

SPECTRAL THEORY. I

CONVERGENCE TO PROJECTIONS

BY
NELSON DUNFORD

Spectral theory has for its origin the classical canonical reduction of a matrix, and it is this classical theory that provides the pattern for the developments presented in this paper. The finite dimensional linear vector space of the classical theory is here replaced by a complex Banach space \mathfrak{X} and the matrix in that theory by a continuous linear operator T in the space \mathfrak{X} . In this situation there are always three and sometimes four types of convergence, namely, the uniform, the strong, the weak, and, in case \mathfrak{X} consists of measurable functions, convergence almost everywhere. While our primary aim is to give necessary and sufficient conditions for convergence (in each of the four types) of a sequence $Q_n(T)$ of polynomials in T to a specified kind of limit operator, we have, and not merely incidentally but of necessity as well as for its own sake, developed a parallel to many of the algebraic aspects of classical matrix theory. In fact the chief tools in the convergence theory here presented are theorems corresponding to (a) the minimal equation theorem for matrices, (b) the theorem of Sylvester concerning the determinant of a polynomial in a matrix, and (c) the various decompositions of the whole space into a direct sum of its subspaces determined by the spectrum of a matrix. This analogue to classical matrix algebra plays its role almost exclusively in the questions involving uniform convergence (which of the four types seems to be the most difficult one to handle; especially in making the proofs of necessity rather than those of sufficiency) and the reader interested only in the other three types of convergence will find the last section fairly complete in itself as far as these questions are concerned.

The type of limit operator $E = \lim_n Q_n(T)$ with which we shall be chiefly concerned is a projection of the whole space \mathfrak{X} on the manifold $\mathfrak{M}[P]$ consisting of all $x \in \mathfrak{X}$ where, for a specified polynomial P ; we have $P(T)x = 0$. In seeking conditions for uniform convergence of a given sequence $Q_n(T)$ towards such a limit, one is faced with two fairly obvious necessary conditions, one of which is $P(T)Q_n(T) \rightarrow 0$. A third condition which will make the set of three both necessary and sufficient may be any one of three types, namely, (a) of *spectral type* and asserting that the roots of $P(\lambda) = 0$ be either in the resolvent set of T or else poles of the resolvent, or (b) of *algebraic type* and asserting that the manifold $\mathfrak{M}[P] = P(T)\mathfrak{X}$ is a closed complement of $\mathfrak{M}[P]$, or (c) of *topological type* and asserting that for every root λ of $P(\lambda) = 0$ of

Presented to the Society, April 24, 1943; received by the editors October 6, 1942, and, in revised form, March 26, 1943.

multiplicity ν we have $(\lambda I - T)^{\nu+1}\mathfrak{X}$ closed. One way of visualizing the difference between strong convergence and uniform is through the corresponding algebraic condition for strong convergence which is $\overline{\mathfrak{N}}[P] \oplus \mathfrak{M}[P] = \mathfrak{X}$.

The ergodic theorem falls within the scope of this theory, for in the ergodic theorem we are dealing with the case of a projection on the fixed points of an operator, that is $P(T) = I - T$. As the general situation shows, the conditions for convergence of the particular sequence $n^{-1}\sum_{m=0}^{n-1}T^m$ of polynomials are, if worded properly, equally valid for any sequence of polynomials $Q_n(T)$ with $Q_n(1) \rightarrow 1$, but for the particular sequence $n^{-1}\sum_{m=0}^{n-1}T^m$ the condition $P(T)Q_n(T) = (I - T)n^{-1}\sum_{m=0}^{n-1}T^m = (I - T^n)/n \rightarrow 0$ takes the relatively simple form of $T^n/n \rightarrow 0$. It is the condition $T^n/n \rightarrow 0$ and not the customary one of $|T^n| \leq M$ which is essential in the ergodic theory.

Since an element of a normed ring (in the sense of Gelfand) when considered as a linear operator on that ring has its norm as an operator the same as its norm in the ring, the purely algebraic results of §2 as well as those results in §3 concerning uniform convergence may be thought of as results concerning an element T of a normed ring.

The indigenous connection between the spectrum of an operator T and questions of convergence of sequences of functions $Q_n(T)$ of T is most readily seen (but not completely comprehended) by an examination of the spectrum of a finite matrix. For this reason and also because the finite case furnishes some insight into the puzzles of the infinite one we shall first present certain aspects of classical matrix theory in a form roughly as they shall appear in the general case.

1. Spectral theory in a finite dimensional space. We shall not discuss this case as completely as we shall the infinite one but enough will be given to indicate the general trend of ideas. In fact we shall merely derive a formula (equivalent to the classical canonical reduction) which will serve as a basis for an operational calculus and show how questions of convergence may be settled by spectral considerations. Except for the first two definitions and lemmas which are not concerned with the dimension or the topology in \mathfrak{X} (and which will be needed in the infinite case) we are concerned here with a linear vector space \mathfrak{X} of dimension $n > 0$ over the field of complex numbers (or any algebraically closed field) and a linear operator T (that is an $n \times n$ matrix) which maps \mathfrak{X} into all or part of itself.

1.1. DEFINITION. Let P be a polynomial, λ a complex number, and ν a positive integer or 0. We define

$$\begin{aligned}\mathfrak{N}[P] &= P(T)\mathfrak{X}, & \mathfrak{M}[P] &= \mathcal{E}_{x \in \mathfrak{X}} [P(T)x = 0], \\ \mathfrak{N}'_{\lambda} &= (\lambda I - T)^{\nu}\mathfrak{X}, & \mathfrak{M}'_{\lambda} &= \mathcal{E}_{x \in \mathfrak{X}} [(\lambda I - T)^{\nu}x = 0],\end{aligned}$$

where I is the identity operator in \mathfrak{X} .

1.2. DEFINITION. If $\mathfrak{M}, \mathfrak{M}_1, \dots, \mathfrak{M}_k$ are linear manifolds in \mathfrak{X} we shall write $\mathfrak{M} = \mathfrak{M}_1 \oplus \dots \oplus \mathfrak{M}_k$ if and only if $\mathfrak{M}_i \subset \mathfrak{M}, i=1, \dots, k$, and every $x \in \mathfrak{M}$ has a unique representation $x = x_1 + \dots + x_k$ with $x_i \in \mathfrak{M}_i, i=1, \dots, k$. We shall write $\mathfrak{M} = \mathfrak{M}_1 + \dots + \mathfrak{M}_k$ if and only if $\mathfrak{M}_i \subset \mathfrak{M}, i=1, \dots, k$, and every $x \in \mathfrak{M}$ can be represented as $x = x_1 + \dots + x_k$ with $x_i \in \mathfrak{M}_i, i=1, \dots, k$.

1.3. LEMMA. If P, Q are polynomials with no common roots then

$$\mathfrak{M}[P \cdot Q] = \mathfrak{M}[P] \oplus \mathfrak{M}[Q].$$

Clearly $\mathfrak{M}[P], \mathfrak{M}[Q] \subset \mathfrak{M}[P \cdot Q]$. Since P and Q have no common factors there are polynomials R, S with

$$R(\lambda)P(\lambda) + S(\lambda)Q(\lambda) \equiv 1,$$

$$R(T)P(T) + S(T)Q(T) = I,$$

$$R(T)P^2(T) + S(T)P(T)Q(T) = P(T).$$

Thus if $P(T)Q(T)x=0$, that is, if $x \in \mathfrak{M}[P \cdot Q]$, then $P(T)(x-y)=0$ where $y=R(T)P(T)x \in \mathfrak{M}[Q]$. This decomposition is unique, for the second of the above identities shows that $u=0$ if $u \in \mathfrak{M}[P] \cdot \mathfrak{M}[Q]$.

1.4. LEMMA. Let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{r_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Let $P_i(\lambda) = P(\lambda)/(\lambda_i - \lambda)^{r_i}$. Then

$$\mathfrak{M}[P] = \mathfrak{M}_{\lambda_1}^{r_1} \oplus \dots \oplus \mathfrak{M}_{\lambda_k}^{r_k}$$

and

$$\mathfrak{X} = \mathfrak{N}[P_1] + \dots + \mathfrak{N}[P_k].$$

The first conclusion is proved by a repeated application of Lemma 1.1 and the second by observing that since $P_1(\lambda), \dots, P_k(\lambda)$ have no common factor there are polynomials R_1, \dots, R_k with

$$R_1(\lambda)P_1(\lambda) + \dots + R_k(\lambda)P_k(\lambda) \equiv 1,$$

$$R_1(T)P_1(T) + \dots + R_k(T)P_k(T) = I.$$

1.5. DEFINITION. The index ν of a complex number λ is the smallest positive integer or zero such that $(T - \lambda I)^{\nu+1}x = 0$ implies $(T - \lambda I)^{\nu}x = 0$.

Thus the index ν of λ is positive if and only if λ is a characteristic number of the matrix T . In what follows we use the symbols $\lambda_1, \dots, \lambda_k$ for the distinct characteristic numbers of T and ν_1, \dots, ν_k for their indices.

1.6. LEMMA. If $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{r_i}$ then $P(T) = 0$ and hence

$$\mathfrak{X} = \mathfrak{M}_{\lambda_1}^{\nu_1} \oplus \dots \oplus \mathfrak{M}_{\lambda_k}^{\nu_k}.$$

To see that $P(T) = 0$ let x be an arbitrary vector in \mathfrak{X} . Since the dimension

of \mathfrak{X} is n there are constants $\alpha_0, \dots, \alpha_n$ with $\alpha_0 x + \alpha_1 T x + \dots + \alpha_n T^n x = 0$, that is, $Q(T)x = 0$ for some polynomial Q . If Q is written in factored form then $\prod (\mu_i I - T)^{m_i} x = 0$ and it is clear that this equation remains valid if all factors with μ_i not a characteristic number are removed, and from the definition of ν_i we see that the other factors may be changed to have the desired exponents. Thus $P(T)x = 0$, $P(T) = 0$. The second conclusion follows from Lemma 1.4.

1.7. DEFINITION. Let E_{λ_i} , $i = 1, \dots, k$, be the projections which, in view of Lemma 1.6, exist and satisfy $I = E_{\lambda_1} + \dots + E_{\lambda_k}$, $E_{\lambda_i}^2 = E_{\lambda_i}$, $E_{\lambda_i} \mathfrak{X} = \mathfrak{M}_{\lambda_i}^{\nu_i}$, $E_{\lambda_i} E_{\lambda_j} = 0$, $i \neq j$.

Note that since $T \mathfrak{M}_{\lambda_i}^{\nu_i} \subset \mathfrak{M}_{\lambda_i}^{\nu_i}$ we have $E_{\lambda_i} T x = E_{\lambda_i} \sum_j T E_{\lambda_j} x = T E_{\lambda_i} x$, that is $T E_{\lambda_i} = E_{\lambda_i} T$.

1.8. DEFINITION. Let $\mathcal{F}(T)$ be the class of all complex functions of a complex variable which are regular at every point of an open set (not necessarily connected) containing the characteristic numbers $\lambda_1, \dots, \lambda_k$ of T . (The open set upon which $f \in \mathcal{F}(T)$ is regular may vary with f .) For $f \in \mathcal{F}(T)$ we define

$$(*) \quad f(T) = \sum_{i=1}^k \sum_{m=0}^{\nu_i-1} \frac{(T - \lambda_i I)^m}{m!} f^{(m)}(\lambda_i) E_{\lambda_i}.$$

1.9. THEOREM. If $f, g \in \mathcal{F}(T)$ and α, β are complex numbers then

- (a) $\alpha f + \beta g \in \mathcal{F}(T)$ and $(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T)$,
- (b) $f \cdot g \in \mathcal{F}(T)$ and $(f \cdot g)(T) = f(T) \cdot g(T)$,
- (c) if $f(\lambda) = \sum_{n=0}^m \alpha_n \lambda^n$ then $f(T) = \sum_{n=0}^m \alpha_n T^n$.

Statements (a) and the first part of (b) are obvious. To prove that $(f \cdot g)(T) = f(T) \cdot g(T)$ let us write, for $h \in \mathcal{F}(T)$,

$$D_{\lambda_i}(h, T) = \sum_{m=0}^{\nu_i-1} \frac{(T - \lambda_i I)^m}{m!} h^{(m)}(\lambda_i).$$

Then since $(T - \lambda_i I)^q E_{\lambda_i} = 0$ for $q \geq \nu_i$ we have

$$\begin{aligned} D_{\lambda_i}(f, T) D_{\lambda_i}(g, T) E_{\lambda_i} &= \sum_{p=0}^{2\nu_i-2} \sum_{n=0}^p \frac{(T - \lambda_i I)^p}{n!(p-n)!} f^{(n)}(\lambda_i) g^{(p-n)}(\lambda_i) E_{\lambda_i} \\ &= \sum_{p=0}^{2\nu_i-2} \frac{(T - \lambda_i I)^p}{p!} \left[\sum_{n=0}^p {}_p C_n f^{(n)}(\lambda_i) \cdot g^{(p-n)}(\lambda_i) \right] E_{\lambda_i} \\ &= \sum_{p=0}^{\nu_i-1} \frac{(T - \lambda_i I)^p}{p!} (f \cdot g)^{(p)}(\lambda_i) E_{\lambda_i} \\ &= D_{\lambda_i}(f \cdot g, T). \end{aligned}$$

Hence

$$\begin{aligned}
 f(T) \cdot g(T) &= \left[\sum_{i=1}^k D_{\lambda_i}(f, T) E_{\lambda_i} \right] \cdot \left[\sum_{i=1}^k D_{\lambda_i}(g, T) E_{\lambda_i} \right] \\
 &= \sum_{i=1}^k D_{\lambda_i}(f, T) D_{\lambda_i}(g, T) E_{\lambda_i} \\
 &= \sum_{i=1}^k D_{\lambda_i}(f \cdot g, T) E_{\lambda_i} = (f \cdot g)(T),
 \end{aligned}$$

which proves (b). In view of (a) and (b) statement (c) need only be verified for the special cases of $f(\lambda)=1$, $f(\lambda)=\lambda$ and (c) for these polynomials is an immediate consequence of Definition 1.7.

1.10. THEOREM. *If $f, g \in \mathcal{F}(T)$ then $f(T)=g(T)$ if and only if $f^{(n)}(\lambda_i)=g^{(n)}(\lambda_i)$, $i=1, \dots, k$; $n=0, \dots, \nu_i-1$. Thus for a polynomial P we have $P(T)=0$ if and only if $P(\lambda)$ contains $\prod_{i=1}^k (\lambda_i - \lambda)^{\nu_i}$ as a factor.*

The sufficiency of the condition is obvious from (*) and its necessity follows from the definition of ν_i as follows. In view of part (a) of Theorem 1.9 we may and shall assume that $f(T)=0$. Let $x \neq 0$, $(T - \lambda_i I)x = 0$ so that by (*) and 1.7 we have $0=f(T)x=f(\lambda_i)x$, $f(\lambda_i)=0$. Next, in case $\nu_i > 1$, pick x so that $(T - \lambda_i I)^2 x = 0$, $(T - \lambda_i I)x \neq 0$ then as before $0=f(T)x=f(\lambda_i)x + (T - \lambda_i I)f^{(1)}(\lambda_i)x = f^{(1)}(\lambda_i)(T - \lambda_i I)x$, and so $f^{(1)}(\lambda_i)=0$. A repetition of this argument clearly yields $f^{(n)}(\lambda_i)=0$, $n=0, 1, \dots, \nu_i-1$.

1.11. THEOREM. *Let $f_n \in \mathcal{F}(T)$, then $f_n(T)$ converges if and only if $f_n^{(j)}(\lambda_i)$ converges for each $i=1, \dots, k$, and $j=0, \dots, \nu_i-1$.*

The sufficiency of the conditions is clear from (*) while their necessity follows from the argument used in the proof of the preceding theorem.

