Return to the self-adjoint transformation.

By E. R. LORCH in New York.

We return once more to the theory of the self-adjoint transformation in Hilbert space. This subject which was born shortly after the turn of the century, has since then attracted wide attention, — particularly so during the last twenty years. The central fact is the structure theorem which asserts the following:

If H is a self-adjoint transformation in Hilbert space $\mathfrak{H}$, then there exists a resolution of the identity $E(\lambda)$ such that the structure of H is completely summarized by the formula

$$H=\int_{-\infty}^{\infty} \lambda \, d E(\lambda). \tag{1}$$

Many proofs have been given of this theorem. Some are concerned with the bounded case. Others apply to the general situation. Still a third variety considers that the unbounded case is best treated by first carrying through a complete discussion for the bounded transformation. We shall not analyse the methods of these proofs, which are after all well known to those interested in this domain.

In the pages which follow we set down a new proof of this fundamental theorem. Our approach is to attack the most general (unbounded) situation directly from the start and to assume no knowledge of transformation theory except the most trivial facts. We should like to believe that our method yields the final result considerably more rapidly than those heretofore advanced. We obtain formula (1) in the following form:

Theorem. *Let H be a self-adjoint transformation in Hilbert space $\mathfrak{H}$. Let $\{\lambda_n\}$ be a set of real numbers $n=0, \pm 1, \pm 2, \ldots$, such that*

a) *for all n, $\lambda_n > \lambda_{n-1}$;* b) $\lim_{n \to \infty} \lambda_n = \infty$; c) $\lim_{n \to -\infty} \lambda_n = -\infty$.

Then there exists in $\mathfrak{H}$ a set of closed linear manifolds $\{\mathfrak{M}_n\}$, $n=0, \pm 1, \pm 2, \ldots$, orthogonal in pairs, spanning $\mathfrak{H}$, and such that H is defined on $\mathfrak{M}_n$ and satisfies

$$\lambda_n I \leqq H \leqq \lambda_{n+1} I. \tag{2}$$

It will be seen that this theorem gives formula (1) directly. In the first place the manifolds $\mathfrak{M}_n$ define in an obvious way the projections $E(\lambda)$. Secondly, the inequality (2) coupled with elementary facts on orthogonality is precisely what makes it possible to define the integral of Riemann-Stieltjes type which is in (1). We remind the reader that $H \geqq \lambda I$ means $(Hf, f) \geqq \lambda (f, f)$.

Our methods are based on integrals around simple curves in the complex plane. The general form of these integrals is

$$K = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta I - H} d\zeta. \tag{3}$$

These expressions are strongly reminiscent of the Cauchy integral formula, except that in the integrand we find instead of the usual $(\zeta - a)^{-1}$ the operator $(\zeta I - H)^{-1}$. It is easy to show that since H is self-adjoint the above operator is bounded providing ζ is not real. This means that our integrals are improper since the curve C cuts the real axis. The reason for this may be put another way. If the contour curve over which the integration is performed lies exclusively in the resolvent set of H we have an operator which was first considered briefly by F. Riesz and subsequently was used by the author to study the reducibility of normed rings, to investigate the theory of their radical, and to define an operational calculus of operators in general vector spaces. In the present work, the path of integration crosses the spectrum of H and hence the behavior of the integrand must be subject to a careful examination. The existence of our integrals rests on very special properties possessed by self-adjoint transformations. The fundamental functional equations satisfied by the integrals (3) are proved with the help of the Neumann expansion for the resolvent and, more important, the functional equation of the resolvent. All that is required to carry through the proofs is a little patience to determine that the operations on iterated integrals are legitimate. One further point may deserve mention: The ideas which underlie the procedure are quite transparent in spite on the fact that at times they may be slightly submerged under a technique and notation which are evoked by the present subject-matter as well as our approach to it. The reader with experience in this subject will wish to omit the reading of sections I and II.

