Stone's Theorem.

Math 212a

December 6, 2001

Recall that if A is a self-adjoint operator on a Hilbert space $\mathbf{H}$ 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 Spec(A) if we use our second proof. In any event, the spectral theorem allows us to write

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

and to verify that

$$U(0) = I, \quad U(s+t) = U(s)U(t)$$

and that U depends continuously on t. 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

$$\int_0^\infty e^{(-z+ix)t} dt = \left. \frac{e^{(-z+ix)t}}{-z+ix} \right|_{t=0}^{\infty} = \frac{1}{z-ix} \text{ if Re } z > 0$$

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

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

Since A is self-adjoint, its spectrum is real. So the spectrum of iA is purely imaginary, and hence any z not on the imaginary axis is in the resolvent set of iA. The above formula gives us an expression for the resolvent in terms of $U(t)$ for z lying in the right half plane. We can obtain a similar formula for the left half plane.

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 and then the one-parameter group $U(t) = e^{iAt}$.

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 **F** 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, Nelson, *Topics in dynamics I: Flows*, and Reed and Simon *Methods of Mathematical Physics, II. Fourier Analysis, Self-Adjointness.*

Contents

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.) Indeed in the expression

$$z = \frac{x-i}{x+i}$$

when x is real, the numerator is the complex conjugate of the denominator and hence $|z| = 1$. Under this transformation, the cardinal points $0, 1, \infty$ of the extended real axis are mapped as follows:

$$0 \mapsto -1, \quad 1 \mapsto -i, \quad \text{and} \quad \infty \mapsto 1.$$

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

$$(A - iI)(A + iI)^{-1}$$

should be unitary. In fact, we can be much more precise. First some definitions:

An operator U, possibly defined only on a subspace of a Hilbert space $\mathbf{H}$ 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 its domain $D(T)$ be dense in $\mathbf{H}$. 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 $\mathbf{H}$) is called **symmetric** if $D(T)$ is dense in $\mathbf{H}$ so that 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 $\mathbf{H}$ 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

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

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

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

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

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 $y = (I - U_T)^{-1}(2ix)$ 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 $z \in \operatorname{im}(I-U)$ 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 $\operatorname{im}(I-U)$ is dense in $\mathbf{H}$, the condition $(y,z) = 0 \ \forall\, z \in \operatorname{im}(I-U)$ implies that $y=0$. Thus $(I-U)^{-1}$ exists, and we may define

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

with

$$D(T) = D\left((I-U)^{-1}\right) = \operatorname{im}(I-U)$$

dense in $\mathbf{H}$. Suppose that $x=(I-U)u,\ y=(I-U)v \ \in D(T) = \operatorname{im}(I-U)$. 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 $D(U_T) = \{2iu \ \ u \in D(U)\} = D(U)$ and

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

Thus $U = U_T$.

We must still show that T is a closed operator. T 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 $\mathbf{H}$. If U is a closed isometry, then $x_n \in D(U)$ and $x_n \to x$ implies 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 $\operatorname{im}(U)^\perp$ 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

$$\mathbf{H}_T^+ = \{x \in \mathbf{H} | T^*x = ix\} \ \text{ and } \ \mathbf{H}_T^- = \{x \in \mathbf{H} | T^*x = -ix\}. \tag{2}$$

The main theorem of this section is

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

$$\mathbf{H}_T^+ = D(U)^\perp \quad \textit{and} \quad \mathbf{H}_T^- = (\mathrm{im}(U))^\perp.$$

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

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

with $x_0 \in D(T)$, $x_+ \in \mathbf{H}_T^+$ and $x_- \in \mathbf{H}_T^-$, so

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

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

Proof. To say that $x \in D(U)^\perp = D\left((T+iI)^{-1}\right)^\perp$ says 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 $\mathbf{H}_T^+ = D(U)^\perp$. Similarly, if $x \in \mathrm{im}(U)^\perp$ then $(x, (T-iI)z) = 0 \ \ \forall z \in D(T)$ implying $T^*x = -ix$ and conversely.

We know that $D(U)$ and $\mathrm{im}(U)$ are closed subspaces of $\mathbf{H}$ so any $w \in \mathbf{H}$ 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

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

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 $x_1 \in D(U)^\perp$ we have

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

So if we set

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

we have

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

so

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

or

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

This implies that $(x - x_0 - x_+) \in \mathbf{H}_T^- = \mathrm{im}(U)^\perp$. 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 $\mathbf{H} = L_2([0,1])$ relative to the standard Lebesgue measure. Consider the operator $\frac{1}{i}\frac{d}{dt}$ which is defined on all elements of $\mathbf{H}$ 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 $\mathbf{H} = L_2$ but if y is *any* function whose derivative is in $\mathbf{H}$, 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 $\mathbf{H}_T^{\pm}$ 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 $\lim_{t \to \pm\infty} x = 0$. 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 $\mathbf{F}$ 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 $\mathbf{F}$ be such a space. We want to consider a one parameter family of operators T_t on $\mathbf{F}$ 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 $x \in \mathbf{F}$.

- For any defining seminorm p there is a defining seminorm q and a constant K such that $p(T_t x) \leq Kq(x)$ for all $t \geq 0$ and all $x \in \mathbf{F}$.

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. In quantum mechanics, where an "observable" is a self-adjoint operator, there is a good reason for emphasizing the self-adjoint operators, and hence including the i. There are many good reasons for deviating from the physicists' notation, not the least having to do with the theory of Lie algebras. I do not want to go into these reasons now. Some will emerge from the ensuing notation. But the presence or absence of the i is a cultural divide between physicists and mathematicians.

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 $\mathbf{F}$. For this we begin as promised with the putative resolvent

$$R(z) := \int_0^\infty e^{-zt} T_t dt \tag{3}$$

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,

$$(zI - A)R(z) = I. \tag{4}$$

This equation says that $R(z)$ is a right inverse for $zI - A$. It will require a lot more work to show that it is also a left inverse.

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

$$\lim_{s\to\infty} sR(s)x = x \quad \forall\ x \in \mathbf{F}. \tag{5}$$

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

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

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

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

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

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

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 $p(T_t x - x) \le p(T_t x) + p(x)$ 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 $\mathbf{F}$.

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

$$T_t x - x = \int_0^t T_s Axds.$$

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 $\ell \in \mathbf{F}^*$ 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 $m = \min g(t) = \min \frac{d^+}{dt}f$ 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 $\mathbf{F}$ into $D(A)$ and satisfies

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

for all z with Re $z > 0$, cf (4).

We shall now show that for this range of z

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

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

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

and choose $\ell \in \mathbf{F}^*$ 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

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

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.

We have from (4) that

$$(zI - A)R(z)(zI - A)x = (zI - A)x$$

and we know that $R(z)(zI - A)x \in D(A)$. From the injectivity of $zI - A$ we conclude that $R(z)(zI - A)x = x$.

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

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

and

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

We have already established 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:

$$\begin{aligned} D(A) &= \operatorname{im}(R(z, A)) && (6)\\ AR(z, A)x = R(z, A)Ax &= (zR(z, A) - I)x \quad x \in D(A) && (7)\\ AR(z, A)x &= (zR(z, A) - I)x \quad \forall\, x \in \mathbf{F} && (8)\\ \lim_{z \nearrow \infty} zR(z, A)x &= x \quad \text{for } z \text{ real } \forall x \in \mathbf{F}. && (9) \end{aligned}$$

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 (6) 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 $\mathbf{H}$.

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 (8) we have

$$AJ_r = r(J_r - I). \tag{10}$$

3.2.1 Translations.

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

$$[U(t)x](s) = x(s - t). \tag{11}$$

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

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

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 (10) that

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

Putting the two equations together gives

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

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

We can now go back and give an intuitive explanation of what goes wrong when considering this same operator A but on $L_2[0,1]$ instead of on $L_2(\mathbf{R})$. If x is a smooth function of compact support lying in $(0,1)$, then x can not tell whether it is to be thought of as lying in $L_2([0,1])$ or $L_2(\mathbf{R})$, so the only choice for a unitary one parameter group acting on x (at least for small $t > 0$) if the shift to the right as given by (11). But once t is large enough that the support of $U(t)x$ hits the right end point, 1, this transformation can not continue as is. The only hope is to have what "goes out" the right hand side come in, in some form, on the left, and unitarity now requires that

$$\int_0^1 |x(s-t)^2 dt = \int_0^1 |x(t)|^2 dt$$

where now the shift in (11) means mod 1. This still allows freedom in the choice of phase between the exiting value of the x and its incoming value. Thus we specify a unitary one parameter group when we fix a choice of phase as the effect of "passing go". This choice of phase is the origin of the θ that we needed to introduce in finding the self adjoint extensions of $\frac{1}{i}\frac{d}{dt}$ acting on functions vanishing at the boundary.