The next theorem is proved in the same way.

1.12. THEOREM. *Let $f, f_n \in \mathcal{F}(T)$ then $f_n(T) \rightarrow f(T)$ if and only if $\lim f_n^{(j)}(\lambda_i) = f^{(j)}(\lambda_i)$, $i=1, \dots, k$; $j=0, \dots, \nu_i-1$.*

1.13. *An illustration.* If $f_n(\lambda) = n^{-1} \sum_{m=0}^{n-1} \lambda^m$ we have $f_n(\lambda)$ converging if and only if $|\lambda| \leq 1$, while $f^{(j)}(\lambda)$ converges for all $j=1, 2, \dots$ providing $|\lambda| < 1$ but the first derivative $f^{(1)}(\lambda)$ fails to converge if $|\lambda|=1$. Hence (1.11) shows that $n^{-1} \sum_{m=0}^{n-1} T^m$ converges if and only if all roots λ_i of T have $|\lambda_i| \leq 1$ and if in addition $\nu_i=1$ for every root λ_i with $|\lambda_i|=1$. Next consider the statement $T^n/n \rightarrow 0$. An application of (1.12) to the functions $f_n(\lambda) = \lambda^n/n$ shows us that $T^n/n \rightarrow 0$ if and only if the roots λ_i of T satisfy the same conditions required in the preceding example. Hence $n^{-1} \sum_{m=0}^{n-1} T^m$ converges if and only if $T^n/n \rightarrow 0$. It is not pure coincidence that these two statements concerning convergence have precisely the same spectral interpretation. It is a consequence of the fact that the statement $T^n/n \rightarrow 0$ is equivalent to $(I - T)n^{-1} \sum_{m=0}^{n-1} T^m$

$= (I - T^n)/n \rightarrow 0$ and is a special case of a general principle which will be developed more completely in §3.

1.14. *Concluding remarks.* The formula

$$(i) \quad f(T) = \sum_{i=1}^k \sum_{m=0}^{\nu_i-1} \frac{(T - \lambda_i I)^m}{m!} f^{(m)}(\lambda_i) E_{\lambda_i}$$

of 1.8 becomes

$$(ii) \quad f(T) = \sum_{i=1}^k f(\lambda_i) E_{\lambda_i}$$

for a matrix T whose characteristic numbers λ_i all have index 1. In particular (ii) holds if T is an Hermitian matrix. This case has been generalized in an extensive literature to the concept of a self-adjoint operator in Hilbert space where (ii) takes the well known form of $f(T) = \int_{-\infty}^{+\infty} f(\lambda) dE_\lambda$, and in recent years has been given, by M. H. Stone⁽¹⁾ and I. Gelfand⁽²⁾, an algebraic topological form. One difference between the ring of polynomials in a general matrix T and the ring of polynomials in an Hermitian matrix T is that (as formulas (i) and (ii) show) the former ring is isomorphic with a direct sum of reduced polynomial rings while the latter is isomorphic with a direct sum of identical fields, that is, isomorphic with a linear space of scalar functions.

If the projections E_{λ_i} in (i) are eliminated from (i) by using the Hermite interpolation formula then (i) takes the very natural form

$$(iii) \quad f(T) = \frac{1}{2\pi i} \int_C \frac{f(\lambda)}{\lambda I - T} d\lambda,$$

where C is the boundary of a set of small circles containing $\lambda_1, \dots, \lambda_k$. If C_i is a small circle about λ_i then (iii) gives

$$(iv) \quad E_{\lambda_i} = \frac{1}{2\pi i} \int_{C_i} \frac{d\lambda}{\lambda I - T}.$$

Formula (iii) was suggested by E. Cartan⁽³⁾ as a possible basis for an operational calculus for infinite matrices and was used (only in an incidental way) by Gelfand. Formulas (iii) and (iv) were discovered by Fantappiè⁽⁴⁾ who was working with finite matrices, and recently (iv) has been extended to the infinite case by Lorch⁽⁵⁾, who proved that an operator E defined by the equa-

⁽¹⁾ M. H. Stone, *A general theory of spectra*. I, Proc. Nat. Acad. Sci. U. S. A. vol. 26 (1940) pp. 280-283.

⁽²⁾ I. Gelfand, *Normierte Ringe*, Rec. Math. (Mat. Sbornik) N. S. vol. 51 (1941) pp. 3-24.

⁽³⁾ The suggestion was made in a letter to G. Giorgi. For this reference as well as for others pertaining to an operational calculus see C. C. MacDuffee, *The theory of matrices*, Berlin, 1933.

⁽⁴⁾ L. Fantappiè, *La calcul des matrices*, C. R. Acad. Sci. Paris vol. 186 (1928) pp. 619-621.

⁽⁵⁾ E. R. Lorch, *The spectrum of linear transformations*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 238-248.

tion $E = (2\pi i)^{-1} \int_C (\lambda I - T)^{-1} d\lambda$ is a projection, and that the algebra of such projections is isomorphic with the Boolean algebra of all open and closed subsets of the spectrum of T .

2. An operational calculus and some algebraic properties of linear operators. The procedure of §2 used to discuss the general linear operator T differs only slightly from that of §1. We first define the class $\mathcal{F}(T)$ and the operator $f(T)$ for $f \in \mathcal{F}(T)$ and then define the projections corresponding to the E_{λ_i} of §1. The fundamental algebraic theorems (Theorems 2.12 through 2.29) may then be formulated in terms of these concepts.

2.1. Preliminary concepts and definitions. From now on the symbol T will stand for a bounded linear operator on a complex Banach space \mathfrak{X} . By the *resolvent set* $\rho(T)$ will be meant the set of all complex numbers λ such that $(\lambda I - T)^{-1}$ exists as an everywhere defined (and hence continuous) linear operator. For $\lambda \in \rho(T)$ we shall write $R_\lambda(T)$ for $(\lambda I - T)^{-1}$. The set $\rho(T)$ is an open set containing all λ with $|\lambda| > |T|$. By the *spectrum* $\sigma(T)$ of T is meant all complex numbers $\lambda \notin \rho(T)$. The spectrum $\sigma(T)$ is a closed point set contained in the circle $|\lambda| \leq |T|$. If the space \mathfrak{X} contains more than one point (and this we assume) the spectrum $\sigma(T)$ is not empty (see Taylor⁽⁶⁾) and we write $|\sigma(T)|$ for l.u.b. $|\lambda|$ where λ varies over $\sigma(T)$. The *point spectrum* $\sigma_p(T)$ consists of all λ such that $(\lambda I - T)^{-1}$ does not exist, the *continuous spectrum* $\sigma_c(T)$ of those λ for which $(\lambda I - T)^{-1}$ exists as an unbounded operator with domain dense in \mathfrak{X} , and the *residual spectrum* $\sigma_r(T)$ of those λ for which $(\lambda I - T)^{-1}$ exists with domain nondense in \mathfrak{X} . It is readily shown that if $(\lambda I - T)^{-1}$ exists as a bounded operator with domain dense in \mathfrak{X} then $\lambda \in \rho(T)$, so that the point sets $\rho(T)$, $\sigma_p(T)$, $\sigma_c(T)$, $\sigma_r(T)$ are disjoint and their union is the whole complex plane. The function $R_\lambda(T)$ satisfies the *resolvent equation* $R_\lambda(T) - R_\mu(T) = (\mu - \lambda)R_\mu(T)R_\lambda(T)$, $\lambda, \mu \in \rho(T)$, and is *regular*, that is, $dR_\lambda(T)/d\lambda$ exists, at every point $\lambda \in \rho(T)$. If $f(\lambda)$ is any vector (or operator) valued function, regular and single-valued at every point of a closed domain D whose boundary C is a finite number of rectifiable Jordan curves then (1) $\int_C f(\lambda) d\lambda = 0$, (2) $f(\lambda) = (2\pi i)^{-1} \int_C f(\xi)(\xi - \lambda)^{-1} d\xi$, for λ in the interior of D , (3) $f(\lambda)$ has derivatives of all orders at every point of D , (4) $f^{(n)}(\lambda) = (2\pi i)^{-1} n! \int_C f(\xi)(\xi - \lambda)^{-n-1} d\xi$, for λ in the interior of D , (5) the Taylor expansion $\sum_{n=0}^{\infty} (\lambda - \xi)^n f^{(n)}(\xi)/n!$ converges uniformly to $f(\lambda)$ for λ in any circle $|\lambda - \xi| \leq r$ contained in the interior of D . If $f(\lambda)$ is regular and single-valued at each point of an annulus enclosed by two concentric circles C_1 and C_2 with center λ_0 and if C is any circle between C_1 and C_2 , then for every λ in this annulus

$$f(\lambda) = \sum_{n=-\infty}^{+\infty} a_n (\lambda - \lambda_0)^n, \quad a_n = \frac{1}{2\pi i} \int_C \frac{f(\lambda) d\lambda}{(\lambda - \lambda_0)^{n+1}}, \quad n = 0, \pm 1, \pm 2, \dots$$

⁽⁶⁾ A. E. Taylor, *The resolvent of a closed transformation*, Bull. Amer. Math. Soc. vol. 44 (1938) pp. 70-74.

The above statements concerning regular functions may be proved as follows. For a linear functional over \mathfrak{X} , that is a point $x^* \in \mathfrak{X}^*$, we have $x^* \int_C f(\lambda) d\lambda = \int_C x^* f(\lambda) d\lambda$ and thus by Cauchy's theorem $x^* \int_C f(\lambda) d\lambda = 0$. Hence by the Hahn-Banach extension theorem (which was proved for complex linear functionals over a complex Banach space by Bohnenblust and Sobczyk⁽⁷⁾) we have $\int_C f(\lambda) d\lambda = 0$. This proves (1) and (2) is proved in the same way while the rest of the statements follow in the usual manner.

Many standard theorems in the theory of linear operators on a Banach space are readily seen to hold in a complex Banach space and these will be used when needed. In particular the various corollaries of the Hahn-Banach extension theorem including the *principle of uniform boundedness* will be used. Also the *interior mapping theorem* or, more specifically, its corollary which asserts that *a closed everywhere defined linear operator is continuous* will be used. This shows that the projections $E, E' = I - E$ determined by *closed* complementary linear manifolds in \mathfrak{X} are continuous.

AN OPERATIONAL CALCULUS

2.2. DEFINITION. By a T -admissible domain is meant an open set D of complex numbers having the following properties:

- (1) D is a finite sum of connected open sets D_i with $\overline{D_i} \overline{D_j} = 0, i \neq j$.
- (2) The boundary C of D consists of a finite number of disjoint closed rectifiable Jordan curves contained in the resolvent set $\rho(T)$ of T . We shall sometimes write $D = D(T)$ to mean that D is a T -admissible domain.

2.3. DEFINITION. The class $\mathcal{F}(T)$ is the class of all complex functions each of which is regular and single-valued at each point of the closure of some $D = D(T)$ which contains the spectrum $\sigma(T)$ of T .

2.4. LEMMA. Let D, D^* be T -admissible domains with $\sigma(T) \subset D^* \subset \overline{D^*} \subset D$. Let $\phi(\lambda)$ be regular and single-valued on \overline{D} . Then

$$\int_C \phi(\lambda) R_\lambda(T) d\lambda = \int_{C^*} \phi(\lambda) R_\lambda(T) d\lambda,$$

where C, C^* are the boundaries of D, D^* respectively.

The domain D is made up of a finite sum of disjoint sets and taking the integral around C is the same as taking the integral around each connected component and adding. Thus, without loss of generality, we may and shall confine our attention to the case where D is connected. Suppose first that D is simply connected and that D^* is connected. Since $\sigma(T) \subset D^*$ none of the inside boundaries of D^* can surround a part of $\sigma(T)$ and hence the integral around every inside boundary of D^* vanishes. Thus the integral \int_{C^*} reduces

⁽⁷⁾ H. F. Bohnenblust and A. Sobczyk, *Extensions of functionals on complex linear spaces*, Bull. Amer. Math. Soc. vol. 44 (1938) pp. 91-93.

to the integral around the outside boundary of D^* which may be joined by a cut to the boundary of D to prove the lemma in this case. Next suppose only that D is simply connected. This case reduces to the first case by making cuts joining up the various parts of D^* . Now suppose the lemma has been proved for the case where D is $(n-1)$ -tuply connected. We now take D to be n -tuply connected. If at least one of the inside boundaries of D lies outside of all the outside boundaries of the various parts of D^* it may be joined by a cut to the outside boundary of D , thus reducing D to a $(n-1)$ -tuply connected domain. If every inside boundary C_i of D lies in one of the holes of some part of D^* and is thus surrounded by some uniquely determined nearest part C_i^* of the boundary C^* then $\int_{C_i} = \int_{C_i^*}$. Also any inside boundary C_i^* of D^* which does not surround such a C_i nor any portion of D^* has the property that $\int_{C_i^*} = 0$ and this type of C_i^* may be dropped. These observations reduce the problem to the case where (i) every inside boundary C_i of C is contained in at least one of the holes in some part of D^* and (ii) every inside boundary of D^* surrounds either (α) an inside boundary of D or (β) one or more parts of D^* . Here parts (α) and (β) are not mutually exclusive. Now consider an inside boundary C_i^* of D^* which comes under (ii)(β). The various portions of D^* within C_i^* may be joined by cuts and one of them to C_i^* . This process reduces the problem to the case where every inside C_i is in one and only one inside C_i^* and every inside C_i^* contains one C_i . Now it is clear that the integral around the remaining (uncut) outside boundaries of D^* is equal to the integral around the outside boundary of D and the integral around such inside boundary of D^* is equal to the integral around the corresponding boundary of D and thus $\int_C = \int_{C^*}$.

2.5. LEMMA. *Let D, D^* be T -admissible domains with $\sigma(T) \subset DD^*$. Then there is a T -admissible domain D_1 with $\sigma(T) \subset D_1 \subset DD^*$. If ϕ is regular and single-valued at each point of \bar{D} and at each point of \bar{D}^* then*

$$\int_C \phi(\lambda) R_\lambda(T) d\lambda = \int_{C^*} \phi(\lambda) R_\lambda(T) d\lambda,$$

where C, C^* are the boundaries of D, D^* respectively.

The first part of this lemma follows from the Heine-Borel Theorem and the second part then follows from the preceding lemma.

2.6. DEFINITION. *Let $f \in \mathcal{F}(T)$, $D = D(T)$, $D \supset \sigma(T)$ and let f be regular and single-valued on \bar{D} . Let C be the boundary of D . We define*

$$f(T) = \frac{1}{2\pi i} \int_C f(\lambda) R_\lambda(T) d\lambda.$$

The preceding lemma shows that $f(T)$ is independent of D .

2.7. THEOREM. If $f, g \in \mathcal{F}(T)$ then

(1) $f+g \in \mathcal{F}(T)$ and $f(T)+g(T)=(f+g)(T)$,

(2) $f \cdot g \in \mathcal{F}(T)$ and $f(T) \cdot g(T)=(f \cdot g)(T)$.