I. Definitions.

Let T be a linear transformation which is defined for every vector f belonging to some linear set Δ which is dense in $\mathfrak{H}$. Consider the set of all pairs $\{g, g^*\}$ such that $(Tf, g) = (f, g^*)$ for all $f \in \Delta$. The relation $g \to g^*$ defines a linear transformation T^*, $T^*g = g^*$, called the adjoint of T. If the domain of T^* is identical with that of T and if $T = T^*$ on this common domain, we say that T is self-adjoint.

A transformation T is said to be closed if it has the following property: Let $\{f_n\}$ be a sequence of vectors belonging to the domain Δ of T. Suppose that $f_n \to f$ and $Tf_n \to g$. Then $f \in \Delta$ and $Tf = g$. It is very easy to see that every adjoint transformation T^* is closed. In particular, all self-adjoint transformations are closed.

II. The resolvent.

Let H be self-adjoint with a domain of definition Δ. Let ζ be a complex number, $\zeta = \alpha + \beta i$, with $\beta \neq 0$. Then if we write $(\zeta I - H) g = f$ for any arbitrary $g \in \Delta$ we have

$$|f|^2 = |(\alpha I - H) g|^2 + \beta^2 |g|^2,$$

hence $|g| \leq |f|/|\beta|$. In particular, if $g \neq 0$, then $f \neq 0$, hence ζ is not in the point spectrum of H. If the totality of elements f obtained by allowing g to vary over Δ were not dense in H, then $\bar{\zeta} = \alpha - \beta i$ would be in the point spectrum of $H^* = H$. Finally, the elements f actually fill $\mathfrak{H}$. For if $\{g_n\}$ is a sequence such that $g_n \in \Delta$ and that $f_n \to f$, then $\{g_n\}$ is a convergent sequence since $|\beta| |g_n - g_m| \leq |f_n - f_m|$. Since H is a closed transformation, the element g to which $\{g_n\}$ converges belongs to Δ and $(\zeta I - H) g = f$. If we consolidate these facts we obtain the critically important result that *if ζ is not a real number, the transformation $(\zeta I - H)^{-1}$ exists and is a bounded linear transformation with a bound equal to or less than $|\beta|^{-1}$ where β is the imaginary part of ζ.*

If ζ is a fixed point in the complex plane which is not on the real axis and if ξ is a point near ζ, then we have

(4) $(\xi I - H)^{-1} = (\zeta I - H)^{-1} + (\zeta - \xi)(\zeta I - H)^{-2} + (\zeta - \xi)^2 (\zeta I - H)^{-3} + \ldots$

This expansion is valid for all ξ such that $|\xi - \zeta| < 1/|(\zeta I - H)^{-1}|$. The relation (4) implies that the bound of $(\xi I - H)^{-1}$ is a continuous function of ξ for ξ lying off the real axis.

The functional equation for the resolvent, valid for non-real ζ and ξ, is

(5) $$(\zeta I - H)^{-1} - (\xi I - H)^{-1} = (\xi - \zeta)(\zeta I - H)^{-1} (\xi I - H)^{-1}.$$

This may be established by multiplying both sides by $(\zeta I - H) = (\zeta - \xi) I + (\xi I - H)$.

III. The point spectrum.

If for a pair $\{\lambda, f\}$ with $f \neq 0$, we have $Hf = \lambda f$, then λ is said to be in the point spectrum of H. If the totality of such vectors f spans $\mathfrak{H}$, then H is said to have a pure point spectrum. Now due to the fact that characteristic vectors f_1 and f_2 which correspond to distinct characteristic values λ_1 and λ_2 are orthogonal to each other, the fundamental formula (1) is trivial for an H with pure point spectrum. In the methods which we employ in what follows the possible existence of the point spectrum causes annoyance.

