Advanced Quantum Mechanics

Alexander Altland

Contents

Chapter 1

Scattering theory

The chapter begins with a phenomenological introduction to the concept of 'scattering' in quantum theory. Specifically, we will define different paradigms of scattering techniques and introduce the concept of a scattering cross section. We then proceeds to develop the theory of elastic quantum scattering, i.e. scattering without energy exchange. (A few words will be said about the generalization to inelastic scattering.) We will solve for wave functions describing scattering setups, first in formal, then in more concrete terms. Emphasis will be put on the theory of scattering targets possessing rotational symmetry. Specifically, we will study how the scattering properties of such systems can be encoded in a small set of numbers, the so called scattering phase shifts. We will discuss the significane of the phase shifts, notably the information they carry about bound sates supported by attractive scattering potentials. We conclude the chapter with the discussion of a number of general concepts of scattering theory, including the so called scattering operator ('S-matrix') and the physics of time reversal.

Scattering techniques represent one of the most powerful and direct ways to obtain information on the microscopic structure of quantum systems. The importance of the concept to fields such as atomic, nuclear, high energy, or condensed matter physics cannot be exaggerated. We start the chapter with a brief introduction to the setups underlying most scattering experiments.

1.1 Introduction to quantum scattering

1.1.1 Basic concepts

The paradigmatic architecture of a scattering experiment is shown in Fig. 1.1, left: a beam of particles p – either massive (electrons, neutrons, muons, atoms, etc.), or the quanta of electromagnetic radiation – is generated at a source, S, and then directed onto a target. The kinematic information about the source beam is stored in a set of quantum numbers $\mathbf{k}$. For example, $\mathbf{k} \equiv (\mathbf{p}, \sigma)$ may comprise the information on the incoming particle's momentum $\mathbf{p}$, and their spin, σ. However, more fancy data sets are conceivable, too. In much of the following, we will put an emphasis on the case where $\mathbf{k}$ is just a momentum. The particles carry energy

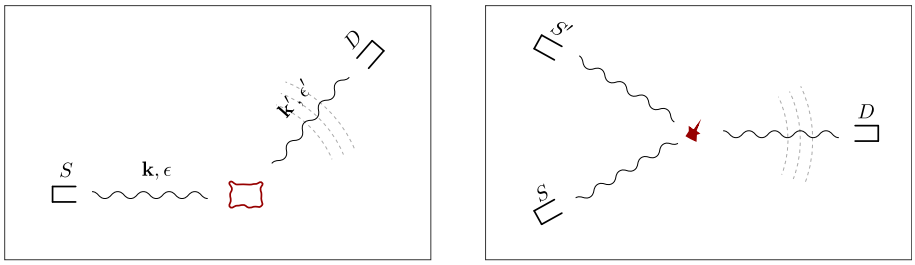


Figure 1.1: Basic setup of scattering experiments. Target scattering (left): a beam of electromagnetic radiation (or massive particles) characterized by a set of quantum numbers $\mathbf{k}$ and energy $\epsilon(\mathbf{k})$ is emitted by a source (S) and directed onto a target sample. The sample responds by emitting radiation according to some distribution $P(\mathbf{k}', \epsilon'(\mathbf{k}')')$, which is, in turn, recorded by a detector (D). Notice that the emitted radiation can, but need not, contain the same type of particles as the source radiation. For example, light quanta may trigger the emission of electrons (photoemission spectroscopy). Collider scattering (right): two particle beams generated by sources S and S', respectively, collide in a scattering region. The detector, placed outside the beam axes, samples the products (particles and/or radiation) generated in the scattering process.

according to some 'dispersion relation' $\epsilon(\mathbf{k})$.[1] The source beam then hits the constituents X of the target to generate a secondary beam of scattered particles p' and modified target particles X'. As with the incoming particles, the target is characterized by quantum numbers $\mathbf{K}$ and a dispersion $E(\mathbf{K})$. Symbolically, we thus have a reaction scheme

$$\begin{array}{ccccccc} p & + & X & \longrightarrow & p' & + & X' \\ \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ \mathbf{k}, \epsilon(\mathbf{k}) & & \mathbf{K}, E(\mathbf{K}) & & \mathbf{k}', \epsilon(\mathbf{k}') & & \mathbf{K}', E'(\mathbf{K}'). \end{array}$$

Notice that the particles p' leaving the sample need not be identical to those incident on the sample. (For example, in photoemission spectroscopy, X-ray quanta displace electrons from the core levels of atoms in a solid. Here p represent the light quanta, and p' electrons.) In such cases, the dispersion ϵ' of the outgoing particles differs from that of the incomings. Further, the dominant scattering process may be elastic (e.g. X-rays scattering off the static lattice structure) or inelastic (e.g. neutrons scattering off phononic excitations). In either case, the accessible information about the scattering process is stored in the frequency–momentum distribution $P(\epsilon(\mathbf{k}'), \mathbf{k}')$ of the scattered particles, as monitored by a detector.

From these data, one would like to restore properties (e.g., the dispersion $E(\mathbf{K})$) of the states inside the target. This is where the detective work of spectroscopy begins. What we know is that the dispersions of the scattered particles and of the sample constituents, $(\mathbf{k}, \epsilon(\mathbf{k}))$

[1] For some sources, e.g. a laser, the preparation of a near-monochromatic source beam of definite $\mathbf{k}$ is (by now) standard. For others, such as neutrons, it requires enormous experimental skills (and a lot of money!).

and $(\mathbf{K}, E(\mathbf{K}))$, respectively, are related through an energy–momentum conservation law, i.e.

$$\begin{aligned} \mathbf{k} + \mathbf{K} &= \mathbf{k}' + \mathbf{K}', \\ \epsilon(\mathbf{k}) + E(\mathbf{K}) &= \epsilon'(\mathbf{k}') + E(\mathbf{K}'). \end{aligned}$$

According to this relation, a "resonant" peak in the recorded distribution $P(\mathbf{k}, \epsilon(\mathbf{k}))$ signals the existence of an internal structure (for example, an excitation, or a lattice structure) of momentum $\Delta\mathbf{K} \equiv \mathbf{K}' - \mathbf{K} = \mathbf{k} - \mathbf{k}'$ and frequency $\Delta E \equiv E' - E = \epsilon - \epsilon'$. However, what sounds like a straightforward recipe in principle may be quite involved in practice: solid state components interact almost exclusively through electromagnetic forces. When charged particles are used as scattering probes, these interactions may actually turn out to be too strong. For example, a beam of electrons may interact strongly with the surface states of a solid (rather than probing its bulk), or the scattering amplitude may be the result of a complicated process of large order in the interaction parameters, and therefore difficult to interpret. The obvious alternative – scattering of neutral particles – is met with its own problems (see below).

Notwithstanding these difficulties, scattering experiments represent one of the most powerful probes of the static and dynamical features of quantum systems. Below, we summarize a number of important sub-categories of scattering setups.

1.1.2 Types of scattering experiments

The following is an incomplete list of the most important families of scattering experiments.

- ▷ **Elastic scattering off immobile targets**: here, the target is effectively so heavy that its mass can be considered infinite. In such cases, scattering is dominantly elastic, $\Delta E = -\Delta\epsilon = 0$. The scattering target can be modelled by some potential distribution, fixed in real space, and arbitrary momentum exchange $\Delta\mathbf{k}$ is possible. Scattering processes of this type form the basis of, e.g., **crystallography**. This is the simplest type of scattering, and we will explore its theory first.

- ▷ **Elastic scattering off general targets**: still no energy exchange, $\Delta E = -\Delta\epsilon = 0$, but we abandon the condition of infinite target mass. **Rutherford scattering**, i.e. the elastic scattering of α-particles off atoms and molecules is an important example.

- ▷ **Inelastic scattering**: when the energies carried by the incident particles become comparable to the energies of excitations of the target, energy exchange becomes an option. The information carried by inelastic scattering data is particularly rich. For example, the existence of quarks in hadronic particles was detected 'indirectly' through the unconventional energy absorption characteristics of the latter in 'deep inelastic' scattering experiments.

- ▷ **Rearrangement scattering**: scattering processes involving an actual 'reaction' of the participating targets. For example, in the collision of two beams of nucleons N_1 and N_2 (Fig. 1.1 right) may be elastic, $N_1 + N_2 \to N_1 + N_2$, or lead to particle production such

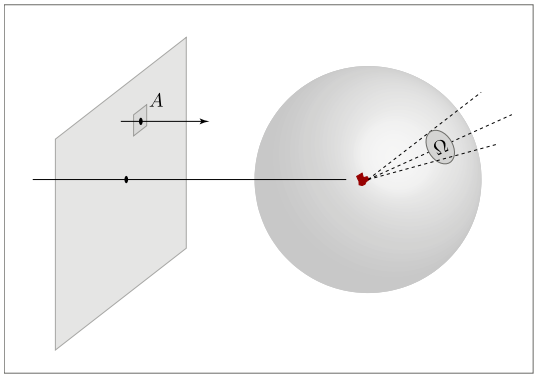
as $N_1 + N_2 \to N_1 + N_2 + \pi$ (pion production), or $N_1 + N_2 \to N_1 + N_2 + K + \bar{K}$ (kaon production).

- ▷ **Resonance scattering**: scattering leads to the formation and subsequent decay of an unstable yet long lived particle. Resonance scattering plays an important role in, e.g., nuclear physics where the scattering of nucleons off heavy nuclei may generate long lived transient states.

Scattering experimentation plays a role in the majority of sub-branches of physics, notably atomic, condensed matter, molecular, nuclear and particle physics. The differences are in the involved energy scales and, relatedly, the cost of scattering experiments. Condensed matter scattering techniques such as Raman or photo emission spectroscopy (meV scales) can be realized in a room sized laboratory and accordingly there are many labs running such types of experiments. By contrast, the few large accelerators of particle physics (TeV scales) cost a fortune. The list of the most renown particle colliders includes:

- ▷ **Fermilab Chicago, Tevatron**: 2 TeV, 6km circumference, 1983. $p + p$ collider, top quark discovery.
- ▷ **Stanford linear accelerator**: 3km, 1966. Electron beam collider, discovery of charm quark, and τ-meson.
- ▷ **Brookhaven, RHIC**: 3.8 km circumference, 2000. Relativistic heavy ion collider, realization of quark gluon plasma.
- ▷ **DESY Hamburg**: ca. 300 GeV, 6.3 km circumference, 1959. e+p and e^++p^+ collider, discovery of gluon.
- ▷ **Large Hadron Collider, LHC:** 27 TeV, 27 km circumference, 2009. World's biggest hadron collider, expect exciting discoveries!

1.1.3 Differential cross section (definition)



One of the most important observables in scattering experiments is called 'differential cross section'. It can be motivated as follows: what a detector will usually record is the number of particles scattered from the target into a certain solid angle. We can characterize this quantity by an angular current *density* j such that the integration over a portion Ω of the unit sphere, $\int_\Omega \sin\theta d\theta d\phi\, j(\theta, \phi)$ gives the current (i.e. the number of particles per unit time) of scattered particles through the angular cone defined by Ω. Also notice that for scattering particles moving radially away from the

target, this quantity does not depend on the radius of the monitoring sphere put around the target. Notice that the physical dimension of the angular current density is given by

$$[j] = \frac{\text{number of particles}}{\text{angle} \times \text{time}}.$$

To give this quantity some meaning, we need to relate it to the incoming current. The latter can be characterized by a current density, i, where we assume the current to be incident along the z-axis. This quantity is defined to give the current (the number of particles per unit time) coming in through an area A in the xy-plane when integrated over A, i.e. $\int_A dxdy\, i(x,y) =$(current through A). For properly prepared particle beams, $i(x,y) \equiv i$ will not depend on the transverse coordinates (x,y). Notice that

$$[i] = \frac{\text{number of particles}}{\text{area} \times \text{time}}.$$

The **differential cross section**, $(d\sigma/d\Omega)$ simply is the ratio between these two quantities:

$$\boxed{\frac{d\sigma}{d\Omega} \equiv \frac{j}{i}.} \tag{1.1}$$

Here, the notation $d\sigma/d\Omega$ is symbolic and does not refer to an actual derivative. Notice that the differential cross section $(d\sigma/d\Omega)(\theta,\phi) = j(\theta,\phi)/i$ depends on angular coordinates. Also notice that

$$[(d\sigma/d\Omega)] = \text{area}.$$

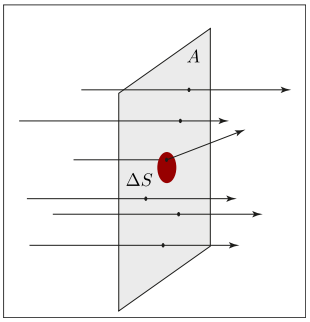
The actual units of area met in a calculation of differential cross sections depend on the physical context. A few typical values are listed in table 1.1.

Let us try to understand the characteristic 'areas' appearing in the table. To this end, we define the **total cross section**

$$\sigma = \int d\Omega\, \frac{d\sigma}{d\Omega}, \tag{1.2}$$

where we used the abbreviation

$$d\Omega \equiv \sin\theta d\theta d\phi$$



for the canonical measure on the sphere. The total cross section relates the current (i.e. the number of scattered particles per unit time) emanating from *an individual scattering center* to the incoming current density. Now, suppose we might attribute a certain geometric area ΔS to our scattering target. Thinking classically, we suspect that incident particles hitting that area will be scattered, i.e. they will contribute to the scattering current sampled by the cross section. Let the total number of particles incident upon an area A in time t be N. The incident current density is then given by $i = \frac{N}{At}$. Of these, a fraction $N\Delta S/A$ will be scattered, i.e. the scattering current is given by $I_{\text{s}} = \frac{N}{t}\frac{\Delta S}{A}$. This means that the total cross section is $\sigma = \frac{I_{\text{s}}}{i} = \Delta S$. We conclude that

Context	characteristic value of $d\sigma/d\Omega$
scattering of visible light	$1\mathrm{cm}^2 = 10^{-4}\mathrm{m}^2$
X-ray scattering	$1\mathring{A}^2 = 10^{-20}\mathrm{m}^2$
neutron scattering off nuclei	$1\mathrm{b}^2 = 10^{-28}\mathrm{m}^2$

Table 1.1: Characteristic values of differential cross sections. The unit 'b' alludes to the phrase 'as big as a barn' (German: 'so groß wie ein Scheunentor'.)

Roughly speaking, the total cross section equals the 'geometric' area of individual scattering centers.

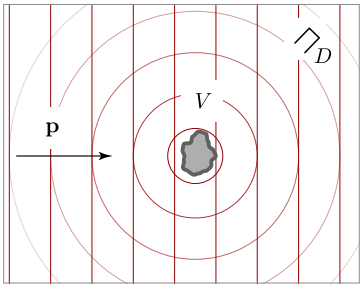
Of course, this simplistic interpretation must be taken with a (major) grain of salt. We certainly have to expect that quantum mechanics will lead to significant modifications to the rigid geometric picture constructed above. However, by order of magnitude, we are doing o.k. Now, as we are going to show below the scattering of particles is the most effective if the wavelength of incident particles is of the order of the extension of the target. This explains, again by order of magnitude, the characteristic values appearing in table 1.1. For example, the wavelength of X-rays is of $\mathcal{O}(1\mathrm{nm})$, which corresponds to areas of $\mathcal{O}(1\mathrm{nm}^2) = \mathcal{O}(10^{-18}\mathrm{m}^2)$.

1.2 General theory of potential scattering

We now turn to the theoretical description of scattering processes. In this section, we will develop the theory of elastic scattering of immobile targets, i.e. the most basic, yet absolutely important setting.[2] In this case, the target may be modeled by a potential distribution V in real space, wherefore one often speaks of **potential scattering**. The basic problem of potential scattering theory is to a) obtain the wave functions describing the scattering particle/target system in such a way that b) boundary conditions corresponding to an *incoming* particle beam are imposed. We begin by constructing the solution of the Schrödinger equation meeting these two criteria in formal terms.

1.2.1 Lippmann-Schwinger equation: derivation

[2] The condition of immobility of the target can, in fact, be dropped if one formulates the theory in the instantaneous rest frame of the target. However, we will not elaborate on this point in explicit terms.



We start out from an Hamilton operator

$$\hat{H} = \hat{H}_0 + \hat{V}, \tag{1.3}$$

where the operator $\hat{H}_0$ describes the dynamics of the incoming particle beam and $\hat{V}$ represents the scattering potential. Without much loss of generality, we may chose

$$\hat{H}_0 \equiv \frac{\hat{p}^2}{2m}, \tag{1.4}$$

i.e. a Hamiltonian describing the free propagation of particles of mass m. However, much of our present introductory discussion below will not depend on this specific choice.

INFO Beware of one point that may cause confusion: the 'natural' eigenstates of the Hamiltonian $\hat{H}_0$ are momentum states $|\mathbf{p}\rangle$ whose real space representation[3]

$$\langle \mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi)^{3/2}} \exp(i\mathbf{x}\cdot\mathbf{p}) \tag{1.5}$$

is infinitely extended in space. This does not sit comfortably with the intuitive picture of localized 'particles' scattering off a target. Formally, a particle is described by a quantum mechanical **wave packet**, i.e. a spatially localized superposition of momentum eigenstates. Once the scattering theory of individual plane wave eigenstates of $\hat{H}_0$ has been developed the scattering of wave packets may, hence, be constructed by superposition principles (cf. M.L. Goldberger and K.M. Watson, *Collision Theory*, Chapter 3 (New York: John Wiley, 1964); R.G. Newton, *Scattering Theory of Waves and Particles*, Chapter 6 (New York: McGraw-Hill, 1966).) However, the complications that come with this after-step can be avoided as long as we restrict ourselves to the consideration of wave packets of (near) definite momentum $\mathbf{p}$ and spatial extension R much larger than the characteristic extension R_V of the scattering region.

The heuristic argument behind this statement goes as follows (cf. Fig. 1.2): as we will see below, the 'interesting' momenta at which actual scattering processes occur are of magnitude $|\mathbf{p}| \sim R_V^{-1}$. To obtain sharply resolved results, we thus need to work with near monochromatic states, whose momentum expectation value is well defined up to some uncertainty $\Delta p \ll R_V^{-1}$. The Heisenberg uncertainty relation $\Delta p \Delta x \geq 1$ then implies a minimum spatial uncertainty of our 'wave packet states' $\Delta x \gg R_V$. As long as we tolerate incoming wave packets much larger than the extension of the scattering center, this degree of uncertainty can be tolerated, i.e. in spite of a tolerance Δx a spatial profile of a wave packet of extension $R \gg \Delta x \gg R_V$ can still be defined. For all practical purposes relating to the description of the scattering of this near monochromatic wave packet of the scattering center it suffices to work with momentum eigenstates $|\mathbf{p}\rangle$.

Our goal is to find a solution $|\psi\rangle$ of the Schrödinger equation

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{1.6}$$

[3] If nothing other is mentioned, we will work in three-dimensional space and set Planck's constant $\hbar$ to unity throughout.

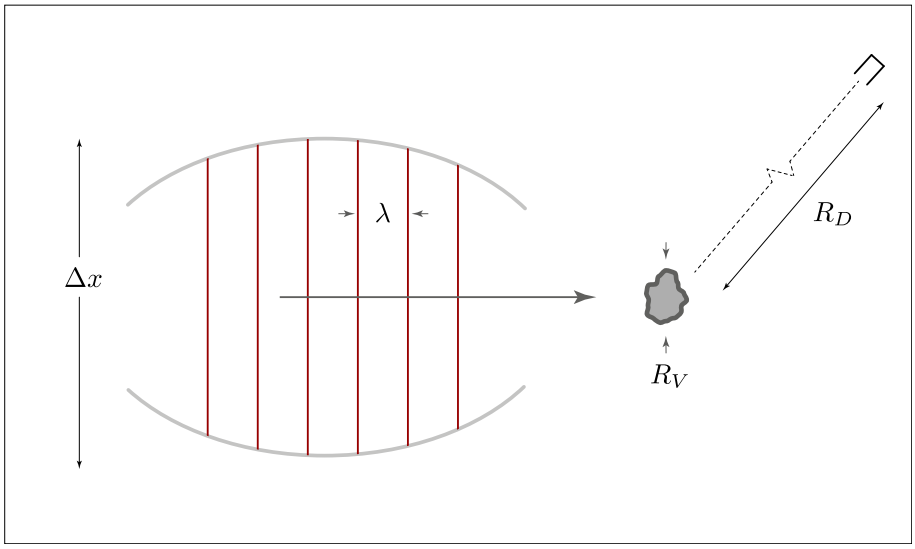


Figure 1.2: On the length scales relevant to a wave package scattering process. Discussion, see text.

such that upon removal of the scattering potential $\hat{V} \to 0$ it smoothly reduces to a solution of the 'unperturbed' equation $\hat{H}_0|\phi\rangle = E|\phi\rangle$ of the same energy.[4] To understand the rational behind this agenda, imagine the scattering process from the perspective of the incoming particles (states). Upon approaching the scattering region, the latter effectively see the emergence (switching on) of a scattering potential, and energy conservation requires that they smoothly evolve into scattering states, i.e. solutions of the full equation at the same energy.

A formal solution to this problem may be proposed as follows:

$$|\psi\rangle = \frac{1}{E - \hat{H}_0}\hat{V}|\psi\rangle + |\phi\rangle, \tag{1.7}$$

where $(E-\hat{H}_0)^{-1}$ is the formal 'inverse' of the operator $(E-\hat{H}_0)$. Acting with $E-\hat{H}_0$ on both sides of the equation, we indeed conclude that (1.7) implies the solution of the Schrödinger equation. Nonetheless, Eq. (1.7) has two serious problems. The first is that both, the left and the right hand side contain the desired solution $|\psi\rangle$, i.e. Eq. (1.7) has the status of an implicit equation which has to be processed, e.g., by recursive means. However, before coming to this complication (which we will be able to get under control) we have to face up to the second problem: the inverse $(E - \hat{H}_0)^{-1}$ does not exist in general. Specifically, if E is an eigenvalue of $\hat{H}_0$ — which clearly *is* a case of physical interest — this operator is not defined.

Now, we may try to be smart and circumvent the above problem as follows: in the original Schrödinger equation (1.6), replace $E \to E^+$, where $E^+ \equiv E + i\delta$, and $\delta > 0$ is an 'infinitesimally small' positive offset. It is tempting to argue that, as long as δ is infinitesimal, this modification is so small, that the solution $|\psi\rangle$ will not be affected. The ansatz (1.7) now

[4] Although it would be natural to designate the eigenstates of the free Hamiltonian (1.4) by $|\mathbf{p}\rangle$, other eigenstates will play a role later on, so we stay with the un-specific designation $|\mathbf{p}\rangle$.

assumes to form of a so-called **Lippmann-Schwinger equation**

$$\boxed{|\psi^+\rangle = \hat{G}_0^+(E)\hat{V}|\psi\rangle + |\phi\rangle,} \tag{1.8}$$

where the so called retarded **Green function (operator)** or **resolvent operator** is defined as

$$\boxed{\hat{G}_0^+(E) = \frac{1}{E^+ - \hat{H}_0}.} \tag{1.9}$$

Thanks to the reality of the spectrum of $\hat{H}_0$, the existence of $\hat{G}_0$ is guaranteed.

EXERCISE Construct an **explicit representation of the Green function** operator. Hint: employ a formal spectral decomposition of $\hat{H}_0$ in terms of its eigenfunctions.

However, the ad hoc manipulation of the Schrödinger equation should leave one with a sense of uneasiness. Specifically, what made us to shift the energy argument into the *positive* half of the complex plane, $E + i\delta$. A negative shift $E - i\delta$ would have worked just as well, but obviously it is not possible to interpolate from one choice to the other once the decision is met, lest the 'dangerous' real axis be crossed.

Before exploring the solution of the Lippmann-Schwinger equation it is well to address these questions.

1.2.2 Retardation

INFO In this text, we will use the **Fourier transform convention**

$$\begin{aligned} f(E) &= \int dt\, e^{iEt} f(t), \\ f(t) &= \int \frac{dE}{2\pi} e^{-iEt} f(E). \end{aligned} \tag{1.10}$$

it will sometimes be important to recall that the Fourier transforms of many functions of physical interest — e.g. oscillatory functions — only exist upon inclusion of an infinitesimal **convergence generating factor**. For instance, the 'proper' interpretation of the first of the Fourier identities in (1.10) reads

$$f(E) = \lim_{\delta \to 0} \int dt\, e^{i(E + i\delta \mathrm{sgn}(t))} f(t).$$

Although the presence of the convergence generating factor is usually not spelled out explicitly, it will be of importance in our discussion below.

Obviously, the sign ambiguity identified above cannot be solved entirely on formal grounds; we need to inject a bit of physical reasoning here. Reassuringly, this will get us to a solution that is satisfactory both from a formal and a conceptual point of view.

The key is to recall the imagined dynamics of a scattering process: in the distant past, incoming states were 'free', i.e. they evolved under the influence of $\hat{H}_0$. We may then imagine the scattering potential to be gradually switched on. As a rule, any interaction with the scattering potential causes an effect only *after* the interaction has taken place. The problem, thus, comes with a sense of **retardation**.

To translate this picture to formal structures, we temporarily turn from an energy to a **time representation**, i.e. we consider the problem of solving the Schrödinger equation (1.6) afresh, this time in a temporal representation: $i\partial_t|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$, which we rewrite as

$$(i\partial_t - \hat{H}_0)|\psi(t)\rangle = \hat{V}|\psi(t)\rangle.$$

Now, interpret this equation as an inhomogeneous linear differential equation in time,

$$(i\partial_t - \hat{H}_0)|\psi(t)\rangle = |\chi(t)\rangle, \tag{1.11}$$

where the inhomogeneity $|\chi(t)\rangle = \hat{V}|\psi(t)\rangle$ is a time dependent element of Hilbert space. (However, the Hilbert space structure will be of secondary importance for the moment, we focus on time dependence.)

EXERCISE Recall what you remember from the method of Green functions in electrodynamics and compare the following construction to that theory.

We may now solve this equation by a procedure similar to the Green function method (!) of the theory of linear differential equations. That is, we first seek for a solution of the differential equation with a particularly simple inhomogeneity, viz. a temporal δ-function, and a unit operator $\mathbb{I}$ in Hilbert space:[5]

$$(i\partial_t - \hat{H}_0)\hat{G}_0(t,t') = \mathbb{I}\delta(t-t'), \tag{1.12}$$

Notice that $\hat{G}_0$ is an operator in Hilbert space, i.e. an object that can be applied to Hilbert states to generate other states. We now claim that the solution of the equation (1.11), is given by

$$|\psi(t)\rangle = \int dt'\, \hat{G}_0(t,t')|\chi(t')\rangle + |\phi(t)\rangle,$$

where $|\phi(t)\rangle$ solves the homogeneous equation $(i\partial_t - \hat{H}_0)|\phi(t)\rangle = 0$. All we have to do to check this assertion is apply the operator $(i\partial_t - \hat{H}_0)$ to the left and the right hand side of the equation and use (1.12). Notice that the real space representation of the solution $\psi(\mathbf{x},t) \equiv \langle\mathbf{x}|\psi(t)\rangle$ is obtained as

$$\langle\mathbf{x}|\psi(t)\rangle = \int dt' \int d^3x'\, \langle\mathbf{x}|\hat{G}_0|\mathbf{x}'\rangle(t,t')\langle\mathbf{x}'|\chi(t')\rangle,$$

where $\langle\mathbf{x}|\hat{G}_0|\mathbf{x}'\rangle(t-t') \equiv G_0(\mathbf{x},\mathbf{x}';t,t')$ defines the matrix elements of the Green function operator.

