Journal of the Mathematical Society of Japan Vol. 4, Nos. 3~4, December, 1952.

On the perturbation theory of closed linear operators.

By Tosio KATO

(Received September 1, 1952)

The perturbation theory of linear operators has been developed by several authors. The most complete results heretofore obtained by Rellich and others[1] are mainly concerned with the "regular" perturbation of self-adjoint operators of a Hilbert space, while some attempts[2] have also been made towards the treatment of "non-regular" cases which are no less important in applications.

Recently another generalization of the theory was given by Sz.-Nagy[3]. By his elegant and powerful method of contour integration, he has been able to transfer most of the theorems for self-adjoint operators to a wider class of closed linear operators of a general Banach space.

In the meantime the present writer was studying the same problem independently and published his main results in Japanese language[4]. It now turned out[5] that there are considerable coincidences between the results as well as methods of Sz.-Nagy and those of the writer.

The purpose of the present paper is to give a further development of the theory based on the fundamental results of Sz.-Nagy and the writer. An important part will also be played in § 2 by a generalization of a method which the writer[6] used in the proof of the adiabatic theorem of quantum mechanics.

It will be pointed out that the perturbation theory of general closed linear operators is not only a generalization of that of self-adjoint operators, but the full significance of the latter is realized only in the light of the former. This is due to the fact that, whereas the function-theoretical behaviour of the eigenvalues and eigenvectors is completely revealed only when we consider the parameter ε as a complex variable, an operator $T(\varepsilon)$ regular in ε cannot in general be self-adjoint or even normal for all values of ε of a complex domain.

We shall see in particular that an essential improvement of the estimation of the convergence radii for eigenvalues and eigenvectors is attained through these considerations.

§ 1. Regularity of the subspace.

Throughout the present paper we follow the definitions and notations of Sz.-Nagy[3]. According to him we consider a closed linear operator T_0 with domain $\mathfrak{D}$ dense in a complex Banach space $\mathfrak{B}$ and with range in $\mathfrak{B}$. We assume that its spectrum $\sigma(T_0)$ consists of two parts σ_0, σ_0' such that a closed rectifiable curve C can be drawn in the resolvent set $\rho(T_0)$ with σ_0 in its interior and σ_0' in its exterior. We now consider the "perturbed" operator

$$T(\varepsilon)=T_0+\varepsilon T_1+\varepsilon^2 T_2+\cdots, \tag{1.1}$$

where T_k's are linear operators with the same domain $\mathfrak{D}$ as T_0 and with ranges in $\mathfrak{B}$. They are assumed to satisfy the inequalities

$$\| T_k f \| \leqq p^{k-1}(a \| f \| + b \| T_0 f \|) \qquad (k=1, 2, \cdots). \tag{1.2}$$

The parameter ε is assumed to be either real or complex. But as $\mathfrak{B}$ is a complex Banach space, we can always extend (1.1) to complex values of ε even if it is initially defined only for real ε. *Thus we may hereafter assume ε to be complex without loss of generality.* This enables us to make use of various theorems of function theory and leads to considerable simplifications and improvements of the results.

It has been shown[3] that the resolvent $R_z(\varepsilon)=[T(\varepsilon)-zI]^{-1}$ with z on C is expressible as a power series of ε absolutely convergent in the circle

$$| \varepsilon | < (p+\alpha)^{-1}, \tag{1.3}$$

where

$$\alpha=aM+bN, \quad M=\operatorname*{Max}_{z \in C} \| R_z(0) \|, \quad N=\operatorname*{Max}_{z \in C} \| T_0 R_z(0) \|. \tag{1.4}$$

In what follows the set (1.3) will be called the *fundamental demain* of ε-plane and denoted by D_0. ε is assumed to belong to D_0 unless the contrary is positively stated.

It has also been shown[3)] that the resolvent set of $T(\varepsilon)$ contains the curve C if ε lies in D_0 and that the spectrum of $T(\varepsilon)$ is separated by C into the interior and exterior parts with the corresponding subspaces $\mathfrak{M}(\varepsilon)$ and $\mathfrak{M}'(\varepsilon)$ respectively. The corresponding projection $P(\varepsilon)$ onto $\mathfrak{M}(\varepsilon)$ is also expressible as a power series of ε convergent in D_0:

$$P(\varepsilon)=\sum_{n=0}^{\infty}\varepsilon^n P_n\,. \tag{1.5}$$

Thus $P(\varepsilon)$ is a regular analytic function[7)] of ε in the fundamental domain D_0. In particular it is continuous in D_0, and it follows that *the dimension m of* $\mathfrak{M}(\varepsilon)$ *is constant throughout* D_0. For, by virtue of the uniform continuity of $P(\varepsilon)$ in each closed subset of D_0, any two points ε', ε'' of D_0 can be joined by a chain $\varepsilon'=\varepsilon_0, \varepsilon_1,\cdots, \varepsilon_n=\varepsilon''$ such that $\|P(\varepsilon_{k-1})-P(\varepsilon_k)\|<1$ $(k=1, 2, \cdots, n)$ and hence $\dim \mathfrak{M}(\varepsilon_0)=\dim \mathfrak{M}(\varepsilon_1)=\cdots=\dim \mathfrak{M}(\varepsilon_n)$ by a lemma of Sz.-Nagy[3)].

§ 2. A regular mapping of $\mathfrak{M}(0)$ onto $\mathfrak{M}(\varepsilon)$.

