CHAPTER 2 : HILBERT SPACE AND LINEAR OPERATORS

In the preceding chapter we have given a description of the scattering process in physical terms. In this chapter we shall begin to transcribe this description into the mathematical language of quantum mechanics. We have seen that in the distant past and the remote future the time evolution of a scattering system is physically indistinguishable from that of a system of free particles. Our first task is then to specify the mathematical concepts and theorems needed for the description of particles, especially of free particles, and of their temporal evolution. This is the topic of the present and the following chapter.

The basic mathematical object which is needed for the description of particles in quantum mechanics is the Hilbert space. In Section 2-1 we introduce the abstract Hilbert space, discuss some of its elementary properties and illustrate these notions in L^2-spaces. In Section 2-2 we present various simple results concerning linear operators in Hilbert space. The emphasis is on getting acquainted with bounded and with self-adjoint operators. Section 2-3 is devoted to some particular classes of operators, namely compact, Hilbert-Schmidt and trace class operators. Finally in Section 2-4 we introduce the tensor product and the direct sum of Hilbert spaces.

The results of Sections 2-3 and 2-4 are not used before Chapter 7, and the reader may skip these sections here and familiarize himself with their material when arriving at Chapter 7.

2-1 THE ABSTRACT HILBERT SPACE AND ITS CONCRETE REALIZATIONS

The abstract Hilbert space $\mathcal{H}$ is a collection of objects called vectors, denoted by f, g,..., which satisfy the following three axioms.

I. $\mathcal{H}$ is a linear vector space with complex coefficients.

This means that to every pair of vectors $f,g\in\mathcal{H}$ there is associated a third vector $(f + g)\in\mathcal{H}$. Furthermore to every vector f and every complex number α there corresponds another vector $\alpha f\in\mathcal{H}$. The following rules are postulated :

$$f + g = g + f, \quad (f + g) + h = f + (g + h), \tag{2.1}$$

$$\alpha(f + g) = \alpha f + \alpha g, \quad (\alpha + \beta)f = \alpha f + \beta f, \quad \alpha(\beta f) = (\alpha\beta)f, \tag{2.2}$$

$$1 \cdot f = f. \tag{2.3}$$

There exists a unique vector θ in $\mathcal{H}$ such that for all $f\in\mathcal{H}$

$$\theta + f = f, \quad 0 \cdot f = \theta. \tag{2.4}$$

II. There exists a strictly positive scalar product in H.

The scalar product (f,g) is a function of pairs of vectors $f,g \in H$ with values in the set C of complex numbers and satisfies the following conditions *) :

$$(f,g) = \overline{(g,f)}, \tag{2.5}$$

$$(f,g + \alpha h) = (f,g) + \alpha(f,h) \text{ for all complex } \alpha, \tag{2.6}$$

$$\|f\| \equiv (f,f)^{\frac{1}{2}} > 0 \quad \text{unless } f = \theta. \tag{2.7}$$

(The bar in (2.5) denotes complex-conjugation.)

III. The space H is complete in the norm defined by (2.7) : Whenever $\{f_n\}$, $n = 1, 2,\ldots$, is a Cauchy sequence in the sense that $\|f_n - f_m\| \to 0$ as $n,m \to \infty$, there exists a vector $f \in H$ such that $\|f_n - f\| \to 0$ as $n \to \infty$.

In this book we shall deal only with separable Hilbert spaces, i.e. we shall postulate also the following axiom :

IV. The space H is separable. This means that there exists a sequence $\{f_n\} \in H$ $(n = 1, 2,\ldots)$ with the property that it is dense in H. We recall that a subset D of H is dense in H if, for any $f \in H$ and any $\eta > 0$, there exists at least one element f_η in D such that $\|f - f_\eta\| < \eta$.

The last two requirements are topological in nature. They limit the size of the space in opposite directions. The first one can always be satisfied by a standard technique of

---

*) In conformity with the practice of most physicists we shall consider the scalar product (f,g) to be linear in g and anti-linear in f, whereas the convention in the mathematical literature is usually the opposite.

adjunction of suitable limit elements; i.e. if a collection of vectors verifies Axioms I and II, one can convert it into a Hilbert space by adding suitable limit elements (this will be used in Section 2-4). The second one is a genuine restriction, requiring that there be a countable dense set $\mathcal{D}$ in $\mathcal{H}$. So far non-separable spaces have not been needed in quantum mechanics. We may add that the property of denseness in a metric space may be visualized by considering the example of the real line R with the usual Euclidean metric where the set of rational numbers is dense.

Before we introduce the L^2-spaces as concrete examples of Hilbert spaces, we collect some additional definitions and a few elementary relations that follow from the above axioms. These will be useful at many later stages of this book. The reader may also look at Problem 2.1 for other examples of Hilbert spaces.

We begin with two simple inequalities. The first one is the Schwarz inequality :

$$|(f,g)| \leq \|f\| \, \|g\| \, . \tag{2.8}$$

For its proof we distinguish two cases. (a) If $f = g$, (2.8) holds with the equality sign by the definition (2.7). (b) If $f \neq g$, we may assume for instance that $g \neq \theta$. For any complex number α one has (cf. also Problem 2.2) $0 \leq \|f + \alpha g\|^2 = (f + \alpha g, f + \alpha g) = \|f\|^2 + |\alpha|^2 \|g\|^2 + \alpha(f,g) + \bar{\alpha}(g,f)$. (2.8) is a particular case of this obtained by setting $\alpha = -(g,f)/\|g\|^2$ and multiplying the resulting inequality by $\|g\|^2$.

As a consequence of (2.8) one obtains the Minkowski inequality which is also called the triangle inequality :

$$\|f + g\| \leq \|f\| + \|g\|. \tag{2.9}$$

Its proof is simple : $\|f + g\|^2 = \|f\|^2 + \|g\|^2 + (f,g) + (g,f) \leq \|f\|^2 + \|g\|^2 + 2|(f,g)| \leq \|f\|^2 + \|g\|^2 + 2\|f\|\,\|g\| = (\|f\| + \|g\|)^2$.

The triangle inequality implies together with axioms I and II that the Hilbert space is a normed linear space. The norm $\|f\|$ of a vector f is a measure of the distance between f and the zero vector θ, or in other words $\|f - g\|$ is a mesure of the distance between f and g. Since the vectors of H will be interpreted as the pure states of some physical system, two states f and g are practically indistinguishable if $\|f - g\|$ is very small. This suggests that one could use this norm in order to express the asymptotic properties of a scattering system, viz. the practical indistinguishability of the real system from a free system when $|t|$ is very large. This will be done in Chapter 4. The precise mathematical statement will be a limit relation : the larger $|t|$ the less the real state should differ from a freely evolving state. For this reason we shall now look at convergence properties of sequences $\{f_n\}$ of elements of H.

In order to define the convergence of a sequence of vectors of H, one resorts to the simpler notion of convergence of a sequence of (real or complex) numbers. So far we have introduced two kinds of numbers constructed from elements of H, the norm of a vector and the scalar product between two vectors. Each of these can be used to define a topology on H.

The convergence of a sequence of vectors in the norm $||\cdot||$ has already been used in formulating Axiom III. In Hilbert space theory this is called strong convergence. A sequence of vectors $\{f_n\}$ converges strongly to a limit vector f if $||f - f_n|| \to 0$ for $n \to \infty$. We then write $f_n \to f$ or $\text{s-lim } f_n = f$ as $n \to \infty$. A necessary and sufficient condition for strong convergence is that the sequence be Cauchy in the sense defined in Axiom III. The sufficiency is nothing but Axiom III, and the necessity is an easy consequence of the triangle inequality : If $f_n \to f$, then $||f_n - f_m|| = ||f_n - f + f - f_m|| \leq ||f_n - f|| + ||f - f_m|| \to 0$ as $n,m\to\infty$. This inequality can also be applied to verify the uniqueness of the limit vector f (cf. Problem 2.3).

The convergence in H obtained by means of the scalar product is called weak convergence. A sequence $\{f_n\}$ converges weakly to a limit f if for every $g\varepsilon H$ the sequence of scalar products $\{(f_n,g)\}$ converges to (f,g). If this is the case we write $\text{w-lim } f_n = f$ as $n \to \infty$. The Cauchy criterion is also valid for weak convergence, i.e. $\{f_n\}$ converges weakly if and only if for every $g\varepsilon H$ the sequence $\{(f_n,g)\}$ is a Cauchy sequence of complex numbers. If $\{f_n\}$ is a Cauchy sequence in this sense, there exists a unique vector $f\varepsilon H$ such that $\text{w-lim } f_n = f$. We shall not use this criterion, but we may add that the uniqueness of the weak limit is an immediate consequence of Proposition 2.2 below.

Strong convergence implies weak convergence, but the converse is not true. In fact one has the following relation which is often very useful :

PROPOSITION 2.1 : $\text{s-}\lim_{n\to\infty} f_n = f$ if and only if $\text{w-}\lim_{n\to\infty} f_n = f$ and $\lim_{n\to\infty} \|f_n\| = \|f\|$.

Proof : It involves the definition of strong and weak convergence and the inequalities of Schwarz and Minkowski.

(i) Suppose $f_n \to f$. We get with (2.8)

$$|(f_n,g) - (f,g)| = |(f_n - f,g)| \leq \|f_n - f\| \|g\| \to 0$$

for every $g \in H$, i.e. $\text{w-}\lim f_n = f$. By using (2.9) we deduce

$$\|f_n\| \leq \|f_n - f\| + \|f\| \quad \text{and} \quad \|f\| \leq \|f - f_n\| + \|f_n\|,$$

and hence $\|f\| - \|f - f_n\| \leq \|f_n\| \leq \|f\| + \|f - f_n\|$. Since $\|f - f_n\| \to 0$ for $n \to \infty$, it follows that the limit of $\|f_n\|$ as $n \to \infty$ exists and is equal to $\|f\|$.

(ii) $\|f - f_n\|^2 = \|f\|^2 + \|f_n\|^2 - (f,f_n) - (f_n,f)$.

If $\text{w-}\lim f_n = f$ and $\|f_n\| \to \|f\|$, the right-hand side converges to $\|f\|^2 + \|f\|^2 - (f,f) - (f,f) = 0$, i.e. $f_n \to f$. #

An example of a weakly convergent sequence which does not converge strongly is an infinite orthonormal sequence. Before we can verify this statement, we have to introduce the notion of orthogonality. Two vectors f and g are said to be orthogonal to each other if $(f,g) = 0$. Similarly two subsets M_1 and M_2 of H are mutually orthogonal if $(f_1,f_2) = 0$ for all $f_1 \in M_1$ and all $f_2 \in M_2$. An important relation concerning mutually orthogonal vectors is the following :

$$\|\sum_{i=1}^{n} f_i\|^2 = \sum_{i=1}^{n} \|f_i\|^2 \quad \text{if} \quad (f_i,f_j) = 0 \quad \text{for all } i \neq j. \tag{2.10}$$

This is easily verified by writing the left-hand side as a scalar product and using the linearity (2.6) of the scalar product.

An orthonormal sequence of vectors $\{h_i\}$ is characterized by the property that $(h_i,h_j) = \delta_{ij}$, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. If $f \in H$, we then have from (2.6) and (2.10)

$$0 \leq \|f - \sum_{i=1}^{n} (h_i,f)h_i\|^2 = \|f\|^2 + \sum_{i=1}^{n} |(h_i,f)|^2 - 2\sum_{i=1}^{n} |(h_i,f)|^2,$$

hence $\sum_{i=1}^{n} |(h_i,f)|^2 \leq \|f\|^2$. Since this holds for each n, we obtain Bessel's inequality

$$\sum_{i=1}^{\infty} |(h_i,f)|^2 \leq \|f\|^2. \tag{2.11}$$

(2.11) implies that $(h_i,f) \to 0$ as $i \to \infty$ for every $f \in H$, i.e. $\{h_i\}$ converges weakly to zero. But $\{h_i\}$ cannot converge strongly, since $\|h_n - h_m\|^2 = \|h_n\|^2 + \|h_m\|^2 = 2$ if $n \neq m$.

An orthonormal set of vectors $\{e_i\}$ is called an orthonormal basis of H if the set of finite linear combinations of vectors belonging to $\{e_i\}$ is dense in H. In a separable Hilbert space an orthonormal basis is always a countable set. This can be seen as follows : Let μ belong to some index set, let $\{e_\mu\}$ be any orthonormal set in H, i.e. $(e_\mu,e_\nu) = 1$ if $\mu = \nu$ and $(e_\mu,e_\nu) = 0$ if $\mu \neq \nu$. Let $D = \{f_i\}$ be a countable dense set in H. For each μ, there exists $i = i(\mu)$ such that $\|e_\mu - f_{i(\mu)}\| < \frac{1}{2}$. Since $\sqrt{2} = \|e_\mu - e_\nu\| \leq \|e_\mu - f_{i(\mu)}\| + \|e_\nu - f_{i(\mu)}\| < \frac{1}{2} + \|e_\nu - f_{i(\mu)}\|$, we must have $i(\nu) \neq i(\mu)$ if $\nu \neq \mu$. Thus there is a one-to-one correspondence between $\{e_\mu\}$ and a subset of the set of positive integers, i.e. $\{e_\mu\}$ is countable.

The existence of an orthonormal basis cán be established by choosing a subset of linearly independent vectors of a countable dense set D and applying to it the Schmidt orthogonalization process [K, Ch. I.6.3.], [RS, Thm. II. 7]. The dimension of a Hilbert space is equal to the number N of vectors of an orthonormal basis (N does not depend on the choice of a particular basis). Our axioms are equally valid for finite and infinite-dimensional spaces. However if the space if finite-dimensional the last two axioms are a consequence of the others. Furthermore in the finite-dimensional case the strong and the weak topology coincide (Problem 2.4).

An arbitrary vector in H can always be expanded in a given basis. A little care is needed though if H is infinite-dimensional, since only finite linear combinations of vectors are admitted by axiom I. Infinite linear combinations are understood as strong limits of finite linear combinations. Thus the expansion of $f \in H$ in an orthonormal basis $\{e_i\}$ means that the sequence of vectors $\{f_n\}$, where $f_n = \sum_{i=1}^{n} (e_i, f) e_i$, converges strongly to f as $n \to \infty$. (f_n is seen to be a linear combination of the first n vectors of the basis $\{e_i\}$, and the coefficient of e_i is nothing but the "component" (e_i, f) of f along e_i.) The fact that the above is a complete expansion is expressed by the Parseval relation which states that the equality sign holds in (2.11) if $\{h_i\}$ is an orthonormal basis (Problem 2.5) :

$$\|f\|^2 = \sum_{i=1}^{\infty} |(e_i, f)|^2. \tag{2.12}$$

(2.12) says that the square of the length of the vector f is equal to the sum of the squares of the absolute values of its components along the vectors e_i of an arbitrary orthonormal basis.

Another simple result in relation with the concept of orthogonality which will frequently be used is the following :

PROPOSITION 2.2 : Let $\mathcal{D}$ be a dense set in $\mathcal{H}$ and $f \epsilon \mathcal{H}$. If $(f,g) = 0$ for all $g \epsilon \mathcal{D}$, then $f = \theta$.

Proof : Suppose $f \neq \theta$. Given $\eta > 0$, there exists $f_\eta \epsilon \mathcal{D}$ such that $\|f - f_\eta\| < \eta \|f\|^{-1}$. Then, using also (2.8), we have $(f,f) = |(f, f - f_\eta) + (f, f_\eta)| = |(f, f - f_\eta)| \leq \|f\| \, \|f - f_\eta\| < \eta$. Since η is arbitrary, this implies $\|f\|^2 = 0$, hence $f = \theta$ by (2.7). #

A concept that we shall need in the next section is that of a linear manifold. This is a subset $\mathcal{M}$ of $\mathcal{H}$ that satisfies Axiom I but not necessarily Axiom III ($\mathcal{M}$ will always verify Axioms II and IV, since it is a subset of $\mathcal{H}$, cf. Problem 2.6). A subset of $\mathcal{H}$ that satisfies all four axioms will be called a subspace*).

In a finite-dimensional Hilbert space a linear manifold is also a subspace, since Axiom III is not an independent postulate. An example of a linear manifold that is not a subspace in an infinite-dimensional Hilbert space is the set $\mathcal{M}$ of all finite linear combinations of a countably infinite number of linearly independent vectors $\{f_i\}$. It is seen that the sum of two elements of $\mathcal{M}$ and the product of an element of $\mathcal{M}$ and a complex number are again finite linear combinations of the vectors $\{f_i\}$, and the other postulates of Axiom I hold because $\mathcal{M}$ is a linear subset of $\mathcal{H}$.

---

*) In some books the designation "subspace" is used also for a manifold, and a subspace in the sense of our definition is then called a closed subspace.

The closure of a linear manifold (i.e. the manifold obtained by adding to M all the limit points, in the sense of Axiom III, of strong Cauchy sequences of vectors belonging to M) is a subspace of H. In the above example the closure of M is strictly bigger than M, since it contains also certain infinite linear combinations of the vectors $\{f_i\}$. (For a specific example, suppose that the f_i form an orthonormal basis $\{e_i\}$, and let $f = \sum_{k=1}^{\infty} k^{-1} e_k$. Then f belongs to H, since by (2.12) $\|f\|^2 = \sum_{k=1}^{\infty} k^{-2} < \infty$, but $f \notin M$.)

An important example of a closed linear manifold (i.e. of a subspace) is the orthogonal complement $N^{\perp}$ of a subset N of H, i.e. the set of all vectors $f \in H$ such that $(f,g) = 0$ for all $g \in N$. The proof that such a set is a subspace is simple and is left to the reader. In this connection it is also worth noticing the following fact known as the projection theorem ([AG],[K],[RS]) : If M is a subspace and $M^{\perp}$ its orthogonal complement, then every vector f in H has a unique decomposition $f = f_1 + f_2$ with $f_1 \in M$ and $f_2 \in M^{\perp}$. A simple consequence is the following : If M is a linear manifold such that the only vector of H that is orthogonal to M is the vector θ, then M is dense in H.

The case where a linear manifold M is dense in H (i.e. where the closure of M is equal to H) will be of particular importance in the next section for the definition of unbounded linear operators in H. An example of a dense linear manifold is the set of all finite linear combinations of vectors of a basis $\{e_i\}$ of H (Problem 2.5).

As a last general remark we mention a theorem which we shall have occasion to use in later chapters. It concerns bounded linear functionals on a Hilbert space H. By definition such a functional is a linear map Φ from H into C (i.e. $\Phi(f)\varepsilon C$ for each $f\varepsilon H$, and $\Phi(\alpha f + g) = \alpha\Phi(f) + \Phi(g)$ for all $\alpha\varepsilon C$ and $f,g\varepsilon H$) which is bounded with respect to the norm in H, i.e.

$$|||\Phi||| \equiv \sup_{f\neq\theta} \frac{|\Phi(f)|}{||f||} < \infty.$$

If g is a fixed vector in H, one may associate with it a bounded linear functional Φ_g on H by $\Phi_g(f) = (g,f)$. The boundedness of Φ_g follows from (2.8) :

$$|||\Phi_g||| = \sup_{f\neq\theta} \frac{|(g,f)|}{||f||} \leq \sup_{f\neq\theta} \frac{||g||\,||f||}{||f||} = ||g||.$$

In fact one has $|||\Phi_g||| = ||g||$, since equality holds in (2.8) for $f = g$. It is an interesting property of Hilbert space that the converse is also true :

PROPOSITION 2.3 : Let $\Phi : H \to C$ be a bounded linear map. Then there exists a uniquely determined vector $g\varepsilon H$ such that $\Phi(f) = (g,f)$ for all $f\varepsilon H$, and $|||\Phi||| = ||g||$.

This result is known as the Riesz representation theorem. We shall omit its proof, since it is given in numerous books on functional analysis (e.g. [K],[RS]). The idea of the proof is indicated in Problem 2.7.

