Stone's Theorem.

Math 212

November 28, 2000

Recall that if A is a self-adjoint operator we can form the one parameter group of unitary operators

$$U(t) = e^{iAt}$$

by virtue of a functional calculus which allows us to construct $f(A)$ for any bounded Borel function defined on $\mathbf{R}$ (if we use our first proof of the spectral theorem using the Gelfand representation theorem) or for any function holomorphic on $\mathrm{Spec}(A)$ if we use our second proof. In any event, the spectral theorem allows us to write

$$U(t) = \int_{-\infty}^{\infty} e^{it\lambda} dE_\lambda.$$

We called this assertion the first half of Stone's theorem. The second half (to be stated more precisely below) asserts the converse: that any one parameter group of unitary transformations can be written in either, hence both, of the above forms.

The idea that we will follow hinges on the following elementary computation

$$\int_0^\infty e^{(z+ix)t} dt = \frac{1}{z+ix} \quad \text{if } \mathrm{Re}\, z < 0$$

valid for any real number x. If we substitute A for x and write $U(t)$ instead of e^{iAt} this suggests that

$$(zI + iA)^{-1} = \int_0^\infty e^{zt} U(t) dt \quad \text{if } \mathrm{Re}\, z < 0.$$

Multiplying by i and setting $w = iz$would give

$$R(w, A) = -i \int_0^\infty e^{zt} U(t) dt \quad \text{if } \mathrm{Im}\, w < 0.$$

Replacing A by $-A$ or, what is the same thing, integrating from 0 to $-\infty$ will give us the resolvent for complex numbers with positive imaginary part. Our previous studies encourage us to believe that once we have found all these putative resolvents, it should not be so hard to reconstruct A.

This program works! But because of some of the subtleties involved in the definition of a self-adjoint operator, we will begin with an important theorem of von-Neumann which we will need, and which will also greatly clarify exactly what it means to be self-adjoint.

A second matter which will lengthen these proceedings is that while we are at it, we will prove a more general version of Stone's theorem valid in an arbitrary Frechet space and for "uniformly bounded semigroups" rather than unitary groups. Stone proved his theorem to meet the needs of quantum mechanics, where a unitary one parameter group corresponds, via *Wigner's theorem* to a one parameter group of symmetries of the logic of quantum mechanics. In more pedestrian terms, unitary one parameter groups arise from solutions of Schrodinger's equation. But many other important equations, for example the heat equations in various settings, require the more general result.

The treatment here will essentially follow that of Yosida, *Functional Analysis* especially Chapter IX.

Contents

1 von Neumann's Cayley transform. 2

1.1 An elementary example. 7

2 Equibounded semi-groups on a Frechet space. 8

2.1 The infinitesimal generator. 9

3 The differential equation 11

3.1 The resolvent. 12

3.1.1 Application to Stone's theorem. 14

3.2 Examples. 14

3.2.1 Translations. 14

3.2.2 The heat equation. 15

3.2.3 Bochner's theorem. 16

1 von Neumann's Cayley transform.

The group $Gl(2, \mathbf{C})$ of all invertible complex two by two matrices acts as "fractional linear transformations" on the plane: the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ sends } \quad z \mapsto \frac{az+b}{cz+d}.$$

Two different matrices M_1 and M_2 give the same fractional linear transformation if and only if $M_1 = \lambda M_2$ for some (non-zero complex) number λ as is clear from the definition. Since

$$\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix} = 2i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

the fractional linear transformations corresponding to $\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$ and $\begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$ are inverse to one another.

It is a theorem in the elementary theory of complex variables that fractional linear transformations are the only orientation preserving transformations of the plane which carry circles and lines into circles and lines. Even without this general theory, an immediate computation shows that $\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$ carries the (extended) real axis onto the unit circle,and hence its inverse carries the unit circle onto the extended real axis. ("Extended" means with the point ∞ added.)

We might think of (multiplication by) a real number as a self-adjoint transformation on a one dimensional Hilbert space, and (multiplication by) a number of absolute value one as a unitary operator on a one dimensional Hilbert space. This suggests in general that if A is a self adjoint operator, then $(H+iI)(H-iI)^{-1}$ should be unitary. In fact, we can be much more precise. First some definitions:

An operator U, possibly defined only on a subspace of H is called **isometric** if $\|Ux\| = \|x\|$ for all x in its domain of definition.

