FUNCTIONS OF SEVERAL SELF-ADJOINT OPERATORS

MICHAEL E. TAYLOR

Hermann Weyl, in [6], defined a functional calculus to deal with the unbounded selfadjoint operators of differentiation and multiplication by a position coordinate. In this paper we examine this calculus in the case of bounded operators.[1] We let x be a vector $(x_1, \cdots, x_n)$ in R^n, $dx = dx_1 \cdots dx_n$, $\langle \lambda, x \rangle = \lambda_1 x_1 + \cdots + \lambda_n x_n$. We let A be an n-tuple of bounded selfadjoint operators $(A_1, \cdots, A_n)$ and $\langle \lambda, A \rangle = \lambda_1 A_1 + \cdots + \lambda_n A_n$. Define $\exp(i\langle \lambda, A \rangle)$ by the usual power series, which converges in norm since $\langle \lambda, A \rangle$ is bounded. If $\hat{f}(\lambda)$ is the Fourier transform of f, $\hat{f}(\lambda) = (2\pi)^{-n/2} \int f(x) \exp(i\langle \lambda, x \rangle) dx$, and if both f and $\hat{f}$ are in L^1, we define $f(A) = (2\pi)^{-n/2} \int \hat{f}(\lambda) \exp(-i\langle \lambda, A \rangle) d\lambda$. This is the Bochner integral, whose convergence in norm is guaranteed since $e^{-i\langle \lambda, A \rangle}$ is a continuous function of λ and is of norm 1.

The questions which shall concern us are: in what way is $f(A)$ continuous as a function of A or as a function of f?, what are some classes of functions for which $f(A)$ is easy to calculate?, to what extent are other functional calculi special cases of this one?, and in what ways can we extend this calculus?

We should point out that the chief interest of this calculus is that the operators involved do not have to commute, which is not true of the calculi based on the Cauchy integral formula or on the spectral theorem. However it is easy to verify that in the case where the A_i commute, $f(A)$ is the same as the operator defined by the spectral theorem. For example, in the one variable case we have

$$f(A) = (2\pi)^{-1/2} \int \hat{f}(\lambda) \exp(-i\langle \lambda, A \rangle) d\lambda = (2\pi)^{-1/2} \int \hat{f}(\lambda) d\lambda \int e^{-i\lambda\theta} dE_\theta$$

$$= (2\pi)^{-1/2} \int dE_\theta \int \hat{f}(\lambda) e^{-i\lambda\theta} d\lambda = \int f(\theta) dE_\theta$$

where the interchange in order of integration is permitted because for each x and y in the Hilbert space H on which the A_i operate, the integral $\int |\hat{f}(\lambda)| d\lambda \int |(dE_\theta x, y)|$ is finite, where $(\ ,\ )$ is the inner product in H. The verification in the case of several commutative operators is the same.

As for the continuity of $f(A)$ as a function of A, we have the following. We say that a sequence A_m of n-tuples converges if each component converges.

PROPOSITION 1. *Let f be fixed. If $A_m \to A$ in norm, then $f(A_m) \to f(A)$ in norm.*

PROOF. If $A_m \to A$ in norm then for each λ, $\exp(-i\langle \lambda, A_m \rangle)$ $\to \exp(-i\langle \lambda, A \rangle)$ in norm. The proposition follows by Lebesgue's dominated convergence theorem.

The same reasoning gives the following.

PROPOSITION 1a. *Let f be fixed. If $A_m \to A$ strongly, then $f(A_m) \to f(A)$ strongly.*

We now want to examine the map $\eta: F \to L(H, H)$ given by $\eta(f)$ $= f(A)$, where A is fixed, $F = \{f \in L^1: \hat{f} \in L^1\}$, $L(H, H)$ is the space of bounded linear operators on H, with the norm topology. We make F into a normed linear space with norm $\|f\| = \int |\hat{f}(x)| \, dx$.

PROPOSITION 2. *η is continuous.*

PROOF. Clearly $\|\eta(f)\| \leq (2\pi)^{-n/2} \int |\hat{f}(x)| \, dx$.