So far we have considered the abstract Hilbert space. While many of the formal developments of scattering theory can be given in the abstract, the interpretation of the theory in terms of observable quantities usually requires the choice of some concrete realization of the space. The most important such realization for scattering theory is the space $L^2(R^n)$. It consists of all Lebesgue measurable*) complex-valued functions defined on n-dimensional Euclidean space R^n which are absolutely square-integrable, i.e. such that**)

$$\int_{R^n} |f(\underline{x})|^2 \, d^n x < \infty. \tag{2.13}$$

The set of all such functions is a linear vector space if one defines addition and multiplication by scalars as follows :

$$(f_1 + f_2)(\underline{x}) = f_1(\underline{x}) + f_2(\underline{x}), \quad (\alpha f)(\underline{x}) = \alpha f(\underline{x}).$$

The scalar product between two such functions is defined by

$$(f,g) = \int_{R^n} \bar{f}(\underline{x}) g(\underline{x}) d^n x. \tag{2.14}$$

This integral is finite, since $|\bar{f}(\underline{x})g(\underline{x})| \leq \frac{1}{2}|f(\underline{x})|^2 + \frac{1}{2}|g(\underline{x})|^2$.

This scalar product verifies (2.5) and (2.6). Concerning its strict positivity there is a certain complication to which one must pay attention in many of the mathematical developments of the theory. This stems from the fact that

$$\|f\|^2 \equiv \int_{R^n} |f(\underline{x})|^2 \, d^n x = 0 \tag{2.15}$$

does not imply that $f(\underline{x}) = 0$ for all $\underline{x}$. It only implies that $f(\underline{x}) = 0$ almost everywhere (a.e.) with respect to the Lebesgue measure on R^n. This means that $f(\underline{x})$ may differ from zero and may in fact assume arbitrary values on a set of Lebesgue measure zero.

---

*) For measure and integration theory the reader may consult e.g. [L],[R].

**) Vectors in R^n will be denoted by $\underline{x}$.

The Hilbert space $L^2(R^n)$ does therefore not consist of the individual functions themselves but rather of classes of equivalent functions. Two functions are defined to be equivalent if they differ only on a set of measure zero. We may formally introduce the following notation : Let V be the set of individual functions satisfying (2.13) and V_o the subset of V satisfying (2.15). Then the Hilbert space we wish to define is the quotient space $V/V_o = L^2(R^n)$. The element θ of $L^2(R^n)$ is given by the class V_o which contains in particular the function $f \equiv 0$. It is quite elementary to verify that the operations of addition, multiplication by scalars, and scalar product can be transferred to the classes since they are independent of the representative elements inside the classes with which the operations are carried out.

These remarks concerning the quotient space V/V_o may strike a physicist as pedantic, since it is in most cases possible to transfer all operations in the Hilbert space $L^2(R^n)$ to individual functions (a practice which we shall frequently follow in the traditional manner). There are occasional situations where the above remark is essential and must be borne in mind.

The completeness of $L^2(R^n)$ is a classical result of analysis known as the Riesz-Fischer Theorem. We shall not give a proof of it in this book (cf. for instance [RN],[R], [RS] or [L]). The separability of $L^2(R^n)$ can be established in different ways (cf. e.g. [RN],[SO]). One possible method is to show that a general function in $L^2(R^n)$ can be approximated arbitrarily well by a finite linear combination of characteristic functions*) of n-dimensional rectangles whose end points have rational coordinates (the details may be read in [RN, Section 32]). Separability then follows since the set of all such characteristic functions is countable.

The characteristic function of an n-dimensional rectangle can be approximated in $L^2(R^n)$-norm arbitrarily well by an infinitely differentiable function vanishing outside the rectangle; this is done by changing the characteristic functions near the edges of the rectangle into a smooth function (Problem 2.8). It follows that the set $C_o^\infty(R^n)$ **) of all infinitely differentiable functions of compact support is dense in $L^2(R^n)$. This result is often useful in scattering theory.

Apart from the linear manifold $C_o^\infty(R^n)$ we shall also need another linear manifold denoted by $S(R^n)$. A function f belongs to $S(R^n)$ if it is infinitely differentiable and if f and its partial derivatives of all orders decrease faster than any negative power of x as $x \to \infty$ ***). More precisely, $f \in S(R^n)$ if f is infinitely differentiable (i.e. $f \in C^\infty(R^n)$) and if for every 2n-tuple of non-negative integers $\{j_1, \ldots, j_n, m_1, \ldots, m_n\}$ one has

$$\sup_{\underline{x} \in R^n} \left| (x_1)^{j_1} \cdots (x_n)^{j_n} \frac{\partial^{|m_1+\cdots+m_n|}}{\partial x_1^{m_1} \cdots \partial x_n^{m_n}} f(x_1, \ldots, x_n) \right| < \infty. \quad (2.16)$$

---

*) The <u>characteristic function</u> of a set S in R^n is defined by $\chi_S(\underline{x}) = 1$ if $\underline{x} \in S$ and $\chi_S(\underline{x}) = 0$ if $\underline{x} \notin S$.

**) If $S \subseteq R^n$ is an open set, then f is said to belong to $C_o^\infty(S)$ if it has partial derivatives to all orders and there exists a compact set in S outside which f is identically zero.

***) Here we have defined $x = |\underline{x}| \equiv (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$.

Such functions are also called functions of rapid decrease. An example is the function $\exp(-x^2)$. $S(R^n)$ has the following interesting properties :

PROPOSITION 2.4 : (a) $S(R^n)$ is dense in $L^2(R^n)$.
(b) $S(R^n)$ is invariant under Fourier transformation.

Proof : The denseness follows because $S(R^n)$ contains $C_0^\infty(R^n)$ and the latter is dense in $L^2(R^n)$. The Fourier transformation will be defined here for functions belonging to $S(R^n)$ and extended to a larger class of functions in the next section. If $\underline{k}$ and $\underline{x}$ are two vectors in R^n, we define $\underline{k}\cdot\underline{x} \equiv k_1x_1 + k_2x_2 + \cdots + k_nx_n$. If $f\varepsilon S(R^n)$, we may define a new function $\tilde{f} : R^n \to C$ by the formula

$$\tilde{f}(\underline{k}) = (2\pi)^{-n/2} \int d^nx e^{-i\underline{k}\cdot\underline{x}} f(\underline{x}). \qquad (\underline{k}\varepsilon R^n) \qquad (2.17)$$

Proposition 2.4(b) asserts that $\tilde{f}$ belongs again to $S(R^n)$. In order to simplify the notation, we indicate the proof of this for the case $n = 1$ and remark that the proof for $n > 1$ is essentially the same (Problem 2.9).

Let $f\varepsilon S(R^1)$. The proof will rely on the following relation which is an immediate consequence of (2.16) :

$$\left|(1 + x^2) \frac{d^r}{dx^r} x^m f(x)\right| \leq c \text{ for all } x\varepsilon R, \qquad (2.18)$$

where c may depend on r and m. From (2.17) one obtains

$$\frac{d^m}{dk^m} \tilde{f}(k) = (-i)^m (2\pi)^{-\frac{1}{2}} \int dx e^{-ikx} x^m f(x). \qquad (2.19)$$

Here we have interchanged the derivatives and the integral on the right-hand side. This is justified provided that for each m the improper integral in (2.19) converges (as a limit of integrals over an increasing sequence of finite intervals) uniformly in k [L, Vol. I, p. 252]. This is the case since (2.18) implies that

$$\left|\int_{|x|\geq R} dx e^{-ikx} x^m f(x)\right| \leq \int_{|x|\geq R} dx \frac{c}{1+x^2} \leq \int_{|x|\geq R} dx \frac{c}{x^2} = \frac{2c}{R},$$

which converges to zero as $R \to \infty$ uniformly in k. It follows that $\tilde{f}$ is infinitely differentiable. One also obtains from (2.19) by integrating by parts

$$k^r \frac{d^m}{dk^m} \tilde{f}(k) = (-i)^{m+r} (2\pi)^{-\frac{1}{2}} \int dx e^{-ikx} \frac{d^r}{dx^r} (x^m f(x)).$$

Hence

$$\sup_{k\in R} \left|k^r \frac{d^m}{dk^m} \tilde{f}(k)\right| \leq (2\pi)^{-\frac{1}{2}} \int dx \left|\frac{d^r}{dx^r} x^m f(x)\right| \leq (2\pi)^{-\frac{1}{2}} \int dx \frac{c}{1+x^2}.$$

The last integral is seen to be finite. Hence $\tilde{f}$ belongs to $S(R^1)$. #

The result of Proposition 2.4(b) can still be strengthened. In fact the Fourier transformation is a mapping from $S(R^n)$ onto $S(R^n)$. This can be seen by defining first the inverse Fourier transformation on $S(R^n)$ by

$$\hat{f}(\underline{x}) = (2\pi)^{-n/2} \int d^n k e^{i\underline{k}\cdot\underline{x}} f(\underline{k}), \quad f\in S(R^n). \qquad (2.20)$$

The only difference between (2.17) and (2.20) is the sign of the exponent. The same reasoning as in the preceding proof implies that $\hat{f}\in S(R^n)$. If we can show that (2.20) is in fact the inverse transformation of (2.17), i.e. that $\hat{\tilde{f}} = f$, we shall have proved that the Fourier transformation is a bijection of $S(R^n)$ onto itself.

For this, let $f, h \in S(R^n)$ and $\alpha > 0$. By using (2.17) and making appropriate substitutions for the integration variables, one gets

$$\int h(\alpha\underline{k})\tilde{f}(\underline{k})e^{i\underline{k}\cdot\underline{y}}d^nk = (2\pi)^{-n/2}\int d^nk h(\alpha\underline{k})\int d^nx e^{-i\underline{k}\cdot(\underline{x}-\underline{y})}f(\underline{x})$$

$$= \alpha^{-n}\int d^nx \tilde{h}(\alpha^{-1}(\underline{x}-\underline{y}))f(\underline{x}) = \int d^nu \tilde{h}(\underline{u})f(\alpha\underline{u} + \underline{y}). \qquad (2.21)$$

As $\alpha \to 0$, the integrand of the left-hand side converges pointwise to $h(0)\tilde{f}(\underline{k})\exp(i\underline{k}\cdot\underline{y})$. It is also absolutely majorized, uniformly in $\alpha \geq 0$, by the supremum over $\underline{y} \in R$ of $|h(\underline{y})||\tilde{f}(\underline{k})|$. Since the latter function is an integrable function of $\underline{k}$, the Lebesgue dominated convergence theorem (Proposition 2.35) implies that

$$\lim_{\alpha\to 0}\int h(\alpha\underline{k})\tilde{f}(\underline{k})e^{i\underline{k}\cdot\underline{y}}d^nk = h(0)\int\tilde{f}(\underline{k})e^{i\underline{k}\cdot\underline{y}}d^nk = h(0)(2\pi)^{n/2}\hat{\tilde{f}}(\underline{y}).$$

Similarly one gets for the last member of (2.21)

$$\lim_{\alpha\to 0}\int\tilde{h}(\underline{u})f(\alpha\underline{u} + \underline{y})d^nu = f(\underline{y})\int\tilde{h}(\underline{u})d^nu.$$

The desired result $\hat{\tilde{f}}(\underline{y}) = f(\underline{y})$ is obtained by verifying for a particular function $h \in S(R^n)$ that $(2\pi)^{n/2}h(0) = \int\tilde{h}(\underline{u})d^nu$. This is possible for instance for $h_o(\underline{x}) = \exp(-\frac{1}{2}x^2)$. Then (Problem 2.10) $\tilde{h}_o(\underline{k}) = \exp(-\frac{1}{2}k^2) = h_o(\underline{k})$, hence

$$\int\tilde{h}_o(\underline{u})d^nu = \int h_o(\underline{u})d^nu = (2\pi)^{n/2}\tilde{h}_o(0) = (2\pi)^{n/2}h_o(0).$$

Thus we have shown that (2.20) defines the inverse of (2.17). We shall henceforth also denote the mapping $f \mapsto \tilde{f}$ by F, i.e. we shall write $\tilde{f} = Ff$. We then have $f = F^{-1}\tilde{f} = \hat{\tilde{f}}$, where F^{-1} is given by (2.20). In the next section we shall extend F and F^{-1} to operators on the entire space $L^2(R^n)$.

A last important result is the remark that F and F^{-1} are isometric on $S(R^n)$, i.e.

$$\|\tilde{f}\| = \|f\| = \|\hat{f}\| \qquad \text{for } f \in S(R^n), \tag{2.22}$$

$$(\tilde{g},\tilde{f}) = (g,f) = (\hat{g},\hat{f}) \qquad \text{for } f,g \in S(R^n). \tag{2.23}$$

(2.22) is a special case of (2.23). The first equality in (2.23) is a particular case of (2.21) : wet set $\alpha = 1$, $\underset{\sim}{y} = 0$ and $h = \tilde{\bar{g}}$; the left-hand side then equals $(\tilde{g},\tilde{f})$. By noticing that $\tilde{\bar{f}} = \bar{\hat{f}}$, one obtains $\tilde{h} = \hat{\bar{\tilde{g}}} = \bar{g}$, i.e. the right-hand side of (2.21) equals (g,f) for the particular choice of α, $\underset{\sim}{y}$ and h made above. This establishes the first part of (2.23). The proof of its second part is left as an exercise (Problem 2.11).

2-2 LINEAR OPERATORS IN HILBERT SPACE

A linear operator in a Hilbert space H is a linear mapping between vectors of H. As an example we have already seen the Fourier transformation F in $L^2(R^n)$ which was defined on all vectors belonging to the dense set $S(R^n)$ and which is obviously linear. We have seen that this operator does not change the L^2-norm of a vector f. Many operators that are important in physical applications do not have such a simple property, and indeed they may be such that the norm of certain image vectors may exceed that of the corresponding initial vector by an arbitrarily large amount (such operators will be called unbounded). A well-known example from elementary quantum mechanics is the position operator Q for a particle in one-dimensional space. Here $H = L^2(R)$, and Q is the operator of multiplication by the variable x in $L^2(R)$:

$$(Qf)(x) = xf(x).$$

If we take for instance $f(x) = (1 + |x|)^{-1}$, then $f \in L^2(R)$ but $\|Qf\| = \infty$, i.e. $Qf \notin L^2(R)$. Thus one expects that in general a linear operator will be defined only on some subset of the Hilbert space. In the above example this domain of definition D(Q) of the operator Q can be taken as the set of those f in $L^2(R)$ such that xf(x) is also an element of $L^2(R)$. It is easily seen that this subset D(Q) is a manifold, and it is also dense in $L^2(R)$ (it contains all functions belonging to $S(R)$ which is dense in $L^2(R)$).

A linear operator*) is defined by giving its domain, i.e. a linear manifold D(A) in H, and a linear mapping A from D(A) into H. Linearity means that, if $f,g \in D(A)$ and $\alpha \in C$, then $A(\alpha f + g) = \alpha Af + Ag$. If A is a linear operator, then D(A) contains the vector θ, since $\theta = 0 \cdot f$ for arbitrary $f \in D(A)$. One then has $A\theta = 0 \cdot (Af) = \theta$. The following notation will also be used : If M is a subset of D(A), then AM is the set of vectors f in H such that f = Ag for some g in M. The set AD(A) will be called the range of the operator A.

Two linear operators A and B are equal if and only if D(A) = D(B) and Af = Bf for all $f \in D(A)$. A linear operator A' is called an extension of A if $D(A) \subseteq D(A')$ and A'f = Af for

*) A linear operator will usually be simply called an operator.

all $f \in D(A)$. Here A' coincides with A on D(A), but it may be defined on a larger domain than A. In this case we shall write $A \subseteq A'$. One may also call A the restriction of A' to D(A).

For some operators A there is a natural way of defining an extension $\bar{A}$. One takes a strong Cauchy sequence $\{f_n\}$ in D(A). If the sequence $\{Af_n\}$ is also Cauchy, and if one denotes by f and g the strong limits of $\{f_n\}$ and $\{Af_n\}$ respectively, it is natural to define $\bar{A}f = g$. Since f is not necessarily in D(A), one may define an extension $\bar{A}$ of A by applying the above definition to all Cauchy sequences $\{f_n\}$ in D(A) which are such that $\{Af_n\}$ is also Cauchy. However, this construction makes sense only if the element g is independent of the choice of a particular Cauchy sequence $\{f_n\}$ converging to f, i.e. if the following condition is satisfied : Whenever $\{f_n\}$ and $\{f'_n\}$ are two Cauchy sequences in D(A) converging strongly to the same limit f and $\{Af_n\}$ and $\{Af'_n\}$ are also Cauchy, then $\text{s-lim}\, Af_n = \text{s-lim}\, Af'_n$. Since A is linear, this condition is easily seen to be equivalent to the following one (Problem 2.12) : Whenever $\{f_n\} \in D(A)$, $f_n \to \theta$ and $\{Af_n\}$ is strongly Cauchy, then $Af_n \to \theta$.

An operator verifying one of these two equivalent conditions is said to be closable, and then the above extension $\bar{A}$ is called the closure of A. A is said to be closed if it is identical with its closure, i.e. if $A = \bar{A}$.

The following is a different criterion for an operator to be closed :

LEMMA 2.5 : Let A be a linear operator in $\mathcal{H}$. The following two statements are equivalent :

(a) A is closed.

(b) Whenever a sequence $\{f_n\}$ verifies (i) $f_n \varepsilon D(A)$, (ii) $f_n \to f$, and (iii) $Af_n \to g$, then $f \varepsilon D(A)$ and $Af = g$.

Proof : The result is a simple consequence of the definition of the closure.

(i) Suppose $A = \bar{A}$, and let $\{f_n\}$ verify (i)-(iii). Since A is closable, one has $f \varepsilon D(\bar{A})$ and $\bar{A}f = g$. Since $A = \bar{A}$, this means that (b) is verified. Hence (a) implies (b).

(ii) Suppose (b) holds. Then A is closable, since the hypotheses $\{f_n\} \varepsilon D(A)$, $f_n \to \theta$ and $Af_n \to g$ then imply $g = A\theta = \theta$. From the construction of $\bar{A}$ and (b) one sees that $D(\bar{A}) \subseteq D(A)$, i.e. $\bar{A} = A$. #

The operators that one encounters in applications are usually closed or closable. An operator A may have more than one closed extension (an example of such an operator will be given further on in this section). If an operator A has a closed extension B, then it is closable. To see this, suppose $\{f_n\} \varepsilon D(A)$, $f_n \to f$ and $Af_n \to g$. Since $A \subseteq B$ and B is closed, Lemma 2.5 implies that $f \varepsilon D(B)$ and $g = Bf$. Since the vector Bf is independent of the sequence $\{f_n\}$ converging to f, this shows that A is closable.

If A is closable, then its closure $\bar{A}$ is its smallest closed extension, i.e. if A' is an arbitrary closed extension of A, then $\bar{A} \subseteq A'$ (Problem 2.13). An example of a non-closable operator is given in Problem 2.14.

We shall make additional comments on closed extensions further on in this section. Here we shall indicate an important class of closable operators, namely the bounded operators. A linear operator A is said to be bounded if there exists a number $M < \infty$ such that $||Af|| \leq M\,||f||$ for all $f \varepsilon D(A)$. If there exists no such M, A is said to be unbounded. For bounded A one defines its norm $||A||$ as*)

$$||A|| = \sup_{\substack{f \varepsilon D(A) \\ f \neq \theta}} \frac{||Af||}{||f||} . \tag{2.24}$$

One then has for every $f \varepsilon D(A)$:

$$||Af|| \leq ||A||\,||f||. \tag{2.25}$$

As a consequence of this inequality, one has the following result : Let A be bounded and $\{f_n\}$ a strong Cauchy sequence in D(A). Then $\{Af_n\}$ is also strongly Cauchy. In fact

$$||Af_n - Af_m|| \leq ||A||\,||f_n - f_m|| \to 0 \text{ as } n,\, m \to \infty.$$

This can be used to prove that a bounded operator is always closable. This result is contained in the following proposition which will prove to be useful in scattering theory.

PROPOSITION 2.6 : If A is a bounded linear operator on a Hilbert space H, it has a unique bounded extension $\bar{A}$ to the subspace spanned by D(A) (i.e. to the closure $\overline{D(A)}$ of D(A)). $\bar{A}$ is closed, and $||\bar{A}|| = ||A||$. In particular, if D(A) is dense in H, then $D(\bar{A}) = H$.

Proof : We verify that the closure $\bar{A}$ of A has the required

---

*) It is customary to use the same symbol $||\cdot||$ for the norm of an operator and for the norm of vectors in H. It is seen from (2.24) that the former is defined in terms of the latter.

properties.