Recall that in order to define the adjoint T^* of an operator T it is necessary that $D(T)$ be dense in H. (Otherwise the equation $(Tx, y) = (x, T^*y)$ $\forall\ x \in D(T)$ does not determine T^*y.) A transformation T (in a Hilbert space H) is called **symmetric** if $D(T)$ is dense in H so T^* is defined and

$$D(T) \subset D(T^*) \text{ and } Tx = T^*x \quad \forall\ x \in D(T).$$

Another way of saying the same thing is T is symmetric if $D(T)$ is dense and

$$(Tx, y) = (x, Ty) \quad \forall\ x, y \in D(T).$$

A self-adjoint transformation is symmetric since $D(T) = D(T^*)$ is one of the requirements of being self-adjoint. Exactly how and why a symmetric operator can fail to be self-adjoint will be clarified in the ensuing discussion. All of the results of this section are due to von Neumann.

Theorem 1 *Let T be a closed symmetric operator. Then $(T+iI)x = 0$ implies that $x = 0$ for any $x \in D(T)$ so $(T+iI)^{-1}$ exists as an operator on its domain*

$$D\left[(T+iI)^{-1}\right] = \operatorname{im}(T+iI).$$

This operator is bounded on its domain and the operator

$$U_T := (T-iI)(T+iI)^{-1} \quad \text{with } D(U_T) = D\left[(T+iI)^{-1}\right] = \operatorname{im}(T+iI)$$

is isometric and closed. The operator $(I-U_T)^{-1}$ exists and

$$T = i(U_T+I)(U_T-I)^{-1}.$$

In particular, $D(T) = \operatorname{im}(I-U_T)$ is dense in H.

Conversely, if U is a closed isometric operator such that $\operatorname{im}(I-U)$ is dense in H then $T = i(U+I)(I-U)^{-1}$ is a symmetric operator with $U = U_T$.

Proof. For any $x \in D(T)$ we have

$$([T \pm iI]x, [T \pm iI]x) = (Tx, Tx) \pm (Tx, ix) \pm (ix, Tx) + (x, x).$$

The middle terms cancel because T is symmetric. Hence

$$\|[T \pm iI]x\|^2 = \|Tx\|^2 + \|x\|^2. \tag{1}$$

Taking the plus sign shows that $(T + iI)x = 0 \Rightarrow x = 0$ and also shows that $\|[T + iI]x\| \geq \|x\|$ so

$$\|[T + iI]^{-1}y\| \leq \|y\| \quad y \in [T + iI](D(T)).$$

If we write $x = [T + iI]^{-1}y$ then (1) shows that

$$\|U_T y\|^2 = \|Tx\|^2 + \|x\|^2 = \|y\|^2$$

so U_T is an isometry with domain consisting of all $y = (T + iI)x$, i.e. with domain $D([T + iI]^{-1})$.

We now show that U_T is closed. So we must show that if $y_n \to y$ and $z_n \to z$ where $z_n = U_T y_n$ then $y \in D(U_T)$ and $U_T y = z$. The y_n form a Cauchy sequence and $y_n = [T + iI]x_n$ since $y_n \in \mathrm{im}(T + iI)$. From (1) we see that the x_n and the Tx_n form a Cauchy sequence, so $x_n \to x$ and $Tx_n \to w$ which implies that $x \in D(T)$ and $Tx = w$ since T is assumed to be closed. But then $(T + iI)x = w + ix = y$ so $y \in D(U_T)$ and $w - ix = z = U_T y$. So we have shown that U_T is closed.

Subtract and add the equations

$$\begin{aligned} y &= (T + iI)x \\ U_T y &= (T - iI)x \quad \text{to get} \\ \frac{1}{2}(I - U_T)y &= ix \quad \text{and} \\ \frac{1}{2}(I + U_T)y &= Tx. \end{aligned}$$

The third equation shows that

$$(I - U_T)y = 0 \Rightarrow x = 0 \Rightarrow Tx = 0 \Rightarrow (I + U_T)y = 0$$

by the fourth equation. So

$$y = \frac{1}{2}([I - U_T]y + [I + U_t]y) = 0.$$

Thus $(I - U_T)^{-1}$ exists, and $(I - U_T)^{-1}y = 2ix$ from the third of the four equations above, and the last equation gives

$$Tx = \frac{1}{2}(I + U_T)y = \frac{1}{2}(I + U_T)(I - U_T)^{-1}2ix$$

or

$$T = i(I + U_T)(I - U_T)^{-1}$$

as required. Furthermore, every $x \in D(T)$ is in $\mathrm{im}(I - U_T)$. This completes the proof of the first half of the theorem.