Statement (1) as well as the first part of (2) is obvious. To see that $f(T) \cdot g(T) = (f \cdot g)(T)$ let C be the boundary of a domain $D = D(T)$ such that $\sigma(T) \subset D$ and such that f and g are both regular and single-valued on \bar{D} . Let C_j be the boundaries of the component parts of D . Then

$$\begin{aligned} f(T) &= \frac{1}{2\pi i} \sum_{j=1}^k \int_{C_j} f(\lambda) R_\lambda(T) d\lambda, & g(T) &= \frac{1}{2\pi i} \sum_{j=1}^k \int_{C_j} g(\lambda) R_\lambda(T) d\lambda, \\ (*) \quad -4\pi^2 f(T) \cdot g(T) &= \sum_{n=1}^k \sum_{m=1}^k \left(\int_{C_n} f(\lambda) R_\lambda(T) d\lambda \right) \left(\int_{C_m} g(\lambda) R_\lambda(T) d\lambda \right). \end{aligned}$$

Now

$$\begin{aligned} \left(\int_{C_n} f(\lambda) R_\lambda(T) d\lambda \right) \left(\int_{C_m} g(\mu) R_\mu(T) d\mu \right) &= \int_{C_n} \int_{C_m} f(\lambda) g(\mu) R_\lambda(T) R_\mu(T) d\lambda d\mu \\ &= - \int_{C_n} \int_{C_m} f(\lambda) g(\mu) \frac{R_\lambda(T) - R_\mu(T)}{\lambda - \mu} d\lambda d\mu \\ &= - \int_{C_m} g(\mu) d\mu \left[\int_{C_n} \frac{f(\lambda) R_\lambda(T)}{\lambda - \mu} d\lambda - R_\mu(T) \int_{C_n} \frac{f(\lambda)}{\lambda - \mu} d\lambda \right]. \end{aligned}$$

Hence if $n \neq m$ and $\mu \in C_m$ we have $\int_{C_n} f(\lambda) (\lambda - \mu)^{-1} d\lambda = 0$ and

$$\begin{aligned} \left(\int_{C_n} f(\lambda) R_\lambda(T) d\lambda \right) \left(\int_{C_m} g(\mu) R_\mu(T) d\mu \right) &= - \int_{C_m} g(\mu) d\mu \int_{C_n} \frac{f(\lambda) R_\lambda(T)}{\lambda - \mu} d\lambda \\ &= - \int_{C_n} f(\lambda) R_\lambda(T) d\lambda \int_{C_m} \frac{g(\mu)}{\lambda - \mu} d\mu = 0, \end{aligned}$$

and thus (*) becomes

$$\begin{aligned} (*) \quad -4\pi^2 f(T) \cdot g(T) &= \sum_{n=1}^k \int_{C_n} f(\lambda) R_\lambda(T) d\lambda \int_{C_n} g(\mu) R_\mu(T) d\mu \\ &= - \sum_{n=1}^k \int_{C_n} \int_{C_n} f(\lambda) g(\mu) \frac{R_\lambda(T) - R_\mu(T)}{\lambda - \mu} d\lambda d\mu. \end{aligned}$$

Now using each point of C_n as a center construct a circle of radius so small that f, g are both regular in the closed circle and also such that the radius is less than half the distance from the point on C_n to the remaining contours

bounding D . A finite number of these circles cover the boundary C_n of the component D_n of D . Let \mathcal{E}_n be the sum of the interiors of this finite number of circles and let $D_n^* = \mathcal{E}_n + D_n$. Let C_n^* be the boundary of D_n^* . Clearly we have $\sigma(T)D_n \subset D_n^*$, $\bar{D}_n \subset D_n^*$. Then

$$\begin{aligned} \int_{C_n} f(\lambda) \frac{R_\lambda(T) - R_\mu(T)}{\lambda - \mu} d\lambda &= \int_{C_n^*} f(\lambda) \frac{R_\lambda(T) - R_\mu(T)}{\lambda - \mu} d\lambda \\ &= \int_{C_n^*} \frac{f(\lambda) R_\lambda(T) d\lambda}{\lambda - \mu} - 2\pi i f(\mu) R_\mu(T) \end{aligned}$$

and (**) may be written as

$$\begin{aligned} 4\pi^2 f(T) \cdot g(T) &= \sum_{n=1}^k \left[2\pi i \int_{C_n} f(\mu) g(\mu) R_\mu(T) d\mu \right. \\ &\quad \left. - \int_{C_n^*} f(\lambda) R_\lambda(T) \left(\int_{C_n} \frac{g(\mu) d\mu}{\lambda - \mu} \right) d\lambda \right]. \end{aligned}$$

For $\lambda \in C_n^*$ we have $\int_{C_n} g(\mu) (\lambda - \mu)^{-1} d\mu = 0$ and so

$$\begin{aligned} f(T) \cdot g(T) &= \frac{1}{2\pi i} \sum_{n=1}^k \int_{C_n} f(\mu) g(\mu) R_\mu(T) d\mu \\ &= \frac{1}{2\pi i} \int_C f(\mu) g(\mu) R_\mu(T) d\mu = (f \cdot g)(T). \end{aligned}$$

2.8. THEOREM. *If $f(\lambda)$ is a polynomial in λ or, more generally, if $f(\lambda) = \sum_{n=0}^{\infty} \alpha_n \lambda^n$ is a series convergent in a circle $|\lambda| \leq \rho$ with $\rho > |T|$ then $f(T) = \sum_{n=0}^{\infty} \alpha_n T^n$.*

Let C be the circle $|\lambda| = \rho$ so that

$$\begin{aligned} f(T) &= \frac{1}{2\pi i} \int_C f(\lambda) R_\lambda(T) d\lambda = \frac{1}{2\pi i} \int_C f(\lambda) \sum_{n=0}^{\infty} T^n \lambda^{-n-1} d\lambda \\ &= \sum_{n=0}^{\infty} T^n \frac{1}{2\pi i} \int_C \frac{f(\lambda)}{\lambda^{n+1}} d\lambda = \sum_{n=0}^{\infty} \alpha_n T^n. \end{aligned}$$

2.9. THEOREM. *If $f \in \mathcal{F}(T)$ then $f(\sigma(T)) = \sigma(f(T))$.*

Let $\lambda \in \sigma(T)$ and let $g(\xi)$ be defined for ξ in the domain of definition of f by the formula $g(\xi) = (f(\lambda) - f(\xi))(\lambda - \xi)^{-1}$ so that $g \in \mathcal{F}(T)$ and, by 2.7 and 2.8, we have $f(\lambda)I - f(T) = (\lambda I - T)g(T)$ which shows that $f(\lambda) \in \sigma(f(T))$. Conversely let $\mu \in \sigma(f(T))$ and suppose that $\mu \notin f(\sigma(T))$. Then the function $h(\xi) = (f(\xi) - \mu)^{-1}$ belongs to $\mathcal{F}(T)$ and, by 2.7 and 2.8, $h(T)(f(T) - \mu I) = I$, which contradicts the fact that $\mu \in \sigma(f(T))$.

2.10. THEOREM. Let $f \in \mathcal{F}(T)$, $g \in \mathcal{F}(f(T))$ and let $F(\xi) = g(f(\xi))$. Then $F \in \mathcal{F}(T)$ and $F(T) = g(f(T))$.

Suppose g is regular and single-valued on \bar{D}^* where $\sigma(f(T)) = f(\sigma(T)) \subset D^*$ and let C^* be the boundary of D^* . Pick $D = D(T)$ such that $\sigma(T) \subset D$, f is regular and single-valued on \bar{D} and $f(\bar{D}) \subset D^*$. Then if C is the boundary of D we have from 2.7

$$R_\lambda(f(T)) = \frac{1}{2\pi i} \int_C \frac{R_\xi(T)}{\lambda - f(\xi)} d\xi, \quad \lambda \in C^*,$$

and

$$\begin{aligned} g(f(T)) &= \frac{1}{2\pi i} \int_{C^*} g(\lambda) R_\lambda(f(T)) d\lambda \\ &= \left(\frac{1}{2\pi i} \right)^2 \int_C \int_{C^*} \frac{g(\lambda) d\lambda}{\lambda - f(\xi)} R_\xi(T) d\xi \\ &= \frac{1}{2\pi i} \int_C g(f(\xi)) R_\xi(T) d\xi = F(T). \end{aligned}$$

The elementary algebraic rules of operation given by 2.7, \dots , 2.10 inclusive will be used from this point on without explicit reference to the theorem in question.

2.11. DEFINITION. By a spectral set of T will be meant any subset σ of $\sigma(T)$ which is both open and closed in $\sigma(T)$. If σ is a spectral set of T the symbol σ' will be used for the complement of σ in $\sigma(T)$ so that σ' is also a spectral set of T .

It is readily seen that a set σ of complex numbers is a spectral set of T if and only if it is of the form $\sigma = \sigma(T)D$ where $D = D(T)$, and that for any spectral set σ of T there are T -admissible domains D, D' with

$$\sigma = \sigma(T)D, \quad \sigma' = \sigma(T)D', \quad \overline{DD'} = 0.$$

Such domains are said to be *complementary domains* and D' will be called a *complement* of D . If C, C' are the boundaries of D, D' respectively and we define

$$E_\sigma[T] = \frac{1}{2\pi i} \int_C R_\lambda(T) d\lambda, \quad E_{\sigma'}[T] = \frac{1}{2\pi i} \int_{C'} R_\lambda(T) d\lambda,$$

then it is seen from 2.7 and 2.8 that $E_\sigma[T], E_{\sigma'}[T]$ are projections reducing T and that

$$I = E_\sigma[T] + E_{\sigma'}[T], \quad E_\sigma[T] \cdot E_{\sigma'}[T] = 0.$$

With every spectral set σ of T is associated the closed linear manifold $\mathfrak{X}_\sigma[T] \subset \mathfrak{X}$ and defined as $\mathfrak{X}_\sigma[T] = E_\sigma[T]\mathfrak{X}$. Thus if $\sigma_1, \dots, \sigma_k$ are disjoint spectral sets of T and $\sigma = \sigma_1 + \dots + \sigma_k$ we have

$$E_\sigma[T] = E_{\sigma_1}[T] + \cdots + E_{\sigma_k}[T], \quad E_{\sigma_i}[T] \cdot E_{\sigma_j}[T] = 0, \quad i \neq j, \\ \mathfrak{X}_\sigma[T] = \mathfrak{X}_{\sigma_1}[T] \oplus \cdots \oplus \mathfrak{X}_{\sigma_k}[T].$$

When there is no chance of confusion we shall sometimes write E_σ and \mathfrak{X}_σ in place of $E_\sigma[T]$ and $\mathfrak{X}_\sigma[T]$ respectively. If the spectral set σ reduces to a single point λ we shall write E_λ , \mathfrak{X}_λ for E_σ , \mathfrak{X}_σ respectively.

2.12. THEOREM. *Let σ be a spectral set of T . Then when T is considered as an operator in the space \mathfrak{X}_σ it has σ for its spectrum. Furthermore for $\lambda \in \sigma$ any one of the following statements is true if and only if it is valid when T is regarded as an operator in \mathfrak{X}_σ : λ is in the point spectrum of T , λ is in the residual spectrum of T , λ is in the continuous spectrum of T , λ is a pole of order ν for $R_\xi(T)$, λ is an essential singularity for $R_\xi(T)$, λ is an isolated point of the spectrum.*

Let $\lambda \notin \sigma$. We shall first show that $(\lambda I - T)$ is a 1-1 bicontinuous map of \mathfrak{X}_σ into all of itself. Let $D = D(T)$ be such that $\sigma = \sigma(T)D$, $\lambda \notin \bar{D}$, and let D' be a complement to D . Define $f \in \mathcal{Y}(T)$ so that $f(\xi) = (\lambda - \xi)^{-1}$ for ξ in a neighborhood of \bar{D} and $f(\xi) = 0$ for ξ in a neighborhood of \bar{D}' . Then

$$(\lambda I - T)f(T) = E_\sigma, \quad f(T)E_\sigma = f(T).$$

Hence if $x \in \mathfrak{X}_\sigma$ we have $y = f(T)x \in \mathfrak{X}_\sigma$ and $(\lambda I - T)y = x$, and if $(\lambda I - T)u = x$ with $u \in \mathfrak{X}_\sigma$ we have $y = f(T)x = f(T)(\lambda I - T)u = E_\sigma u = u$, so that $(\lambda I - T)$ is a 1-1 map of \mathfrak{X}_σ into all of itself. Thus when T is considered as an operator in \mathfrak{X}_σ its spectrum is contained in σ . Now if $\lambda \in \sigma$ then $\lambda \notin \sigma'$ and by what has just been proved we have $(\lambda I - T)$ a 1-1 map of \mathfrak{X}_σ into all of itself. If $(\lambda I - T)$ were also a 1-1 map of \mathfrak{X}_σ into all of itself we see from the equation $\mathfrak{X} = \mathfrak{X}_\sigma \oplus \mathfrak{X}_{\sigma'}$ that $(\lambda I - T)$ would be a 1-1 map of \mathfrak{X} into all of itself, which is impossible since $\lambda \in \sigma$. Thus when T is considered as an operator in \mathfrak{X}_σ it has σ for its spectrum. Next let $\lambda \in \sigma$ be in the point spectrum of T . Let $x \neq 0$ be such that $(\lambda I - T)x = 0$ so that $(\lambda I - T)E_\sigma x = 0$, $(\lambda I - T)E_{\sigma'} x = 0$. Since $(\lambda I - T)$ is 1-1 on $\mathfrak{X}_{\sigma'}$ we have $E_{\sigma'} x = 0$, and hence $x = E_\sigma x$ and λ is in the point spectrum of T when T is considered as operating in \mathfrak{X}_σ . The converse is obvious. Now for any $\lambda \in \sigma$ we have

$$(\lambda I - T)\mathfrak{X} = (\lambda I - T)\mathfrak{X}_\sigma \oplus (\lambda I - T)\mathfrak{X}_{\sigma'} = (\lambda I - T)\mathfrak{X}_\sigma \oplus \mathfrak{X}_{\sigma'}$$

and so $(\lambda I - T)$ is dense in but not equal to \mathfrak{X} if and only if $(\lambda I - T)\mathfrak{X}_\sigma$ is dense in but not equal to \mathfrak{X}_σ . Thus for $\lambda \in \sigma$ either of the following statements is true if and only if it is true when T is regarded as operating in \mathfrak{X}_σ : λ is in the continuous spectrum of T , λ is in the residual spectrum of T . Finally let λ be an isolated point in σ . The Laurent expansion of $R_\xi(T)$ about the point λ is

$$R_\xi(T) = \sum_{n=-\infty}^{+\infty} a_n (\lambda - \xi)^n, \quad a_n = -\frac{1}{2\pi i} \int_C \frac{R_\xi(T) d\xi}{(\lambda - \xi)^{n+1}},$$

where C is a small circle about λ . Thus

$$(\lambda I - T)^k E_\lambda = \frac{1}{2\pi i} \int_C (\lambda - \xi)^k R_\xi(T) d\xi = -a_{-(k+1)},$$

and hence λ is a pole of order ν if and only if

$$(\lambda I - T)^\nu E_\lambda = 0, \quad (\lambda I - T)^{\nu-1} E_\lambda \neq 0.$$

Thus λ will be a pole of order ν when T is regarded as operating in \mathfrak{X}_σ if and only if

$$(\lambda I - T)^\nu E_\lambda E_\sigma = 0, \quad (\lambda I - T)^{\nu-1} E_\lambda E_\sigma \neq 0,$$

but since $\lambda \in \sigma$ we have $E_\lambda E_\sigma = E_\lambda$ and hence the two conditions are equivalent. This completes the proof of the theorem.

The following corollaries will be used frequently.

2.13. THEOREM. *If σ is a spectral set of T and $\lambda \notin \sigma$ then $(\lambda I - T)$ is a 1-1 bicontinuous map of \mathfrak{X}_σ into all of itself.*

2.14. THEOREM. *If σ is a spectral set of T then $E_\sigma[T] = I$ if and only if $\sigma = \sigma(T)$, and $E_\sigma[T] = 0$ if and only if σ is empty.*

This is an immediate consequence of 2.13.

2.15. THEOREM. *If λ is an isolated point in the spectrum of T then λ is a pole of order ν if and only if*

$$(\lambda I - T)^\nu E_\lambda = 0, \quad (\lambda I - T)^{\nu-1} E_\lambda \neq 0.$$

This was established in the proof of 2.12.