(i) Let $\{f_n\}$ and $\{f'_n\}$ be two strong Cauchy sequences in $D(A)$ converging to the same limit f. Then $\{Af_n\}$ and $\{Af'_n\}$ are also Cauchy, and by (2.25) and (2.9)

$$\|Af_n - Af'_n\| \leq \|A\| \|f_n - f'_n\| \leq \|A\| (\|f_n - f\| + \|f - f'_n\|),$$

which converges to zero as $n \to \infty$. Hence $\{Af'_n\}$ converges to the same vector as $\{Af_n\}$, i.e. A is closable. It also follows that every f in $\overline{D(A)}$ belongs to the domain of $\bar{A}$. Hence $D(\bar{A}) = \overline{D(A)}$.

(ii) Suppose A' is another bounded extension of A defined on $D(A') = \overline{D(A)}$. As in (i), A' is closable and $D(\overline{A'}) = \overline{D(A')}$. Since $D(A')$ is closed, this implies $D(\overline{A'}) = D(A')$, i.e. A' is a closed operator. It follows that A' is an extension of $\bar{A}$. But $D(A') = D(\bar{A})$, so that $A' = \bar{A}$. This proves the uniqueness of the extension.

(iii) We have from (2.24) and the fact that $A \subseteq \bar{A}$

$$\|\bar{A}\| = \sup_{f\in D(\bar{A}), f\neq\theta} \frac{\|\bar{A}f\|}{\|f\|} \geq \sup_{f\in D(A), f\neq\theta} \frac{\|Af\|}{\|f\|} = \|A\|.$$

The opposite inequality, i.e. $\|\bar{A}\| \leq \|A\|$, is obtained by letting f and $\{f_n\}$ be as in (i) and by applying twice Proposition 2.1 :

$$\|\bar{A}f\| = \lim_{n\to\infty} \|Af_n\| \leq \|A\| \lim_{n\to\infty} \|f_n\| = \|A\| \|f\|. \#$$

As a consequence of Proposition 2.6, one may always consider a bounded operator to be defined on a subspace. If $D(A)$ in Proposition 2.6 is not dense in $\mathcal{H}$, this subspace is strictly smaller than $\mathcal{H}$. One may then extend A to a bounded operator $\tilde{A}$ defined on all of $\mathcal{H}$ by setting $\tilde{A} = \bar{A}$ on $\overline{D(A)}$ and by identifying $\tilde{A}$ with an arbitrary bounded operator B on the orthogonal complement $D(A)^{\perp}$ of $D(A)$. ($\tilde{A}$ is then defined everywhere : An arbitrary f in $\mathcal{H}$ can be decomposed uniquely into a sum of an element f_1 in $\overline{D(A)}$ and of an element f_2 in $D(A)^{\perp}$, and by linearity one has $\tilde{A}f = \bar{A}f_1 + Bf_2$.) The simplest possibility is to take $B = 0$ on $D(A)^{\perp}$, which we shall do for some operators used in scattering theory.

As an example where Proposition 2.6 can be applied we mention the Fourier transformation $\mathcal{F}$ defined in Section 2-1. We there had $D(\mathcal{F}) = \mathcal{S}(\mathcal{R}^n)$, and $\|\mathcal{F}\| = 1$ from (2.22). Hence $\mathcal{F}$ may be extended to a bounded operator of norm 1 defined on all of $L^2(\mathcal{R}^n)$. We shall henceforth denote this extension also by $\mathcal{F}$. This extension is still an isometric operator, i.e. $\|\mathcal{F}f\| = \|f\|$ for all $f \in L^2(\mathcal{R}^n)$ (Problem 2.15). Also it is still given by (2.17) for vectors f in $L^2(\mathcal{R}^n)$ which are also in $L^1(\mathcal{R}^n)$, i.e. which satisfy in addition to (2.13) the condition

$$\|f\|_1 \equiv \int_{\mathcal{R}^n} |f(\underline{x})| d^n x < \infty.$$

Indeed, for such a vector f the integral in (2.17) exists for each $k \in \mathcal{R}^n$, and one can show that the vector f defined in this way is identical with $\mathcal{F}f$ (the reader may look up details in Section 2-5). For a general f in $L^2(\mathcal{R}^n)$ the integral in (2.17) need not make sense ($f \in L^2(\mathcal{R}^n)$ does not imply that f is integrable); then one has to define $\mathcal{F}f$ as $\mathcal{F}f = \text{s-lim}\, \mathcal{F}f_m$ for $m \to \infty$, where $f_m \in L^2(\mathcal{R}^n) \cap L^1(\mathcal{R}^n)$ and $f_m \to f$. A particular way of choosing f_m is to set $f_m(\underline{x}) = f(\underline{x})$ for $x \leq m$ and $f_m(\underline{x}) = 0$ for $x > m$. The strong convergence as $m \to \infty$ of this particular sequence $\{\mathcal{F}f_m\}$ to $\mathcal{F}f \equiv \tilde{f}$ is also called <u>convergence in the mean</u>, and one uses the following notation for this method of defining the Fourier transformation in $L^2(\mathcal{R}^n)$:

$$\tilde{f}(\underline{k}) = (2\pi)^{-n/2}\, \ell.\text{i.m.} \int d^n x \exp(-i\underline{k}\cdot\underline{x}) f(\underline{x}) \tag{2.26}$$

(some other details about this are given in Section 2-5).

Similarly F^{-1} can be extended to $L^2(R^n)$, and again we shall denote this extension by F^{-1}. It is isometric and given by (2.20) for $f \in L^2(R^n) \cap L^1(R^n)$. Also the extended operator F^{-1} is still the inverse of the extended operator F : FF^{-1} and $F^{-1}F$ are bounded extensions of the identity operator I_o defined on $D(I_o) = S(R^n)$; by Proposition 2.6 the only bounded extension to all of $L^2(R^n)$ of I_o is the identity operator I given by If = f for all f in H. Hence $FF^{-1} = I$ and $F^{-1}F = I$.

The sum of two operators A and B is defined in a natural way as follows : $D(A + B) = D(A) \cap D(B)$, and $(A + B)f = Af + Bf$ for $f \in D(A + B)$. In particular, if $D(B) = H$, then $D(A + B) = D(A)$. From the triangle inequality (2.9) one easily deduces that for bounded operators (Problem 2.16)

$$||A + B|| \leq ||A|| + ||B||. \tag{2.27}$$

One may similarly define the product of A and B : $f \in D(AB)$ if and only if $f \in D(B)$ and $Bf \in D(A)$, and $(AB)f = A(Bf)$ for $f \in D(AB)$. (2.25) implies that for bounded operators

$$||AB|| \leq ||A||\,||B||. \tag{2.28}$$

The operator multiplication is not commutative, i.e. AB may be different from BA (a well-known example is that of the operators $P = -id/dx$ and Q = multiplication by x in $L^2(R)$; another example is the unilateral shift operator and its adjoint which will be discussed later in this section).

It should be remarked that, even if both A and B are densely defined, D(A + B) or D(AB) may consist only of the vector θ. For this reason one has to be very cautious when adding or multiplying unbounded operators. We shall discuss in later chapters some of the difficulties that arise in scattering theory due to the unboundedness of some important operators. When both A and B are bounded and defined everywhere, it is not necessary to worry about domain problems.

A linear operator A is invertible if Af = Ag, f and g in D(A), implies f = g, or equivalently if Af = θ, f$\in$D(A), implies f = θ. The inverse A^{-1} is then well defined and given as follows : $D(A^{-1}) = AD(A)$ (i.e. the range of A) and $A^{-1}(Af) = f$. It is easy to see that A^{-1} is also linear. Also $A^{-1}A$ is then the restriction of the identity operator I to D(A) and AA^{-1} is the restriction of I to $D(A^{-1})$. If A is bounded and D(A) = H, then $A^{-1}A = I$. The following result is a direct consequence of the definitions : If A is closed and invertible, then A^{-1} is also closed (Problem 2.17).

We shall now introduce the concept of the adjoint operator A* of a linear operator A. For this we assume that D(A) is dense in H. We first define the domain D(A*) : a vector g$\in H$ belongs to D(A*) if there exists a vector $g^*\in H$ such that

$$(g,Af) = (g^*,f) \qquad \text{for all } f\in D(A). \tag{2.29}$$

The mapping A* is then defined as $A^*g = g^*$. Thus (2.29) may be rewritten as

$$(g,Af) = (A^*g,f) \text{ for all } f\in D(A) \text{ and all } g\in D(A^*). \tag{2.30}$$

A* is well defined, i.e. the vector g* in (2.29) is unique.

In fact, if g_1^* has the same property as g^*, then $(g_1^* - g^*, f) = 0$ for all $f \varepsilon D(A)$, and since $D(A)$ is dense, Proposition 2.2 implies that $g_1^* = g^*$. Clearly A^* is also linear.

The adjoint of a linear operator A is always a closed operator. In fact, if $g_n \varepsilon D(A^*)$, $g_n \to g$ and $A^* g_n \to h$, then for any f in D(A)

$$(g\,, Af) = \lim_{n\to\infty} (g_n\,, Af) = \lim_{n\to\infty} (A^* g_n\,, f) = (h\,, f),$$

which shows that $g \varepsilon D(A^*)$ and that $A^* g = h$.

Another simple property of the adjoint is the following (Problem 2.18) : If A is closable and D(A) dense, then

$$A^* = (\bar{A})^* \equiv \bar{A}^*. \tag{2.31}$$

If $D(A^*)$ is also dense in $\mathcal{H}$, then $A^{**} \equiv (A^*)^*$ exists. One then has the following result :

PROPOSITION 2.7 : Let A be a linear operator such that D(A) and $D(A^*)$ are dense in $\mathcal{H}$. Then A is closable and $\bar{A} = A^{**}$.

We shall not give a complete proof of this proposition here. Notice that it follows immediately from the definition (2.30) that A^{**} is an extension of A (Problem 2.19). Since A^{**} is the adjoint of an operator, it is closed. Hence A has a closed extension and is therefore closable.

The proof that A^{**} coincides with the closure of A requires the notion of the graph of A. This method is indicated in Section 2-5. The converse of Proposition 2.7 is also true : if A is closable and D(A) dense, then $D(A^*)$ is also dense. We shall not use this result and therefore omit its proof (cf. e.g. [RS],[RN]).

In the next proposition we specify the properties of the adjoint of a bounded operator.

PROPOSITION 2.8 : Let A be a bounded operator with $D(A) = H$. Then A^* is bounded, $D(A^*) = H$ and $\|A^*\| = \|A\|$. In addition $A^{**} = A$.

Proof : Let g be a fixed vector in H. From (2.8) and (2.25) it follows that $|(g\,,Af)| \leq \|A\|\,\|f\|\,\|g\|$. Hence the correspondence $f \mapsto (g\,,Af)$ defines a bounded linear functional Φ on H with $|||\Phi||| \leq \|A\|\,\|g\|$. It then follows from Proposition 2.3 that there exists $g^* \varepsilon H$ such that (2.29) holds, and that $\|g^*\| \leq \|A\|\,\|g\|$.

Since g was arbitrary, we have shown that $D(A^*) = H$, and that $\|A^*\| \leq \|A\|$. To prove the converse inequality, we apply the same reasoning to A^* in the place of A to deduce that $\|A^{**}\| \leq \|A^*\|$. Since A^{**} is an extension of A and $D(A) = H$, one has $A^{**} = A$. Hence the last inequality becomes $\|A\| \leq \|A^*\|$, which proves that $\|A\| = \|A^*\|$. #

If A is bounded with $D(A) = H$ and D(B) is dense, it follows from (2.30) that

$$(\alpha A)^* = \bar{\alpha}A^*, \quad (A + B)^* = A^* + B^* \text{ and } (AB)^* = B^*A^*. \tag{2.32}$$

If both A and B are unbounded, the last two equalities need not hold, since the domains of the respective left-hand and right-hand member may be different. As a good exercise for becoming familiar with the notion of the adjoint the reader may show that in the general case (Problem 2.20) :

$$B^*A^* \subseteq (AB)^*. \tag{2.33}$$

The following is also easy to verify :

$$\text{if } A \subseteq B, \qquad \text{then } B^* \subseteq A^*. \tag{2.34}$$

We shall now discuss some special types of operators that we shall encounter throughout this book. We begin with orthogonal projections*), denoted here by F. They are defined by the requirements

$$D(F) = \mathcal{H} \text{ and } F^2 = F = F^*. \tag{2.35}$$

Their most interesting property is that the set of all orthogonal projections is in one-to-one correspondence with the family of all subspaces of $\mathcal{H}$.

To prove the above assertion, let F be a projection and define $\mathcal{M}_F = F\mathcal{H}$. Thus, if $f \in \mathcal{M}_F$, there exists $g \in \mathcal{H}$ such that $f = Fg$. Hence $Ff = F^2 g = Fg = f$. On the other hand, if f is such that $Ff = f$, then obviously $f \in \mathcal{M}_F$. Thus $\mathcal{M}_F = \{f \in \mathcal{H} | Ff = f\}$. $\mathcal{M}_F$ is clearly a linear manifold. To show that it is a subspace, i.e. that it is strongly closed, let $\{f_n\} \in \mathcal{M}_F$ be a strong Cauchy sequence, $f_n \to f$. Let $g \in \mathcal{H}$. Then $(Ff - f,g) = (f,F^*g) - (f,g) = \lim[(f_n,F^*g) - (f_n,g)] = \lim[(Ff_n,g) - (f_n,g)] = (\theta,g) = 0$ as $n \to \infty$. It follows from Proposition 2.2 that $Ff - f = \theta$. Thus $f \in \mathcal{M}_F$, which proves that $\mathcal{M}_F$ is strongly closed.

If $h \in \mathcal{M}_F^\perp$ and $g \in \mathcal{H}$, we have from (2.30) $(Fh,g) = (F^*h,g) = (h,Fg) = 0$. Thus, by Proposition 2.2, $Fh = \theta$. This shows

*) We shall not use non-orthogonal projections in this book. Therefore an orthogonal projection will simply be called a projection.

that F is the orthogonal projection onto M_F. Conversely, given a subspace M, one may define a linear operator F as follows : Let $\{e_i\}$ be an orthonormal basis of M and

$$Fg = \sum_i (e_i,g)e_i. \tag{2.36}$$

It is easy to verify that this operator is an orthogonal projection (Problem 2.21) with range M.

A particular case is that of a one-dimensional subspace $M = \{\alpha f | \alpha \varepsilon C\}$, where f is a fixed unit vector*) in H. In this case we shall denote by F_f the corresponding orthogonal projection defined by (2.36). The action of F_f may be written as

$$F_f g = (f,g)f.$$

We mention two other simple properties of projections :

LEMMA 2.9 : Let F be an orthogonal projection. Then

(a) $\|F\| = 1$ unless $F = 0$, (2.37)

(b) $(f,Fg) = (Ff,Fg)$ for all f,g in H. (2.38)

Proof : Let $f \varepsilon H$. One may write $f = Ff + (I - F)f$. Here $Ff \varepsilon M_F$ and $(I - F)f \varepsilon M_F^{\perp}$. Thus one has with (2.10) $\|f\|^2 = \|Ff\|^2 + \|(I - F)f\|^2 \geq \|Ff\|^2$, which shows that $\|F\| \leq 1$. The fact that $\|F\| = 1$ follows by taking $\theta \neq f \varepsilon M_F$, which is possible unless F is the zero operator on H.

To prove (b), one uses (2.35) :

$$(f,Fg) = (f,F^2 g) = (f,F^*Fg) = (Ff,Fg). \#$$

---

*) A unit vector is a vector f verifying $\|f\| = 1$.

For an example, let $H = L^2(R^n)$ and let Δ be a measurable subset of R^n. Then $L^2(\Delta)$ is defined to be the set of those equivalence classes of functions in $L^2(R^n)$ which have support in Δ, i.e. whose representatives are zero almost everywhere (with respect to Lebesgue measure) on the complement of Δ. It is easy to see that $L^2(\Delta)$ is a subspace of $L^2(R^n)$, i.e. $L^2(\Delta)$ is itself a Hilbert space. The orthogonal projection F_Δ of $L^2(R^n)$ onto $L^2(\Delta)$ may be written as follows

$$(F_\Delta f)(\underline{x}) \equiv (\chi_\Delta f)(\underline{x}) = \chi_\Delta(\underline{x})f(\underline{x}), \tag{2.39}$$

where χ_Δ is the characteristic function of the set Δ. Clearly $D(F_\Delta) = H$ and $F_\Delta^2 = F_\Delta$. The fact that $F_\Delta^* = F_\Delta$ is obtained by writing (2.30) in the form of integrals.

Next we consider partial isometries, denoted here by Ω. These are important in scattering theory, as we shall see in Chapter 4. They are defined by the requirements

$$D(\Omega) = H \text{ and } \Omega^*\Omega = E,\ E \text{ a projection.} \tag{2.40}$$

For such an operator one has with (2.30) and (2.38)

$$(\Omega f,\Omega g) = (f,\Omega^*\Omega g) = (f,Eg) = (Ef,Eg). \tag{2.41}$$

Hence, if $f,g \in EH$, then $(\Omega f,\Omega g) = (f,g)$. This shows that such an operator is isometric on a part of H, namely on EH, i.e. it preserves the length of vectors of EH and the angles between vectors of EH. This explains the designation "partial isometry". It also follows from (2.41) that Ω is zero on the orthogonal complement of EH : if $f \in (EH)^\perp$, then $\|\Omega f\|^2 = \|Ef\|^2 = 0$, whence $\Omega f = \theta$. This may also be written as $\Omega(I - E)f = \theta$ for all $f \in H$, i.e. $\Omega = \Omega E$.

A partial isometry may also be defined by the property (2.41) :

PROPOSITION 2.10 : If Ω is a linear operator with $D(\Omega) = H$ and E a projection such that $||\Omega f|| = ||Ef||$ for all $f \varepsilon H$, then $\Omega^*\Omega = E$.

Proof : Since $||\Omega f|| = ||Ef||$ for all $f\varepsilon H$, one has $||\Omega|| = ||E||$, i.e. $||\Omega|| = 1$ unless $E = 0$. By Proposition 2.8 $||\Omega^*|| = ||E||$, and Ω^* is defined everywhere. Now one obtains from the polarization identity (Problem 2.22) and (2.38) that for any $f,g\varepsilon H$

$$(\Omega^*\Omega f,g) = (\Omega f,\Omega g) =$$

$$\tfrac{1}{4}\{||\Omega(f+g)||^2 - ||\Omega(f-g)||^2 - i||\Omega(f+ig)||^2 + i||\Omega(f-ig)||^2\} =$$

$$\tfrac{1}{4}\{||E(f+g)||^2 - ||E(f-g)||^2 - i||E(f+ig)||^2 + i||E(f-ig)||^2 =$$

$$(Ef,Eg) = (Ef,g).$$

Hence $\Omega^*\Omega f - Ef$ is orthogonal to every $g\varepsilon H$, i.e. $\Omega^*\Omega f = Ef$ by Proposition 2.2. #

A particular case of a partial isometry is obtained by setting $E = I$. Ω is then isometric on all of H and is called an isometry.

The adjoint Ω^* of a partial isometry is also a partial isometry. In fact $F \equiv \Omega^{**}\Omega^* = \Omega\Omega^*$ is a projection, since $F^2 = \Omega\Omega^*\Omega\Omega^* = \Omega E\Omega^* = \Omega\Omega^* = F$ and $F^* = (\Omega\Omega^*)^* = \Omega^{**}\Omega^* = F$. Some additional properties of partial isometries that are essential for later chapters are collected in the following proposition.

PROPOSITION 2.11 : Let Ω be a partial isometry, $\Omega^*\Omega = E$, and define $F = \Omega\Omega^*$. Then

(a) $\|\Omega\| = \|\Omega^*\| = 1$ unless $E = 0$.

(b) $\Omega E = \Omega$, $E\Omega^* = \Omega^*$, (2.42)

$F\Omega = \Omega$, $\Omega^* F = \Omega^*$. (2.43)

(c) The range of Ω is a subspace, and F is the orthogonal projection onto this subspace.

(d) The restriction of Ω to the subspace EH is invertible, and its inverse is given by Ω^* (more precisely by the restriction of Ω^* to FH).