For this reason, we shall remove it. If the space $\mathfrak{H}$ is separable, this is not necessary since the point spectrum contains at most a denumerable number of real numbers and in any neighborhood one may always cut the real axis at a point not a characteristic value of H. If $\mathfrak{H}$ is not separable, the procedure described below makes it unnecessary later always to amend our statements. We shall prove the following:

If H is a self-adjoint transformation in $\mathfrak{H}$, and if $\mathfrak{M}$ is the closed linear manifold spanned by the totality of characteristic vectors of H, then $\mathfrak{M}^{\perp}$ the orthogonal complement of $\mathfrak{M}$ reduces H, and in the space $\mathfrak{M}^{\perp}$, H is a self-adjoint transformation and has no point spectrum.

Proof: Let $\{\varphi_\alpha\}$ be an orthonormal set which spans $\mathfrak{M}$ and whose members are characteristic vectors of H. Let f be an arbitrary element in the domain Δ of H. Suppose the expansion for f in terms of the set $\{\varphi_\alpha\}$ is $f \sim \Sigma a_\alpha \varphi_\alpha$. Let $f^* = \Sigma a_\alpha \varphi_\alpha$. Then since H is defined for f and for each φ_α it is defined for $f - \Sigma' a_\alpha \varphi_\alpha$ where Σ' denotes a sum containing only a finite number of terms. Since $H^2(\Sigma' a_\alpha \varphi_\alpha)$ is well defined and since

$$(H(f-\Sigma' a_\alpha \varphi_\alpha),\ H(\Sigma' a_\alpha \varphi_\alpha)) = (f-\Sigma' a_\alpha \varphi_\alpha,\ H^2(\Sigma' a_\alpha \varphi_\alpha)) = 0,$$

we have

$$|Hf|^2 = |H(f-\Sigma' a_\alpha \varphi_\alpha)|^2 + |H(\Sigma' a_\alpha \varphi_\alpha)|^2.$$

The sum $\Sigma a_\alpha \varphi_\alpha$ contains at most denumerably many non-zero terms. To simplify notation we shall write $\Sigma a_\alpha \varphi_\alpha = \lim_{n\to\infty} \Sigma_1^n a_\alpha \varphi_\alpha$. The equation established immediately above shows that the sequence $\{H(\Sigma_1^n a_\alpha \varphi_\alpha)\}$ converges. Since H is a closed transformation, f^* and also $f-f^* = f - \Sigma a_\alpha \varphi_\alpha$ belong to the domain of H. Clearly $f-f^* \in \mathfrak{M}^{\perp}$.

Thus every element f in Δ is the sum in a unique way of an element g in $\mathfrak{M}$ and h in $\mathfrak{M}^{\perp}$ where both g and h belong to Δ: $f = g + h$. Clearly, since $Hg \in \mathfrak{M}$ and since $(g, Hh) = (Hg, h) = 0$ we see that $Hh \in \mathfrak{M}^{\perp}$. Since the set Δ of elements f is dense in $\mathfrak{H}$, the set of elements h is dense in $\mathfrak{M}^{\perp}$.

It remains to show that in $\mathfrak{M}^{\perp}$, H is self-adjoint. If there exists a pair $\{k, k^*\}$ in $\mathfrak{M}^{\perp}$ such that for every h in Δ and in $\mathfrak{M}^{\perp}$, $(Hh, k) = (h, k^*)$, then for every f in Δ, $f = g + h$, $g \in \mathfrak{M}$, and $h \in \mathfrak{M}^{\perp}$, we have $(H(g+h), k) = (Hh, k) = (h, k^*) = (g+h, k^*)$. Since H is self-adjoint, $k \in \Delta$ and $Hk = k^*$. This shows that considered in the space $\mathfrak{M}^{\perp}$, H is self adjoint. That H has no point spectrum in $\mathfrak{M}^{\perp}$ is obvious.

