Math2121417
Convergence of semi-groups.
The Trotter product formula.
Feynman path integrals.

Shlomo Sternberg

November 13, 2014

A theorem of Lie

Let A and B be linear operators on a finite dimensional Hilbert space. Lie's formula says that

$$\exp(A+B) = \lim_{n\to\infty} \left[(\exp A/n)(\exp B/n)\right]^n . \tag{1}$$

Let $S_n := \exp(\frac{1}{n}(A+B))$ so that

$$S_n^n = \exp(A+B).$$

Let $T_n = (\exp A/n)(\exp B/n)$. We wish to show that

$$S_n^n - T_n^n \to 0.$$

Proof of Lie's formula, I.

Notice that the constant and the linear terms in the power series expansions for S_n and T_n are the same, so

$$\|S_n - T_n\| \leq \frac{C}{n^2}$$

where $C = C(A, B)$.

Proof of Lie's formula, II.

We have the telescoping sum

$$S_n^n - T_n^n = \sum_{k=0}^{n-1} S_n^k (S_n - T_n) T^{n-1-k}$$

so

$$\|S_n^n - T_n^n\| \leq n \|S_n - T_n\| \left(\max(\|S_n\|, \|T_n\|)\right)^{n-1}.$$

Proof of Lie's formula, III.

$$\|S_n^n - T_n^n\| \leq n\|S_n - T_n\| \left(\max(\|S_n\|, \|T_n\|)\right)^{n-1}.$$ But

$$\|S_n\| \leq \exp\frac{1}{n}(\|A\| + \|B\|) \quad \text{and} \quad \|T_n\| \leq \exp\frac{1}{n}(\|A\| + \|B\|)$$

and

$$\left[\exp\frac{1}{n}(\|A\| + \|B\|)\right]^{n-1} = \exp\frac{n-1}{n}(\|A\|+\|B\|) \leq \exp(\|A\|+\|B\|)$$

so

$$\|S_n^n - T_n^n\| \leq \frac{C}{n}\exp(\|A\| + \|B\|). \quad \square$$

This same proof works if A and B are skew-adjoint operators such that $A+B$ is skew-adjoint on the intersection of their domains. For a proof see Reed-Simon vol. I pages 295-296. For applications this is too restrictive. So we give a more general formulation and proof following Chernoff. To state and prove Chernoff's result, I will need to develop some facts about dissipative operators which logically should have gone into our discussion of the Hille-Yosida theorem.

But before doing so, let me state the Trotter product formula (a generalization of Lie's formula) which will turn out to be a corollary of Chernoff's theorem.

The closure of an operator

I first want to retrench and do some material whose logical place was soon after we defined unbounded operators. I did not do this then because I felt that too many definitions would make for boring material.
Let A an operator from a Banach space **B** to a Banach space **C** and let $\mathcal{L}$ be the domain of A. We say that A is **closable** if it has a closed extension $\tilde{A}$. "Extension" means that the domain $\tilde{\mathcal{L}}$ of $\tilde{A}$ contains $\mathcal{L}$ and that $\tilde{A}$ coincides with A on $\mathcal{L}$.

Lemma

If A is closable then then there is a closed extension $\overline{A}$ of A whose domain $\mathcal{D}$ is is contained in the domain of every closed extension of A. This operator $\overline{A}$ is called the **closure** *of A*

Proof.

Define $\mathcal{D}$ to be the set of all $f \in \mathbf{B}$ such that there exists a sequence $f_n \in \mathcal{L}$ and a $g \in \mathbf{C}$ such that $f_n \to f$ and $Af_n \to g$. Since $\tilde{A}$ is a extension of A it follows that $f_n \in \tilde{\mathcal{L}}$, and since $\tilde{A}$ is closed, $f \in \mathcal{L}$ and $Af = g$. So g is uniquely determined by the above. We may define $\overline{A}$ to be defined on $\mathcal{D}$ as $Af = g$. Clearly $\overline{A}$ is an extension of A and any closed extension of A is an extension of $\overline{A}$. The graph of $\overline{A}$ is the closure of the graph of A, so $\overline{A}$ is closed. □

The definition of a core.