3.2.2 The heat equation.

Let $\mathbf{F}$ 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 $\mathcal{S}$. In fact, for $x \in \mathcal{S}$, we can take the Fourier transform and conclude that

$$[T_t x]\hat{}(\sigma) = e^{-i\sigma^2 t/2} \hat{x}(\sigma).$$

Differentiating this with respect to t and setting $t = 0$ (and taking the inverse Fourier transform) shows that

$$\left[\frac{d}{dt} T_t x\right]_{t=0} = \frac{1}{2}\frac{d^2}{ds^2} x$$

for $x \in \mathcal{S}$. We wish to arrive at the same result for T_t acting on $\mathbf{F}$. It is easy enough to verify that the operators T_t are continuous in the uniform norm and hence extend to an equibounded semigroup on $\mathbf{F}$. We will now verify that the infinitesimal 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

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

since for any $c > 0$ we have

$$\int_0^{\infty} e^{-(\sigma^2+c^2/\sigma^2)}d\sigma = \frac{\sqrt{\pi}}{2}e^{-2c}. \tag{12}$$

Let me postpone the calculation of this integral to the end of the subsection. Assuming the evaluation of this integral we can write

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

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

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

or

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

Comparing this with (10) which says that $Ay_r = r(y_r - x)$ we see that indeed

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

Let us now verify the evaluation of the integral in (12): Start with the known integral

$$\int_0^{\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

Set $x = \sigma - c/\sigma$ so that $dx = (1 + c/\sigma^2)d\sigma$ and $x = 0$ corresponds to $\sigma = \sqrt{c}$. Thus $\frac{\sqrt{\pi}}{2} =$

$$\int_{\sqrt{c}}^{\infty} e^{-(\sigma - c/\sigma)^2}(1 + c/\sigma)^2 d\sigma = e^{2c}\int_{\sqrt{c}}^{\infty} e^{-(\sigma^2 + c^2/\sigma^2)}(1 + c/\sigma^2)d\sigma$$

$$= e^{2c}\left[\int_{\sqrt{c}}^{\infty} e^{-(\sigma^2+c^2/\sigma^2)}d\sigma + \int_{\sqrt{c}}^{\infty} e^{-(\sigma^2+c^2/\sigma^2)}\frac{c}{\sigma^2}d\sigma\right].$$

In the second integral inside the brackets set $t = -c/\sigma$ so $dt = \frac{c}{\sigma^2}d\sigma$ and this second integral becomes

$$\int_0^{\sqrt{c}} e^{-(t^2+c^2/t^2)}dt$$

and hence

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

which is (12).

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

$$\int_{\mathbf{R}}\int_{\mathbf{R}} F(t-s)\phi(t)\overline{\phi(s)}dtds \geq 0. \tag{13}$$

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 by Plancherel 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 positive definite function implies that $(x,x) \geq 0$ for all $x \in \mathcal{F}$.) Passing to the quotient by the subspace of null vectors and completing we obtain a Hilbert space $\mathbf{H}$.

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 $\mathbf{H}$ to define a unitary operator which we shall continue to denote by U_r. We thus obtain a one parameter group of unitary transformations on $\mathbf{H}$. According to Stone's theorem there exists a resolution 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 $\mathbf{H}$. 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

The logic of our argument has been - the Spectral Theorem implies Stone's theorem implies Bochner's theorem. In fact, assuming the Hille-Yosida theorem on the existence of semigroups to be proved below, one can go in the opposite direction. Given a one parameter group $U(t)$ of unitary transformations, it is easy to check that for any $x \in \mathbf{H}$ the function $t \mapsto (U(t)x, x)$ is positive definite, and then use Bochner's theorem to derive the spectral theorem on the cyclic subspace generated by x under $U(t)$. One can then get the full spectral theorem in multiplication operator form as we did in the handout on unbounded self-adjoint operators.

4 The power series expansion of the exponential.

In finite dimensions we have the formula

$$e^{tB} = \sum_0^{\infty} \frac{t^k}{k!} B^k$$

with convergence guaranteed as a result of the convergence of the usual exponential series in one variable. (There are serious problems with this definition from the point of view of numerical implementation which we will not discuss here.)

In infinite dimensional spaces some additional assumptions have to be placed on an operator B before we can conclude that the above series converges. Here is a very stringent condition which nevertheless suffices for our purposes.

Let $\mathbf{F}$ be a Frechet space and B a continuous map of $\mathbf{F} \to \mathbf{F}$. We will assume that the B^k are **equibounded** in the sense that for any defining semi-norm p there is a constant K and a defining semi-norm q such that

$$p(B^k x) \leq Kq(x) \quad \forall\, k = 1, 2, \dots \quad \forall\, x \in \mathbf{F}.$$

Here the K and q are required to be independent of k and x.

Then

$$p(\sum_m^n \frac{t^k}{k!} B^k x) \leq \sum_m^n \frac{t^k}{k!} p(B^k x) \leq Kq(x) \sum_n^n \frac{t^k}{k!}$$

and so

$$\sum_0^n \frac{t^k}{k!} B^k x$$

is a Cauchy sequence for each fixed t and x (and uniformly in any compact interval of t). It therefore converges to a limit. We will denote the map $x \mapsto \sum_0^\infty \frac{t^k}{k!} B^k x$ by

$$\exp(tB).$$

This map is linear, and the computation above shows that

$$p(\exp(tB)x) \leq K \exp(t) q(x).$$

The usual proof (using the binomial formula) shows that $t \mapsto \exp(tB)$ is a one parameter equibounded semi-group. More generally, if B and C are two such operators then if $BC = CB$ then $\exp(t(B + C)) = (\exp tB)(\exp tC)$.

Also, from the power series it follows that the infinitesimal generator of $\exp tB$ is B.

5 The Hille Yosida theorem.

Let us now return to the general case of an equibounded semigroup T_t with infinitesimal generator A on a Frechet space $\mathbf{F}$ where we know that the resolvent $R(z, A)$ for Re $z > 0$ is given by

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

This formula shows that $R(z, A)x$ is continuous in z. The resolvent equation

$$R(z, A) - R(w, A) = (w - z) R(z, A) R(w, A)$$

then shows that $R(z,A)x$ is complex differentiable in z with derivative $-R(z,A)^2x$. It then follows that $R(z,A)x$ has complex derivative of all order given by

$$\frac{d^n R(z,A)x}{dz^n} = (-1)^n n! R(z,A)^{n+1}x.$$

On the other hand, differentiating the integral formula for the resolvent n- times gives

$$\frac{d^n R(z,A)x}{dz^n} = \int_0^\infty e^{-zt}(-t)^n T_t dt$$

where differentiation under the integral sign is justified by the fact that the T_t are equicontinuous in t. Putting the previous two equations together gives

$$(zR(z,A))^{n+1}x = \frac{z^{n+1}}{n!}\int_0^\infty e^{-zt}t^n T_t x dt.$$

This implies that for any semi-norm p we have

$$p((zR(z,A))^{n+1}x) \leq \frac{z^{n+1}}{n!}\int_0^\infty e^{-zt}t^n \sup_{t\geq 0} p(T_t x)dt = \sup_{t\geq 0} p(T_t x)$$

since

$$\int_0^\infty e^{-zt}t^n dt = \frac{n!}{z^{n+1}}.$$

Since the T_t are equibounded by hypothesis, we conclude

Proposition 1 *The family of operators $\{(zR(z,A))^n\}$ is equibounded in Re $z > 0$ and $n = 0, 1, 2, \ldots$.*

We now come to the main result of this section:

Theorem 5 [Hille-Yosida.] *Let A be an operator with dense domain $D(A)$, and such that the resolvents*

$$R(n,A) = (nI - A)^{-1}$$

exist and are bounded operators for $n = 1, 2, \ldots$. Then A is the infinitesimal generator of a uniquely determined equibounded semigroup if and only if the operators

$$\{(I - n^{-1}A)^{-m}\}$$

are equibounded in $m = 0, 1, 2 \ldots$ and $n = 1, 2, \ldots$.

Proof. If A is the infinitesimal generator of an equibounded semi-group then we know that the $\{(I - n^{-1}A)^{-m}\}$ are equibounded by virtue of the preceding proposition. So we must prove the converse.