Proof : (a) has been shown in the proof of Proposition 2.10. (b) $\Omega E = \Omega$ has already been proved. By using (2.35) and (2.32), one obtains from this $E\Omega^* = E^*\Omega^* = (\Omega E)^* = \Omega^*$. This proves (2.42). Next we have $F\Omega = \Omega\Omega^*\Omega = \Omega E = \Omega$. Also, from (2.42) : $\Omega^* F = \Omega^*\Omega\Omega^* = E\Omega^* = \Omega^*$, which proves (2.43).
(c) The definition $F = \Omega\Omega^*$ shows that the range of F is contained in that of Ω : If $f \in H$ and $f = Fg$, then $f = \Omega(\Omega^* g) \in \Omega H$. Thus $FH \subseteq \Omega H$. Similarly it follows from $F\Omega = \Omega$ that $\Omega H \subseteq FH$. Thus $\Omega H = FH$, and (c) is proved.
(d) Suppose $f \in EH$ and $\Omega f = \theta$. From (2.41) one then has $\|f\|^2 = \|\Omega f\|^2 = 0$, i.e. $f = \theta$. Thus the restriction of Ω to EH is invertible. Since $\Omega^*\Omega f = f$ for f in EH, it is clear that its inverse is given by the restriction of Ω^* to FH. #

For a general partial isometry Ω, the projections E onto the so-called initial set EH of Ω and F onto the final set or range FH of Ω are both different from the identity operator. For an isometry, one has $E = I$, but F may still be different from I. In a finite-dimensional Hilbert space, the range of an isometry is the entire space (i.e. $F = I$), since it is a subspace that has the same dimension as the entire space. In an infinite-dimensional Hilbert space an isometry Ω may map the entire space onto a proper subspace of infinite dimension. An example of such an operator is the unilateral shift operator : Let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal basis of H, define $\Omega e_i = e_{i+1}$ and extend this definition to finite linear combinations of the vectors $\{e_i\}$ by linearity. It follows from (2.10) that, if $f = \sum_{i=1}^{N}\alpha_i e_i$, then $\|\Omega f\|^2 = \|f\|^2$. By Problem 2.15 and Proposition 2.10, the closure of the above operator defines an isometric operator Ω with $D(\Omega) = H$. One sees that ΩH is the proper subspace spanned by $\{e_2, e_3, \ldots\}$, and e_1 is orthogonal to ΩH. Thus $F = I - F_{e_1} \neq I$, where F_{e_1} is the orthogonal projection onto the subspace determined by e_1. (One may also explicitly calculate Ω^* : $\Omega^* e_1 = \theta$ since $e_1 \varepsilon (\Omega H)^{\perp}$, and $\Omega^* e_i = e_{i-1}$ if $i \geq 2$ by Proposition 2.11(d).)

An isometry for which F is also the identity operator is called a unitary operator. Thus U is unitary if

$$D(U) = H \text{ and } U^*U = UU^* = I. \quad (2.44)$$

PROPOSITION 2.12 : Let U be unitary. Then

(a) The range of U is equal to H.

(b) U is invertible and $U^{-1} = U^*$. (2.45)

(c) $\|Uf - g\| = \|f - U^*g\|$ for all $f, g \varepsilon H$. (2.46)

Proof : (a) and (b) follow immediately from Proposition 2.11. (c) is easily verified by writing the norms as scalar products and using (2.44). #

We next consider symmetric and self-adjoint operators. A is called symmetric if D(A) is dense in H and $A \subseteq A^*$, i.e. if

$$(Af,g) = (g,Af) \text{ for all } f,g \in D(A). \tag{2.47}$$

A special case is that of a self-adjoint operator which is characterized by $A = A^*$. The condition $A = A^*$ is a severe restriction for the case of unbounded operators, since in addition to (2.47) it requires the domain of A^* to be exactly the same as that of A. For bounded operators this difficulty does not arise, since every bounded symmetric operator with $D(A) = H$ is self-adjoint (this follows immediately from the definitions). If A is bounded and symmetric and D(A) is only dense in H, then the closure $\bar{A}$ of A is self-adjoint. In fact one can easily check that in this case $\bar{A} = A^*$, which then implies together with (2.31) that $\bar{A} = \bar{A}^*$.

Since A^* is closed, the requirement $A \subseteq A^*$ implies that A has a closed extension. This means that a symmetric operator A is always closable. In addition $D(A^*)$ is dense, so that by Proposition 2.7 $\bar{A} = A^{**}$. By applying (2.34) to $A \subseteq A^*$ and using (2.31), one obtains $A^{**} \subseteq A^* = \bar{A}^*$. It follows that the closure $\bar{A}$ of a symmetric operator A is also symmetric. A self-adjoint operator $(A = A^*)$ is always closed. If one extends a symmetric operator A to a larger domain, the domain of the adjoint of the extension A' will be smaller than (or possibly equal to) $D(A^*)$, since a vector belonging to $D(A'^*)$ has to satisfy more conditions than a vector belonging to $D(A^*)$, as can be seen from (2.30). Thus one will have $A \subseteq A'$ and $A'^* \subseteq A^*$. Since $A \subseteq A^*$, it may happen that for certain extensions A' one will have $A' = A'^*$, i.e. that A' will be self-adjoint. For this reason the study of closed extensions of a symmetric operator is an important problem in functional analysis. We shall return to it in Chapter 8.

We wish to point out here that the number of self-adjoint extensions of a symmetric operator may be zero, finite, countably or even uncountably infinite.

The preceding remark may be illustrated by means of differential operators. We wish to associate various symmetric operators with the formal operator id/dx acting on functions defined on various subsets of $\mathcal{R}$. Let $\mathcal{D}$ be the set of all functions in $L^2(\mathcal{R})$ which are absolutely continuous on each finite interval [a,b] and such that $f \varepsilon L^2(\mathcal{R})$. Now consider first the Hilbert space $L^2(0,1)$. We define A_o to be the operator $(A_o f)(x) = if'(x)$, where $f \varepsilon D(A_o)$ if $f \varepsilon L^2(0,1) \cap \mathcal{D}$ and $f(0) = f(1) = 0$. By integrating by parts one may verify that A_o is symmetric; it is also closed, and $D(A_o^*) = L^2(0,1) \cap \mathcal{D}$ (no boundary condition is involved !). Hence $D(A_o^*)$ is strictly larger than $D(A_o)$. This operator A_o has an uncountable number of self-adjoint extensions, which are obtained by replacing the boundary condition $f(0) = f(1) = 0$ by the less restrictive condition $f(0) = \exp(i\phi)f(1)$ with $\phi \varepsilon [0,2\pi)$. Every ϕ determines a different self-adjoint extension of A_o [RN, no. 119],[AG, no. 49].

Similarly one may associate with the formal operator id/dx a symmetric operator A_1 in $L^2(0,\infty)$: f belongs to $D(A_1)$ if $f \varepsilon L^2(0,\infty) \cap \mathcal{D}$ and $f(0) = 0$. This operator has no self-adjoint extension at all [AG, no. 49]. Thirdly one may define A_2 in $L^2(\mathcal{R})$ by $(A_2 f)(x) = if'(x)$ and $D(A_2) = \mathcal{D}$. This symmetric operator is self-adjoint, i.e. it has exactly one self-adjoint extension.

Clearly the restriction of a self-adjoint operator B to a dense subset of its domain is a symmetric operator A. If the dense subset is sufficiently large, B will just be the closure of this restriction. In that case A is said to be essentially self-adjoint. Generally a symmetric operator A is said to be essentially self-adjoint if $\bar{A}$ is self-adjoint. An equivalent definition of essential self-adjointness is easily seen to be $A^* = A^{**}$ (Problem 2.23). An essentially self-adjoint operator has one and only one self-adjoint extension. In fact, assume A' to be a self-adjoint extension of an essentially self-adjoint operator A. Then $A' = A'^* \subseteq A^* = A^{**} = \bar{A}$. Since $\bar{A}$ is the smallest closed extension of A, A' must be equal to the closure $\bar{A}$ of A.

The notion of essential self-adjointness is important since in applications one is often given a non-closed symmetric operator. If such an operator can then be shown to be essentially self-adjoint, it follows that it determines a unique self-adjoint operator.

It is useful to have a criterion for a symmetric operator to be self-adjoint or essentially self-adjoint. Such a criterion can be formulated in terms of the ranges $(A \pm i)D(A)$ of the operators $(A \pm i)$ *). We first prove an auxiliary result :

LEMMA 2.13 : Let A be a symmetric operator, $\bar{A}$ its closure. Then the ranges of $\bar{A} \pm i$ are the closures of the ranges of $A \pm i$ respectively.

---

*) We sometimes use the notation $A + \alpha$ for the operator $A + \alpha I$ $(\alpha \in \mathbb{C})$.

Proof : (i) If $f \in D(\bar{A})$, there exists a sequence $\{f_n\} \in D(A)$ with $f_n \to f$ and $Af_n \to \bar{A}f$. Thus $(A \pm i)f_n \to (\bar{A} \pm i)f$, which shows that $(\bar{A} \pm i)D(\bar{A})$ is contained in the closure of the range of $(A \pm i)$ respectively.

(ii) Since A is symmetric, we have for $f \in D(A)$

$$\|(A \pm i)f\|^2 = \|Af\|^2 + \|f\|^2 \pm i(Af,f) \mp i(f,Af)$$

$$= \|Af\|^2 + \|f\|^2. \tag{2.48}$$

Now suppose for instance that g belongs to the closure of $(A + i)D(A)$. Then there exists a sequence $\{f_n\} \in D(A)$ such that $(A + i)f_n \to g$. It follows from (2.48) that both $\{f_n\}$ and $\{Af_n\}$ are strong Cauchy sequences. Therefore, by Lemma 2.5, $h \equiv$ s-lim $f_n \in D(\bar{A})$, and $g = \bar{A}h + ih$. Thus g belongs to the range of $\bar{A} + i$. This shows that the range of $(\bar{A} + i)$ is closed. The lemma follows by combining this with the result of (i). #

PROPOSITION 2.14 : A symmetric operator A in H is self-adjoint if and only if the range of both of the operators $A \pm i$ is H.

Proof : (i) Suppose $(A \pm i)D(A) = H$. Let $g \in D(A^*)$. Since $(A - i)D(A) = H$, there exists $h \in D(A)$ such that $(A^* - i)g = (A - i)h$. Since A is symmetric, one has $(A - i)h = (A^* - i)h$, which leads to $(A^* - i)(g - h) = \theta$. Now for any $f \in D(A)$

$$0 = ((A^* - i)(g - h),f) = (g - h,(A + i)f).$$

Since $(A + i)D(A) = H$, it follows from Proposition 2.2 that $g - h = \theta$, i.e. $g \in D(A)$. Thus $D(A^*) \subseteq D(A)$. Since $A \subseteq A^*$, we must have $A = A^*$.

(ii) Suppose $A = A^*$. By Lemma 2.13, $(A \pm i)D(A)$ are closed subspaces of H. Thus, if for instance $(A + i)D(A) \neq H$, there exists $g \in \{(A + i)D(A)\}^\perp$. It follows that $g \in D((A + i)^*) =$

$D(A^* - i) = D(A^*)$, and $(A^* - i)g = \theta$. Thus $(g,g) = (g,-iA^*g) = (iAg,g) = (iA^*g,g) = (-g,g) = -(g,g)$, whence $(g,g) = 0$, i.e. $g = \theta$. This proves that $(A + i)D(A) = H$. #

COROLLARY 2.15 : A symmetric operator A in H is essentially self-adjoint if and only if the range of both of the operators $A \pm i$ is dense in H.

Proof : By Lemma 2.13, range $(A \pm i)$ is dense in H if and only if range $(\bar{A} \pm i) = H$, i.e. according to Proposition 2.14 if and only if $\bar{A}$ is self-adjoint. #

In the next proposition we specify a class of operators that are always self-adjoint, namely the maximal multiplication operators by real functions in L^2-spaces. A certain converse of this will be given in Chapter 5 where we shall see that every self-adjoint operator is unitarily equivalent to a multiplication operator in some (more general) L^2-space.

The statement of the proposition involves the notion of the essential supremum of a measurable function $\psi : \Delta \to R$, where Δ is a measurable subset of R^n. ψ is said to be essentially bounded if there exists a number $M < \infty$ such that $|\psi(\underline{x})| \leq M$ for almost all $\underline{x}$ in Δ (with respect to Lebesgue measure). The essential supremum of ψ is the infimum of all numbers M verifying the above condition and will be denoted by $\|\psi\|_\infty$. The set of all essentially bounded measurable func- defined on Δ is denoted by $L^\infty(\Delta)$.

PROPOSITION 2.16 : Let Δ be a measurable set in R^n and ψ a real-valued measurable function defined on Δ which is finite almost everywhere (with respect to Lebesgue measure). Define an operator A in $L^2(\Delta)$ by

$$D(A) = \{f \varepsilon L^2(\Delta) \mid \psi(\underline{x}) f(\underline{x}) \varepsilon L^2(\Delta)\}$$

and
$$(Af)(\underline{x}) = \psi(\underline{x}) f(\underline{x}) \text{ for } f \varepsilon D(A).$$

Then (a) A is self-adjoint.

(b) A is bounded if and only if ψ is essentially bounded, and in that case $\|A\| = \underset{\underline{x}\varepsilon\Delta}{\text{ess sup}} |\psi(\underline{x})| = \|\psi\|_\infty$.

Remarks : (i) A is called maximal since D(A) is the maximal subset in $L^2(\Delta)$ on which multiplication by ψ makes sense in $L^2(\Delta)$.

(ii) The condition that ψ be a measurable function ensures that $\psi(\underline{x}) f(\underline{x})$ is measurable, which is necessary if the latter function is to belong to $L^2(\Delta)$.

Proof : (i) Clearly D(A) is a linear manifold. We show that it is also dense. For this, we define the sets $\Delta_m \equiv \{\underline{x} \varepsilon \Delta \mid |\psi(\underline{x})| \le m\}$, m = 1,2,... . Let $f \varepsilon L^2(\Delta_m)$. Then $f \varepsilon D(A)$, since

$$\int_{\Delta_m} |\psi(\underline{x})|^2 |f(\underline{x})|^2 d^n x \le m^2 \int_{\Delta_m} |f(\underline{x})|^2 d^n x \le m^2 \|f\|^2. \tag{2.49}$$

Thus each of the subspaces $L^2(\Delta_m)$ of $L^2(\Delta)$ belongs to D(A).

Now if $h \varepsilon L^2(\Delta)$, then $\chi_{\Delta_m} h \varepsilon L^2(\Delta_m)$, and if $h \varepsilon [L^2(\Delta_m)]^\perp$, then $\chi_{\Delta_m} h \varepsilon [L^2(\Delta_m)]^\perp$. Thus if h is in $L^2(\Delta)$ as well as in $[L^2(\Delta_m)]^\perp$ for each m, $\chi_{\Delta_m} h$ is the zero vector in $L^2(\Delta_m)$ for each m by Proposition 2.2, i.e. $h(\underline{x}) = 0$ a.e. in $\bigcup_m \Delta_m$. Since the complement of $\bigcup_m \Delta_m$ in Δ is a set of measure zero (the set of points where $\psi(\underline{x})$ is not finite), we have $h(\underline{x}) = 0$ a.e. in Δ, i.e. $h = \theta$. Thus the only vector orthogonal to D(A) in $L^2(\Delta)$ is the vector θ, which means that D(A) is dense in $L^2(\Delta)$ (see the statement following the projection theorem on page 28).

(ii) We now prove (b). If ψ is essentially bounded and M its essential supremum, then one deduces as in (2.49) that $\|A\| \leq M$. If $M_o < M$, let $\Delta(M_o) \equiv \{\underline{x}\varepsilon\Delta \mid |(\psi(x)| \geq M_o\}$. The measure of $\Delta(M_o)$ is positive. If f is a function which is zero outside $\Delta(M_o)$, then $\|Af\| \geq M_o \|f\|$, i.e. $\|A\| \geq M_o$. Hence $M_o \leq \|A\| \leq M$ for each $M_o < M$, i.e. $\|A\| = M$.

If ψ is not essentially bounded, let $m > 0$ and $\Delta(m) = \{\underline{x}\varepsilon\Delta \mid |\psi(x)| \geq m\}$. $\Delta(m)$ has positive measure; hence as above there exists f in $L^2(\Delta)$ such that $\|Af\| \geq m \|f\|$. Since this is true for each $m > 0$, A is unbounded.

(iii) Finally we prove (a). Clearly A is symmetric, i.e. $(f,Ah) = (Af,h)$ for all $f,h\varepsilon D(A)$. For $f\varepsilon L^2(\Delta)$, define $f_\pm(\underline{x}) = f(\underline{x})[\psi(\underline{x}) \pm i]^{-1}$. Since ψ is real, one has $|\psi(\underline{x}) \pm i|^{-1} \leq 1$ and $|\psi(\underline{x})||\psi(\underline{x}) \pm i|^{-1} \leq 1$. These inequalities imply as in (2.49) that $f_\pm\varepsilon L^2(\Delta)$ and $f_\pm\varepsilon D(A)$ respectively. Clearly $(A \pm i)f_\pm = f$. This means that range $(A \pm i) = H$, and by Proposition 2.14 A is self-adjoint. #

To conclude this section we consider convergence properties of sequences of linear operators. We shall mainly be concerned with sequences of bounded operators and therefore restrict the general definitions to this case.

In order to define the notion of convergence of a sequence of operators, one has recourse to the notion of convergence of a sequence of vectors. Thus if $\{A_n\}$ is a sequence of bounded operators with $D(A_n) = H$ for all n, we say that A is the strong limit of A_n if for each $f\varepsilon H$ the sequence of vectors $\{A_nf\}$ converges strongly to the vector Af, i.e. if

$$\lim_{n\to\infty} ||Af - A_n f|| = 0 \quad \text{for all } f\varepsilon H.$$

We then write $A_n \to A$ or $A = \text{s-}\lim_{n\to\infty} A_n$. Similarly A is the <u>weak limit</u> of $\{A_n\}$ if

$$\lim_{n\to\infty} (f,A_n g) = (f,Ag) \quad \text{for all } f,g\varepsilon H.$$

In this case we write $A = \text{w-}\lim_{n\to\infty} A_n$.

A third type of convergence is the <u>convergence in operator norm</u>. A sequence of bounded operators $\{A_n\}$ converges to A in this sense if $||A_n - A||$ converges to zero as $n \to \infty$. This is also called <u>uniform convergence</u> since it is equivalent to the requirement that s-lim $A_n f = Af$ as $n \to \infty$ uniformly on the set $\{f\varepsilon H|\ ||f|| = 1\}$. We shall write u-lim $A_n = A$ as $n \to \infty$. It is clear that the uniform convergence of a sequence $\{A_n\}$ to A implies its strong convergence to A. Also, by Proposition 2.1, the strong convergence of $\{A_n\}$ to A implies weak convergence of $\{A_n\}$ to A.

The Cauchy criterion is valid for each of the three kinds of convergence introduced above; e.g. if $\{A_n\}$ is a sequence of bounded operators with $D(A_n) = H$ and if for each $f\varepsilon H$ the sequence of vectors $\{A_n f\}$ is strongly Cauchy, then there exists a bounded linear operator A such that $A_n \to A$. The proof of this is based on the following interesting fact : if $\{A_n\}$ is a sequence of weakly convergent bounded operators, i.e. such that the sequence of scalar products $\{(f,A_n g)\}$ is a Cauchy sequence of complex numbers for all $f,g\varepsilon H$, then the sequence of norms $\{||A_n||\}$ is bounded, i.e. there exists $M < \infty$ such that $||A_n|| \leq M$ for all $n = 1,2,\ldots$. A fortiori the above property is true for a sequence of strongly convergent and for a sequence of uniformly convergent bounded operators. Similarly a weakly convergent sequence of vectors $\{f_n\}$ is uniformly bounded. This result is known as the <u>uniform boundedness principle</u> and will not be proved here (cf. [AG, no. 29],[RN, no. 84],[RS, Thm. 6.1]). We shall occasionally use it to prove abstract theorems, but in most applications of these theorems the uniform boundedness of the occuring sequences can easily be verified directly.

The formulation of the asymptotic properties of scattering systems is based on strong convergence. For this reason we shall prove here the following two propositions concerning strong convergence. They will be used on various occasions. The first one asserts that, in order to establish the strong convergence of a uniformly bounded sequence of operators $\{A_n\}$, it is sufficient to verify the strong convergence of $\{A_n f\}$ for a fundamental set of vectors f in H (a subset N of H is called <u>fundamental</u> if the linear manifold consisting of all finite linear combinations of vectors belonging to N is dense in H. In particular an orthonormal basis of H and a dense subset of H are fundamental in H).

