

Notes on Entropy Change Over Time

Duncan MacIntyre

February 29, 2024

Contents

1	Time-dependent perturbation theory	2
2	Time evolution of the density operator	3
3	Eigenvalue corrections from perturbation theory	4
4	The change in Von Neumann entropy	6
4(a)	Case where $k = 1$ (first order correction leads)	8
4(b)	Case where $k = 2$ (second order correction leads)	8
4(c)	Case of separable initial state	9
4(d)	Total entropy of a closed system	11
5	Example: Interaction between Qubits	11
A	The first-order approximation for pure states in scattering	13
B	Reduced density operator	14

1 Time-dependent perturbation theory

Suppose we have a Hamiltonian of the form

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}(t)$$

where \hat{H}_0 is a well-understood Hamiltonian that does not depend on time and $\lambda \hat{V}(t)$ is “small”. For example, \hat{H}_0 might be the free particle Hamiltonian $\hat{H}_0 = \frac{m}{2} \nabla^2$. Our equation of motion is the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi, t\rangle = (\hat{H}_0 + \lambda \hat{V}(t)) |\Psi, t\rangle$$

where $|\Psi, t\rangle$ is the usual Schrödinger-picture state at time t .

Let $\hat{U}(t_0, t) = e^{-i\hat{H}_0(t-t_0)/\hbar}$. Then $\hat{U}(t_0, t)$ is the operator that evolves a state from time t_0 to time t according to \hat{H}_0 . We define the interaction-picture state to be

$$|\Psi_I, t\rangle = \hat{U}(t_0, t)^\dagger |\Psi, t\rangle$$

so $|\Psi, t\rangle = \hat{U}(t_0, t) |\Psi_I, t\rangle$. Plugging this in to the Schrödinger equation,

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \hat{U}(t_0, t) |\Psi_I, t\rangle &= (\hat{H}_0 + \lambda \hat{V}(t)) \hat{U}(t_0, t) |\Psi_I, t\rangle \\ i\hbar \left[\frac{\partial}{\partial t} e^{-i\hat{H}_0(t-t_0)/\hbar} \right] |\Psi_I, t\rangle + i\hbar e^{-i\hat{H}_0(t-t_0)/\hbar} \frac{\partial}{\partial t} |\Psi_I, t\rangle &= \hat{H}_0 e^{-i\hat{H}_0(t-t_0)/\hbar} |\Psi_I, t\rangle + \lambda \hat{V}(t) e^{-i\hat{H}_0(t-t_0)/\hbar} |\Psi_I, t\rangle \\ \cancel{-i^2 \hat{H}_0 e^{i\hat{H}_0(t-t_0)/\hbar} |\Psi_I, t\rangle} + i\hbar e^{-i\hat{H}_0(t-t_0)/\hbar} \frac{\partial}{\partial t} |\Psi_I, t\rangle &= \cancel{\hat{H}_0 e^{-i\hat{H}_0(t-t_0)/\hbar} |\Psi_I, t\rangle} + \lambda \hat{V}(t) e^{-i\hat{H}_0(t-t_0)/\hbar} |\Psi_I, t\rangle \\ i\hbar \frac{\partial}{\partial t} |\Psi_I, t\rangle &= e^{i\hat{H}_0(t-t_0)/\hbar} \lambda \hat{V}(t) e^{-i\hat{H}_0(t-t_0)/\hbar} |\Psi_I, t\rangle \\ i\hbar \frac{\partial}{\partial t} |\Psi_I, t\rangle &= \lambda \hat{H}_I(t) |\Psi_I, t\rangle \end{aligned}$$

where we define the interaction Hamiltonian to be

$$\hat{H}_I(t) = \hat{U}(t_0, t)^\dagger \hat{V}(t) \hat{U}(t_0, t) = e^{i\hat{H}_0(t-t_0)/\hbar} \hat{V}(t) e^{-i\hat{H}_0(t-t_0)/\hbar}.$$

This sets up the interaction picture. We have rephrased our problem so that we can continue with quantum mechanics normally without having to worry about the time evolution due to \hat{H}_0 .

We now integrate both sides of our expression.

$$\begin{aligned} \int_{t_0}^t \frac{\partial}{\partial t'} |\Psi_I, t'\rangle &= -\frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') |\Psi_I, t'\rangle dt' \\ |\Psi_I, t\rangle - |\Psi_I, t_0\rangle &= -\frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') |\Psi_I, t'\rangle dt' \\ |\Psi_I, t\rangle &= |\Psi_I, t_0\rangle - \frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') |\Psi_I, t'\rangle dt' \end{aligned} \tag{1}$$

This is called the integral form of the Schrödinger equation.

We can now iteratively calculate perturbative approximations where we assume $\hat{H}_I(t)$ is small. The zero order approximation is simply

$$|\Psi_I, t\rangle = |\Psi_I, t_0\rangle + \mathcal{O}(\lambda).$$

Plugging this in for the state inside the integral in (1), we get the first order approximation

$$|\Psi_I, t\rangle = |\Psi_I, t_0\rangle - \frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') |\Psi_I, t_0\rangle dt' + \mathcal{O}(\lambda^2). \quad (2)$$

Plugging this in to (1) again we get the second order approximation

$$\begin{aligned} |\Psi_I, t\rangle &= |\Psi_I, t_0\rangle - \frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') \left[|\Psi_I, t_0\rangle - \frac{i}{\hbar} \lambda \int_{t_0}^{t'} \hat{H}_I(t'') |\Psi_I, t_0\rangle dt'' \right] dt' + \mathcal{O}(\lambda^3) \\ &= |\Psi_I, t_0\rangle - \frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') |\Psi_I, t_0\rangle dt' - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \hat{H}_I(t') \hat{H}_I(t'') |\Psi_I, t_0\rangle dt'' dt' + \mathcal{O}(\lambda^3). \end{aligned} \quad (3)$$

In general, we can keep going to achieve higher order approximations. For us, however, the second order approximation (3) is enough.

2 Time evolution of the density operator

We now consider mixed states. We will need to describe the system by the density operator $\hat{\rho}(t)$. Let's derive the time evolution of $\hat{\rho}(t)$ in second order perturbation theory based on (3).

Suppose at time t_0 we have a statistical ensemble of interaction-picture states $|\psi_{I_n}, t_0\rangle$ each with probability P_n . Then

$$\hat{\rho}(t_0) = \sum_n P_n |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0|.$$

At time t states will have evolved according to (3), so

$$\hat{\rho}(t) = \sum_n P_n |\Psi_{I_n}, t\rangle \langle \Psi_{I_n}, t|.$$

From (3) we have

$$\begin{aligned} |\Psi_{I_n}, t\rangle &= |\Psi_{I_n}, t_0\rangle - \frac{i}{\hbar} \lambda \int_{t_0}^t \hat{H}_I(t') |\Psi_{I_n}, t_0\rangle dt' - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \hat{H}_I(t') \hat{H}_I(t'') |\Psi_{I_n}, t_0\rangle dt'' dt' + \mathcal{O}(\lambda^3) \\ \langle \Psi_{I_n}, t| &= \langle \Psi_{I_n}, t_0| - \frac{i}{\hbar} \lambda \int_{t_0}^t \langle \Psi_{I_n}, t_0| \hat{H}_I(t') dt' - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \langle \Psi_{I_n}, t_0| \hat{H}_I(t'') \hat{H}_I(t') dt'' dt' + \mathcal{O}(\lambda^3) \end{aligned}$$

so

$$\begin{aligned} |\Psi_{I_n}, t\rangle \langle \Psi_{I_n}, t| &= |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| - \frac{i}{\hbar} \lambda \int_{t_0}^t \left(\hat{H}_I(t') |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| - |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| \hat{H}_I(t') \right) dt' \\ &\quad - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \left(\hat{H}_I(t') \hat{H}_I(t'') |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| + |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| \hat{H}_I(t'') \hat{H}_I(t') \right) dt'' dt' \\ &\quad + \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^t \hat{H}_I(t') |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| \hat{H}_I(t'') dt'' dt' + \mathcal{O}(\lambda^3) \\ &= |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| - \frac{i}{\hbar} \lambda \int_{t_0}^t \left[\hat{H}_I(t'), |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| \right] dt' \\ &\quad - \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \left\{ \hat{H}_I(t') \hat{H}_I(t''), |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| \right\} dt'' dt' \\ &\quad + \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^t \hat{H}_I(t') |\Psi_{I_n}, t_0\rangle \langle \Psi_{I_n}, t_0| \hat{H}_I(t'') dt'' dt' + \mathcal{O}(\lambda^3). \end{aligned}$$