Now suppose we start with an isometry U and suppose that $(I - U)y = 0$ for some $y \in D(U)$. Let $z \in \mathrm{im}(I - U)$ so $z = w - Uw$ for some w. We have

$$(y, z) = (y, w) - (y, Uw) = (Uy, Uw) - (y, Uw) = (Uy - y, Uw) = 0.$$

Since we are assuming that $\mathrm{im}(I - U)$ is dense in H, the condition $(y, z) = 0\ \forall\ z \in \mathrm{im}(I - U)$ implies that $y = 0$. Thus $(I - U)^{-1}$ exists, and we may define

$$T = i(I + U)(I - U)^{-1}$$

with

$$D(T) = D(I - U)^{-1} = \mathrm{im}(I - U)$$

dense in H. Suppose that $x = (I - U)u,\ y = (I - U)v\ \in D(H) = \mathrm{im}(I - U)$. Then

$$(Tx, y) = (i(I + U)u, (I - U)v) = i\left[(Uu, v) - (u, Uv)\right] + i\left[(u, v) - (Uu, Uv)\right].$$

The second expression in brackets vanishes since U is an isometry. So $(Tx, y) =$

$$i(Uu, v) - i(u, Uv) = (-Uu, iv) + (u, iUv) = ([I - U]u, i[I + U]v) = (x, Ty).$$

This shows that T is symmetric.

To see that $U_T = U$ we again write $x = (I - U)u$. We have

$$Tx = i(I + U)u \ \text{ so } \ (T + iI)x = 2iu \ \text{ and } (T - iI)x = 2iUu.$$

Thus $D(U_T) = \{2ix \ \ x \in D(U)\} = D(U)$ and

$$U_T(2iu) = 2iUu = u(2iu).$$

Thus $U = U_T$.

We must still show that T is a closed operator. H maps $x_n = (I - U)u_n$ to $(I + U)u_n$. If both $(I - U)u_n$ and $(I + U)u_n$ converge, then u_n and Uu_n converge. The fact that U is closed implies that if $u = \lim u_n$ then $u \in D(U)$ and $Uu = \lim Uu_n$. But this that $(I - U)u_n \to (I - U)u$ and $i(I + U)u_n \to i(I + U)u$ so T is closed. QED

The map $T \mapsto U_T$ from symmetric operators to isometries is called the **Cayley transform**.

Recall that an isometry is unitary if its domain and image are all of H. If U is a closed isometry, then $x_n \in D(U)$ and $x_n \to x$ that Ux_n is convergent, hence $x \in D(U)$ and $Ux = \lim Ux_n$. Similarly, if $Ux_n \to y$ then the x_n are

Cauchy, hence convergent to an x with $Ux = y$. So for any closed isometry U the spaces $D(U)^\perp$ and $\text{im}(U)^\perp$ measure how far U is from being unitary: If they both reduce to the zero subspace then U is unitary.

For a closed symmetric operator T define

$$H_T^+ = \{x \in H | T^*x = ix\} \text{ and } H_T^- = \{x \in H | T^*x = -ix\}. \qquad (2)$$

The main theorem of this section is

Theorem 2 *Let T be a closed symmetric operator and $U = U_T$ its Cayley transform. Then*

$$H_T^+ = D(U)^\perp \quad \textit{and} \quad H_T^- = (\text{im}(U))^\perp.$$

Every $x \in D(T^*)$ is uniquely expressible as

$$x = x_0 + x_+ + x_-$$

with $x_0 \in D(T)$, $x_+ \in H_T^+$ and $x_- \in H_T^-$, so

$$T^*x = Tx_0 + ix_+ - ix_-.$$

In particular, T is self adjoint if and only if U is unitary.

Proof. To say that $x \in D(U)^\perp = D\left(T + iI)^{-1}\right)^\perp$ that

$$(x, (T + iI)y) = 0 \quad \forall\ y \in D(T).$$

This says that

$$(x, Ty) = -(x, iy) = (ix, y) \quad \forall\ y \in D(T).$$

This is precisely the assertion that $x \in D(T^*)$ and $T^*x = ix$. We can read these equations backwards to conclude that $H_T^+ = D(U)^\perp$. Similarly, if $x \in \text{im}(U)^\perp$ then $(x, (T - iI)z) = 0 \ \ \forall z \in D(H)$ implying $T^*x = -ix$ and conversely.

We know that $D(U)$ and $\text{im}(U)$ are closed subspaces of H so any $w \in H$ can be written as the sum of an element of $D(U)$ and an element of $D(U)^\perp$. Taking $w = (T^* + iI)x$ for some $x \in D(T^*)$ gives

$$(T^* + iI)x = y_0 + x_1, \quad x_0 \in D(U) = \text{im}(T + iI),\ x_1 \in D(U)^\perp.$$

We can write $y_0 = (T + iI)x_0,\ x_0 \in D(T)$ so

$$(T^* + iI)x = (T + iI)x_0 + x_1.$$

Since $T^* = T$ on $D(T)$ and $T^*x_1 = ix_1$as $x_1 \in D(T)^\perp$ we have

$$T^*x_1 + ix_1 = 2ix_1.$$

So if we set

$$x_+ = \frac{1}{2i}x_1$$

we have

$$x_1 = (T^* + iI)x_+, \qquad x_+ \in D(T)^\perp.$$

so

$$(T^* + iI)x = (T^* + iI)(x_0 + x_+)$$

or

$$T^*(x - x_0 - x_+) = -i(x - x_0 - x_+).$$

This implies that $(x - x_0 - x_+) \in H_T^- = \mathrm{im}(U)^\perp$. So if we set

$$x_- := x - x_0 - x_+$$

we get the desired decomposition $x = x_0 + x_+ + x_-$.

