Quantum Mechanics II
(some notes from the lecture)

Adapted by Lídia del Rio, Peter Strassmann

2012

These scribbles are based on the lecture notes of Quantum Mechanics II at ETH Zurich. The idea is to have an English version of the script, including all the steps and explanations that are missing from the original. QM2 for dummies, if you will.

If you want to contribute, drop me an email: `delrio@phys.ethz.ch`.

Notation

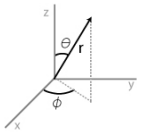
	"Where did that come from?" Well, from many frustrated tentatives...
	Approximation alert
[text]	Comments
OS	Original script
SE	Schrödinger equation
DE	Differential equation
PT	Perturbation theory
h.c.	Hermitian conjugate

Chapter 1

Scattering theory

1.3 Interlude: radial potential

The goal of this section is to solve the SE for systems with spherical symmetry, i.e., with a radial potential, $V(\mathbf{x}) = V(r)$. For instance, there could be a very round ion at position $r = 0$, or maybe a tiny sphere made out of gelatine. It will be useful to represent this world in spherical coordinates, $\mathbf{x} = (r, \theta, \phi)$.



For a potential $V(r)$, the SE reads

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(r,\theta,\phi) = [E - V(r)]\Psi(r,\theta,\phi). \tag{1.1}$$

Expanding the Laplacian ∇^2 in spherical coordinates, we obtain

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{1}{r^2}\underbrace{\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial}{\partial\phi}\right) + \frac{1}{\sin^2}\frac{\partial^2}{\partial\theta^2}\right)}_{M^2(\theta,\phi)}\right]\Psi(r,\theta,\phi) = [E - V(r)]\Psi(r,\theta,\phi). \tag{1.2}$$

The term M^2 only depends on the angles θ and ϕ. Its eigenfunctions are the spherical harmonics $Y_{\ell,m}$ (which you may have seen before in Mathematical Methods for Physics or Electrodynamics),

$$Y_{\ell,m}(\theta,\phi) = \sqrt{\frac{(2\ell+1)(\ell-m)!}{4\pi\ (\ell+m)!}}\ P_\ell^m(\cos\theta)\ e^{im\phi}, \tag{1.3}$$

where ℓ and m are the quantum numbers corresponding to the "angular momentum" of the particle and the z-component of that angular momentum, so they are integers such that $\ell \geq 0, |m| \leq \ell$. P_ℓ^m are the Legendre polynomials,

$$P_\ell^m(x) = \frac{(-1)^m}{2^\ell \ell!}(1-x^2)^{m/2}\frac{d^{\ell+m}}{dx^{\ell+m}}(x^2-1)^\ell \tag{1.4}$$

(don't worry memorizing those formulas). The eigenvalue of M^2 corresponding to each spherical harmonic is given by

$$M^2(\theta,\phi)Y_{\ell,m}(\theta,\phi) = \ell(\ell+1)Y_{\ell,m}(\theta,\phi). \tag{1.5}$$

The spherical harmonics have some neat properties, like being all orthogonal and spanning the space of functions on the sphere.

For now we are going to look for wave functions of the form $\Psi_{\ell m} = R_{\ell m}(r)Y_{\ell m}(\theta,\phi)$. Later on we can generalize this: since the SE is a linear differential equation, any linear combination of solutions, $\Psi = \sum_{\ell,m} c_{\ell m}\Psi_{\ell m}$, is also a solution. So far we have

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{1}{r^2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial}{\partial\phi}\right) + \frac{1}{\sin^2}\frac{\partial^2}{\partial\theta^2}\right)\right]R_{\ell m}(r)Y_{\ell m}(\theta,\phi) = [E - V(r)]R_{\ell m}(r)Y_{\ell m}(\theta,\phi) \tag{1.6}$$

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2}\right]R_{\ell m}(r)Y_{\ell m}(\theta,\phi) = [E - V(r)]R_{\ell m}(r)Y_{\ell m}(\theta,\phi), \tag{1.7}$$

from which we may conclude that $R(r)$ does not depend on m. If we fix θ and ϕ, we can see $Y_{\ell m}(\theta,\phi) \neq 0$ as a constant and get rid of it. We will focus on finding the radial part $R_\ell(r)$ of the wave function,

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2}\right]R_\ell(r) = [E - V(r)]R_\ell(r). \tag{1.8}$$

The potential $V(r)$ is still too general for us. To keep the hope of finding an exact solution for the equation, we will consider only step-like potentials, ,

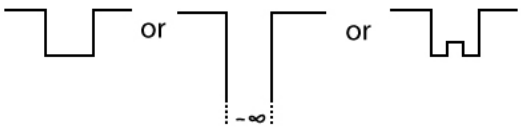
$$V(r) = \begin{cases} V_0, & r \in [0, a_0) \\ V_1, & r \in [a_0, a_1) \\ \vdots & \\ V_n, & r \in [a_{n-1}, \infty). \end{cases} \tag{1.9}$$

You could argue that that is not a very natural form for a physical potential, and you would be absoulutely right. But it is a good first step. Our approach will be to first solve the SE for each region of space where the potential is flat ($r \in [a_{i-1}, a_i), V(r) = V_i$). This will give us a set of solutions for that region, $R_\ell^{i,1}, R_\ell^{i,2}$, etc. We will then combine them to find $R_\ell(r)$ for the whole universe. We will do this by making sure that the global R is still a good candidate for a wave function:

1. $R_\ell(r)$ is continuous (we will enforce this at the borders, $r = a_i$);
2. $\frac{\partial}{\partial r}R_\ell(r)$ is continuous if $|V(r)| < \infty$ (we will enforce this at the borders, $r = a_i$);
3. $R_\ell(r)$ can be normalized (we will discard solutions that diverge).

[Where do these conditions come from? Quite possibly from QM1, but I should investigate this further. To check: Schumacher's book]

For now we will consider only the case where our particule's energy is larger than the potential, $E > V_i, \forall i$ (we will get rid of this assumption in Section 1.3.2). Some examples are



Let us go back to the SE for a flat bit, $V(r) = V_i$,

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2}\right]R_\ell(r) = -\frac{2m}{\hbar^2}[E - V_i]R_\ell(r). \tag{1.10}$$

We will simplify this equation (and our life) by defining a new constant k,

$$k = \frac{\sqrt{2m(E - V_i)}}{h}, \tag{1.11}$$

so that the SE looks neater,

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2} + k^2\right]R_\ell(r) = 0 \tag{1.12}$$

(this only works if $E > V_i$; we will see the other cases in Section 1.3.2). We can simplify this equation further if we get rid of the constant k^2. To do so, we define a new variable,

$$\rho := kr, \qquad r = \frac{\rho}{k}, \qquad \frac{\partial^2}{\partial r^2} = \frac{1}{k^2}\frac{\partial^2}{\partial \rho^2}, \tag{1.13}$$

And now the SE is simply

$$\left[\frac{k}{\rho}k^2\frac{\partial^2}{\partial \rho^2}\frac{\rho}{k} - k^2\frac{\ell(\ell+1)}{\rho^2} + k^2\right]R_\ell(\rho) = 0 \tag{1.14}$$

$$\left[\frac{1}{\rho}\frac{\partial^2}{\partial \rho^2}\rho - \frac{\ell(\ell+1)}{\rho^2} + 1\right]R_\ell(\rho) = 0. \tag{1.15}$$

Good. Now that we have simplified our problem to turn the SE into a rather standard differential equation, we can solve it.

1.3.1 Bessel functions

This differential equation depends on ℓ, so we will do the following:

1. find a solution for the case $\ell = 0$, $R_0(\ell)$;
2. find a recursive solution for $\ell > 0$, $R_{\ell+1} = f(R_\ell)$.

This proof is fastidious, full of miracles (😳), and not particularly enlightening. Feel free to skip it in order to preserve your soul and enthusiasm for quantum mechanics.

Case $\ell = 0$

When $\ell = 0$, the DE is just

$$\frac{1}{\rho}\frac{\partial^2}{\partial\rho^2}\rho R_\ell(\rho) + R_\ell(\rho) = 0. \tag{1.16}$$

This equation could not get more textbook than this, so we can look it up and find that it has two solutions:

$$R_0(\rho) = \frac{\sin\rho}{\rho}, \qquad R_0(\rho) = \frac{-\cos\rho}{\rho}, \tag{1.17}$$

where the sign of each solution is just a convention. At first sight, we can see that the second solution diverges for $\rho \to 0$, so when later we choose the physical solutions of the SE, we will discard that one for small ρ. Later.

Case $\ell > 0$

In this case, our DE is

$$\frac{1}{\rho}\frac{\partial^2}{\partial\rho^2}\rho R_\ell(\rho) + \left[-\frac{\ell(\ell+1)}{\rho^2} + 1\right] R_\ell(\rho) = 0. \tag{1.18}$$

We will start with a substitution, defining a new function $\chi_\ell(\rho)$ such that

$$😳 \quad R_\ell(\rho) = \rho^\ell \chi_\ell(\rho). \tag{1.19}$$

We will omit the "(ρ)" for simplicity's sake. The second derivative of Eq. 1.18 becomes

$$\begin{aligned}\frac{\partial^2}{\partial\rho^2}\rho R_\ell(\rho) &= \frac{\partial^2}{\partial\rho^2}\rho^{\ell+1}\chi_\ell \\ &= \ell(\ell+1)\rho^{\ell-1}\chi_\ell + 2(\ell+1)\rho^\ell\chi'_\ell + \rho^{\ell+1}\chi''_\ell.\end{aligned}$$

Back to Eq. 1.18, we get

$$\begin{aligned}0 &= \rho\left[\ell(\ell+1)\rho^{\ell-1}\chi_\ell + 2(\ell+1)\rho^\ell\chi'_\ell + \rho^{\ell+1}\chi''_\ell\right] + \left[-\frac{\ell(\ell+1)}{\rho^2} + 1\right]\rho^\ell\chi_\ell \\ &= \ell(\ell+1)\rho^{\ell-2}\chi_\ell + 2(\ell+1)\rho^{\ell-1}\chi'_\ell + \rho^\ell\chi''_\ell - \ell(\ell+1)\rho^{\ell-2}\chi_\ell + \rho^\ell\chi_\ell \\ \text{multiply by } \frac{1}{\rho} \hookrightarrow \quad 0 &= \chi_\ell + \frac{2(\ell+1)}{\rho}\chi'_\ell + \chi''_\ell.\end{aligned} \tag{1.20}$$

This is a decent equation, but not yet the recursive solution we are looking for. Hold on.

$$\begin{aligned}😳 \text{ take } \frac{\partial}{\partial\rho} \hookrightarrow \quad 0 &= \chi'_\ell - \frac{2(\ell+1)}{\rho^2}\chi'_\ell + \frac{2(\ell+1)}{\rho}\chi''_\ell + \chi'''_\ell \\ 😳 \text{ substitution: } \chi'_\ell = \chi\rho \hookrightarrow \quad 0 &= \chi\rho - \frac{2(\ell+1)}{\rho^2}\chi\rho + \frac{2(\ell+1)}{\rho}[\chi\rho]' + [\chi\rho]'' \\ &= \chi\rho - \frac{2(\ell+1)}{\rho}\chi + \frac{2(\ell+1)}{\rho}\chi + 2(\ell+1)\chi' + \chi''\rho + 2\chi' + 0 \\ &= \chi\rho + 2(\ell+2)\chi' + \chi''\rho \\ \text{multiply by } \frac{1}{\rho} \hookrightarrow \quad 0 &= \chi + \frac{2(\ell+2)}{\rho}\chi' + \chi''.\end{aligned}$$

But this is exactly Eq. 1.20, when $\ell \to \ell + 1$,

$$0 = \chi_{\ell+1} + \frac{2(\ell+2)}{\rho}\chi'_{\ell+1} + \chi''_{\ell+1}. \tag{1.21}$$

From here we can conclude that

$$\chi'_{\ell+1} = \chi = \frac{\chi'_\ell}{\rho}, \tag{1.22}$$

which is precisely the kind of recursive solution we were looking for. Going back all the way $\ell \to \ell-1 \to \cdots \to 0$, we get

$$\chi_\ell = \left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)^\ell \chi_0 \quad \overrightarrow{R_\ell = \rho^\ell \chi_\ell} \quad R_\ell(\rho) = \rho^\ell \left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)^\ell R_0(\rho). \tag{1.23}$$

Hooray!

Bessel and Neumann functions

Remember how we had two solutions for R_0? This gives origin to two families of solutions: the Bessel functions, $j_\ell(\rho)$, and the Neumann functions, $n_\ell(\rho)$. Here follow their expression and behaviour for small and large ρ.[1] Again, the minus signs are just a practical convention.

	Bessel functions	Neumann functions
$\ell = 0$	$j_0(\rho) = \frac{\sin\rho}{\rho}$	$n_0(\rho) = -\frac{\cos\rho}{\rho}$
general	$j_\ell(\rho) = (-1)^\ell \rho^\ell \left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)^\ell \frac{\sin\rho}{\rho}$	$n_\ell(\rho) = (-1)^{\ell+1} \rho^\ell \left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)^\ell \frac{\cos\rho}{\rho}$
$\rho \to 0$	$j_\ell(\rho) \approx \frac{\rho^\ell}{(2\ell+1)!!}$ (large ℓ : $j_\ell(\rho) \to 0$ quickly)	$n_\ell(\rho) \approx \frac{(2\ell-1)!!}{\rho^{\ell+1}}$ (diverges)
$\rho \to \infty$	$j_\ell(\rho) \approx \frac{1}{\rho}\ \sin\left(\rho - \frac{\pi\ell}{2}\right)$	$n_\ell(\rho) \approx -\frac{1}{\rho}\ \cos\left(\rho - \frac{\pi\ell}{2}\right)$

Hankel functions

We found two families of solutions for our DE. Remember that any linear combination of solutions of a linear differential equations is a solution (it might not be a physical solution, but we will get to that in a moment). Well, there are two linear combinations of Bessel and Neumann functions that will be especially useful for scattering problems. So useful indeed, they get their own name: the Hankel functions.

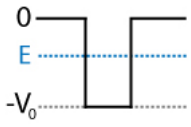
	h_ℓ (or $h_\ell^{(1)}$)	h_ℓ^* (or $h_\ell^{(2)}$)
definition	$h_\ell(\rho) = j_\ell(\rho) + i n_\ell(\rho)$	$h_\ell^*(\rho) = j_\ell(\rho) - i n_\ell(\rho)$
general	$h_\ell(\rho) = -i\ (-1)^\ell \rho^\ell \left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)^\ell \frac{e^{i\rho}}{\rho}$	$h_\ell^*(\rho) = i\ (-1)^\ell \rho^\ell \left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)^\ell \frac{e^{-i\rho}}{\rho}$
$\rho \to \infty$	$h_\ell(\rho) \approx -\frac{i}{\rho}\ e^{i\left(\rho - \frac{\pi\ell}{2}\right)}$	$h_\ell^*(\rho) \approx \frac{i}{\rho} e^{-i\left(\rho - \frac{\pi\ell}{2}\right)}$

Now we are finally ready to look at some concrete, if contrived, physical systems.

[1] For a proof of the asymptotic behaviour, see Exercise 2 of the 1st series.

1.3.2 Bound states

In this first example, our poor particle is trapped in a deep well, and apparently it does not have enough energy to escape!



$$V(r) = \begin{cases} -V_0, & r < a \\ 0, & r \geq a, \end{cases}$$

$$-\infty < -V_0 < E < 0.$$

We will solve the SE inside and outside the well, and then look at the border (the wall, $r = a$), and see what happens.

We are sticking to the notation of the original script, where $V_0 > 0$ and $V = -V_0$, regardless of what we think of it, just because it might be easier to compare these notes with the script that way.

Inside the well

Inside the potential well, the SE reads

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2}\right] R_\ell^{\text{in}}(r) = -\frac{2m}{\hbar^2}[E+V_0]R_\ell^{\text{in}}(r), \qquad E + V_0 > 0. \tag{1.24}$$

We can use the same substitution as before,

$$q = \frac{\sqrt{2m(E+V_0)}}{\hbar}, \tag{1.25}$$

and get the two families of solutions, $j_\ell(qr)$ and $n_\ell(qr)$. This region includes the point $r = 0$, so we cannot have solutions that diverge for $qr \to 0$: we can discard all the Neumann functions, which means that the physical solution inside the well is simply a multiple of j_ℓ:

$$R_\ell^{\text{in}}(r) = A_\ell \; j_\ell(qr). \tag{1.26}$$

We will determine the coefficient A_ℓ later.

Note: Let us go over this again, slowly. How can we be sure that the solution is a multiple of j_ℓ? The solution has to be a linear combination of solutions, and it cannot diverge. Now, j_ℓ and n_ℓ form an orthogonal basis, so any solution has the form $a\, j_\ell(qr) + b\, n_\ell(qr)$. If $|b| > 0$, the solution diverges, and if $b = 0$ the solution converges. Yes, we could expand the solution in a different basis (e.g. $c\, h_\ell(qr) d\, h_\ell*(qr)$), but then both basis elements (h_ℓ and $h_\ell*$) would diverge in different directions, so it would not be easy to find the coefficients... and we would end up with a multiple of j_ℓ anyway.