2.16. DEFINITION. *As in the finite case we say that a complex number λ has index ν (a positive integer or zero) in case $(\lambda I - T)^{\nu+1} x = 0$ implies $(\lambda I - T)^\nu x = 0$ and there is an $x \in \mathfrak{X}$ with $(\lambda I - T)^\nu x = 0$, $(\lambda I - T)^{\nu-1} x \neq 0$. If no such integer ν exists we say that λ is of infinite index. For any complex number λ and integer $n \geq 0$ we define*

$$\mathfrak{N}_\lambda^n = (\lambda I - T)^n \mathfrak{X}, \quad \mathfrak{M}_\lambda^n = \mathcal{E}_{x \in \mathfrak{X}} [(\lambda I - T)^n x = 0].$$

2.17. THEOREM. *If λ is a pole of order ν for $R_\xi(T)$ then λ has index ν and $\mathfrak{X}_\lambda = \mathfrak{M}_\lambda^\nu$, $\mathfrak{X}_{\sigma'} = \mathfrak{N}_\lambda^\nu$, where σ' is the spectral set of T complementary to λ . Furthermore $\mathfrak{N}_\lambda^n = \mathfrak{N}_\lambda^\nu$ for $n \geq \nu$ while for $1 \leq n \leq \nu$, \mathfrak{N}_λ^n is a proper subset of $\mathfrak{N}_\lambda^{n-1}$.*

By 2.15 we have

$$(*) \quad (\lambda I - T)^\nu E_\lambda = 0, \quad (\lambda I - T)^{\nu-1} E_\lambda \neq 0.$$

Thus there is an x with $(\lambda I - T)^\nu E_\lambda x = 0$, $(\lambda I - T)^{\nu-1} E_\lambda x \neq 0$, and so the index of λ is at least ν . Now suppose that for some n and x we have $(\lambda I - T)^n x = 0$. Then since

$$R_{\xi}(T) = - \sum_{n=0}^{\infty} \frac{(\lambda I - T)^n}{(\lambda - \xi)^{n+1}}, \quad |\lambda - \xi| > |\lambda I - T|,$$

we see that the function

$$R_{\xi}(T)x = - \sum_{j=0}^{n-1} \frac{(\lambda I - T)^j}{(\lambda - \xi)^{j+1}} x$$

is regular over the entire plane except perhaps at the point $\xi = \lambda$. Thus if K is any contour surrounding $\sigma(T)$ and C is a small circle about λ we have

$$(**) \quad x = \frac{1}{2\pi i} \int_K R_{\xi}(T) x d\xi = \frac{1}{2\pi i} \int_C R_{\xi}(T) x d\xi = E_{\lambda} x,$$

and hence $(\lambda I - T)^{\nu} x = (\lambda I - T)^{\nu} E_{\lambda} x = 0$, which proves that λ has index at most ν . Thus λ has index ν . Equation (*) shows that $\mathfrak{X}_{\lambda} \subset \mathfrak{M}_{\lambda}^{\nu}$ and (**), which holds for $x \in \mathfrak{M}_{\lambda}^{\nu}$, shows that $\mathfrak{M}_{\lambda}^{\nu} \subset \mathfrak{X}_{\lambda}$. Thus $\mathfrak{M}_{\lambda}^{\nu} = \mathfrak{X}_{\lambda}$. Now

$$\mathfrak{X} = \mathfrak{X}_{\lambda} \oplus \mathfrak{X}_{\sigma'} = \mathfrak{M}_{\lambda}^{\nu} \oplus \mathfrak{X}_{\sigma'},$$

and by 2.13

$$\mathfrak{X}_{\sigma'} = (\lambda I - T)^{\nu} \mathfrak{X}_{\sigma'} \subset (\lambda I - T)^{\nu} \mathfrak{X} = \mathfrak{M}_{\lambda}^{\nu}$$

so that $\mathfrak{X} = \mathfrak{M}_{\lambda}^{\nu} + \mathfrak{M}_{\lambda}^{\nu}$. Since λ has index ν we have $\mathfrak{M}_{\lambda}^{\nu} \cdot \mathfrak{M}_{\lambda}^{\nu} = 0$ and hence $\mathfrak{X} = \mathfrak{M}_{\lambda}^{\nu} \oplus \mathfrak{M}_{\lambda}^{\nu}$. Thus $\mathfrak{X}_{\sigma'}$ and $\mathfrak{M}_{\lambda}^{\nu}$ are both complements to $\mathfrak{M}_{\lambda}^{\nu}$ and $\mathfrak{X}_{\sigma'} \subset \mathfrak{M}_{\lambda}^{\nu}$ which proves that $\mathfrak{M}_{\lambda}^{\nu} = \mathfrak{X}_{\sigma'}$. By 2.13 $\mathfrak{M}_{\lambda}^{\nu+k} = (\lambda I - T)^k \mathfrak{M}_{\lambda}^{\nu} = (\lambda I - T)^k \mathfrak{X}_{\sigma'} = \mathfrak{X}_{\sigma'} = \mathfrak{M}_{\lambda}^{\nu}$ and so $\mathfrak{M}_{\lambda}^n = \mathfrak{M}_{\lambda}^{\nu}$ for $n \geq \nu$. Since if, for some k , $\mathfrak{M}_{\lambda}^{k+1} = \mathfrak{M}_{\lambda}^k$ we have $\mathfrak{M}_{\lambda}^n = \mathfrak{M}_{\lambda}^k$ for all $n \geq k$, to prove the final assertion of the theorem it will suffice to show that $\mathfrak{M}_{\lambda}^{\nu}$ is a proper subset of $\mathfrak{M}_{\lambda}^{\nu-1}$. Since λ has index ν there is an x with $(\lambda I - T)^{\nu} x = 0$, $0 \neq y = (\lambda I - T)^{\nu-1} x \in \mathfrak{M}_{\lambda}^{\nu-1}$. Since $0 \neq y \in \mathfrak{M}_{\lambda}^{\nu}$ and $\mathfrak{M}_{\lambda}^{\nu} \mathfrak{M}_{\lambda}^{\nu} = 0$ we have $y \in \mathfrak{M}_{\lambda}^{\nu-1} - \mathfrak{M}_{\lambda}^{\nu}$ which completes the proof of the theorem.

There is a converse to Theorem 2.17. It states that if for some λ and m we have $\mathfrak{M}_{\lambda}^m \oplus \mathfrak{M}_{\lambda}^m = \mathfrak{X}$ and \mathfrak{M}_{λ}^m closed, then λ is either in the resolvent set $\rho(T)$ or else a pole of order at most m for $R_{\xi}(T)$. This result is a special case of 2.23 which will be proved later.

2.18. THEOREM. *Let λ be either in $\rho(T)$ or an isolated point of the spectrum. Let δ be the distance from λ to the rest of the spectrum. Then the following statements are equivalent:*

- (1) $x \in \mathfrak{X}_{\lambda}$.
- (2) $\lim_n (\lambda I - T)^n \epsilon^{-n} x = 0$ for every $\epsilon > 0$.
- (3) For some positive $\epsilon < \delta$ the equation (2) is satisfied.

To see that (1) implies (2) let C be the circle $|\lambda - \xi| = \rho$ where $0 < \rho < \epsilon$. Then

$$(\lambda I - T)^n \epsilon^{-n} E_\lambda = \frac{1}{2\pi i} \int_C (\lambda - \xi)^n \epsilon^{-n} R_\xi(T) d\xi \rightarrow 0.$$

Conversely suppose (2) is satisfied for some positive $\epsilon < \delta$. Let $0 < \epsilon_1 < \epsilon < \epsilon_2 < \epsilon_3$, and C_i ($i=1, 2, 3$), the circle $|\lambda - \xi| = \epsilon_i$, be chosen so that the spectrum $\sigma(T)$ with the exception of the point $\xi = \lambda$ is in the open annulus bounded by C_2 and C_3 . Then the function $f_n(\xi) = [1 - (\lambda - \xi)^n \epsilon^{-n}]^{-1}$ is in $\mathcal{F}(T)$ and

$$f_n(T) = \frac{1}{2\pi i} \left[\int_{C_1} f_n(\xi) R_\xi(T) d\xi - \int_{C_2} f_n(\xi) R_\xi(T) d\xi + \int_{C_3} f_n(\xi) R_\xi(T) d\xi \right],$$

$$f_n(T) \rightarrow E_\lambda.$$

Since we are assuming (2) to hold for ϵ we have $f_n^{-1}(T)x = x - (\lambda I - T)^n \epsilon^{-n} x \rightarrow x$ and thus for arbitrary $\rho > 0$, $|x - f_n^{-1}(T)x| < \rho$, $|f_n(T)x - x| < \rho |f_n(T)|$ for all large n . Since $f_n(T) \rightarrow E_\lambda$ we have $|E_\lambda x - x| < \rho |E_\lambda|$ and since ρ is arbitrary, $E_\lambda x = x$.

2.19. THEOREM (THE MINIMAL EQUATION THEOREM). *If $f \in \mathcal{F}(T)$ then $f(T) = 0$ if and only if*

(1) *For every pole λ of $R_\xi(T)$ of order ν*

$$f^{(j)}(\lambda) = 0, \quad j = 0, 1, \dots, \nu - 1,$$

(2) *$f(\lambda) \equiv 0$ in a neighborhood of the spectrum $\sigma(T)$ excluding poles of $R_\xi(T)$.*

Suppose (1) and (2) are satisfied. Let $D = D_1 + \dots + D_k$ be a domain upon which f is regular and single-valued and with $D \supset \sigma(T)$. Clearly, in view of (2),

$$(*) \quad f(T) = \frac{1}{2\pi i} \int_{C^*} f(\lambda) R_\lambda(T) d\lambda,$$

where C^* is the boundary of those D_i containing no singularities of $R_\xi(T)$ except a finite number of poles. Let $\lambda_1, \dots, \lambda_k$ be the poles contained within C^* and let ν_1, \dots, ν_k be their orders. Then from (1) we have $f(\lambda) = (\lambda_i - \lambda)^{\nu_i} g_i(\lambda)$ where g_i is regular at λ_i and $g_i \in \mathcal{F}(T)$. Hence by 2.15, $f(T)E_{\lambda_i} = g_i(T)(\lambda_i I - T)^{\nu_i} E_{\lambda_i} = 0$ and hence from (*) we have

$$f(T) = \sum_{i=1}^k f(T)E_{\lambda_i} = 0.$$

Now conversely suppose $f(T) = 0$. Then $0 = \sigma(f(T)) = f(\sigma(T))$ and so $f(\lambda)$ is identically zero in any connected domain (upon which it is regular) containing infinitely many points of $\sigma(T)$. Hence equation (*) holds where C^* is now the boundary of a finite number of domains containing only a finite number of singularities of $R_\xi(T)$. Let $\lambda_1, \dots, \lambda_n$ be the poles within C^* and

ν_1, \dots, ν_n their orders. Let $\lambda_{n+1}, \dots, \lambda_m$ be the essential singularities within C^* . Now suppose that for some $i \leq n$ we have

$$f^{(k)}(\lambda_i) \neq 0, \quad f^{(j)}(\lambda_i) = 0, \quad j = 0, \dots, k-1,$$

where $k < \nu_i$. Define a function $g \in \mathcal{F}(T)$ so that in a neighborhood of λ_i we have $f(\lambda) = (\lambda_i - \lambda)^k g(\lambda)$ while $g(\lambda) \equiv 1$ in a neighborhood of the spectrum $\sigma(T)$ excluding λ_i . Thus g, g^{-1} are both in $\mathcal{F}(T)$. Since $k < \nu_i$ we have by 2.17 a point $x \in \mathfrak{X}_{\lambda_i}$ with $(\lambda_i I - T)^k x \neq 0$, and thus since $g(T)$ has an inverse

$$f(T)x = f(T)E_{\lambda_i}x = g(T)(\lambda_i I - T)^k E_{\lambda_i}x \neq 0,$$

which is a contradiction.

The coefficient a_{-n-1} of $(\lambda_i - \xi)^{-n-1}$ in the Laurent expansion of $R_\xi(T)$ about the point $\xi = \lambda_i$ is (as was shown in the proof of 2.12) $-(\lambda_i I - T)^n E_{\lambda_i}$ and thus for $i > n$ we have $(\lambda_i I - T)^n E_{\lambda_i} \neq 0$ for all $n = 1, 2, \dots$. Just as before suppose that for some $i > n$ and some $k \geq 1$ we have $f^{(k)}(\lambda_i) \neq 0$, $f^{(j)}(\lambda_i) = 0, j = 0, \dots, k-1$, and as before define $g \in \mathcal{F}(T)$ so that $g^{-1}(T)$ exists and $f(T)E_{\lambda_i} = (\lambda_i I - T)^k g(T)E_{\lambda_i} \neq 0$, which again is a contradiction.

An immediate corollary is

2.20. THEOREM. If $f, g \in \mathcal{F}(T)$ then $f(T) = g(T)$ if and only if

(1) For every pole λ of $R_\xi(T)$ of order ν

$$f^{(j)}(\lambda) = g^{(j)}(\lambda), \quad j = 0, \dots, \nu - 1.$$

(2) $f(\lambda) = g(\lambda)$ for every λ in a neighborhood of the spectrum $\sigma(T)$ excluding poles.

2.21. THEOREM. Let $\lambda_1, \dots, \lambda_k$ be poles of $R_\xi(T)$ of orders ν_1, \dots, ν_k respectively. Let σ' be the complement of the spectral set $\sigma = (\lambda_1, \dots, \lambda_k)$. Then for every $f \in \mathcal{F}(T)$ we have

$$f(T) = f(T)E_{\sigma'} + \sum_{i=1}^k \sum_{n=0}^{\nu_i-1} \frac{(T - \lambda_i I)^n}{n!} f^{(n)}(\lambda_i) E_{\lambda_i}.$$

This is an immediate corollary of 2.20.

2.22. DEFINITION. Let $f \in \mathcal{F}(T)$, then as in 1.1 we define the manifolds

$$\mathfrak{N}[f] = f(T)\mathfrak{X}, \quad \mathfrak{M}[f] = \mathcal{E}_{x \in \mathfrak{X}} [f(T)x = 0].$$

2.23. THEOREM. Let $f \in \mathcal{F}(T)$ and suppose that $f(\lambda)$ is not identically zero on any of the domains in which it is regular. Let $\lambda_1, \dots, \lambda_k$ be the roots of $f(\lambda)$ and m_1, \dots, m_k their multiplicities. Then the following statements are equivalent:

(1) $\mathfrak{M}[f] \oplus \mathfrak{N}[f] = \mathfrak{X}$, $\mathfrak{N}[f]$ is closed.

(2) For $i=1, \dots, k$, λ_i is either in the resolvent set $\rho(T)$ or else a pole of order $\nu_i \leq m_i$ of $R_i(T)$.