THEOREM 1. *There is an operator $U(\varepsilon)$ defined for each ε of D_0 with the following properties:*

i) *$U(\varepsilon)$ and its inverse $U^{-1}(\varepsilon)$ are bounded linear operators with domain $\mathfrak{B}$ and range $\mathfrak{B}$;*

ii) *$U(\varepsilon)$ and $U^{-1}(\varepsilon)$ are regular analytic in D_0;*

iii) $P(\varepsilon)=U(\varepsilon)\,P(0)\,U^{-1}(\varepsilon)$, $\quad P(0)=U^{-1}(\varepsilon)\,P(\varepsilon)\,U(\varepsilon)$.

Thus $U(\varepsilon)$ maps $\mathfrak{M}(0)$ onto $\mathfrak{M}(\varepsilon)$ in a one-to-one fashion.

PROOF. I. Since $P(\varepsilon)$ is a projection, we have $P^2(\varepsilon)=P(\varepsilon)$ and hence by differentiation

$$P(\varepsilon)\,P'(\varepsilon)+P'(\varepsilon)\,P(\varepsilon)=P'(\varepsilon)\,, \tag{2.1}$$

where $'$ means $d/d\varepsilon$. Multiplication by $P(\varepsilon)$ from left and right yields

$$P(\varepsilon)\,P'(\varepsilon)\,P(\varepsilon)=0\,. \tag{2.2}$$

We now define the operator

$$Q(\varepsilon)=P'(\varepsilon)\,P(\varepsilon)-P(\varepsilon)\,P'(\varepsilon)\,. \tag{2.3}$$

$Q(\varepsilon)$ as well as $P(\varepsilon)$ and $P'(\varepsilon)$ is a bounded linear operator and regular analytic in ε. We note the following relations which are direct consequences of (2.1), (2.2) and (2.3):

$$P(\varepsilon)\,Q(\varepsilon)=-P(\varepsilon)\,P'(\varepsilon)\,,\quad Q(\varepsilon)\,P(\varepsilon)=P'(\varepsilon)\,P(\varepsilon)\,,$$
$$Q(\varepsilon)\,P(\varepsilon)-P(\varepsilon)\,Q(\varepsilon)=P'(\varepsilon)\,. \tag{2.4}$$

II. Next consider the differential equations

$$X'(\varepsilon)=Q(\varepsilon)\,X(\varepsilon)\,,\ Y'(\varepsilon)=-Y(\varepsilon)\,Q(\varepsilon) \tag{2.5}$$

for unknown operators $X(\varepsilon)$ and $Y(\varepsilon)$. Since these are "linear" differential equations with regular analytic coefficients, they have regular analytic solutions uniquely determined by the initial values $X(0)$, $Y(0)$[8]. Let $X(\varepsilon)=U(\varepsilon)$, $Y(\varepsilon)=V(\varepsilon)$ be the solutions with the initial values $U(0)=V(0)=I$. Then it follows from the uniqueness property that arbitrary solutions of (2.5) are given by

$$X(\varepsilon)=U(\varepsilon)\,X(0)\,,\qquad Y(\varepsilon)=Y(0)\,V(\varepsilon)\,. \tag{2.6}$$

Now we have

$$[V(\varepsilon)\,U(\varepsilon)]'=V'(\varepsilon)\,U(\varepsilon)+V(\varepsilon)\,U'(\varepsilon)=-V(\varepsilon)\,Q(\varepsilon)\,U(\varepsilon)$$
$$+V(\varepsilon)\,Q(\varepsilon)\,U(\varepsilon)=0$$

so that $V(\varepsilon)\,U(\varepsilon)=I$ identically. Similarly we have

$$[U(\varepsilon)\,V(\varepsilon)]'=Q(\varepsilon)\,[U(\varepsilon)\,V(\varepsilon)]-[U(\varepsilon)\,V(\varepsilon)]\,Q(\varepsilon)\,.$$

This shows that $U(\varepsilon)\,V(\varepsilon)$ also satisfies a "linear" differential equation with the initial value $U(0)\,V(0)=I$. But as the constant operator I satisfies the same equation, we must have $U(\varepsilon)\,V(\varepsilon)=I$ by virtue of the uniqueness of the solution. Thus we have shown

$$U(\varepsilon)\,V(\varepsilon)=V(\varepsilon)\,U(\varepsilon)=I\,. \tag{2.7}$$

This implies that the inverse $U^{-1}(\varepsilon)$ of $U(\varepsilon)$ exists and coincides with $V(\varepsilon)$, proving the assertions i) and ii).

III. Next we consider the operator $P(\varepsilon)\,U(\varepsilon)$. We have

$$[P(\varepsilon)\,U(\varepsilon)]'=P'(\varepsilon)\,U(\varepsilon)+P(\varepsilon)\,U'(\varepsilon)$$
$$=[P'(\varepsilon)+P(\varepsilon)\,Q(\varepsilon)]\,U(\varepsilon)=Q(\varepsilon)\,P(\varepsilon)\,U(\varepsilon) \tag{2.8}$$

by (2.5) and (2.4). This shows that $X(\varepsilon)=P(\varepsilon)\,U(\varepsilon)$ is also a solution of the first equation of (2.5) with the initial value $X(0)=P(0)$. Therefore we must have $P(\varepsilon)\,U(\varepsilon)=U(\varepsilon)\,P(0)$ by (2.6). In the same way we can show that $V(\varepsilon)\,P(\varepsilon)=P(0)\,V(\varepsilon)$, thus completing the proof of iii). Incidentally we note the following relation obtained by taking the adjoint of the last equation:

$$P^*(\varepsilon)\,V^*(\varepsilon)=V^*(\varepsilon)\,P^*(0)\,, \tag{2.9}$$

where $V^*(\varepsilon)$ as well as $P^*(\varepsilon)$ is a regular analytic function of $\bar{\varepsilon}$ in the fundamental domain $\overline{D_0}=D_0$.—