Outside the well

Outside, the potential is zero, and the SE becomes

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2}\right] R_\ell^{\text{out}}(r) = -\frac{2m}{\hbar^2}ER_\ell^{\text{out}}(r), \qquad E < 0. \tag{1.27}$$

Now we have a small problem. We cannot define $k = \sqrt{2mE}/\hbar$ if we want k to be real. Let us use instead

$$k = \frac{\sqrt{2m(-E)}}{\hbar}, \tag{1.28}$$

which turns the SE into

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1)}{r^2} - k^2\right] R_\ell^{\text{out}}(r) = 0. \tag{1.29}$$

Now, this is slightly different from our familiar Eq. 1.12. If we want to end up with something that looks like Eq. 1.15, we need to define the variable ρ to be

$$\rho := ikr, \qquad r = \frac{1}{ik}\rho, \qquad \frac{\partial^2}{\partial r^2} = \frac{-1}{k^2}\frac{\partial^2}{\partial \rho^2}, \tag{1.30}$$

and then we have

$$\left[\frac{ik}{\rho} - k^2 \frac{\partial^2}{\partial \rho^2}\frac{\rho}{ik} - k^2 \frac{\ell(\ell+1)}{\rho^2} - k^2\right] R_\ell^{\text{out}}(\rho) = 0 \tag{1.31}$$

$$\left[\frac{1}{\rho}\frac{\partial^2}{\partial \rho^2}\rho - \frac{\ell(\ell+1)}{\rho^2} + 1\right] R_\ell^{\text{out}}(\rho) = 0. \tag{1.32}$$

This is precisely the same equation as before, and therefore has the same (mathematical) solutions, $j_\ell(ikr)$ and $n_\ell(ikr)$. To pick the physical solution, let us look at the asymptotic behaviour of these functions when $r \to \infty$.

$$\sin x = \frac{-i}{2}\left(e^{ix} - e^{-ix}\right) \hookrightarrow$$

$$\begin{aligned} j_\ell(ikr) &\approx \frac{1}{ikr}\sin\left(ikr - \frac{\pi\ell}{2}\right) \\ &= \left(\frac{-i}{kr}\right)\left(\frac{-i}{2}\right)\left(e^{i\left(ikr-\frac{\pi\ell}{2}\right)} - e^{-i\left(ikr-\frac{\pi\ell}{2}\right)}\right) \\ &= \frac{-1}{2kr}\left(\underbrace{e^{-kr}}_{\to 0}e^{-i\frac{\pi\ell}{2}} - \underbrace{e^{kr}}_{\to\infty}e^{i\frac{\pi\ell}{2}}\right), \end{aligned}$$

which diverges. The same is true for $n_\ell(ikr)$,

$$\cos x = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \hookrightarrow$$

$$\begin{aligned} n_\ell(ikr) &\approx \frac{1}{ikr}\cos\left(ikr - \frac{\pi\ell}{2}\right) \\ &= \frac{i}{kr}\frac{1}{2}\left(e^{i\left(ikr-\frac{\pi\ell}{2}\right)} + e^{-i\left(ikr-\frac{\pi\ell}{2}\right)}\right) \\ &= \frac{i}{2kr}\left(\underbrace{e^{-kr}}_{\to 0}e^{-i\frac{\pi\ell}{2}} + \underbrace{e^{kr}}_{\to\infty}e^{i\frac{\pi\ell}{2}}\right). \end{aligned}$$

Fortunately for us, not all solutions diverge. We just have to look for a different basis in which to expand our general solution. For instance, we can try the Hankel functions,

$$h_\ell(ikr) \approx -\frac{i}{ikr}\, e^{i\left(ikr-\frac{\pi\ell}{2}\right)} = \frac{-1}{kr}\,\underbrace{e^{-kr}}_{\to 0}e^{i\frac{\pi\ell}{2}}, \tag{1.33}$$

$$h^*_\ell(ikr) \approx \frac{i}{ikr}\, e^{-i\left(ikr-\frac{\pi\ell}{2}\right)} = \frac{1}{kr}\,\underbrace{e^{kr}}_{\to\infty}e^{-i\frac{\pi\ell}{2}}. \tag{1.34}$$

This is great: $h^*_\ell(ikr)$ diverges, $h_\ell(ikr)$ does not, so the physical solution outside the well is just a multiple of the latter,

$$R^{\text{out}}_\ell(r) = B_\ell\; h_\ell(ikr). \tag{1.35}$$

Continuity at the wall

[Auf der Mauer, auf der Lauer sitzt 'ne kleine Welle.]

Now we can look for the coefficients A_ℓ and B_ℓ. Since $V_0 < 0$, we know three things:

1. The wave function is continuous, so

$$\begin{aligned} R^{\text{in}}_\ell(a) &= R^{\text{out}}_\ell(a) \Leftrightarrow \\ \Leftrightarrow A_\ell\; j_\ell(qa) &= B_\ell\; h_\ell(ika). \end{aligned} \tag{1.36}$$

   From here we will be able to get the ratio A_ℓ/B_ℓ.

2. The first derivative of the wave function is continuous, so

$$\begin{aligned} &\frac{\partial}{\partial r}R^{\text{in}}_\ell(r)\bigg|_{r=a} = \frac{\partial}{\partial r}R^{\text{out}}_\ell(r)\bigg|_{r=a} \Leftrightarrow \\ &\Leftrightarrow A_\ell q\; j'_\ell(qa) = ikB_\ell\; h'_\ell(ika). \end{aligned} \tag{1.37}$$

3. The wave function is normalized. This applies to the total wave function, $\Psi = \sum_{\ell m} c_{\ell m} R_\ell(a) Y_{\ell m}(\theta,\phi)$. Depending on the coefficients $\{c_{\ell m}\}$, we may fix A_ℓ, and get B_ℓ from the first condition.

From the first condition, we get

$$\begin{aligned} A_\ell\; j_\ell(qa) &= B_\ell\; h_\ell(ika) \\ \frac{A_\ell}{B_\ell} &= \frac{h_\ell(ika)}{j_\ell(qa)} \\ &= \left.\frac{-i\;(-1)^\ell(ikr)^\ell\left(\frac{1}{ikr}\frac{\partial}{\partial(ikr)}\right)^\ell\frac{e^{i(ikr)}}{ikr}}{(-1)^\ell(qr)^\ell\left(\frac{1}{qr}\frac{\partial}{\partial(qr)}\right)^\ell\frac{\sin(qr)}{qr}}\right|_{r=a}. \end{aligned}$$

Now, this might be doable, but it is more than a little ugly, and overall not that interesting. Anyway, as an example, let us solve it for $\ell = 0$,

$$\begin{aligned}\frac{A_0}{B_0} &= \left.\frac{-i\frac{e^{i(ikr)}}{ikr}}{\frac{\sin(qr)}{qr}}\right|_{r=a} \\ &= -\frac{q}{k}\frac{e^{-ka}}{\sin(qa)}.\end{aligned} \tag{1.38}$$

Now A_0 will depend on the coefficients $c_{\ell m}$. But there is something more interesting we can do with these continuity conditions.

[it would be cute to have a drawing of $|R_0(r)|$]

How deep is the well?

We saw that the wave function vanishes quickly outside the well (it goes approximately with e^{-kr}/r). This is because the energy of the particle is smaller than the potential outside the well ($E < V = 0$). The only place where the wave function is stable is inside the well, which is why we call it a *bound state*: the wave is bound to the well. It turns out that if the well is not deep enough, there will be no wave function anywhere: we will find a condition for the minimum depth the well such that if $V_0 < V_{\min} \Rightarrow A_0 = B_0 = 0 \Rightarrow R_0(r) = 0 \forall r$. This can be generalized for arbitrary ℓ, but you will get the idea from here.

First a few mathematical tricks. The factor $1/\rho$ in the expressions of $R_0^{\text{in}}(r)$ and $R_0^{\text{out}}(r)$ makes the derivatives a little annoying. It would be lovely to get rid of it, and here is how we do it. 😳 We want $R_0(r) \in C^1$ (i.e. both R_0 and its first derivative are continuous). In general, if we have two functions $f(r), g(r) \in C^1$, then $f(r)\ g(r) \in C^1$. Well, take $f(r) = r, g(r) = R_0(r)$, and the condition for continuity of R_0 implies

$$C^1 \ni r\ R_0(r) = \begin{cases} A_0\ \sin(qr), & r < a \\ B_0\ e^{-kr}, & r > a \end{cases}, \tag{1.39}$$

which is much more treatable. Now we have

$$A_0\ \sin(qa) = B_0\ e^{-ka} \quad \wedge \quad qA_0\ \cos(qa) = -kB_0\ e^{-ka}. \tag{1.40}$$

Dividing the latter by the former, we get

$$\frac{qA_0\ \cos(qa)}{A_0\ \sin(qa)} = -\frac{kB_0\ e^{-ka}}{B_0\ e^{-ka}} \quad \Rightarrow \tan(qa) = -\frac{q}{k} \quad \vee \quad A_0 = B_0 = 0. \tag{1.41}$$

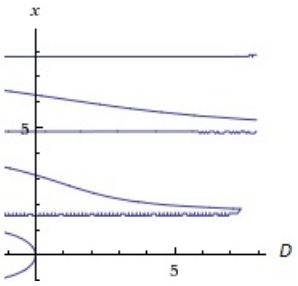
For now we leave the case $A_0 = B_0 = 0$, and try to solve the first case for V_0. Remember that k and q were given by

$$k = \frac{\sqrt{2m(-E)}}{\hbar}, \quad q = \frac{\sqrt{2m(E+V_0)}}{\hbar}, \tag{1.42}$$

so we have

$$\tan\left(\frac{a\sqrt{2m}}{\hbar}\sqrt{E+V_0}\right) = -\frac{\sqrt{2m(E+V_0)}}{\hbar}\frac{\hbar}{\sqrt{2m(-E)}} \tag{1.43}$$

$$= -\sqrt{\frac{-E}{E+V_0}}. \tag{1.44}$$



We can simplify this with two subsitutions, $C = \frac{a\sqrt{2m}}{\hbar} > 0$ and $x = C\sqrt{E+V_0} > 0$, to get

$$x \tan x = -\underbrace{C\sqrt{-E}}_{D>0} \tag{1.45}$$

The plot of solutions is on the left. If $x, D > 0$, there are no solutions for x below a threshold, where $x \tan x = 0$.

Let us find the smallest x for which there are solutions. We know that

$$(x \tan x < 0 \wedge x > 0) \Rightarrow x \in \left[2n\pi + \frac{\pi}{2}, 2n\pi + \pi\right] \cup \left[2n\pi + \frac{3\pi}{2}, 2n\pi\right], \quad 0 \leq n \in \mathbb{Z}, \tag{1.46}$$

The smallest n possible is zero, so we need $x > \pi/2$. Returning to the original variables, we have

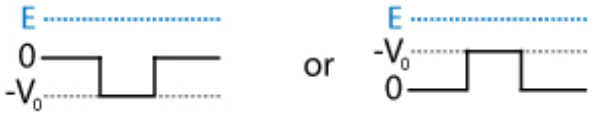
$$\frac{a\sqrt{2m}}{\hbar}\sqrt{E+V_0} > \frac{\pi}{2} \tag{1.47}$$

$$V_0 > \frac{\pi^2\hbar^2}{8a^2m} - E = \underbrace{\frac{\pi^2\hbar^2}{8a^2m}}_{V_{\min}} + |E|. \qquad (E<0) \tag{1.48}$$

In other words, if $V_0 < V_{\min} + |E|$, then we cannot have $\tan(qa) = -q/k$, and so, due to Eq. 1.41 we need $A_0 = B_0 = 0$, which means that $R_0(r) = 0, \forall r$.

1.3.3 Scattering potential

Now the particle has positive energy: the wave function is disturbed by the scattering centre, but should approximate a plave wave for large r.



$$V(r) = \begin{cases} -V_0, & r < a \\ 0, & r \geq a, \end{cases}$$

$$E > 0, E > -V_0.$$

Inside the well

As before,

$$R_\ell^{\text{in}}(r) = A_\ell j_\ell(qr), \quad q = \frac{\sqrt{2m(E+V_0)}}{\hbar} \tag{1.49}$$

Note that, again, $n_\ell(qr) \xrightarrow{r\to 0} -\infty$ and thus n_ℓ cannot appear as part of the wave function.

Outside the well

The general solution has the form

$$R_\ell(r) = B_\ell j_\ell(kr) + C_\ell n_\ell(kr), \quad k = \frac{\sqrt{2mE}}{\hbar} \tag{1.50}$$

Since in this case $r \to 0$ is not included in the domain, n_ℓ can be part of the wave function.

Our goal in this analysis is to find the constants A_ℓ, B_ℓ and C_ℓ. Let us start by looking at the ratio $\frac{C_\ell}{B_\ell}$.

Asymptotic behaviour $r \to \infty$ of $R_\ell(r)$

For large enough r, we can approximate

$$R_\ell(r) \approx \frac{B_\ell}{kr}\sin\left(kr - \frac{\ell\pi}{2}\right) - \frac{c}{kr}\cos\left(kr - \frac{\ell\pi}{2}\right) \tag{1.51}$$

$$= \frac{B}{kr}\left(\sin\left(kr - \frac{\ell\pi}{2}\right) - \frac{C}{B}\cos\left(kr - \frac{\ell\pi}{2}\right)\right) \tag{1.52}$$

We define

$$\tan\delta_\ell = \frac{C}{B} \tag{1.53}$$

With this definition, we have

$$R_\ell(r) = \frac{B}{kr}\left(\sin\left(kr - \frac{\ell\pi}{2}\right) + \tan\delta_\ell\cos\left(kr - \frac{\ell\pi}{2}\right)\right) \tag{1.54}$$

Using

$$\cos(x)\sin(y) + \sin(x)\cos(y) = \sin(x+y) \tag{1.55}$$

we finally get

$$R_\ell(r) = \frac{B}{kr}\frac{1}{\cos\delta_\ell}\sin\left(kr - \frac{\ell\pi}{2} + \delta_\ell\right). \tag{1.56}$$

Remember that this is only true for large r. We can find for instance δ_0,

$$\delta_0 = \tan^{-1}\left(\frac{k}{q}\tan(qa)\right) - ka. \tag{1.57}$$

[For further physical interpretation, check the original script and suggested books (I was getting a bit tired of this kind of problems when I got here).]

1.3.4 Hard sphere

Allow me to introduce the hard sphere, a crude approximation of solid objects. AA In what follows, keep in mind that not even a hard sphere (like a marble) is a hard sphere in the strict sense meant here.

The hard sphere consists of an infinitely high potential barrier,

$$V(r) = \begin{cases}\infty, & r \le a\\ 0, & r > a.\end{cases}$$

The wave function Ψ is thus 0 inside the sphere. On the outside, we can expand R_ℓ in the basis of the Hankel functions $h_\ell := h_\ell^{(1)}$ and $h_\ell^* := h_\ell^{(2)}$:

$$R_\ell(r) = Ah_\ell^*(kr) + Bh_\ell(kr), \qquad k = \frac{\sqrt{smE}}{\hbar} \tag{1.58}$$

$$= A\left[h_\ell^*(kr) + \frac{B}{A}h_\ell(kr)\right]. \tag{1.59}$$

For now we are only interested in the fraction $\frac{B}{A}$ between the two components. Up to a global phase and normalization (which will come later), we can write

$$R_\ell(r) \propto h_\ell^*(kr) + S_\ell(k)h_\ell(kr) \tag{1.60}$$

Now we can reuse the notion of probability flux $\mathbf{j}$, introduced in the beginning of the chapter (see original script). Since no particles are created or destroyed in the scattering process, the radial component $j_r(r) = 0 \forall r$. We have

$$\begin{aligned}
0 = j_r(r) &= \frac{\hbar}{2im}\left(R_\ell^*(r)\frac{\partial}{\partial r}R_\ell(r) - R_\ell(r)\frac{\partial}{\partial r}R_\ell^*(r)\right)\\
&= \frac{\hbar}{8im}\left(\left(h_\ell + S_\ell^*h_\ell^*\right)\left(h_\ell^{*\prime} + S_\ell h_\ell'\right) - \left(h_\ell^* + S_\ell^*h_\ell\right)\left(h_\ell' + S_\ell h_\ell^{*\prime}\right)\right)\\
&= \frac{\hbar}{8im}\left(h_\ell h_\ell^{*\prime} + h_\ell S_\ell h_\ell' + S_\ell^*h_\ell^*h_\ell^{*\prime} + |S_\ell|^2h_\ell^*h_\ell' - h_\ell^*h_\ell' - h_\ell^*S_\ell^*h_\ell^{*\prime} - S_\ell h_\ell h_\ell' - |S_\ell|^2h_\ell h_\ell^{*\prime}\right)\\
&= \frac{\hbar}{8im}\left(h_\ell h_\ell^{*\prime} - h_\ell^*h_\ell' + |S_\ell|^2h_\ell^*h_\ell' - |S_\ell|^2h_\ell h_\ell^{*\prime}\right)\\
&= \frac{\hbar}{8im}\left(h_\ell h_\ell^{*\prime} - h_\ell^*h_\ell'\right)\left(1 - |S_\ell|^2\right).
\end{aligned}$$

Since this has to hold for all r and h_ℓ is r-dependent, we can conclude that

$$|S_\ell|^2 = 1, \tag{1.61}$$

which means that S_ℓ is a phase,and therefore we can define δ_ℓ such that

$$S_\ell = e^{2i\delta_\ell}. \tag{1.62}$$

We will find δ_ℓ by looking at the asymptotic behaviour of $R_\ell(r)$.

Asymptotic behaviour of $R_\ell(r)$ for $r \to \infty$

Using

$$h_\ell(\rho) = -\frac{i}{\rho} e^{i\left(\rho - \frac{\ell\pi}{2}\right)} \qquad h_\ell^*(\rho) = -\frac{i}{\rho} e^{-i\left(\rho - \frac{\ell\pi}{2}\right)}, \tag{1.63}$$

we get

$$R_\ell(r) \propto \frac{i}{kr}\left(e^{-i\left(kr - \frac{\ell\pi}{2}\right)} - e^{2i\delta_\ell} e^{i\left(kr - \frac{\ell\pi}{2}\right)}\right) \tag{1.64}$$

$$= -\frac{i}{kr} e^{i\delta_\ell}\left(e^{i\delta_\ell} e^{i\left(kr - \frac{\ell\pi}{2}\right)}\right) \tag{1.65}$$

$$= -\frac{2e^{i\left(\delta_\ell - \frac{\pi}{2}\right)}}{kr} \sin\left(kr - \frac{\ell\pi}{2} + \delta_\ell\right) \tag{1.66}$$

[Check it yourself: normalization gives a global factor of $\frac{1}{2}$.] Inside the sphere the wave-function is zero, $R_\ell(r) = 0, r \leq a, \forall \ell$. At the border,

$$0 = R_\ell(a) \propto h_\ell^*(ka) + e^{2i\delta_\ell} h_\ell(ka) \tag{1.67}$$

and thus

$$e^{2i\delta_\ell} = -\frac{h_\ell^*(ka)}{h_\ell(ka)} \tag{1.68}$$

$$= -\frac{j_\ell(ka) - i n_\ell(ka)}{j_\ell(ka) + i n_\ell(ka)}. \tag{1.69}$$

We can express this relation using a the tangent [this was in one of the exercise series],

$$\tan\delta_\ell = \frac{j_\ell(ka)}{n_\ell(ka)} \tag{1.70}$$

For small energies, $k \to 0$, the Bessel and Neumann functions become

$$j_\ell(ka) \approx \frac{(ka)^\ell}{(2\ell+1)!!} \tag{1.71}$$

$$n_\ell(ka) \approx \frac{(2\ell-1)!!}{(ka)^{\ell+1}} \tag{1.72}$$

which leads to

$$\tan\delta_\ell \approx \frac{(ka)^{2\ell+1}}{(2\ell+1)[(2\ell-1)!!]^2}. \tag{1.73}$$

For $\ell \geq 1$, this decays rapidly. When computing the cross section σ, we have to sum over all ℓ. In a first approximation, this allows us to only consider $\ell = 0$. [More about this in the exercise series.]