Then

$$\begin{aligned}
\hat{\rho}(t) &= \hat{\rho}(t_0) - \frac{i}{\hbar} \lambda \int_{t_0}^t [\hat{H}(t'), \hat{\rho}(t_0)] dt' \\
&- \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \left\{ \hat{H}_I(t') \hat{H}_I(t''), \hat{\rho}(t_0) \right\} dt'' dt' \\
&+ \frac{\lambda^2}{\hbar^2} \int_{t_0}^t \int_{t_0}^t \hat{H}_I(t') \hat{\rho}(t_0) \hat{H}_I(t'') dt'' dt' + \mathcal{O}(\lambda^3).
\end{aligned} \tag{4}$$

3 Eigenvalue corrections from perturbation theory

The goal of this section is to compute the second-order perturbative corrections to the eigenvalues of an operator. We use the approach known as time-independent perturbation theory.

Suppose $\hat{\rho}$ is an operator with eigenstates $|\Psi_n\rangle$ and eigenvalues σ_n . Suppose we have the asymptotic expansions

$$\begin{aligned}
\hat{\rho} &= \hat{\rho}_0 + \lambda \hat{\rho}_1 + \lambda^2 \hat{\rho}_2 + \mathcal{O}(\lambda^3) \\
|\Psi_n\rangle &= |\Psi_n^{(0)}\rangle + \lambda |\Psi_n^{(1)}\rangle + \lambda^2 |\Psi_n^{(2)}\rangle + \mathcal{O}(\lambda^3) \\
\sigma_n &= \sigma_n^{(0)} + \lambda \sigma_n^{(1)} + \lambda^2 \sigma_n^{(2)} + \mathcal{O}(\lambda^3)
\end{aligned}$$

where $\{|\Psi_n^{(0)}\rangle\}$ is an orthonormal basis for the Hilbert space.

We have $\hat{\rho}|\Psi_n\rangle = \sigma_n|\Psi_n\rangle$. In the zeroth order of λ this gives $\hat{\rho}_0|\Psi_n^{(0)}\rangle = \sigma_n^{(0)}|\Psi_n^{(0)}\rangle$, that is, $|\Psi_n^{(0)}\rangle$ is an eigenstate of $\hat{\rho}_0$ with eigenvalue $\sigma_n^{(0)}$.

Now, we have

$$\begin{aligned}
\sigma_n |\Psi_n\rangle &= \hat{\rho} |\Psi_n\rangle \\
\left(\sigma_n^{(0)} - \sigma_m^{(0)} + \lambda \sigma_n^{(1)} + \lambda^2 \sigma_n^{(2)} + \mathcal{O}(\lambda^3) \right) \langle \Psi_m^{(0)} | \Psi_n \rangle &= \langle \Psi_m^{(0)} | \lambda \hat{\rho}_1 + \lambda^2 \hat{\rho}_2 + \mathcal{O}(\lambda^3) | \Psi_n \rangle \\
\langle \Psi_m^{(0)} | \Psi_n \rangle &= \frac{\langle \Psi_m^{(0)} | \lambda \hat{\rho}_1 + \lambda^2 \hat{\rho}_2 + \mathcal{O}(\lambda^3) | \Psi_n \rangle}{\sigma_n^{(0)} - \sigma_m^{(0)} + \lambda \sigma_n^{(1)} + \lambda^2 \sigma_n^{(2)} + \mathcal{O}(\lambda^3)}.
\end{aligned}$$

We will next proceed to get rid of the denominator by using the Taylor expansion $\frac{1}{1+x} = 1 - x + \mathcal{O}(x^2)$.

Let D_n be the indices degenerate with n . That is, let $D_n = \{m : \sigma_m^{(0)} = \sigma_n^{(0)}\}$. Consider the case

where $m \in D_n$. Let δ be a small real number. In the limit as $\delta \rightarrow 0$,

$$\begin{aligned}
\langle \Psi_m^{(0)} | \Psi_n \rangle &= \frac{\langle \Psi_m^{(0)} | \lambda \hat{\rho}_1 + \lambda^2 \hat{\rho}_2 + \mathcal{O}(\lambda^3) | \Psi_n \rangle}{\lambda \sigma_n^{(1)} + \lambda^2 \sigma_n^{(2)} + \mathcal{O}(\lambda^3) + \lambda \delta} \\
&= \frac{\langle \Psi_m^{(0)} | \lambda \hat{\rho}_1 + \lambda^2 \hat{\rho}_2 + \mathcal{O}(\lambda^3) | \Psi_n \rangle}{\lambda \left(\sigma_n^{(1)} + \delta \right) \left(1 + \lambda \frac{\sigma_n^{(2)}}{\sigma_n^{(1)} + \delta} + \mathcal{O}(\lambda^2) \right)} \\
&= \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 + \lambda \hat{\rho}_2 + \mathcal{O}(\lambda^2) | \Psi_n \rangle}{\sigma_n^{(1)} + \delta} \left(1 - \lambda \frac{\sigma_n^{(2)}}{\sigma_n^{(1)} + \delta} + \mathcal{O}(\lambda^2) \right) \\
&= \frac{\langle \Psi_m^{(0)} | (\hat{\rho}_1 + \lambda \hat{\rho}_2 + \mathcal{O}(\lambda^2)) \left(|\Psi_n^{(0)}\rangle + \lambda |\Psi_n^{(1)}\rangle + \mathcal{O}(\lambda^2) \right)}{\sigma_n^{(1)} + \delta} \left(1 - \lambda \frac{\sigma_n^{(2)}}{\sigma_n^{(1)} + \delta} + \mathcal{O}(\lambda^2) \right) \\
&= \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta} + \lambda \left[\frac{\langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta} + \frac{\langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(1)} \rangle}{\sigma_n^{(1)} + \delta} - \frac{\sigma_n^{(2)} \langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle}{\left(\sigma_n^{(1)} + \delta \right)^2} \right] + \mathcal{O}(\lambda^2)
\end{aligned}$$

(The δ ensured that we did not divide by zero if $\sigma_n^{(1)} = 0$.) Now consider the case where $m \notin D_n$. Then

$$\begin{aligned}
\langle \Psi_m^{(0)} | \Psi_n \rangle &= \frac{\langle \Psi_m^{(0)} | \lambda \hat{\rho}_1 + \mathcal{O}(\lambda^2) | \Psi_n \rangle}{\left(\sigma_n^{(0)} - \sigma_m^{(0)} \right) (1 + \mathcal{O}(\lambda))} \\
&= \frac{\langle \Psi_m^{(0)} | \lambda \hat{\rho}_1 + \mathcal{O}(\lambda^2) | \Psi_n \rangle}{\sigma_n^{(0)} - \sigma_m^{(0)}} (1 + \mathcal{O}(\lambda)) \\
&= \lambda \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n \rangle}{\sigma_n^{(0)} - \sigma_m^{(0)}} + \mathcal{O}(\lambda^2).
\end{aligned}$$