For later use we shall obtain a majorant of $U(\varepsilon)\,P(0)\,f$ where f is an arbitrary element of $\mathfrak{B}$. Since $Q(\varepsilon)\,P(\varepsilon)=P'(\varepsilon)\,P(\varepsilon)$ by (2.4), $Q(\varepsilon)$ in the right side of (2.8) can be replaced by $P'(\varepsilon)$. Then we can replace $P(\varepsilon)\,U(\varepsilon)$ of both sides by $U(\varepsilon)\,P(0)$ according to iii), Theorem 1. In this way we obtain

$$[U(\varepsilon)\,P(0)f\,]'=P'(\varepsilon)\,[U(\varepsilon)\,P(0)f\,]\,,$$

where

$$P'(\varepsilon)=\sum_{n=1}^{\infty} n\varepsilon^{n-1}\,P_n\,.$$

It follows easily that the power series of $U(\varepsilon)\,P(0)f$ is majorized by the expression

$$\|\,P(0)f\,\|\exp\left(\sum_{n=1}^{\infty}\int_0^{\varepsilon} n\varepsilon^{n-1}\,\|\,P_n\,\|\,d\varepsilon\right)=\|\,P(0)f\,\|\exp\left(\sum_{n=1}^{\infty}\varepsilon^n\,\|\,P_n\,\|\right).$$

Putting the inequality[9)]

$$\|\,P_n\,\|\leqq(2\pi)^{-1}\,|\,C\,|\,M\alpha\,(p+\alpha)^{n-1}\qquad(n=1,\,2,\cdots)$$

where $|\,C\,|$ is the length of C, we obtain a majorant of $U(\varepsilon)\,P(0)f$ in the following form

$$\|\,P(0)f\,\|\exp\frac{(2\pi)^{-1}\,|\,C\,|\,M\alpha\varepsilon}{1-(p+\alpha)\varepsilon}\,. \tag{2.10}$$

Finally it will be remarked that $U(\varepsilon)$ is *unitary* for real ε if $\mathfrak{B}$ is a Hilbert space and $T(\varepsilon)$ is self-adjoint or normal for real ε. To see this we have only to note that $P^*(\varepsilon)=P(\varepsilon)$ and hence that $P'^*(\varepsilon)=P'(\varepsilon)$, $Q^*(\varepsilon)=-Q(\varepsilon)$ for real ε. An inspection of the equations (2.5) and their adjoints shows that we must have $U^*(\varepsilon)=V(\varepsilon)$ for real ε. Since

$V(\varepsilon)=U^{-1}(\varepsilon)$, this proves the assertion.

§ 3. Perturbation of the spectrum for finite m.

In what follows we assume that the dimension m of $\mathfrak{M}(0)$ is finite, and choose a base $\{\psi_1, \psi_2, \cdots, \psi_m\}$ of $\mathfrak{M}(0)$. Then there is a base $\{\psi_1^*, \psi_2^*, \cdots\cdots, \psi_m^*\}$ of $\mathfrak{M}^*(0)$ such that[10]

$$(3.1) \qquad (\psi_k, \psi_j^*)=\delta_{jk}=\begin{cases}1 & (j=k)\\ 0 & (j\neq k)\end{cases}.$$

If we set

$$(3.2) \qquad \psi_k(\varepsilon)=U(\varepsilon)\,\psi_k,\ \psi_j^*(\varepsilon)=V^*(\varepsilon)\,\psi_j^* \qquad (j,\ k=1,\ 2,\cdots,\ m),$$

we have by Theorem 1, iii)

$$P(\varepsilon)\,\psi_k(\varepsilon)=P(\varepsilon)\,U(\varepsilon)\,\psi_k=U(\varepsilon)\,P(0)\,\psi_k=U(\varepsilon)\,\psi_k=\psi_k(\varepsilon)$$

and similarly $P^*(\varepsilon)\,\psi_j^*(\varepsilon)=\psi_j^*(\varepsilon)$ by (2.9). Hence $\psi_k(\varepsilon)\in\mathfrak{M}(\varepsilon)$ and $\psi_j^*(\varepsilon)\in\mathfrak{M}^*(\varepsilon)$. Moreover we have

$$(3.3) \qquad (\psi_k(\varepsilon), \psi_j^*(\varepsilon))=(U(\varepsilon)\,\psi_k, V^*(\varepsilon)\,\psi_j^*)=(V(\varepsilon)\,U(\varepsilon)\,\psi_k, \psi_j^*)$$
$$=(\psi_k, \psi_j^*)=\delta_{jk}$$

by (2.7) and (3.1). Since we know that the dimensions of $\mathfrak{M}(\varepsilon)$ and $\mathfrak{M}^*(\varepsilon)$ are equal to m (See § 1), these results show that $\{\psi_1(\varepsilon), \psi_2(\varepsilon),\cdots \psi_m(\varepsilon)\}$ and $\{\psi_1^*(\varepsilon),\cdots, \psi_m^*(\varepsilon)\}$ are bases of $\mathfrak{M}(\varepsilon)$ and $\mathfrak{M}^*(\varepsilon)$ respectively.