It is easy to see that the induced map $S \to L(H, H)$ is continuous, where S is the Schwartz space of rapidly decreasing functions, and since the injection $C_0^\infty \to S$ is continuous, it follows that the composition $\xi: C_0^\infty \to L(H, H)$ is continuous. Hence for each x and y in H, the map $f \to (f(A)x, y)$ is a distribution. The substance of our first theorem will be that for fixed A each such distribution has support in a ball about the origin that does not depend on the elements x and y. We shall state this in the language of spectral theory.

A point p in R^n is said to be in the resolvent set of A if there is a neighborhood U of p such that for any $f \in C_0^\infty$ with support in U, $f(A) = 0$. Otherwise, p is said to be in the spectrum of A. The spectrum of A is denoted $\sigma(A)$. We immediately see that the resolvent set of A is open in R^n, and hence $\sigma(A)$ is closed. Before we state the theorem we shall need a lemma.

LEMMA. *If A_1 and A_2 are bounded selfadjoint operators, then $\|\exp(A_1+iA_2)\| \leq \|\exp(A_1)\|$.*

PROOF. Let $b > \|A_1\|$ and let $B_b = A_1 - bI$. Then $\exp(B_b+iA_2)$ $= \lim_{n\to\infty} (\exp(B_b/n) \exp(iA_2/n))^n$ in norm. This is a special case of the Trotter product formula, and its proof is similar to that of the Theorem 5.3 on page 31 of [3].[2] Hence

[2] See also [3a].

$$\left\| \exp(B_b + iA_2) \right\| \leqq \lim_{n \to \infty} \left\| \exp(B_b/n) \right\|^n \left\| \exp(iA_2/n) \right\|^n = \left\| e^{B_b} \right\|$$

since $\exp(iA_2/n)$ is unitary and $e^{B_b/n}$ is selfadjoint. Thus

$$\left\| \exp(A_1 + iA_2) \right\| = e^b \left\| \exp(B_b + iA_2) \right\| \leqq e^b \left\| e^{B_b} \right\| = \left\| e^{A_1} \right\|.$$

THEOREM 1. *The spectrum of an n-tuple of bounded selfadjoint operators is compact.*

PROOF. Consider the entire function $W_{A,x,y}(z) = (\exp(-i\langle z, A \rangle)x, y)$. Note that $\left| W_{A,x,y}(z) \right| \leqq \left\| \exp(-i\langle z, A \rangle) \right\| \left\| x \right\| \left\| y \right\| \leqq \left\| x \right\| \left\| y \right\| \exp(S \left| \operatorname{Im} z \right|)$, where $S = \left(\sum \left\| A_i \right\|^2 \right)^{1/2}$, by the above lemma. By the Paley-Wiener theorem for distributions, $W_{A,x,y}$ is the inverse Fourier transform $\tilde{T}$ of a distribution T with support in the ball $B_S = \left\{ x \in R^n : \left| x \right| \leqq S \right\}$, if we identify entire functions with tempered distributions in the usual manner. Hence $T(f) = \tilde{T}(\hat{f}) = W_{A,x,y}(\hat{f}) = \int \hat{f}(\lambda)(\exp(-i\langle \lambda, A \rangle)x, y)d\lambda = (2\pi)^{n/2}(\xi(f)x, y)$. Since x and y are arbitrary, we see that the spectrum of A is contained in B_S.

The map $\xi: C_0^\infty \to L(H, H)$ has the property that $f = g$ on a neighborhood of $\sigma(A)$ implies $\xi(f) = \xi(g)$. Hence ξ can be extended to $\zeta: C^\infty \to L(H, H)$, by letting $\zeta(f) = \xi(h)$ where $h \in C_0^\infty$ is any function such that $f = h$ on a neighborhood of $\sigma(A)$.

We thus have two functional calculi, whose domains overlap on $C^\infty \cap F$. We would like to prove that the two definitions of $f(A)$ agree there.

PROPOSITION 3. C_0^∞ *is dense in* F.

PROOF. We first show that S is dense in F. Let $f \in F$. Then $\hat{f} \in F$. Since S is dense in L^1, there is a sequence of functions $g_n \in S$ such that $\int \left| \hat{f} - g_n \right| dx \to 0$ as $n \to \infty$. Let $h_n = \check{g}_n$, the inverse Fourier transform of g_n. Then $h_n \in S$ and we have $\left\| f - h_n \right\| = \int \left| \hat{f} - \hat{h}_n \right| dx = \int \left| \hat{f} - g_n \right| dx \to 0$ as $n \to \infty$. Thus S is dense in F. It is known that C_0^∞ is dense in S, and since the topology of S is stronger than its relative topology as a subspace of F, it follows that C_0^∞ is dense in F.

