Grant Kluber
DRP
December 7, 2020

Trotterization in QM Theory

1 Introduction

One of the major goals of quantum mechanics is finding solutions, called eigenfunctions, to the time-independent Schrodinger wave equation. For a given time-independent Hamiltonian operator $\hat{H}$ on a Hilbert space $\mathcal{H}$, the Schrodinger equation is given by

$$\hat{H} \left|\Psi\right\rangle = E \left|\Psi\right\rangle \tag{1}$$

where $\left|\Psi\right\rangle$ is an eigenfunction and E is an energy eigenvalue. Alone this equation only yields energy values and the stationary states of the physical system. To get the time-dependant eigenvector $\left|\Psi(t)\right\rangle$, one needs the unitary operator $U(t)$:

$$U(t) = e^{-i\hat{H}t/\hbar}, \tag{2}$$

$$\left|\Psi(t)\right\rangle = U(t) \left|\Psi\right\rangle = e^{-i\hat{H}t/\hbar} \left|\Psi\right\rangle . \tag{3}$$

In practice, it can be hard to compute this matrix exponential, so let's focus on a simple example. When the (finite-dimensional in this case) Hamiltonian $\hat{H}$ is given by

$$\hat{H} = \begin{pmatrix} E_1 & 0 & \cdots & 0 \\ 0 & E_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & E_n \end{pmatrix} . \tag{4}$$

Substituting this into Equation 1, the eigenfunctions are given by the standard basis

$$\left|\psi_1\right\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \left|\psi_2\right\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad \left|\psi_n\right\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{5}$$

with corresponding eigenvalues $E_1, \cdots, E_n$, respectively. From this, the arbitrary stationary state $\left|\Psi\right\rangle$ is a complex linear combination:

$$\left|\Psi\right\rangle = \sum_{i=1}^{n} c_i \left|\psi_i\right\rangle , \quad \sum_{i=1}^{n} \left|c_i\right|^2 = 1. \tag{6}$$

Now, to write a formula for how the state $\left|\Psi\right\rangle$ evolves over time, we can use the fact that the exponential of a diagonal matrix is just the diagonal matrix of element exponentials.

Concretely,

$$
\begin{aligned}
U(t) &= \exp\left[\begin{pmatrix} -iE_1t/\hbar & 0 & \cdots & 0 \\ 0 & -iE_2t/\hbar & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -iE_nt/\hbar \end{pmatrix}\right] \\
&= \begin{pmatrix} e^{-iE_1t/\hbar} & 0 & \cdots & 0 \\ 0 & e^{-iE_2t/\hbar} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-iE_nt/\hbar} \end{pmatrix}. \qquad (7)
\end{aligned}
$$

Therefore, the time-dependant state $|\Psi(t)\rangle$ is given by

$$
\begin{aligned}
|\Psi(t)\rangle &= U(t)\,|\Psi\rangle \\
&= \begin{pmatrix} e^{-iE_1t/\hbar} & 0 & \cdots & 0 \\ 0 & e^{-iE_2t/\hbar} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-iE_nt/\hbar} \end{pmatrix} \left(\sum_{i=1}^{n} c_i\,|\psi_i\rangle\right) \\
&= \sum_{i=1}^{n} c_i e^{-iE_it/\hbar}\,|\psi_i\rangle\,. \qquad (8)
\end{aligned}
$$

This calculation was simple because the Hamiltonian $\hat{H}$ was diagonal, but for the general case we need to use the series definition of exponentiation. Given a (potentially infinite dimensional) bounded self-adjoint operator $S : D(S) \to \mathcal{H}$, its exponential is defined by

$$e^S = \sum_{n=0}^{\infty} \frac{S^n}{n!}. \qquad (9)$$

If we have another self-adjoint operator $T : D(T) \to \mathcal{H}$ such that $[S, T] = 0$ (i.e., $ST = TS$) on the domain $D(S) \cap D(T)$, then

$$e^{S+T}\xi = e^S e^T \xi, \quad \xi \in D(S) \cap D(T). \qquad (10)$$

We can actually prove this in a relatively straight-forward manner from the definition of the operator exponential by applying the Binomial Theorem:

$$\begin{aligned}
e^{S+T} &= \sum_{n=0}^{\infty} \frac{(S+T)^n}{n!} \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{n!} \binom{n}{k} S^{n-k} T^k \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{n!} \frac{n!}{(n-k)!k!} S^{n-k} T^k \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{(n-k)!k!} S^{n-k} T^k.
\end{aligned}$$