1.3.5 Plane waves

We now want to take a closer look at plane waves ($V = 0$). For this purpose, we expand them in spherical coordinates (note that $Y_{\ell m}$ form a basis for functions on the sphere):

$$e^{i\mathbf{kx}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell m}(\mathbf{k}) j_\ell(kr) Y_{\ell m}(\theta, \phi) \tag{1.74}$$

Note that n_ℓ cannot be part of this expression since it would diverge at $r = 0$.
We will consider $\mathbf{k} \parallel \hat{\ell}_z$. In the exercises, the addition theorem will be proven which allows us to proof that it works for any $\mathbf{k}$.

Remember that

$$Y_{\ell m} = \sqrt{\frac{2\ell+1}{4\pi}} e^{im\phi} P_\ell^m(\cos\theta) \tag{1.75}$$

Since we are considering the rotationally symmetric case with $\mathbf{kx} = kx\cos\theta$, ϕ cannot be part of the expression. Thus only the $m = 0$ terms will remain. We use the simplified notation

$$P_\ell(x) := P_\ell^0(x) \tag{1.76}$$

Remember that the Legendre polynomials are given by

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(x^2-1)^\ell \tag{1.77}$$

and satisfie the orthogonality relation

$$\int_{-1}^{1}\mathrm{d}x P_\ell(x)P_{\ell'}(x) = \frac{2\delta_{\ell\ell'}}{2\ell+1} \tag{1.78}$$

Defining $A_\ell := c_{\ell 0}$, we thus have

$$e^{ikr\cos\theta} = \sum_{\ell'=0}^{\infty}\sqrt{\frac{2\ell'+1}{4\pi}}A_{\ell'}j_{\ell'}(kr)P_{\ell'}(\cos\theta) \tag{1.79}$$

To find $A_\ell j_\ell(kr)$, we project $e^{ikr\cos\theta}$ onto $P_\ell(\cos\theta)$, which basically means that we multiply both sides of the expression by $P_\ell(\cos\theta)$ and integrate over $\mathrm{d}\cos\theta$. Using the orthogonality relation (1.78) of the Legendre polynomials, we thus get (setting $z := \cos\theta$ in the integral)

$$A_\ell j_\ell(kr) = \frac{\sqrt{4\pi(2\ell+1)}}{2}\int_{-1}^{1}\mathrm{d}z P_\ell(z)e^{ikrz}. \tag{1.80}$$

Let us expand e^{ikrz} around $z = 0$ in a Taylor series:

$$e^{ikrz} = \sum_\ell \frac{(ikrz)^\ell}{\ell!} \tag{1.81}$$

Looking at the leading order term in the Legendre Polynomial from (1.77)

$$P_\ell(z) = \frac{(2\ell)!}{2^\ell(\ell!)^2}z^\ell + O\left(z^{\ell-1}\right) \tag{1.82}$$

which, after solving for z^ℓ and plugging it back into the expansion for e^{ikrz} we get

$$A_\ell j_\ell(kr) = \frac{\sqrt{4\pi(2\ell+1)}}{2}\int_{-1}^{1}\mathrm{d}z P_\ell(z)\left(\sum_{\ell'}\frac{(2\ell)!}{2^\ell\ell!}(ikr)^\ell - O\left((kr)^{\ell-1}\right)\right) \tag{1.83}$$

Using again the orthogonality relation of the Legendre polynomials (1.78), we then find

$$A_\ell j_\ell(kr) = \frac{\sqrt{4\pi(2\ell+1)}}{2}(ikr)^\ell\frac{2^{\ell+1}\ell!}{(2\ell+1)!} + O\left((kr)^{\ell-1}\right) \tag{1.84}$$

Looking at small kr and using

$$j_\ell(kr) \approx \frac{2^\ell\ell!}{(2\ell+1)!}(kr)^\ell \tag{1.85}$$

we finally find

$$A_\ell = i^\ell\sqrt{4\pi(2\ell+1)} \tag{1.86}$$

which gives us our low energy expression for plane waves

$$e^{ikr\cos\theta} = \sum_{\ell=0}^{\infty} i^\ell(2\ell+1)j_\ell(kr)P_\ell(\cos\theta). \tag{1.87}$$

Chapter 2

Interaction between light and matter

2.6 Light scattering reloaded

Our setting is simply a particle in an EM field (light). We want to find out how light is scattered by the paricle. If you shoot a light beam in the direction of the particle, will it change direction? Colour? Will it excite the particle? What is the probability of each process? So far (in previous lectures and original script) we looked at one particular scattering process. We will now look at few more possibilities. First, recap.

Structure of the Hilbert space

The Hilbert space of the particle is $\mathcal{H}_p$ the one of the EM field is $\mathcal{H}_{EM}$. The global Hilbert space is, naturally,

$$\mathcal{H} = \mathcal{H}_p \otimes \mathcal{H}_{EM}. \tag{2.1}$$

We quantized the EM field. Fist, we drew an imaginary cubic box of side L (Fig. 2.1). Then we imposed boundary conditions: we said that wave functions had to be zero at the borders of the box. We can do this by making our box larger than the universe (and assuming that nothing exists outside the universe).

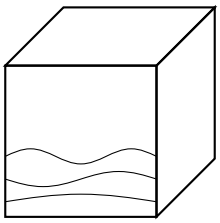


Figure 2.1: Giant box in real space, and some of the possible waves that fulfil the boundary conditions from Eq. 2.2.

So now we have the conditions

$$\psi^{\mathbf{k}}_{EM}(0,0,0) = \psi^{\mathbf{k}}_{EM}(L,y,z) = \psi^{\mathbf{k}}_{EM}(x,L,z) = \psi^{\mathbf{k}}_{EM}(x,y,L) = 0. \tag{2.2}$$

This implies that the wave function of the EM field is a linear combination of plane waves with $\mathbf{k}$ vectors of the form

$$\mathbf{k} = \frac{2\pi}{L}\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}, \quad n_i \in \mathbb{Z}. \tag{2.3}$$

The allowed states in momentum space form an infinite lattice of spacing $\frac{2\pi}{L}$. Photon's polarization, λ, is independent of momentum. We can write the Hilbert space of the EM field as a tensor product of the spaces of all the allowed momenta and polarizarion, $\mathcal{H}_{EM} = \bigotimes_{\mathbf{k},\lambda} \mathcal{H}_{\mathbf{k},\lambda}$.

To describe a particular EM field (i.e., a particular state of this Hilbert space), we populate the allowed sites in momentum space with actual photons. We denote the vacuum (a state where there are no photons) by $|0\rangle$, and then act on this state with creation operators. For instance, $\hat{a}^\dagger_{\mathbf{k}\lambda}|0\rangle$ is a state with exactly one photon with momentum $\mathbf{k}$ and polarization λ. Here is another example: we have an initial field $|\psi\rangle$ with lots of photons. Now it suffers a proccess that should change the momentum of a photon with polarization λ from $\mathbf{k}$ to $\mathbf{k}'$. The new state is $\hat{a}^\dagger_{\mathbf{k}'\lambda}\hat{a}_{\mathbf{k}\lambda}|\psi\rangle$. First we have an anihilation operator that destroys the photon with momentum $\mathbf{k}$, and we create a photon with momentum $\mathbf{k}'$.

The anihilation operator $\hat{a}_{\mathbf{k}\lambda}$ acts only on the Hilbert space $\mathcal{H}_{k,\lambda}$; explicitly, we have

$$\hat{a}_{\mathbf{k},\lambda} = \mathbb{1}_{\mathbf{k}_1,\lambda_1} \otimes \mathbb{1}_{\mathbf{k}_1,\lambda_2} \otimes \cdots \otimes \hat{a}_{\mathbf{k},\lambda} \otimes \mathbb{1} \otimes \dots \ . \tag{2.4}$$

These operators, sometimes referred to as ladder operators, commute as

$$[\hat{a}_{\mathbf{k},\lambda}, \hat{a}^\dagger_{\mathbf{k}',\lambda'}] = \delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda,\lambda'}, \tag{2.5}$$

and are otherwise characterized early in the original script. [If you are not comfortable with these operators, go back to where they are introduced and solve some of the earlier exercise series until you feel you understand them well.] Note that this way of organizing the Hilbert space of the EM field is very different from the way we describe the particle's space. There we act on the Hilbert space with operators like $\hat{r}$ and $\hat{p}$, whereas here everything is done in terms of ladder operators. For instance, to measure the average energy of the field, we define the operator

$$\hat{E} = \sum_{\mathbf{k}\lambda} E_k \ \hat{a}^\dagger_{\mathbf{k},\lambda}\hat{a}_{\mathbf{k},\lambda}, \qquad \langle E\rangle_\psi = \langle\psi|\hat{E}|\psi\rangle, \qquad E_k = \hbar k. \tag{2.6}$$

This operator simply counts the number of photons in each *mode* (a pair $(\mathbf{k},\lambda)$), and multiplies it by the mode's energy. Compare this with the usual way of computing the average energy of a free particle in state $|\phi\rangle$,

$$\langle E\rangle_\phi = \langle\phi|\frac{\hat{p}^2}{2m}|\phi\rangle.$$

There we have to act with the operator $\hat{p}$ (effectively a gradient) on the wave-function, which is usually harder than counting photons in a given mode.

We denote the different states of the particle by $|0\rangle, |1\rangle, \dots, |n\rangle$, etc., in increasing order of energy. We denote the states of the EM field by the number of photons of each mode, like this: $|N_{\mathbf{k}_1,\lambda_1}, \dots, N_{\mathbf{k},\lambda}, \dots\rangle$. Sometimes we are only interested in one or two modes, for instance, if we are looking for processes that change the number of photons in those modes (like photon absorption or deflection). In that case we may label the state only by the relevant modes: for instance, if a process takes one photon from mode $(\mathbf{k},\lambda)$ to mode $(\mathbf{k}',\lambda')$, we can represent this as $|N_{\mathbf{k},\lambda}, N_{\mathbf{k}',\lambda'}\rangle \longrightarrow |N_{\mathbf{k},\lambda} - 1, N_{\mathbf{k}',\lambda'} + 1\rangle$. All the other modes are implicitly left unchanged.

Later on (Chapter 4) we will express the Hilbert space of particles in this new formalism too (second quantization). For now, here is a table with some characteristics of the two spaces.

	Particle	EM field
Hilbert space	$\mathcal{H}_p$	$\mathcal{H}_{EM} = \bigotimes_{\mathbf{k},\lambda} \mathcal{H}_{\mathbf{k},\lambda}$
Operators	$\hat{r}, \hat{p}$	$\hat{a}_{\mathbf{k},\lambda},\ \hat{a}^\dagger_{\mathbf{k},\lambda}$, $\hat{N}_{\mathbf{k},\lambda} = \hat{a}^\dagger_{\mathbf{k},\lambda}\hat{a}_{\mathbf{k},\lambda}$
Typical states	$\lvert n\rangle$ (labelled by energy)	$\lvert N_{\mathbf{k}_1,\lambda_1}, \dots, N_{\mathbf{k},\lambda}, \dots\rangle$ (labelled by # photons/mode)

Global Hamiltonian

The total Hamiltonian was derived earlier in the original script. It can be written as

$$\hat{H} := \underbrace{\hat{H}_p}_{\in \mathrm{End}(\mathcal{H}_p)} + \underbrace{\hat{H}_{EM}}_{\in \mathrm{End}(\mathcal{H}_{EM})} + \underbrace{\hat{H}_{int}}_{\in \mathrm{End}(\mathcal{H}_p \otimes \mathcal{H}_{EM})}, \tag{2.7}$$

where the particle Hamiltonian is

$$\hat{H}_p := \underbrace{\frac{\hat{p}^2}{2m}}_{\text{free particle}} + \underbrace{\hat{V}_0(\hat{\mathbf{r}})}_{\text{e.g. atom}}, \tag{2.8}$$

and the free-field Hamiltonian is

$$\hat{H}_{EM} := \frac{1}{8\pi}\int (\hat{\mathbf{E}}^2 + \hat{\mathbf{B}}^2) d^3\mathbf{r} \ = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}}(\hat{N}_{\mathbf{k},\lambda} + \frac{1}{2}) \tag{2.9}$$

(see for instance Series 5, exercise 1). The interaction term that couples the particle and the radiation is again given by three terms,

$$\hat{H}_{int} = \hat{H}_{int,1} + \hat{H}_{int,2} + \hat{H}_{int,3},$$

$$\hat{H}_{int,1} = -\int \frac{e}{c}\, \hat{\mathbf{j}}(\hat{\mathbf{r}})\ \hat{\mathbf{A}}(\hat{\mathbf{r}})\ d^3\mathbf{r} \tag{2.10}$$

$$\hat{H}_{int,2} = \int \frac{e}{2mc^2}\ \hat{\mathbf{A}}^2(\hat{\mathbf{r}})\ \hat{\rho}(\hat{\mathbf{r}})\ d^3\mathbf{r} \tag{2.11}$$

$$\hat{H}_{int,3} = \int e\ \hat{\rho}(\hat{\mathbf{r}})\ \phi(\hat{\mathbf{r}}, t)\ d^3\mathbf{r}. \tag{2.12}$$

We will only treat the case $\phi = 0$, so $\hat{H}_{int,3}$.

Goal

In general we are interested in scattering processes that excite the particle and change the EM field in some way,

$$\underbrace{|0\rangle_p \otimes |N_{\mathbf{k}_1,\lambda_1}, \ldots, N_{\mathbf{k},\lambda}, \ldots\rangle_{EM}}_{=:|i\rangle} \implies \underbrace{|n\rangle_p \otimes |N'_{\mathbf{k}_1,\lambda_1}, \ldots, N'_{\mathbf{k},\lambda}, \ldots\rangle_{EM}}_{=:|f\rangle}. \tag{2.13}$$

We want to know which of these transitions processes are allowed, and their probability of hapenning. Ideally we would solve the SE for the global system, but that is too hard with this complex Hamiltonian. We can however find the *most likely* transitions if we apply perturbation theory to each of the interaction terms of the Hamiltonian. Our plan is the following:

1. quantize the interaction terms of the Hamiltonian;
2. apply perturbation theory to first order to each term;
3. apply perturbation theory to second order to the first term.

Why don't we apply second order to all the terms? 😳 As far as I can see, it is because it would be a long, tedious calculation that would not bring us any new transitions or insights.

2.6.1 Perturbation theory to first order

How does it work?

Perturbation theory to first order results in Fermi's Golden Rule. There is a summary of the derivation of the rule in the QM1 OS, pages 127–130. Short summary: we write the SE as a function of the interaction Hamiltonian, H_{int}, and Taylor-expand it on time. Then we consider only the first-order terms and compute the probability of a transition from state $|i\rangle$ to state $|f\rangle$. This is called Fermi's Golden Rule, and the probability is given by

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar} |\langle f|H_{int}|i\rangle|^2\, \rho(E_f), \tag{2.14}$$

where $\rho(E_f)$ is the density of states of the final state (usually something like $\delta(E_f - E_i)$). We will explain how perturbation theory works in detail in Section 2.6.2, when we introduce the second-order terms.

First interaction term

We quantized the of the interaction Hamiltonian in previous lectures (Eq. 2.4.17 of the OS),

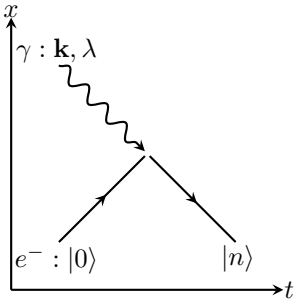
$$\hat{H}_{int,1} = -\frac{e}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar}{\omega_{\mathbf{k}}}} \left(\hat{\mathbf{j}}(-\mathbf{k}) \cdot \mathbf{e}(\mathbf{k},\lambda) \otimes \hat{a}_{\mathbf{k},\lambda} + \hat{\mathbf{j}}(\mathbf{k}) \cdot \mathbf{e}(\mathbf{k},\lambda)^* \otimes \hat{a}^\dagger_{\mathbf{k},\lambda}\right) \tag{2.15}$$

We applied Fermi's Golden Rule and found the transition probability $\Gamma_{i\to f}$ for the transition

$$|0\rangle_p \otimes |N_{\mathbf{k},\lambda}\rangle_{EM} \implies |n\rangle_p \otimes |N_{\mathbf{k},\lambda} - 1\rangle_{EM}. \tag{2.16}$$

What is happening in this transition? The particle is absorbing one photon in mode $(\mathbf{k}, \lambda)$ and using its energy to go to the excited state $|n\rangle$ (see Fig. 2.2a). This process can be represented as a Feynman diagram (introduced in Fig. 2.2b).

(a) Left (before the interaction) , we have a photon about to hit the photon "hits" the particle. On the right, the photon was absorbed by the particle, which is now in an excited state.



(b) Photon absorption described as a Feynman diagram. Time flows from left to right. The vertical axis differentiates between the two Hilbert spaces: EM field on top, particle on bottom. The wobbly line represents the photon, and the straight line the particle (an electron). [Note: sometimes Feynman diagrams are rotated by 90°(e.g., in the script).]

Figure 2.2: A particle absorbs a single photon in mode $(\mathbf{k}, \lambda)$, going from state $|0\rangle$ to $|n\rangle$. Here are two different pictorial representations of the transition.

Second interaction term

Now we are going to write the second interaction term of the Hamiltonian in our new formalism. The original form of this term is

$$\hat{H}_{int,2} = \frac{e}{2mc^2} \int d^3\mathbf{r}\ \hat{\rho}(\hat{\mathbf{r}})\ \hat{\mathbf{A}}^2(\hat{\mathbf{r}}) = \frac{e}{2mc^2} \int d^3\mathbf{r}\ \hat{\rho}(\hat{\mathbf{r}}) \otimes \hat{\mathbf{A}}^2(\hat{\mathbf{r}}),$$

because the operator $\hat{\mathbf{A}}(\mathbf{r}) = \hat{\mathbb{1}}_p \otimes \hat{\mathbf{A}}(\mathbf{r})_{EM}$ acts only on $\mathcal{H}_{EM}$, and $\hat{\rho}(\mathbf{r}) = \hat{\rho}(\mathbf{r})_p \otimes \hat{\mathbb{1}}_{EM}$ acts only on $\mathcal{H}_p$.