<u>PROPOSITION 2.17</u> : Let $\{A_n\}$ be a sequence of linear operators with $D(A_n) = H$ and $\|A_n\| \leq M < \infty$ for all $n = 1, 2, \ldots$. Let N be fundamental in H and suppose that $\{A_n f\}$ converges strongly for every $f \in N$. Then there exists a bounded linear operator A with $D(A) = H$, $\|A\| \leq M$ and $\text{s-lim}\, A_n = A$ as $n \to \infty$.

<u>Proof</u> : Let M be the dense linear manifold consisting of all finite linear combinations of vectors belonging to N. Then $\{A_n f\}$ converges strongly for every $f \in M$. The idea is to prove

strong convergence of $\{A_n\}$ on H by approximating an arbitrary $g \varepsilon H$ by an element of M and then using the strong convergence of $\{A_n\}$ on M.

Let $\hat{A}$ be the linear operator defined by $D(\hat{A}) = M$ and $\hat{A}f = \text{s-lim } A_n f$ $(n \to \infty)$ for $f \varepsilon M$. We notice that $||\hat{A}|| \leq M$, since by Proposition 2.1 $||\hat{A}f|| = \lim ||A_n f|| \leq M \, ||f||$ for all $f \varepsilon M$. Let A be the unique extension of $\hat{A}$ to all of H (Proposition 2.6). Then $||A|| \leq M$.

If $g \varepsilon H$, there exists a sequence $\{f_m\} \varepsilon M$ such that $f_m \to g$. From the triangle inequality (2.9) we find for any m and n

$$||Ag - A_n g|| \leq ||Ag - Af_m|| + ||\hat{A}f_m - A_n f_m|| + ||A_n f_m - A_n g||$$
$$\leq M \, ||g - f_m|| + ||\hat{A}f_m - A_n f_m|| + M \, ||g - f_m||.$$

Given $\delta > 0$, one may choose first m so large that $||g - f_m|| < \delta/4M$. Since $f_m \varepsilon M$, there exists $N < \infty$ such that for all $n \geq N$: $||\hat{A}f_m - A_n f_m|| < \delta/2$. Thus $||Ag - A_n g|| < \delta$ for all $n \geq N$, i.e. $A_n g \to Ag$. #

<u>PROPOSITION 2.18</u> : Let $\{A_n\}$, $\{B_n\}$, A and B be bounded operators defined on all of H. (a) If $\text{s-lim } A_n = A$ and $\text{s-lim } B_n = B$, then $\text{s-lim } A_n B_n = AB$ as $n \to \infty$. (b) If $\text{u-lim } A_n = A$ and $\text{u-lim } B_n = B$, then $\text{u-lim } A_n B_n = AB$ as $n \to \infty$.

<u>Proof</u> : It is based on the triangle inequality. If $f \varepsilon H$, then

$$||ABf - A_n B_n f|| = ||(A - A_n)Bf + A_n(B - B_n)f|| \leq ||(A - A_n)Bf|| + ||A_n|| \, ||(B - B_n)f|| \leq ||(A_n - A)Bf|| + M \, ||(B-B_n)f||. \quad (2.50)$$

Here we have used the fact mentioned before Proposition 2.17 that there exists $M < \infty$ such that $||A_n|| \leq M$ for all n. Each term on the right-hand side of (2.50) converges to zero as $n \to \infty$ by the hypotheses, which proves that $A_n B_n f \to ABf$, establishing (a). The proof of (b) is left as an exercise (Problem 2.24). #

As an immediate consequence of this proposition, one sees that the action of a bounded operator can always be interchanged with strong limits :

COROLLARY 2.19 : Let A be a bounded operator with $D(A) = \mathcal{H}$. Suppose $D(B_n) = \mathcal{H}$ and s-lim $B_n = B$ as $n \to \infty$. Then s-lim $AB_n = AB$ and s-lim $B_n A = BA$ as $n \to \infty$.

The following result is a direct consequence of the definitions (Problem 2.24) : If a sequence of bounded and everywhere defined operators $\{A_n\}$ converges strongly to A, then the adjoint sequence $\{A_n^*\}$ converges weakly to A^*. The adjoint sequence need not be strongly convergent though. We shall encounter this question again in Chapter 4 and discuss some special cases there.

PROPOSITION 2.20 : Let A be everywhere defined and $\|A\| < 1$. Then $I - A$ is invertible. Its inverse is bounded, defined everywhere, given by the uniformly convergent series (called the Neumann series)

$$(I - A)^{-1} = \sum_{n=o}^{\infty} A^n = I + A + A^2 + \ldots \qquad (2.51)$$

and verifies $\|(I-A)^{-1}\| \leq (1 - \|A\|)^{-1}$.

The proof of this result is left as an exercise (Problem 2.25).

Finally, we add a few comments concerning strong convergence when unbounded operators are involved. There are two cases that can be considered here. Firstly, a sequence $\{A_n\}$ of bounded operators may converge strongly to an unbounded operator, i.e. one may have $A_n f \to Af$ for every $f \in D(A)$, where A may be an unbounded operator. In this case the sequence of norms $\{\|A_n\|\}$ cannot be bounded. An example of this will be seen in Section 2-4. Secondly the operators A_n may themselves be unbounded and converge strongly on a subset of $\mathcal{H}$ which is common to their domains, i.e. the sequence of vectors $\{A_n f\}$ may be strongly Cauchy for each f in $\bigcap_{n=N}^{\infty} D(A_n)$ for some N (which may depend on f). The limit may define a bounded or an unbounded operator. Whenever unbounded operators are involved in statements using strong convergence, we shall explicitly specify the set of vectors on which convergence takes place. When we speak simply of strong convergence of operators, it is always understood that only bounded operators occur.

2-3 COMPACT OPERATORS

In this section we define compact operators, Hilbert-Schmidt operators and trace class operators and derive some consequences of these definitions. The aim is not to give a complete theory of these classes of operators but rather to assemble those of their properties that will be needed in later chapters. A few of the lengthier proofs will be given only in Section 2-5.

Consider an operator of the form

$$Tf = \sum_{i=1}^{N}(e_i,f)h_i, \text{ with } D(T) = H, \qquad (2.52)$$

where $\{e_i,h_i\}$ are 2N vectors in H and $N < \infty$. The range of T is the finite-dimensional subspace spanned by $h_1,\ldots,h_N$, and $Tf = \theta$ if f is orthogonal to the subspace spanned by $e_1,\ldots, e_N$. Thus T may be viewed as an operator that acts on the finite-dimensional subspace M spanned by $\{e_i,h_i\}$ in the sense that it is zero on $M^\perp$ and its range also lies in M. T may be described by a matrix acting on vectors in M. An operator of the form (2.52) with $N < \infty$ is called a finite rank operator. (An equivalent definition is obtained by requiring that the range of T be finite-dimensional.)

We first give some simple properties of finite rank operators. Henceforth $B(H)$ will denote the set of all bounded and everywhere defined linear operators in H.

LEMMA 2.21 : Let T be a finite rank operator. Then
(a) T^* is a finite rank operator.
(b) If $B \in B(H)$, then BT and TB are finite rank operators.
(c) If T_1 is of finite rank, then so is $T + \alpha T_1 (\alpha \in C)$.

Proof : (b) and (c) are immediate from the definition. To prove (a), let $f,g \in H$. It follows from (2.30) and (2.52) that
$(f,T^*g) = (Tf,g) = \sum_{i=1}^{N}(f,e_i)(h_i,g) = (f,\sum_{i=1}^{N}(h_i,g)e_i)$.
Hence by Proposition 2.2

$$T^*g = \sum_{i=1}^{N}(h_i,g)e_i, \qquad (2.53)$$

which proves that T^* is of finite rank. #

An operator A in $\mathcal{B}(\mathcal{H})$ is said to be compact*) if there exists a sequence $\{T_N\}$ of finite rank operators such that $\|A - T_N\| \to 0$ as $N \to \infty$. It follows from this definition that each finite rank operator is compact and that in a finite-dimensional Hilbert space every operator is compact. We shall denote by $\mathcal{B}_o$ the set of all compact operators. Let us prove some consequences of the preceding definition.

PROPOSITION 2.22 :

(a) A is compact if and only if A^* is compact.

(b) If $A \in \mathcal{B}_o$ and $B \in \mathcal{B}(\mathcal{H})$, then $AB \in \mathcal{B}_o$ and $BA \in \mathcal{B}_o$.

(c) If A_1 and A_2 are compact, so is $A_1 + \alpha A_2$ $(\alpha \in \mathbb{C})$.

(d) If $\{A_n\} \in \mathcal{B}_o$, $A \in \mathcal{B}(\mathcal{H})$ and $\|A - A_n\| \to 0$ as $n \to \infty$, then $A \in \mathcal{B}_o$.

Proof : Let $\{T_N\}$ be a sequence of finite rank operators converging uniformly to A.

(a) This follows from the fact that $\|A-T_N\| = \|A^*-T_N^*\|$ (cf. Proposition 2.8) and Lemma 2.21(a).

(b) By Lemma 2.21(b) BT_N and T_NB are finite rank operators. Now by (2.28) $\|BA-BT_N\| \le \|B\| \|A-T_N\| \to 0$, which shows that BA is compact; similarly one proves $AB \in \mathcal{B}_o$.

(c) This follows from the triangle inequality (2.27) and Lemma 2.21 (c).

---

*) This definition is equivalent to the customary definition of compactness : A is compact iff for every bounded sequence $\{f_n\}$ the sequence $\{Af_n\}$ has a strongly convergent subsequence [RS, Section VI.5]. This latter characterization of compactness is valid in more general metric spaces than Hilbert space, but the one given above is more convenient for our purposes.

(d) Let $\eta > 0$. For each n we choose a finite rank operator $T_{N(n)}$ such that $||A_n - T_{N(n)}|| < \eta/2$. We next choose n so large that $||A-A_n|| < \eta/2$. Then by (2.27) $||A-T_{N(n)}|| \leq ||A-A_n|| + ||A_n - T_{N(n)}|| < \eta$. Hence $||A-T_{N(n)}|| \to 0$ as $n \to \infty$, which proves that A is compact. #

PROPOSITION 2.23 : Let $\{f_n\}$ be a sequence of vectors which converges weakly to θ, and let A be a compact operator. Then $\{Af_n\}$ converges strongly to θ.

Proof : Since $\{f_n\}$ is uniformly bounded by the uniform boundedness principle, there exists $M < \infty$ such that $||f_n|| \leq M$ for all n. Given $\eta > 0$, choose a finite rank operator T of the form (2.52) such that $||A-T|| < \eta/2M$. It follows from (2.9) and (2.52) that

$$||Af_n|| \leq ||(A-T)f_n|| + ||Tf_n|| < \eta/2 + \sum_{i=1}^{N} |(e_i, f_n)| \, ||h_i|| .$$

Since w-lim $f_n = \theta$, there exists n_o such that $|(e_i, f_n)| \, ||h_i|| < \eta/2N$ for all $i = 1, \ldots, N$ and all $n > n_o$. Hence $||Af_n|| < \eta$ for $n > n_o$, which proves that $||Af_n|| \to 0$ as $n \to \infty$. #

Compact operators share the following property with operators acting in a finite-dimensional space : Consider the equation $f-Af = g$ where g is a given vector, A is compact and f is to be determined. Then either the homogeneous equation $f-Af = \theta$ has a non-trivial solution or else the equation $f-Af = g$ has, for any given $g \in H$, a unique solution $f \in H$, namely $f = (I-A)^{-1}g$ ($(I-A)^{-1}$ is then bounded and defined everywhere). The preceding result is known as the Fredholm alternative and follows immediately from Proposition 2.24

below, the proof of which can be found in Section 2-5. If A is for instance a compact integral operator in an L^2-space (cf. (2.65) for its definition), the above is a result about the existence and the uniqueness of the solution of an integral equation. Of course the operator $(I-A)^{-1}$ cannot in general be written down explicitly. However, if for instance $\|A\| < 1$, the solution may be written as a power series in A by using Proposition 2.20.

For a non-compact operator the above alternative need not hold, since $(I-A)^{-1}$ may exist but be an unbounded operator, so that $f-Af = g$ is not solvable in $\mathcal{H}$ for every g. This happens for instance if A is self-adjoint and the point $\lambda = 1$ belongs to its continuous spectrum (Problem 2.26).

PROPOSITION 2.24 : Let A be a compact operator and $z \epsilon \mathcal{C}$, $z \neq 0$. Then either the equation $Af = zf$ has a solution $f \neq \theta$ in $\mathcal{H}$ or $(zI-A)^{-1}$ exists and belongs to $\mathcal{B}(\mathcal{H})$.

We shall now discuss some spectral properties of compact operators. (A general definition of the spectrum of an operator is given in Section 5-6). In an infinite-dimensional Hilbert space the spectrum of a compact operator consists of isolated*) non-zero eigenvalues**) and of the point $z = 0$. The latter may itself be an eigenvalue or an accumulation point of eigenvalues or both. This result is the content of

---

*) z_o is an isolated eigenvalue if there exists a neighborhood of z_o containing no eigenvalue other than z_o.

**) The number $z \epsilon \mathcal{C}$ is an eigenvalue of a linear operator A if there exists a vector $f \neq \theta$ in D(A) such that $Af = zf$. f is called an eigenvector of A.

Proposition 2.25 the proof of which is also deferred to Section 2-5.

PROPOSITION 2.25 : If $A \varepsilon \mathcal{B}_o$, then each non-zero eigenvalue of A is of finite multiplicity (i.e. the corresponding space of eigenvectors is finite-dimensional). Furthermore the only possible accumulation point of the eigenvalues of A is the point $z = 0$.

If $A \varepsilon \mathcal{B}_o$ is self-adjoint, it is possible to choose an orthonormal basis $\{e_i\}$ of $\mathcal{H}$ such that each e_i is an eigenvector of A, i.e. verifying $Ae_i = \lambda_i e_i$ for some (real) λ_i. This is nothing but the spectral theorem for a self-adjoint operator A for the case where A is compact and will be deduced below. It should be said here that for a non-compact self-adjoint operator B such a basis need not exist, since B may also have continuous spectrum (the relevant details will be explained in Chapter 5). We first give a characterization of the eigenvalues and the eigenvectors of a general self-adjoint operator and some lemmas needed to prove the spectral theorem.

PROPOSITION 2.26 : Let A be self-adjoint. Then
(a) All eigenvalues of A are real.
(b) If $Af_1 = \lambda_1 f_1$ and $Af_2 = \lambda_2 f_2$ with $\lambda_1 \neq \lambda_2$, then f_1 is orthogonal to f_2.

Proof : (a) Suppose $Af = \lambda f$ with $f \neq \theta$. By using (2.30) we then obtain $\lambda(f,f) = (f,\lambda f) = (f,Af) = (Af,f) = (\lambda f,f) = \bar{\lambda}(f,f)$, which implies that $\bar{\lambda} = \lambda$.
(b) One obtains as in (a) that

$$(\lambda_1 - \lambda_2)(f_1,f_2) = (Af_1,f_2) - (f_1,Af_2) = 0.$$

Since $\lambda_1 \neq \lambda_2$, this implies that $(f_1,f_2) = 0$. #

LEMMA 2.27 : Let E be a subset of H and denote by M the subspace spanned by E (i.e. the closure of the set $\mathcal{D}$ of all finite linear combinations of vectors belonging to E). Let $A,B\varepsilon\mathcal{B}(H)$ and suppose that $Af = Bf$ for all $f\varepsilon E$. Then $A = B$ on M.

Proof : The hypotheses imply that $Af = Bf$ for all $f\varepsilon\mathcal{D}$. Hence $Af = Bf$ for all f in M by the uniqueness of the closure (Proposition 2.6). #

LEMMA 2.28 : Let $\{F_k\}$, $k = 1,\ldots,n$ be a set of projections with mutually orthogonal ranges, i.e. verifying $F_jF_k = \delta_{jk}F_k$, and let $\alpha_k\varepsilon C$. Then for all $f\varepsilon H$

$$\|\textstyle\sum_k \alpha_k F_k f\|^2 \leq \sup_k|\alpha_k|^2\,\|f\|^2. \tag{2.54}$$

Proof : This follows by applying (2.10) and Bessel's inequality (2.11) :

$$\|\sum_k \alpha_k F_k f\|^2 = \sum_k |\alpha_k|^2 \|F_k f\|^2 \leq \sup_k|\alpha_k|^2 \sum_j \|F_j f\|^2 \leq \sup_k|\alpha_k|^2 \|f\|^2. \ \#$$

LEMMA 2.29 : Let A be a compact self-adjoint operator, and suppose that A has no non-zero eigenvalue. Then $A = 0$.

The proof of this lemma will be indicated in Section 2-5. We shall use it to establish the spectral theorem for compact self-adjoint operators. Suppose $A = A^*$ and $A\varepsilon\mathcal{B}_o$. Let $\{\lambda_k\}$ be an enumeration of all non-zero eigenvalues of A such that $|\lambda_{k+1}| \leq |\lambda_k|$ for all $k = 1,2,\ldots$. We denote by M_k the subspace spanned by all eigenvectors corresponding to the eigenvalue λ_k, by M_o the subspace $M_o = \{f|Af = \theta\}$ and by

$E\{\lambda_k\}$ the orthogonal projection whose range is M_k. One has dim $M_k < \infty$ for $k \neq 0$ by Proposition 2.25 and $E\{\lambda_j\}E\{\lambda_k\} = \delta_{jk}E\{\lambda_k\}$ by Proposition 2.26 (i.e. M_k is orthogonal to M_j if $k \neq j$). The spectral theorem may now be stated as follows.

PROPOSITION 2.30 : Suppose A is self-adjoint and compact. For each $k = 0,1,\ldots$, let $\{e_k^{(i)}\}$ be an orthonormal basis of M_k. Then the set $\{e_k^{(i)}\}_{i,k}$ is an orthonormal basis of H. Furthermore

$$A = \sum_{k\neq 0} \lambda_k E\{\lambda_k\}, \tag{2.55}$$

the sum being convergent in the uniform operator topology.

Proof : (i) Let M be a subspace which is invariant under A, i.e. such that $Af \in M$ for each $f \in M$. Let $g \in M^{\perp}$, $f \in M$. Then $(Ag,f) = (g,Af) = 0$, which shows that $Ag \in M^{\perp}$, or that $M^{\perp}$ is also invariant under A.

(ii) Let M_+ be the subspace spanned by $\{e_k^{(i)}\}_{i,k}$ with $k \neq 0$, F the orthogonal projection with range M_+ and $F' = I-F$. Clearly A leaves M_+ invariant, hence $AM_+^{\perp} \subseteq M_+^{\perp}$ by (i). This means that

$$F'AF' = AF'. \tag{2.56}$$

Two consequences of (2.56) are : (a) The operator AF' is self-adjoint. (b) If $AF'f = \lambda f$ for some $\lambda \neq 0$, then $F'f = f$ and hence $Af = \lambda f$. Therefore, since all eigenvectors of A corresponding to a non-zero eigenvalue lie in M_+, AF' is a self-adjoint compact operator having no non-zero eigenvalue. Thus $AF' = 0$ by Lemma 2.29, or in other words $M_+^{\perp} = M_o$. This shows that the eigenvectors of A (including those in M_o) span H.

(iii) It is easily seen that the series in (2.55) converges uniformly by using Lemma 2.28 and the fact that $|\lambda_k| \to 0$ as $k \to \infty$ if A has an infinite number of non-zero eigenvalues. Its limit is some operator B in $B(H)$. If e is any eigenvector of A, then clearly Ae = Be. The fact that A = B on H now follows from the result of (ii) and Lemma 2.27. #

We next derive a canonical form for an arbitrary compact operator. If $A \varepsilon B_o$, then $A^*A \varepsilon B_o$ by Proposition 2.22. Furthermore A^*A is self-adjoint and positive, i.e. $(f, A^*Af) = \|Af\|^2 \geq 0$ for all $f \varepsilon H$. The preceding identity also shows that $A^*Af = \theta$ implies $Af = \theta$.