Set

$$J_n = (I - n^{-1}A)^{-1}$$

so $J_n = n(nI - A)^{-1}$ and so for $x \in D(A)$ we have

$$J_n(nI - A)x = nx$$

or

$$J_n Ax = n(J_n - I)x.$$

Similarly $(nI - A)J_n = nI$ so $AJ_n = n(J_n - I)$. Thus we have

$$AJ_n x = J_n Ax = n(J_n - I)x \quad \forall\, x \in D(A). \tag{14}$$

The idea of the proof is now this: By the results of the preceding section, we can construct the one parameter semigroup $s \mapsto \exp(sJ_n)$. Set $s = nt$. We can then form $e^{-nt}\exp(ntJ_n)$ which we can write as $\exp(tn(J_n - I)) = \exp(tAJ_n)$ by virtue of (14). We expect from (5) that

$$\lim_{n\to\infty} J_n x = x \quad \forall \quad x \in \mathbf{F}. \tag{15}$$

This then suggests that the limit of the $\exp(tAJ_n)$ be the desired semi-group.

So we begin by proving (15). We first prove it for $x \in D(A)$. For such x we have $(J_n - I)x = n^{-1}J_n Ax$ by (14) and this approaches zero since the J_n are equibounded. But since $D(A)$ is dense in $\mathbf{F}$ and the J_n are equibounded we conclude that (15) holds for all $x \in \mathbf{F}$.

Now define

$$T_t^{(n)} = \exp(tAJ_n) := \exp(nt(J_n - I)) = e^{-nt}\exp(ntJ_n).$$

We know from the preceding section that

$$p(\exp(ntJ_n)x) \le \sum \frac{(nt)^k}{k!} p(J_n^k x) \le e^{nt}Kq(x)$$

which implies that

$$p(T_t^{(n)}x) \le Kq(x). \tag{16}$$

Thus the family of operators $\{T_t^{(n)}\}$ is equibounded for all $t \ge 0$ and $n = 1, 2, \ldots$. We next want to prove that the $\{T_t^{(n)}\}$ converge as $n \to \infty$ uniformly on each compact interval of t:

The J_n commute with one another by their definition, and hence J_n commutes with $T_t^{(m)}$. By the semi-group property we have

$$\frac{d}{dt}T_t^m x = AJ_m T_t^{(m)}x = T_t^{(m)}AJ_m x$$

so

$$T_t^{(n)}x - T_t^{(m)}x = \int_0^t \frac{d}{ds}(T_{t-s}^{(m)}T_s^{(n)})x\,ds = \int_0^t T_{t-s}^{(m)}(AJ_n - AJ_m)T_s^{(n)}x\,ds.$$

Applying the semi-norm p and using the equiboundedness we see that

$$p(T_t^{(n)}x - T_t^{(m)}x) \leq Ktq((J_n - J_m)Ax).$$

From (15) this implies that the $T_t^{(n)}x$ converge (uniformly in every compact interval of t) for $x \in D(A)$, and hence since $D(A)$ is dense and the $T_t^{(n)}$ are equicontinuous for all $x \in \mathbf{F}$. The limiting family of operators T_t are equicontinuous and form a semi-group because the $T_t^{(n)}$ have this property.

We must show that the infinitesimal generator of this semi-group is A. Let us temporarily denote the infinitesimal generator of this semi-group by B, so that we want to prove that $A = B$. Let $x \in D(A)$. We claim that

$$\lim_{n\to\infty} T_t^{(n)} AJ_n x = T_t Ax \tag{17}$$

uniformly in in any compact interval of t. Indeed, for any semi-norm p we have

$$\begin{aligned} p(T_tAx - T_t^{(n)}AJ_nx) &\leq p(T_tAx - T_t^{(n)}Ax) + p(T_t^{(n)}Ax - T_t^{(n)}AJ_nx) \\ &\leq p((T_t - T_t^{(n)})Ax) + Kq(Ax - J_nAx) \end{aligned}$$

where we have used (16) to get from the second line to the third. The second term on the right tends to zero as $n \to \infty$ and we have already proved that the first term converges to zero uniformly on every compact interval of t. This establishes (17).

We thus have, for $x \in D(A)$,

$$\begin{aligned} T_tx - x &= \lim_{n\to\infty}(T_t^{(n)}x - x) \\ &= \lim_{n\to\infty}\int_0^t T_s^{(n)}AJ_nxds \\ &= \int_0^t (\lim_{n\to\infty} T_s^{(n)}AJ_nx)ds \\ &= \int_0^t T_sAxds \end{aligned}$$

where the passage of the limit under the integral sign is justified by the uniform convergence in t on compact sets. It follows from $T_tx - x = \int_0^t T_sAxds$ that x is in the domain of the infinitesimal operator B of T_t and that $Bx = Ax$. So B is an extension of A in the sense that $D(B) \supset D(A)$ and $Bx = Ax$ on $D(A)$.

But since B is the infinitesimal generator of an equibounded semi-group, we know that $(I - B)$ maps $D(B)$ onto $\mathbf{F}$ bijectively, and we are assuming that $(I - A)$ maps $D(A)$ onto $\mathbf{F}$ bijectively. Hence $D(A) = D(B)$. QED

In case $\mathbf{F}$ is a Banach space, so there is a single norm $p = \|\ \|$, the hypotheses of the theorem read: $D(A)$ is dense in $\mathbf{F}$, the resolvents $R(n, A)$ exist for all integers $n = 1, 2, \ldots$ and there is a constant K independent of n and m such that

$$\|(I - n^{-1}A)^{-m}\| \leq K \quad \forall\ n = 1, 2, \ldots,\ m = 1, 2, \ldots. \tag{18}$$

6 Contraction semigroups.

In particular, if A satisfies

$$\|(I - n^{-1}A)^{-1}\| \leq 1 \tag{19}$$

condition (18) is satisfied, and such an A then generates a semi-group. Under this stronger hypothesis we can draw a stronger conclusion: In (16) we now have $p = q = \|\cdot\|$ and $K = 1$. Since $\lim_{n\to\infty} T_t^n x = T_t x$ we see that under the hypothesis (19) we can conclude that

$$\|T_t\| \leq 1 \quad \forall\, t \geq 0.$$

A semi-group T_t satisfying this condition is called a **contraction semi-group**. We will study another useful condition for recognizing a contraction semigroup in the following subsection.

We have already given a direct proof that if S is a self-adjoint operator on a Hilbert space then the resolvent exists for all non-real z and satisfies

$$\|R(z,S)\| \leq \frac{1}{|\mathrm{Im}\ (z)|}.$$

This implies (19) for $A = iS$ and $-iS$ giving us an independent proof of the existence of $U(t) = \exp(iSt)$ for any self-adjoint operator S. As we mentioned previously, we could then use Bochner's theorem to give a third proof of the spectral theorem for unbounded self-adjoint operators. I might discuss Bochner's theorem in the context of generalized functions later probably next semester if at all. Once we give an independent proof of Bochner's theorem then indeed we will get a third proof of the spectral theorem.

6.1 Dissipation and contraction.

Let $\mathbf{F}$ be a Banach space. Recall that a semi-group T_t is called a **contraction semi-group** if

$$\|T_t\| \leq 1 \quad \forall\, t \geq 0,$$

and that (19) is a sufficient condition on operator with dense domain to generate a contraction semi-group.

The Lumer-Phillips theorem to be stated below gives a necessary and sufficient condition on the infinitesimal generator of a semi-group for the semi-group to be a contraction semi-group. It is generalization of the fact that the resolvent of a self-adjoint operator has $\pm i$ in its resolvent set.

The first step is to introduce a sort of fake scalar product in the Banach space $\mathbf{F}$. A **semi-scalar product** on $\mathbf{F}$ is a rule which assigns a number $\langle\langle x, z\rangle\rangle$ to every pair of elements $x, z \in \mathbf{F}$ in such a way that

$$\begin{aligned}
\langle\langle x+y, z\rangle\rangle &= \langle\langle x, z\rangle\rangle + \langle\langle y, z\rangle\rangle \\
\langle\langle \lambda x, z\rangle\rangle &= \lambda\langle\langle x, z\rangle\rangle \\
\langle\langle x, x\rangle\rangle &= \|x\|^2 \\
|\langle\langle x, z\rangle\rangle| &\leq \|x\| \cdot \|z\|.
\end{aligned}$$

We can always choose a semi-scalar product as follows: by the Hahn-Banach theorem, for each $z \in \mathbf{F}$ we can find an $\ell_z \in \mathbf{F}^*$ such that

$$\|\ell_z\| = \|z\| \ \text{ and } \ell_z(z) = \|z\|^2.$$

Choose one such ℓ_z for each $z \in \mathbf{F}$ and set

$$\langle\langle x, z\rangle\rangle := \ell_z(x).$$

Clearly all the conditions are satisfied. Of course this definition is highly unnatural, unless there is some reasonable way of choosing the ℓ_z other than using the axiom of choice. In a Hilbert space, the scalar product is a semi-scalar product.