Remember that the vector potential $\hat{\mathbf{A}}$ is defined as

$$\hat{\mathbf{A}}(\mathbf{r})_{EM} := \frac{1}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar c^2}{\omega_\mathbf{k}}} \left(\hat{a}_{\mathbf{k},\lambda} \mathbf{e}(\mathbf{k},\lambda) e^{i\mathbf{k}\cdot\mathbf{r}} + \hat{a}^\dagger_{\mathbf{k},\lambda} \mathbf{e}^*(\mathbf{k},\lambda) e^{-i\mathbf{k}\cdot\mathbf{r}} \right). \tag{2.17}$$

We saw earlier in the OS that this definition satisfies $\hat{\mathbf{B}} = \nabla \wedge \hat{\mathbf{A}}$. Squaring the vector potential, we obtain

$$\hat{\mathbf{A}}^2(\mathbf{r}) = \frac{2\pi\hbar c^2}{L^3} \sum_{\mathbf{k},\lambda,\mathbf{k}',\lambda'} \frac{1}{\sqrt{\omega_\mathbf{k}\omega_{\mathbf{k}'}}} \left(\hat{a}_{\mathbf{k},\lambda} \mathbf{e}(\mathbf{k},\lambda) e^{i\mathbf{k}\cdot\mathbf{r}} + h.c. \right) \left(\hat{a}_{\mathbf{k}',\lambda'} \mathbf{e}(\mathbf{k}',\lambda') e^{i\mathbf{k}'\cdot\mathbf{r}} + h.c. \right). \tag{2.18}$$

Now we can write the second interaction term as

$$\hat{H}_{int,2} = \frac{e}{2mc^2} \frac{2\pi\hbar c^2}{L^3} \int d^3\mathbf{r} \hat{\rho}_p(\hat{\mathbf{r}}) \otimes \sum_{\mathbf{k},\lambda,\mathbf{k}',\lambda'} \frac{1}{\sqrt{\omega_\mathbf{k}\omega_{\mathbf{k}'}}} \left(\hat{a}_{\mathbf{k},\lambda} \mathbf{e}(\mathbf{k},\lambda) e^{i\mathbf{k}\cdot\mathbf{r}} + h.c. \right) \left(\hat{a}_{\mathbf{k}',\lambda'} \mathbf{e}(\mathbf{k}',\lambda') e^{i\mathbf{k}'\cdot\mathbf{r}} + h.c. \right). \tag{2.19}$$

We are ready to apply Fermi's Golden Rule. But what initial and final states should we consider? 😳 Right now it is not obvious, but the calculations ahead will show us that this interaction term leads to a transtion where the particle absorbs a photon, goes to an excited state, and emits a different photon (see Figs. 2.3a and 2.3b),

$$|0\rangle_p \otimes |\ldots, N_{\mathbf{k},\lambda}, \ldots, N_{\mathbf{k}',\lambda'} = 0, \ldots\rangle_{EM} \implies |n\rangle_p \otimes |\ldots, N_{\mathbf{k},\lambda} - 1, \ldots, N_{\mathbf{k}',\lambda'} = 1, \ldots\rangle_{EM}. \tag{2.20}$$

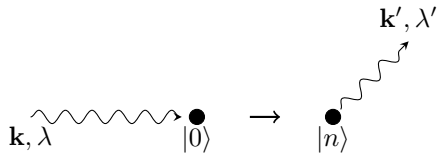
Initial state: particle in ground state, no photons in mode $(\mathbf{k}', \lambda')$ [for simplicity], $N_{\mathbf{k},\lambda}$ photons in mode $(\mathbf{k}, \lambda)$ and whatever in other modes. Final state: particle in excited state $|n\rangle$, one photon in mode $(\mathbf{k}', \lambda')$ and one less photon in mode $(\mathbf{k}, \lambda)$. The other modes are not affected. In short notation,

$$|i\rangle = |0\rangle_p \otimes |\underbrace{N_{\mathbf{k},\lambda}}_{(\mathbf{k},\lambda)}, \underbrace{0}_{(\mathbf{k}',\lambda')}\rangle, \qquad |f\rangle = |n\rangle_p \otimes |\underbrace{N_{\mathbf{k},\lambda} - 1}_{(\mathbf{k},\lambda)}, \underbrace{1}_{(\mathbf{k}',\lambda')}\rangle. \tag{2.21}$$

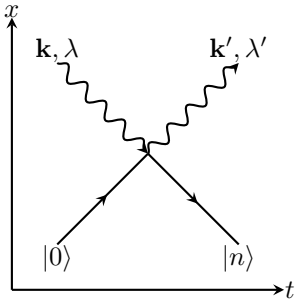
We will see for what values of n, $\mathbf{k}$ and $\mathbf{k}'$ this transition is allowed, and with what probability.

The matrix element from Fermi's Golden Rule (Eq. 2.14) comes

$$\begin{aligned} \langle f|\hat{H}_{int,2}|i\rangle &= \langle n|\langle N_{\mathbf{k},\lambda} - 1, 1| \left(\int d^3\mathbf{r} \frac{e}{2mc^2}\ \hat{\rho}(\hat{\mathbf{r}}) \otimes \hat{\mathbf{A}}^2(\hat{\mathbf{r}}) \right) |0\rangle |N_{\mathbf{k},\lambda}, 0\rangle \\ &= \frac{e}{2mc^2} \int d^3\mathbf{r} \langle n|\hat{\rho}(\hat{\mathbf{r}})|0\rangle\ \langle N_{\mathbf{k},\lambda} - 1, 1|\hat{\mathbf{A}}^2(\hat{\mathbf{r}})|N_{\mathbf{k},\lambda}, 0\rangle. \end{aligned} \tag{2.22}$$



(a) On the left (before), the particle absorbs a photon. On the right (after) the particle is in a excited state and emitted a photon.



(b) The same interaction described as a Feynman diagram. Above, $\mathcal{H}_{EM}$; below, $\mathcal{H}_p$.

Figure 2.3: A particle, originally in state $|0\rangle$, absorbs a photon with momentum $\mathbf{k}$. It goes to excited state $|n\rangle$ and emits a photon with energy $\varepsilon_{k'} = E(n) - E(0) - \varepsilon_k$.

The last term inside that integral expands as

$$\begin{aligned}\langle N_{\mathbf{k},\lambda} - 1, 1|\hat{\mathbf{A}}^2(\hat{\mathbf{r}})|N_{\mathbf{k},\lambda}, 0\rangle \\ = \frac{2\pi\hbar c^2}{L^3}\langle N_{\mathbf{k},\lambda} - 1, 1| \sum_{\mathbf{q},\mu,\mathbf{q}',\mu'} \frac{1}{\sqrt{\omega_{\mathbf{q}}\omega_{\mathbf{q}'}}} \quad &\left(\hat{a}_{\mathbf{q},\mu}\mathbf{e}(\mathbf{q},\mu)e^{i\mathbf{q}\cdot\mathbf{r}} + \hat{a}^\dagger_{\mathbf{q},\mu}\mathbf{e}^*(\mathbf{q},\mu)e^{-i\mathbf{q}\cdot\mathbf{r}}\right) \\ &\left(\hat{a}_{\mathbf{q}',\mu'}\mathbf{e}(\mathbf{q}',\mu')e^{i\mathbf{q}'\cdot\mathbf{r}} + \hat{a}^\dagger_{\mathbf{q}',\mu'}\mathbf{e}^*(\mathbf{q}',\mu')e^{-i\mathbf{q}'\cdot\mathbf{r}}\right)|N_{\mathbf{k},\lambda}, 0\rangle.\end{aligned}$$

You can check that only two terms of that sum are non-vanishing: those with operators $\hat{a}_{\mathbf{k},\lambda}\hat{a}^\dagger_{\mathbf{k}',\lambda'}$ and $\hat{a}^\dagger_{\mathbf{k}',\lambda'}\hat{a}_{\mathbf{k},\lambda}$ (in other words, operators that create a photon in mode $(\mathbf{k}',\lambda')$ and destroy one in mode $(\mathbf{k},\lambda)$). In the end, we are left with

$$\langle N_{\mathbf{k},\lambda} - 1, 1|\hat{\mathbf{A}}^2(\hat{\mathbf{r}})|N_{\mathbf{k},\lambda}, 0\rangle = \frac{1}{L^3}\frac{2\pi\hbar c^2}{\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\, 2\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\,\sqrt{N_{\mathbf{k},\lambda}}.$$

Plugging this in the matrix element of Fermi's Golden Rule, we obtain

$$\begin{aligned}\langle f|\hat{H}_{int,2}|i\rangle &= \underbrace{r_0}_{\frac{e^2}{mc^2}} \frac{2\pi\hbar c^2}{L^3\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\,\sqrt{N_{\mathbf{k},\lambda}}\,\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')\,\langle n|\left(\int d^3\mathbf{r}\;\hat{\rho}(\mathbf{r})\;e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\right)|0\rangle \\ &= \mathbf{r}_0\frac{2\pi\hbar c^2}{L^3\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\,\sqrt{N_{\mathbf{k},\lambda}}\,\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')\,\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle. \qquad (2.23)\end{aligned}$$

Here, r_0 is the constant historically known as the electron radius, and $\hat{\rho}_{\mathbf{k}'-\mathbf{k}}$ is the Fourier transform of $\hat{\rho}(\mathbf{r})$. Finally, we use the Golden Rule (Eq. 2.14) to obtain the transition probability,

$$\begin{aligned}\Gamma_{0\to n,(\mathbf{k},\lambda)\to(\mathbf{k}',\lambda')} &= \frac{2\pi}{\hbar}\,|\langle f|H_{int,2}|i\rangle|^2\,\rho(\varepsilon_f) \\ &= \frac{2\pi}{\hbar}\left(\frac{2\pi\hbar c^2\mathbf{r}_0}{L^3\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\right)^2 N_{\mathbf{k},\lambda}\,|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|^2\,|\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle|^2\,\delta(E_n + \hbar\omega_{\mathbf{k}'} - (E_0 + \hbar\omega_{\mathbf{k}})).\end{aligned} \qquad (2.24)$$

This is Eq. 2.6.7 in the OS.

Now we want to relate Γ, the transition probability, to something that can be measured in the laboratory, as to check the accuracy of this approximation. One good candidate is the cross-section (Fig. 2.4). This measures the percentage of emitted photons that are detected at a given angle,

$$\frac{d\sigma}{d\Omega_{\mathbf{k}'}} = \frac{1}{j_{\text{inc}}}\sum_n \sum_{\mathbf{k}'\in d\Omega_{\mathbf{k}'}} \Gamma_{0\to n,(\mathbf{k},\lambda)\to(\mathbf{k}',\lambda')}. \qquad (2.25)$$

Here, j_{inc} is the flux of incoming photons: the photons that we send in mode $(\mathbf{k},\lambda)$ to excite the particle. We have $N_{\mathbf{k},\lambda}$ photons moving at the speed of light in a box of side L, so the flux is $j_{\text{inc}} = \frac{cN_{\mathbf{k}}}{L^3}$. We sum over all the photons with momentum compatible with the small section $d\Omega_{\mathbf{k}'}$ (photons going in the same direction but with different $|\mathbf{k}|$). Assume that the detectors know which λ' hit them.

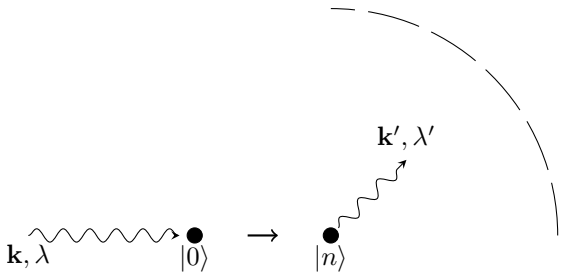


Figure 2.4: We place many small detectors in a circle around the particle. The photon emitted with momentum $\mathbf{k}'$ will hit the detector at section $d\Omega_{\mathbf{k}'}$. We repeat the experiment many times, and find the statistics of photons detected for each angle.

Using Eq. 2.24, we obtain

$$\begin{aligned}\frac{d\sigma}{d\Omega_{\mathbf{k}'}} &= \sum_n \frac{L^3}{cN_{\mathbf{k}}} \sum_{\mathbf{k}' \in d\Omega_{\mathbf{k}'}} \Gamma_{0\to n,(\mathbf{k},\lambda)\to(\mathbf{k}',\lambda')} \\ &= \sum_n \frac{8\pi^3 \hbar r_0^2 c^3}{L^3 \omega_{\mathbf{k}}} \sum_{\mathbf{k}' \in d\Omega_{\mathbf{k}'}} \frac{1}{\omega_{\mathbf{k}'}} \left|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')\right|^2 \left|\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle\right|^2 \delta(E_n + \hbar\omega_{\mathbf{k}'} - (E_0 + \hbar\omega_{\mathbf{k}}))\end{aligned}$$

We need to compute the sum over all $\mathbf{k}' \in d\Omega_{\mathbf{k}'}$. AA We will approximate the sum (in general discrete values) by an integral [Riemann sum from calculus]

$$\sum_{\mathbf{k}'} f(\mathbf{k}') \approx \left(\frac{L}{2\pi}\right)^3 \int d^3\mathbf{k}' f(\mathbf{k}'). \tag{2.26}$$

Remember that $\mathbf{k}' = \frac{2\pi}{L}(n_x, n_y, n_z)$. For large L, neighbouring $\mathbf{k}'$s are very close together, so this approximation is reasonable. Now we can transform the Cartesian coordinates to spherical coordinates, easier to relate to our section,

$$d^3\mathbf{k}' = |\mathbf{k}'|^2 \, d\Omega_{\mathbf{k}'} d\,|\mathbf{k}'| = \frac{\omega_{\mathbf{k}'}^2}{c^3} d\Omega_{\mathbf{k}'} d\omega_{\mathbf{k}'},$$

so we have

$$\sum_{\mathbf{k}' \in d\Omega_{\mathbf{k}'}} f(\mathbf{k}') \approx \frac{L^3}{8\pi^3 c^3} \int_0^\infty d\omega_{\mathbf{k}'} \; \omega_{\mathbf{k}'}^2 \; f(\mathbf{k}'). \tag{2.27}$$

Applying this to the cross section, we obtain

$$\begin{aligned}\frac{d\sigma}{d\Omega_{\mathbf{k}'}} &= \sum_n \frac{8\pi^3 \hbar r_0^2 c^3}{L^3 \omega_{\mathbf{k}}} \; \frac{L^3}{8\pi^3 c^3} \int_0^\infty d\omega_{\mathbf{k}'} \; \omega_{\mathbf{k}'}^2 \; \frac{1}{\omega_{\mathbf{k}'}} \; \left|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')\right|^2 \left|\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle\right|^2 \underbrace{\delta\left(E_n + \hbar\omega_{\mathbf{k}'} - E_0 - \hbar\omega_{\mathbf{k}}\right)}_{\frac{1}{\hbar}\delta\left(\omega_{\mathbf{k}'} + \frac{E_n - E_0}{\hbar} - \omega_{\mathbf{k}}\right)} \\ &= \sum_n \frac{\hbar r_0^2}{\omega_{\mathbf{k}}} \int_0^\infty d\omega_{\mathbf{k}'} \; \omega_{\mathbf{k}'} \; \left|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')\right|^2 \; \left|\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle\right|^2 \; \frac{1}{\hbar} \underbrace{\delta\left(\omega_{\mathbf{k}'} + \frac{E_n - E_0}{\hbar} - \omega_{\mathbf{k}}\right)}_{\text{delta function}} \\ &= \sum_n r_0^2 \; \frac{\omega_{\mathbf{k}'}}{\omega_{\mathbf{k}}} \; \left|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')\right|^2 \; \left|\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle\right|^2 \; \underbrace{\delta_{\hbar\omega_{\mathbf{k}'}+E_n, E_0+\hbar\omega_{\mathbf{k}}}}_{\text{Kronecker delta}}.\end{aligned}$$

We are almost done. All that is left to calculate is the term $|\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle|^2$. This depends on the properties of the particle, namely on states $\{|n\rangle_p\}$. We will compute it for a specific case.

Example Suppose that the particle is a free electron, such that each state is described by a plane wave, $|n\rangle = \frac{1}{\sqrt{L^3}} e^{i\mathbf{q_n}\mathbf{r}}$ AA . We know that $\hat{\rho}(\mathbf{r}) = \delta(\mathbf{r} - \hat{\mathbf{r}})$. The Fourier transform of this delta function is

$$\hat{\rho}_{\mathbf{k}'-\mathbf{k}} = \int_{L^3} d^3\mathbf{r} \delta(\mathbf{r} - \hat{\mathbf{r}}) e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} = e^{-i(\mathbf{k}'-\mathbf{k})\cdot\hat{\mathbf{r}}}. \tag{2.28}$$

The matrix element $\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle$ comes

$$\begin{aligned}\langle n|\hat{\rho}_{\mathbf{k}'-\mathbf{k}}|0\rangle &= \frac{1}{L^3}\int_{L^3} d^3\mathbf{r}\; e^{-i\mathbf{q_n r}}\; e^{i(\mathbf{k}-\mathbf{k}')\hat{\mathbf{r}}}\; e^{i\mathbf{q_0 r}} \\ &= \frac{1}{L^3}\int_{L^3} d^3\mathbf{r} e^{-i\mathbf{r}(\mathbf{q_0}+\mathbf{k}-\mathbf{q_n}-\mathbf{k}')} \\ &= \delta_{\mathbf{k}+\mathbf{q_0},\mathbf{k}'+\mathbf{q_n}}.\end{aligned}$$

This corresponds to momentum conservation (Eq. 2.6.9 in the OS). Finally, the cross-section comes

$$\begin{aligned}\frac{d\sigma_{\mathbf{q_0}\to\mathbf{q_n}}}{d\Omega_{\mathbf{k}'}} &= \sum_n r_0^2\;\frac{\omega_{\mathbf{k}'}}{\omega_{\mathbf{k}}}\;|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|^2\;\delta_{\hbar\omega_{\mathbf{k}'}+E_n,-E_0+\hbar\omega_{\mathbf{k}}}\;\delta_{\mathbf{k}+\mathbf{q_0},\mathbf{k}'+\mathbf{q_n}} \\ &= r_0^2\;\frac{\omega_{\mathbf{k}'}}{\omega_{\mathbf{k}}}\;|\mathbf{e}(\mathbf{k},\lambda)\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|^2. \qquad (2.29)\end{aligned}$$

2.6.2 Perturbation theory to second order

How does it work?

What follows is a general description of perturbation theory to second order in time. Later we will apply this to our concrete problem. A good reference (in German) is Manfred Sigrist's QM2 lecture notes, pages 117–124.

We start from a time-dependent Hamiltonian

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t). \qquad (2.30)$$

Here, $\hat{H}_0$ is the initial time-independent Hamiltonian (imagine something simple, like a free particle or a hydrogen atom). $\hat{V}(t)$ is a small perturbation, compared with $\hat{H}_0$ (e.g. the EM-field), which is "swiched on" at time $t = t_0$ and off again at time $t = \tau$,

$$\hat{V}(t) = \Theta(t_0)[1-\Theta(\tau)]\;\hat{V}.$$

Assume that we know the stationary solution of the initial Hamiltonian,

$$\hat{H}_0|m\rangle = \varepsilon_m|m\rangle, \qquad (2.31)$$

with the time-dependent wave function

$$|m(t)\rangle = e^{-\frac{i\varepsilon_m t}{\hbar}}|m\rangle, \qquad (2.32)$$

for time $t < t_0$ (this can be generalized for $\sum_m c_m|m(t)\rangle$, but let's keep things simple for now). Knowing that the system starts in state $|m(t)\rangle$, we want to find an approximate solution of the full SE, including the potential $\hat{V}$. In particular, we want to compute transition probabilities, e.g. the probability of transition from state $|m\rangle$ to state $|n\rangle$.