As has been shown by Sz.-Nagy[3], the spectrum of $T(\varepsilon)$ contains only a finite number of points inside the curve C. These points are eigenvalues of $T(\varepsilon)$ with the corresponding eigenvectors belonging to the subspace $\mathfrak{M}(\varepsilon)$, and the sum of their principal multiplicities[3] is just equal to m. At first we do not know whether or not the number of these points is independent of ε. In any case, however, let us choose one of them and denote it by $\lambda(\varepsilon)$, and let $\varphi(\varepsilon)$ be one of the eigenvectors associated with $\lambda(\varepsilon)$. Then we have

$$(3.4) \qquad T_0(\varepsilon)\,\varphi(\varepsilon)=\lambda(\varepsilon)\,\varphi(\varepsilon),$$

where $T_0(\varepsilon)\equiv T(\varepsilon)\,P(\varepsilon)$ is a regular analytic function of ε in D_0[3]. Since we know that $\varphi(\varepsilon)\in\mathfrak{M}(\varepsilon)$, we can write

$$(3.5) \qquad \varphi(\varepsilon)=\sum_{k=1}^{m} c_k(\varepsilon)\,\psi_k(\varepsilon), \qquad c_k(\varepsilon)=(\varphi(\varepsilon), \psi_k^*(\varepsilon)),$$

by virtue of (3.3). Putting (3.5) into (3.4) and taking the inner product of the resulting equation with $\psi_j^*(\varepsilon)$, we obtain

$$(3.6) \qquad \sum_{k=1}^{m} c_k(\varepsilon)\,(T_0(\varepsilon)\,\psi_k(\varepsilon),\,\psi_j^*(\varepsilon)) = \lambda(\varepsilon)\,c_j(\varepsilon) \quad (j=1,2,\cdots,m)\,.$$

Conversely (3.6) is also a sufficient condition for $\varphi(\varepsilon)$ and $\lambda(\varepsilon)$ to be a solution of (3.4). For (3.6) implies that the vector $[T_0(\varepsilon)-\lambda(\varepsilon)]\,\varphi(\varepsilon)$ is orthogonal to $\psi_j^*(\varepsilon)\,(j=1,\cdots,m)$; but as $[T_0(\varepsilon)-\lambda(\varepsilon)]\,\varphi(\varepsilon)$ belongs to $\mathfrak{M}(\varepsilon)$[3], it must be zero.

(3.6) is an ordinary eigenvalue problem for the m-dimensional vector $\{c_1(\varepsilon),\cdots,c_m(\varepsilon)\}$. Hence the eigenvalues $\lambda(\varepsilon)$ under consideration are identical with the roots of the secular equation

$$(3.7) \qquad \det\,[(T_0(\varepsilon)\,\psi_k(\varepsilon),\,\psi_j^*(\varepsilon)) - \lambda\delta_{jk}] = 0\,.$$

Since $T_0(\varepsilon)$ is regular analytic in D_0, the coefficients $(T_0(\varepsilon)\,\psi_k(\varepsilon),\,\psi_j^*(\varepsilon))$ are also regular analytic in D_0. Therefore the eigenvalues $\lambda(\varepsilon)$ consist of branches of one or several analytic functions of ε which have only a finite number of *algebraic singularities* in each closed subset of D_0. Furthermore, these analytic functions are *continuous and bounded* throughout D_0, for the coefficient of the highest power λ^m of (3.7) is equal to the constant $(-1)^m$ and, moreover, we know that $\lambda(\varepsilon)$'s lie inside the curve C for $\varepsilon \in D_0$.

Now it is clear that the number s of different eigenvalues is independent of ε except at those *exceptional values* of ε which are either singular points of the analytic functions $\lambda(\varepsilon)$ or for which some of the values of $\lambda(\varepsilon)$ are coincident. Of course there are only a finite number of such exceptional points in each closed subset of D_0. Thus we can denote by $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \cdots, \lambda_s(\varepsilon)$ these different eigenvalues of $T(\varepsilon)$ situated in the interior of C.

The behaviour of the operator $T(\varepsilon)$ in the subspace $\mathfrak{M}(\varepsilon)$ is completely described by the resolvent $R_z(\varepsilon)$. Since the only singular points (as a function of z) of $R_z(\varepsilon)$ inside the curve C are $\lambda_1(\varepsilon),\cdots,\lambda_s(\varepsilon)$, we obtain the expansion of $R_z(\varepsilon)$ into partial fractions in the following form[11]:

$$R_z(\varepsilon)=S_z(\varepsilon)+\sum_{k=1}^{s}\left\{\frac{P_k(\varepsilon)}{\lambda_k(\varepsilon)-z}+\frac{A_k(\varepsilon)}{[\lambda_k(\varepsilon)-z]^2}+\cdots+\frac{A_k^{m-1}(\varepsilon)}{[\lambda_k(\varepsilon)-z]^m}\right\} \tag{3.8}$$

at least except at the exceptional points of ε stated above. Here $S_z(\varepsilon)$ is given by

$$S_z(\varepsilon)=\frac{1}{2\pi i}\int_C \frac{R_{z'}(\varepsilon)}{z'-z}\,dz' \tag{3.9}$$

and is regular analytic for z inside C and $\varepsilon\in D_0$. $P_k(\varepsilon)$ is the projection associated[12)] with the eigenvalue $\lambda_k(\varepsilon)$ and the following relations hold:

$$P_k(\varepsilon)\,P_j(\varepsilon)=\delta_{jk}\,P_k(\varepsilon)\,,\qquad \sum_{k=1}^{s}P_k(\varepsilon)=P(\varepsilon)\,. \tag{3.10}$$

If we denote by $\mathfrak{M}_k(\varepsilon)$ the range of $P_k(\varepsilon)$ and by m_k its dimension, we have

$$\mathfrak{M}(\varepsilon)=\mathfrak{M}_1(\varepsilon)+\cdots\cdots+\mathfrak{M}_s(\varepsilon)\qquad\text{(direct sum)},\qquad m=m_1+\cdots\cdots+m_s\,. \tag{3.11}$$

That m_k are independent of ε will be shown soon below. $A_k(\varepsilon)$ have the following properties:

$$A_k(\varepsilon)=-[T_0(\varepsilon)-\lambda_k(\varepsilon)]\,P_k(\varepsilon)\,,\qquad A_k^{m_k}(\varepsilon)=0\,. \tag{3.12}$$

Hence the expression in { } of (3.8) has actually not more than m_k terms.