(3) The finite set $\sigma = (\lambda_1, \dots, \lambda_k)\sigma(T)$ is a spectral set of T and

$$\mathfrak{X}_\sigma = \mathfrak{M}[f], \quad \mathfrak{X}_{\sigma'} = \mathfrak{N}[f].$$

Assuming (1) we have continuous projections E, E' with $E\mathfrak{X} = \mathfrak{M}[f]$, $E'\mathfrak{X} = \mathfrak{N}[f]$, $EE' = 0$, $I = E + E'$, $TE = ET$. Also for any λ we have $(\lambda I - T)$ a 1-1 map of \mathfrak{X} into all of itself if and only if $(\lambda I - T)$ is a 1-1 map of $E\mathfrak{X}$ into all of itself and of $E'\mathfrak{X}$ into all of itself. Thus the spectrum $\sigma(T)$ of T is $\sigma_1 + \sigma'_1$ where σ_1, σ'_1 are the spectrums of T when it is considered as an operator in $E\mathfrak{X}$, $E'\mathfrak{X}$ respectively. In $E\mathfrak{X}$ we have $f(T) = 0$ and so σ_1 is contained in the point set $(\lambda_1, \dots, \lambda_k)$. Since $f(T)E = 0$ we have $f(T) = f(T)E'$ and hence

$$\mathfrak{N}[f] = f(T)\mathfrak{X} = f(T)E'\mathfrak{X} = f(T)\mathfrak{N}[f] \subset f(T)\mathfrak{X} = \mathfrak{N}[f].$$

Therefore, since $\mathfrak{M}[f] \cdot \mathfrak{N}[f] = 0$, $f(T)$ is a 1-1 map of $\mathfrak{N}[f] = E'\mathfrak{X}$ into all of itself. Hence when regarded as an operator in $E'\mathfrak{X}$, $0 \in \rho(f(T))$ and hence $\lambda_i \in \rho(T)$, $i=1, \dots, k$. Thus σ'_1 is a closed point set not containing $\sigma = \sigma(T)(\lambda_1, \dots, \lambda_k)$. This shows that σ as well as $\lambda_i \in \sigma$ is a spectral set of T . Now let $x = Ex + E'x \in X_{\lambda_i}$, then 2.18 gives

$$\lim_n (\lambda_i I - T)^n \epsilon^{-n} E'x = 0, \quad \epsilon > 0,$$

and hence since $(\lambda_i I - T)$ is a 1-1 bicontinuous map of $E'\mathfrak{X}$ into all of itself we have $E'x = 0$ and $x = Ex \in \mathfrak{M}[f]$. Thus

$$\mathfrak{X}_\sigma = \mathfrak{X}_{\lambda_1} \oplus \dots \oplus \mathfrak{X}_{\lambda_k} \subset \mathfrak{M}[f].$$

On the other hand if $x \in \mathfrak{M}_{\lambda_i}^{m_i}$ we have (by 2.18) $x \in \mathfrak{X}_{\lambda_i}$ and so by 1.4

$$\mathfrak{M}[P] = \mathfrak{M}_{\lambda_1}^{m_1} \oplus \dots \oplus \mathfrak{M}_{\lambda_k}^{m_k} \subset \mathfrak{X}_\sigma,$$

where $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{m_i}$. If $h(\lambda)$ is defined on the same domain as $f(\lambda)$ by the equation $f(\lambda) = P(\lambda) \cdot h(\lambda)$ we see that $h(\lambda)$ has no roots in $\sigma(T)$ and hence $h^{-1}(T)$ exists, which shows that $\mathfrak{M}[P] = \mathfrak{M}[f]$ and thus the above equations give $\mathfrak{M}[f] = \mathfrak{X}_\sigma$. Now since $f(\lambda) \neq 0$ and in $\mathfrak{X}_\sigma = \mathfrak{M}[f]$ we have $f(T) = 0$ we see from the minimal equation theorem that in \mathfrak{X}_σ every spectral point λ_i must be a pole of order $\nu_i \leq m_i$. Thus (2) follows from 2.12. It has therefore been proved that (1) implies (2). During the course of this proof all but the last part of (3) has also been established. To see that $\mathfrak{X}_{\sigma'} = \mathfrak{N}[f]$ note first that since $0 \notin f(\sigma')$ we have (by 2.12)

$$\mathfrak{X}_{\sigma'} = f(T)\mathfrak{X}_{\sigma'} \subset f(T)\mathfrak{X} = \mathfrak{N}[f].$$

Next

$$\mathfrak{X} = \mathfrak{X}_\sigma \oplus \mathfrak{X}_{\sigma'} = \mathfrak{M}[f] \oplus \mathfrak{X}_{\sigma'},$$

and so $\mathfrak{N}[f]$ and $\mathfrak{X}_{\sigma'}$ are both complements of $\mathfrak{M}[f]$ and $\mathfrak{X}_{\sigma'} \subset \mathfrak{N}[f]$. This clearly implies that $\mathfrak{X}_{\sigma'} = \mathfrak{N}[f]$. Thus (1) implies (3). Clearly (3) implies (1). Now assume (2) and let $\nu_i = 0$ if λ_i is in $\rho(T)$, then if $P(\lambda)$ is the polynomial defined above we have by 2.17 and 1.4

$$\begin{aligned}\mathfrak{X}_{\sigma} &= \mathfrak{X}_{\lambda_1} \oplus \cdots \oplus \mathfrak{X}_{\lambda_k} = \mathfrak{M}_{\lambda_1}^{\nu_1} \oplus \cdots \oplus \mathfrak{M}_{\lambda_k}^{\nu_k} \\ &= \mathfrak{M}_{\lambda_1}^{m_1} \oplus \cdots \oplus \mathfrak{M}_{\lambda_k}^{m_k} = \mathfrak{M}[P] = \mathfrak{M}[f].\end{aligned}$$

Also by 2.17 and 1.4 we have, for every x such that $P^2(T)x = 0$,

$$x \in \mathfrak{M}_{\lambda_1}^{2m_1} \oplus \cdots \oplus \mathfrak{M}_{\lambda_k}^{2m_k} = \mathfrak{M}_{\lambda_1}^{\nu_1} \oplus \cdots \oplus \mathfrak{M}_{\lambda_k}^{\nu_k} = \mathfrak{M}[P]$$

and so $P(T)x = 0$, that is $\mathfrak{N}[P] \cdot \mathfrak{M}[P] = 0$. Thus since (using 2.12)

$$\mathfrak{N}[f] = \mathfrak{N}[P] \supset P(T)\mathfrak{X}_{\sigma'} = \mathfrak{X}_{\sigma'}$$

we have

$$\mathfrak{X} = \mathfrak{X}_{\sigma} \oplus \mathfrak{X}_{\sigma'} = \mathfrak{M}[P] \oplus \mathfrak{N}[P] = \mathfrak{M}[f] \oplus \mathfrak{N}[f],$$

and $\mathfrak{X}_{\sigma'} = \mathfrak{N}[f]$ is closed. This completes the proof of 2.23.

The formula $\sigma(f(T)) = f(\sigma(T))$ may be given a much sharper form. To see this let us first examine the finite case of an $n \times n$ matrix T whose distinct characteristic numbers are $\lambda_1, \dots, \lambda_k$. The formula asserts that $f(\lambda_1), \dots, f(\lambda_k)$ are the characteristic numbers of $f(T)$. However a theorem due to Sylvester asserts more; namely, if $\lambda_1, \dots, \lambda_n$ are the characteristic numbers of T , each repeated according to its multiplicity, then $f(\lambda_1), \dots, f(\lambda_n)$ are the characteristic numbers of $f(T)$, where the number of repetitions of a given number in this array is its multiplicity. In the language of determinants this theorem of Sylvester asserts that if $\lambda_1, \dots, \lambda_k$ are the *distinct* characteristic numbers of T and if m_i is the multiplicity of λ_i then $\text{Det}(f(T) - \lambda I) = \prod_{i=1}^k (f(\lambda_i) - \lambda)^{m_i}$. Recalling that the multiplicity $m(\lambda_i, T)$ of a root λ_i of T of index ν_i is the number of linearly independent solutions of $(\lambda_i I - T)^{\nu_i} x = 0$, that is $m(\lambda_i, T)$ is the dimension of $\mathfrak{M}_{\lambda_i}^{m_i} = \mathfrak{M}_{\lambda_i}^{\nu_i}$, it is seen that the Sylvester theorem, that is the formula

$$(i) \quad m(\mu, f(T)) = \sum_{\lambda \in f^{-1}(\mu)} m(\lambda, T)$$

(where $m(\mu, f(T))$ is the multiplicity of μ as a root of $f(T)$ and the sum on the right is taken over all λ such that $f(\lambda) = \mu$) is an immediate consequence of

$$(ii) \quad \mathfrak{X}_{\mu}[f(T)] = \sum_{\lambda \in f^{-1}(\mu)} \mathfrak{X}_{\lambda}[T]$$

(where the sum is a vector sum). Formula (ii) which states that two manifolds are equal is for many purposes more desirable than (i) which merely states that the dimensions of these manifolds are the same. On the other

hand (ii) is readily derived from (i) as follows. Suppose $(\lambda I - T)^{\nu}x = 0$ (that is $x \in \mathfrak{X}_{\lambda}(T)$), where $f(\lambda) = \mu$, then $(\mu I - f(T))^{\nu}x = (f(\lambda)I - f(T))^{\nu}x = (\lambda I - T)^{\nu}g^{\nu}(T)x = 0$ and so $x \in \mathfrak{X}_{\mu}[f(T)]$. Thus the left side of (ii) contains the right but by (i) the dimension of the left side is the same as the dimension of the right side and so the manifolds are the same. A form of this theorem which is more desirable than (ii) is

$$(iii) \quad E_{\mu}[f(T)] = \sum_{\lambda \in f^{-1}(\mu)} E_{\lambda}[T],$$

which states the equality of two projections rather than (as (ii) does) the equality of the manifolds upon which they project.

In Theorems 2.24 and 2.29 to follow we shall give two forms of (iii) applicable to a general linear operator T .

2.24. THEOREM. *Let $f \in \mathcal{F}(T)$ and let τ be a spectral set of $f(T)$. Then $\sigma = \sigma(T) \cdot f^{-1}(\tau)$ is a spectral set of T and*

$$E_{\tau}[f(T)] = E_{\sigma}[T].$$

Let τ' be the complement of τ in $\sigma(f(T))$. Then since $\sigma(f(T)) = f(\sigma(T))$ we have $\sigma + \sigma' = \sigma(T)$ where $\sigma' = \sigma(T) \cdot f^{-1}(\tau')$. Since τ, τ' are closed and disjoint and $f(\lambda)$ is continuous we see that σ, σ' are closed and disjoint and thus they are complementary spectral sets of T . Let D, D' be $f(T)$ -admissible domains with $\overline{D}\overline{D}' = 0, \tau = \sigma(f(T)) \cdot D, \tau' = \sigma(f(T))D'$. Then there are T -admissible domains D_1, D_1' with $\overline{D}_1\overline{D}_1' = 0, \sigma = \sigma(T)D_1, \sigma' = \sigma(T)D_1', f$ is regular on $\overline{D}_1 + \overline{D}_1'$ and $f(\overline{D}_1) \subset D, f(\overline{D}_1') \subset D'$. Let g, h be two characteristic functions with $g \in \mathcal{F}(f(T)), h \in \mathcal{F}(T)$ and such that

$$g(\xi) = \begin{cases} 1, & \xi \in \overline{D}, \\ 0, & \xi \in \overline{D}', \end{cases} \quad h(\xi) = \begin{cases} 1, & \xi \in \overline{D}_1, \\ 0, & \xi \in \overline{D}_1', \end{cases}$$

and let $F(\xi) = g(f(\xi))$. By 2.10 we have $F(T) = g(f(T))$. Since $F(\xi) = h(\xi)$ for $\xi \in \overline{D}_1 + \overline{D}_1'$ we have $F(T) = h(T)$ and

$$E_{\tau}[f(T)] = g(f(T)) = F(T) = h(T) = E_{\sigma}[T].$$

2.25. LEMMA. *Let Γ be a set of complex numbers, and for each $\xi \in \Gamma$ let $B(\xi), A_n(\xi), n = 1, 2, \dots$, be bounded linear operators which have inverses. Let M, M_1, M_2, \dots be constants such that*

$$|B^{-1}(\xi)| \leq M, \quad |A_n^{-1}(\xi)| \leq M_n, \quad \xi \in \Gamma.$$

Then if $\lim_n A_n(\xi) = B(\xi)$ uniformly for $\xi \in \Gamma$ we have also $\lim_n A_n^{-1}(\xi) = B^{-1}(\xi)$ uniformly for $\xi \in \Gamma$.

The proof is the same as that given by Gelfand⁽⁸⁾ for a special case.

(8) Loc. cit.

We shall prove the lemma first for the case $B(\xi) \equiv I$. In this case there is an integer n_0 such that $|A_n(\xi) - I| < 1/2$ for $\xi \in \Gamma$ and $n \geq n_0$. Thus for $n \geq n_0$

$$A_n^{-1}(\xi) = \sum_{p=0}^{\infty} [I - A_n(\xi)]^p, \quad |A_n^{-1}(\xi)| \leq 2, \quad \xi \in \Gamma.$$

Hence $\text{l.u.b.}_{1 \leq n < \infty} \text{l.u.b.}_{\xi \in \Gamma} |A_n^{-1}(\xi)| = K < \infty$ and

$$|A_n^{-1}(\xi) - I| = |A_n^{-1}(\xi)(I - A_n(\xi))| \leq K |I - A_n(\xi)|,$$

which proves the special case of $B(\xi) \equiv I$. Now in the general case we have

$$|A_n(\xi)B^{-1}(\xi) - I| = |(A_n(\xi) - B(\xi))B^{-1}(\xi)| \leq M |A_n(\xi) - B(\xi)|,$$

and so by the first case we have

$$\lim_n B(\xi)A_n^{-1}(\xi) = I$$

uniformly on Γ . Thus

$$|A_n^{-1}(\xi) - B^{-1}(\xi)| = |B^{-1}(\xi)(I - B(\xi)A_n^{-1}(\xi))| \leq M |I - B(\xi)A_n^{-1}(\xi)|,$$

which proves the lemma.

2.26. LEMMA. Let $B, A_n, n = 1, 2, \dots$, be bounded linear operators and let Γ be a closed set of complex numbers with $\Gamma \subset \rho(B)\rho(A_n), n = 1, 2, \dots$. Then if $A_n \rightarrow B$ we have $R_\xi(A_n) \rightarrow R_\xi(B)$ uniformly on Γ .

Let $B(\xi) = \xi I - B, A_n(\xi) = \xi I - A_n$. Since Γ is closed and in $\rho(B)$ it must be at positive distance from $\sigma(B)$. Since $|R_\xi(B)|$ is continuous on Γ and $|R_\xi(B)| \rightarrow 0$ as $\xi \rightarrow \infty$ we see that $\text{l.u.b.}_{\xi \in \Gamma} |B^{-1}(\xi)| < \infty$, similarly for $A_n^{-1}(\xi)$. The lemma thus follows from the preceding one.

2.27. DEFINITION. By $\mathcal{R}(T)$ we shall mean the ring of all operators of the form $f(T)$ where $f \in \mathcal{F}(T)$, and by $\overline{\mathcal{R}}(T)$ the ring of all operators which are limits (in the uniform topology of operators) of elements in $\mathcal{R}(T)$.

2.28. LEMMA. With every $U \in \overline{\mathcal{R}}(T)$ is associated a uniquely defined scalar function f called the spectral function of U , such that:

- (i) The domain of definition of f is $\sigma(T)$ and $f(\lambda)$ is continuous for $\lambda \in \sigma(T)$.
- (ii) If $f_n \in \mathcal{F}(T)$ and $f_n(T) \rightarrow U$ then $f_n(\lambda) \rightarrow f(\lambda)$ uniformly for $\lambda \in \sigma(T)$.
- (iii) $f(\sigma(T)) \subset \sigma(U)$.

If $f_n \in \mathcal{F}(T)$ and $f_n(T) \rightarrow U$ then for every $\lambda \in \sigma(T)$ we have $f_n(\lambda) - f_m(\lambda) \in \sigma(f_n(T) - f_m(T))$ and so

$$\text{l.u.b.}_{\lambda \in \sigma(T)} |f_n(\lambda) - f_m(\lambda)| \leq |f_n(T) - f_m(T)| \rightarrow 0.$$

Thus $f(\lambda) = \lim_n f_n(\lambda)$ exists on $\sigma(T)$ and is continuous there. If also $g_n(T) \rightarrow U$

then by the same argument $|f_n(\lambda) - g_n(\lambda)| \rightarrow 0$ uniformly for $\lambda \in \sigma(T)$, and so $f(\lambda)$ is independent of the particular sequence $f_n(T) \in \mathcal{R}(T)$ used to approach U . To prove (iii) let $\lambda \in \sigma(T)$, so that $f_n(\lambda) \in \sigma(f_n(T))$, and suppose that $f(\lambda) \notin \sigma(U)$. Since $f_n(\lambda)I - f_n(T) \rightarrow f(\lambda)I - U$ and $[f(\lambda)I - U]^{-1}$ exists we have, for all large n , $[f_n(\lambda)I - f_n(T)]^{-1}$ existing, which contradicts the fact that $f_n(\lambda) \in \sigma(f_n(T))$.