Recall that η and hence ξ are continuous on C_0^∞ as a normed linear space with the norm of F. Since C_0^∞ is dense in $C^\infty \cap F$, there is a unique extension of ξ to $C^\infty \cap F$. As we saw, our first definition of $f(A)$ gives such an extension. It suffices to show that so does ζ.

PROPOSITION 4. $\zeta \mid (C^\infty \cap F)$ *is continuous, where* $C^\infty \cap F$ *is normed with the norm of* F.

PROOF. Let $f \in C^\infty \cap F$. $\zeta(f) = \xi(f\phi)$ where ϕ is a function in C_0^∞ equal to 1 in a neighborhood of $\sigma(A)$. If we let $\left\| \ \right\|_1$ be the L^1 norm, we have

$$\left\|\zeta(f)\right\| = \left\|\xi(f\phi)\right\| \leq (2\pi)^{-n/2}\left\|(f\phi)^\wedge\right\|_1 = (2\pi)^{-n}\left\|\hat{f} * \hat{\phi}\right\|_1 \leq (2\pi)^{-n}\left\|\hat{\phi}\right\|_1\left\|\hat{f}\right\|_1.$$

Thus $\zeta \mid (C^\infty \cap F)$ has $(2\pi)^{-n}\left\|\hat{\phi}\right\|_1$ as a bound.

We are now in a position to calculate $f(A)$ for certain important classes of functions f. We first consider the case where f is a polynomial.

LEMMA. *Any derivative of* $\exp(-i\langle\lambda, A\rangle)$ *is bounded as a function of* λ, *for real* λ.

PROOF. Note that $\exp(-i\langle\lambda, A\rangle)$ is an entire function of λ. Hence we can apply the Cauchy integral formula to get

$$\left\|\frac{\partial^{k_1 + \cdots + k_n}}{\partial\lambda_1^{k_1} \cdots \partial\lambda_n^{k_n}} \exp(-i\langle\lambda, A\rangle)\right\| \leq \frac{(k_1!) \cdots (k_n!)}{(2\pi)^n}$$

$$\cdot \int_{|\zeta_i - \lambda_i| = 1} \left\|\exp(-i\zeta_1 A_1 - \cdots - \zeta_n A_n)\right\| \left|d\zeta_1\right| \cdots \left|d\zeta_n\right|.$$

Now if $\zeta_j = x_j + iy_j$ we have

$$\exp(-i\langle\zeta, A\rangle) = \exp(-i(x_1 A_1 + \cdots + x_n A_n) + y_1 A_1 + \cdots + y_n A_n).$$

Since $\left|y_j\right| \leq 1$, we see that $\left\|\exp(-i\langle\zeta, A\rangle)\right\| \leq \exp(\left\|A_1\right\| + \cdots + \left\|A_n\right\|)$ on the domain of integration. Hence

$$\left\|\frac{\partial^{k_1 + \cdots + k_n}}{\partial\lambda_1^{k_1} \cdots \partial\lambda_n^{k_n}} \exp(-i\langle\lambda, A\rangle)\right\| \leq (k_1!) \cdots (k_n!) \exp\left(\sum\left\|A_j\right\|\right).$$

In the following proposition we write $p(i\partial/\partial\lambda)$ as a shorthand for $p(i\partial/\partial\lambda_1, \cdots, i\partial/\partial\lambda_n)$, where $p(x_1, \cdots, x_n)$ is a polynomial.

THEOREM 2. $p(A) = [p(i\partial/\partial\lambda) \exp(-i\langle\lambda, A\rangle)]_{\lambda=0}$.