To show that the decomposition is unique, suppose that

$$x_0 + x_+ + x_- = 0.$$

Applying $(T^* + iI)$ gives

$$0 = (T + iI)x_0 + 2ix_+.$$

But $(T + iI)x_0 \in D(U)$ and $x_+ \in D(U)^\perp$ so both terms above must be zero, so $x_+ = 0$. Also, from the preceding theorem we know that $(T + iI)x_0 = 0 \Rightarrow x_0 = 0$. Hence since $x_0 = 0$ and $x_+ = 0$ we must also have $x_- = 0$. QED

1.1 An elementary example.

Take $H = L_2([0,1])$ relative to the standard Lebesgue measure. Consider the operator $\frac{1}{i}\frac{d}{dt}$ which is defined on all elements of H whose derivative, in the sense of distributions, is again in $L_2([0,1])$. For any two such elements we have the integration by parts formula

$$\left(\frac{1}{i}\frac{d}{dt}x, y\right) = x(1)\overline{y(1)} - x(0)\overline{y(0)} + \left(x, \frac{1}{i}\frac{d}{dt}y\right).$$

(Even though in general the value at a point of an element in L_2 makes no sense, if x is such that $x' \in L_2$ then $\frac{1}{h}\int_0^h x(t)dt$ makes sense, and integration by parts using a continuous representative for x shows that the limit of this expression is well defined and equal to $x(0)$ for our continuous representative.) Suppose we take $T = \frac{1}{i}\frac{d}{dt}$ but with $D(T)$ consisting of those elements whose derivatives belong to L_2 as above, but which in addition satisfy

$$x(0) = x(1) = 0.$$

This space is dense in $H = L_2$ but if y is *any* function whose derivative is in H, we see from the integration by parts formula that

$$(Tx, y) = \left(x, \frac{1}{i}\frac{d}{dt}y\right).$$

In other words, using the Riesz representation theorem, we see that

$$T^* = \frac{1}{i}\frac{d}{dt}$$

defined on *all* y with derivatives in L_2. Notice that

$$T^* e^{\pm t} = \mp i e^{\pm t}$$

so in fact the spaces $H_T^{\pm}$ are both one dimensional.

For each complex number $e^{i\theta}$ of absolute value one we can find a "self adjoint extension" A_θ of T, that is an operator A_θ such that

$$D(T) \subset D(A_\theta) \subset D(T^*)$$

with $D(A_\theta) = D(A_\theta^*)$, $A_\theta = A_\theta^*$ and $A_\theta = T$ on $D(T)$. Indeed, let $D(A_\theta)$ consist of all x with derivatives in L_2 and which satisfy the "boundary condition"

$$x(1) = e^{i\theta}x(0).$$

Let us compute A_θ^* and its domain. Since $D(T) \subset D(A_\theta)$, if $(A_\theta x, y) = (x, A_\theta^* y)$ we must have $y \in D(T^*)$ and $A_\theta^* y = \frac{1}{i}\frac{d}{dt}y$. But then the integration by parts formula gives

$$(Ax, y) - (x, \frac{1}{i}\frac{d}{dt}y) = e^{i\theta}x(0)\overline{y(1)} - x(0)\overline{y(0)}.$$

This will vanish for all $x \in D(A_\theta)$ if and only if $y \in D(A_\theta)$. so we see that A_θ is self adjoint.

The moral is that to construct a self adjoint operator from a differential operator which is symmetric, we may have to supplement it with appropriate boundary conditions.

On the other hand, consider the same operator $\frac{1}{i}\frac{d}{dt}$ considered as an unbounded operator on $L_2(\mathbf{R})$. We take as its domain the set of all elements of $x \in L_2(\mathbf{R})$ whose distributional derivatives belong to $L_2(\mathbf{R})$ and such that $\lim_{t\to\pm\infty} = 0$. The functions $e^{\pm t}$ do not belong to $L_2(\mathbf{R})$ and so our operator is in fact self-adjoint. So the issue of whether or not we must add boundary conditions depends on the nature of the domain where the differential operator is to be defined. A deep analysis of this phenomenon for second order ordinary differential equations was provided by Hermann Weyl in a paper published in 1911. It is safe to say that much of the progress in the theory of self-adjoint operators was in no small measure influenced by a desire to understand and generalize the results of this fundamental paper.