[5] In a real space representation, $\mathbb{I}(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')$.

Thus, all we need to do now to get the inhomogeneous equation under control is to obtain a solution of (1.12) for the Green function. Indeed, it is straightforward to check that the so-called **retarded Green function**

$$\boxed{\hat{G}_0^+(t,t') \equiv -i\Theta(t-t')e^{-i\hat{H}_0(t-t')}} \tag{1.13}$$

solves the equation. To see this, just note that for $t \neq t'$ the relation $i\partial_t \exp(-i\hat{X}t) = \hat{X}\exp(-i\hat{X}t)$, where $\hat{X}$ may be any operator, ensures the solution property. For t approaching t', the identity $\partial_t\Theta(t) = \delta(t)$ generates the δ-distribution on the r.h.s. of the equation. (Exercise: think more about these points!)

Now, $\hat{G}_0^+$ is not the only solution of the differential equation. Any function differing from $\hat{G}_0^+$ by the addition of a constant multiple of the homogeneous solution $\exp(-i\hat{H}_0(t-t'))$ is another solution. For example, the operator function

$$\hat{G}_0^-(t,t') \equiv +i\Theta(t'-t)e^{-i\hat{H}_0(t-t')},$$

the so-called **advanced Green function** solves the equation as well.

The question then is, which solution is the 'right one' and it is here where physics enters the stage. Consider the solution, adapted to the 'real' inhomogeneity $|\chi\rangle = \hat{V}|\psi\rangle$,

$$|\psi(t)\rangle = \int dt'\, \hat{G}_0(t,t')\hat{V}|\psi(t')\rangle + |\phi(t)\rangle.$$

Apparently, the Green function 'kernel' mediates the impact of scattering events at time t' onto the wave function at times t. In the light of our discussion above, this response must be **retarded**, i.e. the time of the cause, t' must precede that of the effect, t. Denoting the retarded scattering solution by $|\psi^+(t)\rangle$[6] we thus have

$$\boxed{|\psi^+(t)\rangle \equiv \int dt'\, \hat{G}_0^+(t,t')\hat{V}|\psi^-(t')\rangle + |\phi(t)\rangle.} \tag{1.15}$$

The fact that $\hat{G}_0^+(t,t') = \hat{G}_0^+(t-t')$ depends only on the difference of its arguments implies that (1.15) has the status of a convolution. This suggests that now is a good time to pass back to the Fourier inverse language of the energy representation. To this end, imagine that the scattering states carry an oscillatory time dependence,

$$|\psi^+(t)\rangle = e^{iEt}|\psi^+(E)\rangle,$$

[6] The notation distinguishes $|\psi^+\rangle$ from an **advanced scattering state**

$$|\psi^-(t)\rangle \equiv \int dt'\, \hat{G}_0^-(t,t')\hat{V}|\psi^-(t')\rangle + |\phi(t)\rangle \tag{1.14}$$

describing an anti-causal effect of the scattering potential. Although unphysical, the advanced states play a formal role in the construction of elements of scattering theory.

i.e. are *eigenstates* at some definite energy E. The Fourier convolution theorem (or a short explicit check) then leads equation

$$|\psi^+(E)\rangle = G_0^+(E)\hat{V}|\psi^+(E)\rangle + |\phi(E)\rangle, \tag{1.16}$$

in which we rediscover the Lippmann-Schwinger equation (1.8). Here, the Fourier representation of the Green function is defined as

$$\hat{G}_0^+(E) = \int_{-\infty}^{\infty} dt\, e^{it(E+i\delta\mathrm{sgn(t)})}\hat{G}_0^+(t) = -i\int_0^{\infty} dt\, e^{i(E^+-\hat{H}_0)t} = \frac{1}{E^+ - \hat{H}_0}. \tag{1.17}$$

These final equation contain the solution to the problems raised in the end of the previous section. Our main conclusions are:

- ▷ The operator $\hat{G}_0^+$ 'naturally' appears upon application of Green function methodology to the Schrödinger equation, the latter understood as an inhomogeneous linear differential equation.
- ▷ Once we use a qualified (means, existing) variant of the Fourier transform, infinitesimally shifted energy arguments as in (1.9) are generated automatically, where
- ▷ the sign of the shift follows from causality, i.e. the condition that the scattering potential affect the scattering states, *after* the scattering process has taken place.

EXERCISE Readers familiar with the **theory of complex functions** may find it instructive to compute $G^+(t)$ by inverse Fourier transform of $\hat{G}^+(E)$. It then becomes very apparent how the infinitesimal shift $i\delta$ causes a large effect, viz. by determining the retardation condition $\Theta(t)$.

1.2.3 Lippmann-Schwinger equation: formal solution

Let us try to construct a 'perturbative' solution of the Lippmann-Schwinger equation (1.7). For a weak scattering potential, $|\psi^+\rangle = |\phi\rangle + \mathcal{O}(V)$. Ignoring contributions of order V^2, we may substitute $|\psi^+\rangle \simeq |\phi\rangle$ on the r.h.s. of the equation to obtain the approximate solution

$$|\psi^+\rangle^{(1)} \equiv \hat{G}_0^+\hat{V}|\phi\rangle + |\phi\rangle, \tag{1.18}$$

where we omitted the energy argument in G_0 for notational clarity. This procedure may be iterated. A solution accurate to $\mathcal{O}(V^2)$ is obtained by substitution of $|\psi^+\rangle^{(1)}$ on the r.h.s.:

$$|\psi^+\rangle^{(2)} \equiv \hat{G}_0^+\hat{V}\left(\hat{G}_0^+\hat{V}|\phi\rangle + |\phi\rangle\right) + |\phi\rangle = \left(\hat{G}_0^+\hat{V}\hat{G}_0^+\hat{V} + \hat{G}_0^+\hat{V} + \mathbb{I}\right)|\phi\rangle.$$

What we observe here is a (geometric) series emerging. Iteration of the program evidently obtains the solution

$$|\psi^+\rangle = \sum_{n=0}^{\infty}\left(\hat{G}_0^+\hat{V}\right)^n|\phi\rangle = \frac{1}{\mathbb{I} - \hat{G}_0^+\hat{V}}|\phi\rangle. \tag{1.19}$$

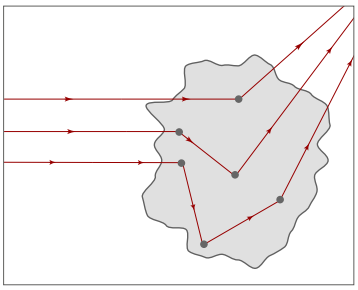
INFO Notice that this relation relies on two non-trivial presumptions at once: the existence of the series, and its re-summability in terms of a geometric series. These conditions have a status far beyond that of mathematical subtleties: many 'real life' scattering problems, can't be solved in terms of infinite series in the scattering potential. The best we can formulate are so-called **asymptotic series**, i.e. series approximating the actual solution of the Schrödinger equation by a finite number of terms (where the accuracy of the approximation is inversely correlated with the strength of the scattering potential.) Likewise, the re-summability of the series has to be critically checked in each problem anew.

Eq. (1.19) may also be formulated as

$$|\psi^+\rangle = \frac{1}{(\hat{G}_0^+)^{-1} - \hat{V}}(\hat{G}_0^+)^{-1}|\phi\rangle = \frac{i\delta}{E^+ - \hat{H}}|\phi\rangle \equiv i\delta\hat{G}^+|\phi\rangle, \tag{1.20}$$

where we defined the **full Green function** of the problem as

$$\boxed{\hat{G}^+(E) \equiv \frac{1}{E^+ - \hat{H}}.}$$



Their formal nature notwithstanding, the identities above shed some light on both, the physics of the scattering process and the physical meaning of Green functions (cf. the figure.) Eq. (1.19) expresses the retarded solution of the scattering problem in terms of a coherent superposition of terms of ascending order in the number of scattering potential 'interactions'. Each term $\hat{V}$ is accompanied by an operator $\hat{G}_0^+$. This suggests an interpretation, where $\hat{V}$ assumes the role of the scattering interaction (naturally), while $\hat{G}_0^+$ 'propagates' the wave function in-between scattering events. The superposition of all these processes then adds to the full scattering state.

The interpretation of $\hat{G}_0^+$ as a carrier of the free quantum dynamics in-between scattering events can ba made more rigorous: imagine you had prepared a quantum state to sit at position $\mathbf{x}$ at time $t = 0$. (Think of $\mathbf{x}$ as the coordinate of a scattering event at time $t = 0$.) According to the laws of quantum mechanics, the amplitude for this state evolving to a coordinate $|\mathbf{x}'\rangle$ in time t is given by

$$A_{\mathbf{x}\to\mathbf{x}'}(t) = \langle\mathbf{x}'|e^{i\hat{H}_0 t}|\mathbf{x}\rangle\Theta(t),$$

where the Θ-function has been added to underpin that the propagation has to be in positive time. It is straightforward to verify that this amplitude obeys the differential equation (1.12), i.e.

$$A_{\mathbf{x}\to\mathbf{x}'}(t) = \langle\mathbf{x}'|\hat{G}_0^+(t)|\mathbf{x}\rangle = G_0^+(\mathbf{x}', \mathbf{x}; t).$$

Alluding to this interpretation the Green function $\hat{G}_0^+$ is often called the **(retarded) propagator** of the theory.

We finally note that it is sometimes useful to condense the multiple scattering processes introduced above into a **compact notation**. We simply declare that the 'complicated object'

$\hat{V}|\psi^+\rangle$ appearing in the right hand side of the Lippmann-Schwinger equation be identical to an operator $\hat{T}$ acting on the zeroth order state $|\phi\rangle$:

$$\hat{V}|\psi^+\rangle \equiv \hat{T}|\phi\rangle, \tag{1.21}$$

which is the definition of the so-called **transition operator**, $\hat{T}$. Comparison with Eq. (1.19) then leads to the series representation

$$\hat{T} = \hat{V} + \hat{V}\hat{G}_0^+\hat{V} + \hat{V}\hat{G}_0^+\hat{V}\hat{G}_0^+\hat{V} + \ldots,$$

which can be 'resummed' in different ways. For example,

$$\hat{T} = \hat{V} + \hat{V}\hat{G}^+\hat{V}, \tag{1.22}$$

where $\hat{G}^+$ is the full Green function, or the implicit representation

$$\hat{T} = \hat{V} + \hat{V}\hat{G}_0^+\hat{T}. \tag{1.23}$$

1.2.4 Differential cross section (computation)

EXERCISE Recapitulate the concept of electromagnetic radiation, notably the notion of 'radiation in the far zone'.

We next aim to go beyond the level of formal solutions, and to this end we need explicit representations of the Green function. In view of the fact that the scattering potential will usually be given in terms of a real space function, it is preferable to work in real space representations.

There are different ways (see the info block below) to conclude that the **real space representation of the retarded Green function** is given by

$$\boxed{G_0^+(\mathbf{x},\mathbf{x}') \equiv \langle \mathbf{x}|\hat{G}_0^+|\mathbf{x}'\rangle = -\frac{m}{2\pi}\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|},} \tag{1.24}$$

where

$$k \equiv \sqrt{2mE}$$

and we have once more omitted the energy argument for notational brevity.

INFO The most cost-efficient way to verify (1.24) is to observe that the real space representation of the defining equation (1.12), $\hat{G}_0^+ = (E^+ - \hat{H}_0)^{-1} \Leftrightarrow (E^+ - \hat{H}_0)\hat{G}_0^+ = \mathbb{I}$ is given by

$$\left(E^+ + \frac{1}{2m}\partial_{\mathbf{x}}^2\right) G_0^+(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}'),$$

which we identify as $\frac{1}{2m}\times$ the Helmholtz equation of electrodynamics. Copying from that theory, we readily get to (1.24). Readers preferring not to leave the realm of quantum mechanics may insert a resolution of unity

$$\int d^3p\, |\mathbf{p}\rangle\langle\mathbf{p}| = \mathbb{I}$$

in terms of $\hat{H}_0$'s eigenfunctions into the definition of (1.24) $G_0^+(\mathbf{x},\mathbf{x}')$ and use (1.5) to obtain

$$G_0^+(\mathbf{x},\mathbf{x}') = \int \frac{d^3p}{(2\pi)^3}\,\langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}|\mathbf{x}'\rangle\frac{1}{E^+-\epsilon_p} = \int \frac{d^3p}{(2\pi)^3}\,\frac{e^{i(\mathbf{x}-\mathbf{x}')\cdot\mathbf{p}}}{E^+-\epsilon_p},$$

where $\epsilon_p = \frac{p^2}{2m}$. Switching to polar coordinates with z-axis aligned with $\mathbf{x}-\mathbf{x}'$, this becomes

$$\begin{aligned} G_0^+(\mathbf{x},\mathbf{x}') &= \frac{1}{(2\pi)^2}\int_0^\infty p^2dp\int_{-1}^1 d\cos\theta\,\frac{e^{i|\mathbf{x}-\mathbf{x}'|p\cos\theta}}{E^+-\epsilon_p} = \\ &= \frac{1}{(2\pi)^2 i|\mathbf{x}-\mathbf{x}'|}\int_0^\infty pdp\frac{e^{i|\mathbf{x}-\mathbf{x}'|p}-e^{-i|\mathbf{x}-\mathbf{x}'|p}}{E^+-\epsilon_p} = \frac{1}{(2\pi)^2 i|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^\infty pdp\frac{e^{i|\mathbf{x}-\mathbf{x}'|p}}{E^+-\frac{p^2}{2m}}. \end{aligned}$$

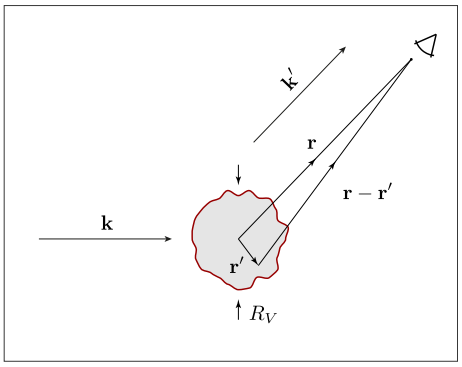
The final one-dimensional integral is tabulated and the result readily gets us to (1.24). (Readers familiar with the theory of complex functions may find it instructive to compute the integral by the theorem of residues.)

Scattering amplitude

Let us, then, consider the Lippmann-Schwinger equation for the retarded scattering state, Eq. (1.7), in a real space representation:

$$\begin{aligned} \psi(\mathbf{x}) &= \int d^3x'\,G_0^+(\mathbf{x},\mathbf{x}')V(\mathbf{x}')\psi(\mathbf{x}') + \phi(\mathbf{x}) = \\ &= -\frac{m}{2\pi}\int d^3x'\,\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}V(\mathbf{x}')\psi(\mathbf{x}') + \phi(\mathbf{x}), \end{aligned} \tag{1.25}$$

where we wrote $\psi(\mathbf{x}) \equiv \psi^+(\mathbf{x}) = \langle\mathbf{x}|\phi^+\rangle$ for notational simplicity. Now, in general the observation point $\mathbf{x}$ at which we wish to read out the scattering amplitude $\psi(\mathbf{x})$ will in general be far from the scattering region, $|\mathbf{x}| \gg R_v$ (cf. the figure.) Choosing coordinates where $\mathbf{x}' = 0$ defines the center of the target, we thus have $|\mathbf{x}| \gg |\mathbf{x}'|$, and this justifies the expansion



$$|\mathbf{x}-\mathbf{x}'| \simeq r - \frac{\mathbf{x}\cdot\mathbf{x}'}{r}, \qquad r \equiv |\mathbf{x}|.$$

Defining

$$\mathbf{k}' \equiv k\frac{\mathbf{x}}{r},$$

i.e. a momentum vector equal in magnitude to the incoming momentum but directed towards the observation point, we may thus approximate

$$\begin{aligned} \psi(\mathbf{x}) &= \frac{1}{(2\pi)^{3/2}}\left(e^{i\mathbf{k}\cdot\mathbf{r}} - (2\pi)^{1/2}m\frac{e^{ikr}}{r}\int d^3x'\,e^{-i\mathbf{k}'\cdot\mathbf{r}'}V(\mathbf{x}')\psi(\mathbf{x}')\right) \equiv \\ &\equiv \frac{1}{(2\pi)^{3/2}}\left(e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{e^{ikr}}{r}f(\mathbf{k},\mathbf{k}')\right), \end{aligned} \tag{1.26}$$

(think what justifies the approximation $1/|\mathbf{x}-\mathbf{x}'| \simeq 1/r$ in the pre-exponential term of the integral in (1.25)) where we defined the **scattering amplitude**

$$\boxed{f(\mathbf{k},\mathbf{k}') \equiv -(2\pi)^{1/2} m \int d^3x'\, e^{-i\mathbf{k}'\cdot\mathbf{r}'} V(\mathbf{x}')\psi(\mathbf{x}') = -(2\pi)^2 m \langle \mathbf{k}'|\hat{V}|\psi^+\rangle,} \tag{1.27}$$

an object of physical dimension $[f]$=length. The designation 'scattering amplitude' is easy enough to understand: Eq. (1.25) states that the wave function assumes the form of a superposition of a planar incoming wave (the first term) and a scattered *spherical* wave (the second term), i.e. a wave whose amplitude decays as $\sim r^{-1}$ away from the scattering center. The quantity $f(\mathbf{k},\mathbf{k}')$ weighs the scattered contribution as a complex 'scattering amplitude'. Don't be fooled by the seeming simplicity of these expressions: the scattering amplitude may contain complex angular dependence, encoded in its dependence on (the angle between) $\mathbf{k}$ and $\mathbf{k}'$. Finally, keep in mind that the scattering amplitude contains the sought for solution $|\psi\rangle$ as an argument, i.e. Eq. (1.25) remains an implicit equation. We may conceal this fact by writing $\hat{V}|\hat{\psi}^+\rangle = \hat{T}|k\rangle$, where $\hat{T}$ is the transition operator defined in (1.21). This leads to the compact representation

$$\boxed{f(\mathbf{k},\mathbf{k}') = -(2\pi)^2 m \langle \mathbf{k}'|\hat{T}|\mathbf{k}\rangle,} \tag{1.28}$$

which expresses the scattering amplitude as a matrix element of the transition operator between the bare incoming and the outgoing state. The series representation of the transition operator makes the multiple-scattering process underlying the formation of the scattering amplitude manifest.

The scattering amplitude really does carry the essential information about the scattering process. To make this observation more apparent, let us relate the scattering amplitude to the differential cross section introduced in section 1.1.3.

To this end, recall the definition of the current density carried by a state ψ,

$$\mathbf{j}(\mathbf{x}) = \frac{1}{2mi}\left(\bar{\psi}(\mathbf{x})(\nabla\psi)(\mathbf{x}) - (\nabla\bar{\psi})(\mathbf{x})\psi(\mathbf{x})\right). \tag{1.29}$$

In our present context, the incoming current density $j_{\rm in} = i\mathbf{e}_z$ is carried by the plane wave contribution to ψ. Application of the formula above readily gives

$$i = \frac{k}{(2\pi)^3 m}.$$

The radial component of the scattered current density is given by $\mathbf{e}_r \cdot \mathbf{j}_{\rm out} = (j/r^2)$, where we noted that the quantity j introduced in (1.1) is the current per angular element $d\Omega = \sin\theta d\theta d\phi$ which differs from the surface element $r^2 d\Omega$ to which $\mathbf{e}_r \cdot \mathbf{j}_{\rm out}$ relates by a factor r^2 (a point to think about!). Noting that $\mathbf{e}_r \cdot \nabla = \partial_r$, we thus obtain

$$j = \frac{r^2}{2mi(2\pi)^3}|f(\mathbf{k},\mathbf{k}')|^2\left[\frac{e^{-ikr}}{r}\partial_r\frac{e^{ikr}}{r} - \text{c.c.}\right] = \frac{|f(\mathbf{k},\mathbf{k}')|^2}{(2\pi)^3 m}k(1+\mathcal{O}(1/rk)).$$

Anticipating that $k \sim R_V^{-1}$ will typically of the order of the inverse extension of the scattering region, we neglect the correction of $\mathcal{O}(R_V/r)$ and obtain the important formula

$$\boxed{\frac{d\sigma}{d\Omega} \simeq |f(\mathbf{k}, \mathbf{k}')|^2.} \tag{1.30}$$

This shows that, truly, the scattering amplitude is our central object of interest.

Optical theorem

EXERCISE Recall the mathematics of the **Dirac identity**

$$\lim_{\delta \searrow 0} \frac{1}{r + i\delta} = P\left(\frac{1}{r}\right) - i\pi\delta(r), \tag{1.31}$$

where r is real and $P(\dots)$ stands for the principal value.

Intuitively, we expect a scattering target to diminish the intensity of the incoming stream of particles. The flux removed from the 'forward direction' will be scattered into other directions. The unitarity (flux conservation) of quantum mechanics requires that the attenuation of the incoming flux must balance against the total flux scattered. The mathematical formulation of this statement is the subject of the 'optical theorem'.[7] The **optical theorem** states that

$$\boxed{\mathrm{Im} f(\mathbf{k}, \mathbf{k}) = \frac{k\sigma}{4\pi},} \tag{1.32}$$

where σ is the total cross section defined in Eq. (1.2). The l.h.s. of the optical theorem contains the scattering amplitude in forward direction, while the r.h.s. makes a statement about the integrated scattering current, in line with the intuition formulated above.

INFO The optical theorem belongs to a class of formulae generally known as **sum rules**. A sum rule relates the total integral ('sum') of some quantity to another expression, often involving the same quantity. For example, the celebrated Kramers-Kronig relations of electrodynamics are sum rules, too. Sum rules are usually derived in rigorous terms (see below) and hold at the full level of generality of a theory. While they do not have much predictive power by themselves, they play an important role as consistency checks. E.g. an experimentor may measure the forward scattering amplitude, and compare to the total cross sections. The two quantities have to be equal and rigorously (no approximation involved) so. If they turn out to be different, the experiment was faulty – a consistency check.

To prove the **prove the optical theorem**, we use the definition (1.27) and the conjugate of the Lippmann-Schwinger equation $\langle \mathbf{k}| = \langle \psi^+| - \langle \psi^+|\hat{V}\hat{G}_0^-$ to obtain

$$\mathrm{Im} f(\mathbf{k}, \mathbf{k}) = -(2\pi)^2 m \,\mathrm{Im}\, \langle \psi^+|\hat{V}\hat{G}_0^-\hat{V}|\psi^+\rangle,$$

[7] E. Feenberg, *The Scattering of Slow Electrons by Neutral Atoms*, Phys. Rev. **40**, 40 (1932).

where the noted that $\mathrm{Im}\,\langle\psi^+|\hat{V}|\psi^+\rangle = 0$ (why?). We now insert a resolution of unity $\mathbb{I} = \int d^3k'\,|\mathbf{k}'\rangle\langle\mathbf{k}'|$ into this expression and use $\hat{G}_0^-|\mathbf{k}'\rangle = (E^- - \epsilon_{k'})^{-1}|\mathbf{k}'\rangle$, where $\epsilon_{k'} = k'^2/2m$. Using a polar measure $d^3k' = k'^2 dk' d\Omega'$ for the $\mathbf{k}'$-integration, we arrive at

$$\begin{aligned}\mathrm{Im} f(\mathbf{k},\mathbf{k}) &= (2\pi)^2 m \int k'^2 dk' \int d\Omega'\,\langle\psi^+|\hat{V}|\mathbf{k}'\rangle\langle\mathbf{k}'\hat{V}|\psi^+\rangle\;\mathrm{Im}\,\frac{1}{E^- - \epsilon_{k'}} = \\ &= (2\pi)^2\pi m^2 k \int d\Omega'\,\langle\psi^+|\hat{V}|\mathbf{k}'\rangle\langle\mathbf{k}'\hat{V}|\psi^+\rangle = \\ &= \frac{k}{4\pi}\int d\Omega'\,|f(\mathbf{k},\mathbf{k}')|^2 = \frac{k}{4\pi}\int d\Omega\,\frac{d\sigma}{d\Omega},\end{aligned}$$

where in getting from the first to the second line we used $\mathrm{Im}\,(E^- - \epsilon_{k'})^{-1} = \pi\delta(E - \epsilon_{k'}) = \pi\delta(k - k')/|\partial_{k'}\epsilon_{k'}| = (m/k)\delta(k - k')$.□

1.2.5 Born approximation

Imagine the scattering of particles weakly interacting with their target. In this case, we might multiple scattering is of lesser importance, and an approximation of the scattering states to first order in the scattering potential might be accurate enough. Formally, this approximation is obtained by setting $|\psi\rangle \simeq |\phi\rangle$ in the integral defining the scattering amplitude (1.27). We thus obtain the approximate expression

$$\boxed{f(\mathbf{k},\mathbf{k}') \simeq -\frac{m}{2\pi}\int d^3x\, e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} V(\mathbf{x}).} \tag{1.33}$$

Eq. (1.33) is known as the **Born approximation**. Inspection of the formula shows that

> The Born approximation of the scattering amplitude is given by the Fourier transform of the scattering potential.

Contrary to the implicit results above, this *is* an explicit, if approximate, expression for the scattering amplitude. In many applications, a first order Born approximation will be the first thing to try. Only if it fails will one move on to more accurate schemes.

INFO For the important class of **centro-symmetric scattering potentials**, $V(\mathbf{x}) = V(r)$, where $r = |\mathbf{x}|$, the Born scattering amplitude can be simplified as

$$\begin{aligned}f(\mathbf{k},\mathbf{k}') &= -\frac{m}{2\pi}\int_0^\infty r^2 dr \int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta)\, e^{-iqr\cos\theta} V(r) = \\ &= -\frac{2m}{q}\int_0^\infty r dr\,\sin(qr) V(r),\end{aligned} \tag{1.34}$$

where we introduced the momentum transfer

$$\mathbf{q} \equiv \mathbf{k} - \mathbf{k}', \quad q \equiv |\mathbf{q}|.$$

The goal of any scattering experiment is to learn as much as is possible about strength and spatial structure of its scattering target. The latter is encoded in the dependence of the scattering amplitude on the *angle* between incident and scattered momentum, $\mathbf{k}$ and $\mathbf{k}'$, resp. Having this principle in mind, let us explore the behavior of the Born scattering amplitude in different physical limits. Specifically, tet us now try to find out, what scattering probes (incident particles) are optimally suited to reveal a maximum of information on the scattering potential. Our scattering setup comes with two intrinsic length scales, viz. the extension of the scattering potential R_V and the wavelength of the incoming particles $\lambda \equiv 2\pi/k$.

In the **scattering of long wavelength particles**, $\lambda/V_R \ll 1$, we may approximate $e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} \simeq 1$ under the integral. The scattering amplitude then collapses to

$$f(\mathbf{k},\mathbf{k}') \simeq -\frac{m}{2\pi}\int d^3x\, V(\mathbf{x}),$$

which contains information about the integrated strength, but not the spatial structure of the target.