To establish (1) for the transformation H operating in $\mathfrak{H}$ it is sufficient to establish it for H in $\mathfrak{M}$ and for H in $\mathfrak{M}^{\perp}$. Since the case H in $\mathfrak{M}$ is trivial, there remains only the case H in $\mathfrak{M}^{\perp}$. In consequence, *from now on, by virtue of the preceding result, we shall assume that H has no point spectrum.*

IV. On a class of improper integrals.

We now introduce the integrals which are the key to the structure theorem. The path of integration is a curve C which lies in the complex plane. For the sake of convenience, C is assumed to be smooth except possibly for a few corners. The curve cuts the axis of reals at a non-zero angle at two point λ and μ. In addition, we shall assume that C is symmetric in the axis of reals. The letters m and n represent positive integers. The integral which concerns us is

$$K_{\lambda\mu}(m,n) = \frac{1}{2\pi i}\int_C (\zeta-\lambda)^m(\mu-\zeta)^n(\zeta I-H)^{-1}\,d\zeta. \tag{6}$$

The integral may be improper at $\zeta=\lambda$ or $\zeta=\mu$. We examine its behavior as follows. We alter C by eliminating from it two short segments which enclose these critical points. The resulting path of integration will be called D. Now the integral (6) over the path D is well defined and exists in the uniform topology of operators This is due to the fact that $(\zeta I-H)^{-1}$ is a continuous function of ζ (see equation (4) in section III). Now since $m\geq 1$ and since C cuts the axis of reals at a non-zero angle the operator $(\zeta-\lambda)^m(\zeta I-H)^{-1}$ is bounded near $\zeta=\lambda$. This type of argument proves that the integral (6) converges in the uniform topology and hence represents a bounded operator. Since the path of integration is symmetric about the real axis and due to the special structure of the integrand, it is clear that $K_{\lambda\mu}(m,n)$ is self-adjoint. Also by virtue of equation (4) in section II, it is clear that the value of the integral is not changed if the path is slightly deformed providing that the points λ and μ remain fixed. We now list these and the other properties of this operator which are of importance to us.

The improper integral $K_{\lambda\mu}(m,n)$ converges in the uniform topology and represents a bounded self-adjoint transformation. The value of the integral is not altered if the path of integration C is deformed slightly providing that the points $\zeta=\lambda$ and $\zeta=\mu$ remain on C.

The transformation $K_{\lambda\mu}(m,n)$ satisfies the further conditions:

a) $K_{\lambda\mu}(m,n)\cdot K_{\lambda\mu}(m',n') = K_{\lambda\mu}(m+m',n+n')$.

b) *If the intervals (λ,μ) and (λ',μ') have no points in common (or at most one end-point in common), then*

$$K_{\lambda\mu}(m,n)\cdot K_{\lambda'\mu'}(m',n')=0.$$

c) *The transformation $H-\lambda I$ is defined for every element in the range of $K_{\lambda\mu}(m,n)$ and furthermore*

$$(H-\lambda I)\cdot K_{\lambda\mu}(m,n) = K_{\lambda\mu}(m+1,n).$$

A similar statement may be made for the operator $\mu I-H$.

d) *For the elements in the range of* $K_{\lambda\mu}(m,n)$, H *is a bounded self-adjoint transformation which satisfies the inequalities* $\lambda I \leq H \leq \mu I$.

We start by proving a). We are concerned with the product of two integrals. We may assume that the associated paths of integration C and C' are such that one lies entirely within the other except at the two points λ and μ. Then using the functional equation for the resolvent (section II, equation (5)) we express the product of these integrals as an iterated integral. Making use of the Cauchy integral formula of ordinary function theory on the innermost integral we obtain a).

The proof of b) is similar to that of a). Since the two paths of integration lie each outside the other, the product of the two integrals yields zero.

To prove c) we note that the operator H is obviously defined for any element $(\zeta I-H)^{-1}f$. Since $H-\lambda I=(H-\zeta I)+(\zeta-\lambda)I$, the product of $K_{\lambda\mu}(m,n)$ by $H-\lambda I$ gives two integrals, one of which is $K_{\lambda\mu}(m+1,n)$ while the other is zero by the Cauchy theorem. Careful examination of this argument shows that use is made of the fact that H is a closed transformation. One may see this as follows. The integral $K_{\lambda\mu}(m,n)$ may be approximated by a finite Riemann sum which we designate by $\Sigma(m,n)$. If we multiply $\Sigma(m,n)$ by $H-\lambda I$ we obtain essentially an approximating sum $\Sigma(m+1,n)$ for $K_{\lambda\mu}(m+1,n)$. The result c) is obtained by taking limits in the sense of integration and using the closure of H.