An operator A with domain $D(A)$ on $\mathbf{F}$ is called **dissipative** relative to a given semi-scalar product $\langle\langle\cdot,\cdot\rangle\rangle$ if

$$\text{Re } \langle\langle Ax, x\rangle\rangle \leq 0 \quad \forall\, x \in D(A).$$

For example, if A is a symmetric operator on a Hilbert space such that

$$(Ax, x) \leq 0 \ \ \forall\, x \in D(A) \tag{20}$$

then A is dissipative relative to the scalar product.

Theorem 6 [**Lumer-Phillips.**] *Let A be an operator on a Banach space $\mathbf{F}$ with $D(A)$ dense in $\mathbf{F}$. Then A generates a contraction semi-group if and only if A is dissipative with respect to any semi-scalar product and*

$$\text{im}(I - A) = \mathbf{F}.$$

Proof. Suppose first that $D(A)$ is dense and that $\text{im}(I - A) = \mathbf{F}$. We wish to show that (19) holds, which will guarantee that A generates a contraction semi-group. Let $s > 0$. Then if $x \in D(A)$ and $y = sx - Ax$ then

$$s\|x\|^2 = s\langle\langle x, x\rangle\rangle \leq s\langle\langle x, x\rangle\rangle - \text{Re } \langle\langle Ax, x\rangle\rangle = \text{Re } \langle\langle y, x\rangle\rangle$$

implying

$$s\|x\|^2 \leq \|y\|\|x\|. \tag{21}$$

We are assuming that $\text{im}(I - A) = \mathbf{F}$. This together with (21) with $s = 1$ implies that $R(1, A)$ exists and

$$\|R(1, A)\| \leq 1.$$

In turn, this implies that for all z with $|z-1|<1$ the resolvent $R(z,A)$ exists and is given by the power series

$$R(z,A)=\sum_{n=0}^{\infty}(z-1)^n R(1,A)^{n+1}$$

by our general power series formula for the resolvent. In particular, for s real and $|s-1|<1$ the resolvent exists, and then (21) implies that $\|R(s,A)\|\leq s^{-1}$. Repeating the process we keep enlarging the resolvent set $\rho(A)$ until it includes the whole real axis and conclude from (21) that $\|R(s,A)\|\leq s^{-1}$ which implies (19). As we are assuming that $D(A)$ is dense we conclude that A generates a contraction semigroup.

Conversely, suppose that T_t is a contraction semi-group with infinitesimal generator A. We know that A is dense. Let $\langle\langle\cdot,\cdot\rangle\rangle$ be any semi-scalar product. Then

$$\text{Re }\langle\langle T_t x-x,x\rangle\rangle=\text{Re }\langle\langle T_t x,x\rangle\rangle-\|x\|^2\leq\|T_t x\|\|x\|-\|x\|^2\leq 0.$$

Dividing by t and letting $t\searrow 0$ we conclude that Re $\langle\langle Ax,x\rangle\rangle\leq 0$ for all $x\in D(A)$, i.e. A is dissipative for $\langle\langle\cdot,\cdot\rangle\rangle$. QED

Once again, this gives a direct proof of the existence of the unitary group generated generated by a skew adjoint operator.

A useful way of verifying the condition $\text{im}(I-A)=\mathbf{F}$ is the following: Let $A^*:\mathbf{F}^*\to\mathbf{F}^*$ be the adjoint operator which is defined if we assume that $D(A)$ is dense.

Proposition 2 *Suppose that A is densely defined and closed, and suppose that both A and A^* are dissipative. Then $\text{im}(I-A)=\mathbf{F}$ and hence A generates a contraction semigroup.*

Proof. The fact that A is closed implies that $(I-A)^{-1}$ is closed, and since we know that $(I-A)^{-1}$ is bounded from the fact that A is dissipative, we conclude that $\text{im}(I-A)$ is a closed subspace of F. If it were not the whole space there would be an $\ell\in F^*$ which vanished on this subspace, i.e.

$$\langle\ell,x-Ax\rangle=0\quad\forall\ x\in D(A).$$

This implies that that $\ell\in D(A^*$ and $A^*\ell=\ell$ which can not happen if A^* is dissipative by (21) applied to A^* and $s=1$. QED

6.2 A special case: $\exp(t(B-I))$ with $\|B\|\leq 1$.

Suppose that $B:F\to F$ is a bounded operator on a Banach space with $\|B\|\leq 1$. Then for any semi-scalar product we have

$$\text{Re }\langle\langle(B-I)x,x\rangle\rangle=\text{Re }\langle\langle Tx,x\rangle\rangle-\|x\|^2\leq\|Bx\|\|x\|-\|x\|^2\leq 0$$

so $B-I$ is dissipative and hence $\exp(t(B-I))$ exists as a contraction semi-group by the Lumer-Phillips theorem. We can prove this directly since we can write

$$\exp(t(B-I)) = e^{-t}\sum_{k=0}^{\infty}\frac{t^k B^k}{k!}.$$

The series converges in the uniform norm and we have

$$\|\exp(t(B-I))\| \le e^{-t}\sum_{k=0}^{\infty}\frac{t^k\|B\|^k}{k!} \le 1.$$

For future use (Chernoff's theorem and the Trotter product formula) we record (and prove) the following inequality:

$$\|[\exp(n(B-I)) - B^n]x\| \le \sqrt{n}\|(B-I)x\| \quad \forall\ x \in \mathbf{F},\ \text{and}\ \forall\ n = 1,2,3\dots. \tag{22}$$

Proof.

$$\begin{aligned}
\|[\exp(n(B-I)) - B^n]x\| &= \|e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}(B^k - B^n)x\| \\
&\le e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}\|(B^k - B^n)x\| \\
&\le e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}\|(B^{|k-n|} - I)x\| \\
&= e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}\|(B-I)(I + B + \cdots + B^{(|k-n|-1})x\| \\
&\le e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}|k-n|\|(B-I)x\|.
\end{aligned}$$

So to prove (22) it is enough establish the inequality

$$e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}|k-n| \le \sqrt{n}. \tag{23}$$

Consider the space of all sequences $\mathbf{a} = \{a_0, a_1, \dots\}$ with scalar product

$$(\mathbf{a},\mathbf{b}) := e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}a_k\overline{b_k}.$$

The Cauchy-Schwartz inequality applied to $\mathbf{a}$ with $a_k = |k-n|$ and $\mathbf{b}$ with $b_k \equiv 1$ gives

$$e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}|k-n| \le \sqrt{e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}(k-n)^2}\cdot\sqrt{e^{-n}\sum_{k=0}^{\infty}\frac{n^k}{k!}}.$$

The second square root is one, and we recognize the sum under the first square root as the variance of the Poisson distribution with parameter n, and we know that this variance is n. QED

7 Convergence of semigroups.

We are going to be interested in the following type of result. We would like to know that if A_n is a sequence of operators generating equibounded one parameter groups $\exp tA_n$ and $A_n \to A$ where A generates an equibounded semi-group $\exp tA$ then the semi-groups converge, i.e. $\exp tA_n \to \exp tA$. We will prove such a result for the case of contractions. But before we can even formulate the result, we have to deal with the fact that each A_n comes equipped with its own domain of definition, $D(A_n)$. We do not want to make the overly restrictive hypothesis that these all coincide, since in many important applications they won't.

For this purpose we make the following definition. Let us assume that $\mathbf{F}$ is a Banach space and that A is an operator on $\mathbf{F}$ defined on a domain $D(A)$. We say that a linear subspace $\mathbf{D} \subset D(A)$ is a **core** for A if the closure $\overline{A}$ of A and the closure of A restricted to $\mathbf{D}$ are the same: $\overline{A} = \overline{A|\mathbf{D}}$. This certainly implies that $D(A)$ is contained in the closure of $A|\mathbf{D}$. In the cases of interest to us $D(A)$ is dense in $\mathbf{F}$, so that every core of A is dense in $\mathbf{F}$.