Let us consider the properties of $P_k(\varepsilon)$ and $A_k(\varepsilon)$ as functions of ε. We first note that $P_k(\varepsilon)$ is regular analytic at each point ε_0 which is not an exceptional point described above. For, since $\lambda_k(\varepsilon)$ is then an isolated eigenvalue of $T(\varepsilon)$ for every ε of a small neighbourhood of ε_0, we can apply to it our results heretofore obtained; we have only to replace T_0 by $T(\varepsilon_0)$, σ_0 by the set composed of a single point $\lambda_k(\varepsilon_0)$, the fundamental domain D_0 by a small neighbourhood $D(\varepsilon_0)$ of ε_0. Then $P(\varepsilon)$ is replaced by $P_k(\varepsilon)$, thus proving that $P_k(\varepsilon)$ is regular analytic in $D(\varepsilon_0)$ and that m_k is constant there. (3.12) then shows that $A_k(\varepsilon)$ is also regular analytic in $D(\varepsilon_0)$. By the process of analytic continuation, it is seen that all $P_k(\varepsilon)$ and $A_k(\varepsilon)$ are branches of respective analytic functions with branch points in common with $\lambda_k(\varepsilon)$. This follows from the fact that $R_z(\varepsilon)$ and $S_z(\varepsilon)$ in (3.8) are regular throughout D_0. By analytic continuation it is also seen that m_k is constant throughout D_0 except for the exceptional values of ε stated above.

To investigate more completely the behaviour of $P_k(\varepsilon)$ and $A_k(\varepsilon)$ in the neighbourhood of an *exceptional point* ε_0, we shall determine a base of $\mathfrak{M}_k(\varepsilon)$. We first note that a necessary and sufficient condition that a $f \in \mathfrak{M}(\varepsilon)$ belong to $\mathfrak{M}_k(\varepsilon)$ is given by[3]

$$[T_0(\varepsilon)-\lambda_k(\varepsilon)]^m f=0\,. \tag{3.13}$$

On setting

$$f=\sum_{l=1}^{m} c_l\, \psi_l(\varepsilon) \tag{3.14}$$

and proceeding in the same way as we deduced (3.6), we obtain

$$\sum_{l=1}^{m} c_l\, ([T_0(\varepsilon)-\lambda_k(\varepsilon)]^m\, \psi_l(\varepsilon), \psi_j^*(\varepsilon))=0 \qquad (j=1, 2, \cdots, m)\,. \tag{3.15}$$

By what is just stated, these linear equations for $c_1, c_2, \cdots, c_m$ must have the rank $m-m_k$ at least for sufficiently small $|\varepsilon-\varepsilon_0|$ and $\varepsilon \neq \varepsilon_0$, for ε is then certainly not an exceptional point. The coefficients of c_l in (3.15) are analytic there with at most an algebraic singularity at $\varepsilon=\varepsilon_0$; hence we can determine a set of m_k independent[13] solutions in such a way that all components c_l are analytic with at most an algebraic singularity at $\varepsilon=\varepsilon_0$. On putting these c_l into (3.14), we obtain a base $\{f_1(\varepsilon), \cdots, f_{m_k}(\varepsilon)\}$ of $\mathfrak{M}_k(\varepsilon)$, each $f_l(\varepsilon)$ being analytic with at most an algebraic singularity at $\varepsilon=\varepsilon_0$.

In quite the same way we can determine a base $\{f_1^*(\bar{\varepsilon}), \cdots, f_{m_k}^*(\varepsilon)\}$ of $\mathfrak{M}_k^*(\varepsilon)$, the range of $P_k^*(\varepsilon)$, such that each $f_l^*(\varepsilon)$ is analytic in $\bar{\varepsilon}$ with at most an algebraic singularity at $\bar{\varepsilon}=\bar{\varepsilon}_0$. Moreover we may assume that

$$(f_l(\varepsilon), f_p^*(\varepsilon))=\delta_{lp} \qquad (l, p=1, 2, \cdots, m_k)\,, \tag{3.16}$$

for the algebraic nature of the singularity is not lost in the process of biorthogonalization.

We can now express $P_k(\varepsilon)$ in terms of these bases $f_l(\varepsilon)$ and $f_p^*(\varepsilon)$. For any $f \in \mathfrak{B}$ we have

$$P_k(\varepsilon)f = \sum_{l=1}^{m_k} (f, f_l^*(\varepsilon)) f_l(\varepsilon)$$

by virtue of (3.16), and this shows that $P_k(\varepsilon)$ has at most an algebraic singularity at $\varepsilon=\varepsilon_0$. Then it follows from (3.12) that the same is also true for $A_k(\varepsilon)$.