In the **scattering of short wavelength particles**, $\lambda/V_R \gg 1$, the oscillatory phase in the Born integral tends to oscillate to zero (at least, if the variation of the scattering potential over the support V_R is smooth.) Roughly, we may approximate

$$|f(\mathbf{k},\mathbf{k}')| \sim mV_0\lambda^3 \sim \lambda(V_0/E_k),$$

where $E_k = k^2/2m$ is the kinetic energy of the incoming particles, and V_0 is the typical amplitude of the scattering potential. In the limit of asymptotically large energies, the scattering region becomes 'invisible' to the incident particles, they just have too high energy to care about the scattering potential. For finite, yet small $\lambda \ll V_R$, we expect the scattering process to become dominantly classical. In the parlor of optics, this is the limit of 'geometric optics' where the wave lengths are small in comparison to the spatial variation of the scattering region, and incident waves may be treated as straight 'rays'. The wave nature becomes irrelevant and quantum mechanics is expected to collapse to its classical limit.

An optimal compromise between the above two scenarios is, thus, reached at $\lambda \sim R_V$. To resolve the quantum nature of a scattering process, we should, thus,

> choose the wavelength of the incoming particles so as to be comparable with the scales of variation of the scattering target.

In this case, we may estimate

$$|f(\mathbf{k},\mathbf{k}')| \sim mV_0R_V^3 \sim R_V(V_0/E_0),$$

where $E_0 = (2\pi/R_V)^2/2m$ is the minimum energy of a particle confined to regions of extension $\sim R_V$ due to quantum uncertainty. The scattering cross section $|d\sigma/d\Omega| \sim R_V^2(V_0/E_0)^2$ is then roughly given by the geometric area of the target, $\sim R_V^2$ times the dimensionless ratio of the two scales 'scattering strength', R_V over 'reference energy', E_0.

1.2.6 Examples

Let us illustrate the concepts introduced above on a few elementary examples. The simplest choice of a scattering potential, surely is a

δ-function potential. Consider the potential $V(\mathbf{x}) = \gamma\delta(\mathbf{x})$. Notice that a potential of infinitely sharp concentration in space would not be 'seen' by impinging classical point particles. However, quantum mechanics is notorious for its smearing of spatial structures and does generate a cross section whose scattering amplitude we calculate in Born approximation as

$$f(\mathbf{k},\mathbf{k}') = -\frac{m}{2\pi}\int d^3x\, e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}\gamma\delta(\mathbf{x}) = -\frac{m\gamma}{2\pi}.$$

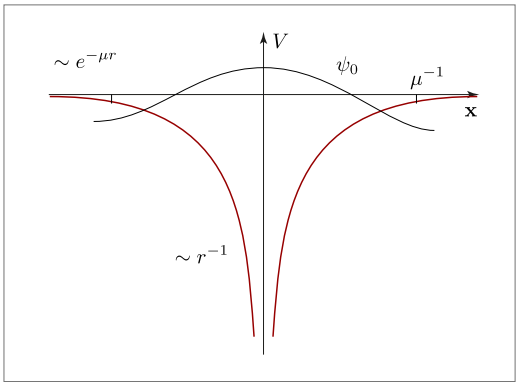
Thus

$$\frac{d\sigma}{d\Omega} = \left(\frac{m\gamma}{2\pi}\right)^2$$

turns out to be independent of angles, i.e. the scattering off a δ-potential is purely isotropic.

Yukawa potential. Let us now consider the somewhat more rewarding example of the **Yukawa potential**

$$V(\mathbf{x}) = V_0\frac{e^{-\mu r}}{\mu r}, \qquad r = |\mathbf{x}|. \tag{1.35}$$



The Yukawa potential enjoys appears in many physical applications. At distance scales smaller than μ^{-1}, $V(\mathbf{x}) \overset{\mu r<1}{\simeq} V_0/(\mu r)$ it asymptotes to the celebrated Coulomb potential. However, unlike with the Coulomb potential it is not spatially long ranged, rather it decays exponentially $V(\mathbf{x}) \overset{\mu r>1}{\sim} \exp(-\mu r)$. In this respect, it resembles many 'real' scattering potentials whose extent is cut off by instances of screening.

Using (1.34), the Born scattering amplitude of the Yukawa potential is calculated as

$$f(\mathbf{k},\mathbf{k}') = -\frac{2mV_0}{q\mu}\int_0^\infty dr\, \sin(qr)e^{-\mu r} = -\frac{2mV_0}{\mu(\mu^2+q^2)}.$$

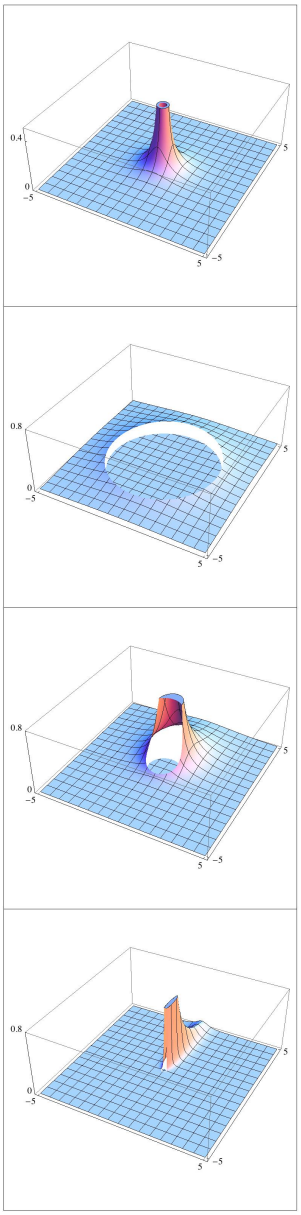
This formula exemplifies the general principles discussed in the end of the foregoing section. For $q \ll \mu \sim R_V^{-1}$, the scattering amplitude $f \simeq -2mV_0/\mu^3$ becomes isotropic, for $q \gg \mu$ it asymptotes to zero, $f \sim q^{-2}$. For fixed $q \gg \mu$, we have

$$f(\mathbf{k},\mathbf{k}') \overset{q\gg\mu}{\simeq} -\frac{2mV_0}{\mu q^2}.$$

Noting that $q^2 = |\mathbf{k}-\mathbf{k}'|^2 = 2k^2(1-\cos\theta) = 4k^2\sin^2(\theta/2)$, this leads to

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{V_0}{4\mu E_k}\right)^2 \frac{1}{\sin^4(\theta/2)}.$$

This equals the celebrated **Rutherford scattering** formula for the scattering of charged particle — at high frequencies, quantum mechanics is gone, in line with our general expectations.



Yukawa scattering provides a testbed for the **limitations of the Born approximation**. Recall that the latter was based on an approximation $|\psi\rangle \simeq |\phi\rangle$ under the integral of the Lippmann-Schwinger equation. In other words, we must require that, for coordinates comparable to the extension of the scattering potential, the scattering state be roughly equal to the incident state. Using Eq. (1.18) and the formula for the Green function (1.24), we find for the first order approximation to the scattering state right in the center of the scattering potential, $\mathbf{x} = 0$,

$$\begin{aligned}\psi^{(1)}(0) &= \frac{1}{(2\pi)^{3/2}}\left(1 - \frac{mV_0}{2\pi\mu}\int d^3x' \frac{e^{ikx'}}{x'}\frac{e^{-\mu x'}}{x'}\right) = \\ &= \frac{1}{(2\pi)^{3/2}}\left(1 - \frac{2mV_0}{\mu}\int_0^\infty dr e^{(ik-\mu)r}\right) = \\ &= \frac{1}{(2\pi)^{3/2}}\left(1 - \frac{2mV_0}{\mu(\mu - ik)}\right).\end{aligned}$$

At $k \sim \mu$, we are thus led to the criterion

$$\frac{mV_0}{\mu^2} \ll 1$$

for the applicability of the first order Born approximation. To make some sense out of this expression, let us estimate the value of V_0 at which the Yukawa potential becomes strong enough to support a quantum **bound state**. (For the Schrödinger equation of a potential strong enough to create a bound state, can surely not be processed in low order perturbation theory.) The energy balance of a fictitious bound state — i.e. an eigenstate of the Schrödinger equation confined to the support of the scattering potential — will contain the competition of a kinetic energy term (positive) and the Yukawa potential energy (negative for $V_0 < 0$). By Heisenberg uncertainty, the minimal kinetic energy of a particle confined to the extension of the potential well is given by $E_{\rm kin} \simeq \frac{1}{2m\mu^2}$. The potential energy gain in the center of the well can be estimated as $E_{\rm pot} \simeq -V_0$. The two terms balance at $|E_{\rm pot}|/E_{\rm kin} \simeq 1$, or $\frac{2mV_0}{\mu^2} \simeq 1$. We thus conclude that, as expected, the criterion for the applicability of the Born approximation coalesces with the bound state criterion.

Finally, notice how the Yukawa scattering amplitude illustrates the general principles on scattering wavelength $2\pi/k$ vs. potential extension $R_V \sim \mu^{-1}$ discussed above. The first panel in the figure above shows a plot of the Yukawa potential in the xz plane, in arbitrary units, where z is the direction of the incident beam. The second, third, and fourth panel, resp., then show the intensity of the scattered wave amplitude, $\frac{1}{(2\pi)^3 r^2}|f(\mathbf{k},\mathbf{k}')|^2$ (again in arbitrary units.) Values $r < k^{-1}$ are blocked out, as the asymptotic spherical wave form of the

scattering amplitude, (1.26) is limited to $kr > 1$. Notice how for $k \ll \mu^{-1}$ (second panel) the scattering is diffuse and isotropic, for resonant values $k \sim \mu^{-1}$ (third panel) resolves angular structures at a strong overall scattering signal while for $k \gg \mu^{-1}$ (fourth panel) the scattered signal becomes weak.

1.3 Centro symmetric scattering

EXERCISE Recall the solution of the Poisson equation in spherical coordinates from electrodynamics. In particular, recall the definition of spherical harmonics and Legendre polynomials. In what sense are these functions 'complete' and 'orthonormal'?

Above, we have seen how the Born scattering amplitude acquired the simplified form of a one-dimensional integral if the scattering potential is symmetric. This, however, has been just the tip of an iceberg; at this point, we have nowhere nearly exploited the full consequences of angular momentum conservation in spherically symmetric scattering problems. In this section we will adjust the general theory of scattering to radially symmetric problems. This will bring us into the position to solve even complex problems at relative ease.

1.3.1 Partial wave decomposition of plane waves

The first idea that comes to mind when thinking about a spherically symmetric problem is a representation in terms of angular momentum eigenstates. However, the actual implementation of that representation is not so obvious, the point being that the starting point of the theory — the incoming plane wave states — are in no manifest way related to angular momentum eigenstates.[8]

Choosing $\mathbf{k} = k\mathbf{e}_z$, the plane wave $\exp(ikz)/(2\pi)^{3/2}$ is an eigenfunction of the Laplace operator, expressed in cartesian coordinates. As you may recall from electrodynamics (or the theory of the hydrogen atom), the Laplacian affords an explicit diagonalization in terms of eigenfunctions expressed in spherical coordinates, cf. the info block below. Further, the set of eigenfunctions of a symmetric linear operator is *complete*, i.e. it must be possible to represent the plane wave in terms of an expansion in spherical eigenfunctions. We are after this expansion.

INFO Let us recall the representation of the **Schrödinger equation**

$$\left(-\frac{1}{2m}\Delta + V - E\right)\psi = 0,$$

in spherical coordinates. Here, $V = V(r)$ is assumed to be a spherically symmetric potential (for otherwise the usage of spherical coordinates would not make much sense.) Note that for $V = 0$,

[8] Of course one might decide to build the theory on more 'angular momentum like' incoming states. However, such states can hardly be realized in an accelerator. By contrast, the plane wave *is* closely tailored to applications and should remain the base point of the theory.

the solution of the equation amounts to a diagonalization of the Laplace operator. We now use the infamous representation of the Laplacian in spherical coordinates,

$$\begin{aligned}\Delta &= \Delta_r + \frac{1}{r^2}\Delta_\Omega,\\ \Delta_r &= \frac{1}{r^2}\partial_r r^2 \partial_r,\\ \Delta_\Omega &= \frac{1}{\sin\theta}\partial_\theta \sin\theta \partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2.\end{aligned} \tag{1.36}$$

Recall (from electrodynamics) that the angular part of the Laplacian, Δ_Ω is diagonalized as,

$$\Delta_\Omega Y_{l,m} = -l(l+1)Y_{l,m}, \tag{1.37}$$

where $Y_{l,m} = Y_{l,m}(\theta,\phi)$ are the **spherical harmonics**. For the sake of completeness, we mention that the spherical harmonics are given by

$$Y_{l,m}(\theta,\phi) = \mathcal{N}_{l,m} e^{im\phi} P_l^m(\cos\theta), \tag{1.38}$$

where P_l^m are the so-called associated Legendre functions and the normalization factor $\mathcal{N}_{l,m}$ is chosen so as to produce the normalization $\int_{S^2} d\Omega\, Y^*_{l,m} Y_{l',m'} = \delta_{ll'}\delta_{mm'}$, where the integral is over the unit sphere with its canonical measure.
Substituting the ansatz $\psi(r,\theta,\phi) = R_l(r)Y_{lm}(\theta,\phi)$ into the Schrödinger equation above, we obtain

$$\left(\frac{1}{r^2}\partial_r r^2 \partial_r - \frac{l(l+1)}{r^2} + 2m(E - V(r))\right) R_l(r) = 0$$

as an effective equation for the radial component of the wave functions. Notice the $(2l+1)$-fold m-degeneracy inherent to the problem. The eigenvalues E may depend on the magnitude, l, of the angular momentum but not on its z-component. We may substitute the ansatz $R = \chi/r$ to transform the radial equation into the more convenient representation[9]

$$\left(\partial_r^2 + 2m\left(E - V(r) - \frac{l(l+1)}{2mr^2}\right)\right)\chi_l(r) = 0. \tag{1.39}$$

Notice that the normalization condition on the three dimensional wave functions, $\int d^3x\, |\psi(x)|^2 = 1$ requires $\int_0^\infty dr\, |\chi_l(r)|^2 = 1$. Also notice that the combination $V + l(l+1)/(2mr^2)$ is the 'effective' radial potential, i.e. the native potential corrected for by an angular momentum contribution, familiar from classical mechanics.

[9] How does on 'guess' such an ansatz? Consider the differential equation $(\mathcal{O} - V)\psi = 0$, where $\mathcal{O}$ is some second order differential operator, and $V = V(r)$ a 'function', i.e. an operator that assumes a diagonal form in the representation where $\mathcal{O}$ is a differential operator. In a more abstract Hilbert space notation, this assumes the form $(\hat{\mathcal{O}} + \hat{V})|\psi\rangle = 0$. We may now subject this equation to a **similarity transformation**, i.e. $\hat{S}(\hat{\mathcal{O}} + \hat{V})\hat{S}^{-1}(\hat{S}|\psi\rangle) = 0$. Choosing $\hat{S}$ to be a linear operator diagonal in the representation where $\hat{V}$ is diagonal (i.e. an operator that is r-diagonal in our problem above, such that the diagonal elements $S(r)$ define a function), and defining $|\chi\rangle \equiv \hat{S}|\psi\rangle$, we obtain the equivalent problem $(\hat{S}\hat{\mathcal{O}}\hat{S}^{-1} + \hat{V})|\chi\rangle = 0$. We may now work out the explicit form of the transformed operator $\hat{S}\hat{\mathcal{O}}\hat{S}^{-1}$, which in practice amounts to letting the derivatives contained in $\mathcal{O}$ act on the function $S(r)$. It can be shown that it is always possible to find a representation $S\mathcal{O}S^{-1} = f(r)\partial_r^2 + g(r)$, i.e. a representation void of first derivatives. In our problem above, it is easy to verify that with the choice $S = r$ we get $r\left(\frac{1}{r^2}\partial_r r^2 \partial_r\right) r^{-1} = \partial_r^2$.

Diagonalization of the Laplacian in spherical coordinates

We now aim to get the complete set of eigenfunctions of the Laplacian in spherical coordinates under control. Once we have these functions, we can use them to represent the plane wave by linear superposition. Mathematically, the problem of diagonalizing the Laplace operator is tantamount to the solution of the free Schrödinger equation $(-\frac{1}{2m}\Delta - E)\psi = 0$.[10] We start from an ansatz

$$\psi_{k,l,m}(r,\theta,\phi) = Y_{l,m}(\theta,\phi)R_{l,k}(r), \tag{1.40}$$

where $k = (2mE)^{1/2}$, and a normalization

$$\int_0^\infty r^2 dr \int d\Omega\, \bar{\psi}_{k',l',m'}\psi_{k,l,m} = \delta_{ll'}\delta_{mm'}2\pi\delta(k-k') \tag{1.41}$$

is implied. Eq. (1.39) for the $l = 0$ radial wave function $\chi_0 = rR_0$ is solved by $\chi_{0,k} = 2\sin(kr)$, which gets us to

$$R_{0,k} = \frac{2\sin(kr)}{r}.$$

For general l, the radial equation may be solved by recursive methods[11] or by noting that the radial equation assumes the form of a **Bessel differential equation** [12]. The solution (compatible with the normalization (1.41) and non-singular at the origin) is given by

$$\begin{aligned} R_{k,l}(r) &= 2kj_l(kr), \\ j_l(x) &= \sqrt{\frac{\pi}{2x}}J_{l+\frac{1}{2}}(x), \end{aligned} \tag{1.42}$$

where j_l are the so-called **spherical Bessel functions**, and $J_{l+\frac{1}{2}}$ Bessel functions of half integer order. These functions are tabulated, however, for practical purposes it is more useful to consider the recursion relation

$$R_{k,l}(r) = 2\left(-\frac{r}{k}\right)^l\left(\frac{1}{r}\partial_r\right)^l\frac{\sin(kr)}{r}. \tag{1.43}$$

In the above, we have normalized the radial functions as

$$\int_0^\infty r^2 dr R_{k',l'}R_{k,l} = 2\pi\delta_{ll'}\delta(k-k'),$$

Which conforms with the normalization (1.41) above.

[10] Upon multiplication by $-2m$, the Schrödinger equation assumes the form of an eigenvalue equation $\Delta\psi = \lambda\psi$, where $\lambda = -2mE$, i.e. the solution of the eigenvalue equation is equivalent to the solution of the free Schrödinger equation.

[11] Cf. LL, §33.

[12] Cf., e.g., `http://mathworld.wolfram.com/BesselDifferentialEquation.html`

Finally, let us discuss two important forms of **asymptotic behavior of the radial functions**: For **large** r, the derivatives in (1.43) dominantly act on the kr combination (why?). We then obtain

$$R_{k,l}(r) \overset{r\gg k^{-1}}{\simeq} 2\frac{\sin\left(kr-\frac{l\pi}{2}\right)}{r}. \tag{1.44}$$

For **small** r, the smallest r-power in the Taylor expansion of the $\sin$-function surviving the action of the l derivatives gives the dominant contribution. Counting powers, we conclude that this is the term of $(2l+1)$th order, i.e. $(-)^l(kr)^{2l+1}/(2l+1)!$. Working out the effect of the subsequent derivatives, we readily find

$$R_{k,l}(r) \overset{r\ll k^{-1}}{\simeq} 2\frac{k^{l+1}}{(2l+1)!!}r^l, \tag{1.45}$$

i.e. the expansion of the lth radial function starts at lth order in l.

Plane wave expansion

We chose a reference frame where the z-axis of the spherical coordinate system is aligned with the incoming momentum. The plane wave then assumes the form

$$\psi = \frac{1}{(2\pi)^{3/2}}e^{ikz} = \frac{1}{(2\pi)^{3/2}}e^{ikr\cos\theta},$$

independent of the azimuthal angle ϕ. This means that only spherical harmonics $Y_{l,m=0}$ of zero azimuthal variation contribute to the expansion. Noting that $Y_{l,0} = i^l\sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta)$, where P_l are the Legendre polynomials, we thus start out from the ansatz,

$$\frac{e^{ikr\cos\theta}}{(2\pi)^{3/2}} = \sum_l C_l P_l(\cos\theta)R_{k,l}(r),$$

where constant numerical factors multiplying the spherical harmonics and the radial wave functions have been absorbed in the expansion coefficients C_l.

The values of these coefficients can be fixed by Taylor expansion in the combination $r\cos\theta$. On the l.h.s. the term of mth order appears as $\frac{1}{(2\pi)^{3/2}}\frac{(ik)^m}{m!}(r\cos\theta)^m$. Turning to the r.h.s., crucially, the combination $(r\cos\theta)^m$ can only be produced by the contribution $l=m$. For $l<m$, the Legendre polynomials contain only powers of $\cos\theta$ of $\mathcal{O}(\leq l<m)$ and for $l>m$, the radial functions start out at order $r^{l>m}$ (cf. Eq. (1.45).) Turning to the contribution $l=m$, we need to pick the term of maximal order from the Legendre polynomial $P_m(x) = \frac{(2l)!}{2^m m!^2}x^m + \mathcal{O}(x^{l<m})$. The term of mth order contributed by the radial function $R_{k,m}$ is given in (1.45). Combining terms, we are led to the condition

$$\begin{aligned}\frac{1}{(2\pi)^{3/2}}\frac{(ik)^m}{m!}(r\cos\theta)^m &\overset{!}{=} C_m\frac{(2m)!}{2^m m!^2}(\cos\theta)^m 2\frac{k^{m+1}}{(2m+1)!!}r^m \Leftrightarrow \\ &\Leftrightarrow C_m = \frac{1}{(2\pi)^{3/2}}\frac{1}{2k}(2i)^m\frac{m!(2m+1)!!}{(2m)!} = (2m+1)\frac{i^m}{2k}.\end{aligned}$$

We thus arrive at the final result for the **plane wave expansion**

$$\boxed{\frac{e^{ikr\cos\theta}}{(2\pi)^{3/2}} = \frac{1}{(2\pi)^{3/2}}\frac{1}{2k}\sum_l (2l+1)i^l P_l(\cos\theta)R_{k,l}(r),} \tag{1.46}$$

where the radial functions$R_{k,l}$ are given by (1.43). For later reference, we spell out the form of the expansion far from the coordinate center,

$$\frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}} \overset{kr\gg 1}{\simeq} \frac{1}{(2\pi)^{3/2}}\frac{1}{2k}\sum_l (2l+1)i^{l+1}P_l(\cos\theta)\left(\frac{e^{-(ikr-\frac{l\pi}{2})}}{r} - \frac{e^{i(kr-\frac{l\pi}{2})}}{r}\right), \tag{1.47}$$

where we used (1.44). This result expresses the plane wave as a superposition of incoming (first term) and outgoing (second term) spherical waves.

1.3.2 Scattering phase shift

Turning back to the full scattering state, $|\psi^+\rangle$, we seek for a representation analogous to (1.47), i.e. an expansion in contributions of definite angular momenta, valid at large separation from the scattering center. As with the incoming state, we start out from an expansion

$$\psi^+(\mathbf{x}) = \sum_l C_l' P_l(\cos\theta)R_{k,l}(r),$$

independent of the azimuthal angle ϕ. Now, far away from the center, the radial Schrödinger equation (1.39) for $\chi = Rr$ collapses to (what is the parameter controlling the approximation below?)

$$\left(\partial_r^2 + 2m\left(E - V(r) - \frac{l(l+1)}{2mr^2}\right)\right)\chi_l \simeq \left(\partial_r^2 + 2mE\right)\chi_l = 0, \tag{1.48}$$

and the general solution of this equation is given by linear superposition of radial plane waves $\exp(\pm ikr)$. We thus conclude that the general form of the expansion reads

$$\psi^+(\mathbf{x}) = \sum_l P_l(\cos\theta)R_{k,l}(r)\left(C_{l,\mathrm{in}}\frac{e^{-ikr}}{r} + C_{l,\mathrm{out}}\frac{e^{ikr}}{r}\right),$$

with expansion coefficients $C_{l,\mathrm{in/out}}$ yet to be determined. Now, by causality the presence of the scattering potential must not affect the *incoming* plane waves contributing to this expansion (think about this point!). Comparison with the expansion of the free state, (1.47) thus leads to the identification $C_{l,\mathrm{in}} = \frac{1}{(2\pi)^{3/2}}\frac{2k+1}{2k}i^{l+1}e^{i\frac{l\pi}{2}}$ and to

$$\psi^+(r,\theta) \simeq \frac{1}{(2\pi)^{3/2}}\frac{1}{2k}\sum_l (2l+1)i^{l+1}P_l(\cos\theta)\left(\frac{e^{-i(kr-\frac{l\pi}{2})}}{r} - \eta_l\frac{e^{i(kr-\frac{l\pi}{2})}}{r}\right), \tag{1.49}$$

with as yet undetermined complex coefficients η_l. Notice that the full complexity of the scattering process is now encoded in the set of numbers $\{\eta_l\}$. An important constraint on the values of these coefficients follows from current conservation, i.e. the fact that all particles incident upon the scattering center must eventually leave it. This means that the total current density carried by the state ψ, integrated over a large surface enclosing the target must be zero.

To be specific, consider a sphere of radius $S \gg R_V$ centered around the target. The current flow through that sphere is calculated as

$$I = \int_S d\mathbf{S} \cdot \mathbf{j} = R^2 \int_{S^2} d\Omega \, j_r(R, \theta),$$

where

$$j_r = \frac{1}{2mi} \left(\bar{\psi}^+ \partial_r \psi^+ - (\partial_r \bar{\psi}^+) \psi^+ \right)$$

Using the **orthonormality relation of Legendre polynomials**,

$$\int_{S^2} d\Omega \, P_l(\cos\theta) P_{l'}(\cos\theta) = \frac{4\pi}{2l+1} \delta_{l,l'}, \tag{1.50}$$

we then readily obtain [13]

$$\boxed{I = \frac{1}{(2\pi)^3} \frac{\pi}{mk} \sum_l (2l+1)(1 - |\eta_l|^2),} \tag{1.51}$$

where we may think of the '1' as the contribution of the incoming current, and the '$|\eta_l|^2$' as the outgoing contribution. This shows that the vanishing of the current (separately, for each angular momentum channel) requires $|\eta_l| = 1$. In recognition of this fact, it is customary to write

$$\boxed{\eta_l = \exp(2i\delta_l),} \tag{1.52}$$

where the real parameters δ_l are known as **scattering phase shifts**.

We aim to give the scattering phase shifts a physical interpretation. To this end, let us add and subtract to $|\psi^+\rangle$ the incoming state, $|\psi^+\rangle = |\phi\rangle + (|\psi^+\rangle - |\phi\rangle)$ and use the spherical wave decomposition (1.47). As a result we obtain

$$\psi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \left(e^{i\mathbf{k}\cdot\mathbf{x}} + \left[\frac{1}{2k} \sum_l (2l+1) i^{l+1} P_l(\cos\theta)(1 - \eta_l) \right] e^{-i\frac{l\pi}{2}} \frac{e^{ikr}}{r} \right). \tag{1.53}$$

[13] The unconventional unit $[I] = \text{Volume}/\text{time}$ has to do with the fact that our wave functions carry unit dimension $[\psi] = 1$, i.e. $|\psi|^2$ is *not* a density, rather it has dimension volume $\times$ density.

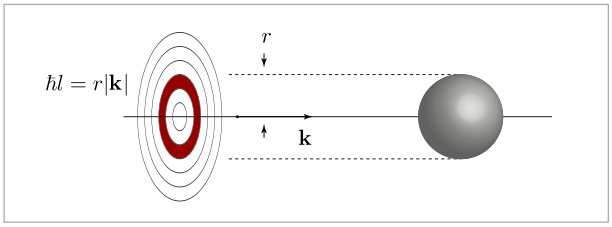


Figure 1.3: On the heuristic interpretation of the partial wave decomposition of the total cross section.