PROOF. Let $\phi \in C_0^\infty$ satisfy $0 \leq \phi \leq 1$, $\phi = 1$ in a neighborhood of B_1, and let $\phi_a(x) = \phi(a^{-1}x)$ for $a \geq 1$. Note that $\hat{\phi}_a(\lambda) = a^n\hat{\phi}(a\lambda)$. Since $\hat{\phi} \in S$, we can pick W so that $\left|\hat{\phi}(\lambda)\right| < W\left|\lambda\right|^{-n-1}$ for all λ. Then $\left|\hat{\phi}_a(\lambda)\right| = a^n\left|\hat{\phi}(a\lambda)\right| < a^n W\left|a\lambda\right|^{-n-1} = a^{-1}W\left|\lambda\right|^{-n-1} \leq W\left|\lambda\right|^{-n-1}$, which is integrable on $R^n - B_r$ for any $r > 0$, and which does not depend on a. For all a so large that $\sigma(A) \subset B_a$ we have

$$p(A) = (2\pi)^{-n/2} \int [p(x)\phi_a]^\wedge(\lambda) \exp(-i\langle\lambda, A\rangle) d\lambda$$

$$= (2\pi)^{-n/2} \int [p(-i\partial/\partial\lambda)\hat{\phi}_a(\lambda)] \exp(-i\langle\lambda, A\rangle) d\lambda$$

$$= (2\pi)^{-n/2} \int \hat{\phi}_a(\lambda)[p(i\partial/\partial\lambda) \exp(-i\langle\lambda, A\rangle)] d\lambda$$

where integration by parts is justified since $p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)$ is bounded. Since for all $a \geq 1$ we have a common bound for $\hat{\phi}_a(\lambda)$ which is integrable on $R^n - B_r$ we can pick N so large that

$$\left\|p(A) - (2\pi)^{-n/2}\int \hat{\phi}_a(\lambda)\phi_N(\lambda)[p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)]d\lambda\right\| < \epsilon$$

for all $a \geq 1$, where ϵ is arbitrary. Now

$$(2\pi)^{-n/2}\int \hat{\phi}_a(\lambda)\phi_N(\lambda)[p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)]d\lambda$$

$$= (2\pi)^{-n}\int \phi_N(\lambda)[p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)]d\lambda \int \phi_a(x)\exp(i\langle\lambda, x\rangle)dx$$

$$= (2\pi)^{-n}\int \phi_a(x)dx \int \phi_N(\lambda)[p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)]\exp(i\langle\lambda, x\rangle)d\lambda$$

$$= (2\pi)^{-n/2}\int \phi_a(x)\hat{h}(x)dx,$$

where $h(\lambda) = \phi_N(\lambda)[p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)]$. Note that $\hat{h} \in L^1$. Pick a so large that this last integral is within ϵ of $(2\pi)^{-n/2}\int\hat{h}(x)dx = h(0)$. Then $p(A)$ is within 2ϵ of $[p(i\partial/\partial\lambda)\exp(-i\langle\lambda, A\rangle)]_{\lambda=0}$. Since ϵ is arbitrary, the theorem follows.

To calculate $p(A)$ more explicitly we look at the power series for $\exp(-i\langle\lambda, A\rangle)$. Hence we see that to evaluate $p(A)$, symmetrize the corresponding polynomial in $A_1, \cdots, A_n$. For instance, $x_j x_k(A) = \frac{1}{2}(A_j A_k + A_k A_j)$.

Note that the proof of Theorem 2 did not make use of the spectral theorem. Since it is easy to prove that the correspondence $A \to f(A)$ is of positive type in the one variable and commutative cases, we can deduce the spectral theorem from Theorem 2. The proof given has the further advantage of being valid when the arguments belong to certain larger classes of Banach algebras, for which the spectral theorem does not hold.

Let us take a quick look at the two variable case. Write $z = x_1 + ix_2$ and $C = A_1 + iA_2$. Then $z(A) = C$. Also $z^2(A) = (x_1^2 + 2ix_1x_2 - x^2)(A) = A_1^2 + iA_1A_2 + iA_2A_1 - A_2^2 = C^2$. More generally, $z^k(A) = C^k$. If $f(z)$ is holomorphic in a neighborhood of a closed disc D containing $\sigma(A)$ in its interior, let $f(z) = a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \cdots$ where z_0 is the center of D. Then $f_n(z) = \sum_{i=0}^{n} a_i(z-z_0)^i \to f(z)$ uniformly in a neighborhood of D, as does each of its (partial) derivatives. Hence $a_0 I + a_1(C-z_0 I) + a_2(C-z_0 I)^2 + \cdots$ converges in norm to $f(A)$