We are now prepared to give the general form of the Sylvester theorem.

2.29. THEOREM. *Let f be the spectral function of an operator $U \in \overline{\mathcal{R}}(T)$ and let τ be a spectral set of U . Then $\sigma = f^{-1}(\tau)$ is a spectral set of T and*

$$E_\sigma[T] = E_\tau[U].$$

Let τ' be the spectral set of U complementary to τ and let $\sigma' = f^{-1}(\tau')$. By the preceding lemma $f(\sigma(T)) \subset \tau + \tau'$ and so $\sigma + \sigma' = \sigma(T)$, $\sigma\sigma' = 0$. Since τ, τ' are both closed in $\sigma(U)$ and f is continuous we see that σ, σ' are both closed and hence are complementary spectral sets of T . Let D, D' be U -admissible domains with $\tau = D \cdot \sigma(U)$, $\tau' = D' \cdot \sigma(U)$, $\overline{D}\overline{D}' = 0$ and let C be the boundary of D . Let $f_n \in \mathcal{F}(T)$ and $f_n(T) \rightarrow U$. Since $f_n(\lambda) \rightarrow f(\lambda)$ uniformly on $\sigma(T)$ and $f(\sigma(T)) \subset \tau + \tau'$ we have for all large n

$$(*) \quad f_n(\sigma) \subset D, \quad f_n(\sigma') \subset D'.$$

Thus $f_n(\sigma) = f_n(\sigma)D$ is a spectral set of $f_n(T)$ and $\sigma(T)f_n^{-1}(f_n(\sigma)) = \sigma$, and so, by 2.24, $E_\sigma[T] = E_{f_n(\sigma)}[f_n(T)]$. Now $(*)$ shows that $C \subset \rho[f_n(T)]$ and thus, by 2.26, $R_\xi(f_n(T)) \rightarrow R_\xi(U)$ uniformly for $\xi \in C$. Hence

$$E_\sigma[T] = E_{f_n(\sigma)}[f_n(T)] = \frac{1}{2\pi i} \int_C R_\xi(f_n(T)) d\xi \rightarrow \frac{1}{2\pi i} \int_C R_\xi(U) d\xi = E_\tau[U],$$

which completes the proof of the theorem.

2.30. THEOREM. *Let $\mu \in \sigma(f(T))$ where $f \in \mathcal{F}(T)$ and suppose that $f(\lambda) \neq \mu$ on any of the domains in which it is regular. Let $\lambda_1, \dots, \lambda_k$ be those points λ in $\sigma(T)$ where $f(\lambda) = \mu$. Then*

- (1) *If μ is a pole of $R_\xi(f(T))$ then $\lambda_1, \dots, \lambda_k$ are poles of $R_\xi(T)$.*
- (2) *Conversely if $\lambda_1, \dots, \lambda_k$ are poles of $R_\xi(T)$ then μ is a pole of $R_\xi(f(T))$.*

Since $f(\lambda)$ is not identically μ on any of the domains in which it is regular it is clear that there are only a finite number $\lambda_1, \dots, \lambda_k$ of roots of $f(\lambda) - \mu = 0$ which lie in the spectrum $\sigma(T)$. Since $f(\sigma(T)) = \sigma(f(T))$ and $\mu \in \sigma(f(T))$ there is at least one root of $f(\lambda) - \mu = 0$ in $\sigma(T)$. Let m_1, \dots, m_k be the multiplicities of $\lambda_1, \dots, \lambda_k$ as roots of $f(\lambda) - \mu = 0$ and let $h \in \mathcal{F}(T)$ be defined on the same domain as f by the equation

$$\mu - f(\lambda) = (\lambda_1 - \lambda)^{m_1} \cdots (\lambda_k - \lambda)^{m_k} h(\lambda)$$

so that $h(\lambda)$ has no roots in $\sigma(T)$ and thus $h^{-1}(T)$ exists. To prove (1) we have,

by 2.24, $\sigma = (\lambda_1, \dots, \lambda_k)$ a spectral set of T and

$$(*) \quad E_\mu[f(T)] = E_{\lambda_1}[T] + \dots + E_{\lambda_k}[T].$$

Now for any integer n we have

$$(**) \quad [\mu I - f(T)]^n E_\mu[f(T)] \\ = (\lambda_1 I - T)^{m_1 n} \dots (\lambda_k I - T)^{m_k n} h^n(T) (E_{\lambda_1}[T] + \dots + E_{\lambda_k}[T]).$$

If $n-1$ is the order of μ as a pole of $R_\xi(f(T))$ we have, by 2.15, the left side of $(**)$ vanishing, and upon operating with $E_{\lambda_i}[T]$ we get

$$(\lambda_i I - T)^{m_i n} \dots (\lambda_k I - T)^{m_k n} h^n(T) E_{\lambda_i}[T] = 0.$$

Since $h^{-1}(T)$ exists and since, by 2.13, $(\lambda_j I - T)$ for $j \neq i$ is a 1-1 map of \mathfrak{X}_{λ_i} into all of itself we have

$$(\lambda_i I - T)^{m_i n} E_{\lambda_i}[T] = 0,$$

which, since λ_i is isolated in $\sigma(T)$, proves (by 2.15) that λ_i is a pole of $R_\xi(T)$. Now conversely suppose that λ_i , $i=1, \dots, k$, is a pole of order ν_i for $R_\xi(T)$ and let n be an integer with $n \geq \nu_i$, $i=1, \dots, k$. Since $f(\lambda)$ is continuous on $\sigma(T)$ it is seen that μ is isolated in $\sigma[f(T)]$ and thus $(*)$ and hence $(**)$ are valid. Since, by 2.15, $(\lambda_i I - T)^{m_i n} E_{\lambda_i}[T] = 0$ this same theorem shows in view of $(**)$ that μ is a pole of $R_\xi(f(T))$.

2.31. DEFINITION. Let S be the unit sphere in \mathfrak{X} . An operator U in \mathfrak{X} is said to be compact in case US is compact in \mathfrak{X} and it is said to be weakly compact in case US is weakly (sequentially) compact in \mathfrak{X} .

We prefer the terminology of 2.31 to the usual terms, that is, completely continuous and weakly completely continuous.

2.32. THEOREM. Let $f \in \mathcal{F}(T)$ be such that $f(T)$ is compact. Then every $\lambda \in \sigma(T)$ with $f(\lambda) \neq 0$ is a pole of $R_\xi(T)$ and the corresponding manifold $\mathfrak{X}_\lambda[T]$ has finite dimension.

By a well known result of F. Riesz^(*) $\mu = f(\lambda)$ is isolated in the spectrum $\sigma(U)$ of the compact operator $U = f(T)$. Thus if $\sigma = \sigma(T)f^{-1}(\mu)$ we have, by 2.24, σ a spectral set of T and

$$\mathfrak{X}_\mu[U] = \mathfrak{X}_\sigma[T].$$

Now by 2.13 we see that U is a 1-1 map of $\mathfrak{X}_\mu[U]$ into all of itself and hence the unit sphere in $\mathfrak{X}_\mu[U]$ is compact and $\mathfrak{X}_\mu[U]$, therefore, has finite dimension. Thus the spectrum of T where it is considered as operating in $\mathfrak{X}_\sigma[T]$ is finite and by 2.12, therefore, σ is finite, which shows that λ is a spectral set of T . Hence $\mathfrak{X}_\lambda[T] \subset \mathfrak{X}_\sigma[T]$ has finite dimension and, in $\mathfrak{X}_\lambda[T]$, $(\lambda I - T)$ has the

(*) F. Riesz, *Über lineare Funktionalgleichungen*, Acta Math. vol. 41 (1918) pp. 71-98.

single point 0 for its spectrum. Thus in $\mathfrak{X}_\lambda[T]$ the matrix $(\lambda I - T)$ is nilpotent, that is, for some n , $(\lambda I - T)^n E_\lambda[T] = 0$ which proves that λ is a pole of $R_\xi[T]$.

2.33. THEOREM. *Let $f_n \in \mathcal{F}(T)$ and let $U = \lim_n f_n(T)$ be a compact operator. Then every $\lambda \in \sigma(T)$ for which $f_n(\lambda)$ does not approach 0 is a pole of $R_\xi(T)$ and the corresponding manifold $\mathfrak{X}_\lambda[T]$ has finite dimension.*

Let $f(\lambda)$ be the spectral function of U so that $f(\lambda) = \lim_n f_n(\lambda)$ for every $\lambda \in \sigma(T)$. Thus if $\mu = f(\lambda) \neq 0$ we have, by 2.28, $\mu \in \sigma(U)$ and by Riesz's result μ is isolated in $\sigma(U)$. The proof now proceeds as it did in the preceding theorem except that we use 2.29 instead of 2.24.

2.34. THEOREM. *If T^n is compact then every spectral point $\mu \neq 0$ of T is a pole of $R_\xi(T)$.*

This is an immediate corollary of 2.32.

2.35. *The Fredholm theory.* If $U = f(T)$ is compact, where $f(\lambda)$ is regular in the circle $|\lambda| \leq |T|$ and has no roots other than $\lambda = 0$ in this circle, then 2.32 shows that all points $\mu \neq 0$ in $\sigma(T)$ are isolated in $\sigma(T)$ and are poles of $R_\xi(T)$. In view of 2.17 it is seen that all of the principal results of the Fredholm theory are valid for the operator T . In particular if T is a weakly compact operator in a Lebesgue space L it is known (Dunford and Pettis⁽¹⁰⁾ and Phillips⁽¹¹⁾) that T^2 is compact and so the Fredholm theory is valid for weakly compact operators in L .

3. Convergence to projections.

3.1. THEOREM. *Every projection $E \in \bar{\mathcal{R}}(T)$ is in $\mathcal{R}(T)$ and $E = E_\sigma[T]$, where σ is a spectral set of T and consists of all $\lambda \in \sigma(T)$ where the spectral function $f(\lambda)$ of E has the value 1. Thus $E = 0$ if and only if there are no such λ .*

The proof will require the following lemma.

3.2. LEMMA. *Every $\lambda \neq 0, 1$ is in the resolvent set $\rho(E)$ of a continuous projection E and*

$$R_\xi(E) = E/(\xi - 1) + (I - E)/\xi, \quad \xi \neq 0, 1.$$

Thus $E_1[E] = E$, $E_0[E] = I - E$, $\sigma(E) = 0$ if and only if $E = 0$; $\sigma(E) = 1$ if and only if $E = I$, and $\sigma(E) = \sigma_p(E)$.

The formula for $R_\xi(E)$ is verified by multiplying by $\xi I - E$ and the remaining conclusions follow immediately.

⁽¹⁰⁾ N. Dunford and B. J. Pettis, *Linear operations on summable functions*, Trans. Amer. Math. Soc. vol. 47 (1940) pp. 323-392.

⁽¹¹⁾ R. S. Phillips, *On linear transformations*, Trans. Amer. Math. Soc. vol. 48 (1940) pp. 516-541.

Theorem 3.1 now follows from 2.29. If we take σ of that theorem to be the point 1 then $E = E_1[E] = E_\sigma[T]$ where σ consists of all $\lambda \in \sigma(T)$ with $f(\lambda) = 1$.

3.3. LEMMA. *If $f_n \in \mathcal{F}(T)$ and if λ is a pole of order ν for $R_\xi(T)$ then the weak convergence of $f_n(T)$ implies the existence of the limits*

$$\lim_n f_n^{(j)}(\lambda), \quad j = 0, \dots, \nu - 1.$$

Furthermore if $f_n(T) \rightarrow f(T)$ weakly where $f \in \mathcal{F}(T)$ then

$$\lim_n f_n^{(j)}(\lambda) = f^{(j)}(\lambda), \quad j = 0, \dots, \nu - 1.$$

By 2.17, λ has index ν and so the lemma follows from 2.21 by the same argument used to derive 1.11 and 1.12 from the formula in 1.8.

3.4. THEOREM. *Let P be a polynomial not identically zero and let $f_n \in \mathcal{F}(T)$, $n = 1, 2, \dots$. If $f_n(T)$ converges to a projection E with $E\mathfrak{X} \subset \mathfrak{M}[P]$ then either there are no $\lambda \in \sigma(T)$ where $f_n(\lambda) \rightarrow 1$, in which case $E = 0$, or else:*

(1) *The set σ of points $\lambda \in \sigma(T)$ where $f_n(\lambda) \rightarrow 1$ consists of a finite number of poles $\lambda_1, \dots, \lambda_k$ of $R_\xi(T)$.*

(2) *If ν_1, \dots, ν_k are the orders of the poles $\lambda_1, \dots, \lambda_k$ then*

$$f_n(\lambda_i) \rightarrow 1, \quad f_n^{(j)}(\lambda_i) \rightarrow 0, \quad j = 1, \dots, \nu_i - 1; \quad i = 1, \dots, k.$$

(3) *For $i = 1, \dots, k$, λ_i is a root of $P(\lambda)$ and its multiplicity is at least ν_i .*

(4) $\prod_{i=1}^k (\lambda_i I - T)^{\nu_i} E = 0$.

If there are no $\lambda \in \sigma(T)$ where $f_n(\lambda) \rightarrow 1$ then $E = 0$ by 3.1. Now suppose the set σ of all $\lambda \in \sigma(T)$ where $f_n(\lambda) \rightarrow 1$ is not empty. We have, by 3.1, σ a spectral set of T and $E = E_\sigma[T]$. Since $E\mathfrak{X} \subset \mathfrak{M}[P]$ we have $P(T)E_\sigma(T) = 0$ and so (1), (3) and (4) follow from 2.19. Statement (2) follows from the preceding lemma.

