilbert space, i.e. any function of class Q[12] can be expanded in a linear combination of the basis set.

Sometimes an eigenvector x of operator $\hat{A}$ (with eigenvalue a) is subject to an operator $f(\hat{A})$, where f is an analytic function. Then,[13]

$$f(\hat{A})x = f(a)x. \tag{B.6}$$

Commutation and eigenvalues

We will sometimes use the theorem that, if two linear and Hermitian operators $\hat{A}$ and $\hat{B}$ commute, they have a common set of eigenfunctions and *vice versa*.

[10] We have the eigenvalue problem $\hat{A}x = ax$. Making a complex conjugate of both sides, we obtain $(\hat{A}x)^* = a^*x^*$. Multiplying the first of the equations by x^* and integrating, and then using x and doing the same with the second equation, we get: $\langle x|\hat{A}x\rangle = a\langle x|x\rangle$ and $\langle \hat{A}x|x\rangle = a^*\langle x|x\rangle$. But $\hat{A}$ is Hermitian, and therefore the left-hand sides of both equations are equal. Subtracting them we have $(a-a^*)\langle x|x\rangle = 0$. Since $\langle x|x\rangle \neq 0$, because $x \neq 0$, then $a = a^*$. This is what we wanted to show.

The orthogonality of the eigenfunctions of a Hermitian operator (corresponding to different eigenvalues) may be proved as follows. We have $\hat{A}x_1 = a_1x_1$, $\hat{A}x_2 = a_2x_2$, with $a_1 \neq a_2$. Multiplying the first equation by x_2^* and integrating, we obtain $\langle x_2|\hat{A}x_1\rangle = a_1\langle x_2|x_1\rangle$. Then, let us make the complex conjugate of the second equation: $(\hat{A}x_2)^* = a_2x_2^*$, where we have used $a_2 = a_2^*$ (this was proved above). Then let us multiply by x_1 and integrate: $\langle \hat{A}x_2|x_1\rangle = a_2\langle x_2|x_1\rangle$. Subtracting the two equations, we have $0 = (a_1 - a_2)\langle x_2|x_1\rangle$, and taking into account $a_1 - a_2 \neq 0$ this gives $\langle x_2|x_1\rangle = 0$.

[11] This basis set may be assumed to be orthonormal, because the eigenfunctions

- as square-integrable can be normalized;
- if they correspond to different eigenvalues, are automatically orthogonal;
- if they correspond to the same eigenvalue, they can be orthogonalized (still remaining eigenfunctions) by a method described in Appendix J.

[12] That is, continuous, single-valued and square integrable, see Fig. 2.5.

[13] The operator $f(\hat{A})$ is defined through the Taylor expansion of the function f: $f(\hat{A}) = c_0 + c_1\hat{A} + c_2\hat{A}^2 + \cdots$. If the operator $f(\hat{A})$ now acts on an eigenfunction of $\hat{A}$, then, because $\hat{A}^n x = a^n x$, we obtain the result.

We will prove this theorem in the case of no degeneracy (i.e. there is only one linearly independent vector for a given eigenvalue). We have an eigenvalue equation $\hat{B}y_n = b_n y_n$. Applying this to both sides of operator $\hat{A}$ and using the commutation relation $\hat{A}\hat{B} = \hat{B}\hat{A}$ we have: $\hat{B}(\hat{A}y_n) = b_n(\hat{A}y_n)$. This means that $\hat{A}y_n$ is an eigenvector of $\hat{B}$ corresponding to the eigenvalue b_n. But, we already know such a vector, this is y_n. The two vectors have to be proportional: $\hat{A}y_n = a_n y_n$, which means that y_n is an eigenvector of $\hat{A}$.

Now, the inverse theorem. We have two operators and any eigenvector of $\hat{A}$ is also an eigenvector of $\hat{B}$. We want to show that the two operators commute. Let us write the two eigenvalue equations: $\hat{A}y_n = a_n y_n$ and $\hat{B}y_n = b_n y_n$. Let us take a vector ϕ. Since the eigenvectors $\{y_n\}$ form the complete set, then

$$\phi = \sum_n c_n y_n.$$

Applying the commutator $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ we have

$$\begin{aligned}
[\hat{A}, \hat{B}]\phi &= \hat{A}\hat{B}\phi - \hat{B}\hat{A}\phi = \hat{A}\hat{B}\sum_n c_n y_n - \hat{B}\hat{A}\sum_n c_n y_n \\
&= \hat{A}\sum_n c_n \hat{B}y_n - \hat{B}\sum_n c_n \hat{A}y_n = \hat{A}\sum_n c_n b_n y_n - \hat{B}\sum_n c_n a_n y_n \\
&= \sum_n c_n b_n \hat{A}y_n - \sum_n c_n a_n \hat{B}y_n = \sum_n c_n b_n a_n y_n - \sum_n c_n a_n b_n y_n = 0.
\end{aligned}$$

This means that the two operators commute.