Summary of assumptions: $(\tau - t_0)$ small, $||\hat{V}|| \ll ||\hat{H}_0||$, solution for $\hat{H}_0$ known, initial state $|m(t)\rangle$.

We define the *interaction potential*

$$\hat{V}_i(t) := e^{\frac{i\hat{H}_0 t}{\hbar}}\hat{V}(t)e^{-\frac{i\hat{H}_0 t}{\hbar}} \qquad \Rightarrow \hat{V}(t) = e^{-\frac{i\hat{H}_0 t}{\hbar}}\hat{V}_i(t)e^{\frac{i\hat{H}_0 t}{\hbar}} \qquad (2.33)$$

and *interaction states*

$$|\tilde{\psi}(t)\rangle := e^{\frac{i\hat{H}_0 t}{\hbar}}|\psi(t)\rangle \qquad \Rightarrow |\psi(t)\rangle = e^{-\frac{i\hat{H}_0 t}{\hbar}}|\tilde{\psi}(t)\rangle, \qquad (2.34)$$

where $|\psi(t)\rangle$ is the actual state of the system (the solution of the SE). In particular, just before we turn on $\hat{V}$ (at $t = t_0$), the interaction state is

$$|\tilde{\psi}(t_0)\rangle = e^{\frac{i\hat{H}_0 t_0}{\hbar}}|m(t_0)\rangle = |m\rangle. \qquad (2.35)$$

Now we can write the SE in the so-called *interaction picture*,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle &= (\hat{H}_0+\hat{V}(t))|\psi(t)\rangle \\ i\hbar\frac{\partial}{\partial t}\left(e^{-\frac{i\hat{H}_0 t}{\hbar}}|\tilde{\psi}(t)\rangle\right) &= \left(\hat{H}_0 + e^{-\frac{i\hat{H}_0 t}{\hbar}}\hat{V}_i(t)e^{\frac{i\hat{H}_0 t}{\hbar}}\right)e^{-\frac{i\hat{H}_0 t}{\hbar}}|\tilde{\psi}(t)\rangle \\ i\hbar\left(-\frac{i\hat{H}_0}{\hbar}\right)e^{-\frac{i\hat{H}_0 t}{\hbar}}|\tilde{\psi}(t)\rangle + i\hbar e^{-\frac{i\hat{H}_0 t}{\hbar}}\frac{\partial}{\partial t}|\tilde{\psi}(t)\rangle &= \hat{H}_0 e^{-\frac{i\hat{H}_0 t}{\hbar}}|\tilde{\psi}(t)\rangle + e^{-\frac{i\hat{H}_0 t}{\hbar}}\hat{V}_i(t)|\tilde{\psi}(t)\rangle \\ i\hbar\;\frac{\partial}{\partial t}|\tilde{\psi}(t)\rangle &= \hat{V}_i(t)\;|\tilde{\psi}(t)\rangle. \qquad (2.36)\end{aligned}$$

Integrating the SE leads to

$$|\tilde{\psi}(t)\rangle = \left(\hat{T}_t e^{-\frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{V}_i(t')}\right) \underbrace{|\tilde{\psi}(t_0)\rangle}_{|m\rangle}, \tag{2.37}$$

where $\hat{T}_t$ is the time-ordering operator. Now we expand the exponential, obtaining

$$|\tilde{\psi}(t)\rangle = \left(\hat{\mathbb{1}} + \frac{1}{i\hbar}\left[\int_{t_0}^{t} dt' \hat{V}_i(t')\right] + \frac{1}{(i\hbar)^2}\left[\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \hat{V}_i(t')\hat{V}_i(t'')\right] + \mathcal{O}(\hat{V}^3)\right)|m\rangle. \tag{2.38}$$

This means that the final solution is a linear combination of terms $|\tilde{\psi}(t)\rangle = \sum_j |\tilde{\psi}_j(t)\rangle$, where $|\tilde{\psi}_j(t)\rangle$ is the solution for the jth term of the expansion. Remember that we are looking for the probability of transition from state $|m\rangle$ to state $|n\rangle$:

$$\begin{aligned}
\text{Prob}(|m\rangle \to |n\rangle) &= |\langle n(t)|\psi(t)\rangle|^2 \\
&= \left|\langle n|\tilde{\psi}(t)\rangle\right|^2 \\
&= \left|\langle n| \sum_j |\tilde{\psi}_j(t)\rangle\right|^2 \\
&= \left|\sum_j \langle n|\tilde{\psi}_j(t)\rangle\right|^2 .
\end{aligned} \tag{2.39}$$

The contribution from

- the 0th-order term, $|\tilde{\psi}_0(t)\rangle = \hat{\mathbb{1}}|m\rangle$, is trivial;
- the first-order term, $|\tilde{\psi}_1(t)\rangle = \frac{1}{i\hbar}\int_{t_0}^{t} dt' \hat{V}_i(t')\ |m\rangle$, results in Fermi's Golden Rule;
- the **second-order term**, $|\tilde{\psi}_2(t)\rangle = \frac{1}{(i\hbar)^2}\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \hat{V}_i(t')\hat{V}_i(t'')\ |m\rangle$, will be examined here;
- higher-order terms, $|\tilde{\psi}_{3+}(t)\rangle = O(\hat{V}^3)\ |m\rangle$, is dismissed (see assumptions) AA .

A quick note about taking only the second-order term. When we apply this to our problem of light-matter interaction, we will see that the contribution from the 0th- and first-order terms is zero for the relevant transition and interaction potential. In general,

$$\underbrace{\text{Prob}^{(2)}(|m\rangle \to |n\rangle) := \left|\langle n|\tilde{\psi}_2(t)\rangle\right|^2}_{\text{contribution from 2nd term}} \begin{cases} = \text{Prob}(|m\rangle \to |n\rangle) & \text{if } \langle n|\tilde{\psi}_j(t)\rangle = 0,\ \forall j \neq 2 \\ \approx \text{Prob}(|m\rangle \to |n\rangle) & \text{if } \langle n|\tilde{\psi}_0(t)\rangle = \langle n|\tilde{\psi}_1(t)\rangle = 0 \text{ and } ||\hat{V}|| \text{ small.} \end{cases} \tag{2.40}$$

The matrix element in Eq. 2.40 is

$$\begin{aligned}
\langle n|\tilde{\psi}_2(t)\rangle &= \frac{1}{(i\hbar)^2}\langle n|\left(\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \hat{V}_i(t')\hat{V}_i(t'')\right)|m\rangle \\
&= \frac{1}{(i\hbar)^2}\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \langle n|\hat{V}_i(t')\hat{V}_i(t'')|m\rangle,
\end{aligned} \tag{2.41}$$

since $|n\rangle, |m\rangle$ are time-independent. We will use two tricks to solve this integral 😵 .

Sort of a dirty trick. AA We assume $\hat{V}(t > t_0) = \hat{V}e^{\eta t}$, where $\eta = lim_{x\to 0}\ x =: 0^+$. This will make our integrals converge and it does not really change $\hat{V}$ by much, because η and t are very small (remember that we turn off $\hat{V}$ quickly); in addition $\hat{V}$ should be time-independent. It follows that

$$\hat{V}_i(t) := e^{\frac{i\hat{H}_0 t}{\hbar}}\hat{V}(t)e^{-\frac{i\hat{H}_0 t}{\hbar}} = e^{\frac{i\hat{H}_0 t}{\hbar}}\hat{V}e^{\eta t}e^{-\frac{i\hat{H}_0 t}{\hbar}}, \tag{2.42}$$

where $\hat{V}$ is time-independent. The matrix element becomes

$$\begin{aligned}
\langle n|\hat{V}_i(t)|l\rangle &= \langle n|e^{\frac{i\hat{H}_0 t}{\hbar}}\hat{V}e^{\eta t}e^{-\frac{i\hat{H}_0 t}{\hbar}}|l\rangle \\
&= \langle n|e^{\frac{i\varepsilon_n t}{\hbar}}\hat{V}e^{\eta t}e^{-\frac{i\varepsilon_l t}{\hbar}}|l\rangle \\
&= \exp\left[\frac{i}{\hbar}(\varepsilon_n - \varepsilon_l)t + \eta t\right] \underbrace{\langle n|\hat{V}|l\rangle}_{\text{time-independent}} .
\end{aligned} \tag{2.43}$$

We will use this soon.

Quite a clever, elegant trick. In Eq. 2.41, the term $\langle n|\hat{V}_i(t')\hat{V}_i(t'')|m\rangle$ is hard to compute directly (in general we may not know how to write $\hat{V}_i^2$). So we insert an identity operator between the two $\hat{V}_i$'s. We can write the identity as $\hat{\mathbb{1}} = \sum_\ell |\ell\rangle\langle\ell|$ for any orthonormal basis ($\langle\ell|\ell'\rangle = \delta_{\ell,\ell'}$) of our Hilbert space. We will choose this basis to be the eigenvalues of $\hat{H}_0$ (our stationary solutions), $\left\{|\ell\rangle \in \mathcal{H} : \hat{H}_0|\ell\rangle = \varepsilon_\ell|\ell\rangle\right\}$. The matrix element becomes

$$\begin{aligned}\langle n|\hat{V}_i(t')\mathbb{1}\hat{V}_i(t'')|m\rangle &= \langle n|\hat{V}_i(t')\left(\sum_\ell |\ell\rangle\langle\ell|\right)\hat{V}_i(t'')|m\rangle \\ &= \sum_\ell \langle n|\hat{V}_i(t')|\ell\rangle\ \langle\ell|\hat{V}_i(t'')|m\rangle, \end{aligned} \tag{2.44}$$

which is easier to calculate, because now we only have matrix elements of a single operator.

Using both tricks, and setting $t_0 = 0$ for simplicity, we can compute the total matrix element in Eq. 2.41,

$$\begin{aligned}\langle n|\tilde{\psi}_2(t)\rangle &= \frac{1}{(i\hbar)^2}\sum_\ell \langle n|\hat{V}|\ell\rangle\langle\ell|\hat{V}|m\rangle \int_{t_0}^{t} dt' \exp\left[\frac{i}{\hbar}(\varepsilon_n - \varepsilon_\ell)t' + \eta t'\right] \int_{t_0}^{t'} dt'' \exp\left[\frac{i}{\hbar}(\varepsilon_\ell - \varepsilon_m)t'' + \eta t''\right] \\ &= e^{\frac{i}{\hbar}(\varepsilon_m - \varepsilon_n)t}\frac{e^{2\eta t}}{\varepsilon_m - \varepsilon_n + 2i\hbar\eta}\sum_\ell \frac{\langle n|\hat{V}|\ell\rangle\langle\ell|\hat{V}|m\rangle}{\varepsilon_m - \varepsilon_\ell + i\hbar\eta}.\end{aligned}$$

So, the probability of transition at time t (at least the contribution from the second-order term) is

$$\mathrm{Prob}^{(2)}(|m\rangle \to |n\rangle) = \left|e^{\frac{i}{\hbar}(\varepsilon_m - \varepsilon_n)t}\right|^2 \frac{e^{4\eta t}}{|\varepsilon_m - \varepsilon_n + 2i\hbar\eta|^2} \underbrace{\left|\sum_\ell \frac{\langle n|\hat{V}|\ell\rangle\langle\ell|\hat{V}|m\rangle}{\varepsilon_m - \varepsilon_\ell + i\hbar\eta}\right|^2}_{=:A}. \tag{2.45}$$

In general, we are more interested in the transition rate

$$\begin{aligned}\Gamma_{m\to n} &= \frac{d}{dt}\mathrm{Prob}(|m\rangle \to |n\rangle) \\ &= \frac{d}{dt}\frac{e^{4\eta t}}{|\varepsilon_m - \varepsilon_n + 2i\hbar\eta|^2}A \\ &= \frac{4\eta e^{4\eta t}}{|\varepsilon_m - \varepsilon_n + 2i\hbar\eta|^2}A \\ &\approx A\frac{4\eta e^{4\eta t}}{(\varepsilon_m - \varepsilon_n)^2 + (2i\hbar\eta)^2} \qquad \text{since } \eta \to 0 \\ &= \frac{A}{\hbar^2}\lim_{\eta\to 0^+}\frac{\eta}{(\frac{\varepsilon_m - \varepsilon_n}{2\hbar})^2 + \eta^2} \\ &= \frac{A}{\hbar^2}\delta\left(\frac{\varepsilon_m - \varepsilon_n}{2\hbar}\right) \\ &= \frac{2A}{\hbar}\delta\left(\varepsilon_m - \varepsilon_n\right) \\ &= \frac{2}{\hbar}\left|\sum_\ell \frac{\langle n|\hat{V}|\ell\rangle\langle\ell|\hat{V}|m\rangle}{\varepsilon_m - \varepsilon_\ell + i\hbar\eta}\right| \delta\left(\varepsilon_m - \varepsilon_n\right).\end{aligned} \tag{2.46}$$

We are finished. What can we conclude? What do those terms with $|\ell\rangle$ mean? Here are a couple of ways to think about it:

1. Our system goes from $|m\rangle$ to $|n\rangle$ through intermediate states $|\ell\rangle$.

   The $|\ell\rangle$'s that more likely occur are the ones with energy close to that of $|m\rangle$ due to the term

   $$\frac{1}{\varepsilon_m - \varepsilon_\ell + i\hbar\eta}, \tag{2.47}$$

   but energy violations are possible from $|m\rangle$ to $|\ell\rangle$. You might hear in particle physics something along the lines of:

   > "but that's ok, because of the energy-time uncertainty relation and we can never detect $|\ell\rangle$ in the laboratory anyway..."

or as we say back home

"rebebebeu pardais ao ninho."

This interpretation makes it easy to visualise the transition going through the intermediate states, and you can think about it like that, **so long as you remember it is all in your mind**.

2. The $|\ell\rangle$'s are a nice mathematical trick we inserted to solve a huge integral, which is anyway only the second-order term of the exact expression of the propagator. They are an extremely useful tool, and we will use Feynman diagrams to represent them, but they do not mean anything "real", in the sense that there is no way to determine whether the system actually goes through $|\ell\rangle$.

   I like this interpretation.

3. What is reality anyway?

   I like you. Try taking Quantum information Theory in the next semester. I promise you will learn more about the meaning of reality there than in any philosophy class.

First interaction term

Now we are going to apply this *elegant theory* to the case of light-matter interaction (Section 2.6.2 in the OS). By the way, the OS' approximations are not so obvious, so we will do it in a slightly different way.

In Section 2.6.1 we applied PT to first order in time to the first and second interaction terms of our Hamiltonian. Now, we will apply PT to second order to the first term (Eq. 2.15),

$$\hat{V} = \hat{H}_{int,1} = -\frac{e}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar}{\omega_{\mathbf{k}}}} \left(\hat{\mathbf{j}}(-\mathbf{k}) \cdot \mathbf{e}(\mathbf{k},\lambda) \otimes \hat{a}_{\mathbf{k},\lambda} + \hat{\mathbf{j}}(\mathbf{k}) \cdot \mathbf{e}^*(\mathbf{k},\lambda) \otimes \hat{a}^\dagger_{\mathbf{k},\lambda} \right).$$

Again, we are looking for transitions of the form

$$|i\rangle = |0\rangle_p | \underbrace{N_{\mathbf{k},\lambda}}_{(\mathbf{k},\lambda)}, \underbrace{0}_{(\mathbf{k},\lambda)} \rangle_{EM} \qquad \Rightarrow \qquad |f\rangle = |n\rangle_p |N_{\mathbf{k},\lambda} - 1, 1\rangle_{EM}. \tag{2.48}$$

We already know that PT to first order on this term does not give us any of these transitions, $\Gamma^{(1)}_{i\to f} = 0$, so all the contributions to the transition rate will come from the second-order term (and higher, but we dismiss those). From Eq. 2.46, we have

$$\Gamma_{i\to f} \approx \Gamma^{(2)}_{i\to f} = \frac{2}{\hbar} \left| \sum_{\ell} \frac{\langle f|\hat{V}|\ell\rangle\langle\ell|\hat{V}|i\rangle}{E(i) - E(\ell) + i\hbar\eta} \right|^2 \delta E(i) - E(f). \tag{2.49}$$

The intermediate states $|\ell\rangle$ have the form $|\ell\rangle_p \otimes |\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle$ [sorry, I am running out of symbols here...]. Let us focus on the term

$$\langle f|\hat{V}|\ell\rangle\langle\ell|\hat{V}|i\rangle = \langle n|\langle N_{\mathbf{k},\lambda} - 1, 1|\ \hat{V}\ |\ell\rangle|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle\langle\ell|\langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\ \hat{V}\ |0\rangle|N_{\mathbf{k},\lambda}, 0\rangle$$

$$= -\frac{e\sqrt{2\pi\hbar}}{c\sqrt{L^3}} \left[\sum_{\mathbf{q},\mu} \frac{1}{\sqrt{\omega_{\mathbf{q}}}} \left(\langle n|\hat{\mathbf{j}}(-\mathbf{q}) \cdot \mathbf{e}(\mathbf{q},\mu)|\ell\rangle\langle N_{\mathbf{k},\lambda} - 1, 1|\hat{a}_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle + \langle n|\hat{\mathbf{j}}(\mathbf{q}) \cdot \mathbf{e}^*(\mathbf{q},\mu)|\ell\rangle\langle N_{\mathbf{k},\lambda} - 1, 1|\hat{a}^\dagger_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle \right) \right]$$

$$\left[\sum_{\mathbf{q}',\mu'} \frac{1}{\sqrt{\omega_{\mathbf{q}'}}} \left(\langle \ell|\hat{\mathbf{j}}(-\mathbf{q}') \cdot \mathbf{e}(\mathbf{q}',\mu')|0\rangle\langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\hat{a}_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda}, 0\rangle + \langle \ell|\hat{\mathbf{j}}(\mathbf{q}') \cdot \mathbf{e}^*(\mathbf{q}',\mu')|0\rangle\langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\hat{a}^\dagger_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda}, 0\rangle \right) \right]$$

$$= -\frac{e\sqrt{2\pi\hbar}}{c\sqrt{L^3}} \sum_{\mathbf{q},\mu} \sum_{\mathbf{q}',\mu'} \frac{1}{\sqrt{\omega_{\mathbf{q}}\ \omega_{\mathbf{q}'}}}$$

$$\begin{pmatrix} & \langle n|\hat{\mathbf{j}}(-\mathbf{q}) \cdot \mathbf{e}(\mathbf{q},\mu)|\ell\rangle & \langle \ell|\hat{\mathbf{j}}(-\mathbf{q}') \cdot \mathbf{e}(\mathbf{q}',\mu')|0\rangle & \langle N_{\mathbf{k},\lambda} - 1, 1|\hat{a}_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle & \langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\hat{a}_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda}, 0\rangle \\ + & \langle n|\hat{\mathbf{j}}(-\mathbf{q}) \cdot \mathbf{e}(\mathbf{q},\mu)|\ell\rangle & \langle \ell|\hat{\mathbf{j}}(\mathbf{q}') \cdot \mathbf{e}^*(\mathbf{q}',\mu')|0\rangle & \langle N_{\mathbf{k},\lambda} - 1, 1|\hat{a}_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle & \langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\hat{a}^\dagger_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda}, 0\rangle \\ + & \langle n|\hat{\mathbf{j}}(\mathbf{q}) \cdot \mathbf{e}^*(\mathbf{q},\mu)|\ell\rangle & \langle \ell|\hat{\mathbf{j}}(-\mathbf{q}') \cdot \mathbf{e}(\mathbf{q}',\mu')|0\rangle & \langle N_{\mathbf{k},\lambda} - 1, 1|\hat{a}^\dagger_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle & \langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\hat{a}_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda}, 0\rangle \\ + & \langle n|\hat{\mathbf{j}}(\mathbf{q}) \cdot \mathbf{e}^*(\mathbf{q},\mu)|\ell\rangle & \langle \ell|\hat{\mathbf{j}}(\mathbf{q}') \cdot \mathbf{e}^*(\mathbf{q}',\mu')|0\rangle & \langle N_{\mathbf{k},\lambda} - 1, 1|\hat{a}^\dagger_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda}, \ldots\rangle & \langle\tilde{N}_{\mathbf{k},\lambda}, \ldots|\hat{a}^\dagger_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda}, 0\rangle \end{pmatrix}.$$