With this result, notice that every possible product of S^m with T^n occurs for $m, n \in \mathbb{Z}^+ \cap \{0\}$. Thus, rewrite the sum as follows:

$$\begin{aligned}
\sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{(n-k)!k!} S^{n-k} T^k &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{m!n!} S^m T^n \\
&= \left(\sum_{m=0}^{\infty} \frac{1}{m!} S^m \right) \left(\sum_{n=0}^{\infty} \frac{1}{n!} T^n \right) \\
&= e^S e^T.
\end{aligned}$$

In the case when S and T do not commute, this argument fails because the Binomial Theorem no longer applies. It seems like there's no way to generalize this argument for the non-commutative case. For example, in the binomial expansion of $(S+T)^3$ with $[S, T] \neq 0$, $STS \neq S^2 T$, and so it's impossible to collect terms on the left and right sides of the overall sum. However, a surprising result called the Trotter Product Formula tells gives a limit where the formula does hold in a weak sense, even when S and T are non-commutative. This is what we explore for the remainder of this presentation.

Before the next section, it's important to bring up one prerequisite: unitary evolution groups. Put simply, a unitary evolution group is a group of unitary operators $G(t)$ for $t \in \mathbb{R}$ given by a homomorphism of $(\mathbb{R}, +)$ (i.e., for all $s, t \in \mathbb{R}$, $G(s+t) = G(s)G(t)$). An important notion is the infinitesimal generator T of a unitary evolution group, given pointwise by

$$T\xi = i \lim_{h \longrightarrow 0} \frac{G(h) - I}{h} \xi. \tag{11}$$

It turns out that T is self-adjoint (by Stone's Theorem), and it also turns out that *any* self-adjoint operator T corresponds to the (strongly/pointwise continuous) unitary evolution group $U_T(t) = e^{-itT}$.

2 The Trotter Product Formula

Claim. *Let S and T be bounded self-adjoint operators on $\mathcal{H}$ with domains $D(S)$ and $D(T)$, respectively. Then, for every $t \in \mathbb{R}$ and for all $\xi \in D(S+T) = D(S) \cap D(T) = \mathcal{D}$,*

$$\lim_{n \longrightarrow \infty} \left\| e^{-it(S+T)}\xi - \left(e^{-i\frac{t}{n}S} e^{-i\frac{t}{n}T}\right)^n \xi \right\| = 0. \tag{12}$$

In other words, the following strong (pointwise) operator limit holds:

$$\operatorname*{s-lim}_{n \to \infty} \left(e^{-i\frac{t}{n}S} e^{-i\frac{t}{n}T}\right)^n = e^{-it(S+T)}.$$

Proof. Our approach is somewhat technical and is adapted from [1]: we first construct an infinitesimal difference for Equation 12 and show it's zero, then use the nice properties of unitary evolution groups to extrapolate.

Let $h \in \mathbb{R}$ such that $h \neq 0$ and $\xi \in \mathcal{D}$; then, define $u_h(\xi)$ as

$$u_h(\xi) = \frac{1}{h}\left(e^{-ihS}e^{-ihT} - e^{-it(S+T)}\right).$$

Now, note this can be rewritten as

$$u_h(\xi) = \frac{e^{-ihS} - I}{h}\xi + e^{-ihS}\frac{e^{-ihT} - I}{h}\xi - \frac{e^{-ih(S+T)} - I}{h}\xi$$

where I is the identity operator on $\mathcal{H}$. Taking the limit $h \to 0$ for the first and third terms yields $S\xi$ and $-(S+T)\xi$, respectively, using the definition of the infinitesimal generator (Equation 11). For the second term, the Dominated Convergence Theorem applies, so that it yields $T\xi$ as $h \to 0$. Thus, $\lim_{h \longrightarrow 0} u_h(\xi) = 0$.