To prove d), we note first that we may write the following relation for inner products in $\mathfrak{H}$; here f is an arbitrary vector in $\mathfrak{H}$:

$$(7)\quad \begin{aligned}((H-\lambda I)\,K_{\lambda\mu}(m,n)f,\ K_{\lambda\mu}(m,n)f) &= (K_{\lambda\mu}(m+1,n)f,\ K_{\lambda\mu}(m,n)f) = \\ &= (K_{\lambda\mu}(m,n)\,K_{\lambda\mu}(m+1,n)f,f) = (K_{\lambda\mu}(2m+1,2n)f,f).\end{aligned}$$

In writing this down, use has been made of relations a) and c). We shall show that there exists a self-adjoint transformation $L_{\lambda\mu}(2m+1,2n)$ such that

$$(8)\qquad L^2_{\lambda\mu}(2m+1,2n)=K_{\lambda\mu}(2m+1,2n).$$

This will prove that the inner product in equation (7) is non-negative and hence that for $g=K_{\lambda\mu}(m,n)f$, $H\geq\lambda I$.

Consider to this effect the integral

$$(9)\qquad L_{\lambda\mu}(m,n)=\frac{1}{2\pi i}\int_C(\zeta-\lambda)^{m/2}(\mu-\zeta)^{n/2}(\zeta I-H)^{-1}d\zeta.$$

The square root of $(\zeta-\lambda)^m$ is understood to be that which is positive when $\zeta-\lambda$ is real and positive; that of $(\mu-\zeta)^n$ is positive when $\mu-\zeta$ is real and positive. By arguments similar to those used above, one may prove that the integral $L_{\lambda\mu}(m,n)$ exists and represents a bounded self-adjoint transformation; furthermore one may also establish equation (8). The integral (9) introduces a new feature in case $m=1$; for in that case $(\zeta-\lambda)^{m/2}(\zeta I-H)^{-1}$

may have an infinite bound at $\zeta=\lambda$. A routine argument disposes of this difficulty.

This concludes the proof of the fact that $H\geq\lambda I$. The steps which demonstrate that $H\leq\mu I$ are similar.

V. On the linear manifolds associated with $K_{\lambda\mu}(m,n)$.

Associated with every transformation $K_{\lambda\mu}(m,n)$ we consider two closed linear manifolds: the manifold of the zeros of the transformation and the manifold of the closure of the range of the transformation. It is an obvious fact that each is the orthogonal complement of the other. In this section we develop some properties of these manifolds.

Let the closure of the range of $K_{\lambda\mu}(m,n)$ be denoted by $\mathfrak{M}_{\lambda\mu}(m,n)$. Let the set of its zeros be denoted by $\mathfrak{N}_{\lambda\mu}(m,n)$. We prove:

a) *$\mathfrak{M}_{\lambda\mu}(m,n)$ is independent of m and n; that is, $\mathfrak{M}_{\lambda\mu}(m,n)=\mathfrak{M}_{\lambda\mu}(m',n')$. This common manifold is denoted by $\mathfrak{M}_{\lambda\mu}$.*

b) *If λ,μ,ν are real numbers subject to $\lambda<\mu<\nu$ then $\mathfrak{M}_{\lambda\mu}+\mathfrak{M}_{\mu\nu}=\mathfrak{M}_{\lambda\nu}$ where the '+' operation is performed in the sense of addition of linear manifolds.*

To prove a), consider $\mathfrak{N}_{\lambda\mu}(m,n)$. We have $(H-\lambda I)K_{\lambda\mu}(m,n)=$ $=K_{\lambda\mu}(m+1,n)$. Since H has no point spectrum, this means that $\mathfrak{N}_{\lambda\mu}(m,n)=\mathfrak{N}_{\lambda\mu}(m+1,n)$. This type of reasoning yields a).