Now, remember that to go from $|N_{\mathbf{k},\lambda}, 0\rangle$ to $N_{\mathbf{k},\lambda} - 1, 1$ we need to add a photon in mode $(\mathbf{k}', \lambda')$ and destroy a photon in mode $(\mathbf{k}, \lambda)$. In other words, we need an operator $\hat{a}^\dagger_{\mathbf{k}',\lambda'}$ and an operator $\hat{a}_{\mathbf{k},\lambda}$. We only have one operator between initial and intermediate states, and again one operator between intermediate and final states. This implies that those intermediate states must represent one step of the process: either we already have a photon in mode $(\mathbf{k}', \lambda')$, with $|N_{\mathbf{k},\lambda}, 1\rangle$, or we already took one photon from mode $(\mathbf{k}, \lambda)$, with $|N_{\mathbf{k},\lambda} - 1, 0\rangle$. This also means that, out of those four terms in our ginourmous equation above, only two will survive:

- $\langle N_{\mathbf{k},\lambda}-1,1|\hat{a}_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle\langle\tilde{N}_{\mathbf{k},\lambda},\ldots|\hat{a}_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda},0\rangle = 0$, because we cannot destroy two photons to go from the initial to the final state.
- $\langle N_{\mathbf{k},\lambda}-1,1|\hat{a}_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle\langle\tilde{N}_{\mathbf{k},\lambda},\ldots|\hat{a}^\dagger_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda},0\rangle = \sqrt{N_{\mathbf{k},\lambda}}$ if $|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle = |N_{\mathbf{k},\lambda},1\rangle$, $(\mathbf{q}',\mu') = (\mathbf{k}',\lambda')$, and $(\mathbf{q},\mu) = (\mathbf{k},\lambda)$; and 0 otherwise. Here, first we create a photon in mode $(\mathbf{k}',\lambda')$ and then destroy a photon from mode $(\mathbf{k},\lambda)$.
- $\langle N_{\mathbf{k},\lambda}-1,1|\hat{a}^\dagger_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle\langle\tilde{N}_{\mathbf{k},\lambda},\ldots|\hat{a}_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda},0\rangle = \sqrt{N_{\mathbf{k},\lambda}}$ if $|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle = |N_{\mathbf{k},\lambda}-1,0\rangle$, $(\mathbf{q}',\mu') = (\mathbf{k},\lambda)$, and $(\mathbf{q},\mu) = (\mathbf{k}',\lambda')$; and 0 otherwise. Here, first we destroy a photon from mode $(\mathbf{k},\lambda)$ and then create a photon in mode $(\mathbf{k}',\lambda')$.
- $\langle N_{\mathbf{k},\lambda}-1,1|\hat{a}^\dagger_{\mathbf{q},\mu}|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle\langle\tilde{N}_{\mathbf{k},\lambda},\ldots|\hat{a}^\dagger_{\mathbf{q}',\mu'}|N_{\mathbf{k},\lambda},0\rangle = 0$, because we cannot create two photons to go from the initial to the final state.

Note that we impose no restrictions on the intermediate state of the particle, $|\ell\rangle_p$ Our equation has become much more managable,

$$\langle f|\hat{V}|\ell\rangle\langle\ell|\hat{V}|i\rangle = -\frac{e\sqrt{2\pi\hbar}}{c\sqrt{L^3}}\sqrt{\frac{N_{\mathbf{k},\lambda}}{\omega_{\mathbf{k}}\ \omega_{\mathbf{k}'}}} \left(\begin{array}{lll} & \langle n|\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)|\ell\rangle & \langle\ell|\hat{\mathbf{j}}(\mathbf{k}')\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|0\rangle \quad \delta(|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle,|N_{\mathbf{k},\lambda},1\rangle) \\ + & \langle n|\hat{\mathbf{j}}(\mathbf{k}')\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|\ell\rangle & \langle\ell|\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)|0\rangle \quad \delta(|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle,|N_{\mathbf{k},\lambda}-1,0\rangle) \end{array}\right), \tag{2.50}$$

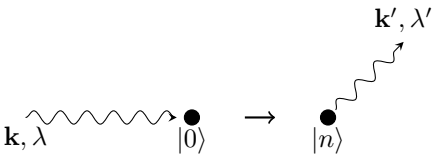
The transition rate comes

$$\begin{aligned}\Gamma_{i\to f} \approx \Gamma^{(2)}_{i\to f} &= \frac{2}{\hbar}\left|\sum_\ell \frac{\langle f|\hat{V}|\ell\rangle\langle\ell|\hat{V}|i\rangle}{E(i)-E(\ell)+i\hbar\eta}\right|^2 \delta(E(i)-E(f)) \\ &= \frac{4\pi e^2}{c^2L^3}\frac{N_{\mathbf{k},\lambda}}{\omega_{\mathbf{k}}\ \omega_{\mathbf{k}'}}\delta(\hbar\omega_{\mathbf{k}'}+E_n-E_0-\hbar\omega_{\mathbf{k}}) \\ &\qquad \left|\sum_\ell \underbrace{\frac{\langle n|\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)|\ell\rangle\langle\ell|\hat{\mathbf{j}}(\mathbf{k}')\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|0\rangle}{E_0-E_\ell-\hbar\omega_{\mathbf{k}'}+i\hbar\eta}}_{|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle=|N_{\mathbf{k},\lambda},1\rangle} + \underbrace{\frac{\langle n|\hat{\mathbf{j}}(\mathbf{k}')\cdot\mathbf{e}^*(\mathbf{k}',\lambda')|\ell\rangle\langle\ell|\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)|0\rangle}{E_0-E_\ell+\hbar\omega_{\mathbf{k}}+i\hbar\eta}}_{|\tilde{N}_{\mathbf{k},\lambda},\ldots\rangle=|N_{\mathbf{k},\lambda}-1,0\rangle}\right|^2.\end{aligned} \tag{2.51}$$

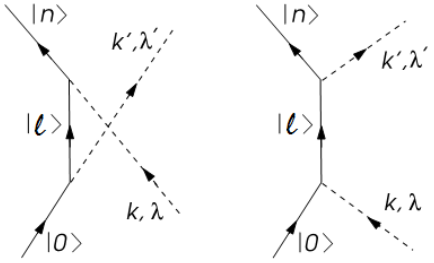
Now you can calculate the cross-section. [sorry, running out of time] The total rate of photons detected at a given angle, by the way, is the sum of the the cross-sections given by all interaction terms, with PT going as far as we find reasonable.

The Feynman diagram for this transition is represented in Fig. 2.5b. You get two diagrams, representing the two terms inside that sum: one diagram for the first term, where the intermediate state has one less photon in mode $(\mathbf{k},\lambda)$, and one diagram for the second term, where the intermediate state has one more photon in mode $(\mathbf{k}',\lambda')$.

Very Important Note: Feynman diagrams represent the matrix elements that appear when we apply PT to an interaction. They are a simple way to visualize all the ugly maths behind the final result for transition rates and cross-sections. They do not represent what happens "in reality". As we see in Fig. 2.5a, all that we can observe in the lab is that the particle absorbed a photon and emitted a different photon, going to an excited state, just like in the previous case (Section 2.6.1). Intermediate states are only in our mind.



(a) On the left (before), the particle absorbs a photon. On the right (after) the particle is in a excited state and emitted a photon in a different mode.



(b) Feynman diagrams for the interaction. On the left, a particle goes to state $|\ell\rangle$, emmiting a photon in mode $(\mathbf{k}', \lambda')$, and later absorbs a photon in mode $(\mathbf{k}, \lambda)$, going to the final state $|n\rangle$. On the right, first the particle absorbs the photon in mode $(\mathbf{k}, \lambda)$ and goes to state $|\ell\rangle$, and later emits a photon in mode $(\mathbf{k}', \lambda')$, going to the final state.

Figure 2.5: Transition obtained from PT to second order on time, applied to the first interaction term.

Chapter 6

Atoms (and the periodic table)

[This chapter is a teaser for quantum chemistry, and a motivation for computational physics.]

The goal: To understand the behaviour of atoms and ions with several electrons.

The problem: Interaction between the electrons makes things horribly complicated. There is no known analytical solution for two or more electrons. Current numerical simulations are quite good, matching experimental results, but we won't go there. Instead, we will look at a few approximations and attempts to solve them.

Notation: Z is number of protons in the nucleus, N the number of electrons. They are the same for atoms, different for ions.

The atom's Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{\hat{p}_i}{2m} - \sum_{i=1}^{N} \frac{Ze^2}{|\hat{x}_i|} + \sum_{i>j} \frac{e^2}{|\hat{x}_i - \hat{x}_j|}, \tag{6.1}$$

where e is the electron's charge, and all coordinates are relative to the nucleus. The first term is just the kinetic energy of every electron, and the second term is the Coulomb attraction to the protons in the nucleus. (AA) This Hamiltonian is already an approximation: we implicitly assume a point-like nucleus with charge Ze, instead of considering Z protons with their own wave functions (to say nothing of neutrons). In the last term, we have the interaction between electrons. We sum over $j, i > j$ to avoid double-counting of pairs. This term was not present in the hydrogen atom, and makes it impossible to find an exact analytical solution for the Hamiltonian (think of the three-body problem in classical mechanics). We will go over some approximations now, but you have been warned: it ain't pretty.

6.1 Thomas-Fermi approximation

This is a very old approximation that we keep learning mostly out of historical interest. It is crude and classical, and doesn't really allow us to compute anything very interesting or correct. However, some of the techniques used pop up again and again in solid-state physics, for example, so it is worth a look.

(AA) We assume,

1. $N \gg 1$ very large;
2. electrons independent;
3. the electrons create a classical density of charge $e\rho(\mathbf{r})$;
4. the density of electrons has spherical symmetry, $\rho(\mathbf{r}) = \rho(r)$ (this does not happen in reality: some spherical harmonics do not have this symmetry);
5. there are no electrons near the nucleus, so $\rho(r \to 0) = 0$ (this comes naturally if you do a full QM analysis of the problem, but has to be forced by hand in a classical approximation);
6. the electrons only feel $\rho(\mathbf{r})$ and not the other individual electrons (mean-field approximation; see solid-state physics)

We can normalize the charge density,

$$N = \int d^3\mathbf{r}\ \rho(r) = 4\pi \int_0^\infty dr\ r^2\ \rho(r). \tag{6.2}$$

Now we assume that each electron only feels the scalar potential created by this charge density, and not the individual electrons. This is the so-called *mean-field approximation*. The next step is to use Maxwell's equations to find the total potential,

$$-\nabla^2\phi(\mathbf{r}) = \begin{cases} 4\pi(-e)\rho(\mathbf{r}), & |r| > 0 \\ Ze, & |r| \to 0, \end{cases} \tag{6.3}$$

which tells us that the potential also has spherical symmetry. We can write the classical Hamiltonian of each individual electron as

$$H_i = \frac{\mathbf{p}_i}{2m} - e\phi(\mathbf{r_i}). \tag{6.4}$$

We are looking for the ground state of the global Hamiltonian $\hat{H} = \sum_i \hat{H}_i = \sum_i \frac{\mathbf{p}_i}{2m} - e\varphi(\mathbf{r}_i)$. In quantum mechanics we build the ground state of the hydrogen atom by putting two electrons of opposite spin per state, starting at the state with the lowest energy continuing until we ran out of electrons. The energy of the last occupied state is called the Fermi level, E_F.

In this classical approximation, we do something analogous: we populate the states of phase-space $(\mathbf{r}, \mathbf{p}, s)$ with a uniform density of electrons, up to E_F. AA We assume that the energy is degenerate in spin (which is not true), so we can forget about the degree of freedom s if we multiply the density of $(\mathbf{r}, \mathbf{p})$ by two. The volume of a minimal cell of phase space is $(2\pi\hbar)^3$, so we get an electron density in phase space of

$$\tilde{\rho}(\mathbf{r}, \mathbf{p}) = \begin{cases} \frac{2}{(2\pi\hbar)^3}, & E_F > E(\mathbf{r}, \mathbf{p}) := \frac{\mathbf{p}}{2m} - e\phi(r) \\ 0 & E_F < E(\mathbf{r}, \mathbf{p}). \end{cases} \tag{6.5}$$

To obtain the density of electrons *in space*, we integrate over all $\mathbf{p}$,

$$\begin{aligned} \rho(r) &= \int d^3\mathbf{p}\tilde{\rho}(\mathbf{r}, \mathbf{p}) && (6.6) \\ &= \frac{2}{(2\pi\hbar)^3} \int_{E<E_F} d^3\mathbf{p} \\ &= \frac{2}{(2\pi\hbar)^3} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\ \sin\theta \int_0^{\sqrt{2m(E_F + e\phi(r))}} dp\ p^2 \\ &= \frac{8\pi}{(2\pi\hbar)^3} \frac{1}{3} \left(2m[E_F + e\phi(r)]\right)^{3/2}. && (6.7) \end{aligned}$$

Here we used the fact that $\phi(r) \geq 0, \forall r > 0$ (the potential is negative only in the centre where the protons are, but there $\rho = 0$).

Now we found ourselves a system of differential equations,

$$\rho(r) = \begin{cases} 0, & r \to 0 \\ \frac{(2m[E_F + e\phi(r)])^{3/2}}{3\pi^2\hbar^3}, & r > 0 \end{cases} \tag{6.8}$$

$$\nabla^2\phi(r) = \begin{cases} -Ze, & r \to 0, \\ 4\pi e\rho(\mathbf{r}), & r > 0, \end{cases} \tag{6.9}$$

which people in the twenties had moderate fun solving.[1] We will not do that because with all those classical assumptions and approximations, the final density of electrons has little to do with the real wave-function density in an atom.

Nevertheless, this approximation was used to try to estimate the size of an atom for the first time. "Size" was defined as the radius of a ball that contains most of the electron density (sort of like the half-life time for radioactive elements). They obtained two results: when the percentage of charge inside this ball was set to be a constant, say 98%, then the radius grew with $Z^{-1/3}$. When the percentage changed with Z, like $(Z-1)/Z$, meaning "all the electrons except one are inside the ball", then the radius was constant [there is a mistake in the original script here!]. In other words: depending on how you measure it, the size of the atom decreases or stays the same as you increase the number of electrons. This is not complete nonsense, because we are also increasing the number of protons, and therefore the attraction to the nucleus. But it should be a warning that we seriously misestimated the interaction between electrons.

[1] If you want to solve this, add the normalization condition, 6.2, and find an expression for E_F.

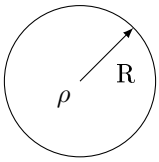


Figure 6.1: [Example: the atom is has 98% of the total electron density inside a radius of R.]

6.2 Hartree approximation

This is the first quantum-mechanical approximation. AA Assumptions:

1. The electrons are independent, distinguishable particles; in other words, the global wave function of all the electrons is a product state

$$|\psi\rangle_{1,2,\ldots,N} = |\psi_1\rangle_1 |\psi_2\rangle_2 \ldots |\psi_N\rangle_N, \tag{6.10}$$

   where $|\psi_j\rangle_i$ is the ket corresponding to the wave function ψ_j, living in Hilbert space $\mathcal{H}_i$.

2. The Pauli principle is introduced by hand by saying that we cannot have more than one electron with the same wave function ψ_j.

The next step is to solve the SE under these assumptions, to find the effective potential and the density of particles. However, as you may imagine, the results are not very accurate, since we stipulated that electrons are not fermions.

6.3 Hartree-Fock approximation

This is a more reasonable approximation, and we will study it in detail. It is widely used both in solid-state physics and in quantum chemistry.

AA Assumptions:

1. The electrons are non-interacting fermions; in other words, the global wave function is the anti-permutation of a product state,

$$|\psi\rangle_{1,2,\ldots,N} = \frac{1}{\sqrt{N!}} \sum_{P\in\mathcal{S}_N} (-1)^{|P|} |\psi_{P(1)}\rangle_1 |\psi_{P(2)}\rangle_2 \ldots |\psi_{P(N)}\rangle_N \tag{6.11}$$

$$= a_N^\dagger \ldots a_2^\dagger a_1^\dagger |0\rangle, \tag{6.12}$$

   where $|\psi_j\rangle_i$ is the ket corresponding to the wave function of particle i (includes the spatial wave function $\phi_j(\mathbf{x})$ and spin s_j). **Important remark:** In general, the global wave function is the anti-permutation of an arbitrary state. With this restriction, we are excluding many possible correlations between electrons.