We need to deal with the possibility that A and B are operators with different domains of definition, and their sum $A+B$ has a third domain of definition. So we introduce the following

Definition

Let $\mathbf{F}$ be a Banach space and A 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 $\mathbf{F}$.

The product formula, assumptions.

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.

(Recall that a semigroup T_t is called a **contraction semi-group** if $\|T_t\| \leq 1$ for all t.)

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 $\overline{A+B}$.

The product formula, statement.

Theorem

[Trotter.] *Under the above hypotheses*

$$R_t y = \lim \left(P_{\frac{t}{n}} Q_{\frac{t}{n}} \right)^n y \quad \forall\, y \in \mathbf{F} \tag{2}$$

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

We will prove this theorem as a corollary of Chernoff's theorem to be stated and proved below, and this will involve us in a detour to prove some facts about dissipative operators. But let me first give a striking illustration of this theorem:

Feynman "path integrals", 1.

Consider the operator

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

given by

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

Recall that 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 $L_2(\mathbb{R}^3)$. The operator H_0 is called the "free Hamiltonian of non-relativistic quantum mechanics".

Feynman "path integrals"

The one parameter group generated by the free Hamiltonian.

The Fourier transform $\mathcal{F}$ is a unitary isomorphism of $L_2(\mathbb{R}^3)$ into $L_2(\mathbb{R}_3)$ and carries H_0 into multiplication by ξ^2 whose domain consists of those $\hat{\phi} \in L_2(\mathbb{R}_3)$ such that $\xi^2\hat{\phi}(\xi)$ belongs to $L_2(\mathbb{R}_3)$. 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 $L_2(\mathbb{R}^3)$ via the Fourier transform gives us a one parameter group of unitary transformations whose infinitesimal generator is $-iH_0$.

Feynman "path integrals"

The one parameter group generated by the free Hamiltonian in x-space.

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 integrals.

Feynman "path integrals", 2

Using the Trotter product formula

Let V be a function on $\mathbb{R}^3$. 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

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

Feynman "path integrals", 3

Using the Trotter product formula, 2

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:

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

where

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

Feynman "path integrals", 4

The action

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).

Feynman "path integrals", 5

Integrating the action over polygonal paths

Take $m = 2$ 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.

Feynman "path integrals" ?, 6

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 now turn to the task of proving the Trotter product formula (among other things).

The question to be studied

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 semi-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*. Recall that a semi-group T_t is called a contraction semi-group if $\|T_t\| \leq 1$ for all $t \geq 0$.

Some facts about contraction semi-groups

Recall that if an operator A on a Banach space satisfies

$$\|(I - n^{-1}A)^{-1}\| \leq 1 \qquad (3)$$

for all $n \in \mathbb{N}$ then the Hille-Yosida condition is satisfied, and the semi-group it generates is a contraction semi-group.
We will study another useful condition for recognizing a contraction semigroup in what follows.
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.

A fake scalar product.

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

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

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

$$\|\ell_u\| = \|u\| \text{ and } \ell_u(u) = \|u\|^2.$$

Choose one such ℓ_u for each $u \in \mathbf{F}$ and set

$$\langle\langle x, u \rangle\rangle := \ell_u(x).$$

Clearly all the conditions are satisfied.

Dissipative operators.

Of course this definition is highly unnatural, unless there is some reasonable way of choosing the ℓ_u 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

$$(Ax, x) \leq 0 \ \ \forall \ x \in D(A) \tag{4}$$

then A is dissipative relative to the scalar product.

Theorem

[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*

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

Proof.