We summarize our results as

THEOREM 2. *Let the dimension m of $\mathfrak{M}(0)$ be finite. Then the spectrum of $T(\varepsilon)$ inside the curve C consists of a finite number s of eigenvalues $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \cdots, \lambda_s(\varepsilon)$. The set of functions $\lambda_1(\varepsilon), \cdots, \lambda_s(\varepsilon)$ comprise the total branches of one or several analytic functions which are continuous and bounded in the fundamental damain D_0 and which possess only a finite number of algebraic singularities in each closed subset of D_0. Except at the values of ε which are either branch points of $\lambda_k(\varepsilon)$ or at which the value of $\lambda_k(\varepsilon)$ coincides with some other ones $\lambda_j(\varepsilon)$, the principal multiplicity m_k of each $\lambda_k(\varepsilon)$ is constant, and we have the decomposition* (3.8) *of the resolvent $R_z(\varepsilon)$, where $S_z(\varepsilon)$ is regular analytic for $\varepsilon \in D_0$, $P_k(\varepsilon)$ and $A_k(\varepsilon)$ are branches of analytic functions with only algebraic singularities at most at the exceptional points just described. $P_k(\varepsilon)$ are projections with ranges $\mathfrak{M}_k(\varepsilon)$ which are the principal subspaces corresponding to the respective eigenvalues $\lambda_k(\varepsilon)$, and the relations* (3.10), (3.11) *and* (3.12) *hold.*

Remark 1. Whereas $\lambda_k(\varepsilon)$ have no other singularities than branch points and are continuous even at such points, $P_k(\varepsilon)$ and $A_k(\varepsilon)$ are not necessarily continuous there and, moreover, may have other singularities at points where $\lambda_k(\varepsilon)$ are regular but some of their values are coincident. This is seen from the examples given below.

Remark 2. If $s=1$ (no splitting of eigenvalue!) $\lambda_1(\varepsilon)$ has no branch point and hence must be regular throughout D_0. Since we know that the same is true for $P_1(\varepsilon)=P(\varepsilon)$, $A_1(\varepsilon)$ is also regular by (3.12).

Remark 3. If $\mathfrak{B}$ is a Hilbert space and $T(\varepsilon)$ is self-adjoint or normal for real ε, all $P_k(\varepsilon)$ are orthogonal projections for real ε. Hence follows that $P_k(\varepsilon)$ have no branch point on the real axis, for it is easily seen[14] that $P_k^*(\varepsilon)=P_k(\varepsilon)$ cannot hold for both positive and negative values of $\varepsilon-\varepsilon_0$ if ε_0 is a real algebraic branch point of $P_k(\varepsilon)$.

Moreover, since $\|P_k(\varepsilon)\|=1$ holds for real ε, $P_k(\varepsilon)$ cannot have a pole on the real axis. Hence $P_k(\varepsilon)$ must be regular for real values of ε. Then $\lambda_k(\varepsilon)$ too must be regular for real ε, for a branch point of $\lambda_k(\varepsilon)$ should be also a branch point of $P_k(\varepsilon)$. Finally $A_k(\varepsilon)$ vanish[15] for a normal operator $T(\varepsilon)$. Thus we have

$$R_z(\varepsilon)=S_z(\varepsilon)+\sum_{k=1}^{s}[\lambda_k(\varepsilon)-z]^{-1}P_k(\varepsilon)$$

where $\lambda_k(\varepsilon)$, $P_k(\varepsilon)$ and $S_z(\varepsilon)$ are regular for real ε. In this way we have obtained again the main results of the perturbation theory of self-adjoint operators due to Rellich and others[1]. It will be noted that $\lambda_k(\varepsilon)$ may well have non-real singularities.

Example 1[16]. Let $\mathfrak{B}$ be two-dimensional and let

$$T(\varepsilon)=\begin{pmatrix}0 & 1\\ \varepsilon & 0\end{pmatrix}.$$

Then we have

$$R_z(\varepsilon)=\frac{P_1(\varepsilon)}{\lambda_1(\varepsilon)-z}+\frac{P_2(\varepsilon)}{\lambda_2(\varepsilon)-z}, \qquad \lambda_1(\varepsilon)=\varepsilon^{\frac{1}{2}},\quad \lambda_2(\varepsilon)=-\varepsilon^{\frac{1}{2}},$$

$$P_1(\varepsilon)=\frac{1}{2}\begin{pmatrix}1 & \varepsilon^{-\frac{1}{2}}\\ \varepsilon^{\frac{1}{2}} & 1\end{pmatrix}, \qquad P_2(\varepsilon)=\frac{1}{2}\begin{pmatrix}1 & -\varepsilon^{-\frac{1}{2}}\\ -\varepsilon^{\frac{1}{2}} & 1\end{pmatrix}$$

for $\varepsilon\neq 0$ and

$$R_z(0)=\frac{P}{\lambda_0-z}+\frac{A}{(\lambda_0-z)^2}, \qquad \lambda_0=0, P=I, A=-\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$$

for $\varepsilon=0$. Thus $R_z(\varepsilon)$ takes on quite different forms for $\varepsilon\neq 0$ and $\varepsilon=0$.

Example 2. Let $\mathfrak{B}$ be as above and let

$$T(\varepsilon)=\begin{pmatrix}0 & 1\\ \varepsilon^2 & 0\end{pmatrix}.$$

Then we have the same expression for $R_z(\varepsilon)$ with

$$\begin{matrix}\lambda_1(\varepsilon)=\varepsilon,\\ \lambda_2(\varepsilon)=-\varepsilon,\end{matrix} \qquad P_1(\varepsilon)=\frac{1}{2}\begin{pmatrix}1 & \varepsilon^{-1}\\ \varepsilon & 1\end{pmatrix}, \qquad P_2(\varepsilon)=\frac{1}{2}\begin{pmatrix}1 & -\varepsilon^{-1}\\ -\varepsilon & 1\end{pmatrix}.$$

Thus $P_1(\varepsilon)$, $P_2(\varepsilon)$ are single-valued and yet have a pole at $\varepsilon=0$ where $\lambda_1(\varepsilon)$, $\lambda_2(\varepsilon)$ are regular.