Comparison with the definition of the scattering amplitude, Eq. (1.26) then leads to the identification

$$\boxed{f(\mathbf{k},\mathbf{k}') = \frac{1}{2k}\sum_l (2l+1) i^{l+1} P_l(\cos\theta)(1-\eta_l) e^{-i\frac{l\pi}{2}}.} \tag{1.54}$$

We have managed to encode the information about the scattering amplitude — a function on the two sphere — in a set of *numbers*, $\{\eta_l\}$. The practical importance of this formula lies in the fact that, usually, the first few values $\eta_0, \eta_1, \eta_2, \ldots$ suffice to describe the properties of a scattering target with sufficient precision. Indeed, we have argued in section 1.2.5 that for incident wavelengths $\lambda > R_V$, the scattering amplitudes are nearly isotropic. In this case, only the $l = 0$ channel — the so called s-**wave scattering** channel — will have a scattering parameter η_0 significantly different from unity. Model calculations on exactly solvable problems, e.g. the scattering off an infinitely strong potential of radius R_V (a 'hard ball'), indeed show that scattering phase shifts tend to scale as $\delta_l \sim (R_V/\lambda)^{\alpha_l}$, where α_l is a power increasing in l. In the next section we will discuss an example of s-wave scattering at a level beyond the Born approximation.

INFO To better understand the **physical meaning of the scattering phase shifts**, it is instructive to consider the total cross section, $\sigma = \int d\Omega\, |f(\mathbf{k},\mathbf{k}'|^2$. Using the orthonormality relation (1.50), we readily obtain

$$\sigma = \sum_l \sigma_l, \qquad \sigma_l \equiv \frac{\pi}{k^2}(2l+1)|1-\eta_l|^2. \tag{1.55}$$

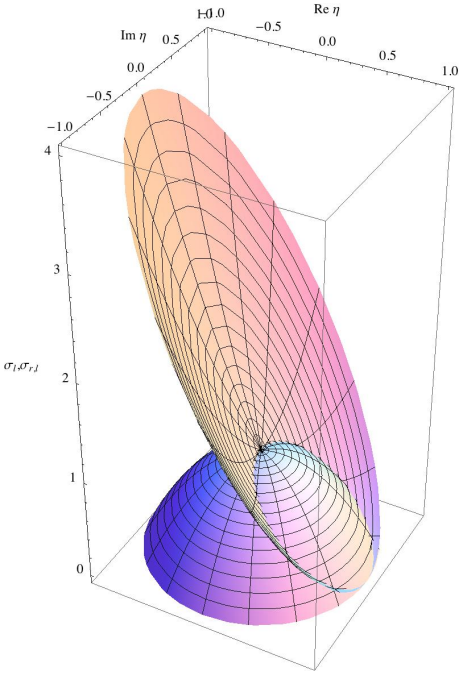
This formula affords a simple semiclassical interpretation (cf. Fig. 1.3): consider a state of angular momentum component $\hbar l$ in a plane perpendicular to the incident beam axis.[14] Using the classical relation $\hbar l = |\mathbf{k}|r$, where r is the distance to the beam axis, we conclude that this state corresponds to particles incident in a cylindrical shell of radius r around the axis. Now, angular momentum is quantized, meaning that to particles corresponding to $\hbar l$ we should attribute a 'ring' of radius $\hbar(l+1)/k$ and thickness $\hbar/p$. The geometric area corresponding to the particles incident in that angular momentum sector is given by $\frac{\pi}{k^2}((l+1)^2 - l^2) = \frac{\pi(2l+1)}{k^2} \equiv \sigma_{l,\mathrm{cl}}$, where we dropped the constant $\hbar$ again. Here, we introduced $\sigma_{l,\mathrm{cl}}$ as an abbreviation for a 'maximally efficient' classical cross section, which scatters all the particles incoming in the angular momentum sector l.

[14] Accounting for the semiclassical nature of the argument, we reinstall $\hbar$ here.

On this basis, let's take a look at the cross section 1.55. Apparently, the scattering is the most effective for $\eta_l = -1$, in which case $\sigma_l = 4\sigma_{l,\mathrm{cl}}$ is four times more effective than the perfect classical scattering center. The origin of the multiplicative enhancement $4 = 2^2$ is quantum interference: the incoming and the scattered amplitude may constructively interfere, $2 = 1 + 1$, to an amplitude squaring to a factor 4 enhancement in the intensity. By contrast, for $\eta_l = 1$, we have perfect destructive interference and a vanishing cross section $\sigma_l = 0$.

Let us briefly explore what happens if *inelastic scattering* is an option. Here, we *define* the term inelastic scattering as all types of scattering where the condition $|\eta_l| = 1$ need no longer hold. In this case, if $|\eta_l| < 1$, there will be net current flow into the target region (cf. Eq. (1.51).) We write this **reaction current** as

$$I = j_{\mathrm{in}} \sum_l \sigma_{r,l}, \qquad \sigma_{r,l} = \frac{\pi}{k^2}(2l+1)(1-|\eta_l|^2),$$



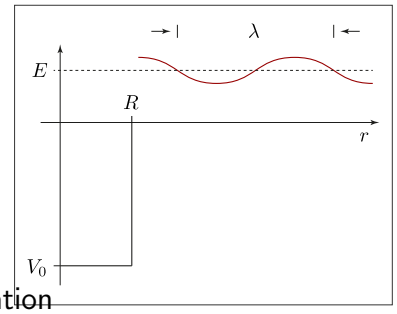
where $j_{\mathrm{in}} = \frac{1}{(2\pi)^3}\frac{k}{m}$ is the incoming current density, and $\sigma_{r,l}$ the reaction cross section in channel l. The reaction current reaches its maximum at $\eta_l = 0$, when $\sigma_{r,l}$ coincides with the full geometric extension of the angular momentum sector. At this value, all incoming current is swallowed by the target, according to its geometric extension. The process is 'inelastic', i.e. no interference corrections here.

The cross section, σ_l and the reaction cross section $\sigma_{r,l}$ are shown in the figure as a function of the complex parameter η_l.

1.3.3 Example: scattering off a potential well

In this section, we will consider the scattering of low energy particles, $\lambda \ll R_V$, off a spherically symmetric potential well of depth V_0. Thanks to the condition $\lambda/R_V \ll 1$, scattering in the s-wave channel will be dominant, i.e. we wish to compute the scattering phase shift δ_0. In this section, no weakness assumptions on the scattering potential will be made, and we will go beyond the level of the first order Born approximation.

INFO The situation outlined above is relevant to, e.g., the **scattering of low energy neutrons** in nuclear physics. Due to the short-rangedness of nuclear forces, the potential describing a large nucleus can to a reasonable approximation be modeled by a rectangular potential well. As we will see, the scattering of low energetic particles subject to short range interaction with the nucleus — slow neutrons — reveals telling information on effective extension and depths of the nuclear potential.



The energy and length scales relevant to the problem are shown in the well. Denoting the extension of the well by $R_V = R$ for simplicity, our task is to solve the radial Schrödinger equation (1.39) for a potential $V(r) = -V_0\Theta(R-r)$, where Θ is the Heaviside step function. For $l = 0$, the equation assumes the form

$$\left(\partial_r^2 + 2m(E + V_0\Theta(R-r))\right)\chi = 0,$$

where we wrote $\chi_0 \equiv \chi$ for simplicity. Introducing the abbreviation

$$k_i \equiv 2m\sqrt{E+V_0},$$

this equation has the general solution

$$\chi(r) = c_{+,i}e^{ik_i r} + c_{-,i}e^{-ik_i r}, \qquad r < R,$$
$$\chi(r) = c_{+,o}e^{ikr} + c_{-,o}e^{-ikr}, \qquad r > R.$$

Without loss of generality, we may choose coefficients $c_{\pm,o}$ such that the wave function in outer space, $r > R$, equals the radial component of the $l = 0$ contribution to the scattering state (1.49). A straightforward comparison then leads to the alternative representation

$$\chi(r) = \frac{i}{(2\pi)^{3/2}2k}\left(e^{-ikr} - \eta_0 e^{ikr}\right) =$$
$$= \frac{e^{i\delta_0}}{(2\pi)^{3/2}k}\sin(kr+\delta_0), \qquad r > R. \tag{1.56}$$

Turning to the 'inner contribution', the ratio between the coefficients $c_{\pm,i}$ is determined by the condition that the radial wave function $R(r) = \chi(r)/r$ be finite at the origin. This requires $c_{+,i} = -c_{-,i} \equiv a/2i$, or

$$\chi(r) = a\sin(k_i r), \qquad r < R.$$

The remaining two coefficients a and η_0 are now fixed by the requirement of continuity of the wave function and of its derivative at $r = R$. A short calculation shows that this is equivalent to the conditions:

$$\frac{e^{i\delta_0}}{(2\pi)^{3/2}k}\sin(kR+\delta_0) = a\sin(k_i R),$$
$$\frac{e^{i\delta_0}}{(2\pi)^{3/2}k_i}\cos(kR+\delta_0) = a\cos(k_i R).$$

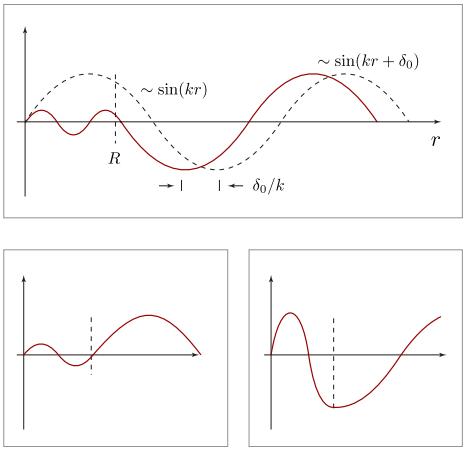
Division of these equations leads to

$$\boxed{\delta_0 = \arctan\left(\frac{k}{k_i}\tan(k_i R)\right) - kR.} \tag{1.57}$$

The differential cross section corresponding to this phase shift is given by (1.55), or

$$\sigma \simeq \sigma_0 = \frac{4\pi}{k^2}\sin^2(\delta_0).$$

EXERCISE Check that this result conforms with the **optical theorem** Eq. (1.32).



The conceptual meaning of the phase scattering phase shift is shown in the figure. Within the potential well, the wave function oscillates at a faster rate than in the outside region. As a result, the wave function gets 'dragged' towards the inner region, and this manifests itself in a phase shift relative to the wave function of the potential-less problem. An optimal contact between the inner and the outer wave function is established if the inner wave function approaches the boundary at a vanishing derivative (bottom right). In this case, the inner amplitude assumes its maximum possible value. The phase shift corresponding to this **resonance scattering** condition is given by $\delta_0 = kR + \frac{\pi}{2} \overset{kR \ll 1}{\simeq} \frac{\pi}{2}$. At resonance, the cross section assumes its maximal value

$$\sigma_{\max} = \frac{4\pi}{k^2}.$$

Conversely, if the wave functions match close to a zero crossing (bottom left), one speaks of **potential scattering**. The difference in phase shift to the resonant value is $\pi/2$, i.e. $\delta_0 \simeq \pi \mathrm{mod}(2\pi)$, and the scattering cross section approaches 0. In this limit, the target is next to 'invisible' to the incoming beam. Starting from a resonant configuration, somewhat paradoxically, a potential scattering configuration may be reached by *increasing* the scattering potential, i.e. an

> An increase in the scattering potential may lead to a reduced, or even vanishing scattering cross section.

A manifestation of this phenomenon is known as the **Ramsauer-Townsend effect**. It occurs, e.g. in the scattering of electrons off rare gas targets such as argon, krypton or xenon. The Ramsauer-Townsend effect was first observed in 1923, before the advent of wave mechanics, and represented a mystery at the time.

1.3.4 Scattering at ultra-low energies: scattering length and bound states

In the previous section, we have explored what happens if the energy of the incoming particles is low, and s-wave scattering dominates. We next consider the regime of the lowest conceivable energies, where the wavelength $2\pi/k$ represents the largest length scale in the problem (barring the detector distance, of course.) Perhaps surprisingly, we will find that this limit contains a wealth of information on the microscopic structure of the target.

In the limit of vanishingly small energies, the Schrödinger equation (1.48) is solved by

$$\chi_0(r) = \text{const.} \times (r - a),$$

where a is a constant. We can think of this as a sin-profile (1.56) in the limit of infinitely large wave length. Indeed, for small k the sin function in (1.56) can be expanded as

$$\chi_0 = \text{const.} \times \left(r + \frac{\delta_0}{k} \right) + \mathcal{O}(k),$$

where we *anticipated* that the scattering phase shift will be proportional to k, i.e. $\delta_0/k = \mathcal{O}(k^0)$. For the rectangular potential well, our result (1.57) makes this proportionality manifest — just expand the formula in small k to reach this conclusion — however, for the time being we do not wish to make assumptions on the details of the scattering potential.

Comparison of these expressions leads to the identification

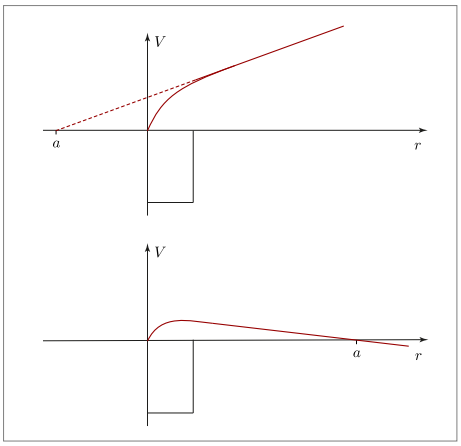
$$\lim_{k\to 0} \frac{\delta_0}{k} = -a. \tag{1.58}$$

Eq. (1.54) for the scattering amplitude further implies

$$f_0 = \frac{1}{k} e^{i\delta_0} \sin(\delta_0) \overset{k\to 0}{\longrightarrow} \frac{\delta_0}{k} = -a,$$

i.e. the s-wave cross section becomes

$$\boxed{\sigma_0 \overset{k\to 0}{=} 4\pi a^2.} \tag{1.59}$$



The length scale a is known as the **scattering length**. Eq. (1.59) tells us that in the limit of ultra-low energies, the target scatters as if it were a scattering center of geometric area (four times) a disk of radius a. Indeed, a sets the length scale at which the scattering wave function vanishes. In very loose terms, we may compare this to the vanishing of a wave function at a hard potential barrier. I.e. in a sense, a plays the role of the effective geometric extension of the scattering target, as seen from the perspective of very low energetic scattering particles.

However, as we will see momentarily, this interpretation must be taken with a grain of salt: the scattering length, can become infinitely large, or even assume negative values! Let us explore this phenomenon on the paradigmatic example of the rectangular well. (The essentials of the discussion generalize to other types of potentials.) Substitution of (1.57) into (1.58) gives

$$a = -\lim_{k\to 0} \frac{1}{k} \left(\arctan\left(\frac{k}{k_i} \tan(k_i R) \right) - kR \right) = -\frac{\tan(k_i R)}{k_i} + R.$$

Apparently, the scattering length shows strong dependence on $k_i = \sqrt{2m(-V_0)}$, whenever $k_i R = \frac{\pi}{2} \bmod \pi$. To understand the origin of this observation, consider the schematic form of the wave function shown in the upper pannel of the figure above. Here, $k_i R \lesssim \frac{\pi}{2}$ is just a little smaller than $\pi/2$, which means that the wave function $\sim \sin(k_i r)$ describing the inner of the potential well connects to the outside wave function at slightly positive derivative $\sim \cos(k_i R) > 0$. The linear extrapolation of the outside wave function has its zero at a negative value a, and this explains the appearance of a negative scattering length. Increasing $k_i R \nearrow \frac{\pi}{2}$ leads to a divergence of a towards large negative values. As we cross the critical value $\pi/2$, the derivative turns to small negative values (the lower panel), meaning that the crossing point jumps to large positive values. This is the mathematical origin of the singular behavior of the formula above. It obviously repeats itself whenever $k_i R = \frac{\pi}{2} \bmod \pi$. However, from a physical point of view, the divergence of the scattering length signals an important qualitative phenomenon. In general, an attractive target potential will support **bound state** solutions, i.e. wave functions of finite spatial support. The bound state spectrum is discrete, i.e. bound state solutions exist only for discrete energy values ϵ_i. Now, imagine the situation where one of the values $\epsilon_i \nearrow 0$ approaches zero from below. In the outside region, the bound state must connect to an exponentially decaying profile, $\sim \exp(-\kappa r)$, where $\kappa = \sqrt{2m|\epsilon_i|}$. For vanishingly small $|\epsilon_i|$, the decaying wave function $\sim \exp(-\kappa r) \simeq 1 - \kappa r$ is locally indistinguishable from a linear downturn of very gentle slope. At $\epsilon_i = 0$, the bound state of negative infinitesimal energy must smoothly connect to a scattering state of positive infinitesimal energy, and the spatial profile of that state is just the one shown in the lower panel of the figure above. In other words,

> A singularity $a \to \infty$ of the scattering length signals the appearance of a zero energy bound state.

This interpretation nicely connects with the divergence $\sigma \sim a^2$ of the cross section. In a way, the target supporting a zero energy bound state has the capacity to 'capture' zero energy wave functions. For the particular example of the rectangular well, new bound states always appear when the condition $k_i R = \frac{\pi}{2} \bmod \pi$ is met. However, the qualitative phenomenon discussion above is of general nature and will show for other types of attractive scattering potentials as well.

EXERCISE Generalize the discussion above to **repulsive** scattering potentials. Consider a rectangular potential $V(r) = V_0 \Theta(R - r)$ with $V_0 > 0$ as an example.

1.4 Dynamics of scattering

We conclude this chapter with an introduction of a number of general concepts relevant to quantum scattering. Specifically, we will introduce the concept of the so-called S-matrix, and discuss the role of time reversal in scattering. Both topics relate to the dynamics of scattering processes, hence the title of this section.

1.4.1 Scattering matrix

In scattering theory we aim to connect between 'incoming' states evolving according to some 'free' kinematics and equally free 'outgoing' states via non-trivial scattering processes. In quantum mechanics 'connecting states' means acting with a suitably defined operator. It is then a natural question to ask whether there exists an operator that incorporates all the relevant information on quantum scattering in any conceivable setup. This operator, let's call it $\hat{S}$, should meet a number of criteria, summarized in the following wishlist:

- ▷ The operator $\hat{S}$ should contain the full information on all aspects of the scattering dynamics.
- ▷ In the absence of a scattering target, the operator $\hat{S} = \mathbb{I}$ should reduce to the unit operator. I.e. we aim to define $\hat{S}$ in such a way that the information on the free propagation of states is split off.
- ▷ The operator $\hat{S}$ should be definable for any scattering scenario (including inelastic scattering.) However, in the cases discussed before, it should relate in simple ways to the carriers of scattering information introduced above, e.g. to scattering phase shifts.
- ▷ The operator $\hat{S}$ ought to be useful even in situations where one won't restrict oneself to a single incoming state (as we did so far.) For example, in the angle resolved scattering off non-spherically symmetric targets, one would like to probe the scattering properties of particles coming in from arbitrary directions. In such cases, $\hat{S}$ should carry the information on the scattering amplitude from arbitrary incoming states to arbitrary outgoing states.

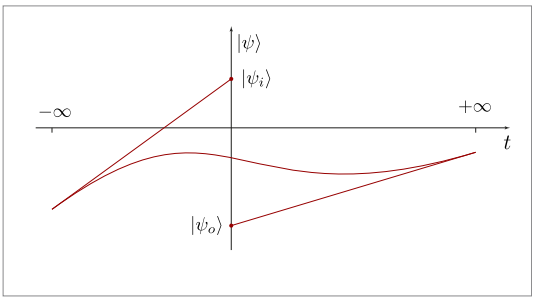
Let us, then, consider some 'incoming' state $|\phi_i\rangle$. In practice, $|\phi_i\rangle$ might be a momentum eigenstate, at a characteristic momentum directed towards the target. We may also think of a superposition of momentum states to wave packets approaching the target, or something altogether different. We aim to connect this state to $|\phi_o\rangle$, where $|\phi_o\rangle$ is taken from a set of states spanning the 'outgoing' sector of Hilbert space.[15]

By way of a first guess, might define the scattering operator $\hat{S}$ through

$$\hat{S} \stackrel{?}{\equiv} e^{-i\hat{H}\Delta t},$$

where Δt is chosen to be a 'very large time'. Here, the operator $\hat{S}$ is identified with the full time evolution operator of the theory and this means that the first criterion above is met. However, the definition has its problems: even for $\hat{H} = \hat{H}_0$, it describes non-empty dynamical evolution which, for $\Delta t \to \infty$, will generally 'transport' incoming states to spatial infinity. Relatedly, the definition depends in an ad hoc way on the choice of the reference time Δt.

[15] Strictly speaking, the scattering states considered so far are not normalizable (their normalizing integrals involve δ-functions), whence they do not lie in a proper Hilbert space. However, in practice this complication does not play much of a role — just put the system into a giant box, and all states *do* become normalizable — whence we will pretend the existence of a state 'Hilbert' space possessing all properties required by quantum mechanics, scalar product, normalizability, etc. For a mathematically sound discussion of the state spaces of scattering theory, we refer to the expert literature.



A smarter choice of $\hat{S}$ is defined in the cartoon in the figure to the right. Think of the state $|\psi_i\rangle$ as a wave package approaching the scattering target. We first ask where this wave package came from, assuming that it had not yet been influenced by the scattering target. The answer is, of course, $e^{-i(-\Delta t)\hat{H}_0}|\psi_i\rangle$ (cf the upper straight line in the figure which represents free propagation to large negative times.) The initial reference state thus obtained we expose to the full dynamics, i.e. we build a state $e^{-i\hat{H}\Delta t}e^{-i\hat{H}_0(-\Delta t)}|\psi_i\rangle$, where in the figure the evolution under $\hat{H}$ is represented by a curved line. We now send Δt to infinity to build a state $\hat{\Omega}_-|\psi_i\rangle$, where

$$\hat{\Omega}_+ \equiv \lim_{\Delta t\to\infty} e^{-i\hat{H}\Delta t}e^{i\hat{H}_0\Delta t},$$

is the so-called incoming **Møller operator**. In an analogous way we construct a state $\hat{\Omega}_0|\psi_o\rangle$,

$$\hat{\Omega}_- \equiv \lim_{\Delta t\to\infty} e^{i\hat{H}\Delta t}e^{-i\hat{H}_0\Delta t}, \tag{1.60}$$

i.e. a state first evolved to the infinite future under the free dynamics, and then evolved back under the full dynamics. We now claim that the physical information on the scattering process is encoded in the matrix element

$$\langle\psi_o|\hat{\Omega}_-^\dagger\hat{\Omega}_+|\psi_i\rangle \equiv \langle\psi_o|\hat{S}|\psi_i\rangle, \tag{1.61}$$

where

$$\boxed{\hat{S} \equiv \lim_{\Delta t\to\infty} e^{i\hat{H}_0\Delta t}e^{-2i\hat{H}\Delta t}e^{i\hat{H}_0\Delta t}} \tag{1.62}$$

is the **scattering operator**. The scattering operator arguably meets the criteria formulated above: the presence of the time evolution operator, $\exp(-2i\hat{H}t)$, suggests that it contains the full information on the scattering dynamics. For $\hat{H}_0 \to \hat{H}$ it reduces to the unit operator, and it is defined in full generality.[16] Also notice that

$$\boxed{\hat{S} \text{ is a unitary operator, } \hat{S}^\dagger\hat{S} = \mathbb{I}.}$$

To connect the scattering operator to concepts introduced earlier in this text, we use the auxiliary formula

$$\lim_{t\to\infty} f(t) = \delta\int_0^\infty dt\, e^{-\delta t} f(t),$$

[16] A more thorough discussion of the scattering operator would need to address the existence of the temporal limits above, the nontrivial issue of the domain of definition of the Møller operators, and more. We do not do this here, so our treatment remains heuristic.

according to which the limit of a function can be expressed as an integral, provided the limit exists. (Convince yourself of the validity of this expression.) Let us now apply this formula to the action of a Møller operator on an eigenstate of the free Hamiltonian. With $|\psi_i\rangle \equiv |\phi_\alpha\rangle$, where $\hat{H}_0|\phi_\alpha\rangle = E_\alpha|\phi_\alpha\rangle$ and $\{\alpha\}$ is a set of quantum numbers, we have

$$\begin{aligned}\hat{\Omega}_+|\phi_\alpha\rangle &= \lim_{\Delta t\to\infty} e^{-i\hat{H}\Delta t}e^{i\hat{H}_0\Delta t}|\phi_\alpha\rangle = \\ &= \lim_{\Delta t\to\infty} e^{-i(\hat{H}-E_\alpha)\Delta t}|\phi_\alpha\rangle = \\ &= \lim_{\delta\to 0}\delta\int_0^\infty dt e^{-i(\hat{H}-E_\alpha-i\delta)\Delta t}|\phi_\alpha\rangle = \\ &= \frac{i\delta}{E_\alpha+i\delta-\hat{H}}|\phi_\alpha\rangle = i\delta\hat{G}^+(E_\alpha)|\phi_\alpha\rangle = |\psi_\alpha^+\rangle,\end{aligned}$$

where $|\psi_\alpha^+\rangle$ is the scattering state corresponding to $|\phi_\alpha\rangle$ according to Eq. (1.20). In an analogous way we obtain $\langle\phi_\alpha|\hat{\Omega}_-^\dagger = \langle\psi_\alpha^-|$, where

$$|\psi_\alpha^-\rangle = -i\delta\hat{G}^-(E_\alpha)|\phi_\alpha\rangle,$$

and $\hat{G}^-(E) = (E-i\delta-\hat{H})^{-1}$ is the advanced Green function. According to these results, the S-**matrix element**

$$\boxed{S_{\alpha\alpha'} = \langle\phi_\alpha|\hat{S}|\phi_{\alpha'}\rangle = \langle\psi_\alpha^-|\psi_{\alpha'}^+\rangle} \tag{1.63}$$

is given by the overlap of two scattering states corresponding to the reference states $|\phi_{\alpha/\alpha'}\rangle$. The overlap is built between an outgoing scattering state, and an — admittedly less easy to imagine — fictitious anti-causal scattering state $|\psi_\alpha^-\rangle$. One may think of the latter as a state whose time evolution is obtained by time reversal (like in a movie played in reverse direction) of the time evolution of the outgoing state. Eq. (1.62) underpins the fact that the scattering matrix carries the full information of the scattering process, i.e. knowing all matrix elements of $\hat{S}$ is equivalent to knowledge of the scattering states.