Our primary goal is to find the atom's ground state. However, in general the ground state may have correlated electrons, so we will settle for the state of non-interacting electrons (of the form of Eq. 6.12) that minimizes the atom's energy. Formally, we want to find the single-particle wave-functions $\{|\psi_i\rangle\}$ that minimize $\langle\psi|H|\psi\rangle$ (where $|\psi\rangle$ is given by Eq. 6.12). This optimization is subject to a normalization constraint on the single-particle wave-functions, $\langle\psi|\psi\rangle - 1 = 0, \forall i$. We will solve this using the method of Lagrange multipliers: we define the Lagrange function

$$\Lambda(|\psi_1\rangle, \ldots, |\psi_N\rangle, \underbrace{\lambda_1, \ldots, \lambda_N}_{\text{Lagrange multipliers}}) = \langle\psi|H|\psi\rangle - \sum_i \lambda_i \left(\langle\psi_i|\psi_i\rangle - 1\right), \tag{6.13}$$

and try to minimize it, ie, find the arguments of $\nabla\Lambda = 0$. Before we do it, though, we have to write the term $\langle\psi|H|\psi\rangle$ as a function of the single-particle wave functions. This will require us to write both H and $|\psi\rangle$ in the second quantization formalism. The state is easy,

$$|\psi\rangle = a_N^\dagger \ldots a_2^\dagger a_1^\dagger |0\rangle, \tag{6.14}$$

where $|0\rangle$ is the vacuum. Now the Hamiltonian. Recall that, in general, we can create a many-fermion operator from a single-particle operator M,

$$M_{\text{1-part}} \longrightarrow \sum_{i,j} a_i^\dagger \, \langle\psi_i|M|\psi_j\rangle \, a_j, \tag{6.15}$$

for some basis $\{|\psi_i\rangle\}_i$ of the single-particle Hilbert space. Note that the operators $\left\{a_j, a_i^\dagger\right\}$ act on the global fermionic state, while M and $\{|\psi_i\rangle\}$ refer to a single particle. This new operator has the following physical interpretation: we see if M takes state $|\psi_j\rangle$ to state $|\psi_i\rangle$ (the weight of this transtion is given by $\langle\psi_i|M|\psi_j\rangle$). If that happens, we destroy a particle in state $|\psi_j\rangle$ and create a particle in state $|\psi_i\rangle$, using the fermionic operators. For two-particle operators, this generalizes to

$$M_{\text{2-part}} \longrightarrow \sum_{i,j,k,m} a_i^\dagger a_j^\dagger \ \langle\psi_i|\langle\psi_j|M|\psi_k\rangle|\psi_m\rangle \ a_k a_m. \tag{6.16}$$

Applying this to our Hamiltonian, we obtain

$$H = \sum_{ij} a_i^\dagger \ \langle\psi_i|\frac{\hat{p}^2}{2m}|\psi_j\rangle \ a_j - \sum_{ij} a_i^\dagger \ \langle\psi_i|\frac{Ze^2}{|\hat{x}_j|}|\psi_j\rangle \ a_j + \frac{1}{2}\sum_{ijkm} a_i^\dagger a_j^\dagger \ \langle\psi_i|\langle\psi_j|\frac{e^2}{|\hat{x}_k - \hat{x}_m|}|\psi_k\rangle|\psi_m\rangle \ a_k a_m. \tag{6.17}$$

Now we can compute the energy of this state, $\langle\psi|H|\psi\rangle$, one term at the time.
For the two first terms, with one-particle operators, we have

$$\langle\psi|\left(\sum_{ij} a_i^\dagger \ \langle\psi_i|M|\psi_j\rangle \ a_j\right)|\psi\rangle \ = \sum_{ij}\langle\psi_i|M|\psi_j\rangle \ \underbrace{\langle\psi|a_i^\dagger a_j|\psi\rangle}_{=\delta_{ij}\Theta(N-i)} = \sum_{i=1}^{M}\langle\psi_i|M|\psi_i\rangle \tag{6.18}$$

For two-particle operators, we have

$$\begin{aligned}
&\langle\psi|\left(\sum_{ijkm} a_i^\dagger a_j^\dagger \ \langle\psi_i|\langle\psi_j|M|\psi_k\rangle|\psi_m\rangle \ a_k a_m\right)|\psi\rangle \\
&= \sum_{ijkm}\langle\psi_i|\langle\psi_j|M|\psi_k\rangle|\psi_m\rangle \ \langle\psi|a_i^\dagger a_j^\dagger a_k a_m|\psi\rangle \\
&= \sum_{i,j=1}^{N}\langle\psi_i|\langle\psi_j|M|\psi_i\rangle|\psi_j\rangle \ \underbrace{\langle\psi|a_i^\dagger a_j^\dagger a_i a_j|\psi\rangle}_{i=k,j=m} + \langle\psi_i|\langle\psi_j|M|\psi_j\rangle|\psi_i\rangle \ \underbrace{\langle\psi|a_i^\dagger a_j^\dagger a_j a_i|\psi\rangle}_{i=k,j=m} \\
&= \sum_{i,j=1}^{N} -\langle\psi_i|\langle\psi_j|M|\psi_i\rangle|\psi_j\rangle \ \underbrace{\langle\psi|a_i^\dagger a_i a_j^\dagger a_j|\psi\rangle}_{=1} + \langle\psi_i|\langle\psi_j|M|\psi_j\rangle|\psi_i\rangle \ \underbrace{\langle\psi|a_i^\dagger a_i a_j^\dagger a_j|\psi\rangle}_{=1} \\
&= \sum_{i,j=1}^{N} -\langle\psi_i|\langle\psi_j|M|\psi_i\rangle|\psi_j\rangle + \langle\psi_i|\langle\psi_j|M|\psi_j\rangle|\psi_i\rangle.
\end{aligned}$$

Applying these results to our Hamiltonian, we obtain

$$\langle\psi|H|\psi\rangle = \sum_{i=1}^{N}\langle\psi_i|\frac{-\hbar^2\nabla^2}{2m}|\psi_i\rangle - \sum_{i=1}^{N}\langle\psi_i|\frac{Ze^2}{|\hat{x}_i|}|\psi_i\rangle - \sum_{i,j=1}^{N}\langle\psi_i|\langle\psi_j|\frac{e^2}{|\hat{x}_i - \hat{x}_j|}|\psi_i\rangle|\psi_j\rangle + \sum_{i,j=1}^{N}\langle\psi_i|\langle\psi_j|\frac{e^2}{|\hat{x}_j - \hat{x}_i|}|\psi_j\rangle|\psi_i\rangle. \tag{6.19}$$

6.3.1 Minimizing the Lagrange function

Remember that the Lagrange function was defined as $\Lambda = \langle\psi|H|\psi\rangle - \sum_i \lambda_i \ (\langle\psi_i|\psi_i\rangle - 1)$. Using the reults above, it becomes

$$\begin{aligned}
\Lambda(|\psi_1\rangle,\ldots,|\psi_N\rangle,\lambda_1,\ldots,\lambda_N) = -\sum_{i=1}^{N}\Bigg(&\frac{\hbar^2}{2m}\langle\psi_i|\nabla^2|\psi_i\rangle + Ze^2\langle\psi_i|\frac{1}{|\hat{x}_i|}|\psi_i\rangle \\
&+ e^2\sum_{j=1}^{N}\langle\psi_j|\langle\psi_i|\frac{1}{|\hat{x}_j - \hat{x}_i|}|\psi_j\rangle|\psi_i\rangle - e^2\sum_{j=1}^{N}\langle\psi_i|\langle\psi_j|\frac{e^2}{|\hat{x}_i - \hat{x}_j|}|\psi_j\rangle|\psi_i\rangle \\
&+ \lambda_i \ (\langle\psi_i|\psi_i\rangle - 1)\Bigg),
\end{aligned} \tag{6.20}$$

where we rearranged the indices of the two-particle terms for later convenience. We are looking for the arguments $(|\psi_1\rangle,\ldots,|\psi_N\rangle,\lambda_1,\ldots,\lambda_N)$ that minimize it. We can split the minimization condition in two parts:

1. Derivatives in order to all scalar multipliers $\{\lambda_i\}$ must be zero, giving us

$$\bigwedge_i \frac{\partial}{\partial \lambda_i}\Lambda = 0 \Leftrightarrow \forall i \ \langle\psi_i|\psi_i\rangle = 1, \tag{6.21}$$

2. All the $\{|\psi_i\rangle\}$ must also be local minima. Let us see what that implies in a more general case. Say we want to find a local minimum $|\phi_0\rangle$ of a matrix element of the form $\langle\phi|A|\phi\rangle$, where A is any Hermitian operator. This means that if you change the function ϕ_0 even a tiny bit, the value of the matrix element will increase. We can express that in the following way: if instead of $|\phi_0\rangle$ we have $|\phi_0\rangle + t|\eta\rangle$, where η is any other function and t is very small, then the matrix element increases. In other words, the derivative in order to t of the above expression must be zero for t tiny,

$$\frac{\partial}{\partial t}\Big(\langle\phi_0| + t\langle\eta|\Big)A\Big(|\phi_0\rangle + t|\eta\rangle\Big)\Big|_{t=0} = 0, \qquad \forall|\eta\rangle. \tag{6.22}$$

This will give us a very nice expression in the end. Stick by as we compute it,

$$\begin{aligned}
& \frac{\partial}{\partial t}\Big(\langle\phi_0|A|\phi_0\rangle + t\langle\eta|A|\phi_0\rangle + t\langle\phi_0|A|\eta\rangle + t^2\langle\eta|A|\eta\rangle\Big)\Big|_{t=0} = 0, \qquad \forall|\eta\rangle \\
\Leftrightarrow \quad & \Big(0 + \langle\eta|A|\phi_0\rangle + \langle\phi_0|A|\eta\rangle + 2t\langle\eta|A|\eta\rangle\Big)\Big|_{t=0} = 0, \qquad \forall|\eta\rangle \\
\Leftrightarrow \quad & \langle\eta|A|\phi_0\rangle + \langle\phi_0|A|\eta\rangle = 0, \qquad \forall|\eta\rangle.
\end{aligned}$$

In particular, this is true for $|\eta'\rangle = i|\eta\rangle$, so we get

$$\begin{aligned}
& \forall|\eta\rangle \quad \langle\eta|A|\phi_0\rangle + \langle\phi_0|A|\eta\rangle = 0 \quad \wedge \quad -i\langle\eta|A|\phi_0\rangle + i\langle\phi_0|A|\eta\rangle = 0 \\
\Leftrightarrow \quad & \forall|\eta\rangle \quad \langle\eta|A|\phi_0\rangle + \langle\phi_0|A|\eta\rangle = 0 \quad \wedge \quad -\langle\eta|A|\phi_0\rangle + \langle\phi_0|A|\eta\rangle = 0 \\
\Leftrightarrow \quad & \forall|\eta\rangle \quad \langle\eta|A|\phi_0\rangle = \langle\phi_0|A|\eta\rangle = 0 \\
\Rightarrow \quad & A|\phi_0\rangle = \langle\phi_0|A = 0.
\end{aligned}$$

[This is slightly different from the way I did it in the lecture. This version is more rigorous, and works for finite-dimensional Hilbert spaces. If we have infinitely many dimensions, the calculation becomes more delicate (though the result is similar). We won't treat that case here.]

Now we have to reduce our problem to the minimization of terms of the form $\langle\psi|A|\psi\rangle$. That is fairly direct, as

$$\Lambda = -\sum_{i=1}^{N}\langle\psi_i|\underbrace{\begin{pmatrix} \frac{\hbar^2}{2m}\nabla^2 + Ze^2\frac{1}{|\hat{x}_i|} \\ +e^2\sum_{j=1}^{N}(\langle\psi_j|\otimes\mathcal{I})\frac{1}{|\hat{x}_j-\hat{x}_i|}(|\psi_j\rangle\otimes\mathcal{I}) \\ -e^2\sum_{j=1}^{N}(\mathcal{I}\otimes\langle\psi_j|)\frac{e^2}{|\hat{x}_i-\hat{x}_j|}(|\psi_j\rangle\otimes\mathcal{I}) \\ +\lambda_i \end{pmatrix}}_{A}|\psi_i\rangle - \lambda_i. \tag{6.23}$$

All the $\{|\psi_i\rangle\}$ and $\{\lambda_i\}$ count as independent variables, so for each $|\psi_i\rangle$ we have to minimize $\langle\psi_i|A|\psi_i\rangle$ (the final λ_i is a constant). This gives us the minimization condition

$$\begin{aligned}
& 0 = A|\psi_i\rangle \\
\Leftrightarrow \quad & \lambda_i\,|\psi_i\rangle = -\frac{\hbar^2}{2m}\nabla^2|\psi_i\rangle - Ze^2\frac{1}{|\hat{x}_i|}|\psi_i\rangle - e^2\sum_{j=1}^{N}\langle\psi_j|\frac{1}{|\hat{x}_i-\hat{x}_j|}|\psi_j\rangle|\psi_i\rangle + e^2\sum_{j=1}^{N}\langle\psi_j|\frac{e^2}{|\hat{x}_j-\hat{x}_i|}|\psi_i\rangle|\psi_j\rangle.
\end{aligned}$$

We can expand the inner products of the last two terms. Remember that the single-particle wave-functions included the spatial wave function and spin, $|\psi_i\rangle = |\phi_i\rangle|s_i\rangle$, where $\phi_i(\mathbf{x})$ depends on the position. We obtain

$$\begin{aligned}
\lambda_i\,|\phi_i\rangle|s_i\rangle = & -\left(\frac{\hbar^2}{2m}\nabla_i^2 + Ze^2\frac{1}{|\hat{x}_i|}\right)|\phi_i\rangle|s_i\rangle \\
& - e^2\sum_{j=1}^{N}\left[\int d^3\mathbf{x_j}\phi_j^*(\mathbf{x}_j)\frac{1}{|\hat{x}_i-\mathbf{x}_j|}\phi_j(\mathbf{x}_j)\right]\underbrace{\langle s_j|s_j\rangle}_{=1}|\phi_i\rangle|s_i\rangle \\
& + e^2\sum_{j=1}^{N}\left[\int d^3\mathbf{x_j}\phi_j^*(\mathbf{x}_j)\frac{1}{|\hat{x}_i-\mathbf{x}_j|}\phi_i(\mathbf{x}_j)\right]\underbrace{\langle s_j|s_i\rangle}_{=\delta_{s_is_j}}|\phi_j\rangle|s_i\rangle.
\end{aligned}$$

The spin component is not doing much there (since none of those operators acts on it). Without it, the equation looks like

$$\begin{aligned}\lambda_i\ |\phi_i\rangle = &-\left(\frac{\hbar^2}{2m}\nabla_i^2 + Ze^2\frac{1}{|\hat{x_i}|}\right)|\phi_i\rangle \\ &- e^2\sum_{j=1}^{N}\left(\int d^3\mathbf{y}\phi_j^*(\mathbf{y})\frac{1}{|\hat{x}-\mathbf{y}|}\phi_j(\mathbf{y})\right)|\phi_i\rangle \\ &+ e^2\sum_{j=1}^{N}\left(\int d^3\mathbf{y}\phi_j^*(\mathbf{y})\frac{1}{|\hat{x}-\mathbf{y}|}\phi_i(\mathbf{y})\right)|\phi_j\rangle\ \delta_{s_i s_j}.\end{aligned}$$

If we are interested in the spatial wave function at a particular point $\mathbf{x}$, we get, using $\mathbf{y} := \mathbf{x_j}$ inside the integrals for simplicity,

$$\begin{aligned}\lambda_i\ \phi_i(\mathbf{x}) = &-\left(\frac{\hbar^2}{2m}\nabla_x^2 + Ze^2\frac{1}{|\mathbf{x}|}\right)\phi_i(\mathbf{x}) \\ &- e^2\sum_{j=1}^{N}\int d^3\mathbf{y}\phi_j^*(\mathbf{y})\frac{1}{|\mathbf{x}-\mathbf{y}|}\phi_j(\mathbf{y})\phi_i(\mathbf{x}) + e^2\sum_{j=1}^{N}\int d^3\mathbf{y}\phi_j^*(\mathbf{y})\frac{1}{|\mathbf{x}-\mathbf{y}|}\phi_i(\mathbf{y})\phi_i(\mathbf{x})\ \delta_{s_i s_j}. \\ = &\underbrace{-\left(\frac{\hbar^2}{2m}\nabla_x^2 + Ze^2\frac{1}{|\mathbf{x}|}\right)\phi_i(\mathbf{x})}_{\text{single particle}} + e^2\sum_{j=1}^{N}\int d^3\mathbf{y}\frac{1}{|\mathbf{x}-\mathbf{y}|}\left[\underbrace{-\phi_j^*(\mathbf{y})\phi_j(\mathbf{y})\phi_i(\mathbf{x})}_{\text{Hartree term}} + \underbrace{\delta_{s_i s_j}\ \phi_j^*(\mathbf{y})\phi_i(\mathbf{y})\phi_i(\mathbf{x})}_{\text{Fock term}}\right].\end{aligned}$$

This equation looks a lot like a SE, where λ_i acts almost as an energy eigenvalue. Now we have a system of DE which can be solved numerically, for instance using as initial $\{\lambda_i, |\psi_i\rangle\}_i$ the energy eigenvalues and eigenstates of a single-electron atom. Once we have the non-correlated state that minimizes the Hamiltonian, we can apply small numerical perturbations to that state and see if the energy decreases. The perturbated state will in general have correlations between electrons.

[See in the script/wikipedia the accuracy of this approximation.]

Chapter 7

Statistical quantum mechanics

I am going to review some basic concepts of quantum mechanics. Bear with me, it will become interesting very soon. For more details, see the Quantum Information Theory script, by Profs. Renato Renner and Matthias Christandl,
`http://www.itp.phys.ethz.ch/education/lectures_hs11/qit/resources/script`
(beginning of chapter 4).

7.1 Pure states: evolution and measurements

Say you have a quantum state $|\psi\rangle$ on a Hilbert space $\mathcal{H}$. We saw in QM1 that quantum states evolve under unitary operators, $|\psi(t)\rangle = U_t|\psi(0)\rangle$, with $U_t = e^{-iHt}$.

There are a few conventions to describe measurements performed on quantum states. A special case is a measurement where your device can distinguish all the elements of a basis of $\mathcal{H}$. For instance, you measure a qubit in the computational basis, and your machine gives you two outcomes: "0" when it measures $|0\rangle$, and "1" when it measures $|1\rangle$. The probability of obtaining outcome "x" is given by

$$\Pr[x]_\psi = |\langle\psi|x\rangle|^2 = \langle\psi|x\rangle\langle x|\psi\rangle, \tag{7.1}$$

which we can rewrite with the help of the matrix trace as

$$\Pr[x]_\psi = \mathrm{Tr}(|x\rangle\langle x|\ |\psi\rangle\langle\psi|). \tag{7.2}$$

To see this, recall that the trace is the sum of the diagonal elements of a matrix, with respect to any basis $\{|i\rangle\}_i$, $\mathrm{Tr}(A) = \sum_i \langle i|A|i\rangle$. We have

$$\begin{aligned}
\mathrm{Tr}(|x\rangle\langle x|\ |\psi\rangle\langle\psi|) &= \sum_i \langle i|x\rangle\langle x|\psi\rangle\langle\psi|i\rangle \\
&= \langle x|\psi\rangle \sum_i \langle\psi|i\rangle\langle i|x\rangle \\
&= \langle x|\psi\rangle|\psi\rangle \underbrace{\left(\sum_i |i\rangle\langle i|\right)}_{\mathbb{1}} \langle x| \\
&= \langle x|\psi\rangle\langle\psi|x\rangle.
\end{aligned}$$

Recap: trace. The trace is linear, $\alpha\mathrm{Tr}(A) + \beta\mathrm{Tr}(B) = \mathrm{Tr}(\alpha A + \beta B)$, cyclic, $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB)$, and invariant under unitary transformations, i.e., basis-independent, $\mathrm{Tr}(A) = \mathrm{Tr}(UAU^\dagger)$.