Putting this all together,

$$\begin{aligned}
|\Psi_n\rangle &= \sum_{m \in D_n} |\Psi_m^{(0)}\rangle \langle \Psi_m^{(0)} | \Psi_n \rangle + \sum_{m \notin D_n} |\Psi_m^{(0)}\rangle \langle \Psi_m^{(0)} | \Psi_n \rangle \\
&= \sum_{m \in D_n} |\Psi_m^{(0)}\rangle \lim_{\delta \rightarrow 0} \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta} \\
&\quad + \lambda \left(\sum_{m \in D_n} |\Psi_m^{(0)}\rangle \lim_{\delta \rightarrow 0} \left[\frac{\langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta} + \frac{\langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(1)} \rangle}{\sigma_n^{(1)} + \delta} - \frac{\sigma_n^{(2)} \langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle}{\left(\sigma_n^{(1)} + \delta \right)^2} \right] \right. \\
&\quad \left. + \sum_{m \notin D_n} |\Psi_m^{(0)}\rangle \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n \rangle}{\sigma_n^{(0)} - \sigma_m^{(0)}} \right) \\
&\quad + \mathcal{O}(\lambda^2)
\end{aligned} \tag{5}$$

In the zeroth order of λ , equation (5) gives

$$|\Psi_n^{(0)}\rangle = \sum_{m \in D_n} |\Psi_m^{(0)}\rangle \lim_{\delta \rightarrow 0} \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta}.$$

Multiplying by $\langle \Psi_n^{(0)} |$ we get

$$\sigma_n^{(1)} = \langle \Psi_n^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle. \quad (6)$$

If we instead multiply by $\langle \Psi_k^{(0)} |$ where $k \in D_n$ but $k \neq n$ we get

$$0 = \langle \Psi_k^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle, \quad k \in D_n, \quad m \neq n. \quad (7)$$

In other words, the matrix given by $\langle \Psi_k^{(0)} | \hat{\rho}_1 | \Psi_j^{(0)} \rangle$ is diagonal on $k, j \in D_n$, for any n .

In the first order of λ , equation (5) says

$$| \Psi_n^{(1)} \rangle = \sum_{m \in D_n} | \Psi_m^{(0)} \rangle \lim_{\delta \rightarrow 0} \left[\frac{\langle \Psi_m^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta} + \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n^{(1)} \rangle}{\sigma_n^{(1)} + \delta} - \frac{\sigma_n^{(2)} \langle \Psi_m^{(0)} | \hat{\rho}_0 | \Psi_n^{(0)} \rangle}{(\sigma_n^{(1)} + \delta)^2} \right] + \sum_{m \notin D_n} | \Psi_m^{(0)} \rangle \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n \rangle}{\sigma_n^{(0)} - \sigma_m^{(0)}}.$$

Multiplying by $\langle \Psi_k^{(0)} |$ where $k \notin D_n$ gives

$$\langle \Psi_k^{(0)} | \Psi_n^{(1)} \rangle = \frac{\langle \Psi_k^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle}{\sigma_n^{(0)} - \sigma_k^{(0)}}.$$

If we instead multiply by $\langle \Psi_n^{(0)} |$ we get

$$0 = \langle \Psi_n^{(0)} | \Psi_n^{(1)} \rangle = \lim_{\delta \rightarrow 0} \left[\frac{\langle \Psi_n^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle}{\sigma_n^{(1)} + \delta} + \frac{\langle \Psi_n^{(0)} | \hat{\rho}_1 | \Psi_n^{(1)} \rangle}{\sigma_n^{(1)} + \delta} - \frac{\sigma_n^{(2)} (\sigma_n^{(1)} + \delta)}{(\sigma_n^{(1)} + \delta)^2} \right]$$

so

$$\begin{aligned} \sigma_n^{(2)} &= \langle \Psi_n^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle + \langle \Psi_n^{(0)} | \hat{\rho}_1 | \Psi_n^{(1)} \rangle - \langle \Psi_n^{(0)} | \Psi_n^{(1)} \rangle \sigma_n^{(1)} \\ &= \langle \Psi_n^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle + \sum_{m \in D_n} \langle \Psi_n^{(0)} | \hat{\rho}_1 | \Psi_m^{(0)} \rangle \langle \Psi_m^{(0)} | \Psi_n^{(1)} \rangle + \sum_{m \notin D_n} \langle \Psi_n^{(0)} | \hat{\rho}_1 | \Psi_m^{(0)} \rangle \langle \Psi_m^{(0)} | \Psi_n^{(1)} \rangle - \langle \Psi_n^{(0)} | \Psi_n^{(1)} \rangle \sigma_n^{(1)} \\ &= \langle \Psi_n^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle + \sigma_n^{(1)} \langle \Psi_n^{(0)} | \Psi_n^{(1)} \rangle + \sum_{m \notin D_n} \langle \Psi_n^{(0)} | \hat{\rho}_1 | \Psi_m^{(0)} \rangle \frac{\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle}{\sigma_n^{(0)} - \sigma_m^{(0)}} - \langle \Psi_n^{(0)} | \Psi_n^{(1)} \rangle \sigma_n^{(1)} \end{aligned}$$

hence

$$\sigma_n^{(2)} = \langle \Psi_n^{(0)} | \hat{\rho}_2 | \Psi_n^{(0)} \rangle + \sum_{m \notin D_n} \frac{|\langle \Psi_m^{(0)} | \hat{\rho}_1 | \Psi_n^{(0)} \rangle|^2}{\sigma_n^{(0)} - \sigma_m^{(0)}}. \quad (8)$$

4 The change in Von Neumann entropy

Suppose we have a Hilbert space of the form $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. The \mathcal{H}_A -reduced density operator of a density operator $\hat{\rho}$ is defined to be

$$\bar{\rho} = \sum_m (\cdot \otimes \langle \beta_m |) \hat{\rho} (\cdot \otimes | \beta_m \rangle)$$

where $\{|\beta_m\rangle\}$ is any orthonormal basis for \mathcal{H}_B . (One can verify that all choices of basis give the same $\bar{\rho}$.) We can consider $\bar{\rho}$ to act on the Hilbert space \mathcal{H}_A .

Define the \mathcal{H}_A -subsystem entropy to be $S = -\text{Tr} \bar{\rho} \log \bar{\rho}$. If $\bar{\rho}$ is diagonalized like

$$\bar{\rho} = \sum_n \bar{\sigma}_n |\alpha_n\rangle \langle \alpha_n|, \quad (9)$$

where $\{|\alpha_n\rangle\}$ is an orthonormal basis of \mathcal{H}_A , then

$$S = -\sum_n \bar{\sigma}_n \log \bar{\sigma}_n.$$

Now, looking at equation (4), we see that at some fixed time t we can form asymptotic expansions

$$\begin{aligned} \bar{\rho}(t) &= \bar{\rho}_0 + \lambda \bar{\rho}_1 + \lambda^2 \bar{\rho}_2 + \mathcal{O}(\lambda^3) \\ |\alpha_n\rangle &= |\alpha_n^{(0)}\rangle + \lambda |\alpha_n^{(1)}\rangle + \lambda^2 |\alpha_n^{(2)}\rangle + \mathcal{O}(\lambda^3) \\ \bar{\sigma}_n &= \bar{\sigma}_n^{(0)} + \lambda \bar{\sigma}_n^{(1)} + \lambda^2 \bar{\sigma}_n^{(2)} + \mathcal{O}(\lambda^3) \end{aligned}$$

by taking

$$\begin{aligned} \hat{\rho}_0 &= \hat{\rho}(t_0) \\ \bar{\rho}_0 &= \bar{\rho}(t_0) = \sum_m (\cdot \otimes \langle \beta_m |) \hat{\rho}_0 (\cdot \otimes | \beta_m \rangle) \\ \bar{\rho}_1 &= -\frac{i}{\hbar} \int_{t_0}^t \sum_m (\cdot \otimes \langle \beta_m |) [\hat{H}(t'), \hat{\rho}_0] (\cdot \otimes | \beta_m \rangle) dt' \end{aligned} \quad (10)$$

$$\begin{aligned} \bar{\rho}_2 &= -\frac{1}{\hbar^2} \int_{t_0}^t \int_{t_0}^{t'} \sum_m (\cdot \otimes \langle \beta_m |) \left\{ \hat{H}_I(t') \hat{H}_I(t''), \hat{\rho}_0 \right\} (\cdot \otimes | \beta_m \rangle) dt'' dt' \\ &\quad + \frac{1}{\hbar^2} \int_{t_0}^t \int_{t_0}^t \sum_m (\cdot \otimes \langle \beta_m |) \hat{H}_I(t') \hat{\rho}_0 \hat{H}_I(t'') (\cdot \otimes | \beta_m \rangle) dt'' dt'. \end{aligned} \quad (11)$$