It is instructive to formulate this statement in an alternative way. We first manipulate Eq. (1.20) to obtain an alternative representation of the Lippmann-Schwinger equation (indices α temporarily suppressed):

$$\begin{aligned}|\psi^+\rangle &= \hat{G}^+(i\delta)|\phi\rangle = \\ &= \hat{G}^+(\underbrace{E+i\delta-\hat{H}}_{(\hat{G}^+)^{-1}} - \underbrace{(E-\hat{H}_0)}_{\to 0} + \hat{V})|\phi\rangle = \hat{G}^+\hat{V}|\phi\rangle + |\phi\rangle.\end{aligned}$$

Subtracting the plus and the minus variant of this equation,

$$|\psi_\alpha^\pm\rangle = \hat{G}^\pm(E_\alpha)\hat{V}|\phi_\alpha\rangle + |\phi_\alpha\rangle,$$

we obtain

$$\begin{aligned}|\psi_\alpha^-\rangle &= |\psi_\alpha^+\rangle + (\hat{G}^-(E_\alpha) - \hat{G}^+(E_\alpha))\hat{V}|\phi_\alpha\rangle \Leftrightarrow \\ \langle\psi_\alpha^-| &= \langle\psi_\alpha^+| + \langle\phi_\alpha|\hat{V}(\hat{G}^+(E_\alpha) - \hat{G}^-(E_\alpha)).\end{aligned}$$

Using this result, and the identity $(\hat{G}^+(E_\alpha) - \hat{G}^-(E_\alpha))|\psi^+_{\alpha'}\rangle = |\psi^+_{\alpha'}\rangle \left(\frac{1}{E^+_\alpha - E_{\alpha'}} - \frac{1}{E^-_\alpha - E_{\alpha'}}\right) = |\psi^+_{\alpha'}\rangle(-2\pi i)\delta(E_\alpha - E_{\alpha'})$, we express the scattering matrix as

$$S_{\alpha\alpha'} = \langle\psi^+_\alpha|\psi^+_{\alpha'}\rangle - 2\pi i\delta(E_\alpha - E_{\alpha'})\langle\phi_\alpha|\hat{V}|\psi^+_{\alpha'}\rangle.$$

Finally, using $\langle\psi^+_\alpha|\psi^+_{\alpha'}\rangle = \langle\phi_\alpha|\phi_{\alpha'}\rangle = \delta(\alpha - \alpha')$ (why?) and remembering the definition of the transition operator (1.21) we arrive at the important formula

$$\boxed{S_{\alpha\alpha'} = \delta(\alpha - \alpha') - 2\pi i\delta(E_\alpha - E_{\alpha'})\langle\phi_\alpha|\hat{T}|\phi_{\alpha'}\rangle}. \tag{1.64}$$

The first term in this expression is kind of trivial. It expresses the reduction of the S-matrix to a 'unit matrix' in the absence of scattering. The second term makes the conservation of energy in elastic scattering manifest. The residual expression $\langle\phi_\alpha|\hat{T}|\phi_{\alpha'}\rangle$, sometimes designated **on-shell S-matrix** describes the essential scattering processes.

Notice that the frist term $\delta(\alpha - \alpha')$ implies $E_\alpha - E'_\alpha$. In other words, the scattering matrix is 'diagonal' in energy space (overall proportional to $E_\alpha - E_{\alpha'}$.) We can trace this fact back to an important feature of the scattering operator, and of the constituting Møller operators. The representation (1.60) implies (prove it!) $\hat{H}\Omega_\pm = -\hat{\Omega}_\pm\hat{H}_0$. Using these identities in the representation $\hat{S} = \hat{\Omega}^\dagger_-\hat{\Omega}_+$, we readily obtain

$$[\hat{H}_0, \hat{S}] = 0. \tag{1.65}$$

This means that the scattering operator can be diagonalized in the space of eigenfunctions of $\hat{H}_0$. In practice, it will thus be advantageous to represent the scattering matrix in a basis of eigenfunctions of the unperturbed problem.

An advantage of the expression (1.64) above is that it can readily be adjusted to different physical contexts. For example, in the scattering of plane wave states, $\alpha = \mathbf{k}$, the on-shell scattering matrix reduces the scattering amplitudes (cf. Eq. (1.28))

$$\langle\mathbf{k}|\hat{T}|\mathbf{k}'\rangle = -\frac{1}{(2\pi)^2 m} f(\mathbf{k}', \mathbf{k}).$$

EXERCISE Prove the unitarity of the S-matrix on the basis of Eq. (1.64). (This is a tricky one, and all the more instructive. Apply tricks similar to those used in the proof of the optical theorem above.)

1.4.2 Time reversal in scattering theory

Much in this chapter has had to do with the chronology of scattering processes, states incoming and outgoing, causal Green functions etc. In this section, we will explore how the theory behaves under one of the most important discrete symmetry operations of physics, '*time reversal symmetry*'. The name suggests that we are to think about the consequences of an operation that 'reverses the direction of time'. Unfortunately, this allusion can be a source of serious confusion. A symmetry operation is meaningful if we can implement it in practical

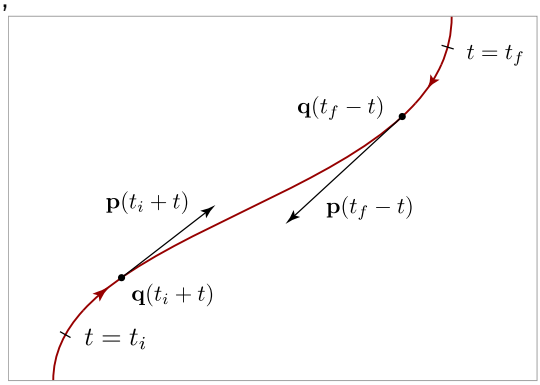
terms, or at least as a matter of principle. No problem with translations, rotation, even parity (space reflection) is a conceivable operation. But time reversal? From a philosophical point of view it may be desirable to emphasize the unity of space and time, and to treat space and time inversion on equal footings. In practice, however, we have no means to invert the flow of time, and hence shouldn't think of a fictitious $t \to -t$ operation in terms of a symmetry, at least not literally.

Yet there is a symmetry that comes close to 'time reversal' in that it expresses the absence of an intrinsic 'arrow of time' in classical or quantum mechanics. What this means is that a movie showing some mechanical motion (think of a ball thrown, or similar) makes sense no matter in which direction you play it.[17] Better than calling it 'time reversal', we might refer to this as 'symmetry under reversal of motion', a term introduced by Wigner in the early thirties. However, the shorter designation '**time reversal**' has stuck, whence we will omit the quotation marks henceforth.

Time reversal: classical

Consider a phase space trajectory $\mathbf{x}(t) \equiv (\mathbf{q}(t), \mathbf{p}(t))$, i.e. a solution of the Hamilton equations of motion $d_t \mathbf{x}(t) = -\{H(\mathbf{x}(t)), \mathbf{x}(t)\}$, where $\{f, g\} \equiv \partial_q f \partial_p g - \partial_p f \partial_q g$ are the Poisson brackets, and $H = H(\mathbf{x}) = \frac{p^2}{2m} + V(q)$ is a Hamilton function.[18] Associated to this trajectory, we may formulate a time reversed partner

$$(T\mathbf{x})(t) \equiv (\mathbf{q}(-t), -\mathbf{p}(-t)).$$



Within our 'thrown ball' analogy, the $(-t)$ means that we play the movie backwards in time, and the $-\mathbf{p}$ means that we invert the direction of motion, cf. the picture, where the time reversal $t_f \to t_i$ of a trajectory segment $t_i \to t_f$ is shown. We may express the observation above by saying that classical mechanics is invariant under the **time reversal symmetry operation**,

$$\begin{cases} t & \to -t, \\ \mathbf{q} & \to +\mathbf{q}, \\ \mathbf{p} & \to -\mathbf{p}. \end{cases} \tag{1.66}$$

Notice that the transformation above changes the sign of the Poisson brackets invariant, $\{\mathbf{q}, (-\mathbf{p})\} = -\{\mathbf{q}, \mathbf{p}\}$, i.e. it violates the otherwise sacrosanct principle of invariance of Poisson brackets under symmetry transformations (reflect on this point.)

[17] To this statement we have to add the important disclaimer that it applies to Newtonian mechanics in the absence of **irreversible forces** such as friction. A rolling ball coming to a halt on a sticky surface makes sense, the reverse does not. How to reconcile irreversibility with the apparent absence of a time arrow in the laws of mechanical motion is a deep and philosophical question (literally) which we dare not touch.

[18] To be precise, the notation above is defined as $\{H(\mathbf{x}(t)), \mathbf{x}(t)\} \equiv \{H(\mathbf{x}), \mathbf{x}\}\big|_{\mathbf{x}=\mathbf{x}(t)}$.

Also notice that there are situations where the time reversal, or reversal of motion symmetry gets lost. For example, a charged particle injected into a region of transverse **magnetic field** will have its motion bended according to a 'right hand rule'. The reverse motion — bending according to a left hand rule — is not physical.

EXERCISE Stop reading and think how to reconcile this observation with the absence of a time arrow in mechanics.

There is no conflict between this observation and what has been said above. If we extended the time reversal operation to the circulating electric currents causing the magnetic field, time reversal would change the sign of the latter, too, and the reverse motion would be physical. However, it is often useful to consider the physics of time reversal in a restricted setting where 'external' magnetic fields are left untouched. One then addresses a scenario where 'time reversal symmetry is broken by an external field'.

Time reversal: quantum

How do we generalized the notion of time reversal to the quantum world. We first observe that a unitary transformation, $\hat{U}$, acting in Hilbert space, $\hat{Q} \to \hat{U}^\dagger \hat{Q} \hat{U}$ will preserve the commutator between operators. Specifically, $[\hat{q}, \hat{p}] \to [(\hat{U}\hat{q}\hat{U}^\dagger), (\hat{U}\hat{p}\hat{U}^\dagger)] = [\hat{q}, \hat{p}]$. In view of the quantum–to–classical correspondence $[\ ,\ ] \leftrightarrow i\hbar\{\ ,\ \}$ between commutators and Poisson brackets, and the *inversion* of the Poisson bracket under time reversal, we anticipate that the time reversal operation will not be describable in terms of a unitary operator.

To get an idea of what the time reversal operation in quantum mechanics will look like, consider the time dependent Schrödinger equation corresponding to the a generic potential Hamilton operator

$$i\partial_t \psi(\mathbf{x}, t) = \left(\frac{1}{2m} (i\partial_{\mathbf{x}})^2 + V(\mathbf{x}) \right) \psi(\mathbf{x}, t). \tag{1.67}$$

Taking the complex conjugate of both sides, and mapping $t \to -t$, we obtain another solution:

$$i\partial_t \overline{\psi(\mathbf{x}, -t)} = \left(\frac{1}{2m} (-i\partial_{\mathbf{x}})^2 + V(\mathbf{x}) \right) \overline{\psi(\mathbf{x}, -t)}. \tag{1.68}$$

Thus, if the mapping $(\mathbf{x}, t) \to \psi(\mathbf{x}, t)$ defines a solution of the Schrödinger equation, then so does $(\mathbf{x}, t) \to \overline{\psi(\mathbf{x}, -t)}$. We thus conclude that

> In the time dependent quantum theory of spinless particles and in a real space coordinate representation, the operation of time reversal is described by Eq. (1.69).

$$\begin{cases} \psi(\mathbf{x}, t) & \to \overline{\psi(\mathbf{x}, -t)} \\ t & \to -t. \end{cases} \tag{1.69}$$

As we will see momentarily, the disclaimers 'time dependent', 'spinless', and 'coordinate representation' in the statement above all have to be taken seriously. We first note that in an energy representation time reversal acts just by complex conjugation $\psi(\mathbf{x}, E) \to \overline{\psi(\mathbf{x}, E)}$, for in the Fourier transform $\overline{\psi(\mathbf{x}, E)} \to \overline{\psi(\mathbf{x}, t)}$ the complex conjugation automatically accounts for the inversion of sign $\overline{\psi(\mathbf{x}, E)} \to \overline{\psi(\mathbf{x}, -t)}$. Second, notice that the *real space* Schrödinger equation above nicely transforms in accord with our classical understanding of time reversal: the real function $V(\mathbf{x})$ remains unaffected, in particular $\mathbf{x} \to \mathbf{x}$, while the momentum operator $(\hat{\mathbf{p}} \leftrightarrow -i\partial_{\mathbf{x}}) \to (-(-i\partial_{\mathbf{x}}) \leftrightarrow -\mathbf{p})$ changes sign. All these features would be messed up, had we subjected the Schrödinger equation in the conjugate momentum representation to a complex conjugation operation. In this representation, $(\hat{\mathbf{p}} \leftrightarrow \mathbf{p})$ is described by a real vector, which would remain unaffected. Also, by the definition of Fourier transforms $\psi(\mathbf{x}, E) \to \overline{\psi(\mathbf{x}, E)}$ corresponds to the transformation $\psi(\mathbf{p}, E) \to \overline{\psi(-\mathbf{p}, E)}$, different from an ordinary complex conjugation of $\psi(\mathbf{p}, E)$ by a sign inversion of the argument. These statements show that the identification of complex conjugation holds only in the coordinate representation of the theory.

It is possible to formulate the action of complex conjugation in terms of an operator $\hat{K}$ acting in Hilbert space. The action of $\hat{K}$ is defined as follows. First declare the coordinate basis $\{|\mathbf{x}\rangle\}$ as a basis which remains invariant under $\hat{K}$:

$$\forall \mathbf{x}: \qquad \hat{K}|\mathbf{x}\rangle = |\mathbf{x}\rangle. \tag{1.70}$$

Next declare that K, unlike conventional operators, acts non-trivially on complex numbers, viz. by complex conjugation $\hat{K}c = \bar{c}\hat{K}$, for $c \in \mathbb{C}$. Since all states in Hilbert space can be represented by linear combination of coordinate basis states — the definition of a basis — the statements above define the action of $\hat{K}$. For example

$$\hat{K}|\mathbf{p}\rangle = \int d^3x \hat{K}\langle \mathbf{x}|\mathbf{p}\rangle|\mathbf{x}\rangle = \int d^3x \overline{\langle \mathbf{x}|\mathbf{p}\rangle}\hat{K}|\mathbf{x}\rangle = \int d^3x \langle \mathbf{x}|-\mathbf{p}\rangle|\mathbf{x}\rangle = |-\mathbf{p}\rangle.$$

where we used $\overline{\langle \mathbf{x}|\mathbf{p}\rangle} = (2\pi)^{-3/2}\overline{\exp(i\mathbf{x}\cdot\mathbf{p})} = (2\pi)^{-3/2}\exp(-i\mathbf{x}\cdot\mathbf{p}) = \langle \mathbf{x}|-\mathbf{p}\rangle$. Also notice that for a general state $|\psi\rangle$,

$$\hat{K}|\psi\rangle = \int d^3x\, \hat{K}\psi(\mathbf{x})|\mathbf{x}\rangle = \overline{\psi(\mathbf{x})}|\mathbf{x}\rangle,$$

i.e. the expansion coefficients of the complex conjugated state simply obtain by complex conjugation of the 'coefficients' $\psi(\mathbf{x})$, in accord with our discussion above.

The definitions above imply that $\hat{K}$ preserves the scalar product of states

$$\langle \psi|\phi\rangle = \int d^3x\, \overline{\psi(\mathbf{x})}\phi(\mathbf{x})$$

up to complex conjugation:

$$\langle \hat{K}\psi|\hat{K}\phi\rangle = \int d^3x\, \overline{\overline{\psi(\mathbf{x})}}\,\overline{\phi(\mathbf{x})} = \int d^3x\, \overline{\phi(\mathbf{x})}\psi(\mathbf{x}) = \langle \phi|\psi\rangle = \overline{\langle \psi|\phi\rangle}. \tag{1.71}$$

An operator with the property (1.71) is called an **anti-unitary operator**. We may, thus, summarize our so far findings by saying that for a solution $|\psi\rangle$ of the Schrödinger equation (of spinless particles) we can construct another solution $\hat{K}|\psi\rangle$ by application of the anti-unitary operator $\hat{K}$=(complex conjugation). This operator acts on the basis states in accord with our classical understanding of time reversal. If we are working in a time dependent representation, the action of $\hat{K}$ has to be augmented with an explicit inversion of time, $t \to -t$, and this, too, conforms with classical intuition. Have we, then, got the quantum description of time reversal under control? The answer is almost, but not quite.

The discussion above applies to spinless particles. Now, **spin** is a quantum analog of classical angular momentum, and the latter we expect to change sign under time reversal! We thus anticipate that spin will behave non-trivially under time reversal operations, and the question is how. We next aim to generalize the operation of time reversal so as to encompasses spin. To this end, we postulate that the general time reversal operation in quantum mechanics is mediated by an anti-unitary operator, $\hat{\Theta}$ (for the proof, see section 4.4 of S.) We second observe that the product of two anti-unitary operators is unitary, an easy consequence of the definition (1.71). Now, the action of $|\psi\rangle \to \hat{\Theta}|\psi\rangle$ on states implies that operators transform as $\hat{A} \to \hat{\Theta}\hat{A}\hat{\Theta}^{-1}$. We require that a two-fold time reversal, mediated by $\hat{\Theta}^2$ must bring any operator back to its original form[19] $\hat{\Theta}^2\hat{A}\hat{\Theta}^{-2} \stackrel{!}{=} \hat{A}$. Now, this condition requires $\hat{\Theta}^2 = z\mathbb{I}$ and, since $\hat{\Theta}^2$ is unitary, $z = \exp(i\phi)$ with a real phase ϕ. Above, we have seen that for spinless quantum states (as described by the conventional Schödinger equation), $\hat{\Theta} = \hat{K}$ does the job. Here, $\hat{\Theta}^2 = \hat{K}^2 = \mathbb{I}$ and the issue of a phase does not arise.

From here on, optional reading. Readers not familiar with the quantum theory of angular momentum may skip the rest of the section.

The situation gets more interesting if we consider **particles with spin**. For simplicity, consider a spin $1/2$ state, as represented by a spinor

$$\psi = \begin{pmatrix} \psi_{\uparrow}, \\ \psi_{\downarrow} \end{pmatrix}$$

Spin is a variant of angular momentum, and angular momentum is to change direction under time reversal. We thus expect $(\hat{\Theta}\psi_{\uparrow/\downarrow}) = \eta_{\downarrow/\uparrow}\overline{\psi_{\downarrow/\uparrow}}$, where $\eta_{\downarrow/\uparrow}$ are phases. Or, in a matrix representation

$$\hat{\Theta}\psi = \begin{pmatrix} & \eta_{\downarrow} \\ \eta_{\uparrow} & \end{pmatrix} \begin{pmatrix} \overline{\psi_{\uparrow}}, \\ \overline{\psi_{\downarrow}} \end{pmatrix},$$

where the anti-unitarity of time reversal requires the η-matrix to be unitary. We verify that this enforces $|\eta_{\uparrow,\downarrow}|^2 = 1$. The above condition $\hat{\Theta}^2 = e^{i\phi}\mathbb{I}$ translates to a second condition (check!)

[19] While the self-involutary action of time reversal on operators is based on solid physical reasoning (e.g. on the requirement that the phase space functions corresponding to operators in the classical limit do not change under double inversion of motion), with wave functions we can be not so sure. We always have to account for the appearance of an (unobservable) global phase different from unity, i.e. $\hat{\Theta}^2|\psi\rangle = (\text{phase}) \times |\psi\rangle$ remains an option. We will soon see that this option is, in fact, realized in nature.

$\overline{\eta_\downarrow}\eta_\uparrow = \eta_\downarrow\overline{\eta_\uparrow} = e^{i\phi}$, which in turn requires $\phi = 0$, like in the spinless case, or $\phi = \pi$. In chapter 3 below, we will see that the latter option is realized in nature. The above conditions are then met by the choice $\eta_\downarrow = 1$, $\eta_\uparrow = -1$, or

$$\boxed{\hat{\Theta}\psi = \hat{U}\hat{K}\psi, \quad \hat{U} = i\sigma_2,} \tag{1.72}$$

where $\sigma_2 = \left(\begin{smallmatrix} & -i \\ i & \end{smallmatrix}\right)$ is a Pauli matrix acting in spin space. Notice that this matrix can be represented as $\hat{U} = \exp(i\pi\hat{J}_y)$, where $\hat{J}_y = \frac{1}{2}\sigma_2$ is the rotation generator around the y-axis. Thus, time reversal amounts to complex conjugation followed by a π-rotation around the y-axis, which is a way to *flip* the spin. It can be shown (cf. S section 4.4) that this statement generalizes to higher angular momentum states:

> The time reversal operation of quantum states carrying angular momentum j is mediated by $\hat{\Theta} = \hat{U}\hat{K}$, where $\hat{U} = \exp(i\pi\hat{J}_y)$.

Now irrespective of the magnitude of j, the operators $\hat{J}_y$ are represented in terms of matrices with purely imaginary entries, i.e. $\hat{K}\hat{U} = \hat{K}\exp(i\pi\hat{J}_y) = \exp(-i\pi(-\hat{J}_y))\hat{K} = \hat{U}\hat{K}$ and this means $\hat{\Theta}^2 = \hat{U}\hat{K}\hat{U}\hat{K} = \hat{U}^2 = \exp(2\pi i\hat{J}_y) = \pm\mathbb{I}$ for j =integer (+) or j =half-integer (-), respectively. (Verify the final equality: consider $\hat{J}_y$ in its eigenbasis, use that the eigenvalues of $\hat{J}_y$ are integer or half-integer valued, and you are almost there.)

Time reversal in scattering

We are now in a position to discuss the ramifications of time reversal in scattering. Heuristically, one expects time reversal to revert a scattering process, and in the following we will give this expectation a quantitative basis. We call a **Hamilton operator time reversal invariant**, if

$$\boxed{\hat{\Theta}\hat{H}\hat{\Theta}^{-1} = \hat{H}.} \tag{1.73}$$

Notice that this equation implies $(i\partial_t - \hat{H})\psi = 0 \Rightarrow \hat{\Theta}(i\partial_t - \hat{H})\psi = \hat{\Theta}(i\partial_t - \hat{H})\hat{\Theta}^{-1}(\hat{\Theta}\psi) = (-i\partial_t - \hat{H})(\hat{\Theta}\psi) = 0$. In other words, the time reversed state obeys a Schrödinger equation governed by the same Hamilton operator, but evolving backwards in time. Also notice that in a real space representation, where $\hat{\Theta} = \hat{K}$ acts by complex conjugation, time reversal implies $\hat{H} = \overline{\hat{H}} = \hat{H}^T$, i.e. the 'matrix' $\hat{H}_{\mathbf{x},\mathbf{x}'} \equiv \langle\mathbf{x}|\hat{H}|\mathbf{x}'\rangle$ becomes real symmetric.

EXERCISE Consider a Hamilton operator in the presence of an external magnetic field and explore how time reversal invariance goes lost.

It is now an easy matter to describe the impact of time reversal on the scattering operator. From the representation (1.62), we obtain

$$\hat{\Theta}\hat{S}\hat{\Theta} = \lim_{\Delta t\to\infty} \hat{\Theta}\left(e^{i\hat{H}_0\Delta t}e^{-2i\hat{H}\Delta t}e^{i\hat{H}_0\Delta t}\right)\hat{\Theta}^{-1} = e^{-i(\hat{\Theta}\hat{H}_0\hat{\Theta}^{-1})\Delta t}e^{2i(\hat{\Theta}\hat{H}\hat{\Theta}^{-1})\Delta t}e^{-i(\hat{\Theta}\hat{H}_0\hat{\Theta}^{-1})\Delta t}.$$

For a time reversal invariant Hamilton operator, this implies

$$\boxed{\hat{\Theta}\hat{S}\hat{\Theta}^{-1} = \hat{S}^{-1}.} \tag{1.74}$$

This conforms with the intuition that time reversal should 'invert' the scattering process.

It's appealing form notwithstanding, the result (1.74) must be processed with caution. In practical applications, one will mostly work with the transition operator (cf. Eq. (1.64)) rather than with the abstract scattering operator. The action of time reversal on the underlying basis states must then be taken into account.

EXERCISE Show that for a time reversal invariant Hamilton operator, and in a basis of plane wave states $\{|\mathbf{k}\rangle\}$,

$$\langle \mathbf{k}|\hat{\Theta}|\mathbf{k}'\rangle = \langle -\mathbf{k}'|\hat{\Theta}|-\mathbf{k}\rangle.$$

(Hint: this is again a tricky one. Use the anti-unitarity relation (1.71), and the representation (1.22).)

1.5 Summary & outlook

This concludes our survey of quantum scattering. We have introduced the conceptual foundations of scattering methods in general, to then develop the basic theory of *elastic* scattering in general. Key concepts such as scattering –cross sections, –amplitudes, –phase shifts, and –length have been introduced. Of course, there is much more that could be said about elastic scattering theory, but the most severe omission of the present chapter is the theory of *inelastic scattering*. Inelastic scattering represents the most important approach to probing the physics of sub-atomic particles. While our discussion above assumed the scattering target to be described by a static potential, 'realistic' targets usually contain internal degrees of freedom — think of a hadron comprising quarks as fundamental constituents. In a scattering process, these constituents may be excited, and the corresponding exchange of energy contains telling information on the structure of the target. Or, the scattering process may trigger a reaction such as the breakup of a nucleus, or the creation of emergent particles in the high energy collision of fundamental scattering targets. While the theory of such processes goes beyond what has been discussed in this chapter, the basic concepts above retain their usefulness. For further information, interested readers are referred to the particle physics literature.

Chapter 2

Second quantization

The purpose of this chapter is to introduce and apply the method of second quantization, a technique that underpins the formulation of quantum many-particle theories. The first part of the chapter focuses on methodology and notation, while the remainder is devoted to the development of applications designed to engender familiarity with and fluency in the technique. Specifically, we will investigate the physics of the interacting electron gas, charge density wave propagation in one-dimensional quantum wires, and spin waves in a quantum Heisenberg (anti)ferromagnet. Indeed, many of these examples and their descendants will reappear as applications in our discussion of methods of quantum field theory in subsequent chapters.

Recapitulate the basic theory of many particle quantum systems

In introductory courses we learn that the quantum mechanics of many particle systems is formulated in 'tensor products'[1] of single particle Hilbert spaces. Wave functions representing $n > 1$ indistinguishable bosons and fermions are then represented in terms of symmetrized or anti-symmetrized sums over products of compound single particle wave amplitudes. The ensuing expressions arguably have a certain formal appeal to them. However, the moment you actually turn to bread–and–butter computations of quantum mechanics — matrix elements of operators, or between states, traces of operators, etc. — it becomes painfully clear that working with many particle wave functions is cumbersome. This is true for few-body wave functions, and the more so for the macroscopically populated wave functions of quantum statistical mechanics and solid state theory. The framework of second quantization provides a language whereby these technical hardships can be reduced to a minimum.

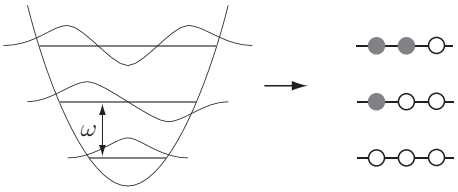
We introduce the basic ideas underlying second quantization on an example which may seem barely related to the points alluded to above – the single particle harmonic oscillator. Only then will we turn to the actual framework of many particle quantum mechanics.

[1] For a revision of these concepts, see below.

Prelude: revision of the quantum harmonic oscillator

Consider a standard harmonic oscillator (HO) Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2.$$



The first few energy levels $\epsilon_n = \omega\left(n + \frac{1}{2}\right)$ and the associated Hermite polynomial eigenfunctions are displayed schematically in the figure.

The HO has, of course, the status of a single-particle problem. However, the equidistance of its energy levels suggests an alternative interpretation. One can think of a given energy state ϵ_n as an accumulation of n elementary entities, or **quasi-particles**, each having energy ω. What can be said about the features of these new objects? First, they are structureless, i.e. the only "quantum number" identifying the quasi-particles is their energy ω (otherwise n-particle states formed of the quasi-particles would not be equidistant). This implies that the quasi-particles must be *bosons*. (The same state ω can be occupied by more than one particle, cf. the figure.)