3.5. LEMMA. *Let $f_n \in \mathcal{F}(T)$, $n = 1, 2, \dots$, and let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Then if $f_n(T)P(T) \rightarrow 0$ weakly and $\lim_n f_n(\lambda_i) = \alpha_i \neq 0$, $i = 1, \dots, k$, we have*

$$\mathfrak{M}_{\lambda_i}^{\nu_i} = \mathfrak{M}_{\lambda_i}^n, \quad \overline{\mathfrak{M}}_{\lambda_i}^{\nu_i} = \overline{\mathfrak{M}}_{\lambda_i}^n, \quad i = 1, \dots, k; \quad n \geq \nu_i.$$

If $\lambda_i \in \rho(T)$ the conclusion is obvious. If $\lambda_i \in \sigma(T)$ then $f_n(\lambda)$ is regular at λ_i and we may expand $P(\lambda)f_n(\lambda)$ about $\lambda = \lambda_i$,

$$P(\lambda)f_n(\lambda) = \sum_{m=0}^{\nu_i} \frac{(\lambda - \lambda_i)^m}{m!} [P(\lambda) \cdot f_n(\lambda)]_{\lambda=\lambda_i}^{(m)} + (\lambda - \lambda_i)^{\nu_i+1} g_{n,i}(\lambda).$$

If the function $g_{n,i}$ which is regular in a neighborhood of λ_i is defined by the

above formula on the whole of the T -admissible domain D_n upon which f_n is defined and regular, then $g_{n,i} \in \mathcal{F}(T)$. Now since $P^{(r)}(\lambda_i) = 0$, $r < \nu_i$, and $P^{(\nu_i)}(\lambda_i) \neq 0$ we have

$$[P(\lambda)f_n(\lambda)]_{\lambda=\lambda_i}^{(m)} = \begin{cases} \sum_{r=0}^m m C_r f_n^{(m-r)}(\lambda_i) P^{(r)}(\lambda_i) = 0, & m < \nu_i, \\ f_n(\lambda_i) P^{(\nu_i)}(\lambda_i), & m = \nu_i, \end{cases}$$

and so

$$(*) \quad P(T)f_n(T) = \frac{(T - \lambda_i I)^{\nu_i}}{\nu_i!} f_n(\lambda_i) P^{(\nu_i)}(\lambda_i) + (T - \lambda_i I)^{\nu_i+1} g_{n,i}(T).$$

Since $P(T)f_n(T) \rightarrow 0$ weakly, this equation shows that for $x \in \mathfrak{M}_{\lambda_i}^{\nu_i+1}$ we have

$$\alpha_i P^{(\nu_i)}(\lambda_i) x^*(T - \lambda_i I)^{\nu_i} x = 0, \quad x^* \in \mathfrak{X}^*,$$

and since $\alpha_i P^{(\nu_i)}(\lambda_i) \neq 0$ we have $x \in \mathfrak{M}_{\lambda_i}^{\nu_i}$. Thus $\mathfrak{M}_{\lambda_i}^{\nu_i} = \mathfrak{M}_{\lambda_i}^{\nu_i+1} = \mathfrak{M}_{\lambda_i}^n$, $n \geq \nu_i$. Since $P(T)f_n(T) \rightarrow 0$ weakly, (*) gives

$$\nu_i!(T - \lambda_i I)^{\nu_i+1} g_{n,i}(T) \rightarrow -\alpha_i P^{(\nu_i)}(\lambda_i)(T - \lambda_i I)^{\nu_i} \text{ weakly,}$$

and since $\alpha_i P^{(\nu_i)}(\lambda_i) \neq 0$ we see that every $x \in \mathfrak{M}_{\lambda_i}^{\nu_i}$ is the weak limit of a sequence $x_n \in \mathfrak{M}_{\lambda_i}^{\nu_i+1}$. Thus $\mathfrak{M}_{\lambda_i}^{\nu_i} \subset \overline{\mathfrak{M}_{\lambda_i}^{\nu_i+1}}$ and hence

$$\mathfrak{M}_{\lambda_i}^{\nu_i+1} \subset \mathfrak{M}_{\lambda_i}^{\nu_i} \subset \overline{\mathfrak{M}_{\lambda_i}^{\nu_i+1}}, \quad \overline{\mathfrak{M}_{\lambda_i}^{\nu_i+1}} = \overline{\mathfrak{M}_{\lambda_i}^{\nu_i}}.$$

Now since $(\lambda_i I - T)P(T)f_n(T) \rightarrow 0$ weakly we conclude, by what has just been proved, that $\overline{\mathfrak{M}_{\lambda_i}^{\nu_i+2}} = \overline{\mathfrak{M}_{\lambda_i}^{\nu_i+1}}$. A repetition of this argument shows that $\overline{\mathfrak{M}_{\lambda_i}^n} = \overline{\mathfrak{M}_{\lambda_i}^{\nu_i}}$ for all $n \geq \nu_i$.

3.6. THEOREM. Let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Let $f_n \in \mathcal{F}(T)$ satisfy

- (1) $f_n(\lambda_i) \rightarrow 1$, $f_n^{(j)}(\lambda_i) \rightarrow 0$, $i = 1, \dots, k$; $j = 1, \dots, \nu_i - 1$,
- (2) $P(T)f_n(T) \rightarrow 0$.

Then the following statements are equivalent.

- (3) $f_n(T) \rightarrow E$, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$.
- (4) Each λ_i , $i = 1, \dots, k$, is either in $\rho(T)$ or else a pole of $R_i(T)$.
- (5) For $i = 1, \dots, k$, λ_i is either in $\rho(T)$ or else a pole of $R_i(T)$ of order at most ν_i .
- (6) $\mathfrak{M}[P] \oplus \mathfrak{M}[P] = \mathfrak{X}$, $\mathfrak{M}[P]$ is closed.
- (7) $\mathfrak{M}_{\lambda_i}^{\nu_i+1}$, $i = 1, \dots, k$, is closed.

Furthermore when the limit E exists the set $\sigma = (\lambda_1, \dots, \lambda_k)\sigma(T)$ is a spectral set of T and

$$E = E_\sigma[T], \quad \mathfrak{M}[P] = \mathfrak{X}_\sigma, \quad \mathfrak{M}[P] = \mathfrak{X}_{\sigma'}.$$

In view of 3.4 statement (3) implies (4) and (4) in view of 3.5 and 2.17 implies (5). From 2.23 it is seen that (5) and (6) are equivalent. Also 2.17

shows that (5) implies (7). Now assume (7). Lemma 3.5 gives

$$\mathfrak{N}_{\lambda_i}^{\nu_i+1} = \overline{\mathfrak{N}_{\lambda_i}^{\nu_i}} \supset \mathfrak{N}_{\lambda_i}^{\nu_i} \supset \mathfrak{N}_{\lambda_i}^{\nu_i+1},$$

so that $\mathfrak{N}_{\lambda_i}^{\nu_i} = \mathfrak{N}_{\lambda_i}^n$ for $n \geq \nu_i$. Since $\mathfrak{N}_{\lambda_i}^{2\nu_i} = \mathfrak{N}_{\lambda_i}^{\nu_i}$ we have for every $x \in \mathfrak{X}$ a $z \in \mathfrak{N}_{\lambda_i}^{\nu_i}$ such that $(\lambda_i I - T)^{\nu_i} x = (\lambda_i I - T)^{\nu_i} z$, that is $x - z \in \mathfrak{M}_{\lambda_i}^{\nu_i}$ and so

$$(*) \quad \mathfrak{X} = \mathfrak{M}_{\lambda_i}^{\nu_i} + \mathfrak{N}_{\lambda_i}^{\nu_i} \text{ and } \mathfrak{N}_{\lambda_i}^{\nu_i} \text{ is closed.}$$

Now if $\lambda_i \in \sigma(T)$, f_n is regular at λ_i and so for λ near λ_i

$$f_n(\lambda) = \sum_{j=0}^{\nu_i-1} \frac{f_n^{(j)}(\lambda_i)}{j!} (\lambda - \lambda_i)^j + (\lambda - \lambda_i)^{\nu_i} g_{n,i}(\lambda).$$

If $g_{n,i}(\lambda)$ is defined by the above equation for all λ in the domain D_n where f_n is defined and regular then $g_{n,i} \in \mathcal{F}(T)$ and

$$f_n(T) = \sum_{j=0}^{\nu_i-1} \frac{f_n^{(j)}(\lambda_i)}{j!} (T - \lambda_i I)^j + (T - \lambda_i I)^{\nu_i} g_{n,i}(T).$$

Thus for $x \in \mathfrak{M}_{\lambda_i}^{\nu_i}$ we have, using (1),

$$f_n(T)x = \sum_{j=0}^{\nu_i-1} \frac{f_n^{(j)}(\lambda_i)}{j!} (T - \lambda_i I)^j x \rightarrow x.$$

Now if $P_i(\lambda) = P(\lambda)/(\lambda - \lambda_i)^{\nu_i}$ we see from Lemma 1.4 that $P_i(T)x \neq 0$ providing $0 \neq x \in \mathfrak{M}_{\lambda_i}^{\nu_i}$. Thus if $x = (\lambda_i I - T)^{\nu_i} y \in \mathfrak{M}_{\lambda_i}^{\nu_i}$ we have $f_n(T)x \rightarrow x$, $P_i(T)f_n(T)x = P(T)f_n(T)y \rightarrow 0$ (by (2)) and so $P_i(T)x = 0$ and hence $x = 0$. This shows that $\mathfrak{N}_{\lambda_i}^{\nu_i} \cdot \mathfrak{M}_{\lambda_i}^{\nu_i} = 0$. Thus (*) gives

$$\mathfrak{X} = \mathfrak{M}_{\lambda_i}^{\nu_i} \oplus \mathfrak{N}_{\lambda_i}^{\nu_i} \text{ is closed.}$$

Thus 2.23 applied to the function $f(\lambda) = (\lambda_i - \lambda)^{\nu_i}$ shows us that λ_i is a pole of order at most ν_i for $R_i(T)$. Hence (7) implies (5). We have now shown the equivalence of (4), (5), (6), (7) and also the fact that (3) implies any one of them. We shall now show that these conditions imply (3). Assuming (5) we have by 2.21

$$f_n(T) = f_n(T)E_{\sigma'} + \sum_{\lambda_i \in \sigma} \sum_{j=1}^{m_i-1} \frac{(T - \lambda_i I)^j}{j!} f_n^{(j)}(\lambda_i) E_{\lambda_i}[T],$$

where $\sigma = (\lambda_1, \dots, \lambda_k)\sigma(T)$ and m_i is the order of $\lambda_i \in \sigma$ as a pole of $R_i(T)$. Now in $\mathfrak{X}_{\sigma'}$ the spectrum of T is σ' (2.12) and hence since $0 \notin P(\sigma')$ we see that $P(T)$ is a 1-1 bicontinuous map of $\mathfrak{X}_{\sigma'}$ into all of itself. Thus $f_n(T)E_{\sigma'} = f_n(T)P(T)P^{-1}(T)E_{\sigma'} \rightarrow 0$ (by (2)). The above formula for $f_n(T)$ now shows (using (1)) that $f_n(T) \rightarrow \sum_{\lambda_i \in \sigma} E_{\lambda_i}[T] = E_{\sigma}[T]$. The final statements of the theorem follow from 2.23.

3.7. THEOREM. Let $P(\lambda)$, $f_n(\lambda)$ be as in the preceding theorem and satisfy (1) and (2) of that theorem. For $i=1, \dots, k$ let $(\lambda_i I - T)$ take bounded closed sets into closed sets. Then statements (3), \dots , (7) of the preceding theorem are all true.

Since $(\lambda_i I - T)$ takes bounded closed sets into closed sets it may be proved (see Banach⁽¹²⁾) where the corresponding statement is proved for a compact T) that $(\lambda_i I - T)\mathfrak{X}$ is closed. Since $(\lambda_i I - T)^2\mathfrak{X} \subset (\lambda_i I - T)\mathfrak{X}$ a repetition of this principle gives the fact that $(\lambda_i I - T)^2\mathfrak{X}$ is closed, and so on, that is, $\mathfrak{N}_{\lambda_i}^n$ is closed for all $n=1, 2, \dots$, and thus (7) of 3.6 is satisfied.

3.8. THEOREM. Let $P(\lambda)$, $f_n(\lambda)$ be as in 3.6 and satisfy (1), (2) of that theorem. Suppose that 0 is not a root of $P(\lambda)$ and that T^n is compact for some integer n . Then (3), \dots , (7) of 3.6 are all true.

By 2.34 all spectral points $\lambda \neq 0$ of T are poles of $R_\xi(T)$ and so 3.8 follows from 3.6. Theorem 3.8 is also an immediate corollary of 3.7.

3.9. THEOREM. Let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{v_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Let $f_n \in \mathcal{F}(T)$ satisfy the following:

- (1) $f_n(\lambda_i) \rightarrow 1$, $f_n^{(j)}(\lambda_i) \rightarrow 0$, $i=1, \dots, k$; $j=1, \dots, v_i-1$.
- (2) $P(T)f_n(T) \rightarrow 0$ strongly.

Then the following conditions are equivalent.

- (3) $f_n(T) \rightarrow E$ strongly, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$.
- (4) $f_n(T)x$ is weakly compact, $x \in \mathfrak{X}$.
- (5) $\mathfrak{N}[P] \oplus \mathfrak{M}[P] = \mathfrak{X}$, $|f_n(T)| \leq M$.

As in the proof of 3.6 we have for $y \in \mathfrak{M}_{\lambda_i}^{v_i}$

$$f_n(T)y = \sum_{j=0}^{v_i-1} \frac{f_n^{(j)}(\lambda_i)}{j!} (T - \lambda_i I)^j y \rightarrow y,$$

and so 1.4 gives the fact that $f_n(T)y \rightarrow y$ for $y \in \mathfrak{M}[P]$. Also (2) shows that $f_n(T)y \rightarrow 0$ for $y \in \mathfrak{N}[P]$. Now if we assume (4) we see that $|f_n(T)| \leq M$ and hence $f_n(T)y \rightarrow 0$ for $y \in \overline{\mathfrak{N}}[P]$. Thus $\overline{\mathfrak{N}}[P] \cdot \mathfrak{M}[P] = 0$. Since $|f_n(T)| \leq M$ it is seen that $f_n(T)x$ converges for every x in the closure of $\mathfrak{N}[P] + \mathfrak{M}[P]$. It will now be shown that

$$(*) \quad \overline{\mathfrak{N}}[P] \oplus \mathfrak{M}[P] = \mathfrak{X}.$$

If this is not true there is an $x \in X$ and an $x^* \in X^*$ with $x^*x = 1$, $x^*\mathfrak{M}[P] = 0 = x^*\mathfrak{N}[P]$. By 1.4 there are points $x_i \in \mathfrak{M}[P_i]$ with $x = x_1 + \dots + x_k$ and, since $x^*x = 1$, not all x^*x_i are zero. Suppose $x^*x_1 \neq 0$ and pick a sequence n_i of integers so that $f_{n_i}(T)x_1 \rightarrow y_1$ weakly. Since $x_1 \in \mathfrak{N}[P_1]$ it is of the form $x_1 = P_1(T)z_1$ and $(\lambda_1 I - T)^{v_1}x_1 = P(T)z_1$. Thus $(\lambda_1 I - T)^{v_1}f_{n_i}(T)x_1 = f_{n_i}(T)P(T)z_1$

⁽¹²⁾ S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932; in particular Théoreme 11, p. 151.

$\rightarrow 0$ which shows that $(\lambda_1 I - T)^{r_1} y_1 = 0$, that is, $y_1 \in \mathfrak{M}_{\lambda_1}^{r_1} \subset \mathfrak{M}[P]$. Since $x^* \mathfrak{M}[P] = 0$ we have therefore

$$(**) \quad x^* f_{n_i}(T) x_1 \rightarrow 0.$$

Now it is clear that it is no loss of generality to assume that $\lambda_1 \in \sigma(T)$, for if $\lambda_1 \in \rho(T)$ hypothesis (2) remains valid if $P(T)$ is replaced by $P_1(T)$ and also $\mathfrak{M}[P] = \mathfrak{M}[P_1]$. Thus we shall assume $\lambda_1 \in \sigma(T)$ and hence $f_n(\lambda)$ is regular at $\lambda = \lambda_1$. As in the proof of 3.5 we have

$$f_n(T) = \sum_{j=0}^{r_1-1} \frac{(T - \lambda_1 I)^j}{j!} f_n^{(j)}(\lambda_1) + (T - \lambda_1 I)^{r_1} g_n(T)$$

and hence, since $(T - \lambda_1 I)^{r_1} x_1 = (-1)^{r_1} P(T) z_1$ and $x^* \mathfrak{M}[P] = 0$, we have from (1)

$$x^* f_n(T) x_1 = x^* \sum_{j=0}^{r_1-1} \frac{f_n^{(j)}(\lambda_1)}{j!} (T - \lambda_1 I)^j x_1 \rightarrow x^* x_1 \neq 0.$$

The fact contradicts (**) and proves (*). Thus on the basis of (4) we have shown that

$$Ex = \lim_n f_n(T) x, \quad x \in \mathfrak{X},$$

exists. Statement (2) shows that $P(T)E = 0$ and hence $E\mathfrak{X} \subset \mathfrak{M}[P]$. But, as was shown above, $f_n(T)x \rightarrow x$ for $x \in \mathfrak{M}[P]$ and hence $E^2x = \lim_n f_n(T)Ex = Ex$ so that $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$. Thus we have shown that (4) implies (3) and (5). It remains to be shown that (5) implies (3). But, as was shown at the beginning of the proof, (1), (2) and $|f_n(T)| \leq M$ imply the convergence of $f_n(T)x$ for every $x \in \overline{\mathfrak{M}[P]} + \mathfrak{M}[P]$ and hence $Ex = \lim_n f_n(T)x$ exists for every $x \in \mathfrak{X}$. The above argument shows that E is a projection of the whole space onto the manifold $\mathfrak{M}[P]$. This completes the proof of the theorem.