From (9) we see that

$$\bar{\rho}_0 = \sum_n \bar{\sigma}_n^{(0)} |\alpha_n^{(0)}\rangle \langle \alpha_n^{(0)}|. \quad (12)$$

Therefore we can take $\{|\alpha_n^{(0)}\rangle\}$ to be an orthonormal basis for \mathcal{H}_A and we must have $\sum_n \bar{\sigma}_n^{(0)} = 1$.

We can now compute the change in entropy. Let k be the first non-zero order for which some eigenvalue correction $\bar{\sigma}_n^{(k)}$ does not vanish. For example, if some $\bar{\sigma}_n^{(1)} \neq 0$, then $k = 1$. If all $\bar{\sigma}_n^{(1)} = 0$ but some $\bar{\sigma}_n^{(2)} \neq 0$, then $k = 2$. Now let

$$I = \{n : \bar{\sigma}_n^{(0)} \neq 0\} \quad J = \{n : \bar{\sigma}_n^{(0)} = 0 \text{ and } \bar{\sigma}_n^{(k)} \neq 0\} \quad K = \{n : \bar{\sigma}_n^{(0)} = 0 \text{ and } \bar{\sigma}_n^{(k)} = 0\}.$$

Then

$$\begin{aligned}
S &= - \sum_{n \in I} \overline{\sigma_n} \log \overline{\sigma_n} - \sum_{n \in J} \overline{\sigma_n} \log \overline{\sigma_n} - \sum_{n \in K} \overline{\sigma_n} \log \overline{\sigma_n} \\
&= - \sum_{n \in I} \overline{\sigma_n} \log \left(\overline{\sigma_n^{(0)}} (1 + \mathcal{O}(\lambda^k)) \right) - \sum_{n \in J} \overline{\sigma_n} \log \left(\lambda^k \overline{\sigma_n^{(k)}} (1 + \mathcal{O}(\lambda)) \right) + \mathcal{O}(\lambda^k) \\
&= - \sum_{n \in I} \left(\overline{\sigma_n^{(0)}} + \mathcal{O}(\lambda^k) \right) \left(\log \overline{\sigma_n^{(0)}} + \mathcal{O}(\lambda^k) \right) - \sum_{n \in J} \overline{\sigma_n} \left(\log \lambda^k + \log \overline{\sigma_n^{(k)}} + \mathcal{O}(\lambda) \right) + \mathcal{O}(\lambda^k) \\
&= - \sum_{n \in I} \overline{\sigma_n^{(0)}} \log \overline{\sigma_n^{(0)}} - \log \lambda^k \sum_{n \in J} \overline{\sigma_n} + \mathcal{O}(\lambda^k) \\
&= - \sum_{n \in I} \overline{\sigma_n^{(0)}} \log \overline{\sigma_n^{(0)}} - \log \lambda^k \sum_{n \in J \cup K} \overline{\sigma_n} + \mathcal{O}(\lambda^k) \\
&= - \sum_{n \in I} \overline{\sigma_n^{(0)}} \log \overline{\sigma_n^{(0)}} - \log \lambda^k \left(1 - \sum_{n \in I} \overline{\sigma_n} \right) + \mathcal{O}(\lambda^k) \\
&= - \sum_{n \in I} \overline{\sigma_n^{(0)}} \log \overline{\sigma_n^{(0)}} - \log \lambda^k \left(1 - \sum_{n \in I} \overline{\sigma_n^{(0)}} - \sum_{n \in I} \lambda^k \overline{\sigma_n^{(k)}} \right) + \mathcal{O}(\lambda^k) \\
&= - \sum_{n \in I} \overline{\sigma_n^{(0)}} \log \overline{\sigma_n^{(0)}} + \lambda^k \log \frac{1}{\lambda^k} \sum_{n \in I} \overline{\sigma_n^{(k)}} + \mathcal{O}(\lambda^k)
\end{aligned}$$

Here we used that $\sum_n \overline{\sigma_n^{(0)}} = 1$. [The above should be carefully checked. In particular, what happens if J is empty? Is it true that we can replace J with $J \cup K$ at that mid step?] Now, the first term, $-\sum_{n \in I} \overline{\sigma_n^{(0)}} \log \overline{\sigma_n^{(0)}}$, would be the entropy if $\lambda = 0$, that is, with no perturbation. Thus the change in entropy due to the perturbation is

$$\boxed{\Delta S = \left(\lambda^k \log \frac{1}{\lambda^k} \right) \sum_{n \in I} \overline{\sigma_n^{(k)}} + \mathcal{O}(\lambda^k).} \quad (13)$$

4(a) Case where $k = 1$ (first order correction leads)

We can compute $\overline{\sigma_n^{(1)}}$ by combining equations (6) and (10). We get

$$\overline{\sigma_n^{(1)}} = -\frac{i}{\hbar} \sum_m \int_{t_0}^t \left(\langle \alpha_n^{(0)} | \otimes \langle \beta_m | \right) \left[\hat{H}(t'), \hat{\rho}_0 \right] \left(| \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle \right) dt'. \quad (14)$$

If this is nonzero for some n , (13) gives

$$\Delta S = \left(-\frac{i}{\hbar} \lambda \log \frac{1}{\lambda} \right) \sum_{n \in I} \sum_m \int_{t_0}^t \left(\langle \alpha_n^{(0)} | \otimes \langle \beta_m | \right) \left[\hat{H}(t'), \hat{\rho}_0 \right] \left(| \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle \right) dt' + \mathcal{O}(\lambda). \quad (15)$$

4(b) Case where $k = 2$ (second order correction leads)

On the other hand, (14) could be zero for all n . Then we turn to the second order correction. Let us assume that $\overline{\sigma_n^{(2)}} \neq 0$ for some n .

Substituting (8) into (13),

$$\Delta S = -\lambda^2 \log \frac{1}{\lambda^2} \left(\sum_{n \in I} \langle \alpha_n^{(0)} | \bar{\rho}_2 | \alpha_n^{(0)} \rangle + \sum_{n \in I} \sum_{m \notin D_n} \frac{|\langle \alpha_m^{(0)} | \bar{\rho}_1 | \alpha_n^{(0)} \rangle|^2}{\overline{\sigma_n^{(0)}} - \overline{\sigma_m^{(0)}}} \right) + \mathcal{O}(\lambda^2).$$

Now, in the second term with the double sum, for every term $\frac{|\langle \alpha_m^{(0)} | \bar{\rho}_1 | \alpha_n^{(0)} \rangle|^2}{\overline{\sigma_n^{(0)}} - \overline{\sigma_m^{(0)}}}$ there is an equal but opposite term $\frac{|\langle \alpha_n^{(0)} | \bar{\rho}_1 | \alpha_m^{(0)} \rangle|^2}{\overline{\sigma_m^{(0)}} - \overline{\sigma_n^{(0)}}}$. Thus this double sum evaluates to zero. We end up with

$$\boxed{\Delta S = -\lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \langle \alpha_n^{(0)} | \bar{\rho}_2 | \alpha_n^{(0)} \rangle + \mathcal{O}(\lambda^2)} \quad (16)$$

where $\bar{\rho}_2$ is given by (11).

