functionals in the dual space V' are called *bra* vectors, and are denoted as $\langle F|$. The numerical value of the functional is denoted as

$$F(\phi) = \langle F|\phi\rangle \,. \tag{1.6}$$

According to the Riesz theorem, there is a one-to-one correspondence between bras and kets. Therefore we can use the same alphabetic character for the functional (a member of V') and the vector (in V) to which it corresponds, relying on the bra, $\langle F|$, or ket, $|F\rangle$, notation to determine which space is referred to. Equation (1.5) would then be written as

$$\langle F|\phi\rangle = (F, \phi) \,, \tag{1.7}$$

$|F\rangle$ being the vector previously denoted as f. Note, however, that the Riesz theorem establishes, by construction, an *antilinear* correspondence between bras and kets. If $\langle F| \leftrightarrow |F\rangle$, then

$$c_1^*\langle F| + c_2^*\langle F| \leftrightarrow c_1|F\rangle + c_2|F\rangle \,. \tag{1.8}$$

Because of the relation (1.7), it is possible to regard the “braket” $\langle F|\phi\rangle$ as merely another notation for the inner product. But the reader is advised that there are situations in which it is important to remember that the primary definition of the bra vector is as a linear functional on the space of ket vectors.

[[In his original presentation, Dirac *assumed* a one-to-one correspondence between bras and kets, and it was not entirely clear whether this was a mathematical or a physical assumption. The Riesz theorem shows that there is no need, and indeed no room, for any such assumption. Moreover, we shall eventually need to consideer more general spaces (rigged-Hilbert-space triplets) for which the one-to-one correspondence between bras and kets does not hold.]]

1.2 *Linear Operators*

An *operator* on a vector space maps vectors onto vectors; that is to say, if A is an opetator and ψ is a vector, then $\phi = A\psi$ is another vector. An operator is fully defined by specifying its action on every vector in the space (or in its *domain*, which is the name given to the subspace on which the operator can meaningfully act, should that be smaller than the whole space).

A *linear operator* satisfies

$$A(c_1\psi_1 + c_2\psi_2) = c_1(A\psi_1) + c_2(A\psi_2) \,. \tag{1.9}$$

It is sufficient to define a linear operator on a set of basis vectors, since everly vector can be expressed as a linear combination of the basis vectors. We shall be treating only linear operators, and so shall henceforth refer to them simply as operators.

To assert the *equality* of two operators, $A = B$, means that $A\psi = B\psi$ for *all* vectors (more precisely, for all vectors in the common domain of A and B, this qualification will usually be omitted for brevity). Thus we can define the sum and product of operators,

$$(A + B)\psi = A\psi + B\psi\,,$$
$$AB\psi = A(B\psi)\,,$$

both equations holding for all ψ. It follows from this definition that operator mulitplication is necessarily *associative*, $A(BC) = (AB)C$. But it need not be *commutative*, AB being unequal to BA in general.

Example (i). In a space of discrete vectors represented as columns, a linear operator is a square matrix. In fact, any operator equation in a space of N dimensions can be transformed into a matrix equation. Consider, for example, the equation

$$M|\psi\rangle = |\phi\rangle\,. \tag{1.10}$$

Choose some orthonormal basis $\{|u_i\rangle, i = 1 \ldots N\}$ in which to expand the vectors,

$$|\psi\rangle = \sum_j a_j|u_j\rangle\,,\ |\phi\rangle = \sum_k b_k|u_k\rangle\,.$$

Operating on (1.10) with $\langle u_i|$ yields

$$\begin{aligned}\sum_j \langle u_i|M|u_j\rangle a_j &= \sum_k \langle u_i|u_k\rangle b_k \\ &= b_i\,,\end{aligned}$$

which has the form of a matrix equation,

$$\sum_j M_{ij} a_j = b_i\,, \tag{1.11}$$

with $M_{ij} = \langle u_i|M|u_j\rangle$ being known as a *matrix element* of the operator M. In this way any problem in an N-dimensional linear vector space, no matter how it arises, can be transformed into a matrix problem.

The same thing can be done formally for an infinite-dimensional vector space if it has a denumerable orthonormal basis, but one must then deal with the problem of convergence of the infinite sums, which we postpone to a later section.

Example (ii). Operators in function spaces frequently take the form of differential or integral operators. An operator equation such as

$$\frac{\partial}{\partial x} x = 1 + x \frac{\partial}{\partial x}$$

may appear strange if one forgets that operators are only defined by their action on vectors. Thus the above example means that

$$\frac{\partial}{\partial x}[x \ \psi(x)] = \psi(x) + x \frac{\partial \psi(x)}{\partial x} \quad \text{for all} \ \ \psi(x)\,.$$

So far we have only defined operators as acting to the right on ket vectors. We may define their *action to the left* on bra vectors as

$$(\langle\phi|A)|\psi\rangle = \langle\phi|(A|\psi\rangle) \tag{1.12}$$

for all ϕ and ψ. This appears trivial in Dirac's notation, and indeed this triviality contributes to the practival utility of his notation. However, it is worthwhile to examine the mathematical content of (1.12) in more detail.