As Theorem 3.4 shows, the condition (1) of 3.6, that is, $f_n(\lambda_i) \rightarrow 1$, $f_n^{(j)}(\lambda_i) \rightarrow 0$, $1 \leq j \leq r_i - 1$, is a necessary condition for the uniform convergence $f_n(T) \rightarrow E$, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$ providing $P(\lambda)$ is the polynomial of smallest degree for which $P(T)f_n(T) \rightarrow 0$. As the following theorem shows, a completely analogous situation holds in the case of strong convergence.

3.10. THEOREM. *Let $P(\lambda)$ be a polynomial and $f_n \in \mathcal{F}(T)$ be such that*

(1) *$f_n(T) \rightarrow E$ strongly, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$.*

Let $P'(\lambda)$ be the polynomial of smallest degree such that $\mathfrak{M}[P] = \mathfrak{M}[P']$. Let $P''(\lambda)$ be the polynomial of smallest degree such that $P''(T)f_n(T) \rightarrow 0$ strongly. Then

(2) *$P'(\lambda) \equiv P''(\lambda)$.*

(3) *If λ is a root of $P'(\lambda)$ of multiplicity ν then λ is in the point spectrum of T and has index ν and*

$$f_n(\lambda) \rightarrow 1, \quad f_n^{(j)}(\lambda) \rightarrow 0, \quad j = 1, \dots, \nu - 1.$$

Let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu_i}$, $P'(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu'_i}$, $P''(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu''_i}$. Clearly $0 \leq \nu'_i \leq \nu_i$ and since (1) implies $P(T)f_n(T) \rightarrow 0$ strongly we see that $0 \leq \nu''_i \leq \nu_i$. Also since, in the sense of strong convergence, we have $P'(T)f_n(T) \rightarrow P'(T)E = 0$ we see that $0 \leq \nu''_i \leq \nu'_i \leq \nu_i$. Let μ_i be the index of λ_i . By 1.4 we have $\mathfrak{M}_{\lambda_1}^{\nu_1} \oplus \cdots \oplus \mathfrak{M}_{\lambda_k}^{\nu_k} = \mathfrak{M}_{\lambda_1}^{\nu'_1} \oplus \cdots \oplus \mathfrak{M}_{\lambda_k}^{\nu'_k}$ from which it is clear that every root λ_i of $P'(\lambda)$ has index $\mu_i \geq \nu'_i$ and hence every root of $P'(\lambda)$ is in the point spectrum. Since f_n is regular at such a root we may write

$$f_n(T) = \sum_{j=0}^{\nu'_i-1} \frac{f_n^{(j)}(\lambda_i)}{j!} (T - \lambda_i I)^j + (T - \lambda_i I)^{\nu'_i} g_{n,i}(T).$$

If $x \neq 0$ and $(\lambda_i I - T)x = 0$ we have $f_n(T)x = f_n(\lambda_i)x$ and since $x \in \mathfrak{M}_{\lambda_i}^{\nu'_i} \subset \mathfrak{M}[P]$ we have $f_n(T)x \rightarrow x$, that is $f_n(\lambda_i) \rightarrow 1$. If $\nu'_i > 1$ then there is an x with $(T - \lambda_i I)x \neq 0$, $(T - \lambda_i I)^2 x = 0$ so that $x \in \mathfrak{M}[P]$ and

$$f_n(T)x = f_n(\lambda_i)x + f_n^{(1)}(\lambda_i)(T - \lambda_i I)x,$$

which shows that $f_n^{(1)}(\lambda_i) \rightarrow 0$. This argument may be repeated to give

$$f_n^{(j)}(\lambda_i) \rightarrow 0, \quad j = 1, \dots, \nu'_i - 1.$$

Since $0 \leq \nu''_i \leq \nu'_i$ we see therefore that (1), (2), and (4) of Theorem 3.9 are satisfied with $P(\lambda)$ replaced by $P''(\lambda)$ and hence, by that theorem, $f_n(T)$ converges strongly to a projection of \mathfrak{X} on $\mathfrak{M}[P'']$. This shows that $\mathfrak{M}[P''] = \mathfrak{M}[P]$, which proves that $\nu'_i \leq \nu''_i$ and hence $P'(\lambda) \equiv P''(\lambda)$. It remains to be shown that the index μ_i of a root λ_i of P' is equal to its multiplicity ν'_i . We have already seen that $\mu_i \geq \nu'_i$, and 3.5 applied to the polynomial $P''(\lambda) = P'(\lambda)$ gives $\mu_i \leq \nu'_i$. This completes the proof of the theorem.

The discussion of weak convergence is almost identical with strong convergence, and so we shall not discuss it in full but merely state the following two theorems. These theorems may be proved by using the methods of proof employed in 3.9 and 3.10.

3.11. THEOREM. Let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Let $f_n \in \mathcal{F}(T)$ satisfy

(1) $f_n(\lambda_i) \rightarrow 1$, $f_n^{(j)}(\lambda_i) \rightarrow 0$, $i = 1, \dots, k$; $j = 1, \dots, \nu_i - 1$.

(2) $P(T)f_n(T) \rightarrow 0$ weakly.

Then the following assertions are equivalent.

(3) $f_n(T) \rightarrow E$ weakly, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$.

(4) $f_n(T)x$ is weakly compact, $x \in \mathfrak{X}$.

(5) $\overline{\mathfrak{M}[P] \oplus \mathfrak{M}[P]} = \mathfrak{X}$, $|f_n(T)| \leq M$.

3.12. THEOREM. Let $P(\lambda)$ be a polynomial and $f_n \in \mathcal{F}(T)$ be such that

(1) $f_n(T) \rightarrow E$ weakly, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}[P]$.

Let $P'(\lambda)$ be the polynomial of smallest degree such that $\mathfrak{M}[P] = \mathfrak{M}[P']$. Let

$P''(\lambda)$ be the polynomial of smallest degree such that $P''(T)f_n(T) \rightarrow 0$ weakly. Then

$$(2) P'(\lambda) \equiv P''(\lambda).$$

(3) If λ is a root of $P'(\lambda)$ of multiplicity ν then λ is in the point spectrum of T and has index ν and

$$f_n(\lambda) \rightarrow 1, \quad f_n^{(j)}(\lambda) \rightarrow 0, \quad j = 1, \dots, \nu - 1.$$

3.13. THEOREM. Let $P(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\nu_i}$ be a polynomial whose distinct roots are $\lambda_1, \dots, \lambda_k$. Let $f_n \in \mathcal{F}(T)$ be such that

$$(1) f_n(\lambda_i) \rightarrow 1, f_n^{(j)}(\lambda_i) \rightarrow 0, i = 1, \dots, k; j = 1, \dots, \nu_i - 1.$$

$$(2) f_n(T) \rightarrow E \text{ weakly, } E^2 = E, E\mathfrak{X} = \mathfrak{M}[P].$$

Then

$$(3) f_n(T) \rightarrow E \text{ strongly if and only if } P(T)f_n(T) \rightarrow 0 \text{ strongly.}$$

(4) $f_n(T) \rightarrow E$ uniformly if and only if $P(T)f_n(T) \rightarrow 0$ uniformly and $\mathfrak{N}[P]$ is closed.

This is an immediate consequence of Theorems 3.6, 3.9 and 3.11.

In the following discussion of almost everywhere convergence we suppose that \mathfrak{X} is composed of measurable scalar functions $x(t)$ defined for t in a set S . The functions $x \in \mathfrak{X}$ are measurable with respect to a completely additive non-negative set function Φ which is defined on a Borel field containing S . We suppose also that S is a denumerable union of sets of finite measure and that $x = 0$ if and only if $x(t) = 0$ almost everywhere. Addition and scalar multiplication of the functions $x(t)$ correspond to the similar operations in \mathfrak{X} . It is also assumed that if $x_n \rightarrow x$ in \mathfrak{X} and $x_n(t) \rightarrow y(t)$ almost everywhere then $x(t) = y(t)$ almost everywhere. For a linear operator U in \mathfrak{X} we shall write $U(x, t)$ for the value of the function Ux at the point t .

The following theorem of Banach⁽¹³⁾ is fundamental.

3.14. THEOREM. Let U_n be a sequence of linear operators in \mathfrak{X} satisfying

$$(1) \text{ For every } x \in \mathfrak{X}, \limsup_n |U_n(x, t)| < \infty \text{ almost everywhere,}$$

(2) For every x in a fundamental set in \mathfrak{X} the $\lim U_n(x, t)$ exists almost everywhere.

Then

$$(3) \text{ For every } x \in \mathfrak{X} \text{ the } \lim_n U_n(x, t) \text{ exists almost everywhere.}$$

3.15. THEOREM. Let P, f_n, E be as in 3.9 and satisfy (1), (2), (4) of that theorem. Then

(1) for every $x \in \mathfrak{X}$, $\lim_n f_n(T)(x, t) = E(x, t)$ almost everywhere if and only if

⁽¹³⁾ S. Banach, *Sur la convergence presque partout des fonctionnelles linéaires*, Bull. Sci. Math. (2) vol. 50 (1926) pp. 36-43. This theorem of Banach has also been used by Yosida in generalizing the ergodic theorem of G. D. Birkhoff, see K. Yosida, *Ergodic theorems of the Birkhoff-Khinchine's type*, Jap. J. Math. vol. 17 (1940) pp. 31-36.

(2) for every $x \in \mathfrak{X}$, $\limsup_n |f_n(T)(x, t)| < \infty$ almost everywhere, and

(3) for every $x \in \mathfrak{X}$, $\lim_n P(T)f_n(T)(x, t)$ exists almost everywhere.

Statements (2) and (3) are clearly necessary for (1). Now conversely (3) shows that $f_n(T)(x, t)$ converges almost everywhere for $x \in \mathfrak{N}[P]$, and for $x \in \mathfrak{M}_{\lambda_i}^{\nu_i}$ we have

$$f_n(T)x = \sum_{j=0}^{\nu_i-1} \frac{f_n^{(j)}(\lambda_i)}{j!} (T - \lambda_i I)^j x,$$

and so $f_n(T)(x, t)$ converges almost everywhere for $x \in \mathfrak{M}[P] = \sum_{i=1}^k \mathfrak{M}_{\lambda_i}^{\nu_i}$. Thus 3.9(5) and 3.14 give the desired conclusion.

THE ERGODIC THEOREM

The following theorems are obtained from the preceding ones by taking the polynomial $P(T)$ to be $I - T$, that is, we are here concerned with convergence to a projection of the whole space on the fixed points of T . For brevity we shall write \mathfrak{M} , \mathfrak{N} in place of $\mathfrak{M}[P]$, $\mathfrak{N}[P]$ so that \mathfrak{M} is the manifold of fixed points and $\mathfrak{N} = (I - T)\mathfrak{X}$. While the ergodic theorem is concerned with the convergence of $f_n(T)$ where $f_n(\lambda) = n^{-1} \sum_{\nu=0}^{n-1} \lambda^\nu$, we shall word the theorems for any sequence of functions $f_n \in \mathcal{F}(T)$ with $f_n(1) \rightarrow 1$. It should be noted that the condition $(I - T)f_n(T) \rightarrow 0$ in any of the various meanings of convergence reduces, for the particular sequence of averages mentioned above, to the condition $T^n/n \rightarrow 0$.

3.16. THEOREM. Let $f_n \in \mathcal{F}(T)$ satisfy $f_n(1) \rightarrow 1$ and $(I - T)f_n(T) \rightarrow 0$. Then the following statements are equivalent.

- (1) $f_n(T) \rightarrow E$, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}$.
- (2) The point $\lambda = 1$ is either in $\rho(T)$ or else a pole of $R_\xi(T)$.
- (3) The point $\lambda = 1$ is either in $\rho(T)$ or else a simple pole of $R_\xi(T)$.
- (4) $\mathfrak{N} \oplus \mathfrak{M} = \mathfrak{X}$, \mathfrak{N} is closed.
- (5) $(I - T)^2\mathfrak{X}$ is closed.

This⁽¹⁴⁾ is a corollary of 3.6.

3.17. THEOREM. Let $f_n \in \mathcal{F}(T)$ satisfy $f_n(1) \rightarrow 1$, $(I - T)f_n(T) \rightarrow 0$, and let $I - T$ take bounded closed sets into closed sets. Then statements (1), \dots , (5) of 3.16 are all true.

This is a corollary of 3.7.

3.18. THEOREM. Let $f_n \in \mathcal{F}(T)$ satisfy $f_n(1) \rightarrow 1$, $(I - T)f_n(T) \rightarrow 0$ and let T^n

⁽¹⁴⁾ In the original statement of this theorem the assertion was made that if $\lambda = 1$ is isolated in $\sigma(T)$ and if $|T^n| \leq M$ then $\lambda = 1$ is necessarily a simple pole of $R_\xi(T)$. We are indebted to Professor Einar Hille for detecting an error in the attempted proof of this statement.

be compact for some integer n . Then the statements (1), \dots , (5) of 3.16 are all true.

This is a corollary of 3.8.

3.19. THEOREM. Let $f_n \in \mathcal{F}(T)$ satisfy $f_n(1) \rightarrow 1$ and $(I - T)f_n(T) \rightarrow 0$ strongly. Then the following conditions are equivalent.

- (1) $f_n(T) \rightarrow E$ strongly, $E^2 = E$, $E\mathfrak{X} = \mathfrak{M}$.
- (2) $f_n(T)x$ is weakly compact, $x \in \mathfrak{X}$.
- (3) $\mathfrak{N} \oplus \mathfrak{M} = \mathfrak{X}$, $|f_n(T)| \leq \mathfrak{M}$.

This follows from 3.9.

Similar theorems concerning weak convergence and almost everywhere convergence may be obtained from 3.10, \dots , 3.15, and these are left to the reader.

It should be mentioned however in connection with the ergodic theorem of G. D. Birkhoff⁽¹⁵⁾, which is concerned with the almost everywhere convergence of $n^{-1} \sum_{\nu=0}^{n-1} x(\phi^\nu t)$, that the condition (3) of 3.15 is redundant. That is, even with no assumption concerning the measure preserving character of the point map ϕ other than the assertion that the strong (that is mean) theorem holds, we can say that $n^{-1} \sum_{\nu=0}^{n-1} x(\phi^\nu t)$ converges almost everywhere if and only if

$$(*) \quad \limsup_n \left| \frac{1}{n} \sum_{\nu=0}^{n-1} x(\phi^\nu t) \right| < \infty \text{ almost everywhere.}$$

To see this note that (*) implies

$$(**) \quad \limsup_n |x(\phi^n t)/n| < \infty \text{ almost everywhere.}$$

Clearly $x(\phi^n t)/n \rightarrow 0$ for a bounded function $x(t)$ and hence 3.14 shows that $x(\phi^n(t))/n \rightarrow 0$ almost everywhere for every summable function $x(t)$. Thus 3.15 shows that (*) is necessary and sufficient for

$$(***) \quad \lim_n \frac{1}{n} \sum_{\nu=0}^{n-1} x(\phi^\nu t) = E(x, t) \text{ almost everywhere.}$$

Hence, one important problem remaining is to determine the conditions on the point map ϕ which are necessary and sufficient for (*) and hence for (***). This problem will be discussed elsewhere.

⁽¹⁵⁾ G. D. Birkhoff, *Proof of the ergodic theorem*, Proc. Nat. Acad. Sci. U. S. A. vol. 17 (1931) pp. 656-660.