2 Equibounded semi-groups on a Frechet space.

A Frechet space is a vector space with a topology defined by a sequence of semi-norms and which is complete. An important example is the Schwartz space $\mathcal{S}$. Let F be such a space. We want to consider a one parameter family of operators T_t on F defined for all $t \geq 0$ and which satisfy the following conditions:

- $T_0 = I$
- $T_t \circ T_s = T_{t+s}$
- $\lim_{t \to t_0} T_t x = T_{t_0} x \ \forall t_0 \geq 0$ and $x \in F$.
- For any defining seminorm p there is a defining seminorm q such that $p(T_t x) \leq K q(x)$ for all $t \geq 0$ and all $x \in X$.

We call such a family an **equibounded continuous semigroup**. We will usually drop the adjective "continuous" and even "equibounded" since we will not be considering any other kind of semigroup.

2.1 The infinitesimal generator.

We are going to begin by showing that every such semigroup has an " infinitesimal generator", i.e. can be written in some sense as $T_t = e^{At}$. It is important to observe that we have made a serious change of convention in that we are dropping the i that we have used until now. With this new notation, for example, the infinitesimal generator of a group of unitary transformations will be a skew-adjoint operator rather than a self-adjoint operator. There are many good reasons for this change in notation, but I do not want to go into them now.

So we define the operator A as

$$Ax = \lim_{t \searrow 0} \frac{1}{t}(T_t - I)x.$$

That is A is the operator defined on the domain $D(A)$ consisting of those x for which the limit exists.

Our first task is to show that $D(A)$ is dense in F. For this we begin as promised with the putative resolvent

$$R(z) := \int_0^\infty e^{-zt} T_t dt$$

which is defined (by the boundedness and continuity properties of T_t) for all z with Re $z > 0$. We begin by checking that every element of $\operatorname{im} R(z)$ belongs to $D(A)$: We have

$$\frac{1}{h}(T_h - I)R(z)x = \frac{1}{h}\int_0^\infty e^{-zt} T_{t+h} x dt - \frac{1}{h}\int_0^\infty e^{-zt} T_t x dt =$$

$$\frac{1}{h}\int_h^\infty e^{-z(r-h)} T_r x dr - \frac{1}{h}\int_0^\infty e^{-zt} T_t x dt \quad = \frac{e^{zh}-1}{h}\int_h^\infty e^{-zt} T_t x dt - \frac{1}{h}\int_0^h e^{-zt} T_t x dt$$

$$= \frac{e^{zh}-1}{h}\left[R(z)x - \int_0^h e^{-zt} T_t dt\right] - \frac{1}{h}\int_0^h e^{-zt} T_t x dt.$$

If we now let $h \to 0$, the integral inside the bracket tends to zero, and the expression on the right tends to x since $T_0 = I$. We thus see that

$$R(z)x \in D(A)$$

and

$$AR(z) = zR(z) - I,$$

or, rewriting this in a more familiar form,

$$(zI - A)R(z) = I.$$

We will now show that $D(A)$ is dense in F by showing that the span of the $\mathrm{im}(R(z))$ is dense. In fact, taking s to be real, we will show that

$$\lim_{s \to \infty} sR(s)x = x \quad \forall \; x \in F.$$

Indeed,

$$\int_0^\infty se^{-st} dt = 1$$

for any $s > 0$. So we can write

$$sR(s)x - x = s \int_0^\infty e^{-st}[T_t x - x]dt.$$

Applying any seminorm p we obtain

$$p(sR(s)x - x) \leq s \int_0^\infty e^{-st} p(T_t x - x)dt.$$

For any $\epsilon > 0$ we can, by the continuity of T_t, find a $\delta > 0$ such that

$$p(T_t x - x) < \epsilon \;\; \forall \;\; 0 \leq t \leq \delta.$$

Now let us write

$$s \int_0^\infty e^{-st} p(T_t x - x)dt = s \int_0^\delta e^{-st} p(T_t x - x)dt + s \int_\delta^\infty e^{-st} p(T_t x - x)dt.$$

The first integral is bounded by

$$\epsilon s \int_0^\delta e^{-st} dt \leq \epsilon s \int_0^\infty e^{-st} dt = \epsilon.$$

As to the second integral, let M be a bound for $p(T_t x) + p(x)$ which exists by the uniform boundedness of T_t. The triangle inequality says that $p(T_t x - x) \leq p(T_t x) + p(x)$ so the second integral is bounded by

$$M \int_\delta^\infty se^{-st} dt = Me^{-s\delta}.$$

This tends to 0 as $s \to \infty$, completing the proof that $sR(s)x \to x$ and hence that $D(A)$ is dense in F.