Toward b): By a) we may assume that all integers m,n, etc. are equal to 1. Also since the point spectrum is absent we may without changing the character of $\mathfrak{M}_{\lambda\mu}$ assume that it is defined by an integral operator whose integrand is $(2\pi i)^{-1}(\zeta-\lambda)(\zeta-\mu)(\zeta-\nu)(\zeta I-H)^{-1}$. The same is true of $\mathfrak{M}_{\mu\nu}$ and $\mathfrak{M}_{\lambda\nu}$. We denote the three operators by $T_{\lambda\mu}$, $T_{\mu\nu}$, and $T_{\lambda\nu}$. Now it is clear that a possible path of integration which defines $T_{\lambda\nu}$ is the sum of the paths defining $T_{\lambda\mu}$ and $T_{\mu\nu}$. Thus $T_{\lambda\nu}=T_{\lambda\mu}+T_{\mu\nu}$. Also, we have $T_{\lambda\mu}\cdot T_{\mu\nu}=0$. These facts lead to the conclusion in b).

VI. Conclusion of proof.

We recall the method for proving our principal theorem. Let $\{\lambda_n\}$ be a monotone increasing set of real numbers such that $\lim_{n\to\infty}\lambda_n=\infty$ and $\lim_{n\to-\infty}\lambda_n=-\infty$. For every integer s and for some fixed m construct the transformation $K_{\lambda_s\lambda_{s+1}}(m,m)$. Determine the closure of its range $\mathfrak{M}_s$. The manifolds $\mathfrak{M}_s$ are orthogonal in pairs (in virtue of b) in section IV) and on $\mathfrak{M}_s$, $\lambda_s I\leq H\leq\lambda_{s+1}I$. To complete our proof it remains to show that the manifolds $\mathfrak{M}_s$ span $\mathfrak{H}$. This we proceed to do.

In the first place, the result b) in section V shows that instead of considering many manifolds, we need consider only one. Going beyond this, it is clear that we need consider only the single manifold $\mathfrak{P}_r$ which corresponds to the interval $(-r,r)$.

If the manifolds $\mathfrak{P}_r$, $0 < r < \infty$, do not span $\mathfrak{H}$, there is a non-zero element f orthogonal to each $\mathfrak{P}_r$. We shall explicitly assume that f is orthogonal to each $\mathfrak{P}_r$ and subsequently prove that $f=0$. Now $\mathfrak{P}_r$ is the closure of the range of $K_{-r,r}(1, 1)$. Since the latter operator is self-adjoint, we have $K_{-r,r}(1, 1)f=0$ for all values of r. Thus we have for all r

$$f=\frac{1}{2\pi i r^2}\int_C [(r^2-\zeta^2)\zeta^{-1}-(r^2-\zeta^2)(\zeta I-H)^{-1}]\,d\zeta\cdot f. \tag{10}$$

Here C is a circle with radius r and center the origin; the first term in the integrand yields f by Cauchy's theorem. Now a simple calculation shows that except for a constant factor the integrand in (10) is of the form $H(r^2-\zeta^2)\cdot\zeta^{-1}\cdot(\zeta I-H)^{-1}f$. If we remove the factor H from under the integral sign and perform the integration we obtain an element which we may denote by g_r. Hence equation (10) may be written in the form $f=Hg_r$. Now, by using (3) and standard techniques of evaluation of integrals, it is an easy matter to show that $|g_r|\leq 2r^{-1}|f|$. As $r\to\infty$, $g_r\to 0$ while $Hg_r\to f$. Since H is a closed transformation, $f=0$.

BARNARD COLLEGE, COLUMBIA UNIVERSITY.

(Received August 1, 1949.)