Suppose first that $D(A)$ is dense and that $\mathrm{im}(I-A) = \mathbf{F}$. We wish to show that

$$\|(I - n^{-1}A)^{-1}\| \leq 1 \tag{3}$$

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

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

implying

$$s\|x\|^2 \leq \|y\|\|x\|. \tag{5}$$

We are assuming that $\mathrm{im}(I-A) = \mathbf{F}$. This together with (5) with $s = 1$ implies that $R(1, A)$ exists and

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

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

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.

$$y = sx - Ax.$$

$$\|(I - n^{-1}A)^{-1}\| \leq 1. \quad (3)$$

$$s\|x\|^2 \leq \|y\|\|x\|. \quad (5).$$

In particular, for s real and $|s-1| < 1$ the resolvent exists, and then (5) implies that $\|R(s,A)\| \leq s^{-1}$. Repeating the process we keep enlarging the resolvent set $\rho(A)$ until it includes the whole positive real axis and conclude from (5) that $\|R(s,A)\| \leq s^{-1}$ which implies (3). 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 $\mathrm{Dom}(A)$ is dense. Let $\langle\langle\cdot,\cdot\rangle\rangle$ be any semi-scalar product. Then

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

Dividing by t and letting $t \searrow 0$ we conclude that $\mathrm{Re}\ \langle\langle Ax, x\rangle\rangle \leq 0$ for all $x \in D(A)$, i.e. A is dissipative for $\langle\langle\cdot,\cdot\rangle\rangle$, completing the proof of the Lumer-Phillips theorem. □

A bound on the resolvent

Suppose that A satisfies the condition of the Lumer Phillips theorem, and T_t is the one parameter semi-group it generates. We know from the general theory of equibounded semigroups that the resolvent $R(z, A)$ exists for all z with Re $z > 0$ (and is given as the Laplace transform of T_t). Let Re $z > 0$ and $x = R(z, A)y$ so that $x \in D(A)$ and $y = zx - Ax$. Repeating an argument we just gave

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

implying

$$(\text{Re } z\|x\|^2 \leq \|y\|\|x\|, \tag{6}$$

i.e.

$$\|R(z, A)\| \leq \frac{1}{\text{Re } z}.$$

Resolvent convergence.

Proposition.

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

$$A_n x \to Ax. \tag{7}$$

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

$$R(z, A_n)y \to R(z, A)y. \tag{8}$$

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 $(zI - A)\mathbf{D}$ is dense in $\mathbf{F}$ since A is closed. So in proving (8) we may assume that $y = (zI - A)x$ with $x \in \mathbf{D}$.

Proof.

Then

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

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

Theorem

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, I.

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 $x \in \mathbf{D}$.

Proof, II.

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.

Proof, III.

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 $(\phi_n \star \rho)(t)$.

Proof, IV.

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.

Proof, V.

Hence using the Fourier inversion formula and, say, the dominated convergence theorem (for Banach space valued functions),

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

uniformly in t. $\square$

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 **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 [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

$$\lim_{n\to\infty} R(z_0, A_n) \to R(z_0) \quad \text{and} \quad \operatorname{im} R(z_0) \text{ is dense in } \mathbf{F}.$$

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 $t \geq 0$.

Theorem

[Chernoff.] *Let* $f : [0, \infty) \to$ *bounded operators on* $\mathbf{F}$ *be a continuous map with* $\|f(t)\| \leq 1 \quad \forall\ t$ *and* $\quad f(0) = I$. *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 $y \in \mathbf{F}$

$$\lim \left[f\left(\frac{t}{n}\right)\right]^n y = (\exp tA)y \tag{9}$$

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

Before proceeding to the proof of Chernoff's theorem, we need two facts:

Suppose that $B : F \to F$ is a bounded operator on a Banach space with $\|B\| \leq 1$. Then we know that $\exp tB$ exists and is given by the convergent power series. We also know that $\exp t(B - I)$ exists and has the expression

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

We have

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

so $\exp t(B - I)$ is a contraction semi-group.

We also will use the following inequality:

$$\|[\exp(n(B-I))-B^n]x\| \leq \sqrt{n}\|(B-I)x\| \quad \forall\, x \in \mathbf{F}, \text{ and } \forall\, n = 1, 2, 3 \ldots. \tag{10}$$

Proof.