4(c) Case of separable initial state

We will now consider the special case where, at the time t_0 , the density operator takes the form

$$\hat{\rho}_0 = \hat{\rho}(t_0) = \sum_{n,m} \sigma_{n,m}^{(0)} \left(|\alpha_n^{(0)}\rangle \otimes |\beta_m\rangle \right) \left(\langle \alpha_n^{(0)}| \otimes \langle \beta_m| \right) \quad (17)$$

where $\{|\alpha_n^{(0)}\rangle\}$ is *any* orthonormal basis for \mathcal{H}_A and $\{|\beta_m\rangle\}$ is *any* orthonormal basis for \mathcal{H}_B . Then the reduced density operator at time t_0 is

$$\bar{\rho}_0 = \sum_n \overline{\sigma_n^{(0)}} |\alpha_n^{(0)}\rangle \langle \alpha_n^{(0)}|$$

where

$$\overline{\sigma_n^{(0)}} = \sum_m \sigma_{n,m}^{(0)}.$$

Comparing this to (12), we see that, if we fix a time t , there should be some basis $\{|\alpha_n\rangle\}$ of \mathcal{H}_A and some quantities $\overline{\sigma_n}$ such that

$$\bar{\rho}(t) = \sum_n \overline{\sigma_n} |\alpha_n\rangle \langle \alpha_n|$$

and our above analysis holds. *[Is this true?]*

Now,

$$\begin{aligned} \left(\langle \alpha_n^{(0)} | \otimes \langle \beta_m | \right) \hat{H}(t') \hat{\rho}_0 \left(|\alpha_n^{(0)}\rangle \otimes |\alpha_m^{(0)}\rangle \right) &= \sigma_{n,m}^{(0)} \left(\langle \alpha_n^{(0)} | \otimes \langle \beta_m | \right) \hat{H}(t') |\Psi_n^{(0)}\rangle \\ &= \left(\langle \alpha_n^{(0)} | \otimes \langle \beta_m | \right) \hat{\rho}_0 \hat{H}(t') \left(|\alpha_n^{(0)}\rangle \otimes |\alpha_m^{(0)}\rangle \right) \end{aligned}$$

so the first-order eigenvalue correction (14) vanishes. We will assume that the second-order eigenvalue correction $\overline{\sigma_n^{(2)}}$ does not vanish and turn to (16).

Substituting (11) into (16) and using that

$$\langle \alpha_n^{(0)} | \otimes \langle \beta_m | \left\{ \hat{H}_I(t') \hat{H}_I(t''), \hat{\rho}_0 \right\} |\alpha_n^{(0)}\rangle \otimes |\beta_m\rangle = 2\sigma_{n,m}^{(0)} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') |\alpha_n^{(0)}\rangle \otimes |\beta_m\rangle$$

we get

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sum_m \left(2\sigma_{n,m}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \\ \left. - \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{\rho}_0 \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right) + \mathcal{O}(\lambda^2). \quad (18)$$

Putting in (17),

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sum_m \left(2\sigma_{n,m}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \\ \left. - \sum_{n^*} \sum_{m^*} \sigma_{n^*,m^*}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') | \alpha_{n^*}^{(0)} \rangle \otimes | \beta_{m^*}^{(0)} \rangle \langle \alpha_{n^*}^{(0)} | \otimes \langle \beta_{m^*}^{(0)} | \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right) + \mathcal{O}(\lambda^2)$$

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sum_m \left(2\sigma_{n,m}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \\ \left. - \sum_{n^*} \sum_{m^*} \sigma_{n^*,m^*}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_{n^*}^{(0)} | \otimes \langle \beta_{m^*}^{(0)} | \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') | \alpha_{n^*}^{(0)} \rangle \otimes | \beta_{m^*}^{(0)} \rangle dt'' dt' \right) + \mathcal{O}(\lambda^2)$$

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \left(\sum_{n \in I} \sum_m 2\sigma_{n,m}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \\ \left. - \sum_{n^*} \sum_{m^*} \sigma_{n^*,m^*}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_{n^*}^{(0)} | \otimes \langle \beta_{m^*}^{(0)} | \hat{H}_I(t'') \left[\left(\sum_{n \in I} | \alpha_n^{(0)} \rangle \langle \alpha_n^{(0)} | \right) \otimes \cdot \right] \hat{H}_I(t') | \alpha_{n^*}^{(0)} \rangle \otimes | \beta_{m^*}^{(0)} \rangle dt'' dt' \right) + \mathcal{O}(\lambda^2)$$

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \left(\sum_{n \in I} \sum_m 2\sigma_{n,m}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \\ \left. - \sum_{n^*} \sum_{m^*} \sigma_{n^*,m^*}^{(0)} \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_{n^*}^{(0)} | \otimes \langle \beta_{m^*}^{(0)} | \hat{H}_I(t'') \left[\left(1 - \sum_{n' \notin I} | \alpha_{n'}^{(0)} \rangle \langle \alpha_{n'}^{(0)} | \right) \otimes \cdot \right] \hat{H}_I(t') | \alpha_{n^*}^{(0)} \rangle \otimes | \beta_{m^*}^{(0)} \rangle dt'' dt' \right) + \mathcal{O}(\lambda^2)$$

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sum_m \sigma_{n,m}^{(0)} \left(\left[2 \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \right. \\ \left. \left. - \int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') \hat{H}_I(t'') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right] \right. \\ \left. + \sum_{n' \notin I} \sum_{m'} \int_{t_0}^t \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t'') | \alpha_{n'}^{(0)} \rangle \otimes | \beta_{m'} \rangle dt'' \int_{t_0}^t \langle \alpha_{n'}^{(0)} | \otimes \langle \beta_{m'} | \hat{H}_I(t') | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt' \right) + \mathcal{O}(\lambda^2)$$

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sum_m \sigma_{n,m}^{(0)} \left(\int_{t_0}^t \int_{t_0}^{t'} \langle \alpha_n^{(0)} | \otimes \langle \beta_m | [\hat{H}_I(t'), \hat{H}_I(t'')] | \alpha_n^{(0)} \rangle \otimes | \beta_m \rangle dt'' dt' \right. \\ \left. + \sum_{n' \notin I} \sum_{m'} \left| \int_{t_0}^t \langle \alpha_n^{(0)} | \otimes \langle \beta_m | \hat{H}_I(t') | \alpha_{n'}^{(0)} \rangle \otimes | \beta_{m'} \rangle dt' \right|^2 \right) + \mathcal{O}(\lambda^2).$$

[The above derivation should be double checked.] We have derived the rather interesting result that

$$\Delta S = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sum_m \sigma_{n,m}^{(0)} \left(\int_{t_0}^t \int_{t_0}^{t'} \left\langle \alpha_n^{(0)} \right| \otimes \left\langle \beta_m \right| \left[\hat{H}_I(t'), \hat{H}_I(t'') \right] \left| \alpha_n^{(0)} \right\rangle \otimes \left| \beta_m \right\rangle dt'' dt' \right. \\ \left. + \sum_{n' \notin I} \sum_{m'} \left| \int_{t_0}^t \left\langle \alpha_n^{(0)} \right| \otimes \left\langle \beta_m \right| \hat{H}_I(t') \left| \alpha_{n'}^{(0)} \right\rangle \otimes \left| \beta_{m'} \right\rangle dt' \right|^2 \right) + \mathcal{O}(\lambda^2). \quad (19)$$

The second term is fascinating; it looks like a transition amplitude.

Assuming that these terms do not both vanish, we see that the entropy grows proportionally to $(t - t_0)^2$.