We begin with an important preliminary result:

Proposition 3 *Suppose that A_n and A are dissipative operators, i.e. generators of contraction semi-groups. Let $\mathbf{D}$ be a core of A. Suppose that for each $x \in \mathbf{D}$ we have that $x \in D(A_n)$ for sufficiently large n (depending on x) and that*

$$A_n x \to Ax. \tag{24}$$

Then for any z with $\mathrm{Re}\ z > 0$ *and for all $y \in \mathbf{F}$*

$$R(z, A_n)y \to R(z, A)y. \tag{25}$$

Proof. We know that the $R(z, A_n)$ and $R(z, A)$ are all bounded in norm by $1/\mathrm{Re}\ z$. So it is enough for us to prove convergence on a dense set. Since $(zI - A)D(A) = \mathbf{F}$, it follows that $(zI - A)D$ is dense in $\mathbf{F}$. So in proving (25) we may assume that $y = (zI - A)x$ with $x \in D$. Then

$$\begin{aligned} \|R(z, A_n)y - R(z, A)y\| &= \|R(z, A_n)(zI - A)x - x\| \\ &= \|R(z, A_n)(zI - A_n)x + R(z, A_n)(A_n x - Ax) - x\| \\ &= \|R(z, A_n)(A_n - A)x\| \\ &\leq \frac{1}{\mathrm{Re}\ z}\|(A_n - A)x\| \to 0, \end{aligned}$$

where, in passing from the first line to the second we are assuming that n is chosen sufficiently large that $x \in D(A_n)$. QED

Theorem 7 *Under the hypotheses of the preceding proposition,*

$$(\exp(tA_n))x \to (\exp(tA))x$$

for each $x \in \mathbf{F}$ *uniformly on every compact interval of* t.

Proof. Let

$$\phi_n(t) := e^{-t}[((\exp(tA_n))x - (\exp(tA))x)] \quad \text{for } t \geq 0$$

and set $\phi(t) = 0$ for $t < 0$. It will be enough to prove that these $\mathbf{F}$ valued functions converge uniformly in t to 0, and since $\mathbf{D}$ is dense and since the operators entering into the definition of ϕ_n are uniformly bounded in n, it is enough to prove this convergence for $x \in \mathbf{D}$ which is dense. We claim that for fixed $x \in \mathbf{D}$ the functions $\phi_n(t)$ are uniformly equi-continuous. To see this observe that

$$\frac{d}{dt}\phi_n(t) = e^{-t}[(\exp(tA_n))A_n x - (\exp(tA))Ax] - e^{-t}[(\exp(tA_n))x - (\exp(tA))x]$$

for $t \geq 0$ and the right hand side is uniformly bounded in $t \geq 0$ and n.

So to prove that $\phi_n(t)$ converges uniformly in t to 0, it is enough to prove this fact for the convolution $\phi_n \star \rho$ where ρ is any smooth function of compact support, since we can choose the ρ to have small support and integral $\sqrt{2\pi}$, and then $\phi_n(t)$ is close to $(\phi_n \star \rho)(t)$.

Now the Fourier transform of $\phi_n \star \rho$ is the product of their Fourier transforms: $\hat{\phi_n}\hat{\rho}$. We have $\hat{\phi_n}(s) =$

$$\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{(-1-is)t}[(\exp tA_n)x - (\exp(tA))x]dt = \frac{1}{\sqrt{2\pi}}[R(1+is, A_n)x - R(1+is, A)x].$$

Thus by the proposition

$$\hat{\phi_n}(s) \to 0,$$

in fact uniformly in s. Hence using the Fourier inversion formula and, say, the dominated convergence theorem (for Banach space valued functions),

$$(\phi_n \star \phi)(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \hat{\phi_n}(s)\hat{\rho}(s)e^{ist}ds \to 0$$

uniformly in t. QED

The preceding theorem is the limit theorem that we will use in what follows. However, there is an important theorem valid in an arbitrary Frechet space, and which does not assume that the A_n converge, or the existence of the limit A, but only the convergence of the resolvent at a single point z_0 in the right hand plane!

In the following $\mathbf{F}$ is a Frechet space and $\{\exp(tA_n)\}$ is a family of of equibounded semi-groups which is also equibounded in n, so for every semi-norm p there is a semi-norm q and a constant K such that

$$p(\exp(tA_n)x) \leq Kq(x) \quad \forall\, x \in F$$

where K and q are independent of t and n. I will state the theorem here, and refer you to Yosida pp.269-271 for the proof.

Theorem 8 **[Trotter-Kato.]** *Suppose that* $\{\exp(tA_n)\}$ *is an equibounded family of semi-groups as above, and suppose that for some* z_0 *with positive real part there exist an operator* $R(z_0)$ *such that*

$$\lim_{n\to\infty} R(z_0, A_n) \to R(z_0)$$

and

$$\mathrm{im}\, R(z_0) \quad \textit{is dense in } \mathbf{F}.$$

Then there exists an equibounded semi-group $\exp(tA)$ *such that*

$$R(z_0) = R(z_0, A)$$

and

$$\exp(tA_n) \to \exp(tA)$$

uniformly on every compact interval of $t \geq 0$.

8 The Trotter product formula.

In what follows, $\mathbf{F}$ is a Banach space. Eventually we will restrict attention to a Hilbert space.

8.1 Chernoff's theorem.

Theorem 9 **[Chernoff.]** *Let* $f : [0, \infty) \to$ *bounded operators on* $\mathbf{F}$ *with*

$$\|f(t)\| \leq 1 \quad \forall\, t$$

and

$$f(0) = I.$$

Let A *be a dissipative operator and* $\exp tA$ *the contraction semi-group it generates. Let* $\mathbf{D}$ *be a core of* A. *Suppose that*

$$\lim_{h\searrow 0} \frac{1}{h}[f(h) - I]x = Ax \quad \forall\, x \in \mathbf{D}.$$

Then for all $y \in \mathbf{F}$

$$\lim \left[f\left(\frac{t}{n}\right)\right]^n y = (\exp tA)y \tag{26}$$

uniformly in any compact interval of $t \geq 0$.

Proof. For fixed $t > 0$ let

$$C_n := \frac{n}{t}\left[f\left(\frac{t}{n}\right) - I\right].$$

Then $\frac{t}{n}C_n$ generates a contraction semi-group by the special case of the Lumer-Phillips theorem discussed in Section 6.2, and therefore (by change of variable), so does C_n. So C_n is the generator of a semi-group

$$\exp tC_n$$

and the hypothesis of the theorem is that $C_n x \to Ax$ for $x \in \mathbf{D}$. Hence by the limit theorem in the preceding section

$$(\exp tC_n)y \to (\exp tA)y$$

for each $y \in \mathbf{F}$ uniformly on any compact interval of t. Now

$$\exp(tC_n) = \exp n\left[f\left(\frac{t}{n}\right) - I\right]$$

so we may apply (22) to conclude that

$$\| \left(\exp(tC_n) - f\left(\frac{t}{n}\right)^n\right) x\| \leq \sqrt{n}\| \left(f\left(\frac{t}{n}\right) - I\right) x\| = \frac{t}{\sqrt{n}}\|\frac{n}{t}\left(f\left(\frac{t}{n}\right) - I\right) x\|.$$

The expression inside the $\|\cdot\|$ on the right tends to Ax so the whole expression tends to zero. This proves (26) for all x in $\mathbf{D}$. But since $\mathbf{D}$ is dense in $\mathbf{F}$ and $f(t/n)$ and $\exp tA$ are bounded in norm by 1 if follows that (26) holds for all $y \in \mathbf{F}$. QED

8.2 The product formula.

Let A and B be the infinitesimal generators of the contraction semi-groups $P_t = \exp tA$ and $Q_t = \exp tB$ on the Banach space F. Then $A + B$ is only defined on $D(A) \cap D(B)$ and in general we know nothing about this intersection. However let us *assume* that $D(A) \cap D(B)$ is sufficiently large that the closure $\overline{A + B}$ is a densely defined operator and $\overline{A + B}$ is in fact the generator of a contraction semi-group R_t. So $\mathbf{D} := D(A) \cap D(B)$ is a core for $\overline{A + B}$.

Theorem 10 [Trotter.] *Under the above hypotheses*

$$R_t y = \lim \left(P_{\frac{t}{n}} Q_{\frac{t}{n}}\right)^n y \quad \forall\, y \in \mathbf{F} \tag{27}$$

uniformly on any compact interval of $t \geq 0$.