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

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

We have proved that

$$\|[\exp(n(B-I)) - B^n]x\| \leq e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!} |k-n| \|(B-I)x\|.$$

So to prove

$$\|[\exp(n(B-I)) - B^n]x\| \leq \sqrt{n} \|(B-I)x\|. \quad (10).$$

it is enough establish the inequality

$$e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!} |k-n| \leq \sqrt{n}. \tag{11}$$

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

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

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

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

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. $\square$

Proof of Chernoff's theorem. For fixed $t > 0$ let

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

So $\frac{t}{n}C_n$ generates a contraction semi-group by what we just proved, 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 $x \in \mathbf{D}$.

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

Hence by the limit theorem we proved above,

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

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 (10) 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 (9) 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 it follows that (9) holds for all $y \in \mathbf{F}$. □

We recall the hypotheses and the statement of the theorem:

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

[Trotter.] *Under the above hypotheses*

$$R_t y = \lim \left(P_{\frac{t}{n}} Q_{\frac{t}{n}} \right)^n y \quad \forall\ y \in \mathbf{F} \tag{12}$$

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 (12). □

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

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 $y \in H$

$$\exp(it((\overline{S+T}))y = \lim_{n\to\infty}\left(\exp(\frac{t}{n}iS)(\exp\frac{t}{n}iT)\right)^n y \qquad (13)$$

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

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

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 $y \in \mathbf{H}$

$$\exp t\overline{[A, B]}y =$$

$$\lim_{n\to\infty} \left((\exp -\sqrt{\frac{t}{n}}A)(\exp -\sqrt{\frac{t}{n}}B)(\exp \sqrt{\frac{t}{n}}A)(\exp \sqrt{\frac{t}{n}}B) \right)^n y \tag{14}$$

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

Proof

The restriction of $[A, B]$ to **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 and yields

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

so (14) is a consequence of Chernoff's theorem with $s = \sqrt{t}$. □

Enter Mark Kac

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.

The integral of V over paths

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

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

The map

$$\omega \mapsto \int_0^t V(\omega(s))ds \tag{15}$$

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

$$\frac{t}{m}\sum_{j=1}^{m} V\left(\omega\left(\frac{jt}{m}\right)\right) \to \int_0^t V(\omega(s))ds \tag{16}$$

for each fixed ω.

The Feynman-Kac formula

Theorem

The Feynman-Kac formula. *Let V be a continuous real valued function of compact support on $\mathbb{R}^n$. Let $H = \Delta + V$ as an operator on $\mathfrak{H} = L^2(\mathbb{R}^n)$. Then H is self-adjoint and for every $f \in \mathfrak{H}$*

$$\left(e^{-tH}f\right)(x) = \int_{\Omega_x} f(\omega(t)) \exp\left(\int_0^t V(\omega(s))ds\right) d_x\omega \qquad (17)$$

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

The following proof is taken from Reed-Simon II page 280. It is due to Nelson.

We know from Kato-Relich that H is a self adjoint operator with the same domain as Δ. We may apply Trotter to conclude that

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

This convergence is in L_2. But by passing to a subsequence we may assume that the convergence is almost everywhere.

Now

$$\left(e^{(t/k)\Delta}e^{-(t/k)V}\right)^k f(x)$$
$$= \int \cdots \int f(x_k)e^{-(t/k)V(x_k)}\, p\big(\frac{t}{k}, x_k - x_{k-1}\big)\, e^{(t/k)V(x_{k-1})} \ldots$$
$$\cdot\ e^{-(t/k)V(x_1)}\, p\big(\frac{t}{k}, x - x_1\big)\, dx_1 \cdots dx_k.$$

By the very definition of Wiener measure this is

$$\left(e^{(t/k)\Delta}e^{-(t/k)V}\right)^k f(x) = E_x(\varphi_k),$$

$$\varphi_k(\omega) = f\big(\omega(t)\big)\ e^{-S_k(\omega)}, \quad S_k(\omega) = \frac{t}{k}\sum_{j=1}^{k} V\Big(\omega\big(\frac{jt}{k}\big)\Big).$$

The integrand (with respect to Wiener measure) converges on all continuous paths to the integrand on the right hand side of (17). We can then apply the dominated converges theorem to conclude the truth of the theorem.