It follows that all eigenvalues of A^*A are real and non-negative, and each non-zero eigenvalue has finite multiplicity. Let $\mu_1 \geq \mu_2 \geq \ldots$ be an enumeration of the non-zero eigenvalues of A^*A such that each of them appears as many times as its multiplicity. Let $\{e_{j,o}\}$ be an orthonormal basis of the subspace $M_o = \{f | A^*Af = \theta\}$. By Proposition 2.30 there exists an orthonormal set $\{e_k\}$ such that $\{e_k, e_{j,o}\}$ is a basis of H and such that

$$A^*Ae_k = \mu_k e_k \ , \quad A^*Ae_{k,o} = \theta . \tag{2.57}$$

Let $\lambda_k = \mu_k^{1/2}$. The numbers $\{\lambda_k\}$ are called the <u>singular values</u> of the compact operator A. The following characterization of compact operators is a generalization of (2.52).

PROPOSITION 2.31 (Canonical expansion of compact operators) : Let $A \in \mathcal{B}_o$. Then there exist two orthonormal sets $\{e_k\}$ and $\{h_k\}$ such that for all $f \in H$

$$Af = \sum_k \lambda_k (e_k, f) h_k, \tag{2.58}$$

where $\{\lambda_k\}$ are the singular values of A and the infinite sum (viewed as the limit of a sequence of operators) converges in operator norm.

Proof : Let $\{e_k, e_{j,o}\}$ be the orthonormal basis used in (2.57), and define $h_k = \lambda_k^{-1} A e_k$. One has

$$(h_j, h_k) = \lambda_j^{-1} \lambda_k^{-1} (e_j, A^* A e_k) = \lambda_j^{-1} \lambda_k (e_j, e_k) = \delta_{jk}.$$

Thus $\{h_k\}$ is an orthonormal set.

Define A_N by (2.58) with the sum running from k = 1 to k = N. Let M > N. By using (2.10) and (2.11) one obtains

$$\|(A_M - A_N) f\|^2 = \sum_{k=N+1}^{M} \lambda_k^2 |(e_k, f)|^2 \leq \lambda_{N+1}^2 \sum_{k=N+1}^{M} |(e_k, f)|^2$$
$$\leq \lambda_{N+1}^2 \|f\|^2.$$

Since $\lambda_N \to 0$ as $N \to \infty$, $\{A_N\}$ is a Cauchy sequence in the uniform operator topology. Denote by B its limit. Then

$$Be_j = \lim_{N\to\infty} \sum_{k=1}^{N} \lambda_k (e_k, e_j) h_k = \lambda_j h_j = Ae_j.$$

Clearly $Be_{j,o} = \theta$, and we have already seen that $Ae_{j,o} = \theta$. Hence B and A coincide on an orthonormal basis, i.e. B = A by Lemma 2.27. #

In order to prove that a given operator is compact, one may try to approximate it uniformly by a sequence of finite rank operators or by a sequence of operators that are already known to be compact (Proposition 2.22(d)). For this the Hilbert-Schmidt operators are very useful. Firstly they form a subset of the class of compact operators, and secondly in an L^2-space they have a simple characterization as integral operators (cf. Proposition 2.33), so that in many instances it is easy to decide whether a given operator belongs to the Hilbert-Schmidt class or not.

To define this class of operators, we introduce the Hilbert-Schmidt norm $\|A\|_{HS}$ of an operator A in $\mathcal{B}(\mathcal{H})$:

$$\|A\|_{HS}^2 = \sum_k \|Ae_k\|^2, \tag{2.59}$$

where $\{e_k\}$ is an orthonormal basis of $\mathcal{H}$. A is said to be a Hilbert-Schmidt operator if $\|A\|_{HS} < \infty$, and the set of all Hilbert-Schmidt operators will be denoted by $\mathcal{B}_2$.

In the above definition the quantity $\|A\|_{HS}$ appears to depend on the choice of an orthonormal basis $\{e_k\}$. We shall now show that the sum in (2.59) is the same for each orthonormal basis of $\mathcal{H}$. For this, let $\{g_k\}$ be an arbitrary orthonormal basis. By using the Parseval relation (2.12), (2.30) and again (2.12) one obtains

$$\sum_k \|Ag_k\|^2 = \sum_k \sum_i |(e_i, Ag_k)|^2 = \sum_i \sum_k |(A^*e_i, g_k)|^2$$
$$= \sum_i \|A^*e_i\|^2, \tag{2.60}$$

where the change of the order of summation is permitted because only non-negative terms are involved. Since $\{e_i\}$ may be thought of as being fixed, (2.60) shows that the value of $\sum_k \|Ag_k\|^2$ is the same for each orthonormal basis $\{g_k\}$. (2.60) also implies that

$$\|A\|_{HS} = \|A^*\|_{HS}. \tag{2.61}$$

PROPOSITION 2.32 :

(a) $A\varepsilon\mathcal{B}_2$ if and only if $A^*\varepsilon\mathcal{B}_2$.

(b)
$$\|A\| \leq \|A\|_{HS}. \tag{2.62}$$

(c) Every Hilbert-Schmidt operator is compact, i.e. $\mathcal{B}_2 \subseteq \mathcal{B}_o$.

(d) If $A\varepsilon\mathcal{B}_2$ and $B\varepsilon\mathcal{B}(H)$, then $AB\varepsilon\mathcal{B}_2$ and $BA\varepsilon\mathcal{B}_2$.

(e) If $A_1\varepsilon\mathcal{B}_2$ and $A_2\varepsilon\mathcal{B}_2$, then $(A_1 + \alpha A_2)\varepsilon\mathcal{B}_2$ $(\alpha\varepsilon C)$.

Proof : (a) follows from (2.61). To prove (b), fix $f \neq \theta$ and choose an orthonormal basis $\{e_k\}$ such that $e_1 = f/\|f\|$. Then

$$\frac{\|Af\|^2}{\|f\|^2} = \|Ae_1\|^2 \leq \sum_k \|Ae_k\|^2 = \|A\|_{HS}^2 .$$

Since this inequality holds for each $f \neq \theta$, (b) follows from (2.24).

(c) Let $A\varepsilon\mathcal{B}_2$ and fix an orthonormal basis $\{e_k\}$. For each $N < \infty$ we define a finite rank operator T_N by $T_Nf = \sum_{i=1}^N (e_i,f)Ae_i$. Then $T_Ne_k = 0$ if $k > N$ and $T_Ne_k = Ae_k$ if $k \leq N$. Thus $\|A-T_N\|_{HS}^2 = \sum_{k=N+1}^\infty \|Ae_k\|^2$. It follows from (2.62) and the hypothesis $A\varepsilon\mathcal{B}_2$ that

$$\lim_{N\to\infty} \|A-T_N\|^2 \leq \lim_{N\to\infty} \sum_{k=N+1}^\infty \|Ae_k\|^2 = 0,$$

which proves that A is compact.

(d) We have

$$\|BA\|_{HS}^2 = \sum_k \|BAe_k\|^2 \le \|B\|^2 \sum_k \|Ae_k\|^2 = \|B\|^2 \|A\|_{HS}^2. \quad (2.63)$$

Hence $BA\varepsilon\mathcal{B}_2$. Since $A^*\varepsilon\mathcal{B}_2$ by part (a) and $B^*\varepsilon\mathcal{B}(\mathcal{H})$ by Proposition 2.8, this also implies that $B^*A^*\varepsilon\mathcal{B}_2$. By applying part (a) again, we get $(B^*A^*)^* = AB\varepsilon\mathcal{B}_2$.

(e) One has for $f,g\varepsilon\mathcal{H}$

$$\|f + g\|^2 \le \|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2. \quad (2.64)$$

It follows that $\|A_1 + \alpha A_2\|_{HS}^2 \le 2\|A_1\|_{HS}^2 + 2|\alpha|^2 \|A_2\|_{HS}^2$. #

We now consider integral operators in $\mathcal{H} = L^2(\mathcal{R}^n)$. Let $K_A : \mathcal{R}^{2n} \to \mathcal{C}$ be a measurable function. Then the operator A given by

$$(Af)(\underline{x}) = \int d^n y K_A(\underline{x},\underline{y}) f(\underline{y}) \quad (2.65)$$

is called an integral operator and K_A the kernel of A. The domain of A consists of those functions f in $L^2(\mathcal{R}^n)$ for which the integral in (2.65) exists for almost every $\underline{x}$ and for which the function $(Af)(\underline{x})$ defined by (2.65) belongs to $L^2(\mathcal{R}^n)$.

The function K_A is said to be a Hilbert-Schmidt kernel if

$$M_A \equiv \int d^n x d^n y |K_A(\underline{x},\underline{y})|^2 < \infty. \quad (2.66)$$

One has the following interesting result :

PROPOSITION 2.33 : Let $\mathcal{H} = L^2(\mathcal{R}^n)$.

(a) If A is an integral operator with Hilbert-Schmidt kernel K_A, then A is a Hilbert-Schmidt operator and

$$\|A\|_{HS}^2 = \int d^n x d^n y |K_A(\underline{x},\underline{y})|^2. \quad (2.67)$$

(b) If A is a Hilbert-Schmidt operator, then A is an integral operator with Hilbert-Schmidt kernel.

The proof of this result is given in Section 2-5. One can see from it that an analogous theorem is valid in a general L^2-space. If one evaluates $\|A\|_{HS}$ in the basis used in (2.57), one obtains

$$\|A\|_{HS}^2 = \sum_k |(e_k, A^*Ae_k)| = \sum_k \lambda_k^2. \quad (2.68)$$

Thus $\|A\|_{HS}^2$ is the sum of the squares of the singular values of A.

We now pass on to the trace class operators. For their definition we first introduce the absolute value $|A|$ of a compact operator A. This is done by starting from the orthonormal basis $\{e_k, e_{j,o}\}$ used in (2.57) and defining $|A|e_k = \lambda_k e_k$, $|A|e_{j,o} = \theta$, where $\{\lambda_k\}$ are the singular values of A. This definition is extended by linearity to the set $\mathcal{D}$ of finite linear combinations of these basis vectors, and $|A|$ is the closure of the operator defined in this manner on $\mathcal{D}$ (Proposition 2.6).

We now define the <u>trace norm</u> of $A \in \mathcal{B}_o$ by

$$|||A|||_1 = \sum_k \lambda_k = \sum_k (e_k, |A|e_k) \equiv \mathrm{Tr}|A|. \quad (2.69)$$

The <u>trace</u> of a positive operator B is defined as $\mathrm{Tr}B = \sum_k (g_k, Bg_k)$, where $\{g_k\}$ is an orthonormal basis of $\mathcal{H}$. Since $\mathrm{Tr}B = \|B^{1/2}\|_{HS}^2$, it is independent of the basis $\{g_k\}$ *). An

*) The square-root of a bounded positive operator will be defined in Lemma 5.5.

operator A in $\mathcal{B}_0$ is said to be of trace class if $|||A|||_1$ is finite. The set of all trace class operators will be denoted by $\mathcal{B}_1$.

The derivation of a fair number of the deeper results in scattering theory involves trace class operators. Since we shall not reproduce such arguments in this book, we refrain from proving here the basic properties of trace class operators and simply mention a few of them. A notable one is that an operator belongs to the trace class if and only if it can be written as the product of two Hilbert-Schmidt operators. Such a factorization is the usual method of showing that an operator is trace class. We collect this factorization property in the following proposition. Its proof is based on the canonical expansion of compact operators and can be found in Section 2-5.

PROPOSITION 2.34 :
(a) A belongs to $\mathcal{B}_1$ if and only if $A = BC$ with $B, C \varepsilon \mathcal{B}_2$.
(b) If $A\varepsilon\mathcal{B}_1$ and $T\varepsilon\mathcal{B}(\mathcal{H})$, then $AT\varepsilon\mathcal{B}_1$ and $TA\varepsilon\mathcal{B}_1$.

We may now define the trace of an arbitrary trace class operator by

$$\text{Tr } A = \sum_k (g_k, Ag_k) \tag{2.70}$$

where $\{g_k\}$ is an arbitrary orthonormal basis of $\mathcal{H}$. We have to show that Tr A is finite and independent of the basis $\{g_k\}$. For this we write $A = BC$ with $B, C\varepsilon\mathcal{B}_2$. By using the polarization identity (2.97) one easily gets that

$$4 \sum_k (g_k, Ag_k) = 4 \sum_k (B^*g_k, Cg_k)$$
$$= \|B^*+C\|^2_{HS} - \|B^*-C\|^2_{HS} - i\|B^*+iC\|^2_{HS} + i\|B^*-iC\|^2_{HS}.$$

One sees that the last member of this equation is independent of the basis $\{g_k\}$ and finite by Proposition 2.32. Together with (2.61) the above identity also implies that, if $B,C\varepsilon\mathcal{B}_2$, then Tr BC = Tr CB.

For an operator A in $\mathcal{B}(\mathcal{H})$ to be of trace class it is not sufficient that the sum in (2.70) be convergent (or even absolutely convergent) for some orthonormal basis $\{g_k\}$. A counter-example is the unilateral shift operator in an infinite-dimensional Hilbert space defined by $\Omega g_k = g_{k+1}$, where $\{g_k\}$ is a fixed orthonormal basis. One has $(g_k, \Omega g_k) = 0$ for all k, so that the sum in (2.70) is absolutely convergent. However $\Omega\notin\mathcal{B}_1$ ($\Omega\varepsilon\mathcal{B}_1$ would imply $\Omega^*\Omega = I\varepsilon\mathcal{B}_1$). In order to conclude that $A\varepsilon\mathcal{B}_1$ the sum in (2.73) must be absolutely convergent for every orthonormal basis of $\mathcal{H}$.

We end this section with the following definitions : A $\mathcal{B}(\mathcal{H})$-valued functions $s \mapsto A(s)$ is called <u>strongly continuous</u>, <u>weakly continuous</u>, <u>norm continuous</u> or <u>continuous in Hilbert-Schmidt</u> norm if s-lim A(s) = A(t), w-lim A(s) = A(t), u-lim A(s) = A(t) or $\|A(s)-A(t)\|_{HS} \to 0$ respectively as $s \to t$.

2-4 DIRECT SUMS AND TENSOR PRODUCTS OF HILBERT SPACES

In this section we indicate two methods of constructing from a given family $\mathcal{H}_1,\ldots,\mathcal{H}_n$ of Hilbert spaces a new

Hilbert space H. The idea is essentially to take as elements of H the set $H_1 \times H_2 \times \cdots \times H_n$ of n-tuples $\{f_1, \ldots, f_n\}$ formed of elements f_i of H_i and to define a scalar product between such n-tuples by taking either the sum or the product of the respective scalar products in H_i. This leads to the direct sum and the tensor product of $H_1, \ldots, H_n$. The former is important in multichannel scattering theory, and the latter is involved whenever one deals with a quantum mechanical system which is composed of several subsystems.

Let us begin with two Hilbert spaces H_1 and H_2. The direct sum of H_1 and H_2, denoted by $H_1 \oplus H_2$, is defined as follows : The elements of $H_1 \oplus H_2$ are pairs of vectors $\{f_1, f_2\}$ with $f_i \varepsilon H_i$, and the scalar product between two such pairs is

$$(\{f_1,f_2\},\{g_1,g_2\}) = (f_1,g_1)_{H_1} + (f_2,g_2)_{H_2}. \tag{2.71}$$

Addition and multiplication by scalars are defined by

$$\{f_1,f_2\} + \{g_1,g_2\} = \{f_1 + g_1, f_2 + g_2\}, \tag{2.72}$$

$$\alpha\{f_1,f_2\} = \{\alpha f_1,\alpha f_2\}. \tag{2.73}$$

We leave it to the reader to check that $H_1 \oplus H_2$ is a Hilbert space, i.e. that Axioms I-IV of Section 2-1 are verified (Problem 2.30). One may remark that both H_1 and H_2 may be considered as subspaces of $H_1 \oplus H_2$: H_1 may be identified with the set of pairs of the form $\{f_1, \theta_2\}$ with $f_1 \varepsilon H_1$ and θ_2 the zero vector of H_2, and similarly for H_2. $H_1 \oplus H_2$ is simply the orthogonal sum of H_1 and H_2. In particular, if $\{e_i\}$ is an orthonormal basis of H_1 and $\{h_j\}$ an orthonormal basis of H_2, then the set of pairs $\{\{e_i, \theta_2\}, \{\theta_1, h_j\}\}$ is an orthonormal basis of $H_1 \oplus H_2$. Thus dim $(H_1 \oplus H_2)$ = dim H_1 + dim H_2. We also remark that a Hilbert space H may always be viewed as the direct sum of a subspace M of H and of its orthogonal complement : $H = M \oplus M^{\perp}$.

One may similarly define the direct sum of a finite or countable number of Hilbert spaces H_k. We shall denote this direct sum by $H = \oplus H_k$ $(k = 1,2,\ldots)$. Its elements are sequences $\{f_1, f_2, \ldots\}$ with $f_k \varepsilon H_k$ such that

$$\sum_k \|f_k\|^2_{H_k} < \infty. \tag{2.74}$$

Addition and multiplication by scalars is defined componentwise as in (2.72) and (2.73), and each H_k may again be considered to be a subspace of H. The proof of the completeness of a countably infinite direct sum of Hilbert spaces is similar to that for ℓ^2 (Problem 2.1). The space ℓ^2 gives an example for a countably infinite direct sum, since it may be viewed as a direct sum of one-dimensional Hilbert spaces.

For each k, let A_k be a bounded linear operator in H_k with $\|A_k\| \leq M < \infty$ for all k. We may define an operator A in $H = \oplus H_k$ by

$$A\{f_1, f_2, \ldots\} = \{A_1 f_1, A_2 f_2, \ldots\}. \tag{2.75}$$

We shall use the notation $A = A_1 \oplus A_2 \oplus \cdots = \oplus A_k$ for such an operator. Of course the operators of this form do not exhaust the set of bounded operators on $\oplus H_k$. They are characterized by the property that they leave each H_k invariant. The following rules follow immediately from the above definitions :

$$[\oplus A_k] + [\oplus B_k] = \oplus(A_k + B_k), \tag{2.76}$$

$$[\oplus A_k][\oplus B_k] = \oplus A_k B_k \text{ and } [\oplus A_k]^* = \oplus A_k^*. \qquad (2.77)$$

If the sequence $\{A_k\}$ is not uniformly bounded with respect to k or, more generally, if the operators A_k are unbounded, one may similarly define an operator $A = \oplus A_k$ by (2.75) with $D(A) = \{\{f_1,f_2,\dots\} | f_k \varepsilon D(A_k)$ for each k and $\sum_k \|A_k f_k\|^2_{H_k} < \infty\}$.

Let us now turn to the tensor product $G = H_1 \otimes H_2$ of two Hilbert spaces H_1 and H_2. For this we again consider pairs $\{f_1,f_2\}$ with $f_i \varepsilon H_i$ and try to define a scalar product between two such pairs by

$$(\{f_1,f_2\},\{g_1,g_2\})_G = (f_1,g_1)_{H_1}(f_2,g_2)_{H_2}. \qquad (2.78)$$

Here one is faced with two difficulties. The first one has to do with the linear structure of a Hilbert space. In fact, since $\{g_1,g_2\} + \alpha\{h_1,h_2\}$ has to belong to G, (2.6) requires that

$$(\{f_1,f_2\},\{g_1,g_2\} + \alpha\{h_1,h_2\}) =$$
$$(f_1,g_1)(f_2,g_2) + \alpha(f_1,h_1)(f_2,h_2).$$

Now in general the right hand side cannot be written in the form (2.78), i.e. as the scalar product between two elements of $H_1 \times H_2$. Thus, in order to obtain the linear structure of the tensor product, one has to introduce new elements not contained in $H_1 \times H_2$. This is done by first adding to $H_1 \times H_2$ all finite linear combinations of elements of $H_1 \times H_2$ and extending the scalar product (2.78) by linearity to these new elements.

The second difficulty turns up when one considers (2.7). It is seen that for instance $\|\{f_1,\theta_2\}\|_G = \|\{\theta_1,f_2\}\|_G =$

0 and $\|\alpha\{f_1,f_2\} + \beta\{g_1,f_2\} - \{\alpha f_1 + \beta g_1,f_2\}\|_G = 0$ for any $f_1,g_1 \varepsilon H_1$, $f_2 \varepsilon H_2$ and $\alpha,\beta\varepsilon C$. In order to ensure that Axiom II is verified, one will therefore consider as vectors of $H_1 \otimes H_2$ the equivalence classes of the elements already introduced, two elements being equivalent if their difference has norm zero (that this does define an equivalence relation follows from (2.9) which is valid for the norm $\|\cdot\|_G$ since the Schwarz inequality (2.8) used in its proof can be established wihout using (2.7) [RN, no. 83]). The equivalence class of the pair $\{f_1,f_2\}$ will be denoted by $f_1 \otimes f_2$ and the set of all equivalence classes by $H_1 \hat{\otimes} H_2$. Finally one completes $H_1 \hat{\otimes} H_2$ by the standard method of completing a metric space [RS, Theorem I.1.3] to obtain the Hilbert space $H_1 \otimes H_2$.