Proof. Define

$$f(t) = P_t Q_t.$$

For $x \in D$ we have

$$f(t)x = P_t(I + tB + o(t))x = (I + At + Bt + o(t))x$$

so the hypotheses of Chernoff's theorem are satisfied. The conclusion of Chernoff's theorem asserts (27). QED

A symmetric operator on a Hilbert space is called **essentially self adjoint** if its closure is self-adjoint. So a reformulation of the preceding theorem in the case of self-adjoint operators on a Hilbert space says

Theorem 11 *Suppose that S and T are self-adjoint operators on a Hilbert space H and suppose that $S+T$ (defined on $D(S) \cap D(T)$) is essentially self-adjoint. Then for every $y \in H$*

$$\exp(it((\overline{S+T}))y = \lim_{n\to\infty} \left(\exp(\frac{t}{n}iA)(\exp\frac{t}{n}iB)\right)^n y \tag{28}$$

where the convergence is uniform on any compact interval of t.

8.3 Commutators.

An operator A on a Hilbert space is called skew-symmetric if $A^* = -A$ on $D(A)$. This is the same as saying that iA is symmetric. So we call an operator skew adjoint if iA is self-adjoint. We call an operator A **essentially skew adjoint** if iA is essentially self-adjoint.

If A and B are bounded skew adjoint operators then their Lie bracket

$$[A, B] := AB - BA$$

is well defined and again skew adjoint.

In general, we can only define the Lie bracket on $D(AB) \cap D(BA)$ so we again must make some rather stringent hypotheses in stating the following theorem.

Theorem 12 *Let A and B be skew adjoint operators on a Hilbert space H and let*

$$\mathbf{D} := D(A^2) \cap D(B^2) \cap D(AB) \cap D(BA).$$

Suppose that the restriction of $[A, B]$ to $\mathbf{D}$ is essentially skew-adjoint. Then for every $y \in \mathbf{H}$

$$\exp t\overline{[A,B]}y = \lim_{n\to\infty} \left((\exp -\sqrt{\frac{t}{n}}A)(\exp -\sqrt{\frac{t}{n}}B)(\exp\sqrt{\frac{t}{n}}A)(\exp\sqrt{\frac{t}{n}}B)\right)^n y \tag{29}$$

uniformly in any compact interval of $t \geq 0$.

Proof. The restriction of $[A, B]$ to $\mathbf{D}$ is assumed to be essentially skew-adjoint, so $[A, B]$ itself (which has the same closure) is also essentially skew adjoint.

We have

$$\exp(tA)x = (I + tA + \frac{t^2}{2}A^2)x + o(t^2)$$

for $x \in D$ with similar formulas for $\exp(-tA)$ etc.

Let

$$f(t) := (\exp -tA)(\exp -tB)(\exp tA)(\exp tB).$$

Multiplying out $f(t)x$ for $x \in D$ gives a whole lot of cancellations in and yields

$$f(s)x = (I + s^2[A, B])x + o(s^2)$$

so (29) is a consequence of Chernoff's theorem with $s = \sqrt{t}$. QED

We still need to develop some methods which allow us to check the hypotheses of the last three theorems.

8.4 The Kato-Rellich theorem.

This is the starting point of a class of theorems which asserts that that if A is self-adjoint and if B is a symmetric operator which is "small" in comparison to A then $A + B$ is self adjoint.

Theorem 13 **[Kato-Rellich.]** *Let A be a self-adjoint operator and B a symmetric operator with*

$$D(B) \supset D(A)$$

and

$$\|Bx\| \leq a\|Ax\| + b\|x\| \quad 0 \leq a < 1, \quad \forall\ x \in D(A).$$

Then $A + B$ is self-adjoint, and is essentially self-adjoint on any core of A.

Proof. [Following Reed and Simon II page 162.] To prove that $A + B$ is self adjoint, it is enough to prove that $\mathrm{im}(A + B \pm i\mu_0) = H$. We do this for $A + B + i\mu_0$. The proof for $A + B - i\mu_0$ is identical.

Let $\mu > 0$. Since A is self-adjoint, we know that

$$\|(A + i\mu)x\|^2 = \|Ax\|^2 + \mu^2\|x\|^2$$

from which we concluded that $(A + i\mu)^{-1}$ maps $\mathbf{H}$ onto $D(A)$ and

$$\|(A + i\mu)^{-1}\| \leq \frac{1}{\mu}, \quad \|A(A + i\mu)^{-1}\| \leq 1.$$

Applying the hypothesis of the theorem to $x = (A + i\mu)^{-1}y$ we conclude that

$$\|B(A + i\mu)^{-1}y\| \leq a\|A((A + i\mu)^{-1}y\| + b\|(A + i\mu)^{-1}y\| \leq \left(a + \frac{b}{\mu}\right)\|y\|.$$

Thus for $\mu >> 1$, the operator

$$C := B(A + i\mu)^{-1}$$

satisfies

$$\|C\| < 1$$

since $a < 1$. Thus $-1 \notin \mathrm{Spec}(A)$ so $\mathrm{im}(I + C) = \mathbf{H}$. We know that $\mathrm{im}(A + \mu I) = \mathbf{H}$ hence

$$\mathbf{H} = \mathrm{im}(I + C) \circ (A + i\mu I) = \mathrm{im}(A + B + i\mu I)$$

proving that $A + B$ is self-adjoint.

If $\mathbf{D}$ is any core for A, it follows immediately from the inequality in the hypothesis of the theorem that the closure of $A + B$ restricted to $\mathbf{D}$ contains $D(A)$ in its domain. Thus $A + B$ is essentially self-adjoint on any core of A. QED

8.5 Feynman path integrals.

Consider the operator

$$H_0 : \quad L_2(\mathbf{R}^3) \to L_2(\mathbf{R}^3)$$

given by

$$H_0 := -\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right).$$

Here the domain of H_0 is taken to be those $\phi \in L_2(\mathbf{R}^3)$ for which the differential operator on the right, taken in the distributional sense, when applied to ϕ gives an element of $L_2(\mathbf{R}^3)$.

The operator H_0 is called the "free Hamiltonian of non-relativistic quantum mechanics". The Fourier transform $\mathcal{F}$ is a unitary isomorphism of $L_2(\mathbf{R}^3)$ into $L_2(\mathbf{R}_3)$ and carries H_0 into multiplication by ξ^2 whose domain consists of those $\hat{\phi} \in L_2(\mathbf{R}_3)$ such that $\xi^2\hat{\phi}(\xi)$ belongs to $L_2(\mathbf{R}_3)$. The operator consisting of multiplication by $e^{-it\xi^2}$ is clearly unitary, and provides us with a unitary one parameter group. Transferring this one parameter group back to $L_2(\mathbf{R}^3)$ via the Fourier transform gives us a one parameter group of unitary transformations whose infinitesimal generator is $-iH_0$.

Now the Fourier transform carries multiplication into convolution, and the inverse Fourier transform (in the distributional sense) of $e^{-i\xi^2 t}$ is $(2it)^{-3/2}e^{ix^2/4t}$. Hence we can write, in a formal sense,

$$(\exp(-itH_0)f)(x) = (4\pi it)^{-3/2}\int_{\mathbf{R}^3} \exp\left(\frac{i(x-y)^2}{4t}\right) f(y)dy.$$

Here the right hand side is to be understood as a long winded way of writing the left hand side which is well defined as a mathematical object. The right hand side can also be regarded as an actual integral for certain classes of f, and as the L_2 limit of such such integrals. We shall discuss this interpretation in Section 10.

Let V be a function on $\mathbf{R}^3$. We denote the operator on $L_2(\mathbf{R}^3)$ consisting of multiplication by V also by V. Suppose that V is such that H_0+V is again self-adjoint. For example, if V were continuous and of compact support this would certainly be the case by the Kato-Rellich theorem. (Realistic "potentials" V will not be of compact support or be bounded, but nevertheless in many important cases the Kato-Rellich theorem does apply.)