4(d) Total entropy of a closed system

What if we want to consider the entropy of the entire universe, rather than the subsystem entropy? This is equivalent to taking \mathcal{H}_B to consist of a single possible state. When we consider the entire universe, the change in entropy should be zero. Thus (19) becomes

$$0 = \frac{1}{\hbar^2} \lambda^2 \log \frac{1}{\lambda^2} \sum_{n \in I} \sigma_n^{(0)} \left(\int_{t_0}^t \int_{t_0}^{t'} \left\langle \alpha_n^{(0)} \right| \left[\hat{H}_I(t'), \hat{H}_I(t'') \right] \left| \alpha_n^{(0)} \right\rangle dt'' dt' + \sum_{n' \notin I} \left| \int_{t_0}^t \left\langle \alpha_n^{(0)} \right| \hat{H}_I(t') \left| \alpha_{n'}^{(0)} \right\rangle dt' \right|^2 \right) + \mathcal{O}(\lambda^2).$$

This equation should remain true regardless of our initial state. In particular, we should have

$$- \int_{t_0}^t \int_{t_0}^{t'} \langle \Psi_n | \left[\hat{H}_I(t'), \hat{H}_I(t'') \right] | \Psi_n \rangle dt'' dt' = \sum_{n' \neq n} \left| \int_{t_0}^t \langle \Psi_n | \hat{H}_I(t') | \Psi_{n'} \rangle dt' \right|^2 \quad (20)$$

for all possible bases $\{\Psi_n\}$ of \mathcal{H} and for all times t .

5 Example: Interaction between Qubits

In this section we will study a simple bipartite system and calculate the growth in subsystem entropy due to a weak interaction. Consider the Hilbert space $\{|0\rangle, |1\rangle\} \otimes \{|0\rangle, |1\rangle\}$ describing two qubits. To simplify notation, write $|ab\rangle$ to mean $|a\rangle \otimes |b\rangle$.

Suppose the qubits interact according to the Hamiltonian

$$\hat{H} = \hat{S}_z^A \otimes \hat{S}_z^B + \lambda \hat{S}_x^A \otimes \hat{S}_x^B.$$

One can calculate that this Hamiltonian has eigenstates and eigenvalues

$$\begin{aligned} |E_0\rangle &= \frac{1}{\sqrt{2}} (|10\rangle - |01\rangle) & E_0 &= \frac{\hbar}{2} (-1 - \lambda) \\ |E_1\rangle &= \frac{1}{\sqrt{2}} (|10\rangle + |01\rangle) & E_1 &= \frac{\hbar}{2} (-1 + \lambda) \\ |E_2\rangle &= \frac{1}{\sqrt{2}} (|11\rangle - |00\rangle) & E_2 &= \frac{\hbar}{2} (1 - \lambda) \\ |E_3\rangle &= \frac{1}{\sqrt{2}} (|11\rangle + |00\rangle) & E_3 &= \frac{\hbar}{2} (1 + \lambda). \end{aligned}$$

Let's start by calculating the exact change in entropy; then we can calculate the perturbative result and compare.

Suppose at $t = 0$ we start in the initial state $|11\rangle = \frac{1}{\sqrt{2}}(|E_2\rangle + |E_3\rangle)$. At time t , the state is then

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \left(e^{-itE_2/\hbar} |E_2\rangle + e^{-itE_3/\hbar} |E_3\rangle \right)$$

and the density operator is

$$\hat{\rho} = \frac{1}{2} |E_2\rangle \langle E_2| + \frac{1}{2} |E_3\rangle \langle E_3| + \frac{1}{2} e^{it(E_3-E_2)/\hbar} |E_2\rangle \langle E_3| + \frac{1}{2} e^{-it(E_3-E_2)/\hbar} |E_3\rangle \langle E_2|.$$

Then the reduced density operator for the first qubit is

$$\begin{aligned} \bar{\rho} &= \cdot \otimes \langle 0| \hat{\rho} \cdot \otimes |0\rangle + \cdot \otimes \langle 1| \hat{\rho} \cdot \otimes |1\rangle \\ &= \frac{1}{2} \left(1 + \cos \left[(E_3 - E_2) \frac{t}{\hbar} \right] \right) |1\rangle \langle 1| + \frac{1}{2} \left(1 - \cos \left[(E_3 - E_2) \frac{t}{\hbar} \right] \right) |0\rangle \langle 0|. \end{aligned}$$

Recognizing that $\frac{E_3-E_2}{\hbar} = \lambda$, we see that the Von Neumann subsystem entropy is

$$S = \frac{1}{2} (1 + \cos \lambda t) \log \left(1 + \cos \frac{\lambda t}{2} \right) + \frac{1}{2} (1 - \cos \lambda t) \log \left(1 - \cos \frac{\lambda t}{2} \right).$$

Approximating $\cos x \approx 1 - x^2$, this becomes

$$\begin{aligned} S &\approx \frac{1}{2} (2 - \lambda^2 t^2) \log (2 - \lambda^2 t^2/4) + \frac{1}{2} \lambda^2 t^2 \log (\lambda^2 t^2) \\ &\approx \log 2 - \frac{1}{2} \lambda^2 t^2 \log \left(\frac{1}{\lambda^2 t^2} \right). \end{aligned}$$

We observe that the leading order term for the change in entropy is of order $\lambda^2 t^2 \log \frac{1}{\lambda^2 t^2}$. Interestingly, the entropy decreases with time. This makes sense given that we started in a maximally entangled state (i.e. a Bell state). *[Actually, I'm not sure that this is a good explanation of the decreasing entropy...]*

We did the derivation above by solving the Schrödinger equation exactly. Let's now apply our perturbative calculation and see that we get the same result. The Schrödinger picture Hamiltonian is $\hat{H} = \hat{H}_0 + \lambda \hat{V}$ where

$$\begin{aligned} \hat{H}_0 &= \frac{\hbar}{2} (-|E_0\rangle \langle E_0| - |E_1\rangle \langle E_1| + |E_2\rangle \langle E_2| + |E_3\rangle \langle E_3|) \\ \hat{V} &= \frac{\hbar}{2} (-|E_0\rangle \langle E_0| + |E_1\rangle \langle E_1| - |E_2\rangle \langle E_2| + |E_3\rangle \langle E_3|). \end{aligned}$$

Because $[\hat{H}_0, \hat{V}] = 0$ we have $[\hat{U}(0, t), \hat{V}] = 0$. Thus the interaction picture Hamiltonian is just

$$\hat{H}_I = \hat{V} = \frac{\hbar}{2} (-|E_0\rangle \langle E_0| + |E_1\rangle \langle E_1| - |E_2\rangle \langle E_2| + |E_3\rangle \langle E_3|).$$

We must now diagonalize the initial density operator in a basis that easily separates into bases for the two subspaces. Indeed, we get

$$\begin{aligned} \hat{\rho} &= \frac{1}{2} |E_2\rangle \langle E_2| + \frac{1}{2} |E_3\rangle \langle E_3| + \frac{1}{2} e^{it\lambda} |E_2\rangle \langle E_3| + \frac{1}{2} e^{-it\lambda} |E_3\rangle \langle E_2| \\ &= \left(\frac{1}{\sqrt{2}} e^{it\lambda} |E_2\rangle + \frac{1}{\sqrt{2}} |E_3\rangle \right) \left(\frac{1}{\sqrt{2}} e^{-it\lambda} \langle E_2| + \frac{1}{\sqrt{2}} \langle E_3| \right) \\ &= \left(\frac{1}{2} (1 - e^{it\lambda}) |00\rangle + \frac{1}{2} (1 + e^{it\lambda}) |11\rangle \right) \left(\frac{1}{2} (1 + e^{-it\lambda}) \langle 00| + \frac{1}{2} (1 + e^{-it\lambda}) \langle 11| \right) \end{aligned}$$

[This is not a separable state, so we cannot use the perturbation formula. Should maybe change the example to a separable state.]

Appendices

These are extra calculations that are not relevant to the above discussion. They should be removed from documents to be shared. (I don't delete them now because they could be useful in future work.)