We shall now indicate a more explicit way of constructing $H_1 \otimes H_2$. Let $\{e_i\}$ and $\{h_j\}$ be orthonormal bases of H_1 and H_2 respectively. Consider the set of pairs $\{e_i,h_j\}_{i,j}$. They clearly form an orthonormal set with respect to the scalar product (2.78). The tensor product $G = H_1 \otimes H_2$ may then be defined as a Hilbert space in which the above orthonormal set forms an orthonormal basis.

Up to an isomorphism, this latter definition coincides with the first one and is also independent of the choice of a particular basis in H_1 or H_2. To see this, it suffices to verify that each vector of the form $f_1 \otimes f_2$ with $f_i \varepsilon H_i$ can be completely expanded with respect to the orthonormal set $\{e_i \otimes h_j\}$. Now one has $f_1 = \sum\alpha_i e_i$, $f_2 = \sum\beta_j h_j$ with $\sum_{i,j}|\alpha_i\beta_j|^2 = \|f_1\|^2 \|f_2\|^2$, and (2.78) implies that $(e_i \otimes h_j, f_1 \otimes f_2)_G = \alpha_i\beta_j$. Hence

$$\sum_{i,j} |(e_i \otimes h_j, f_1 \otimes f_2)_G|^2 = \|f_1\|^2 \|f_2\|^2 = \|f_1 \otimes f_2\|_G^2,$$

which proves the Parseval relation. It follows that we must have $f_1 \otimes f_2 = \sum_{i,j} \alpha_i \beta_j e_i \otimes h_j$.

There is no canonical way of identifying H_1 or H_2 with subspaces of $H_1 \otimes H_2$. However, if $f_2 \varepsilon H_2$ is a fixed vector such that $\|f_2\| = 1$, the set of vectors $f_1 \otimes f_2$ with f_1 ranging over H_1 is a subspace of $H_1 \otimes H_2$ that is isomorphic to H_1. It will be denoted by $H_1 \otimes f_2$. We also remark that $\dim H_1 \otimes H_2 = \dim H_1 \cdot \dim H_2$.

As an example we consider the tensor product $G_{mn} = L^2(R^m) \otimes L^2(R^n)$. Let $\{e_i(\underline{x})\}$ and $\{h_j(\underline{y})\}$ be orthonormal bases of $L^2(R^m)$ and $L^2(R^n)$ respectively $(\underline{x} \varepsilon R^m, \underline{y} \varepsilon R^n)$. Then the set of functions $\{e_i(\underline{x}) h_j(\underline{y})\}$ forms an orthonormal set in $L^2(R^{m+n})$; indeed

$$(e_i h_j, e_r h_s)_{L^2(R^{m+n})} = \int d^m x d^n y \bar{e}_i(\underline{x}) \bar{h}_j(\underline{y}) e_r(\underline{x}) h_s(\underline{y})$$

$$= (e_i \otimes h_j, e_r \otimes h_s)_{G_{mn}}.$$

It is an interesting fact that the above set of functions is indeed an orthonormal basis of $L^2(R^{m+n})$, which means that $L^2(R^m) \otimes L^2(R^n)$ is naturally isomorphic to $L^2(R^{m+n})$ (identify $f_1 \otimes f_2 \varepsilon L^2(R^m) \otimes L^2(R^n)$ with $f_1 f_2 \varepsilon L^2(R^{m+n})$.) We shall use this result on various occasions, but we leave it to the reader to verify that $g = \theta$ is the only vector in $L^2(R^{m+n})$ which is orthogonal to the set $\{e_i(\underline{x}) h_j(\underline{y})\}$ (Problem 2.31).

Let A_k be linear operators in H_k $(k = 1,2)$. One may define an operator denoted $A_1 \hat{\otimes} A_2$ in $G = H_1 \otimes H_2$ by

$$(A_1 \hat{\otimes} A_2)(f_1 \otimes f_2) = A_1 f_1 \otimes A_2 f_2 \quad \text{for } f_k \varepsilon D(A_k) \tag{2.79}$$

and extending this definition by linearity to the set of all finite linear combinations of vectors of the form $f_1 \otimes f_2$ with $f_k \varepsilon D(A_k)$. This set will be written as $D(A_1) \hat{\otimes} D(A_2)$. If each $D(A_k)$ is dense in the respective Hilbert space $\mathcal{H}_k$, $A_1 \hat{\otimes} A_2$ is densely defined. A particular case is an operator of the form $A_1 \hat{\otimes} I$ whose action differs from the identity only in the first component. In order that the above definition makes sense, one has to verify that, whenever $\sum_{i=1}^m \alpha_i \{f_1^i, f_2^i\}$ and $\sum_{k=1}^n \beta_k \{g_1^k, g_2^k\}$ define the same vector in $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$, then so do $\sum_{i=1}^m \alpha_i \{A_1 f_1^i, A_2 f_2^i\}$ and $\sum_{k=1}^n \beta_k \{A_1 g_1^k, A_2 g_2^k\}$, where $f_j^i, g_j^k \varepsilon D(A_j)$. For this, let $u_k \varepsilon D(A_k)$. (2.30) and (2.78) then imply that

$$\sum_i \alpha_i (\{u_1, u_2\}, \{A_1 f_1^i, A_2 f_2^i\}) = \sum_i \alpha_i (\{A_1^* u_1, A_2^* u_2\}, \{f_1^i, f_2^i\})$$
$$= \sum_k \beta_k (\{A_1^* u_1, A_2^* u_2\}, \{g_1^k, g_2^k\}) = \sum_k \beta_k (\{u_1, u_2\}, \{A_1 g_1^k, A_2 g_2^k\}).$$

Since the set of vectors $\{u_1 \otimes u_2 | u_k \varepsilon D(A_k)\}$ is fundamental in $\mathcal{H}_1 \otimes \mathcal{H}_2$, $\sum_i \alpha_i \{A_1 f_1^i, A_2 f_2^i\} - \sum_k \beta_k \{A_1 g_1^k, A_2 g_2^k\}$ must be in the equivalence class of $\{\theta, \theta\}$ by virtue of Proposition 2.2.

If $A_1 \hat{\otimes} A_2$ is closable, we denote its closure by $A_1 \otimes A_2$. If $A_1 \varepsilon \mathcal{B}(\mathcal{H}_1)$ and $A_2 \varepsilon \mathcal{B}(\mathcal{H}_2)$, then $A_1 \hat{\otimes} A_2$ is bounded (see Problem 2.35), hence $A_1 \otimes A_2 \varepsilon \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. If only bounded operators are involved, one has the following rules :

$$\alpha(A_1 \otimes A_2) = (\alpha A_1) \otimes A_2 = A_1 \otimes (\alpha A_2), \tag{2.80}$$

$$(A_1 \otimes A_2)(B_1 \otimes B_2) = A_1 B_1 \otimes A_2 B_2, \tag{2.81}$$

$$(A_1 \otimes A_2)^* = A_1^* \otimes A_2^*. \tag{2.82}$$

If A_1 is self-adjoint, so is $A_1 \otimes I$ (Problem 2.39). For other properties of $A_1 \hat{\otimes} A_2$, see Problems 7.7, 11.3 and 14.8.

The construction of the tensor product $G = \otimes H_k$ (k = 1,...,n) of a finite number of Hilbert spaces is done in complete analogy with that given above for n = 2 by using instead of (2.78) the definition

$$(\{f_1,\ldots,f_n\},\{g_1,\ldots,g_n\})_G = (f_1,g_1)(f_2,g_2)\ldots(f_n,g_n). \quad (2.83)$$

Operators of the form $A_1 \hat{\otimes} \ldots \hat{\otimes} A_n$ can be defined by an obvious modification of (2.79). Infinite tensor products will not be used in this book.

2-5 NOTES AND SUPPLEMENTARY MATERIAL

A. We add some comments regarding the extension of the Fourier transformation from $S(R^n)$ to $L^2(R^n)$. To simplify the notation we set n = 1. Suppose $f\varepsilon L^1(R)\cap L^2(R)$, and let $\tilde{f}$ be defined by (2.17). Then if $g\varepsilon S \equiv S(R)$,

$$|\int dk\bar{\tilde{f}}(k)\tilde{g}(k)| = (2\pi)^{-\frac{1}{2}}|\int dk\int dx\bar{f}(x)e^{ikx}\tilde{g}(k)|$$

$$= |\int dx\bar{f}(x)g(x)| = |(f,g)| \leq \|f\|\,\|g\|, \quad (2.84)$$

where the interchange of the order of integration is permitted since the integrand is absolutely integrable. Hence the first integral in (2.84) defines a bounded linear functional on S which can be extended by continuity to H. By Proposition 2.3 there exists $f_o\varepsilon L^2(R)$ such that $\int dk\{\bar{\tilde{f}}(k) - \bar{\tilde{f}}_o(k)\}\tilde{g}(k) = 0$ for all $g\varepsilon S$.

Let [a,b] be a bounded interval. By taking a uniformly bounded sequence $\{\tilde{g}_n\}\varepsilon S$ with support in [a-1, b+1] converging pointwise to the characteristic function of [a,b], one gets from the dominated convergence theorem that

$\int_a^b dk\{\bar{\tilde{f}}(k) - \bar{\tilde{f}}_o(k)\} = 0$, which implies $\tilde{f}(k) = \tilde{f}_o(k)$ on $[a,b]$ (cf. [R, Lemma 5.7]). Hence $f = f_o \varepsilon L^2(\mathcal{R})$.

Let $\{f_m\} \varepsilon S$ be such that $f_m \to f$. Then by (2.84) and (2.23), $(\tilde{f},\tilde{g}) = (f,g) = \lim(f_m,g) = \lim(\tilde{f}_m,\tilde{g}) = (\mathcal{F}f,\tilde{g})$ for all $g\varepsilon S$, so that $\tilde{f}$ defined by (2.17) is identical with the vector $\mathcal{F}f$ by Proposition 2.2.

If one requires only $f\varepsilon L^2(\mathcal{R})$, then, as pointed out in Section 2-2, its Fourier transform has to be defined through convergence in the mean. One writes $f = \text{s-lim } f_m$ with $f_m(x) = f(x)$ for $|x| \leq m$ and $f(x) = 0$ for $|x| > m$ and defines $\tilde{f} = \text{s-lim } \tilde{f}_m$ $(m \to \infty)$. $\tilde{f}_m$ is well defined, since $f_m \varepsilon L^1(\mathcal{R}) \cap L^2(\mathcal{R})$ as a consequence of (2.8) :

$$\|f_m\|_{L^1} = \int_{-m}^{m} dx|f(x)| \leq \{\int_{-m}^{m} dx|f(x)|^2\}^{\frac{1}{2}}\{\int_{-m}^{m} ds\}^{\frac{1}{2}}$$
$$\leq \|f\| (2m)^{\frac{1}{2}} < \infty. \qquad (2.85)$$

<u>B</u>. As the <u>Lebesgue dominated convergence theorem</u> is often invoked in this book, we shall state here a simple version of it for the convenience of the reader. For its proof one may consult e.g. [MS, page 169], [R, Thm. 4.15].

<u>PROPOSITION 2.35</u> : Let Δ be a Lebesgue measurable set in $\mathcal{R}^n$. Let $\{f_t\}_{t\varepsilon\mathcal{R}}$ be a family of measurable functions from $\mathcal{R}^n$ to $\mathcal{C}$ and $g\varepsilon L^1(\Delta)$ such that $|f_t(\underline{x})| \leq g(\underline{x})$ on Δ and $\lim f_t(\underline{x}) = f(\underline{x})$ for almost all $\underline{x}$ in Δ as $t \to \tau$, $\tau\varepsilon[-\infty,\infty]$. Then $f\varepsilon L^1(\Delta)$ and

$$\lim_{t\to\tau} \int_\Delta f_t(\underline{x}) d^n x = \int_\Delta f(\underline{x}) d^n x.$$

<u>C</u>. <u>The graph of an operator</u>. The definition of the closure of an operator A involves sequences $\{f_n\}\varepsilon D(A)$ such

that both $\{f_n\}$ and $\{Af_n\}$ are strongly Cauchy. It is useful to combine these two Cauchy sequences into one mathematical object. For this one introduces the <u>graph</u> of A which is defined to be the set $\Gamma(A)$ of all pairs $\{f,Af\}$ with f ranging over D(A). It is natural to regard $\Gamma(A)$ as a subset of $G \equiv H \oplus H$ (this is similar to plotting the graph of a function $f : R \to R$ in the plane R^2). Since A is linear, $\Gamma(A)$ is a linear manifold in G. In general a linear manifold M in G will be the graph of some operator if and only if the elements $\{f_1,f_2\}$ of M are uniquely determined by their first argument f_1. Since M is linear, this condition is equivalent to the requirement that M contain no element of the form $\{\theta,g\}$ with $g \neq \theta$.

The usefulness of regarding $\Gamma(A)$ as a subset of G resides in the following identity :

$$\|\{f,g\} - \{f_n,Af_n\}\|_G^2 = \|f-f_n\|^2 + \|g-Af_n\|^2.$$

One may deduce from this that, if A is closable, then $\Gamma(\bar{A}) = \overline{\Gamma(A)}$, and that A is a closed operator if and only if $\Gamma(A)$ is a closed subspace of $H \oplus H$.

The adjoint operator may also be specified in terms of graphs. To do this one introduces the following unitary operator U in G : $U\{f,g\} = \{-g,f\}$. One then sees that the pair $\{g,g^*\}$ verifies (2.29) if and only if $\{g,g^*\}$ is orthogonal to $U\Gamma(A)$ in G. Since $\Gamma(\bar{A})$ is just the closure of $\Gamma(A)$, this means that $\Gamma(A^*) \oplus U\Gamma(\bar{A}) = G$ if A is closable.

Proposition 2.7 can now be proved by applying twice the preceding identity. (i) Upon replacing A by A* and using

(2.31), we obtain $\Gamma(A^{**}) \oplus U\Gamma(A^*) = G$. (ii) Upon applying U to both of its members, using the identity $U^2 = -I$ and the fact that $\Gamma(\bar{A}) = -\Gamma(\bar{A})$ as sets, we find that $U\Gamma(A^*) \oplus \Gamma(\bar{A}) = G$. By comparing the two equations thus obtained, we infer that $\Gamma(A^{**}) = \Gamma(\bar{A})$, whence $A^{**} = \bar{A}$.

D. We conclude by giving a number of proofs that were omitted in Section 2-3.

Proof of Proposition 2.24 : Since A/z is compact, it suffices to consider the case $z = 1$. So let us assume that the only solution of $Af = f$ is $f = \theta$. Then $I-A$ is invertible, and it remains to show that $(I-A)^{-1} \in \mathcal{B}(H)$.

The idea of the proof is to reduce the problem to a finite-dimensional subspace. For this one chooses a finite rank operator T such that $\|A-T\| < 1$. Then $(I-A+T)^{-1} \in \mathcal{B}(H)$ by Proposition 2.20. Define $Y = T(I-A+T)^{-1}$ and denote by M its range. Clearly Y is a finite rank operator, so that M is a subspace (Problem 2.38) of dimension $n < \infty$. The following identity is also easily checked :

$$(I-A) = (I-Y)(I-A+T). \tag{2.86}$$

Denote by F the projection whose range is M. Then $FY = Y$, implying that $F-YF = F(I-YF)$. Thus the operator $F-YF$ maps M into itself. Suppose there exists $g \in M$ such that $(I-Y)g = (F-YF)g = \theta$. Define $f = (I-A+T)^{-1}g$. It then follows from (2.86) that $(I-A)f = (I-Y)g = \theta$. Hence $f = \theta$ by the original assumption, which implies $g = (I-A+T)f = \theta$. This shows that the operator $F-YF$, considered as a map in M, is invertible. Since this operator is given by a $n \times n$ matrix, the inverse matrix defines a bounded operator in M which will be denoted by $(F-YF)_M^{-1}$. It satisfies

$$(I-Y)(F-YF)_M^{-1} = F \text{ and } (F-YF)_M^{-1}(F-YF) = F. \qquad (2.87)$$

Now define $Z = (F-YF)_M^{-1}F(I+Y-YF) + (I-F)$. By what we have shown above, $Z\varepsilon B(H)$. By using $Y = FY$, $Y(I-Y) = (I-Y)Y$ and (2.87), one obtains

$$Z(I-Y) = (F-YF)_M^{-1}(F-YF)(I-Y) + (F-YF)_M^{-1}(I-Y)FY + (I-F)(I-Y) =$$

$$= F(I-Y) + FY + (I-F)(I-Y) = I.$$

Thus $I-Y$ is invertible, and Z is an extension of $(I-Y)^{-1}$. Similarly one finds $(I-Y)Z = I$, implying that [range $(I-Y)$] = $H = D(Z)$. Hence $Z = (I-Y)^{-1}$. It now follows from (2.86) that $(I-A)^{-1} = (I-A+T)^{-1}(I-Y)^{-1}$, which is in $B(H)$. #

Proof of Proposition 2.25 : Let $\{f_n\}$ be an infinite sequence of linearly independent eigenvectors of A, $\{z_n\}$ the corresponding sequence of eigenvalues. Choose an orthonormal set $\{e_i\}$ such that e_n is a linear combination of $f_1,\dots,f_n$, which can be done by the Schmidt orthogonalization procedure [K, Ch. I.6.3]. Thus

$$e_n = \sum_{i=1}^{n} \alpha_{ni} f_i,$$

and f_k is a linear combination of $e_1,\dots,e_k$. Now

$$Ae_n = \sum_{i=1}^{n} \alpha_{ni} z_i f_i = \sum_{i=1}^{n-1} \alpha_{ni}(z_i - z_n)f_i + z_n e_n$$

$$= \sum_{k=1}^{n-1} \beta_{nk} e_k + z_n e_n$$

and

$$\|Ae_n\|^2 = \sum_{k=1}^{n-1} |\beta_{nk}|^2 + |z_n|^2.$$

Since w-lim $e_n = \theta$ as $n \to \infty$, we have by Proposition 2.23 that $\|Ae_n\|^2 \to 0$ as $n \to \infty$, i.e. $\lim z_n = 0$ as $n \to \infty$.

This shows that an infinite sequence of different eigenvalues necessarily converges to $z = 0$. If z is an eigenvalue of A of infinite multiplicity, one takes for $\{f_n\}$ an infinite sequence of linearly independent vectors verifying $Af = zf$. Then $z_n = z$ for all n, and by the above we have $z = 0$. Thus the only possible eigenvalue of infinite multiplicity is $z = 0$. #

Proof of Lemma 2.29 : (i) We first notice the following consequences of (2.9) and (2.27) respectively

$$\left| \|f\| - \|g\| \right| \leq \|f - g\| \quad \text{for} \quad \text{all } f,g\varepsilon H, \qquad (2.88)$$

$$\left| \|A\| - \|B\| \right| \leq \|A - B\| \quad \text{for} \quad A,B\varepsilon \mathcal{B}(H), \qquad (2.89)$$

which are obtained by arguments similar to those in part (i) of the proof of Proposition 2.1.

(ii) Let $\{T_N\}$ be a sequence of finite rank operators converging uniformly to A. It may be seen as in (2.50) that $\|A^*A - T_N^*T_N\| \to 0$ as $N \to \infty$. Hence $A^*A = A^2$ is the uniform limit of a sequence $Y_N \equiv T_N^*T_N$ of self-adjoint finite rank positive operators.

It follows from (2.30) that $Y_N^*f = \theta$ for all f in $(\text{range } Y_N)^\perp$. Hence Y_N is zero on $(\text{range } Y_N)^\perp$. By part (i) of the proof of Proposition 2.30, range Y_N is invariant under Y_N, i.e. Y_N can be decomposed into the sum of a self-adjoint operator on range Y_N and the zero operator on $(\text{range } Y_N)^\perp$.