This idea can be formulated in quantitative terms by employing the formalism of ladder operators in which the operators $\hat{p}$ and $\hat{x}$ are traded for the pair of Hermitian adjoint operators $\hat{a} \equiv \sqrt{\frac{m\omega}{2}}(\hat{x} + \frac{i}{m\omega}\hat{p})$, $\hat{a}^\dagger \equiv \sqrt{\frac{m\omega}{2}}(\hat{x} - \frac{i}{m\omega}\hat{p})$. Up to a factor of i, the transformation $(\hat{x}, \hat{p}) \to (\hat{a}, \hat{a}^\dagger)$ is canonical, i.e. the new operators obey the canonical commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1. \tag{2.1}$$

More importantly, the a-representation of the Hamiltonian is very simple, namely

$$\hat{H} = \omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \tag{2.2}$$

as can be checked by direct substitution. Suppose, now, we had been given a zero eigenvalue state $|0\rangle$ of the operator $\hat{a}$: $\hat{a}|0\rangle = 0$. As a direct consequence, $\hat{H}|0\rangle = (\omega/2)|0\rangle$, i.e. $|0\rangle$ is identified as the ground state of the oscillator.[2] The complete hierarchy of higher energy states can now be generated by setting $|n\rangle \equiv (n!)^{-1/2}\,(\hat{a}^\dagger)^n|0\rangle$.

EXERCISE Using the canonical commutation relation, verify that $\hat{H}|n\rangle = \omega(n + 1/2)|n\rangle$ and $\langle n|n\rangle = 1$.

Formally, the construction above represents yet another way of constructing eigenstates of the quantum HO. However, its "real" advantage is that it naturally affords a many-particle interpretation. To this end, let us *declare* $|0\rangle$ to represent a "vacuum" state, i.e. a state with zero particles present. Next, imagine that $\hat{a}^\dagger|0\rangle$ is a state with a single featureless particle

[2] This can be verified by explicit construction. Switching to a real-space representation, the solution of the equation $[x + \partial_x/(m\omega)]\langle x|0\rangle = 0$ obtains the familiar ground state wavefunction $\langle x|0\rangle = \sqrt{m\omega/2\pi}e^{-m\omega x^2/2}$.

(the operator $\hat{a}^\dagger$ does not carry any quantum number labels) of energy ω. Similarly, $(\hat{a}^\dagger)^n|0\rangle$ is considered as a many-body state with n particles, i.e. within the new picture, $\hat{a}^\dagger$ is an operator that creates particles. The total energy of these states is given by $\omega \times$ (occupation number). Indeed, it is straightforward to verify (see exercise above) that $\hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle$, i.e. the Hamiltonian basically counts the number of particles. While, at first sight, this may look unfamiliar, the new interpretation is internally consistent. Moreover, it achieves what we had asked for above, i.e. it allows an interpretation of the HO states as a superposition of independent structureless entities.

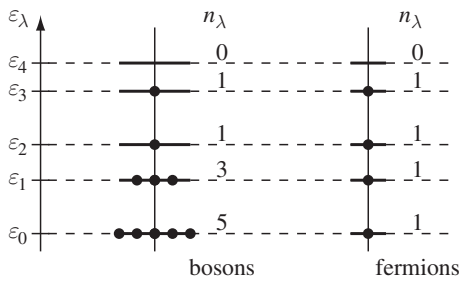
INFO The representation above illustrates the capacity to think about individual quantum problems in **complementary pictures**. This principle finds innumerable applications in modern physics. The existence of different interpretations of a given system is by no means heretical but, rather, reflects a principle of quantum mechanics: there is no "absolute" system that underpins the phenomenology. The only thing that matters is observable phenomena. For example, we will see later that the "fictitious" quasi-particle states of oscillator systems *behave* as "real" particles, i.e. they have dynamics, can interact, be detected experimentally, etc. From a quantum point of view these object are, then, real particles.

2.1 Introduction to second quantization

Apart from a certain aesthetic appeal, the above discussion would not be of much relevance if it were not for the fact that it can be developed to a comprehensive and highly efficient formulation of many-body quantum mechanics in general — the apparatus of second quantization[3]. In fact, second quantization can be considered the first major cornerstone on which the theoretical framework of quantum field theory was built. This being so, extensive introductions to the concept can be found throughout the literature. We will therefore not develop the formalism in full mathematical rigor but rather proceed pragmatically by first motivating and introducing its basic elements, followed by a discussion of the "second quantized" version of standard operations of quantum mechanics (taking matrix elements, changing bases, representing operators, etc.). The second part of the chapter is concerned with developing fluency in the method by addressing a few applications.

Motivation

[3] The term "second quantization" is unfortunate. Historically, this terminology was motivated by the observation that the ladder operator algebra fosters an interpretation of quantum excitations as discrete "quantized" units. Fundamentally, however, there is nothing like "two" superimposed quantization steps in single- or many-particle quantum mechanics. Rather, one is dealing with a particular representation of the "first and only quantized" theory tailored to the particular problem at hand.



We begin our discussion by recapitulating some fundamental notions of many-body quantum mechanics, as formulated in the traditional language of symmetrized or anti-symmetrized wavefunctions. Consider the (normalized) set of wavefunctions $|\lambda\rangle$ of some single-particle Hamiltonian $\hat{H} : \hat{H}|\lambda\rangle = \epsilon_\lambda|\lambda\rangle$, where ϵ_λ are the eigenvalues. With this definition, the normalized two-particle wavefunction $\psi_{\rm F}(\psi_{\rm B})$ of two fermions (bosons) populating levels λ_1 and λ_2 is given by the anti-symmetrized (symmetrized) product

$$\begin{aligned}\psi_{\rm F}(x_1,x_2) &= \frac{1}{\sqrt{2}}\left(\langle x_1|\lambda_1\rangle\langle x_2|\lambda_2\rangle - \langle x_1|\lambda_2\rangle\langle x_2|\lambda_1\rangle\right),\\ \psi_{\rm B}(x_1,x_2) &= \frac{1}{\sqrt{2}}\left(\langle x_1|\lambda_1\rangle\langle x_2|\lambda_2\rangle + \langle x_1|\lambda_2\rangle\langle x_2|\lambda_1\rangle\right).\end{aligned}$$

In the Dirac bracket representation, the two-body states $|\lambda_1,\lambda_2\rangle_{\rm F(B)}$ corresponding to the wave functions $\psi_{\rm F(B)}(x_1,x_2) = \left(\langle x_1|\otimes\langle x_2|\right)|\lambda_1,\lambda_2\rangle_{\rm F(B)}$ above can be presented as

$$|\lambda_1,\lambda_2\rangle_{\rm F(B)} \equiv \frac{1}{\sqrt{2}}\left(|\lambda_1\rangle\otimes|\lambda_2\rangle + \zeta|\lambda_2\rangle\otimes|\lambda_1\rangle\right),$$

where $\zeta = -1$ for fermions while $\zeta = 1$ for bosons.

Note that the explicit symmetrization of the wavefunctions is necessitated by quantum mechanical **indistinguishability**: for fermions (bosons) the wave function has to be anti-symmetric (symmetric) under particle exchange.[4] More generally, an appropriately symmetrized N-particle wavefunction can be expressed in the form

$$\boxed{|\lambda_1,\lambda_2,\ldots,\lambda_N\rangle \equiv \frac{1}{\sqrt{N!\prod_{\lambda=0}^{\infty}(n_\lambda!)}}\sum_{\mathcal{P}}\zeta^{(1-\mathrm{sgn}\,\mathcal{P})/2}|\lambda_{\mathcal{P}1}\rangle\otimes|\lambda_{\mathcal{P}2}\rangle\otimes\cdots\otimes|\lambda_{\mathcal{P}N}\rangle,} \quad (2.3)$$

where n_λ represents the total number of particles in state λ (for fermions, Pauli exclusion enforces the constraint $n_\lambda \leq 1$) – see the schematic figure above. The summation runs over all $N!$ permutations of the set of quantum numbers $\{\lambda_1,\ldots,\lambda_N\}$, and $\mathrm{sgn}\,\mathcal{P}$ denotes the sign of the permutation $\mathcal{P}$. ($\mathrm{sgn}\,\mathcal{P} = 1\,[-1]$ if the number of transpositions of two elements which brings the permutation $(\mathcal{P}_1,\mathcal{P}_2,\ldots,\mathcal{P}_N)$ back to its original form $(1,2,\ldots,N)$ is even [odd].) The prefactor $1/\sqrt{N!\prod_\lambda(n_\lambda!)}$ normalizes the many-body wavefunction. In the fermionic case, wave functions corresponding to the states above are known as **Slater determinants**.

Finally, notice that it will be useful to assume that the quantum numbers $\{\lambda_i\}$ defining the state $|\lambda_1,\lambda_2,\ldots,\lambda_N\rangle$ are ordered according to some convention. (For example, for $\lambda_i = x_i$ a one-dimensional coordinate representation, we might order according to the rule $x_1 \leq x_2 \leq \cdots \leq x_N$.) Once an ordered sequence of states has been fixed we may – for notational

[4] Notice, however, that in two-dimensional systems the standard doctrine of fully symmetric/anti-symmetric many particle wave functions is too narrow and more general types of exchange statistics can be realized, cf. our discussion on page 49.

convenience – label our quantum states by integers, $\lambda_i = 1, 2, \ldots$. Any initially non-ordered state (e.g. $|2,1,3\rangle$) can be brought into an ordered form ($|1,2,3\rangle$) at the cost of, at most, a change of sign.

INFO For the sake of completeness, let us spell out the connection between the **permutation group and many-body quantum mechanics** in a more mathematical language. The basic arena wherein N-body quantum mechanics takes place is the product space,

$$\mathcal{H}^N \equiv \underbrace{\mathcal{H} \otimes \cdots \otimes \mathcal{H}}_{N \text{ copies}},$$

of N single-particle Hilbert spaces. In this space, we have a linear representation of the permutation group, S^N,[5] assigning to each $\mathcal{P} \in S^N$ the permutation (no ordering of the λs implied at this stage),

$$\mathcal{P} : \mathcal{H}^N \quad \rightarrow \quad \mathcal{H}^N, \quad |\lambda_1\rangle \otimes \cdots \otimes |\lambda_N\rangle \mapsto |\lambda_{\mathcal{P}1}\rangle \otimes \cdots \otimes |\lambda_{\mathcal{P}N}\rangle.$$

The identification of all irreducible subspaces of this representation is a formidable task that, thanks to a fundamental axiom of quantum mechanics, we need not address in full. All we need to know is that S^N has two particularly simple one-dimensional irreducible representations: one wherein each $\mathcal{P} \in S^N$ acts as the identity transform $\mathcal{P}(\Psi) \equiv \Psi$ and, another, the alternating representation $\mathcal{P}(\Psi) = \mathrm{sgn}\,\mathcal{P} \cdot \Psi$. According to a basic postulate of quantum mechanics, the state vectors $\Psi \in \mathcal{H}^N$ describing bosons/fermions must transform according to the identity/alternating representation. The subset $\mathcal{F}^N \subset \mathcal{H}^N$ of all states showing this transformation behavior defines the physical N-body Hilbert space. To construct a basis of $\mathcal{F}^N$, one may apply the symmetrization operator $P^s \equiv \sum_{\mathcal{P}} \mathcal{P}$ (anti-symmetrization operator $P^a \equiv \sum_{\mathcal{P}} (\mathrm{sgn}\,\mathcal{P})\mathcal{P}$) to the basis vectors $|\lambda_1\rangle \otimes \cdots \otimes |\lambda_N\rangle$ of $\mathcal{H}^N$. Up to normalization, this operation obtains the states (2.3).

Some readers may wonder why we mention these representation-theoretic aspects. Being pragmatic, all we really need to know is the symmetrization/anti-symmetrization postulate, and its implementation through Eq. (2.3). Notice, however, that one may justly question what we actually mean when we talk about the permutation exchange of quantum numbers. For example, when we compare wavefunctions that differ by an exchange of coordinates, we should, at least in principle, be able to tell by what physical operation we effect this exchange (for, otherwise, we cannot really compare them other than in a formal and, in fact, in an ambiguous sense).

Oddly enough, decades passed before this crucial issue in quantum mechanics was critically addressed. In a now seminal work by Leinaas and Myrheim[6] it was shown that the standard paradigm of permutation exchange is far from innocent. In particular, it turned out that its applicability is

[5] Recall that a **linear representation** of a group G is a mapping that assigns to each $g \in G$ a linear mapping $\rho_g : V \rightarrow V$ of some vector space V. For notational convenience one usually writes $g : V \rightarrow V$ instead of $\rho_g : V \rightarrow V$. Conceptually, however, it is often important to carefully discriminate between the abstract group elements g and the *matrices* (also g) assigned to them by a given representation. (Consider, for example the symmetry group $G = \mathrm{SU}(2)$ of quantum mechanical spin. $\mathrm{SU}(2)$ is the two-dimensional group of unitary matrices with determinant one. However, when acting in the Hilbert space of a quantum spin $S = 5$, say, elements of $\mathrm{SU}(2)$ are represented by $(2S + 1 = 11)$-dimensional matrices.) Two representations ρ and ρ' that differ only by a unitary transformation, $\forall g \in G : \rho_g = U \rho'_g U^{-1}$, are called unitary equivalent. If a transformation U can be found such that all representation matrices ρ_g assume a block structure, the representation is called reducible, and otherwise **irreducible**. Notice that the different sub-blocks of a reducible representation by themselves form irreducible representation spaces. The identification of all distinct irreducible representations of a given group is one of the most important objectives of group theory.

[6] J. M. Leinaas and J. Myrheim, On the theory of identical particles, Il *Nuovo Cimento* B **37** (1977), 1–23.

tied to the dimensionality of space! Specifically, in two-dimensional spaces (in a sense, also in $d = 1$) a more elaborate scheme is needed. (Still one may use representation-theoretic concepts to describe particle exchange. However, the relevant group – the **braid group** – now differs from the permutation group.) Physically, these phenomena manifest themselves in the emergence of quantum particles different from both bosons and fermions.

While representations like Eq. (2.3) *can* be used to represent the full Hilbert space of many-body quantum mechanics, a moment's thought shows that this formulation is not at all convenient:

- ▷ It takes little imagination to anticipate that practical computation in the language of Eq. (2.3) will be cumbersome. For example, to compute the overlap of two wavefunctions one needs to form no less than $(N!)^2$ different products.
- ▷ The representation is tailor-made for problems with fixed particle number N. However, we know from statistical mechanics that for $N = \mathcal{O}(10^{23})$ it is much more convenient to work in a grand canonical formulation where N is allowed to fluctuate.
- ▷ Closely related to the above, in applications one will often ask questions such as, "What is the amplitude for injection of a particle into the system at a certain space-time coordinate (x_1, t_1) followed by annihilation at some later time (x_2, t_2)?" Ideally, one would work with a representation that supports the intuition afforded by thinking in terms of such processes: i.e. a representation where the quantum numbers of individual quasi-particles rather than the entangled set of quantum numbers of all constituents are fundamental.

The "second quantized" formulation of many-body quantum mechanics, as introduced in the next subsection, will remove all these difficulties in an elegant and efficient manner.

The apparatus of second quantization

Occupation number representation and Fock space

Some of the disadvantages of the representation (2.3) can be avoided with relatively little effort. In our present notation, quantum states are represented by "N-letter words" of the form $|1, 1, 1, 1, 2, 2, 3, 3, 3, 4, 6, 6, \ldots\rangle$. Obviously, this notation contains a lot of redundancy. A more efficient encoding of the state above might read $|4, 2, 3, 1, 0, 2, \ldots\rangle$, where the ith number signals how many particles occupy state number i; no more information is needed to characterize a symmetrized state. (For fermions, these occupation numbers take a value of either zero or one.) This defines the "**occupation number representation**." In the new representation, the basis states of $\mathcal{F}^N$ are specified by $|n_1, n_2, \ldots\rangle$, where $\sum_i n_i = N$. Any state $|\Psi\rangle$ in $\mathcal{F}^N$ can be obtained by a linear superposition

$$|\Psi\rangle = \sum_{\substack{n_1, n_2, \ldots, \\ \sum n_i = N}} c_{n_1, n_2, \ldots} |n_1, n_2, \ldots\rangle.$$

As pointed out above, eventually we will want to emancipate ourselves from the condition of a fixed particle number N. A Hilbert space large enough to accommodate a state with an undetermined number of particles is given by

$$\mathcal{F} \equiv \bigoplus_{N=0}^{\infty} \mathcal{F}^N. \tag{2.4}$$

Notice that the direct sum contains a curious contribution $\mathcal{F}^0$, the "vacuum space." This is a one-dimensional Hilbert space which describes the sector of the theory with no particles present. Its single normalized basis state, the **vacuum state**, is denoted by $|0\rangle$. We will soon see why it is convenient to add this strange animal to our family of basis states. The space $\mathcal{F}$ is called **Fock space** and it defines the principal arena of quantum many-body theory.

To obtain a basis of $\mathcal{F}$, we need only take the totality of our previous basis states $\{|n_1, n_2, \ldots\rangle\}$, and drop the condition $\sum_i n_i = N$ on the occupation numbers. A general many-body state $|\Psi\rangle$ can then be represented by a linear superposition $|\Psi\rangle = \sum_{n_1,n_2,\ldots} c_{n_1,n_2,\ldots}|n_1, n_2, \ldots\rangle$. Notice that states of different particle numbers may contribute to the linear superposition forming $|\Psi\rangle$.

Foundations of second quantization

The occupation number representation introduced above provides a step in the right direction, but it does not yet solve our main problem, the need for explicit symmetrization/antisymmetrization of a large number of states in each quantum operation.

As a first step towards the construction of a more powerful representation, let us recall an elementary fact of linear algebra: a linear map $A : V \to V$ of a vector space into itself is fully determined by defining the images $w_i \equiv Av_i$ of the action of A on a basis $\{v_i\}$. Now let us use this scheme to introduce a set of linear operators acting in Fock space. For every $i = 1, 2, \ldots$, we define operators $a_i^\dagger : \mathcal{F} \to \mathcal{F}$ through

$$a_i^\dagger |n_1, \ldots, n_i, \ldots\rangle \equiv (n_i + 1)^{1/2} \zeta^{s_i} |n_1, \ldots, n_i + 1, \ldots\rangle, \tag{2.5}$$

where $s_i = \sum_{j=1}^{i-1} n_j$. In the fermionic case, the occupation numbers n_i have to be understood $\mathrm{mod}\, 2$. Specifically, $(1 + 1) = 0 \ \mathrm{mod}\, 2$, i.e. the application of $a_i^\dagger$ to a state with $n_i = 1$ annihilates this state.

Notice that by virtue of this definition we are able to generate every basis state of $\mathcal{F}$ by repeated application of $a_i^\dagger$s to the vacuum state. (From a formal point of view, this fact alone is motivation enough to add the vacuum space to the definition of Fock space.) Indeed, repeated application of Eq. (2.5) leads to the important relation

$$\boxed{|n_1, n_2, \ldots\rangle = \prod_i \frac{1}{(n_i!)^{1/2}} (a_i^\dagger)^{n_i} |0\rangle.} \tag{2.6}$$

Notice that Eq. (2.6) presents a strong statement: the complicated permutation "entanglement" implied in the definition (2.3) of the Fock states can be generated by straightforward

application of a set of linear operators to a single reference state. Physically, N-fold application of operators $a^\dagger$ to the empty vacuum state generates an N-particle state, which is why the $a^\dagger$s are commonly called **creation operators**. Of course, the introduction of creation operators might still turn out to be useless, i.e. consistency with the properties of the Fock states (such as the fact that, in the fermionic case, the numbers $n_i = 0, 1$ are defined only $\mathrm{mod}\, 2$), might invalidate the simple relation (2.5) with its (n_i-independent!) operator $a_i^\dagger$. However, as we shall demonstrate below, this is not the case.

Consider two operators $a_i^\dagger$ and $a_j^\dagger$ for $i \neq j$. From the definition (2.5), one may readily verify that $(a_i^\dagger a_j^\dagger - \zeta a_j^\dagger a_i^\dagger)|n_1, n_2, \ldots\rangle = 0$. Holding for every basis vector, this relation implies that $[a_i^\dagger, a_j^\dagger]_\zeta = 0$, where

$$[\hat{A}, \hat{B}]_\zeta \equiv \hat{A}\hat{B} - \zeta \hat{B}\hat{A},$$

i.e. $[\,,\,]_{\zeta=1} \equiv [\,,\,]$ is the commutator and $[\,,\,]_{\zeta=-1} \equiv \{\,,\,\} \equiv [\,,\,]_+$ the anti-commutator. Turning to the case $i = j$, we note that, for fermions, the two-fold application of $a_i^\dagger$ to any state leads to its annihilation. Thus, $a_i^{\dagger 2} = 0$ is nilpotent, a fact that can be formulated as $[a_i^\dagger, a_i^\dagger]_+ = 0$. For bosons we have, of course, $[a_i^\dagger, a_i^\dagger] = 0$ (identical operators commute!). Summarizing, we have found that the creation operators obey the commutation relation

$$\forall i, j: \quad \left[a_i^\dagger, a_j^\dagger\right]_\zeta = 0. \tag{2.7}$$

Now, quantum mechanics is a unitary theory so, whenever one meets a new operator $\hat{A}$, one should determine its Hermitian adjoint $\hat{A}^\dagger$. To understand the action of the Hermitian adjoints $\left(a_i^\dagger\right)^\dagger = a_i$ of the creation operators we may take the complex conjugates of all basis matrix elements of Eq. (2.5):

$$\begin{aligned} & \langle n_1, \ldots, n_i, \ldots | a_i^\dagger | n_1', \ldots, n_i', \ldots\rangle &&= \left(n_i' + 1\right)^{1/2} \zeta^{s_i'} \delta_{n_1, n_1'} \ldots \delta_{n_i, n_i'+1} \ldots \\ \Rightarrow \quad & \langle n_1', \ldots, n_i', \ldots | a_i | n_1, \ldots, n_i, \ldots\rangle^* &&= n_i^{1/2} \zeta^{s_i} \delta_{n_1', n_1} \ldots \delta_{n_i', n_i - 1} \ldots \end{aligned}$$

Holding for every bra $\langle n_1', \ldots, n_i', \ldots|$, the last line tells us that

$$a_i |n_1, \ldots, n_i, \ldots\rangle = n_i^{1/2} \zeta^{s_i} |n_1, \ldots, n_i - 1, \ldots\rangle, \tag{2.8}$$

a relation that identifies a_i as an operator that "annihilates" particles. The action of creation and **annihilation operators** in Fock space is illustrated in Fig. 2.1. Creation operators $a^\dagger : \mathcal{F}^N \to \mathcal{F}^{N+1}$ increase the particle number by one, while annihilation operators $a : \mathcal{F}^N \to \mathcal{F}^{N-1}$ lower it by one; the application of an annihilation operator to the vacuum state, $a_i|0\rangle = 0$, annihilates it. (Do not confuse the vector $|0\rangle$ with the number zero.)

Taking the Hermitian adjoint of Eq. (2.7) we obtain $[a_i, a_j]_\zeta = 0$. Further, a straightforward calculation based on the definitions (2.5) and (2.8) shows that $[a_i, a_j^\dagger]_\zeta = \delta_{ij}$. Altogether, we have shown that the creation and annihilation operators satisfy the algebraic closure relation

$$\boxed{[a_i, a_j^\dagger]_\zeta = \delta_{ij}, \quad [a_i, a_j]_\zeta = 0, \quad [a_i^\dagger, a_j^\dagger]_\zeta = 0.} \tag{2.9}$$

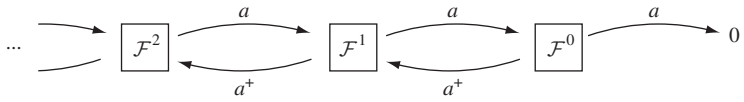


Figure 2.1: Visualization of the generation of the Fock subspaces $\mathcal{F}^N$ by repeated action of creation operators on the vacuum space $\mathcal{F}^0$.

Given that the full complexity of Fock space is generated by application of $a_i^\dagger$s to a single reference state, the simplicity of the relations obeyed by these operators seems remarkable and surprising.

INFO Perhaps less surprising is that, behind this phenomenon, there lingers some mathematical structure. Suppose we are given an abstract algebra $\mathcal{A}$ of objects $a_i, a_i^\dagger$ satisfying the relation (2.9). (Recall that an algebra is a vector space whose elements can be multiplied by each other.) Further suppose that $\mathcal{A}$ is irreducibly represented in some vector space V, i.e. that there is a mapping assigning to each $a_i \in \mathcal{A}$ a linear mapping $a_i : V \to V$, such that every vector $|v\rangle \in V$ can be reached from any other $|w\rangle \in V$ by (possibly iterated) application of operators a_i and $a_i^\dagger$ (irreducibility).[7] According to the **Stone–von Neumann theorem** (a) such a representation is unique (up to unitary equivalence), and (b) there is a unique state $|0\rangle \in V$ that is annihilated by every a_i. All other states can then be reached by repeated application of $a_i^\dagger$s. The precise formulation of this theorem, and its proof – a good practical exercise in working with creation/annihilation operators – are left as a challenging exercise. From the Stone–von Neumann theorem, we can infer that the Fock space basis could have been constructed in reverse. Not knowing the basis $\{|n_1, n_2, \ldots\rangle\}$, we could have started from a set of operators obeying the commutation relations (2.9) acting in some a priori unknown space $\mathcal{F}$. Outgoing from the unique state $|0\rangle$, the prescription (2.6) would then have yielded an equally unique basis of the entire space $\mathcal{F}$ (up to unitary transformations). In other words, the algebra (2.9) fully characterizes the operator action and provides information equivalent to the definitions (2.5) and (2.8).

Practical aspects

Our next task will be to promote the characterization of Fock space bases introduced above to a full reformulation of many-body quantum mechanics. To this end, we need to find out how changes from one single-particle basis $\{|\lambda\rangle\}$ to another $\{|\tilde{\lambda}\rangle\}$ affect the operator algebra $\{a_\lambda\}$. (In this section we shall no longer use integers to identify different elements of a given single-particle basis. Rather, we use Greek labels λ, i.e. $a_\lambda^\dagger$ creates a particle in state λ.) Equally important, we need to understand in what way generic operators acting in many-particle Hilbert spaces can be represented in terms of creation and annihilation operators.

[7] To appropriately characterize the representation, we need to be a bit more precise. Within $\mathcal{A}$, a_i and $a_i^\dagger$ are independent objects, i.e. in general there exists no notion of Hermitian adjointness in $\mathcal{A}$. We require, though, that the representation assigns to $a_i^\dagger$ the Hermitian adjoint (in V) of the image of a_i. Also, we have to require that $[a_i, a_j^\dagger] \in \mathcal{A}$ be mapped onto $[a_i, a_j^\dagger] : V \to V$ where, in the latter expression, the commutator involves the ordinary product of matrices $a_i, a_j^\dagger : V \to V$.