Since u_h in general is linear and bounded, and since the operator $S + T$ is closed (because every self-adjoint operator is closed), there is a technical argument involving the Uniform Boundedness Principle that eventually shows (through some topology) that

$$\lim_{h \longrightarrow 0} \max_{|s|<|t|} \|u_h(\xi_s)\| = 0 \tag{13}$$

where the symbol ξ_s is given by the unitary evolution group $\xi_s = e^{-is(S+T)}$. Because the unitary evolution group is strongly continuous, it follows that $J_{\xi,t} = \{\, \xi_s \mid |s| \leq |t| \,\}$ is compact and totally bounded in $J_{\xi,t}$ under a suitable norm, the graph norm of $S+T$. Totally bounded means that $J_{\xi,t}$ can be covered by a finite number of open balls, which naturally leads to an interpolation arguments: any $\xi_s \in J_{\xi,t}$ lies inside one of those balls, and from that we can upper bound the graph norm with a distance term that vanishes as $h \to 0$. Thus, since $\xi_s \in J_{\xi,t}$ is arbitrary, it also applies to the ξ_s that maximizes $\|u_h(\xi_s)\|$.

Now that we've shown this, we need to relate $u_h(\xi)$ with the difference in Equation 12. For bounded operators A, B and any $n \in \mathbb{Z}^+$, one can show that

$$A^n - B^n = \sum_{j=0}^{n-1} A^j (A - B) B^{n-1-j}.$$

Substituting $A = e^{-it(S+T)/n}$ (so that $A^n = e^{-it(S+T)}$) and $B = e^{-i\frac{t}{n}S} e^{-i\frac{t}{n}T}$,

$$\begin{aligned}
&\left(e^{-it(S+T)/n}\right)^n - \left(e^{-itS/n} e^{-itT/n}\right)^n \\
&= \sum_{j=0}^{n-1} \left(e^{-itS/n} e^{-itT/n}\right)^j \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) \left(e^{-it(S+T)/n}\right)^{n-1-j}
\end{aligned}$$

Thus, taking the norm and upper-bounding it applied to an arbitrary $\xi \in \mathcal{D}$ (using the triangle inequality and the fact that unitary operators on the left don't affect the norm),

$$\begin{aligned}
&\left\| \sum_{j=0}^{n-1} \left(e^{-itS/n} e^{-itT/n}\right)^j \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) \left(e^{-it(S+T)/n}\right)^{n-1-j} \xi \right\| \\
&\leq \sum_{j=0}^{n-1} \left\| \left(e^{-itS/n} e^{-itT/n}\right)^j \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) \left(e^{-it(S+T)/n}\right)^{n-1-j} \xi \right\| \\
&= \sum_{j=0}^{n-1} \left\| \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) \left(e^{-it(S+T)/n}\right)^{n-1-j} \xi \right\|
\end{aligned}$$

Now, the only dependence we have on j is in the last term of the product. If we define $s_j = t(n-1-j)/n$, then that term becomes $e^{-is_j(S+T)}$. Because $|s_j| < |t|$, $\{s_j\}$ is a subset of the interval $[-|t|, |t|]$. Therefore, we upper-bound the previous expression with a maximum over the entire interval:

$$\begin{aligned}
&\sum_{j=0}^{n-1} \left\| \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) e^{-is_j(S+T)} \xi \right\| \\
&\leq n \max_{|s|<|t|} \left\| \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) e^{-is(S+T)} \xi \right\|
\end{aligned}$$