Example 3[17]. Let $\mathfrak{B}$ be as above and let

$$T(\varepsilon)=\begin{pmatrix}0 & \varepsilon\\ 0 & 0\end{pmatrix}.$$

Then we have

$$R_z(\varepsilon)=\frac{P(\varepsilon)}{\lambda(\varepsilon)-z}+\frac{A(\varepsilon)}{[\lambda(\varepsilon)-z]^2}, \qquad \lambda(\varepsilon)=0,\ P(\varepsilon)=I,\ A(\varepsilon)=-T(\varepsilon).$$

Here we have $A(\varepsilon)\neq 0$ for $\varepsilon\neq 0$ and $A(0)=0$, in contrast to Example 1.

§ 4. Estimation of convergence radii and coefficients.

In this section we shall obtain some estimates[18] of the convergence radii and the coefficients of the eigenvalues and eigenvectors of $T(\varepsilon)$ as power series of ε. For simplicity we restrict ourselves to the case $m=1$. Then we have $s=1$ a fortiori, and it follows from Remark 2 of the preceding section that $\lambda_1(\varepsilon)\equiv\lambda(\varepsilon)$ and $P_1(\varepsilon)=P(\varepsilon)$ are regular analytic throughout the fundamental domain D_0. Thus the power series

$$\lambda(\varepsilon)=\sum_{n=0}^{\infty}\varepsilon^n\lambda_n, \qquad P(\varepsilon)=\sum_{n=0}^{\infty}\varepsilon^n P_n \tag{4.1}$$

are convergent in D_0, that is, for[19) 20)] $|\varepsilon|<(p+\alpha)^{-1}$.

Furthermore, since $\lambda(\varepsilon)$ lies in the interior of C for $\varepsilon\in D_0$, we have $|\lambda(\varepsilon)-\lambda_0|<\delta$, where $\delta=\mathrm{Max}\,|z-\lambda_0|$ for $z\in C$. It follows from Cauchy's inequality in function theory that[21)]

$$|\lambda_n|\leqq\delta(p+\alpha)^n \qquad (n=1,2,\cdots). \tag{4.2}$$

The vectors $\psi_1(\varepsilon)\equiv\psi(\varepsilon)$ and $\psi_1^*(\varepsilon)\equiv\psi^*(\varepsilon)$ constructed in § 3 are respectively the eigenvectors of $T(\varepsilon)$ and $T^*(\varepsilon)$ associated with $\lambda(\varepsilon)$ and $\overline{\lambda(\varepsilon)}$. Since we have

$$(\psi(\varepsilon),\psi^*(\varepsilon))=1 \tag{4.3}$$

by (3.3), $\psi(\varepsilon)$ is regular analytic and $\neq 0$ throughout D_0. Hence the expansion of the eigenvector $\psi(\varepsilon)$:

$$\psi(\varepsilon)=\sum_{n=0}^{\infty}\varepsilon^n\psi^{(n)} \tag{4.4}$$

is convergent also for[20) 22)] $|\varepsilon|<(p+\alpha)^{-1}$.

On setting $f=\psi^{(0)}$ in (2.10), we obtain a majorant of (4.4) and therefrom we can derive an estimate of $\| \psi^{(n)} \|$. Without aiming at the utmost accuracy, we note the following simple estimate obtained by further replacing (2.10) by its majorant

$$\omega\left(1-\frac{rM\alpha\varepsilon}{1-(p+\alpha)\varepsilon}\right)^{-1}=\omega\frac{1-(p+\alpha)\varepsilon}{1-(p+\alpha+rM\alpha)\varepsilon}$$

(where we have set $\omega=\| P(0)\psi^{(0)}\|=\|\psi^{(0)}\|$ and $|C|=2\pi r$):

$$\|\psi^{(n)}\| \leqq \omega rM\alpha(p+\alpha+rM\alpha)^{n-1} \qquad (n=1,2,\cdots)\text{[23]}. \tag{4.5}$$

In conclusion we note that the estimate of the convergence radius for $\lambda(\varepsilon)$ as given above is the best possible one. This is shown by the following

Example 4. Let $\mathfrak{B}$ be a two-dimensional unitary space and let

$$T(\varepsilon)=\begin{pmatrix}1 & \varepsilon\\ \varepsilon & -1\end{pmatrix},\quad T_0=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\quad T_1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad T_2=T_3=\cdots=0\,.$$

Then we have (we consider the eigenvalue $\lambda_0=1$ of T_0)

$$\begin{array}{lllll} p=0, & a=1, & b=0, & \lambda_0=1, & r=1,\\ M=1, & N=1, & \alpha=1, & (p+\alpha)^{-1}=1 & \end{array}$$

(C is chosen as a circle with center $\lambda_0=1$ and radius $r=1$). The exact eigenvalue is $\lambda(\varepsilon)=(1+\varepsilon^2)^{\frac{1}{2}}$, for which the convergence radius is just equal to $1=(p+\alpha)^{-1}$.

§ 5. General regular perturbation.

In the foregoing sections we started from the assumptions (1.1) and (1.2) for $T(\varepsilon)$. But this is only a special case of "regular" perturbations. Following the definition of Rellich[24] in the case of self-adjoint operators, we can define a non-bounded regular operator $T(\varepsilon)$ as follows. Let $T(\varepsilon)$ be a linear operator, depending on a complex parameter ε, with domain $\mathfrak{D}(\varepsilon)$ dense in a Banach space $\mathfrak{B}$ and with range in $\mathfrak{B}$. $T(\varepsilon)$ is said to be regular in a neighbourhood of $\varepsilon=0$ if the following conditions are fulfilled:

i) there is a bounded operator $W(\varepsilon)$ with domain $\mathfrak{B}$ and range

$\mathfrak{D}(\varepsilon)$ and which is regular analytic in a neighbourhood of $\varepsilon=0$;

ii) the operator $T(\varepsilon)\,W(\varepsilon)$ with domain $\mathfrak{B}$ is bounded and regular analytic in ε.