▷ **Change of basis:** Using the resolution of identity $\mathrm{id} = \sum_{\lambda=0}^{\infty} |\lambda\rangle\langle\lambda|$, the relations $|\tilde{\lambda}\rangle = \sum_\lambda |\lambda\rangle\langle\lambda|\tilde{\lambda}\rangle$, $|\lambda\rangle \equiv a_\lambda^\dagger|0\rangle$, and $|\tilde{\lambda}\rangle \equiv a_{\tilde{\lambda}}^\dagger|0\rangle$ immediately give rise to the transformation law

$$\boxed{a_{\tilde{\lambda}}^\dagger = \sum_\lambda \langle\lambda|\tilde{\lambda}\rangle a_\lambda^\dagger, \qquad a_{\tilde{\lambda}} = \sum_\lambda \langle\tilde{\lambda}|\lambda\rangle a_\lambda.} \tag{2.10}$$

In applications, one is often dealing with continuous sets of quantum numbers (such as position coordinates). In these cases, the quantum numbers are commonly denoted by a bracket notation $a_\lambda \to a(x) = \sum_\lambda \langle x|\lambda\rangle a_\lambda$ and the summations appearing in the transformation formula above translate to integrals: $a_\lambda = \int dx \langle\lambda|x\rangle a(x)$.

EXAMPLE The transformation from the coordinate to the Fourier momentum representation in a finite one-dimensional system of length L would read

$$a_k = \int_0^L dx\, \langle k|x\rangle a(x), \qquad a(x) = \sum_k \langle x|k\rangle a_k, \tag{2.11}$$

where $\langle k|x\rangle \equiv \langle x|k\rangle^* = e^{-ikx}/\sqrt{L}$. In a three-dimensional extended system, the transformation assumes the form

$$a(\mathbf{k}) = \int d^3x\, \langle \mathbf{k}|\mathbf{x}\rangle a(\mathbf{x}), \qquad a(\mathbf{x}) = \int d^3k\, \langle \mathbf{x}|\mathbf{k}\rangle a(\mathbf{k}), \tag{2.12}$$

where $\langle \mathbf{k}|\mathbf{x}\rangle \equiv \langle \mathbf{x}|\mathbf{k}\rangle^* = e^{-i\mathbf{k}\cdot\mathbf{x}}/(2\pi)^{3/2}$, etc.

▷ **Representation of operators (one-body):** Single-particle or one-body operators $\hat{\mathcal{O}}_1$ acting in the N-particle Hilbert space $\mathcal{F}^N$ generally take the form $\hat{\mathcal{O}}_1 = \sum_{n=1}^N \hat{o}_n$, where $\hat{o}_n$ is an ordinary single-particle operator acting on the nth particle. A typical example is the kinetic energy operator $\hat{T} = \sum_n \hat{p}_n^2/2m$, where $\hat{p}_n$ is the momentum operator acting on the nth particle. Other examples include the one-particle potential operator $\hat{V} = \sum_n V(\hat{x}_n)$, where $V(x)$ is a scalar potential, the total spin operator $\sum_n \hat{\mathbf{S}}_n$, etc. Since we have seen that, by applying field operators to the vacuum space, we can generate the Fock space in general and any N-particle Hilbert space in particular, it must be possible to represent any operator $\hat{\mathcal{O}}_1$ in an a-representation.

Now, although the representation of n-body operators is, after all, quite straightforward, the construction can, at first sight, seem daunting. A convenient way of finding such a representation is to express the operator in terms of a basis in which it is diagonal, and only later transform to an arbitrary basis. For this purpose it is useful to define the **occupation number operator**

$$\boxed{\hat{n}_\lambda = a_\lambda^\dagger a_\lambda}, \tag{2.13}$$

with the property that, for bosons or fermions (exercise), $\hat{n}_\lambda \left(a_\lambda^\dagger\right)^n|0\rangle = n\left(a_\lambda^\dagger\right)^n|0\rangle$. Since $\hat{n}_\lambda$ commutes with all $a_{\lambda'\neq\lambda}^\dagger$, Eq. (2.6) readily implies that $\hat{n}_{\lambda_j}|n_{\lambda_1}, n_{\lambda_2}, \ldots\rangle = n_{\lambda_j}|n_{\lambda_1}, n_{\lambda_2}, \ldots\rangle$, i.e., $\hat{n}_\lambda$ simply counts the number of particles in state λ (hence the

name "occupation number operator"). Let us now consider a one-body operator, $\hat{\mathcal{O}}_1$, which is diagonal in the basis $|\lambda\rangle$, with $\hat{o} = \sum_i o_{\lambda_i}|\lambda_i\rangle\langle\lambda_i|$, $o_{\lambda_i} = \langle\lambda_i|\hat{o}|\lambda_i\rangle$. With this definition, one finds that

$$\begin{aligned}\langle n'_{\lambda_1}, n'_{\lambda_2}, \ldots|\hat{\mathcal{O}}_1|n_{\lambda_1}, n_{\lambda_2}, \ldots\rangle &= \sum_i o_{\lambda_i} n_{\lambda_i}\langle n'_{\lambda_1}, n'_{\lambda_2}, \ldots|n_{\lambda_1}, n_{\lambda_2}, \ldots\rangle \\ &= \langle n'_{\lambda_1}, n'_{\lambda_2}, \ldots|\sum_i o_{\lambda_i}\hat{n}_{\lambda_i}|n_{\lambda_1}, n_{\lambda_2}, \ldots\rangle.\end{aligned}$$

Since this equality holds for any set of states, one can infer the second quantized representation of the operator $\hat{\mathcal{O}}_1$,

$$\hat{\mathcal{O}}_1 = \sum_{\lambda=0}^{\infty} o_\lambda \hat{n}_\lambda = \sum_{\lambda=0}^{\infty}\langle\lambda|\hat{o}|\lambda\rangle a^\dagger_\lambda a_\lambda.$$

The result is straightforward: a one-body operator engages a single particle at a time – the others are just spectators. In the diagonal representation, one simply counts the number of particles in a state λ and multiplies by the corresponding eigenvalue of the one-body operator. Finally, by transforming from the diagonal representation to a general basis, one obtains the general result,

$$\boxed{\hat{\mathcal{O}}_1 = \sum_{\mu\nu}\langle\mu|\hat{o}|\nu\rangle a^\dagger_\mu a_\nu.} \tag{2.14}$$

To cement these ideas, let us consider some specific examples: representing the matrix elements of the single-particle spin operator as $(S_i)_{\alpha\alpha'} = \frac{1}{2}(\sigma_i)_{\alpha\alpha'}$, where α, α' is a two-component spin index and σ_i are the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.15}$$

the **spin operator** of a many-body system assumes the form

$$\hat{\mathbf{S}} = \sum_\lambda a^\dagger_{\lambda\alpha'}\mathbf{S}_{\alpha'\alpha}a_{\lambda\alpha}. \tag{2.16}$$

(Here, λ denotes the set of additional quantum numbers, e.g. a lattice site index.) When second quantized in the position representation, one can show that the **one-body Hamiltonian** for a free particle is given as a sum of kinetic and potential energy as

$$\boxed{\hat{H} = \int d^d r\ a^\dagger(\mathbf{r})\left[\frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{r})\right]a(\mathbf{r}),} \tag{2.17}$$

where $\hat{\mathbf{p}} = -i\hbar\partial$.

EXERCISE Starting with momentum representation (in which the kinetic energy is diagonal), transform to the position representation and thereby establish Eq. (2.17).

The **local density operator** $\hat{\rho}(\mathbf{r})$, measuring the particle density at a certain coordinate $\mathbf{r}$, is simply given by

$$\hat{\rho}(\mathbf{r}) = a^\dagger(\mathbf{r})a(\mathbf{r}). \tag{2.18}$$

Finally, the **total occupation number operator**, obtained by integrating over the particle density, is defined by $\hat{N} = \int d^d r\ a^\dagger(\mathbf{r})a(\mathbf{r})$. In a theory with discrete quantum numbers, this operator assumes the form $\hat{N} = \sum_\lambda a_\lambda^\dagger a_\lambda$.

▷ **Representation of operators (two-body):** Two-body operators $\hat{\mathcal{O}}_2$ are needed to describe pairwise interactions between particles. Although pair-interaction potentials are straightforwardly included in classical many-body theories, their embedding into conventional many-body quantum mechanics is made cumbersome by particle indistinguishability. The formulation of interaction processes within the language of second quantization is considerably more straightforward:

Initially, let us consider particles subject to the symmetric two-body potential $V(\mathbf{r}_m, \mathbf{r}_n) \equiv V(\mathbf{r}_n, \mathbf{r}_m)$ between two particles at position $\mathbf{r}_m$ and $\mathbf{r}_n$. Our aim is to find an operator $\hat{V}$ in second quantized form whose action on a many-body state gives (presently, it is more convenient to use the original representation Eq. (2.3) rather than the occupation number representation)

$$\hat{V}|\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N\rangle = \sum_{n<m}^{N} V(\mathbf{r}_n, \mathbf{r}_m)|\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N\rangle = \frac{1}{2}\sum_{n\neq m}^{N} V(\mathbf{r}_n, \mathbf{r}_m)|\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N\rangle.$$

When this is compared with the one-point function, one might immediately guess that

$$\hat{V} = \frac{1}{2}\int d^d r \int d^d r' a^\dagger(\mathbf{r})a^\dagger(\mathbf{r}')V(\mathbf{r}, \mathbf{r}')a(\mathbf{r}')a(\mathbf{r}). \tag{2.19}$$

That this is the correct answer can be confirmed by applying the operator to a many-body state:

$$\begin{aligned}
a^\dagger(\mathbf{r})a^\dagger(\mathbf{r}')a(\mathbf{r}')a(\mathbf{r})|\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\rangle &= a^\dagger(\mathbf{r})a^\dagger(\mathbf{r}')a(\mathbf{r}')a(\mathbf{r})a^\dagger(\mathbf{r}_1)\cdots a^\dagger(\mathbf{r}_N)|0\rangle \\
&= \sum_{n=1}^{N} \zeta^{n-1}\delta(\mathbf{r}-\mathbf{r}_n)a^\dagger(\mathbf{r}_n)\overbrace{a^\dagger(\mathbf{r}')a(\mathbf{r}')}^{\hat{\rho}(\mathbf{r}')}a^\dagger(\mathbf{r}_1)\cdots a^\dagger(\mathbf{r}_{n-1})a^\dagger(\mathbf{r}_{n+1})\cdots a^\dagger(\mathbf{r}_N)|0\rangle \\
&= \sum_{n=1}^{N} \zeta^{n-1}\delta(\mathbf{r}-\mathbf{r}_n)\sum_{m(\neq n)}^{N}\delta(\mathbf{r}'-\mathbf{r}_m)a^\dagger(\mathbf{r}_n)a^\dagger(\mathbf{r}_1)\cdots a^\dagger(\mathbf{r}_{n-1})a^\dagger(\mathbf{r}_{n+1})\cdots a^\dagger(\mathbf{r}_N)|0\rangle \\
&= \sum_{n,m\neq n}^{N}\delta(\mathbf{r}-\mathbf{r}_n)\delta(\mathbf{r}'-\mathbf{r}_m)|\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_N\rangle.
\end{aligned}$$

Multiplying by $V(\mathbf{r},\mathbf{r}')/2$, and integrating over $\mathbf{r}$ and $\mathbf{r}'$, one confirms the validity of the expression. It is left as an exercise to confirm that the naive expression $\frac{1}{2}\int d^dr \int d^dr' V(\mathbf{r},\mathbf{r}')\hat{\rho}(\mathbf{r})\hat{\rho}(\mathbf{r}')$ does not reproduce the two-body operator. More generally, turning to a non-diagonal basis, it is straightforward to confirm that a general two-body operator can be expressed in the form

$$\boxed{\hat{\mathcal{O}}_2 = \sum_{\lambda\lambda'\mu\mu'} \mathcal{O}_{\mu,\mu',\lambda,\lambda'} a^\dagger_\mu a^\dagger_{\mu'} a_\lambda a_{\lambda'}}, \tag{2.20}$$

where $\mathcal{O}_{\mu,\mu',\lambda,\lambda'} \equiv \langle \mu,\mu'|\hat{\mathcal{O}}_2|\lambda,\lambda'\rangle$.

As well as the pairwise **Coulomb interaction** formulated above, another important interaction, frequently encountered in problems of quantum magnetism, is the **spin–spin interaction**. From our discussion of the second-quantized representation of spin $\hat{\mathbf{S}}$ above, we can infer that the general spin–spin interaction can be presented in second-quantized form as

$$\hat{V} = \frac{1}{2}\int d^dr \int d^dr' \sum_{\alpha\alpha'\beta\beta'} J(\mathbf{r},\mathbf{r}')\mathbf{S}_{\alpha\beta}\cdot\mathbf{S}_{\alpha'\beta'} a^\dagger_\alpha(\mathbf{r}) a^\dagger_{\alpha'}(\mathbf{r}') a_{\beta'}(\mathbf{r}') a_\beta(\mathbf{r}),$$

where $J(\mathbf{r},\mathbf{r}')$ denotes the exchange interaction.

In principle, one may proceed in the same manner and represent general n-body interactions in terms of second-quantized operators. However, as $n > 2$ interactions appear infrequently, we refer to the literature for discussion.

▷ **Interaction representation:** second quantization unfolds its true power in problems with particle interactions (usually described by operators quartic in creation/annihilation operators.) Formally, one may split an interacting Hamiltonian as $\hat{H} = \hat{H}_0 + \hat{H}_{\rm int}$, where $\hat{H}_0$ is a one-body operator. Often, it will be convenient to approach such problems in the interaction representation, wherein operators dynamically evolve under the free Hamiltonian $\hat{H}_0$:

$$\begin{aligned} a_\mu &\to e^{i\hat{H}_0 t} a_\mu e^{-i\hat{H}_0 t}, \\ a^\dagger_\mu &\to e^{i\hat{H}_0 t} a^\dagger_\mu e^{-i\hat{H}_0 t}. \end{aligned}$$

The interaction representation assumes a particularly simple form in the eigenbasis of $\hat{H}_0$. Assume, we had diagonalized $\hat{H}_0$ as

$$\hat{H}_0 = \sum_a a^\dagger_a a_a \epsilon_a,$$

where ϵ_a are the single particle energies of the problem. The commutation relation $[a^\dagger_a a_a, a_b] = -\delta_{ab}$ then readily leads to the result (prove it!)

$$\begin{aligned} a_a &\to e^{-i\epsilon_a t} a_a, \\ a^\dagger_a &\to e^{i\epsilon_a t} a^\dagger_a. \end{aligned} \tag{2.21}$$

This completes our formal introduction to the method of second quantization. To develop fluency in the operation of the method, we will continue by addressing a few problems chosen from the realm of condensed matter and quantum optics. In doing so, we will see that second quantization can lead to considerable simplification of the analysis of many-particle systems.

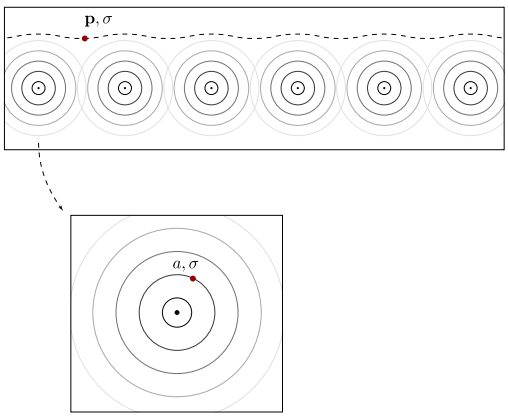
2.2 Applications of second quantization

Second quantization is *the* standard approach to tackling problems in many body physics. However it has to be kept in mind that second quantization is but a *reformulation* of many particle quantum mechanics, it doesn't solve problems by it's own. Indeed, problems in many particle physics tend to be notoriously difficult to solve and cases of exact solvability are previously few.

In view of this general situation, we aim to demonstrate how second quantization can be applied to bring a number of prototype problems of many body quantum mechanics into a form that supports intuition and is optimally adjusted to further analytical or numerical processing. Examples we will mention include plasma Hamiltonians, the Hamiltonian of electrons bound to atoms, and the quantization of the electromagnetic field.

2.2.1 Second quantized model Hamiltonians: two examples

In this short section we discuss two representatives of second quantized Hamiltonians. At this point, the examples mainly serve illustrative purposes. However, later in the text, we will apply the model Hamiltonians below in some applications of the formalism.



Fermionic plasma

A (Coulomb) **plasma** is a system of charged (and hence Coulomb-interacting) particles in a gaseous phase. By way of example, we here consider a plasma of identical and spinful fermions. We assume the fermions to be embedded into a uniform positively charged substrate, so that electro-neutrality of the total system is maintained.

Characterizing the state of each fermion in terms of a momentum quantum number $\mathbf{k}$ and a spin variable $\sigma = \uparrow, \downarrow$, the kinetics of the fermion system is described by the Hamiltonian

$$\hat{H}_0 = \sum_{\mathbf{k},\sigma} a^\dagger_\sigma(\mathbf{k}) \left(\frac{\mathbf{k}^2}{2m} - \mu \right) a_\sigma(\mathbf{k}),$$

where we have included a chemical potential in the definition of $\hat{H}_0$. To this Hamiltonian, we might now add the interaction term (2.19). However, in view of the fact that the free kinetic term assumes a diagonal form in the momentum representation, it is favorable to express the interaction in a momentum representation, too. To this end, notice that

$$V(\mathbf{r}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{q}\cdot\mathbf{x}} V(\mathbf{q}), \qquad V(\mathbf{q}) = \frac{e^2}{|\mathbf{q}|^2}$$

is the Fourier transform of the Coulomb potential $V(\mathbf{r}) = \frac{e^2}{4\pi}\frac{1}{r}$. Using this result, an application of the transform identities (2.12) readily gets us to the representation

$$\hat{V} = \frac{1}{2}\int d^3k d^3k' d^3q\, a^\dagger_\sigma(\mathbf{k}) a^\dagger_{\sigma'}(\mathbf{k}') V(\mathbf{q}) a_{\sigma'}(\mathbf{k}'+\mathbf{q}) a_\sigma(\mathbf{k}-\mathbf{q}). \tag{2.22}$$

To the plasma Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$ one would need to add an electron-substrate interaction to ensure the overall neutrality of the system (think how you would describe this contribution in second quantization.) However, this term does significantly affect the dynamics of the system (why?) and is usually kept implicit in the notation.

Multi electron atoms

Consider an atom, either in isolation or embedded into the crystalline matrix of a solid. Assume you have managed to compute the electronic eigenstates $|\psi_{a,\sigma}\rangle$ describing the shell structure of the system, where $a = (n, l, m)$ comprises principal and angular momentum quantum numbers, and σ is spin as above. The Hamiltonian describing the system of *noninteracting* shell electrons then reads

$$\hat{H}_0 = \sum_{a,\sigma} a^\dagger_{a,\sigma}(\epsilon_a - \mu) a_{a,\sigma}, \tag{2.23}$$

where we assumed spin-degeneracy of the atomic energy levels ϵ_a. To this Hamiltonian, one might now add an interaction contribution (2.20), where $\mu = (a, \sigma)$ would comprise orbital and spin quantum numbers, and $\mathcal{O}_{\mu\mu',\lambda\lambda'}$ would be the matrix elements of the electron-electron interaction in the basis of the orbital wave functions. In general, the diagonalization of the interacting generalization of the atom Hamiltonian represents a formidable task, which requires the application of numerical methods. The realistic modeling of an atom would also require inclusion of a spin-orbit fine structure interaction, and of hyperfine interactions. However, the discussion of these complications are beyond the scope of the present text. (Readers familiar with the concept of fine structure interactions may want to think about the modeling of these interactions in the language of second quantization.)

INFO The Hamiltonians discussed above represent limiting cases of a more general **solid state Hamiltonian**. To understand this, consider the one-dimensional cartoon of a solid shown in the figure. The solid is formed by a crystalline array of atoms, each comprising electrons subject to the attractive potential of a positively charged ion core. The behavior of individual electrons crucially

depends on whether we are dealing with a 'core' electron buried in the inner shells of the atom's central potential, or a 'nearly free' outer shell electron.
Elements of the groups I-IV of the periodic table contain s or p wave electrons outside a closed noble gas shell. Due to the relatively large separation from the nucleus, and the screening of the nuclear potential by core electrons, the outer electrons experience only a moderate core potential – they behave as 'nearly free' itinerant electrons. In such systems, the residual interaction of the electrons with lattice potential can be described by representing the former in a basis of so-called Bloch functions. However, in many cases of interest, Bloch functions effectively behave similar to momentum eigenstates and the plasma Hamiltonian above makes for a good description. Conversely, many aspects of the physical properties of inner shell solid state electron (bands) are captured by atomic model Hamiltonians, where corrections due to the interaction between adjacent atoms are described by so-called tight-binding Hamiltonians. The discussion of the ensuing physical phenomena is the subject of solid state theory, where the methods of second quantization are pervasive.

2.2.2 Quantization of the electromagnetic field

The goal of this section is to 'quantize' the electromagnetic field. We will see that a canonical quantization procedure applied to the Hamiltonian of the electromagnetic field gets us to a representation of the field in terms of a superposition of quantum harmonic oscillators. The quanta of these oscillators – cf. the discussion of section 2 – will then be identified as photons, i.e. the quantum constituents of the electromagnetic field.

Rather than addressing the photon field in full generality, we will here consider photons propagating in a wave guide, a simpler setup that, nonetheless, enjoys great importance in applications.

Lagrangian of an electromagnetic waveguide

To quantize the electromagnetic field, we need to identify its 'coordinates' and 'momenta', i.e. canonical classical variables for which quantum commutation relations can ge identified. While the structure of the Hamiltonian action of the field – i.e. a functional comprising coordinates and momenta – may not be entirely obvious, we may use its ally, the Lagrangian action as a starting point.

In the absence of currents, the **Lagrangian action of the electromagnetic field** is given by

$$S[A] = \frac{1}{4}\int d^4x\, F_{\mu\nu}F^{\mu\nu},$$

where $F = \{F_{\mu\nu}\}$ with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor, $A = \{A_\mu\} = (A_0, \mathbf{A})$ the electromagnetic potential, and an Einstein summation convention is implied: $f_\mu g^\mu \equiv f_0 g_0 - \sum_{i=1}^3 f_i g_i$ where $\mu = 0$ and $\mu = i = 1, 2, 3$ are time and space like components, resp. It would be tempting to think of $A_\mu(\mathbf{x}, t)$ as the 'coordinates' and $\partial_0 A_\mu(\mathbf{x}, t)$ as generalized

'velocities' of the Lagrangian action $S[A]$.[8] However, this view is premature in that it ignores the presence of gauge constraints to be imposed on the potential A.

Presently, it will be convenient to employ the Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$, $A_0 = 0$. (Within the framework of Lagrangian mechanics, $\nabla \cdot \mathbf{A}(\mathbf{x}, t) = 0$ has the status of a holonomic constraint imposed on the variables $\mathbf{A} = \{A_i\}$.) It is straightforward to verify that in this representation the action assumes the form

$$S[A] = \int dt\, L[\mathbf{A}, \partial_t \mathbf{A}],$$
$$L[\mathbf{A}, \partial_t \mathbf{A}] = \frac{1}{2} \int d^3x \, \left[(\partial_t \mathbf{A})^2 + \mathbf{A} \cdot \Delta \mathbf{A} \right]. \tag{2.24}$$

Progress with this expression can be made, if we expand vector valued functions $\mathbf{A}$ in terms of eigenfunctions of the Laplace operator, Δ, i.e. we consider an ansatz

$$\mathbf{A}(\mathbf{x}, t) = \sum_{\mathbf{k}} \alpha_{\mathbf{k}}(t) R_{\mathbf{k}}(\mathbf{x}), \tag{2.25}$$

where

$$-\Delta \mathbf{R}_{\mathbf{k}}(\mathbf{x}) = \lambda_{\mathbf{k}} \mathbf{R}_{\mathbf{k}}(\mathbf{x}), \qquad \nabla \cdot \mathbf{R}_{\mathbf{k}}(\mathbf{x}) = 0. \tag{2.26}$$

Here, $\lambda_{\mathbf{k}} \in \mathbb{R}$ are the eigenvalues and $\mathbf{k}$ is an index. The second equation in the equation above satisfies the gauge condition $\nabla \cdot \mathbf{A} = 0$. Thanks to the completeness of the eigenfunctions of the Laplace operator, we may then adopt the time dependent coefficients $\alpha_{\mathbf{k}}$ as the new coordinates of the problem.[9]

For general geometries, the solution of Eq. (2.26) is made cumbersome by the presence of the gauge constraint. We may circumvent these difficulties by considering cases where the geometry of the system reduces the complexity of the eigenvalue problem. This restriction is less artificial than it might appear. For example, in anisotropic electromagnetic waveguides, the solutions of the eigenvalue equation can be formulated as[10] where $\mathbf{k} = k\mathbf{e}, k \in \mathbb{R}$ becomes

[8] Beware of a possible source of confusion: within the framework of Lagrangian mechanics, we are accustomed to using the symbol '$\mathbf{x}$' for the coordinates (of a point particle.) However, the action $S[A]$ defines a system of *continuum* Lagrangian mechanics, and here '$\mathbf{x}$' plays an altogether different role: the coordinates of the problem are $A_\mu(\mathbf{x})$ and $\mathbf{x}$ assumes the role of an index. (Much like the index 'i' of a multi-dimensional generalized coordinate $\mathbf{q} = \{q_i\}$ Lagrangian mechanics.) There are 'infinitely many' different values of $\mathbf{x}$, i.e. we are dealing with an infinite dimensional generalization of finite dimensional Lagrangian mechanics. It is always possible to stir back towards the more familiar territory of the finite dimensional framework by introducing a lattice discretization $\mathbf{x} \to \mathbf{x}_i$, $i \in$ (some finite index space), whence $A_\mu(\mathbf{x}) \to A_\mu(\mathbf{x}_i) \equiv A_{\mu,i}$ becomes a finite dimensional object.

[9] The rational behind this reasoning: if you have a problem where functions $f(x)$, $x \in U \subset \mathbb{R}^n$ play the role of coordinates (again, you may find it helpful to think of $f(x)$ in terms of the continuum limit of a coordinate 'vector' $f(x_i) \equiv f_i$ obtained by discretization of U), and $\{g_k\}$, $k \in X$, where X is some index space, is a *complete* set of functions in U, you may expand $f = \sum_{k \in X} \alpha_k g_k$. Since any function can be represented in this way, the set $\{\alpha_k | k \in X\}$ carries the same information as the set $\{f(x) | x \in U\}$. In practice, α_k may be Fourier transform coefficients, the coefficients of angular momentum eigenfunctions, or else.

[10] More precisely, one should say that Eq. (2.26) defines the set of eigenfunctions relevant for the low-energy dynamics of the waveguide. More-complex eigenfunctions of the Laplace operator exist but they carry much higher energy.

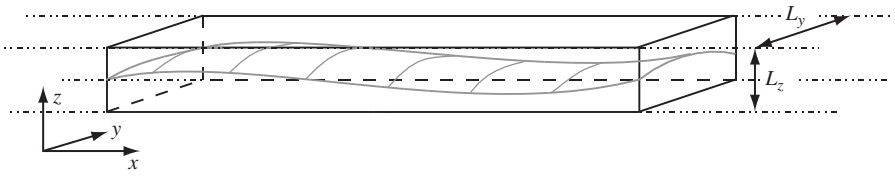


Figure 2.2: EM waveguide with rectangular cross-section. The structure of the eigenmodes of the EM field is determined by boundary conditions at the walls of the cavity.

a *one-dimensional* index and $\mathbf{e}$ is a unit vector along the extended axis of the system. The vector-valued functions $\mathbf{R}_k$ can be chosen real and orthonormalized, $\int \mathbf{R}_k \cdot \mathbf{R}_{k'} = \delta_{kk'}$. The dependence of the eigenvalues λ_k on k depends on details of the geometry (see Eq. (2.28) below) and need not be specified for the moment.