A bra vector is in fact a linear functional on the space of ket vectors, and in a more detailed notation the bra $\langle\phi|$ is the functional

$$F_\phi(\cdot) = (\phi, \cdot)\,, \tag{1.13}$$

where ϕ is the vector that corresponds to F_ϕ via the Riesz theorem, and the dot indicates the place for the vector argument. We may define the operation of A on the bra space of functionals as

$$AF_\phi(\psi) = F_\phi(A\psi) \quad \text{for all} \ \ \psi\,. \tag{1.14}$$

The right hand side of (1.14) satisfies the definition of a linear functional of the vector ψ (not merely of the vector $A\psi$), and hence it does indeed define a new functional, called AF_ϕ. According to the Riesz theorem there must exist a ket vector χ such that

$$\begin{aligned} AF_\phi(\psi) &= (\chi, \psi) \\ &= F_\chi(\psi)\,. \end{aligned} \tag{1.15}$$

Since χ is uniquely determined by ϕ (given A), there must exist an operator $A^\dagger$ such that $\chi = A^\dagger\phi$. Thus (1.15) can be written as

$$AF_\phi = F_{A^\dagger\phi}\,. \tag{1.16}$$

From (1.14) and (1.15) we have $(\phi, A\psi) = (\chi, \psi)$, and therefore

$$(A^\dagger\phi, \psi) = (\phi, A\psi) \quad \text{for all} \quad \phi \text{ and } \psi\,. \tag{1.17}$$

This is the usual definition of the *adjoint*, $A^\dagger$, of the operator A. All of this nontrivial mathematics is implicit in Dirac's simple equation (1.12)!

The adjoint operator can be formally defined within the Dirac notation by demanding that if $\langle\phi|$ and $|\phi\rangle$ are corresponding bras and kets, then $\langle\phi|A^\dagger \equiv \langle\omega|$ and $A|\phi\rangle \equiv |\omega\rangle$ should also be corresponding bras and kets. From the fact that $\langle\omega|\psi\rangle^* = \langle\psi|\omega\rangle$, it follows that

$$\langle\phi|A^\dagger|\psi\rangle^* = \langle\psi|A|\phi\rangle \quad \text{for all} \quad \phi \text{ and } \psi\,, \tag{1.18}$$

this relation being equivalent to (1.17). Although simpler than the previous introduction of $A^\dagger$ via the Riesz theorem, this formal method fails to prove the existence of the operator $A^\dagger$.

Several useful properties of the adjoint operator that follow directly from (1.17) are

$$\begin{aligned}
(cA)^\dagger &= c^* A^\dagger\,, \quad \text{where } c \text{ is a complex number,}\\
(A+B)^\dagger &= A^\dagger + B^\dagger\,,\\
(AB)^\dagger &= B^\dagger A^\dagger\,.
\end{aligned}$$

In addition to the inner product of a bra and a ket, $\langle\phi|\psi\rangle$, which is a scalar, we may define an *outer product*, $|\psi\rangle\langle\phi|$. This object is an operator because, assuming associative multiplication, we have

$$(|\psi\rangle\langle\phi|)|\lambda\rangle = |\psi\rangle(\langle\phi|\lambda\rangle)\,. \tag{1.19}$$

Since an operator is defined by specifying its action on an arbitrary vector to produce another vector, this equation fully defines $|\psi\rangle\langle\phi|$ as an operator. From (1.18) it follows that

$$(|\psi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\psi|\,. \tag{1.20}$$

In view of this relation, it is tempting to write $(|\psi\rangle)^\dagger = \langle\psi|$. Although no real harm comes from such a notation, it should not be encouraged because it uses the "adjoint" symbol, †, for something that is not an operator, and so cannot satisfy the fundamental definition (1.16).

A useful characteristic of an operator A is its *trace*, defined as

$$\mathrm{Tr}\ A = \sum_j \langle u_j|A|u_j\rangle\,,$$

where $\{|u_j\rangle\}$ may be any orthonormal basis. It can be shown [see Problem (1.3)] that the value of Tr A is independent of the particular orthonormal basis that is chosen for its evaluation. The trace of a matrix is just the sum of its diagonal elements. For an operator in an infinite-dimensional space, the trace exists only if the infinite sum is convergent.

1.3 *Self-Adjoint Operators*

An operator A that is equal to its adjoint $A^\dagger$ is called *self-adjoint*. This means that it satisfies

$$\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle* \tag{1.21}$$

and that the domain of A (i.e. the set of vectors ϕ on which $A\phi$ is well defined) coincides with the domain of $A^\dagger$. An operator that only satisfies (1.21) is called *Hermitian*, in analogy with a Hermitian matrix, for which $M_{ij} = M_{ji}*$.

[[The distinction between Hermitian and self-adjoint operators is relevant only for operators in infinite-dimensional vector spaces, and we shall make such a distinction only when it is essential to do so. The operators that we call "Hermitian" are often called "symmetric" in the mathematical literature. That terminology is objectionable because it conflicts with the corresponding properties of matrices.]]

The following theorem is useful in identifying Hermitian operators on a vector space with complex scalars.

Theorem 1. If $\langle\psi|A|\psi\rangle = \langle\psi|A|\psi\rangle^*$ for all $|\psi\rangle$, then it follows that $\langle\phi_1|A|\phi_2\rangle = \langle\phi_2|A|\phi_1\rangle^*$ for all $|\phi_1\rangle$ and $|\phi_2\rangle$, and hence that $A = A^\dagger$.

Proof. Let $|\psi\rangle = a|\phi_1\rangle + b|\phi_2\rangle$ for arbitrary a, b, $|\phi_1\rangle$, and $|\phi_2\rangle$.

Then

$$\begin{aligned}\langle\psi|A|\psi\rangle &= |a|^2\langle\phi_1|A|\phi_1\rangle + |b|^2\langle\phi_2|A|\phi_2\rangle \\ &\quad + a^*b\langle\phi_1|A|\phi_2\rangle + b^*a\langle\phi_2|A|\phi_1\rangle\end{aligned}$$