A The first-order approximation for pure states in scattering

We will continue for scattering between two particles. We will assume that the particles are distinguishable so that we don't need to restrict ourselves to symmetric or antisymmetric states. We will also assume that the particles interact in a way that only depends on the distance between them. That is, in the position basis, we can write

$$\langle x_1, x_2 | \hat{V}(t) | \psi \rangle = \int_{x_1, x_2} d^3x_1 d^3x_2 V(x_1 - x_2) \langle x_1, x_2 | \psi \rangle. \quad (21)$$

(Note that this is not time dependent.) Take \hat{H}_0 to be the free particle Hamiltonian.

It will be easiest to work in the momentum basis. We will do our calculations for momentum eigenstates—that is, plane waves. Write momentum eigenstates as $|k_1, k_2\rangle$. In the position basis these have wavefunctions

$$\langle x_1, x_2 | k_1, k_2 \rangle = \frac{1}{(2\pi)^3} e^{ik_1 \cdot x_1 + ik_2 \cdot x_2}. \quad (22)$$

Here x_1 and x_2 are the (3-vector) positions of the first and second particles and k_1 and k_2 are the (3-vector) momenta of the first and second particle. Also, we note that $|k_1, k_2\rangle$ has energy $\frac{1}{2m_1}k_1^2 + \frac{1}{2m_2}k_2^2$ and so

$$\hat{U}(t_0, t) |k_1, k_2\rangle = e^{-\frac{i}{\hbar}(t-t_0)\left(\frac{1}{2m_1}k_1^2 + \frac{1}{2m_2}k_2^2\right)} |k_1, k_2\rangle. \quad (23)$$

Let $\tilde{V}(p)$ be the Fourier transform of $V(x_1 - x_2)$ so that

$$V(x_1 - x_2) = \int_{-\infty}^{\infty} \frac{d^3p}{(2\pi)^3} \tilde{V}(p) e^{ip \cdot (x_1 - x_2)}. \quad (24)$$

We now consider (2) and take the limits $t_0 \rightarrow -\infty$ and $t \rightarrow \infty$ because, in scattering, we measure states long before and long after the scattering occurs. Then, dropping the higher-order terms,

$$|\Psi_I, \infty\rangle = |\Psi_I, -\infty\rangle - \frac{i}{\hbar} \int_{-\infty}^{\infty} \hat{H}_I(t) |\Psi_I, -\infty\rangle dt.$$

Take the initial state to be $|\Psi_I, -\infty\rangle = |k_1, k_2\rangle$. Multiplying by the bra $\langle k'_1, k'_2|$,

$$\begin{aligned} \langle k'_1, k'_2 | \Psi_I, \infty \rangle &= \langle k'_1, k'_2 | k_1, k_2 \rangle - \frac{i}{\hbar} \int_{-\infty}^{\infty} \langle k'_1, k'_2 | \hat{H}_I(t) | k_1, k_2 \rangle dt \\ &= \delta^3(k'_1 - k_1) \delta^3(k'_2 - k_2) - \frac{i}{\hbar} \int_{-\infty}^{\infty} \langle k'_1, k'_2 | \hat{H}_I(t) | k_1, k_2 \rangle dt. \end{aligned} \quad (25)$$

This motivates us to compute the matrix element $\langle k'_1, k'_2 | \hat{H}_I(t) | k_1, k_2 \rangle$. We can start by using (23):

$$\begin{aligned} \langle k'_1, k'_2 | \hat{H}_I(t) | k_1, k_2 \rangle &= \langle k'_1, k'_2 | \hat{U}(t_0, t)^\dagger \hat{V}(t) \hat{U}(t_0, t) | k_1, k_2 \rangle \\ &= \langle k'_1, k'_2 | e^{\frac{i}{\hbar}(t-t_0)\left(\frac{1}{2m_1}k_1'^2 + \frac{1}{2m_2}k_2'^2\right)} \hat{V}(t) e^{-\frac{i}{\hbar}(t-t_0)\left(\frac{1}{2m_1}k_1^2 + \frac{1}{2m_2}k_2^2\right)} | k_1, k_2 \rangle \\ &= e^{-\frac{i}{\hbar}(t-t_0)\left(\frac{1}{2m_1}k_1'^2 + \frac{1}{2m_2}k_2'^2 - \frac{1}{2m_1}k_1^2 - \frac{1}{2m_2}k_2^2\right)} \langle k'_1, k'_2 | \hat{V}(t) | k_1, k_2 \rangle. \end{aligned}$$

We see that

$$\begin{aligned} \langle k'_1, k'_2 | \hat{V}(t) | k_1, k_2 \rangle &= \langle k'_1, k'_2 | \left\{ \int_{x_1, x_2} d^3x_1 d^3x_2 |x_1, x_2\rangle \langle x_1, x_2| \right\} \hat{V}(t) | k_1, k_2 \rangle \\ &= \int_{x_1, x_2} d^3x_1 d^3x_2 \langle k'_1, k'_2 | x_1, x_2 \rangle \langle x_1, x_2 | \hat{V}(t) | k_1, k_2 \rangle \\ &= \int_{x_1, x_2} \frac{d^3x_1 d^3x_2}{(2\pi)^6} e^{-ik'_1 \cdot x_1 - ik'_2 \cdot x_2} V(x_1 - x_2) e^{ik_1 \cdot x_1 + ik_2 \cdot x_2} \quad (\text{by (21) and (22)}) \\ &= \int_{-\infty}^{\infty} \frac{d^3p}{(2\pi)^9} \int_{x_1, x_2} d^3x_1 d^3x_2 e^{-ik'_1 \cdot x_1 - ik'_2 \cdot x_2} \tilde{V}(p) e^{ip \cdot (x_1 - x_2)} e^{ik_1 \cdot x_1 + ik_2 \cdot x_2} \quad (\text{by (24)}) \\ &= \int_{-\infty}^{\infty} \frac{d^3p}{(2\pi)^9} \tilde{V}(p) \int_{x_1, x_2} d^3x_1 d^3x_2 e^{i(k_1 - k'_1 + p) \cdot x_1} e^{i(k_2 - k'_2 - p) \cdot x_2} \\ &= \int_{-\infty}^{\infty} \frac{d^3p}{(2\pi)^9} \tilde{V}(p) (2\pi)^6 \delta^3(k_1 - k'_1 + p) \delta^3(k_2 - k'_2 - p) \\ &= \frac{1}{(2\pi)^3} \tilde{V}(k'_1 - k_1) \delta^3(k'_1 - k_1 - k'_2 + k_2). \end{aligned}$$

so

$$\langle k'_1, k'_2 | \hat{H}_I(t) | k_1, k_2 \rangle = e^{-\frac{i}{\hbar}(t-t_0)\left(\frac{1}{2m_1}k_1'^2 + \frac{1}{2m_2}k_2'^2 - \frac{1}{2m_1}k_1^2 - \frac{1}{2m_2}k_2^2\right)} \frac{1}{(2\pi)^3} \tilde{V}(k'_1 - k_1) \delta^3(k'_1 - k_1 - k'_2 + k_2)$$

If we integrate over all time, we get a delta function:

$$\int_{-\infty}^{\infty} dt \langle k'_1, k'_2 | \hat{H}_I(t) | k_1, k_2 \rangle = \frac{\hbar}{(2\pi)^2} \tilde{V}(k'_1 - k_1) \delta\left(\frac{1}{2m_1}k_1'^2 + \frac{1}{2m_2}k_2'^2 - \frac{1}{2m_1}k_1^2 - \frac{1}{2m_2}k_2^2\right) \delta^3(k'_1 - k_1 - k'_2 + k_2). \quad (26)$$

Putting this into (25),

$$\begin{aligned} \langle k'_1, k'_2 | \Psi_I, \infty \rangle &= \delta^3(k'_1 - k_1) \delta^3(k'_2 - k_2) \\ &\quad - \frac{i}{(2\pi)^2} \tilde{V}(k'_1 - k_1) \delta^3(k'_1 - k_1 - k'_2 + k_2) \delta\left(\frac{k_1'^2}{2m_1} + \frac{k_2'^2}{2m_2} - \frac{k_1^2}{2m_1} - \frac{k_2^2}{2m_2}\right). \quad (27) \end{aligned}$$

Interestingly, we see that the delta functions cause energy and momentum to be conserved.