3 The differential equation

Theorem 3 *If $x \in D(A)$ then for any $t > 0$*

$$\lim_{h \to 0} \frac{1}{h}[T_{t+h} - T_t]x = AT_t x = T_t Ax.$$

In colloquial terms, we can formulate the theorem as saying that

$$\frac{d}{dt} T_t = AT_t = T_t A$$

in the sense that the appropriate limits exist when applied to $x \in D(A)$.

Proof. Since T_t is continuous in t, we have

$$T_t Ax = T_t \lim_{h \searrow 0} \frac{1}{h}[T_h - I]x = \lim_{h \searrow 0} \frac{1}{h}[T_t T_h - T_t]x =$$

$$\lim_{h \searrow 0} \frac{1}{h}[T_{t+h} - T_t]x = \lim_{h \searrow 0} \frac{1}{h}[T_h - I]T_t x$$

for $x \in D(A)$. This shows that $T_t x \in D(A)$ and

$$\lim_{h \searrow 0} \frac{1}{h}[T_{t+h} - T_t]x = AT_t x = T_t Ax.$$

To prove the theorem we must show that we can replace $h \searrow 0$ by $h \to 0$. Our strategy is to show that with the information that we already have about the existence of right handed derivatives, we can conclude that

$$T_t x - x = \int_0^t T_s Ax ds.$$

Since T_t is continuous, this is enough to give the desired result. In order to establish the above equality, it is enough, by the Hahn-Banach theorem to prove that for any $\ell \in F^*$ we have

$$\ell(T_t x) - \ell(x) = \int_0^t \ell(T_s Ax) ds.$$

In turn, it is enough to prove this equality for the real and imaginary parts of ℓ.

So it all boils down to a lemma in the theory of functions of a real variable:

Lemma 1 *Suppose that f is a continuous real valued function of t with the property that the right hand derivative*

$$\frac{d^+}{dt} f := \lim_{h \searrow 0} \frac{f(t+h) - f(t)}{h} = g(t)$$

exists for all t and $g(t)$ is continuous. Then f is differentiable with $f' = g$.

Proof. We first prove that $\frac{d^+}{dt}f \geq 0$ on an interval $[a, b]$ implies that $f(b) \geq f(a)$. Suppose not. Then there exists an $\epsilon > 0$ such that

$$f(b) - f(a) < -\epsilon(b - a).$$

Set

$$F(t) := f(t) - f(a) + \epsilon(t - a).$$

Then $F(a) = 0$ and

$$\frac{d^+}{dt}F > 0.$$

At a this implies that there is some $c > a$ near a with $F(c) > 0$. On the other hand, since $F(b) < 0$, and F is continuous, there will be some point $s < b$ with $F(s) = 0$ and $F(t) < 0$ for $s < t \leq b$. This contradicts the fact that $[\frac{d^+}{dt}F](s) > 0$.

Thus if $\frac{d^+}{dt}f \geq m$ on an interval $[t_1, t_2]$ we may apply the above result to $f(t) - mt$ to conclude that

$$f(t_2) - f(t_1) \geq m(t_2 - t_1),$$

and if $\frac{d^+}{dt}f(t) \leq M$ we can apply the above result to $Mt - f(t)$ to conclude that $f(t_2) - f(t_1) \leq M(t_2 - t_1)$. So if $m = \min g(t) = \frac{d^+}{dt}f$ on the interval $[t_1, t_2]$ and M is the maximum, we have

$$m \leq \frac{f(t_2) - f(t_1)}{t_2 - t_1} \leq M.$$

Since we are assuming that g is continuous, this is enough to prove that f is indeed differentiable with derivative g. QED

3.1 The resolvent.

We have already verified that

$$R(z) = \int_0^\infty e^{-zt}T_t dt$$

maps F into $D(A)$ and satisfies

$$(zI - A)R_z = I$$

for all z with $\mathrm{Re} z > 0$.

We shall now show that for this range of z

$$(zI - A)x = 0 \quad \Rightarrow \quad x = 0 \quad x \in D(A)$$

so that $(zI - A)^{-1}$ exists and is given by $R(z)$. Suppose that

$$Ax = zx \qquad x \in D(A)$$

and choose $\ell \in F^*$ with $\ell(x) = 1$. Consider

$$\phi(t) := \ell(T_t x).$$

By the result of the preceding section we know that ϕ is a differentiable function of t and satisfies the differential equation

$$\phi'(t) = \ell(T_t A x) = z\phi(t), \quad \phi(0) = 1.$$

So

$$\phi(t) = e^{zt}$$

which is impossible since $\phi(t)$ is a bounded function of t and the right hand side of the above equation is not bounded for $t \geq 0$ since the real part of z is positive.