Hence for such f we see that $f(A) = (2\pi i)^{-1}\int_{\partial D}f(\zeta)(\zeta I - C)^{-1}d\zeta$, which is the usual formula for $f(C)$.[3]

We move now to another rule of calculation. Let $Cx = Tx + x_0$ be an affine transformation of R^n, where T is a nonsingular linear transformation. $Cx = (\sum_j t_{1j}x_j + x_{0,1}, \cdots, \sum_j t_{nj}x_j + x_{0,n})$, where (t_{ij}) is the matrix of T, $x = (x_1, \cdots, x_n)$, and $x_0 = (x_{0,1}, \cdots, x_{0,n})$. Let $C(A) = (\sum_j t_{1j}A_j + x_{0,1}I, \cdots, \sum_j t_{nj}A_j + x_{0,n}I)$.

PROPOSITION 5. *Let $f \in C^\infty$ and let $g(x) = f(Cx)$. Then $g(A) = f(C(A))$.*

PROOF. If suffices to consider the case when $f \in C_0^\infty$. We have

$$g(A) = (2\pi)^{-n/2} \int \hat{g}(\lambda) \exp(-i\langle \lambda, A\rangle)d\lambda$$

$$= (2\pi)^{-n} \int\int g(x) \exp(i\langle \lambda, x\rangle) \exp(-i\langle \lambda, A\rangle)dxd\lambda$$

$$= (2\pi)^{-n} \int\int f(Tx + x_0) \exp(i\langle \lambda, x\rangle) \exp(-i\langle \lambda, A\rangle)dxd\lambda$$

$$= (2\pi)^{-n} \int\int f(y) \exp(i\langle \lambda, T^{-1}(y - x_0)\rangle) \exp(-i\langle \lambda, {}^{\cdot}A\rangle) \cdot \big| (\det T)^{-1}\big| \, dyd\lambda$$

$$= (2\pi)^{-n} \int\int f(y) \exp(i\langle T^{-1*}\lambda, y - x_0\rangle) \exp(-i\langle \lambda, {}^{\cdot}A\rangle) \cdot \big| (\det T)^{-1}\big| \, dyd\lambda$$

$$= (2\pi)^{-n} \int\int f(y) \exp(i\langle \gamma, y - x_0\rangle) \exp(-i\langle T^*\gamma, A\rangle) \cdot \big| (\det T)^{-1}(\det T^*)\big| \, dyd\gamma$$

$$= (2\pi)^{-n} \int\int f(y) \exp(i\langle \gamma, y\rangle) \exp(-i\langle \gamma, C(A)\rangle)dyd\gamma,$$

since multiples of I commute with all operators. Regard the above as an iterated integral, first with respect to the x variables and then with respect to the λ variables. Each such integral is absolutely convergent, and so the changes of variables made are justified. The last integral is equal to $(2\pi)^{-n/2}\int\hat{f}(\gamma)e^{-i\langle\gamma, C(A)\rangle}d\gamma = f(C(A))$.

We can apply this proposition to calculate $f(A)$ where f depends on only one variable.

PROPOSITION 6. *Let $f(x) = b(x_1) \in C^\infty$. Then $f(A) = b(A_1)$.*

---

[3] See [**4**, p. 422].

PROOF. Let $f_m(x) = f(x_1, x_2/m, \cdots, x_n/m)$. Then $f_m(x) = f(x) = b(x_1)$. By the previous proposition, $f_m(A) = f(A_1, A_2/m, \cdots, A_n/m)$. However, previous results show that $f_m(A) = f(A)$ and that $f(A_1, A_2/m, \cdots, A_n/m) \to f(A_1, 0, \cdots, 0)$ in norm as $m \to \infty$. Hence $f(A) = f(A_1, 0, \cdots, 0) = b(A_1)$.

We end this paper with a few results of a different nature.

PROPOSITION 7. *The adjoint of $f(A)$ is $\bar{f}(A)$.*

PROOF. Straightforward.