B Reduced density operator

We are interested in the situation where the second particle exits the system and becomes lost. We therefore consider the reduced density operator

$$\begin{aligned} \hat{\rho}_{\text{reduced}}(t) &= \int d^3k_2 \langle k_2 | \hat{\rho}(t) | k_2 \rangle \\ &= \hat{\rho}_{\text{reduced}}(t_0) - \frac{i}{\hbar} \int_{t_0}^t dt' \int_{-\infty}^{\infty} d^3k_2 \langle k_2 | \left[\hat{H}_I(t'), \hat{\rho}(t_0) \right] | k_2 \rangle + \mathcal{O}(\hat{H}_I(t)^2). \end{aligned}$$

In the momentum basis the matrix elements are

$$\langle k_1' | \hat{\rho}_{\text{reduced}}(t) | k_1 \rangle = \langle k_1' | \hat{\rho}_{\text{reduced}}(t_0) | k_1 \rangle - \frac{i}{\hbar} \int_{t_0}^t dt' \int_{-\infty}^{\infty} d^3 k_2 \langle k_1', k_2 | [\hat{H}_I(t'), \hat{\rho}(t_0)] | k_1, k_2 \rangle + \mathcal{O}(\hat{H}_I(t)^2).$$

As before, we drop the higher order terms and take $t \rightarrow \infty$ and $t_0 \rightarrow -\infty$. Then

$$\begin{aligned} \langle k_1' | \hat{\rho}_{\text{reduced}}(\infty) | k_1 \rangle &= \langle k_1' | \hat{\rho}_{\text{reduced}}(-\infty) | k_1 \rangle - \frac{i}{\hbar} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d^3 k_2 \langle k_1', k_2 | [\hat{H}_I(t), \hat{\rho}(-\infty)] | k_1, k_2 \rangle \\ &= \langle k_1' | \hat{\rho}_{\text{reduced}}(-\infty) | k_1 \rangle \\ &\quad - \frac{i}{\hbar} \int_{-\infty}^{\infty} d^3 k_2 d^3 \tilde{k}_1 d^3 \tilde{k}_2 \left\{ \int_{-\infty}^{\infty} dt \langle k_1', k_2 | \hat{H}_I(t) | \tilde{k}_1, \tilde{k}_2 \rangle \right\} \langle \tilde{k}_1, \tilde{k}_2 | \hat{\rho}(-\infty) | k_1, k_2 \rangle \\ &\quad + \frac{i}{\hbar} \int_{-\infty}^{\infty} d^3 k_2 d^3 \tilde{k}_1 d^3 \tilde{k}_2 \langle k_1', k_2 | \hat{\rho}(-\infty) | \tilde{k}_1, \tilde{k}_2 \rangle \left\{ \int_{-\infty}^{\infty} dt \langle \tilde{k}_1, \tilde{k}_2 | \hat{H}_I(t) | k_1, k_2 \rangle \right\}. \end{aligned}$$

Plugging in our result (26),

$$\begin{aligned} \langle k_1' | \hat{\rho}_{\text{reduced}}(\infty) | k_1 \rangle &= \langle k_1' | \hat{\rho}_{\text{reduced}}(-\infty) | k_1 \rangle \\ &\quad - i \int_{-\infty}^{\infty} d^3 k_2 d^3 \tilde{k}_1 d^3 \tilde{k}_2 \left(\frac{1}{(2\pi)^2} \tilde{V}(k_1' - \tilde{k}_1) \delta\left(\frac{k_1'^2}{2m_1} + \frac{k_2^2}{2m_2} - \frac{\tilde{k}_1^2}{2m_1} - \frac{\tilde{k}_2^2}{2m_2}\right) \right. \\ &\quad \left. \delta^3(k_1' - \tilde{k}_1 - k_2 + \tilde{k}_2) \langle \tilde{k}_1, \tilde{k}_2 | \hat{\rho}(-\infty) | k_1, k_2 \rangle \right) \\ &\quad + i \int_{-\infty}^{\infty} d^3 k_2 d^3 \tilde{k}_1 d^3 \tilde{k}_2 \left(\frac{1}{(2\pi)^2} \tilde{V}(\tilde{k}_1 - k_1) \delta\left(\frac{\tilde{k}_1^2}{2m_1} + \frac{\tilde{k}_2^2}{2m_2} - \frac{k_1^2}{2m_1} - \frac{k_2^2}{2m_2}\right) \right. \\ &\quad \left. \delta^3(\tilde{k}_1 - k_1 - \tilde{k}_2 + k_2) \langle k_1', k_2 | \hat{\rho}(-\infty) | \tilde{k}_1, \tilde{k}_2 \rangle \right). \end{aligned}$$

Finally, we use the delta functions to get rid of the \tilde{k}_1 integrals. This gives us

$$\begin{aligned} \langle k_1' | \hat{\rho}_{\text{reduced}}(\infty) | k_1 \rangle &= \langle k_1' | \hat{\rho}_{\text{reduced}}(-\infty) | k_1 \rangle \\ &\quad - i \int_{-\infty}^{\infty} d^3 k_2 d^3 \tilde{k}_2 \frac{1}{(2\pi)^2} \tilde{V}(k_2 - \tilde{k}_2) \langle k_1' + \tilde{k}_2 - k_2, \tilde{k}_2 | \hat{\rho}(-\infty) | k_1, k_2 \rangle D(k_1', k_2, \tilde{k}_2) \\ &\quad + i \int_{-\infty}^{\infty} d^3 k_2 d^3 \tilde{k}_2 \frac{1}{(2\pi)^2} \tilde{V}(\tilde{k}_2 - k_2) \langle k_1', k_2 | \hat{\rho}(-\infty) | k_1 + \tilde{k}_2 - k_2, \tilde{k}_2 \rangle D(k_1, k_2, \tilde{k}_2). \end{aligned}$$

where

$$D(k_1, k_2, \tilde{k}_2) = \delta\left(\frac{1}{2} k_2^2 \left(\frac{1}{m_1} - \frac{1}{m_2}\right) + \frac{1}{2} \tilde{k}_2^2 \left(\frac{1}{m_1} + \frac{1}{m_2}\right) - \frac{1}{2m_1} (-k_1 \cdot \tilde{k}_2 + k_1 \cdot k_2 + \tilde{k}_2 \cdot k_2)\right).$$

Then, assuming \tilde{V} is an even function,

$$\begin{aligned} \langle k_1' | \hat{\rho}_{\text{reduced}}(\infty) | k_1 \rangle &= \left(\langle k_1' | \hat{\rho}_{\text{reduced}}(-\infty) | k_1 \rangle + \right. \\ &\quad \left. \int_{-\infty}^{\infty} \frac{d^3 k_2 d^3 \tilde{k}_2}{(2\pi)^2} \tilde{V}(k_2 - \tilde{k}_2) \left(P(k_1, k_1', k_2, \tilde{k}_2) D(k_1, k_2, \tilde{k}_2) + P^*(k_1', k_1, k_2, \tilde{k}_2) D(k_1', k_2, \tilde{k}_2) \right) \right) \end{aligned} \quad (28)$$

where

$$P(k_1, k_1', k_2, \tilde{k}_2) = i \langle k_1', k_2 | \hat{\rho}(-\infty) | k_1 + \tilde{k}_2 - k_2, \tilde{k}_2 \rangle.$$