If we let $h = |t|/n$, then the last expression becomes

$$\begin{aligned}
&\frac{|t|}{h} \max_{|s|<|t|} \left\| \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) e^{-is(S+T)} \xi \right\| \\
&= |t| \max_{|s|<|t|} \left\| \frac{1}{h} \left(e^{-it(S+T)/n} - e^{-itS/n} e^{-itT/n}\right) e^{-is(S+T)} \xi \right\| \\
&= |t| \max_{|s|<|t|} \left\| u_h(\xi_s) \right\|.
\end{aligned}$$

Therefore, as $n \to \infty$, $h \to 0$ so that

$$\left(e^{-i\frac{t}{n}S}e^{-i\frac{t}{n}T}\right)^n \xi \to e^{-it(S+T)}\xi, \quad \xi \in \mathcal{D}.$$

$\square$

3 Applications and Implications

Given a finite-dimensional $m \times m$ (complex) Hermitian matrix A, its matrix exponential can be computed exactly by diagonalizing. By the complex spectral theorem for finite-dimensional matrices, A can be decomposed as $A = UDU^\dagger$, where U is a unitary matrix, D is a diagonal matrix, and $(\cdot)^\dagger$ denotes the conjugate transpose. Using the fact that $A^n = (UDU^\dagger)^n = UD^nU^\dagger$ for any $n \in \mathbb{Z}^+$, it turns out that e^A is given by

$$e^A = Ue^DU^\dagger. \tag{14}$$

However, when m is very large diagonalizing A is computational hard! In fact, it is at least as hard as inverting A, since $A^{-1} = UD^{-1}U^\dagger$, and the best algorithms for matrix inversion have approximately $O(m^3)$ time complexity. Therefore, an approximation is needed to compute matrix exponentials in general instead of using Equation 14 directly, including the matrix exponential of sums.

We can efficiently compute the matrix exponential e^{A+B} by Trotterization for non-commuting matrices A and B. Using the series definition of the matrix exponential for sufficiently large N,

$$e^{A/N} \approx I + \frac{A}{N}, \quad e^{B/N} \approx I + \frac{B}{N}.$$

As a consequence of the Trotter Product Formula,

$$e^{A+B} \approx \left(e^{A/N}e^{B/N}\right)^N \approx \left[\left(I + \frac{A}{N}\right)\left(I + \frac{B}{N}\right)\right]^N = \left[I + \frac{A}{N} + \frac{B}{N} + \frac{AB}{N^2}\right]^N \approx \left[I + \frac{A+B}{N}\right]^N.$$

Note the similarities between this equation and the limit definition for e^x for $x \in \mathbb{R}$:

$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n.$$

Care must be taken to choose an N that is not too large (causing numerical instability from $I + (A+B)/N$ being too close to I); additionally, N must also be sufficiently large, otherwise the approximation won't be very good. Therefore, some trial and error is needed to determine an appropriate value of N; ideally, N should be a power of 2 (i.e., $N = 2^p$ for some $p \in \mathbb{Z}^+$) so that repeated squaring can be applied [2].

Trotterization is also useful for analyzing operators that individually have "nice" operator exponentials, but when summed together have a complicated operator exponential. A good example of this is the unitary evolution group $e^{it(\hat{p}^2/2m+\hat{x})}$ corresponding to the Hamiltonian operator $\hat{H} = \frac{\hat{p}^2}{2m} + \hat{x}$ (i.e., $V(x) = x$). It's difficult to compute this exponential directly, but the exponential of $it\hat{p}^2/2m$ individually corresponds to a spatial translation and the exponential of $it\hat{x}$ individually corresponds to a momentum translation. Thus, the exponential of the sum can be reduced to terms much easier to analyze.

References

[1] de Oliveira César R. *Intermediate spectral theory and quantum dynamics*. Birkhäuser, 2009.

[2] *Repeated squaring*. URL: https://algorithmist.com/wiki/Repeated_squaring.