(iii) Let $\{e_{iN}\}$, $i = 1,\dots, M(N)$ be a set of mutually orthogonal eigenvectors of Y_N in range Y_N (i.e. $Y_N e_{iN} = \lambda_{iN} e_{iN}$ with $\lambda_{iN} \geq 0$) which span range Y_N (for a self-adjoint operator in a finite-dimensional space, such a set of eigenvectors always exists, cf. [H]). Then Y_N may be written as

$$Y_N f = \sum_{i=1}^{M(N)} \lambda_{iN}(e_{iN},f)e_{iN} . \qquad (2.90)$$

Let us denote by λ_N the largest eigenvalue of Y_N and by e_N one of the corresponding eigenvectors ($\|e_N\| = 1$). As in the proof of Lemma 2.28 one deduces from (2.90) that $\|Y_N\| \leq \lambda_N$. Since $\|Y_N e_N\| = \lambda_N$, we have in fact $\|Y_N\| = \lambda_N$. Since $\|A^2 - Y_N\| \to 0$ as $N \to \infty$, one obtains from (2.89) that $\lim \lambda_N = \lim \|Y_N\| = \|A^2\| \equiv \rho$ as $N \to \infty$.

(iv) We have by (2.88) and (2.9)

$$\begin{aligned}\left|\|A^2e_N\| - \rho\right| &= \left|\|A^2e_N\| - \|\rho e_N\|\right| \leq \|(A^2 - \rho)e_N\| \\ &\leq \|(A^2 - Y_N)e_N\| + \|(Y_N - \lambda_N)e_N\| + |\rho - \lambda_N|\ \|e_N\| \\ &\leq \|A^2 - Y_N\| + |\rho - \lambda_N| \to 0 \text{ as } N \to \infty. \qquad (2.91)\end{aligned}$$

This implies that $\|A^2e_N\| \to \rho$ as $N \to \infty$.

(v) Assume $\rho \neq 0$. Since A is assumed to have no non-zero eigenvalues, we have $(A \pm \rho^{\frac{1}{2}})^{-1} \varepsilon \mathcal{B}(\mathcal{H})$ by virtue of Proposition 2.24. Hence $(A^2 - \rho)^{-1} = (A - \rho^{\frac{1}{2}})^{-1} \cdot (A + \rho^{\frac{1}{2}})^{-1} \varepsilon \mathcal{B}(\mathcal{H})$, i.e. range $(A^2 - \rho) = \mathcal{H}$.

Let $f \varepsilon \mathcal{H}$. Then there exists $g \varepsilon \mathcal{H}$ such that $f = (A^2 - \rho)g$. As $\|(A^2 - \rho)e_N\| \to 0$ by one of the inequalities in (2.91), we get $(f,e_N) = ((A^2 - \rho)g,e_N) = (g,(A^2 - \rho)e_N) \to 0$ as $N \to \infty$. Thus e_N converges weakly to zero. By Proposition 2.23, $\|A^2e_N\| \to 0$ as $N \to \infty$, a contradiction. Hence $\rho = 0$, i.e. $\|A^*A\| = 0$. Therefore $\|Ag\|^2 = (g,A^*Ag) = 0$ for every $g \varepsilon \mathcal{H}$, i.e. $A = 0$. #

<u>Proof of Proposition 2.33</u> : (a) (i) Let

$$\Delta = \{\underline{x} \varepsilon \mathcal{R}^n \,|\, \int |K_A(\underline{x},\underline{y})|^2 \, d^n y = \infty\}.$$

The Lebesgue measure of Δ is zero as a consequence of (2.66). Let $f \in L^2(R^n)$ and $\underline{x} \notin \Delta$. The integral in (2.65) may be viewed as a scalar product in $L^2(R^n)$, so that by applying to it the Schwarz inequality (2.8) one gets

$$|(Af)(\underline{x})|^2 \leq \int d^n y |K_A(\underline{x},\underline{y})|^2 \int d^n y |f(\underline{y})|^2.$$

This implies $\|Af\|^2 \leq M_A \|f\|^2$, i.e. $D(A) = H$ and $A \in B(H)$.

(ii) Let $A \in B(H)$, let $\{e_k\}$ be an orthonormal basis of H and define $\alpha_{jk} = (e_j, Ae_k)$. Then $Ae_k = \sum_j \alpha_{jk} e_j$, and by (2.59) and (2.12)

$$\|A\|^2_{HS} = \sum_k \|Ae_k\|^2 = \sum_{k,j} |\alpha_{jk}|^2. \qquad (2.92)$$

(iii) If $\{e_k\}$ is an orthonormal basis of $L^2(R^n)$, the complex conjugate functions $\{\bar{e}_k\}$ also form an orthonormal basis. Hence the functions $\{e_j(\underline{x})\bar{e}_k(\underline{y})\}$ form an orthonormal basis of $L^2(R^n) \otimes L^2(R^n) = L^2(R^{2n})$ (cf. Section 2-4). Since K_A is in $L^2(R^{2n})$, it may be expanded with respect to the latter basis :

$$K_A = \sum_{j,k} \beta_{jk} e_j \bar{e}_k.$$

The Parseval relation in $L^2(R^{2n})$ gives

$$M_A = \|K_A\|^2_{L^2(R^{2n})} = \sum_{j,k} |\beta_{jk}|^2. \qquad (2.93)$$

Now

$$\alpha_{jk} \equiv (e_j, Ae_k)_{L^2(R^n)} = \int d^n x d^n y \bar{e}_j(\underline{x}) K_A(\underline{x},\underline{y}) e_k(\underline{y})$$

$$= (e_j \bar{e}_k, K_A)_{L^2(R^{2n})} = \beta_{jk}. \qquad (2.94)$$

By combining (2.92) - (2.94) we get (2.67), and since $M_A < \infty$ by hypothesis we have $A \in B_2$.

(b) Define α_{jk} as in part (ii) above, and let

$$K_N(\underline{x},\underline{y}) = \sum_{j,k=1}^{N} \alpha_{jk} e_j(\underline{x}) \bar{e}_k(\underline{y}).$$

K_N is a finite linear combination of functions in $L^2(\mathcal{R}^{2n})$, i.e. $K_N \varepsilon L^2(\mathcal{R}^{2n})$. Let $N' > N$. One has as in (2.93) that

$$\int d^n x d^n y |K_N(\underline{x},\underline{y}) - K_{N'}(\underline{x},\underline{y})|^2$$

$$= \sum_{j=1}^{N} \sum_{k=N+1}^{N'} |\alpha_{jk}|^2 + \sum_{j=N+1}^{N'} \sum_{k=1}^{N'} |\alpha_{jk}|^2$$

$$\leq \sum_{j=1}^{\infty} \sum_{k=N+1}^{\infty} |\alpha_{jk}|^2 + \sum_{j=N+1}^{\infty} \sum_{k=1}^{\infty} |\alpha_{jk}|^2.$$

Since $A \varepsilon \mathcal{B}_2$, the double sum in (2.92) is convergent. This implies that $\{K_N\}$ is a strong Cauchy sequence in $L^2(\mathcal{R}^{2n})$. Therefore it has a limit function $K(\underline{x},\underline{y})$ in $L^2(\mathcal{R}^{2n})$, and by part (a), $K(\underline{x},\underline{y})$ defines a Hilbert-Schmidt operator B in $L^2(\mathcal{R}^n)$.

It remains to show that B = A. For this one deduces as in (2.94) that

$$(e_j, Be_k)_{L^2(\mathcal{R}^n)} = (e_j \bar{e}_k, K)_{L^2(\mathcal{R}^{2n})} = \lim_{N\to\infty} (e_j \bar{e}_k, K_N)_{L^2(\mathcal{R}^{2n})} = \alpha_{jk}.$$

Hence $Be_k = \sum_j \alpha_{jk} e_j$, i.e. $Be_k = Ae_k$ for all k. The fact that B = A is now a consequence of Lemma 2.27. #

<u>Proof of Proposition 2.34</u> : (i) Suppose $A \varepsilon \mathcal{B}_1$. Let $\{e_k\}, \{h_k\}$, $\{\lambda_k\}$ be as in Proposition 2.31. Define B and C by

$$Bf = \sum_j \lambda_j^{\frac{1}{2}} (e_j, f) h_j, \qquad Cf = \sum_k \lambda_k^{\frac{1}{2}} (e_k, f) e_k.$$

Then

$$BCf = \sum_j \lambda_j^{\frac{1}{2}} (e_j, Cf) h_j = \sum_{j,k} \lambda_j^{\frac{1}{2}} \lambda_k^{\frac{1}{2}} (e_j, e_k)(e_k, f) h_j$$

$$= \sum_k \lambda_k (e_k, f) h_k = Af.$$

Since $\sum \lambda_k < \infty$, $\lambda_k \to 0$ as $k \to \infty$. Thus one sees as in the proof of Proposition 2.31 that B as well as C is the uniform limit of a sequence of finite rank operators, i.e. $B,C\varepsilon\mathcal{B}_o$. By using also (2.68) we then get

$$\|B\|_{HS}^2 = \|C\|_{HS}^2 = \sum_k(\lambda_k^{\frac{1}{2}})^2 = |||A|||_1 < \infty,$$

which proves the "only if" part of (a).

(ii) Suppose A = BC with $B,C\varepsilon\mathcal{B}_2$. A is compact by Proposition 2.32. Consider its canonical expansion (2.58). Then one gets from (2.69), (2.58) and (2.8) that

$$|||A|||_1 = \sum_k\lambda_k = \sum_k(h_k,Ae_k) = \sum_k(B^*h_k,Ce_k)$$

$$\leq \sum_k \|B^*h_k\|\,\|Ce_k\| \leq (\sum_k \|B^*h_k\|^2)^{\frac{1}{2}}(\sum_j \|Ce_j\|^2)^{\frac{1}{2}}$$

$$\leq \|B^*\|_{HS}\,\|C\|_{HS} = \|B\|_{HS}\,\|C\|_{HS} < \infty, \tag{2.95}$$

which proves (a). (b) follows from (a) and Proposition 2.32(d). #

PROBLEMS

2.1 : Verify that the following are Hilbert spaces :
(i) The set C^n of all n-tupels $\alpha = \{\alpha_1,\dots,\alpha_n\}$ of complex numbers ($\alpha_i\varepsilon C$) with the scalar product

$$(\alpha,\beta) = \sum_{i=1}^n \bar{\alpha}_i\beta_i . \quad (\alpha,\beta\varepsilon C^n) \tag{2.96}$$

(ii) The set ℓ^2 of all infinite sequences $\alpha = \{\alpha_1,\alpha_2,\dots\}$ of complex numbers which satisfy $\sum_i|\alpha_i|^2 < \infty$, with the scalar product given by (2.96) for $n = \infty$ [AG, Section 4].

2.2 : Verify that (i) $\|\theta\| = 0$, $(f,\theta) = 0$ for all $f\varepsilon H$, (ii) $(\alpha f,g) = \bar{\alpha}(f,g)$, $(f+\alpha h,g) = (f,g) + \bar{\alpha}(h,g)$.

2.3 : Prove that a sequence $\{f_n\}$ can converge strongly to at most one vector f.

2.4 : (a) If $f_n \to f$, $g_n \to g$, then $(f_n,g_n) \to (f,g)$ as $n\to\infty$.

(b) If w-lim $f_n = f$ and $\|f_n\| \leq \|f\|$ for all n, then s-lim $f_n = f$ as $n \to \infty$. (c) Show that in a finite-dimensional Hilbert space weak convergence implies strong convergence.

2.5 : Let $\{e_i\}$ be an orthonormal basis of $\mathcal{H}$ and $f \varepsilon \mathcal{H}$. Define $f_n = \sum_{k=1}^{n} (e_k, f) e_k$. (i) Show that $f_n \to f$ as $n \to \infty$. (ii) Prove Parseval's relation (2.12). (Hint : Use (2.11) to show that $\{f_n\}$ is a Cauchy sequence. Use Proposition 2.2 to prove that its limit is f. Use Proposition 2.1 for (ii).)

2.6 : Verify that a linear manifold $\mathcal{M}$ satisfies axioms II and IV. (Hint : Let $\{f_n\}$ be a dense sequence in $\mathcal{M}$, $\bar{\mathcal{M}}$ the closure of $\mathcal{M}$ and $\eta_m = 1/m$. Define $f_{n,1}$ from the projection theorem and show that the sequence $\{f_{n,1}\}$ is dense in $\bar{\mathcal{M}}$. For each n choose a sequence $\{g_{nm}\} \varepsilon \mathcal{M}$ such that $\|g_{nm} - f_{n,1}\| < \eta_m$).

2.7† : Prove Proposition 2.3. (Hint : Show that the set $\mathcal{N} = \{f | \Phi(f) = 0\}$ is a subspace. If $\Phi \not\equiv 0$, there exists $h \varepsilon \mathcal{N}^\perp$ such that $\Phi(h) = 1$. Define $g = h/\|h\|^2$. If $f \varepsilon \mathcal{H}$, then $f - \Phi(f) h \varepsilon \mathcal{N}$, which implies $(g,f) = \Phi(f)$.)

2.8† : In $L^2(\mathbb{R})$, let $\chi_{[a,b]}$ be the characteristic function of the interval [a,b], and let $\eta > o$. Show that there exists a function g in $C_o^\infty(\mathbb{R})$ such that $\|\chi_{[a,b]} - g\| < \eta$. (Hint : Replace $\chi_{[a,b]}$ near a and b by a function similar to $\exp[x^2(x^2-M^2)^{-1}]$, $|x| \leq M$.)

2.9 : Prove Proposition 2.4(b) for n > 1.

2.10 : Let α be a complex number with Re $\alpha > 0$. Make the appropriate choice of the branch of $\sqrt{\alpha}$ and show that the Fourier transform of the function $\exp(-\alpha|\underline{x}|^2/2)$ is $\alpha^{-n/2}\exp(-|\underline{k}|^2/2\alpha)$.

2.11 : Prove the second equality (2.23).

2.12 : Verify the equivalence of the two conditions for an operator to be closable given before Lemma 2.5.

2.13 : Let A be closable, $\bar{A}$ its closure and A' a closed extension of A. Prove that $\bar{A} \subseteq A'$.

2.14 : Let $\{e_i\}$ be an orthonormal basis of an infinite-dimensional Hilbert space $\mathcal{H}$, and define a linear operator A as follows : D(A) is the set of all finite linear combinations of vectors of $\{e_i\}$, and $Ae_k = ke_1$. Show that A is not closable. Verify also that D(A*) is not dense in $\mathcal{H}$. Find a sequence of bounded operators converging strongly to A on D(A). (Hint : Consider the sequence $\{f_n\}$ with $f_n = n^{-1}e_n$).

2.15 : Let Ω be a linear operator such that $\|\Omega f\| = \|f\|$ for all $f \in D(\Omega)$, and denote by $\bar{\Omega}$ the closure of Ω. Verify that $\|\bar{\Omega}f\| = \|f\|$ for all $f \in D(\bar{\Omega})$.

2.16 : Prove equations (2.27) and (2.28).

2.17 : Prove the following statements. (a) If A is closed and invertible, then A^{-1} is closed. (b) If in addition A^{-1} is bounded and defined everywhere and B is closed, then AB is closed. (c) If A is closed and C bounded with $D(C) = \mathcal{H}$, then A+C is closed. (d) If $\bar{A}$ and B are bounded with $D(\bar{A}) = D(B) = \mathcal{H}$ and AB is densely defined, then $\overline{AB} = \bar{A}B$.

2.18 : Verify the assertion (2.31).

2.19 : Show that, if A** exists, then $A \subseteq A^{**}$.

2.20 : Verify (2.32), (2.33) and (2.34).

2.21 : Show that (2.36) defines an orthogonal projection.

2.22 : Prove the polarization identity

$$4(f,g) = \|f+g\|^2 - \|f-g\|^2 - i\|f+ig\|^2 + i\|f-ig\|^2. \qquad (2.97)$$

Show that, for $f,g \in D(A)$, (f,Ag) can be similarly expressed as a sum of four terms of the form $\beta(h,Ah)$ with $\beta \in \mathbb{C}$.

2.23 : Suppose $A \subseteq A^*$. Show that $\bar{A}$ is self-adjoint if and only if $A^* = A^{**}$.

2.24 : (a) Suppose that $D(A_n) = D(A) = \mathcal{H}$, A_n and A are bounded and $A_n \to A$. Show that $\{A_n^*\}$ converges weakly to A^* as $n \to \infty$. (b) Prove Proposition 2.18(b).

2.25 : Prove Proposition 2.20. (Hint : Use (2.28) and the identity $I-A^{n+1} = (I+A+\ldots+A^n)(I-A)$.)

2.26 : Let $\mathcal{H} = L^2(\mathbb{R})$ and let Q be the maximal multiplication operator defined by $(Qf)(x) = xf(x)$. Show that $(I-Q)^{-1}$ exists but is unbounded.

2.27 : Give an example of a compact operator which is not in the Hilbert-Schmidt class.

2.28 : Find necessary and sufficient conditions for a projection and for a partial isometry to be compact, Hilbert-Schmidt or trace class.

2.29 : (Polar decomposition) : Let A be closed. Then there exists a positive self-adjoint operator $|A|$ with $D(|A|) = D(A)$ and a partial isometry Ω with initial set $\overline{|A|\mathcal{H}}$ and final set $\overline{A\mathcal{H}}$ such that $A = \Omega|A|$. (This may be viewed as a generalization of the polar decomposition $z = \exp(i\phi)|z|$ of complex numbers). Verify the above statement for the case where A is compact by using Proposition 2.31.

2.30 : Verify that the space $\mathcal{H}_1 \oplus \mathcal{H}_2$ satisfies axioms I-IV.

2.31†: Prove that there is an isomorphism from $L^2(\mathcal{R}^m) \otimes L^2(\mathcal{R}^n)$ onto $L^2(\mathcal{R}^{m+n})$, see Section 2-4. (By an isomorphism between two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ we mean a linear map j from $\mathcal{H}_1$ onto $\mathcal{H}_2$ such that $(jf,jg)_2 = (f,g)_1$ for all $f,g\varepsilon\mathcal{H}_1$.)

2.32†: Show that
(a) the direct sum $A = \oplus A_k$ of a sequence of self-adjoint operators $\{A_k\}_{k=1}^{\infty}$ is self-adjoint,
(b) the direct sum $F = \oplus F_k$ of a sequence of projections $\{F_k\}_{k=1}^{\infty}$ is a projection,
(c) $\otimes F_k$ is a projection in $\otimes \mathcal{H}_k$ $(k=1,\ldots,n)$.

2.33 : Let E and F be two projections. Show that the following three statements are equivalent : (i) $E\mathcal{H} \subseteq F\mathcal{H}$, (ii) $EF = E = FE$, (iii) $E \leq F$ (i.e. $F-E \geq 0$).

2.34 : Show that $A^*A\varepsilon\mathcal{B}_o$ if and only if $A\varepsilon\mathcal{B}_o$. (Hint : Use Proposition 2.31.)

2.35 : (a) Show that $\|A \hat{\otimes} I\| = \|A\|$. (Hint : Use an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$). (b) If $A\varepsilon\mathcal{B}(\mathcal{H}_1)$, $B\varepsilon\mathcal{B}(\mathcal{H}_2)$, then $\|A \otimes B\| = \|A\|\,\|B\|$. [Hint : Use (a)]. (c) Suppose that $\|A_n\| \leq M$, $\|B_n\| \leq M$, $A_n \to A$ and $B_n \to B$. Show that $A_n \otimes B_n \to A \otimes B$. (Hint : Use Proposition 2.17.)

2.36 : (Cyclicity of the trace) : Let $A\varepsilon\mathcal{B}_1$, $B,C\varepsilon\mathcal{B}_2$, $D\varepsilon\mathcal{B}(\mathcal{H})$. Then TrBCD = TrDBC = TrCDB and TrAD = TrDA.

2.37 : (a) Let A be Hilbert-Schmidt or trace class. Show that the infinite sum in (2.58), viewed as the limit of a sequence of operators, converges in Hilbert-Schmidt or trace norm respectively. (b) Let $B \varepsilon \mathcal{B}_1$. Then $\|B\| \leq \|B\|_{HS} \leq |||B|||_1$.

2.38 : Show that the range of a finite rank operator is a (closed) subspace.

2.39 : If A is self-adjoint in $\mathcal{H}_1$, then $A \hat{\otimes} I$ is essentially self-adjoint in $\mathcal{H}_1 \otimes \mathcal{H}_2$. If furthermore $\mathcal{H}_2$ is finite-dimensional, then $A \hat{\otimes} I = A \otimes I$.