PROPOSITION 8. *If $f(0) = 0$ and the A_i all have finite dimensional range, then so does $f(A)$.*

PROOF. Let V be a subspace of H such that the range of each A_i is in V. For $x \in V^{\perp}$, $(A_iy, x) = 0$ for all $y \in H$. Hence $(y, A_ix) = 0$ for all $y \in H$, so $A_ix = 0$. Hence, looking at the power series, we see that $\exp(-i\langle \lambda, A \rangle)x = x$ for $x \in V^{\perp}$. Now

$$\bar{f}(A)x = (2\pi)^{-n/2} \int \bar{f}^{\wedge}(\lambda) \exp(-i\langle \lambda, A \rangle) x \, d\lambda$$

$$= x \, (2\pi)^{-n/2} \int \bar{f}^{\wedge}(\lambda) \exp(-i\langle \lambda, 0 \rangle) d\lambda = \bar{f}(0)x \quad \text{for } x \in V^{\perp}.$$

If $f(0) = 0$, then $x \in V^{\perp}$ implies $\bar{f}(A)x = 0$, so $(\bar{f}(A)x, y) = 0$ for all $y \in H$, and so $(x, f(A)y) = 0$ for all $y \in H$, i.e. the range of $f(A)$ is in $V^{\perp\perp} = V$.

Since an operator is compact if and only if it is the limit in norm of a sequence of operators with finite dimensional range, using the fact that $f(A)$ is a continuous function of A, we have the following.

COROLLARY. *If $f(0) = 0$ and all the A_i are compact, then so is $f(A)$.*

As a final application, let A be a single selfadjoint operator. The correspondence $A \to f(A)$ is a map $g: S \to L(H, H)$ from the real Banach space of bounded selfadjoint operators to the Banach space of all bounded operators on H. We assume $f \in C^{\infty}$, as a function of a real variable, and ask what is the derivative g_* of the map g.[4]

PROPOSITION 9. $g_*(A)A_0 = h(A, A_0)$ *where* $h(x, y) = yf'(x)$.

PROOF. A straightforward calculation will show that if $f(A) = A^n$, then $g_*(A)A_0 = A_0A^{n-1} + AA_0A^{n-2} + \cdots + A^{n-1}A_0 = u(A, A_0)$ where $u(x, y) = nyx^{n-1}$. Hence if $f(A) = e^{-iA\lambda}$, $g_*(A)A_0 = -i\lambda v_{\lambda}(A, A_0)$ where

[4] We refer here to the derivative defined in Dieudonné's *Foundations of modern analysis*, p. 143.

$v_\lambda(x, \quad y) = y e^{i\lambda x}.$[5] Let $f(A) = (2\pi)^{-1/2} \int \hat{f}(\lambda) e^{-i\lambda A} d\lambda$. Then $g_*(A) A_0$ $= (2\pi)^{-1/2} \int \hat{f}(\lambda)(-i\lambda) v_\lambda(A, \quad A_0) d\lambda$ (see also the second footnote) $= -\frac{1}{2} W(A, \quad A_0)$ where $W(x, y) = y(2\pi)^{-1/2} \int -i\lambda \hat{f}(\lambda) e^{-i\lambda x} d\lambda$, since the corresponding Riemann sums of these two expressions in A and A_0 are equal and the Riemann sums for the last integral, functions of x and y, converge locally uniformly, together with all the partial derivatives. But $W(x, y) = h(x, y)$.

Bibliography

1. S. Bochner and K. Chandrasekharan, *Fourier transforms*, Princeton Univ. Press, Princeton, N. J., 1949.

2. L. Hörmander, *Linear partial differential operators*, Springer-Verlag, New York, 1963.

3. E. Nelson, *Operator differential equations*, Mimeographed lecture notes, Princeton Univ., Princeton, N. J., 1964.

3a. ———, *Feynman integral and the Schrodinger equation*, J. Math. Phys. 5 (1964), 332–343.

4. F. Riesz and Bela Sz.-Nagy, *Functional analysis*, Ungar, New York, 1955.

5. L. Schwartz, *Théorie des distributions*, Hermann, Paris, 1957.

6. H. Weyl, *The theory of groups and quantum mechanics*, Dover, New York, 1931.

7. K. Yosida, *Functional analysis*, Springer-Verlag, New York, 1965.

PRINCETON UNIVERSITY

---

[5] See the theorem on p. 157 of the above mentioned text.