Then the Trotter product formula says that

$$\exp -it(H_0+V) = \lim_{n\to\infty}\left(\exp(-i\frac{t}{n}H_0)(\exp -i\frac{t}{n}V)\right)^n .$$

We have

$$\left((\exp -i\frac{t}{n}V)f\right)(x) = e^{-i\frac{t}{n}V(x)}f(x).$$

Hence we can write the expression under the limit sign in the Trotter product formula, when applied to f and evaluated at x_0 as the following formal expression:

$$\left(\frac{4\pi it}{n}\right)^{-3n/2}\int_{\mathbf{R}^3}\cdots\int_{\mathbf{R}^3}\exp(iS_n(x_0,\ldots,x_n))f(x_n)dx_n\cdots dx_1$$

where

$$S_n(x_0,x_1,\ldots,x_n,t) := \sum_{i=1}^{n}\frac{t}{n}\left[\frac{1}{4}\left(\frac{(x_i-x_{i-1})}{t/n}\right)^2 - V(x_i)\right].$$

If $X : s\mapsto X(s),\ 0\le s\le t$ is a piecewise differentiable curve, then the **action** of a particle of mass m moving along this curve is defined in classical mechanics as

$$S(X) := \int_0^t\left(\frac{m}{2}\dot{X}(s)^2 - V(X(s))\right)ds$$

where $\dot{X}$ is the velocity (defined at all but finitely many points).

Take $m=\frac{1}{2}$ and let X be the polygonal path which goes from x_0 to x_1, from x_1 to x_2 etc., each in time t/n so that the velocity is $|x_i-x_{i-1}|/(t/n)$ on the i-th segment. Also, the integral of $V(X(s))$ over this segment is approximately $\frac{t}{n}V(x_i)$. The formal expression written above for the Trotter product formula can be thought of as an integral over polygonal paths (with step length t/n) of $e^{iS_n(X)}f(X(t))d_nX$ where S_n approximates the classical action and where d_nX is a measure on this space of polygonal paths.

This suggests that an intuitive way of thinking about the Trotter product formula in this context is to imagine that there is some kind of "measure" dX on the space Ω_{x_0} of *all* continuous paths emanating from x_0 and such that

$$\exp(-it(H_0+V)f)(x) = \int_{\Omega_{x_0}}e^{iS(X)}f(X(t))dX.$$

This formula was suggested in 1942 by Feynman in his thesis (Trotter's paper was in 1959), and has been the basis of an enormous number of important calculations in physics, many of which have given rise to exciting mathematical theorems which were then proved by other means. I am unaware of any general mathematical justification of these "path integral" methods in the form that they are used.

9 The Feynman-Kac formula.

An important advance was introduced by Mark Kac in 1951 where the unitary group $\exp -it(H_0+V)$ is replaced by the contraction semi-group $\exp -t(H_0+V)$. Then the techniques of probability theory (in particular the existence of Wiener measure on the space of continuous paths) can be brought to bear to justify a formula for the contractive semi-group as an integral over path space. I will state and prove an elementary version of this formula which follows directly from what we have done. The assumptions about the potential are physically unrealistic, but I choose to regard the extension to a more realistic potential as a technical issue rather than a conceptual one.

Let V be a continuous real valued function of compact support. To each continuous path ω on $\mathbf{R}^n$ and for each fixed time $t \geq 0$ we can consider the integral

$$\int_0^t V(\omega(s))ds.$$

The map

$$\omega \mapsto \int_0^t V(\omega(s))ds \tag{30}$$

is a continuous function on the space of continuous paths, and we have

$$\frac{t}{m}\sum_{j=1}^{m} V\left(\omega\left(\frac{jt}{m}\right)\right) \to \int_0^t V(\omega(s))ds \tag{31}$$

for each fixed ω.

Theorem 14 The Feynman-Kac formula. *Let V be a continuous real valued function of compact support on $\mathbf{R}^n$. Let*

$$H = \Delta + V$$

as an operator on $\mathbf{H} = L^2(\mathbf{R}^n)$. Then H is self-adjoint and for every $f \in \mathbf{H}$

$$\left(e^{-tH}f\right)(x) = \int_{\Omega_x} f(\omega(t)) \exp\left(\int_0^t V(\omega(s))ds\right) d_x\omega \tag{32}$$

where Ω_x is the space of continuous paths emanating from x and $d_x\omega$ is the associated Wiener measure.

Proof. [From Reed-Simon II page 280.] Since multiplication by V is a bounded self-adjoint operator, we can apply the Kato-Rellich theorem (with $a = 0$!) to conclude that H is self-adjoint, and with the same domain as Δ. So we may apply the Trotter product formula to conclude that

$$(e^{-Ht})f = \lim_{m\to\infty} \left(e^{-\frac{t}{m}\Delta} e^{-\frac{t}{m}V}\right)^m f.$$

This convergence is in L^2, but by passing to a subsequence we may also assume that the convergence is almost everywhere. Now

$$\left[\left(e^{-\frac{t}{m}\Delta} e^{-\frac{t}{m}V}\right)^m f\right](x)$$

$$= \int_{\mathbf{R}^n} \cdots \int_{\mathbf{R}^n} p\left(x, x_m, \frac{t}{m}\right) \cdots p\left(x, x_m, \frac{t}{m}\right) f(x)_1 \exp\left(-\sum_{j=1}^m \frac{t}{m} V(x_j)\right) dx_1 \cdots dx_m.$$

By the very definition of Wiener measure, this last expression is

$$\int_{\Omega_x} \exp\left(\frac{t}{m}\sum_{j=1}^m V\left(\omega\left(\frac{jt}{m}\right)\right)\right) f(\omega(t)) d_x\omega.$$

The integrand (with respect to the Wiener measure $d_x\omega$) converges on all continuous paths, that is to say almost everywhere with respect to $d_x\mu$ to the integrand on right hand side of (32). So to justify (32) we must prove that the integral of the limit is the limit of the integral. We will do this by the dominated convergence theorem:

$$\int_{\Omega_x} \left|\exp\left(\frac{t}{m}\sum_{j=1}^m V\left(\omega\left(\frac{jt}{m}\right)\right)\right) f(\omega(t))\right| d_x\omega$$

$$\leq e^{t\max|V|} \int_{\Omega_x} |f(\omega(t))| d_x\omega = e^{t\max|V|} \left(e^{-t\Delta}|f|\right)(x) < \infty$$

for almost all x. Hence, by the dominated convergence theorem, (32) holds for almost all x. QED

10 The free Hamiltonian and the Yukawa potential.

In this section I want to discuss the following circle of ideas. Consider the operator

$$H_0 : \quad L_2(\mathbf{R}^3) \to L_2(\mathbf{R}^3)$$

given by

$$H_0 := -\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right).$$

Here the domain of H_0 is taken to be those $\phi \in L_2(\mathbf{R}^3)$ for which the differential operator on the right, taken in the distributional sense, when applied to ϕ gives an element of $L_2(\mathbf{R}^3)$.

The operator H_0 has a fancy name. It is called the "free Hamiltonian of non-relativistic quantum mechanics". Strictly speaking we should add "for particles of mass one in units where Planck's constant is one".

The Fourier transform is a unitary isomorphism of $L_2(\mathbf{R}^3)$ into $L_2(\mathbf{R}_3)$ and carries H_0 into multiplication by ξ^2 whose domain consists of those $\hat{\phi} \in L_2(\mathbf{R}_3)$ such that $\xi^2 \hat{\phi}(\xi)$ belongs to $L_2(\mathbf{R}_3)$. The operators

$$V(t) : L_2(\mathbf{R}_3) \to L_2(\mathbf{R}_3), \qquad \hat{\phi}(\xi) \mapsto e^{-it\xi^2} \hat{\phi}$$

form a one parameter group of unitary transformations whose infinitesimal generator in the sense of Stone' theorem is operator consisting of multiplication by ξ^2 with domain as given above. [The minus sign before the i in the exponential is the convention used in quantum mechanics. So we write $\exp -itA$ for the one-parameter group associated to the self-adjoint operator A. I apologize for this (rather irrelevant) notational change, but I want to make the notation in this problem set consistent with what you will see in physics books.]

Thus the operator of multiplication by ξ^2, and hence the operator H_0 is a self-adjoint transformation. The operator of multiplication by ξ^2 is clearly non-negative and so every point on the negative real axis belongs to its resolvent set. Let us write a point on the negative real axis as $-\mu^2$ where $\mu > 0$. Then the resolvent is given by multiplication by $-f$ where

$$f(\xi) = f_\mu(\xi) := \frac{1}{\mu^2 + \xi^2}.$$

We can summarize what we "know" so far as follows:

1. The operator H_0 is self adjoint.

2. The one parameter group of unitary transformations it generates via Stone's theorem is
$$U(t) = \mathcal{F}^{-1} V(t) \mathcal{F}$$
where $V(t)$ is multiplication by $e^{-it\xi^2}$.

3. Any point $-\mu^2$, $\mu > 0$ lies in the resolvent set of H_0 and
$$R(-\mu^2, H_0) = -\mathcal{F}^{-1} m_f \mathcal{F}$$
where m_f denotes the operation of multiplication by f and f is as given above.