From $(zI - A)R_z = I$ we see that $zI - A$ maps $\operatorname{im} R(z) \subset D(A)$ onto F so certainly $zI - A$ maps $D(A)$ onto F bijectively. Hence

$$\operatorname{im}(R(z)) = D(A), \quad \operatorname{im}(zI - A) = F$$

and

$$R(z) = (zI - A)^{-1}.$$

We have aleady estabished the following:

The resolvent $R_z = R(z, A) = \int_0^\infty e^{-zt} T_t dt$ is defined as a strong limit for Re $z > 0$ and, for this range of z:

$$\begin{aligned} D(A) &= \operatorname{im}(R(z, A) & (3) \\ AR(z,A)x = R(z,A)Ax &= (zR(z,A) - I)x \quad x \in D(A) & (4) \\ AR(z,A)x &= (zR(z,A) - I)x \quad \forall\, x \in F & (5) \\ \lim_{z \nearrow \infty} zR(z,A)x &= x \quad \text{for } z \text{ real } \; \forall x \in F. & (6) \end{aligned}$$

We also have

Theorem 4 *The operator A is closed.*

Proof. Suppose that $x_n \in D(A)$, $x_n \to x$ and $y_n \to y$ where $y_n = Ax_n$. We must show that $x \in D(A)$ and $Ax = y$. Set

$$z_n := (I - A)x_n \quad \text{so} \quad z_n \to x - y.$$

Since $R(1, A) = (I - A)^{-1}$ is a bounded operator, we conclude that

$$x = \lim x_n = \lim (I - A)^{-1} z_n = (I - A)^{-1}(x - y).$$

From (3) we see that $x \in D(A)$ and from the preceding equation that $(I - A)x = x - y$ so $Ax = y$. QED

3.1.1 Application to Stone's theorem.

We now have enough information to complete the proof of Stone's theorem:

Suppose that $U(t)$ is a one-parameter group of unitary transformations on a Hilbert space. We have $(U(t)x, y) = (x, U(t)^{-1}y) = (x, U(-t)y)$ and so differentiating at the origin shows that the infinitesimal generator A, which we know to be closed, is skew-symmetric:

$$(Ax, y) = (x, Ay) \quad \forall\, x, y \in D(A).$$

Also the resolvents $(zI - A)^{-1}$ exist for all z which are not purely imaginary, and $(zI - A)$ maps $D(A)$ onto the whole Hilbert space H.

Writing $A = iT$ we see that T is symmetric and that its Cayley transform U_T has zero kernel and is surjective, i.e. is unitary. Hence T is self-adjoint. This proves Stone's theorem that every one parameter group of unitary transformations is of the form e^{iTt} with T self-adjoint.

3.2 Examples.

For $r > 0$ let

$$J_r := (I - r^{-1}A)^{-1} = rR(r, A)$$

so by (4) we have

$$AJ_r = r(J_r - I). \tag{7}$$

3.2.1 Translations.

Consider the one parameter group of translations acting on $L_2(\mathbf{R})$:

$$[U(t)x](s) = x(s - t).$$

This is defined for all $x \in \mathcal{S}$ and is an isometric isomorphism there, so extends to a unitary one parameter group acting on $L_2(\mathbf{R})$. Equally well, we can take the above equation in the sense of distributions, where it makes sense for all elements of $\mathcal{S}'$, in particular for all elements of $L_2(\mathbf{R})$. We know that we can differentiate in the distributional sense to obtain

$$A = -\frac{d}{dx}$$

as the "infinitesimal generator" in the distributional sense. Let us see what the general theory gives. Let $y_r := J_r x$ so

$$y_r(s) = r\int_0^\infty e^{-rt}x(s-t)dt = r\int_{-\infty}^s e^{-r(s-u)}x(u)du.$$

The right hand expression is a differentiable function of s and

$$y_r'(s) = rx(s) - r^2\int_s^\infty e^{-r(s-u)}x(u)du = rx(s) - ry_r(s).$$

On the other hand we know from (7) that

$$Ay_r = AJ_r x = r(y_r - x).$$

Putting the two equations together gives

$$A = -\frac{d}{dx}$$

as expected. This is a skew-adjoint operator in accordance with Stone's theorem.

3.2.2 The heat equation.

Let F consist of the bounded uniformly continuous functions on $\mathbf{R}$. For $t > 0$ define

$$[T_t x](s) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-(s-v)/2t} x(v) dv.$$

In other words, T_t is convolution with

$$n_t(u) = \frac{1}{\sqrt{2\pi t}} e^{-u^2/2t}.$$

We have already verified in our study of the Fourier transform that this is a continuous semi-group (when we set $T_0 = I$) when acting on $\mathcal{S}$. It is easy enough to verify that these operators are continuous in the uniform norm and hence extend to an equibounded semigroup on F. We will now verify that the infiitesimal generator A of this semigroup is

$$A = \frac{1}{2}\frac{d^2}{ds^2}$$

with domain consisting of all twice differentiable functions.