If we further assume that $T(0)$ is closed and has a non-empty resolvent set $\rho(T(0))$, we can show by the method of Rellich[24] that $T(\varepsilon)$ is also closed and that every point of $\rho(T(0))$ belongs to $\rho(T(\varepsilon))$, provided ε is sufficiently small. Then the argument of Sz.-Nagy[3] can be applied without change, and all the results of Sz.-Nagy and ours are valid for this more general case. It is not necessary to enter into these details here.

Department of Physics, University of Tokyo.

Bibliography

[1] Heinz, E.: Beiträge zur Störungstheorie der Spektralzerlegung, Math. Ann. **123** (1951), 415-438.

[2] Hille, E.: Functional analysis and semi-groups, New York, 1948.

[3] Kato, T.: On the convergence of the perturbation method, I, Progress of Theoretical Physics **4** (1949), 514-523.

[4] Kato, T.: II, ibid **5** (1950), 95-101 and 207-212.

[5] Kato, T.: Perturbation theory of linear operators (in Japanese), Sugaku (Mathematics) **2** (1950), 201-208.

[6] Kato, T.: On the adiabatic theorem of quantum mechanics, J. Phys. Soc. Japan **5** (1950), 435-439.

[7] Kato, T.: On the convergence of the perturbation method, J. Fac. Sci. Univ. Tokyo, Section I, vol. **6** (1951), 145-226.

[8] Kato, T.: Perturbation theory of semi-bounded operators, Math. Ann. (in press).

[9] Nagumo, M.: Einige alalytische Untersuchungen in linearen metrischen Ringen, Jap. J. Math. **13** (1936), 61-80.

[10] Rellich, F.: Störungstheorie der Spektralzerlegung, I. Mitt., Math. Ann. **113** (1937), 600-619.

[11] Rellich, F.: II. Mitt., ibid. **113** (1937), 677-685.

[12] Rellich, F.: III. Mitt., ibid. **116** (1939), 555-570.

[13] Rellich, F.: IV. Mitt., ibid. **117** (1940), 356-382.

[14] Rellich, F.: V. Mitt., ibid. **118** (1941/43), 462-484.

[15] Sz.-Nagy, B. v.: Spektraldarstellung linearer Transformationen des Hilbertschen Raumes, Erg. Math. **5** (1942), 416-496.

[16] Sz.-Nagy, B. v.: Perturbations des transformations autoadjointes dans l'espace de Hilbert, Comment. Math. Helv. **19** (1946/47), 347-366.

[17] Sz.-Nagy, B. v.: Perturbations des transformations linéaires fermées, Acta Sci. Math. Szeged **14** (1951), 125-137.

[18] Titchmarsh, E. C.: Some theorems on perturbation theory, Proc. Roy. Soc. **A200** (1949), 34-46.
[19] Titchmarsh, E. C.: II, ibid **A201** (1950), 473-479.
[20] Titchmarsh, E. C.: III, ibid. **A207** (1951), 321-328.

Notes

1) Rellich [10]—[14]; Sz.-Nagy [15], [16]; Heinz [1]; Kato [3], [7].
2) Titchmarsh [18]—[20]; Kato [4], [7], [8].
3) Sz.-Nagy [17].
4) Kato [5].
5) The writer is indebted to Prof. Rellich for a chance of seeing the paper of Prof. Sz.-Nagy.
6) Kato [6].
7) Except $T(\varepsilon)$ we have no occasion of considering a non-bounded operators. So all operators of the form $A(\varepsilon)$ are assumed to be bounded and have domain $\mathfrak{B}$, unless the contrary is positively stated.
8) This is proved by the method of successive approximation; there is no difficulty since the fundamental domain D_0 is simply connected.
9) See Eq. (20) of Sz.-Nagy [17].
10) This is implied by Eq. (11) of Sz.-Nagy [17]. $\mathfrak{M}^*(\varepsilon)$ is the range of $P^*(\varepsilon)$.
11) Cf. Nagumo [9]; Hille [2], Chap. V.
12) This means that the range $\mathfrak{M}_k(\varepsilon)$ of $P_k(\varepsilon)$ is the *principal subspace* corresponding to $\lambda_k(\varepsilon)$ in the sense of Sz.-Nagy [17].
13) Independent at least for sufficiently small $|\varepsilon-\varepsilon_0|$ and $\varepsilon \neq \varepsilon_0$.
14) Cf. Rellich [10].
15) $A_k(\varepsilon)=0$ holds at first for real ε; then it holds identically by analytic continuation.
16) This is the example a) of Sz.-Nagy [17].
17) This is the example c) of Sz.-Nagy [17].
18) For the application of these results to practical problems, see Kato [7], § 5, where they are applied, in particular, to the Mathieu equation and to the helium wave equation.
19) It should be noted that this result and (4.2) are valid even if $m>1$, provided that $s=1$
20) This estimate is simpler and more precise than the corresponding ones of Sz.-Nagy [16] and [17].
21) There seems to be no simple relation between (4.2) and the corresponding estimates of Sz.-Nagy [16] and [17]. However, (4.2) is more favourable at least if $p=0$.
22) It will be noted that $\|\psi(\varepsilon)\|=1$ for real ε if $T(\varepsilon)$ is self-adjoint or normal for real ε and $\|\psi(0)\|=1$. This follows from the fact that $U(\varepsilon)$ is unitary as we remarked at the end of § 2.
23) This is somewhat more precise than the corresponding estimates of Sz.-Nagy [16] and [17].
24) Rellich [12].