4. If $g \in \mathcal{S}$ and m_g denotes multiplication by g, then the the operator $\mathcal{F}^{-1} m_g \mathcal{F}$ consists of convolution by $\check{g}$. Neither the function $e^{-it\xi^2}$ nor the function f belongs to $\mathcal{S}$, so the operators $U(t)$ and $R(-\mu^2, H_0)$ can only be thought of as convolutions in the sense of generalized functions.

10.1 The Yukawa potential and the resolvent.

Nevertheless, we will be able to give some slightly more explicit (and very instructive) representations of these operators as convolutions. For example, we will use the Cauchy residue calculus to compute $\check{f}$ and we will find, up to factors of powers of 2π that $\check{f}$ is the function

$$Y_\mu(x) := \frac{e^{-\mu r}}{r}$$

where r denotes the distance from the origin, i.e. $r^2 = x^2$. This function has an integrable singularity at the origin, and vanishes rapidly at infinity. So convolution by Y_μ will be well defined and given by the usual formula on elements of $\mathcal{S}$ and extends to an operator on $L_2(\mathbf{R}^3)$.

The function Y_μ is known as the **Yukawa potential**. Yukawa introduced this function in 1934 to explain the forces that hold the nucleus together. The exponential decay with distance contrasts with that of the ordinary electromagnetic or gravitational potential $1/r$ and, in Yukawa's theory, accounts for the fact that the nuclear forces are short range. In fact, Yukawa introduced a "heavy boson" to account for the nuclear forces. The role of mesons in nuclear physics was predicted by brilliant theoretical speculation well before any experimental discovery. Here are the details:

Since $f \in L_2$ we can compute its inverse Fourier transform as

$$(2\pi)^{-3/2}\check{f} = \lim_{R\to\infty}(2\pi)^{-3}\int_{|\xi|\leq R}\frac{e^{i\xi\cdot x}}{\mu^2+\xi^2}d\xi. \tag{33}$$

Here lim means the L_2 limit and $|\xi|$ denotes the length of the vector ξ, i.e. $|\xi| = \sqrt{\xi^2}$ and we will use similar notation $|x| = r$ for the length of x. Assume $x \neq 0$. Let

$$u := \frac{\xi\cdot x}{|\xi||x|}$$

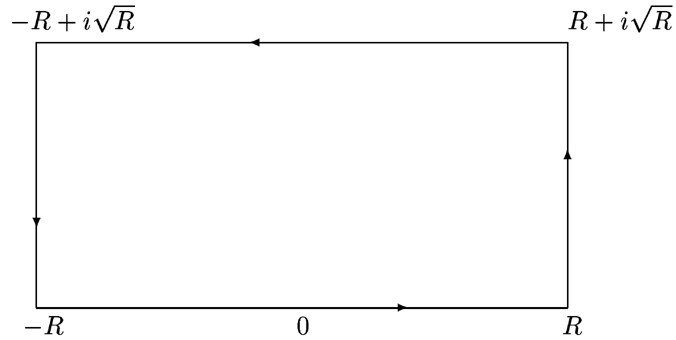
so u is the cosine of the angle between x and ξ. Fix x and introduce spherical coordinates in ξ space with x at the north pole and $s = |\xi|$ so that

$$(2\pi)^{-3}\int_{|\xi|\leq R}\frac{e^{i\xi\cdot x}}{\mu^2+\xi^2}d\xi = (2\pi)^{-2}\int_0^R\int_{-1}^1\frac{e^{is|x|u}}{s^2+\mu^2}s^2duds$$

$$= \frac{1}{(2\pi)^2 i|x|}\int_{-R}^R\frac{se^{is|x|}}{(s+i\mu)(s-i\mu)}ds.$$

This last integral is along the bottom of the path in the complex s-plane consisting of the boundary of the rectangle as drawn in the figure.

On the two vertical sides of the rectangle, the integrand is bounded by some constant time $1/R$, so the contribution of the vertical sides is $O(1/\sqrt{R})$. On the top the integrand is $O(e^{-\sqrt{R}})$. So the limits of these integrals are zero. There



is only one pole in the upper half plane at $s = i\mu$, so the integral is given by $2\pi i\times$ this residue which equals

$$2\pi i \frac{i\mu e^{-\mu|x|}}{2i\mu} = \pi i e^{-\mu|x|}$$

. Inserting this back into (33) we see that the limit exists and is equal to

$$(2\pi)^{-3/2}\hat{f} = \frac{1}{4\pi}\frac{e^{-\mu|x|}}{|x|}.$$

We conclude that for $\phi \in \mathcal{S}$

$$[(H_0+\mu^2)^{-1}\phi](x) = \frac{1}{4\pi}\int_{\mathbf{R}^3}\frac{e^{-\mu|x-y|}}{|x-y|}\phi(y)dy,$$

and since $(H_0+\mu^2)^{-1}$ is a bounded operator on L_2 this formula extends in the L_2 sense to L_2.

10.2 The time evolution of the free Hamiltonian.

The "explicit" calculation of the operator $U(t)$ is slightly more tricky. The function $\xi \mapsto e^{-it\xi^2}$ is an "imaginary Gaussian", so we expect is inverse Fourier transform to also be an imaginary Gaussian, and then we would have to make sense of convolution by a function which has absolute value one at all points. There are several ways to proceed. One involves integration by parts, and I hope to explain how this works later on in the course in conjunction with the method of stationary phase.

Here I will follow Reed-Simon vol II p.59 and add a little positive term to t and then pass to the limit. In other words, let α be a complex number with positive real part and consider the function

$$\xi \mapsto e^{-\xi^2\alpha}$$

This function belongs to $\mathcal{S}$ and its inverse Fourier transform is given by the function

$$x \mapsto (2\alpha)^{-3/2}e^{-x^2/4\alpha}.$$

(In fact, we verified this when α is real, but the integral defining the inverse Fourier transform converges in the entire half plane Re $\alpha > 0$ uniformly in any Re $\alpha > \epsilon$ and so is holomorphic in the right half plane. So the formula for real positive α implies the formula for α in the half plane.)

We thus have

$$(e^{-H_0\alpha}\phi)(x) = \left(\frac{1}{4\pi\alpha}\right)^{3/2}\int_{\mathbf{R}^3} e^{-|x-y|^2/4\alpha}\phi(y)dy.$$

Here the square root in the coefficient in front of the integral is obtained by continuation from the positive square root on the positive axis. For example, if we take $\alpha = \epsilon + it$ so that $-\alpha = -i(t - i\epsilon)$ we get

$$(U(t)\phi)(x) = \lim_{\epsilon \searrow 0}(U(t-i\epsilon)\phi)(x) = \lim_{\epsilon \searrow 0}(4\pi i(t-i\epsilon))^{-\frac{3}{2}}\int e^{-|x-y|^2/4i(t-i\epsilon)}\phi(y)dy.$$

Here the limit is in the sense of L_2. We thus could write

$$(U(t))(\phi)(x) = (4\pi i)^{-3/2}\int e^{i|x-y|^2/4t}\phi(y)dy$$

if we understand the right hand side to mean the $\epsilon \searrow 0$ limit of the preceding expression.

Actually, as Reed and Simon point out, if $\phi \in L_1$ the above integral exists for any $t \neq 0$, so if $\phi \in L_1 \cap L_2$ we should expect that the above integral is indeed the expression for $U(t)\phi$. Here is their argument: We know that

$$\exp(-i(t - i\epsilon))\phi \to U(t)\phi$$

in the sense of L_2 convergence as $\epsilon \searrow 0$. Here we a theorem from measure theory which says that if you have an L_2 convergent sequence you can choose a subsequence which also converges pointwise almost everywhere. So choose a subsequence of ϵ for which this happens. But then the dominated convergence theorem kicks in to guarantee that the integral of the limit is the limit of the integrals.

To sum up: The function

$$P_0(x, y; t) := (4\pi it)^{-3/2}e^{i|x-y|/4t}$$

is called the **free propagator**. For $\phi \in L_1 \cap L_2$

$$[U(t)\phi](x) = \int_{\mathbf{R}^3} P_0(x, y; t)\phi(y)dy$$

and the integral converges. For general elements ψ of L_2 the operator $U(t)\psi$ is obtained by taking the L_2 limit of the above expression for any sequence of elements of $L_1 \cap L_2$ which approximate ψ in L_2. Alternatively, we could interpret the above integral as the $\epsilon \searrow$ limit of the corresponding expression with t replaced by $t - i\epsilon$.