Let us set $y_r = J_r x$ so

$$\begin{aligned}
y_r(s) &= \int_{-\infty}^{\infty} x(v) \left[\int_0^{\infty} r \frac{1}{\sqrt{2\pi t}} e^{-rt-(s-v)^2/2t} dt \right] dv \\
&= \int_{-\infty}^{\infty} x(v) \left[2\sqrt{r} \frac{1}{\sqrt{2\pi}} e^{-\sigma^2 - r(s-v)^2/2\sigma^2} d\sigma \right] dv \quad \text{setting } t = \sigma^2/r \\
&= \int_{-\infty}^{\infty} x(v)(r/2)^{\frac{1}{2}} e^{-\sqrt{2n}|s-v|} dv
\end{aligned}$$

since for any $c > 0$ we have

$$\int_0^{\infty} e^{-(\sigma^2 + c/\sigma^2)} d\sigma = \frac{\sqrt{\pi}}{2} e^{-2c}$$

as can be seen by completing the square.

So we can write

$$y_r(s) = \left(\frac{r}{2}\right)^{\frac{1}{2}} \left[\int_s^{\infty} x(v) e^{-\sqrt{2r}(v-s)} dv + \int_{-\infty}^{s} e^{-\sqrt{2r}(s-v)} dv \right].$$

This is a differentiable function of s and we can differentiate to obtain

$$y_r'(s) = r\left[\int_s^\infty x(v))e^{-\sqrt{2r}(v-s)}dv - \int_{-\infty}^s x(v))e^{-\sqrt{2r}(s-v)}dv\right].$$

This is also differentiable and compute its derivative to obtain

$$y_r''(s) = -2rx(s) + r^{3/2}\sqrt{2}\int_{-\infty}^\infty x(v))e^{-\sqrt{2n}(v-s)}dv,$$

or

$$y_r'' = 2r(y_r - x).$$

Comparing this with (7) which says that $Ay_r = r(y_r - x)$ we see that indeed

$$A = \frac{1}{2}\frac{d^2}{ds^2}.$$

3.2.3 Bochner's theorem.

A complex valued continuous function F is called **positive definite** if for every continuous function ϕ of compact support we have

$$\int_{\mathbf{R}}\int_{\mathbf{R}} F(t-s)\phi(t)\overline{\phi(s)}dtds \geq 0. \tag{8}$$

We can write this as

$$(F\star\overline{\phi},\overline{\phi}) \geq 0$$

where the convolution is taken in the sense of generalized functions. If we write $F = \hat{G}$ and $\overline{\phi} = \hat{\psi}$ then this equation becomes

$$(G\psi,\psi) \geq 0$$

or

$$\langle G, |\psi|^2\rangle \geq 0$$

which will certainly be true if G is a finite non-negative measure. Bochner's theorem asserts the converse: that any positive definite function is the Fourier transform of a finite non-negative measure. We shall follow Yosida pp. 346-347 in showing that Stone's theorem implies Bochner's theorem.

Let $\mathcal{F}$ denote the space of functions on $\mathbf{R}$ which have finite support, i.e. vanish outside a finite set. This is a complex vector space, and has the semi-scalar product

$$(x,y) := \sum_{t,s} F(t-s)x(t)\overline{y(s)}.$$

(It is easy to see that the fact that F is a postive definite function implies that $(x,x) \geq 0$ for all $x \in \mathcal{F}$.) Passing to the quotient by the subspce of null vectors and completing we obtain a Hilbert space H.

Let U_r be defined by $[U_r x](t) = x(t - r)$ as usual. Then

$$(U_r x, U_r y) = \sum_{t,s} F(t-s)x(t-r)\overline{y(s-r)} = \sum_{t,s} F(t+r-(s+r))x(t)\overline{y(s)} = (x, y).$$

So U_r descends to H to define a unitary operator which we shal continue to denote by U_r. We thus obtain a one parameter group of unitary transformations on H. According to Stone's theorem there exists a resolutions E_λ of the identity such that

$$U_t = \int_{-\infty}^{\infty} e^{it\lambda} dE_\lambda.$$

Now choose $\delta \in \mathcal{F}$ to be defined by

$$\delta(t) = \begin{cases} 1 & \text{if} \quad t = 0 \\ 0 & \text{if} \quad t \neq 0 \end{cases} .$$

Let x be the image of δ in H. Then

$$(U_r x, x) = \sum F(t - s)\delta(t - r)\delta(s) = F(r).$$

But by Stone we have

$$F(r) = \int_{-\infty}^{\infty} e^{ir\lambda} d\mu_{x,x} = \int_{-\infty}^{\infty} e^{ir\lambda} d\|E_\lambda x\|^2$$

so we have represented F as the Fourier transform of a finite non-negative measure. QED