After a measurement with outcome x, the state of the system *collapses* to $|x\rangle$. [1] We call the operator $|x\rangle\langle x|$ the projector onto state $|x\rangle$. We will generalise it in the next section. Before that, a word about vectors and matrices in finite-dimensional systems.

7.1.1 Finite systems: kets and bras as vectors

If the Hilbert space $\mathcal{H}$ is finite (e.g. a qubit), you can use vectors to represent kets and bras, and matrices for operators. This is probably explained in detail in QM1, but I will repeat it here for the sake of completeness.

[1] Again, for more about this, come to QIT!

Pick a basis $\{\phi_i\}_{i=1}^N$ of your Hilbert space. We identify kets with the elements of the Hilbert space, for instance $|i\rangle : c \mapsto x\phi_i$. Now any ket $|\psi\rangle$ can be expanded in this basis:

$$|\psi\rangle = \sum_i \langle \phi_i, \psi\rangle \ |i\rangle, \tag{7.3}$$

where $\langle \psi, \phi_i\rangle$ is the inner product. We represent this as a vertical vector,

$$|\psi\rangle = \begin{pmatrix} \langle \phi_0, \psi\rangle \\ \langle \phi_1, \psi\rangle \\ \vdots \\ \langle \phi_N, \psi\rangle \end{pmatrix}, \tag{7.4}$$

such that row i corresponds to ket $|i\rangle$. What about bras? They are simply horizontal vectors,

$$\langle\tau| = (\langle \tau, \psi_0\rangle, \ \langle \tau, \phi_1\rangle, \ \cdots, \ \langle \tau, \phi_N\rangle) \tag{7.5}$$

This is compatible with our previous definition of bra as a function that takes a wave function and returns an inner product. In fact,

$$\begin{aligned} \langle\tau| \ |\psi\rangle &= (\langle \tau, \psi_0\rangle, \ \langle \tau, \phi_1\rangle, \ \cdots, \ \langle \tau, \phi_N\rangle) \begin{pmatrix} \langle \phi_0, \psi\rangle \\ \langle \phi_1, \psi\rangle \\ \vdots \\ \langle \phi_N, \psi\rangle \end{pmatrix} \\ &= \sum_i \langle \tau, \phi_i\rangle \ \langle \phi_i, \psi\rangle \\ &= \langle \tau, \psi\rangle, \end{aligned}$$

hence we can write inner products as contractions between bras and kets, $\langle \tau, \psi\rangle = \langle\tau| \ |\psi\rangle =: \langle\tau|\psi\rangle$. Similarly, operators that act on the Hilbert space (endomorphisms on $\mathcal{H}$) can be represented as matrices,

$$\begin{aligned} A &= \underbrace{\left(\sum_i |i\rangle\langle i|\right)}_{\mathbb{1}} A \underbrace{\left(\sum_j |j\rangle\langle j|\right)}_{\mathbb{1}} \\ &= \sum_{i,j} \underbrace{\langle i|A|j\rangle}_{\langle \phi_i, A\phi_j\rangle} \ |i\rangle\langle j| \\ &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{N1} & & & a_{NN} \end{pmatrix}, \quad a_{ij} = \langle i|A|j\rangle. \end{aligned} \tag{7.6}$$

Examples. Say that $\mathcal{H}$ is the Hilbert space of a qubit. Here go a few quick examples of kets, bras, and operators in the computational basis, $\{|0\rangle, |1\rangle\}$.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad |-\rangle\langle -| = \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| - \langle 1|}{\sqrt{2}}\right) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\Pr[1]_{|-\rangle} = \mathrm{Tr}\Big[|1\rangle\langle 1| \ |-\rangle\langle -|\Big] = \mathrm{Tr}\left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right] = \mathrm{Tr}\left[\frac{1}{2}\begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}\right] = \frac{1}{2}.$$

7.2 Density operator

Imagine that you are given a quantum state to play with. The only problem is that you are not sure what state. For instance, you may have a flawed source of quantum states, which was supposed to produce state $|\psi\rangle$, but

will produce state $|\tau\rangle$ instead with probability p. What is the probability of obtaining $|x\rangle$ when you measure your unknown state in that basis? Well, it has to be [probability of having $|\psi\rangle$] $*$ [probability of measuring $|x\rangle$ on $|\psi\rangle$] + [probability of having $|\tau\rangle$] $*$ [probability of measuring $|x\rangle$ on $|\tau\rangle$],

$$\begin{aligned}
\Pr(x) &= (1-p)\Pr(x)_\psi + p\ \Pr(x)_\tau \\
&= (1-p)\mathrm{Tr}(|x\rangle\langle x|\ |\psi\rangle\langle\psi|) + p\ \mathrm{Tr}(|x\rangle\langle x|\ |\tau\rangle\langle\tau|) \\
&= \mathrm{Tr}\big((1-p)|x\rangle\langle x|\ |\psi\rangle\langle\psi| + p\ |x\rangle\langle x|\ |\tau\rangle\langle\tau|\big) \\
&= \mathrm{Tr}\Big(|x\rangle\langle x|\ \underbrace{\Big[(1-p)\ |\psi\rangle\langle\psi| + p\ |\tau\rangle\langle\tau|\Big]}_{=:\rho}\Big) \\
&= \mathrm{Tr}(|x\rangle\langle x|\ \rho). \qquad (7.7)
\end{aligned}$$

Here, we defined the density operator $\rho = (1-p)\ |\psi\rangle\langle\psi| + p\ |\tau\rangle\langle\tau|$. This is a mathematical object that reflects our ignorance about the actual state of the system. For infinite dimensions, it becomes a matrix (the density matrix). For instance, if $|\psi\rangle = |1\rangle$ and $|\tau\rangle = |-\rangle$, we have

$$\begin{aligned}
\rho &= (1-p)\ |1\rangle\langle 1| + p\ |-\rangle\langle -| \\
&= (1-p)\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + p\ \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{p}{2} & -\frac{p}{2} \\ -\frac{p}{2} & 1-\frac{p}{2} \end{pmatrix}.
\end{aligned}$$

The probability of obtaining a "1" when measuring this state in the computational basis is

$$\begin{aligned}
\Pr(1)_\rho &= \mathrm{Tr}[|1\rangle\langle 1|\ \rho] \\
&= \mathrm{Tr}\left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{p}{2} & -\frac{p}{2} \\ -\frac{p}{2} & 1-\frac{p}{2} \end{pmatrix}\right] \\
&= \mathrm{Tr}\left[\begin{pmatrix} 0 & 0 \\ -\frac{p}{2} & 1-\frac{p}{2} \end{pmatrix}\right] = 1 - \frac{p}{2}.
\end{aligned}$$

Note that we could also have computed this quantity the old way,

$$\begin{aligned}
\Pr(1)_\rho &= (1-p)\ \Pr(1)_{|1\rangle} + p\ \Pr(1)_{|-\rangle} \\
&= (1-p)\ 1 + p\ \frac{1}{2} = 1 - \frac{p}{2}.
\end{aligned}$$

It practice, however, it is generally more convenient to use density matrices. Now let us look at time evolution of density operators. At time t, the state is either $U_t|\psi\rangle$, with probability $1-p$, or $U_t|\tau\rangle$, with probability p. The probability of measuring "x" at time t is therefore

$$\begin{aligned}
\Pr(x) &= \mathrm{Tr}\left[|x\rangle\langle x|\ (\ (1-p)\ U_t|\psi\rangle\langle\psi|U_t^\dagger + p\ U_t|\tau\rangle\langle\tau|U_t^\dagger)\right] \\
&= \mathrm{Tr}\left[|x\rangle\langle x|\ U_t\ (\ \underbrace{(1-p)\ |\psi\rangle\langle\psi| + p\ |\tau\rangle\langle\tau|}_{\rho_0})\ U_t^\dagger\right] \\
&= \mathrm{Tr}\left[|x\rangle\langle x|\ \underbrace{U_t\ \rho_0\ U_t^\dagger}_{=:\rho(t)}\right]. \qquad (7.8)
\end{aligned}$$

We have just showed that the time evolution of a density operator is simply given by

$$\rho(t) = U_t\ \rho(0)\ U_t^\dagger. \qquad (7.9)$$

Now, this is more relevant than it might appear at first sight. Think about what you can do in a lab: you can let systems evolve, by manipulating Hamiltonians, and you can measure them. You have showed that both time evolution and measurement statistics of a system depend only on the density operator (and the Hamiltonian). This implies that it is impossible to distinguish two systems with the same density operator, even if they were produced in different ways. Let me give you a somewhat trivial example: suppose that your friend Alice has two machines that produce polarized photons. The only problem is that those machines are not reliable at all: machine A produces state $|0\rangle$ with probability $1/2$ and state $|1\rangle$ with probability $1/2$, while machine B produces B produces state $|+\rangle$ with probability $1/2$ and state $|-\rangle$ with probability $1/2$. The density operators of a state coming from each machine are

$$\begin{aligned}
\rho_A &= \frac{1}{2}\ |0\rangle\langle 0| + \frac{1}{2}\ |1\rangle\langle 1| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\mathbb{1}}{2}, \\
\rho_B &= \frac{1}{2}\ |+\rangle\langle +| + \frac{1}{2}\ |+\rangle\langle +| = \frac{1}{2}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \frac{\mathbb{1}}{2}.
\end{aligned}$$

Now imagine that Alice gives you a million photons, and tells you that she used the same machine to produce all of them. But she won't tell you which machine. You would think that it would be easy to find out, right? Here are a million photons, either half are $|0\rangle$ and half are $|1\rangle$ or half are $|+\rangle$ and half are $|-\rangle$. You can do whatever you want to your photons: rotate the states, measure them, let them evolve under convoluted Hamiltonians, make them interact with an external system, let your local soothsayer examine them. And yet you will never be able to determine which machine was used, because photons from A and from B have the same density matrix. The origins of a quantum state do not matter.[2]

7.2.1 Properties and terminology of density operators

- In general,

$$\rho = \sum_i p_i \ |\psi_i\rangle\langle\psi_i|, \tag{7.10}$$

  where the $\{p_i\}$ form a probability distribution (meaning $\forall p_i \geq 0$ and $\sum_i p_i = 1$), and the $\{|\psi_i\rangle\}$ are quantum states. This has the following physical interpretation: the system is in state $|\psi_i\rangle$ with probability p_i.

- Density operators are endomorphisms on the system's Hilbert space, $\rho : \mathcal{H} \to \mathcal{H}$.

- By definition, the $\{p_i\}$ are the eigenvalues of ρ, therefore $\mathrm{Tr}(\rho) = \sum_i p_i = 1$, ρ is positive semi-definite, and Hermitian.

- A system is said to be in a *pure* state if the corresponding density operator only has one non-zero eigenvalue, $\{p_i\} = \{1, 0, , 0 \ldots\}$, and therefore $\rho = |\psi\rangle\langle\psi|$. Otherwise a state is considered *mixed* (meaning that there is more than one possibility for the exact state of the system).

- If $\rho = \mathbb{1}/N$, where N is the dimension of the Hilbert space of interest, it is called *fully mixed*.

7.2.2 Superposition *vs* mixture

It is easy to confuse mixtures of quantum states with quantum superpositions in the begininning. For instance, you may be wondering about the difference between states

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad \text{and} \quad \rho = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}.$$

In a nutshell, the difference is that the former is a pure state (you know that the system is in that exact state), while the latter is a probabilistic mixture of two possible states. To clarify things, we can write the density matrices corresponding to both states, in two different bases:

$$\begin{aligned}
\sigma :&= |+\rangle\langle +| \\
&= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \ \frac{\langle 0| + \langle 1|}{\sqrt{2}} \\
&= \frac{|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|}{2}
\end{aligned}
\qquad
\begin{aligned}
\rho &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\
&= \frac{1}{2}\begin{pmatrix} \frac{|+\rangle+|-\rangle}{\sqrt{2}}\frac{\langle +|+\langle -|}{\sqrt{2}} \\ +\frac{|+\rangle-|-\rangle}{\sqrt{2}}\frac{\langle +|-\langle -|}{\sqrt{2}} \end{pmatrix} \\
&= \frac{|+\rangle\langle +| + |-\rangle\langle -|}{2}
\end{aligned}$$

basis	$\{\|0\rangle, \|1\rangle\}$	$\{\|+\rangle, \|-\rangle\}$
σ	$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
ρ	$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$

There, completely different matrices. But does that have any operational meaning? For instance, would we get different results if we measured the two states? Let's check. If you measure each state in basis $\{|0\rangle, |1\rangle\}$, the probability of obtaining $|0\rangle$ would be

$$\begin{aligned}
\Pr(0)_\sigma &= \mathrm{Tr}[\ |0\rangle\langle 0|\ \sigma] = \mathrm{Tr}\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\right] = \frac{1}{2}, \\
\Pr(0)_\rho &= \mathrm{Tr}[\ |0\rangle\langle 0|\ \rho] = \mathrm{Tr}\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\right] = \frac{1}{2}.
\end{aligned}$$

This does not look very promising: we could not distinguish the two states solely from the statistics of this measurement. But what happens if instead we measure them in basis $\{|+\rangle, |-\rangle\}$? The probability of obtaining

[2] This proves particularly handy in quantum cryptography.

$|+\rangle$ is, for each of them,

$$\Pr(+)_\sigma = \mathrm{Tr}[\,|+\rangle\langle+|\,\sigma] = \mathrm{Tr}\left[\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}\right] = 1,$$
$$\Pr(+)_\rho = \mathrm{Tr}[\,|+\rangle\langle+|\,\rho] = \mathrm{Tr}\left[\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}\frac{1}{2}&0\\0&\frac{1}{2}\end{pmatrix}\right] = \frac{1}{2}.$$

The states have different measurement statistics in this basis. If you have to find out whether you have fifty copies of $|+\rangle$ or fifty copies of ρ, you can just measure all your systems in basis $\{|+\rangle, |-\rangle\}$. If you get at least one outcome "$-$", then you know that you had state ρ. Question for bonus points: what if you are only given one copy of the state?

Here is another difference between the two states: $|+\rangle$ is a pure state, which means that you can in principle transform it into any other pure state via unitary evolution, while $\rho = \frac{\mathbb{1}}{2}$ is fully mixed, and invariant under unitary transformations. See, for instance, of happens to each state under the following change of basis,

$$U = |+\rangle\langle 0| + |0\rangle\langle+| + |1\rangle\langle-| + |-\rangle\langle 1| = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix},$$
$$U|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = |0\rangle,$$
$$U\rho U^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}\frac{1}{2}&0\\0&\frac{1}{2}\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix} = \begin{pmatrix}\frac{1}{2}&0\\0&\frac{1}{2}\end{pmatrix} = \frac{\mathbb{1}}{2}.$$

7.2.3 Measurements generalised

Imagine that you have a four-dimensional Hilbert space, with basis $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle\}$. You also have have a measurement device, but not a very good one: it only distinguishes the first two from the last two states. It has two measurement outcomes, "$<$" for $|1\rangle$ and $|2\rangle$, and "$>$" for $|3\rangle$ and $|4\rangle$. The probability of obtaining outcome $<$ when measuring a state ρ is given by

$$\begin{aligned}\Pr(<)_\rho &= \Pr(1 \vee 2)_\rho \\ &= \Pr(1)_\rho + \Pr(2)_\rho \\ &= \mathrm{Tr}(|1\rangle\langle 1|\,\rho) + \mathrm{Tr}(|2\rangle\langle 2|\,\rho) \\ &= \mathrm{Tr}\Big(\underbrace{(|1\rangle\langle 1| + |2\rangle\langle 2|)}_{P_<}\,\rho\Big).\end{aligned}$$

We can represent this measurement using projectors ,

$$P_< = |1\rangle\langle 1| + |2\rangle\langle 2|, \qquad \Pr(<)_\rho = \mathrm{Tr}(P_<\,\rho),$$
$$P_> = |3\rangle\langle 3| + |4\rangle\langle 4|, \qquad \Pr(>)_\rho = \mathrm{Tr}(P_>\,\rho).$$

and, not surprisingly, people call it a *projective measurement*. What happens to the system after a measurement with outcome "$<$"? Well, it goes to either $|1\rangle$ or $|2\rangle$, we just don't know which. In other words, it becomes a probabilistic mixture of $|1\rangle$ and $|2\rangle$,

$$\begin{aligned}\rho^< &= \frac{\Pr(1)_\rho|1\rangle\langle 1| + \Pr(2)_\rho|2\rangle\langle 2|}{\Pr(1)_\rho + \Pr(2)_\rho} \\ &= \frac{\mathrm{Tr}(|1\rangle\langle 1|\,\rho)\;|1\rangle\langle 1| + \mathrm{Tr}(|2\rangle\langle 2|\,\rho)\;|2\rangle\langle 2|}{\mathrm{Tr}(P_<\rho)} \\ &= \frac{\langle 1|\rho|1\rangle\;|1\rangle\langle 1| + \langle 2|\rho|2\rangle\;|2\rangle\langle 2|}{\mathrm{Tr}(P_<\rho)} \\ &= \frac{|1\rangle\langle 1|\rho|1\rangle\langle 1| + |2\rangle\langle 2|\rho|2\rangle\langle 2|}{\mathrm{Tr}(P_<\,\rho)} \\ &= \frac{(|1\rangle\langle 1| + |2\rangle\langle 2|)\rho(|1\rangle\langle 1| + |2\rangle\langle 2|)}{\mathrm{Tr}(P_<\rho)} \\ &= \frac{P_<\rho P_<}{\mathrm{Tr}(P_<\rho)}.\end{aligned}$$

This generalizes to any projective measurement with outcomes $\{i\}$ and corresponding projectors $\{P_i\}$,

$$\Pr(i)_\rho = \mathrm{Tr}(\Pr_i\,\rho), \qquad \rho^i = \frac{P_i\rho P_i}{\mathrm{Tr}(\Pr_i\,\rho)}.$$

For further generalizations... I can't emphasize it enough, come to the Quantum Information Theory next semester! 😳