INFO An **electromagnetic waveguide** is a quasi-one-dimensional cavity with metallic boundaries (see Fig. 2.2). The practical advantage of waveguides is that they are good at confining EM waves. At large frequencies, where the wavelengths are of order meters or less, radiation loss in conventional conductors is high. In these frequency domains, hollow conductors provide the only practical way of transmitting radiation.
EM field propagation inside a waveguide is constrained by boundary conditions. Assuming the walls of the system to be perfectly conducting,

$$\mathbf{E}_{\parallel}(\mathbf{x}_{\mathrm{b}}) = 0, \qquad \mathbf{B}_{\perp}(\mathbf{x}_{\mathrm{b}}) = 0, \tag{2.27}$$

where $\mathbf{x}_{\mathrm{b}}$ is a point at the system boundary and $\mathbf{E}_{\parallel}$ ($\mathbf{B}_{\perp}$) is the parallel (perpendicular) component of the electric (magnetic) field.
For concreteness, returning to the problem of field quantization, let us consider a cavity with uniform rectangular cross-section $L_y \times L_z$. To conveniently represent the Lagrangian of the system, we wish to express the vector potential in terms of eigenfunctions $\mathbf{R}_{\mathbf{k}}$ that are consistent with the boundary conditions (2.27). A complete set of functions fulfilling this condition is given by

$$\mathbf{R}_{\mathbf{k}} = \mathcal{N}_k \begin{pmatrix} c_1 \cos(k_x x)\sin(k_y y)\sin(k_z z) \\ c_2 \sin(k_x x)\cos(k_y y)\sin(k_z z) \\ c_3 \sin(k_x x)\sin(k_y y)\cos(k_z z) \end{pmatrix}.$$

Here, $k_i = n_i \pi / L_i$, $n_i \in \mathbb{N}$, $i = x, y, z$, $\mathcal{N}_k$ is a factor normalizing $\mathbf{R}_{\mathbf{k}}$ to unit modulus, and the coefficients c_i are subject to the condition $c_1 k_x + c_2 k_y + c_3 k_z = 0$. Indeed, it is straightforward to verify that a general superposition of the type $\mathbf{A}(\mathbf{x}, t) \equiv \sum_{\mathbf{k}} \alpha_{\mathbf{k}}(t) \mathbf{R}_{\mathbf{k}}(\mathbf{x})$, $\alpha_{\mathbf{k}}(t) \in \mathbb{R}$, is divergenceless, and generates an EM field compatible with (2.27). Substitution of $\mathbf{R}_{\mathbf{k}}$ into Eq. (2.26) identifies the eigenvalues as $\lambda_{\mathbf{k}} = k_x^2 + k_y^2 + k_z^2$. In the physics and electronic engineering literature, eigenfunctions of the Laplace operator in a quasi-one-dimensional geometry are commonly described as **modes**. As we will see shortly, the energy of a mode (i.e. the Hamiltonian evaluated on a specific mode configuration) grows with $|\lambda_{\mathbf{k}}|$. In cases where one is interested in the low-energy dynamics of the EM field, only configurations with small $|\lambda_{\mathbf{k}}|$ are relevant. For example, let us consider a massively anisotropic waveguide with $L_z < L_y \ll L_x$. In this case the modes with smallest $|\lambda_k|$ are those with $k_z = 0$, $k_y = \pi / L_y$, and $k_x \equiv k \ll L_{z,y}^{-1}$. (Why is it not possible to set both k_y and k_z to zero?) With this choice,

$$\lambda_k = k^2 + \left(\frac{\pi}{L_y}\right)^2, \tag{2.28}$$

and a scalar index k suffices to label both eigenvalues and eigenfunctions $\mathbf{R}_k$. A caricature of the spatial structure of the functions $\mathbf{R}_k$ is shown in Fig. 2.2. The dynamical properties of these configurations will be discussed in the text.

We may now proceed to expand the vector potential in terms of eigenfunctions $\mathbf{R}_k$ as $\mathbf{A}(\mathbf{x},t) = \sum_k \alpha_k(t)\mathbf{R}_k(\mathbf{x})$, where the sum runs over all allowed values of the index parameter k. (In a waveguide, $k = \pi n/L$, $n \in \mathbb{N}$, where L is the length of the guide.) Substituting this expansion into Eq. (2.24) and using the normalization properties of $\mathbf{R}_k$, we obtain,

$$L(\alpha, \partial_t \alpha) = \frac{1}{2}\sum_k \left(\dot{\alpha}_k^2 - \lambda_k \alpha_k^2\right)$$

i.e. a decoupled representation where the system is described in terms of independent dynamical systems with coordinates α_k.

Quantization

From this point on, quantization proceeds along the lines of the standard algorithm. Firstly, define momenta through the relation $\pi_k = \partial_{\dot{\alpha}_k} L = \dot{\alpha}_k$. This produces the Hamiltonian $H(\alpha,\pi) = \frac{1}{2}\sum_k (\pi_k\pi_k + \lambda_k\alpha_k\alpha_k)$. Notice that H is a sum over independent harmonic oscillators, one for each value of k. Next quantize the theory by promoting fields to operators $\alpha_k \to \hat{\alpha}_k$ and $\pi_k \to \hat{\pi}_k$, and declaring $[\hat{\pi}_k, \hat{\alpha}_{k'}] = -i\delta_{kk'}$. The quantum harmonic oscillator Hamiltonian then reads

$$\hat{H} = \frac{1}{2}\sum_k \left(\hat{\pi}_k\hat{\pi}_k + \omega_k^2\hat{\alpha}_k\hat{\alpha}_k\right),$$

where $\omega_k^2 = \lambda_k$. Following the same logic as marshaled in Section 2, we then define ladder operators $b_k \equiv \sqrt{\frac{\omega_k}{2}}(\hat{\alpha}_k + \frac{i}{\omega_k}\hat{\pi}_k)$, $b_k^\dagger \equiv \sqrt{\frac{\omega_k}{2}}(\hat{\alpha}_k - \frac{i}{\omega_k}\hat{\pi}_k)$, whereupon the Hamiltonian assumes the now familiar form

$$\hat{H} = \sum_k \omega_k \left(b_k^\dagger b_k + \frac{1}{2}\right). \tag{2.29}$$

For the specific problem of the first excited mode in a waveguide of width L_y, $\omega_k = [k^2 + (\pi/L_y)^2]^{1/2}$. Equation (2.29) represents our final result for the quantum Hamiltonian of the EM waveguide. Before concluding this section let us make a few comments on the structure of the result:

- ▷ That we obtained a simple *one-dimensional* superposition of oscillators is due to the boundary conditions specific to a narrow waveguide. For less restrictive geometries, e.g. free space, a more complex superposition of vectorial degrees of freedom in three-dimensional space would have been obtained. However, the principal mapping of the free EM field onto a superposition of oscillators is independent of geometry.

- ▷ Physically, the quantum excitations described by Eq. (2.29) are, of course, the **photons** of the EM field. The unfamiliar appearance of the dispersion relation ω_k is again a peculiarity of the waveguide. However, in the limit of large longitudinal wave numbers $k \gg L_y^{-1}$, the dispersion approaches $\omega_k \sim |k|$, i.e. the relativistic dispersion of the photon field.

▷ The oscillators described by Eq. (2.29) exhibit zero-point fluctuations. It is a fascinating aspect of quantum electrodynamics that these oscillations, caused by quantization of *the* most relativistic field, surface at various points of non-relativistic physics. Examples include the Casimir effects, or the physics of van der Waals forces. For the discussion of these phenomena, we refer to the literature.

2.2.3 Atom-field Hamiltonian

The examples discussed thus far, all have in common that they were described by theories *quadratic* in the field operators. Such theories are termed **free theories**. They are free, in the sense that their solution is reducible to the solution of a single particle Schrödinger equation. The situation gets infinitely more complicated, and physically rewarding, the moment terms of higher order (third and beyond), in field operators appear. Theories of this type are called **interacting theories**, or nonlinear theories. There are only preciously few examples of exactly solvable interacting theories, in most cases one has to resort to approximation strategies. Primarily to illustrate the efficiency of second quantized formulations, we here discuss a simple case of two nonlinearly coupled free systems.

In first quantized language, the Hamiltonian describing the interaction between a single electron and an electromagnetic field is given by

$$\hat{H} = \frac{1}{2m}(\hat{\mathbf{p}} - \hat{\mathbf{A}})^2 + \hat{V} + \hat{H}_{\rm F},$$

where we have set the electron charge to unity, $\hat{V}$ represents a static potential, $\hat{H}_{\rm F}$ is the Hamiltonian of the electromagnetic field, and we assume a Coulomb gauge, $A = (0, \mathbf{A})$ for the dynamic parts of the Coulomb field. Hamiltonians of this structure can describe, e.g., the interaction of atomic shell electrons with an external field ($\hat{V}$=core potential), or the interaction of electrons inside a cavity with a field confined to the cavity ($\hat{V}$=boundary potential). The corresponding many electron system is obtained by upgrading the electronic Hamiltonian to the second quantized operator (2.17), i.e.

$$\hat{H} = \int d^3r\, a^\dagger(\mathbf{r}) \left(\frac{1}{2m}(\hat{\mathbf{p}} - \hat{\mathbf{A}})^2 + \hat{V} \right) a(\mathbf{r}) + \hat{H}_{\rm F},$$

where we consider spinless electrons for simplicity. We now split the Hamiltonian into three contributions

$$\hat{H} \simeq \hat{H}_{\rm e} + \hat{H}_{\rm I} + \hat{H}_{\rm F}, \tag{2.30}$$

where

$$\hat{H}_{\rm e} = \int d^3r\, a^\dagger(\mathbf{r}) \left(\frac{1}{2m}\hat{\mathbf{p}}^2 + \hat{V} \right) a(\mathbf{r}),$$
$$\hat{H}_{\rm I} = -\frac{1}{2m} \int d^3r\, a^\dagger(\mathbf{r}) (\hat{\mathbf{A}} \cdot \hat{\mathbf{p}} + \hat{\mathbf{p}} \cdot \mathbf{A})\, a(\mathbf{r}),$$

and we ignored the $\mathbf{A}^2$-contribution to the field-electron coupling Hamiltonian $\hat{H}_{\rm I}$.

Let us now consider the case where $H_{\rm e}$ models an atomic Hamiltonian. In this case, it will be appropriate to expand in a basis of atomic eigenfunctions (cf. Eq. (2.23)),

$$\hat{H}_{\rm e} = \sum_a a_a^\dagger(\epsilon_a - \mu)a_a, \tag{2.31}$$

where, again, we ignore spin. According to the general transformation rule (2.10),

$$a_a = \int d^3r\, \overline{\psi_a(\mathbf{r})}a(\mathbf{r}), \qquad a(\mathbf{r}) = \sum_a \psi_a(\mathbf{r})a_a,$$

where ψ_a is the ath atomic eigenfunction. Similarly, we expand the vector potential in a basis of vector-valued eigenfunctions of the Laplacian (cf. Eq. (2.25)),

$$\begin{aligned}\hat{A}(\mathbf{x}) &= \sum_{\mathbf{k}} \hat{\alpha}_{\mathbf{k}} R_{\mathbf{k}}(\mathbf{x}) = \\ &= \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}}(b_{\mathbf{k}} + b_{\mathbf{k}}^\dagger)R_{\mathbf{k}}(\mathbf{x}),\end{aligned}$$

where in the second line, we denoted the oscillator operators of the electromagnetic field (cf. Eq. (2.29)) by $b_{\mathbf{k}}$. In this representation

$$\hat{H}_{\rm F} = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + \frac{1}{2} \right), \tag{2.32}$$

becomes a sum over oscillator modes.

We next substitute the operator representations of $(a_a, b_{\mathbf{k}})$ into the interaction Hamiltonian to obtain

$$\begin{aligned}H_{\rm I} &= \int d^3r \sum_{a,b,\mathbf{k}} \frac{(-i)}{2m\sqrt{2\omega_{\mathbf{k}}}} \overline{\psi_a(\mathbf{r})}(\partial_{\mathbf{r}} R_{\mathbf{k}}(\mathbf{r}) + R_{\mathbf{k}}(\mathbf{r})\partial_{\mathbf{r}})\psi_b(\mathbf{r}))\, a_a^\dagger a_b (b_{\mathbf{k}} + b_{\mathbf{k}}^\dagger) \simeq \\ &\simeq \sum_{a,b,\mathbf{k}} \frac{R_{\mathbf{k}}(0)}{m\sqrt{2\omega_{\mathbf{k}}}}\, a_a^\dagger a_b (b_{\mathbf{k}} + b_{\mathbf{k}}^\dagger) \int d^3r\, \overline{\psi_a(\mathbf{r})}(-i\partial_{\mathbf{r}})\psi_b(\mathbf{r}) = \\ &= \sum_{a,b,\mathbf{k}} \frac{R_{\mathbf{k}}(0)}{m\sqrt{2\omega_{\mathbf{k}}}}\, a_a^\dagger a_b (b_{\mathbf{k}} + b_{\mathbf{k}}^\dagger) \langle a|\hat{p}|b\rangle = \\ &= \sum_{a,b,\mathbf{k}} g_{a,b,\mathbf{k}}\, a_a^\dagger a_b (b_{\mathbf{k}} + b_{\mathbf{k}}^\dagger),\end{aligned} \tag{2.33}$$

where

$$g_{a,b,\mathbf{k}} \equiv -i\frac{(\epsilon_b - \epsilon_a)}{\sqrt{2\omega_{\mathbf{k}}}} R_{\mathbf{k}}(0)\langle a|\hat{\mathbf{x}}|b\rangle,$$

is the so-called **dipole matrix element**. In the crucial second equality above, we used that for radiation energies resonant with intra-atomic transitions, electromagnetic wavelengths

$\lambda \sim 10^{-7}$m are about three orders of magnitude larger than the characteristic extension of atoms $\sim 10^{-10}$m. Thus, $R_{\mathbf{k}}(\mathbf{r})$ is approximately constant across the atom and can be replaced by the amplitude $R_{\mathbf{k}}(0)$ at the atom's center. In the last equality we used $\langle a|\mathbf{p}|b\rangle = im\langle a|[\hat{H}, \hat{\mathbf{x}}]|b\rangle = im(\epsilon_a - \epsilon_b)\langle a|\hat{\mathbf{x}}|b\rangle$.

2.2.4 Rabi oscillations

The Hamiltonian defined by Eqs. (2.30), (2.31) (2.32), and (2.33) provides the formal basis for the description of many light/matter interaction phenomena. For example, the quantum theory of the laser can be formulated in this description. In this section, we will explore the physics of light/matter interaction on a simple phenomenon: when an atom is excited by a coherent light field, it will cyclically absorb and emit photons. The ensuing oscillatory emission profiles are called Rabi oscillations. They play a fundamental role in many phenomena of quantum optics.

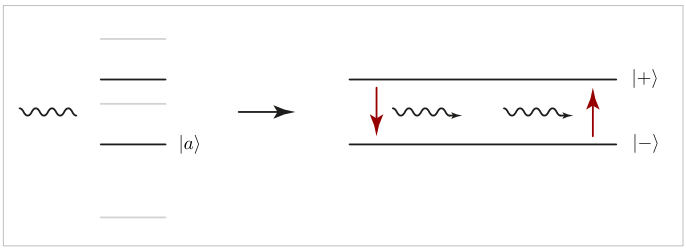


Figure 2.3: On the definition of the James-Cummings Hamiltonian. Discussion, see text.

James-Cummings Hamiltonian

To begin with, let us transform the light/matter Hamiltonian to an interaction representation. Application of (2.21) readily leads to

$$H_{\rm I} \to \sum_{a,b,\mathbf{k}} g_{a,b,\mathbf{k}}\, a_a^\dagger a_b\, e^{i\epsilon_{ab}t}(b_{\mathbf{k}} e^{-i\omega_{\mathbf{k}}t} + b_{\mathbf{k}}^\dagger e^{i\omega_{\mathbf{k}}t}),$$

where $\epsilon_{ab} = \epsilon_a - \epsilon_b$. We expect the light-matter interaction to be strongest, if the energy of the photon modes $\omega_{\mathbf{k}} \simeq \epsilon_{ab}$ is in resonance with inter-atomic energy differences. For transitions failing this condition, the Hamiltonian $H_{\rm I}$ is rapidly oscillatory in time, which is a signature of classically forbidden transitions. Let us, then, assume that the system of atomic energy levels contains two states $|a\rangle$ and $|b\rangle$ whose energy difference $\epsilon \equiv \epsilon_{ab}$ is in resonance with an incoming photon mode, cf. Fig. 2.3. For notational simplicity, we denote the upper/lower of the resonant levels by $|+\rangle/|-\rangle$, and we drop the index $\mathbf{k}$ of the resonant mode, cf. Fig. 2.3, right. Neglecting all other states, and keeping only contributions where the energy consuming transition $|-\rangle \to |+\rangle$ absorbs a photon (b), while the energy releasing transition $|+\rangle \to |-\rangle$ emits a photon ($b^\dagger$) (what is the formal justification of this, so-called

rotating wave approximation?), we arrive at the simplified Hamiltonian

$$\hat{H} = \frac{\epsilon}{2}\sum_{s=\pm} s a_s^\dagger a_s + \omega b^\dagger b + \sum_s \left(g e^{i(\epsilon-\omega)t}\, a_+^\dagger a_- b + g^* e^{-i(\epsilon-\omega)t}\, a_-^\dagger a_+ b^\dagger \right),$$

where, we have defined the chemical potential $\mu \equiv (\epsilon_a + \epsilon_b)/2$ to lie in the middle of the two resonant levels. We defined $g \equiv g_{a,b,\mathbf{k}}$ and $g^* = g_{b,a,\mathbf{k}}$, and we observed that the intra-state dipole matrix elements $g_{a,a,\mathbf{k}} = g_{b,b,\mathbf{k}} = 0$ (why?).

At this point, we may observe that the fermionic Fock space of the reduced problem decomposes into a trivial sector $\mathcal{F}^0 \oplus \mathcal{F}^2$ where the two levels are both empty, or both occupied. In this sector, no transitions are possible, either because there aren't any particles ($\mathcal{F}^0$), or because of Pauli blocking ($\mathcal{F}^2$).[11] The non-trivial space is $\mathcal{F}^1$ with its two single occupied states $|-\rangle$ and $|+\rangle$. Rather than using the heavy artillery of second quantized operators, this ordinary two-dimensional Hilbert space is preferably represented in the language of single particle quantum mechanics. Specifically, we introduce a 'pseudo-spin' representation

$$|+\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Within this representation, the second quantized operators spanning the Hamiltonian correspond to

$$\begin{aligned} a_+^\dagger a_+ - a_-^\dagger a_- &\leftrightarrow \sigma_3, \\ a_+^\dagger a_- &\leftrightarrow \sigma_+, \\ a_-^\dagger a_+ &\leftrightarrow \sigma_-, \end{aligned}$$

where σ_i are the Pauli matrices defined in (2.15), $\sigma_+ = \frac{1}{2}(\sigma_1 + i\sigma_2) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, and $\sigma_- = \sigma_+^T$. Assuming perfect resonance conditions, $\omega = \epsilon$, the effective single-mode/two-level Hamiltonian

$$\boxed{\hat{H} = \frac{\omega}{2}\sigma_3 + \omega\omega b^\dagger b + g\sigma_+ b + g\sigma_- b^\dagger} \tag{2.34}$$

is known as the **James-Cummings Hamiltonian**. Here, we assumed the coupling constant g to be real.

EXERCISE Show that for complex g a gauge transformation can be employed to turn g into a real constant.

Rabi oscillations

The Hamiltonian (2.34) acts in the product Hilbert space $\mathcal{H}_2 \otimes \mathcal{F}$ of the two-state space $\mathcal{H}_2$ and the Fock space of the quantized mode, $\mathcal{F}$. However, a moments thought shows

[11] Remember that we ignore the electron spin, i.e. our discussion applies to electrons of a fixed spin polarization.

that, for any $n \in \mathbb{N}$, the Hamiltonian couples only the two states $|+,n\rangle \equiv |+\rangle \otimes |n\rangle$ and $|-,n+1\rangle \equiv |-\rangle \otimes |n+1\rangle$.

Suppose, we had prepared an atomic excited state $|+,n\rangle$. We may then ask for the probability that the system remains in the upper state, $P_n(t) \equiv |\langle +,n| \exp(-i\hat{H}t)|+,n\rangle|^2$. Splitting the Hamiltonian as $\hat{H} = \hat{H}_0 + \hat{H}_\mathrm{I}$, where $\hat{H}_0 = \frac{\omega}{2}\sigma_3 + \omega b^\dagger b$, we note that $\hat{H}_0 = (n+\frac{1}{2})\omega\mathbb{I}$ acts trivially in our Hilbert space; it merely generates a phase which does not affect the probability, i.e. $P_n(t) \equiv |\langle +,n| \exp(-i\hat{H}_\mathrm{I}t)|+,n\rangle|^2$. From (2.5), we infer that $\hat{H}_\mathrm{I}|+,n\rangle = g\sqrt{n}|-,n+1\rangle$, i.e. the coupling Hamiltonian affords the matrix representation $\hat{H}_\mathrm{I} = g\sqrt{n}\sigma_1$. The identity $\exp(ic\sigma_1) = \cos(c)\mathbb{I} + i\sin(c)\sigma_1$ then readily gets us to the result

$$P_n(t) = \cos^2(g\sqrt{n}t).$$

These are the celebrated **Rabi oscillations**. Contrary to naive expectation, an atom prepared in an excited state does not just 'decay' into the lower state. Rather, we are met with an oscillatory pattern of emission and absorption into the photon field.

Coherent states

Now, one may object that the above result for P_n is academic in that it presumes a precisely determined number of photons, n. But what would be a 'non-academic' choice of a photon field? The platonic ideal of a perfectly coherent light source would be of precisely defined intensity and phase. However, much as with position and momentum of a quantum particle, intensity and phase of a quantum photon field cannot be simultaneously fixed. To understand this, notice that the **number operator** $\hat{n} \equiv b^\dagger b$ is the quantum operator determining the photon number ($\leftrightarrow$ intensity) of our light mode. We now define a **phase operator**, $\hat{\phi}$ through

$$b = \hat{n}^{1/2}\exp(i\hat{\phi}),$$
$$b^\dagger = \exp(-i\hat{\phi})\hat{n}^{1/2}.$$

EXERCISE Show that the canonical commutation relation $[b, b^\dagger] = 1$ implies $[\hat{n}, \hat{\phi}] = i$.

(Up to a factor of i), the transformation $(b, b^\dagger) \to (\hat{n}, \hat{\phi})$ is canonical, i.e. number and phase operator fully characterize the electromagnetic field.

EXERCISE Trace back the definition of the field constituents $(\hat{n}, \hat{\rho}) \to (b, b^\dagger) \to \mathbf{A} \to (\mathbf{E}, \mathbf{B})$ to establish a connection between the phase operator and the electric and magnetic field. Explore how the phase ϕ determines the phase and polarization of the electromagnetic wave described by the vector potential. (In doing so, you may treat (n, ϕ) as classical variables.)

Much like a gaussian wave package minimized the relative quantum uncertainty of coordinate and momentum of a point particle, we may define quantum states that minimize the relative uncertainty of intensity and phase of a quantum light field. These states, the so-called **coherent states** define the closest approximation to an optimally defined light field. Without proof, we state that the coherent states are given by

$$|\alpha\rangle \equiv e^{-|\alpha|^2/2}\exp(\alpha b^\dagger)|0\rangle, \tag{2.35}$$

where $\alpha \in \mathbb{C}$ is a complex number. It is a good exercise to check that the expectation value of the field intensity in the coherent state is given by $\langle\alpha|\hat{n}|\alpha\rangle = |\alpha|^2$. The computation of the phase expectation value is more tricky. However, it is not difficult to verify that $\langle\alpha|b|\alpha\rangle = \alpha$. The above definition of the phase operator then shows that $\frac{1}{2i}\ln(\alpha/\bar{\alpha}) \simeq \phi$, i.e. modulus and phase of α determine intensity and phase of the coherent state light field.

EXERCISE Prove the relations

$$\begin{aligned}
&\langle\alpha|\alpha\rangle = 1,\\
&b|\alpha\rangle = \alpha|\alpha\rangle,\\
&\langle\alpha|\hat{n}|\alpha\rangle = |\alpha|^2,\\
&\langle\alpha|(\hat{n} - \langle n\rangle)^2|\alpha\rangle = |\alpha|^2.
\end{aligned} \tag{2.36}$$

Collapse and revival

Let us now explore the probability

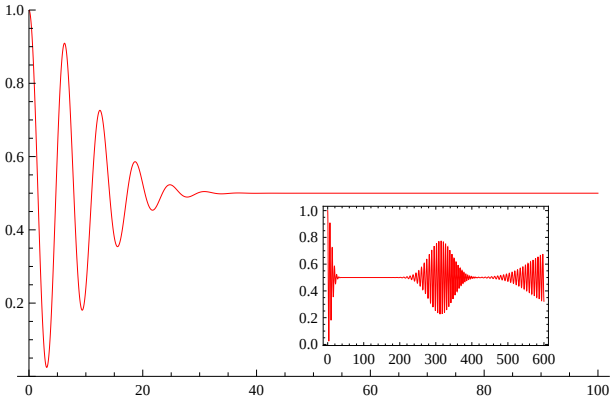
$$P_\alpha(t) = |\langle +, \alpha|e^{-i\hat{H}t}|+, \alpha\rangle|^2$$

that a state initially prepared as a product state of the excited atomic level and a coherent light field stays in the excited state $|+, \alpha\rangle \equiv |+\rangle \otimes |\alpha\rangle$. Using that

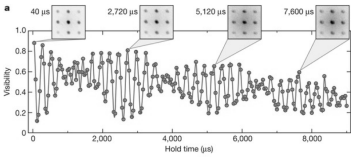
$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{n!}(b^\dagger)^n|0\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{(n!)^{1/2}}|n\rangle,$$

and the result obtained in the previous section, we obtain

$$\begin{aligned}
P_\alpha(t) &= e^{-|\alpha|^2}\sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!}P_n(t) =\\
&= e^{-|\alpha|^2}\sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!}\frac{1}{2}(1+\cos(2g\sqrt{n}t) =\\
&= \frac{1}{2}\left(1 + e^{-|\alpha|^{2n}}\sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!}\cos(2g\sqrt{n}t)\right).
\end{aligned}$$



Unlike with the fictitious case of a fixed photon number field, we observe a decay of the excited state, which is modulated by a rapidly oscillatory pattern. The reason for this is that we now superimpose oscillatory cosines of different frequency $\propto \sqrt{n}$, which leads to destructive interference. Eq. (2.36) shows that the coherent state contains $\bar{n} = |\alpha|^2$ photons on average, with fluctuations $\delta n \sim \sqrt{\langle (n-\bar{n})^2 \rangle} = |\alpha|$. We expect decay of the envelope profile for times $t 2g(\sqrt{\bar{n}+\delta n} - \sqrt{\bar{n}-\delta n}) \overset{\bar{n}\gg 1}{\sim} 2tg \sim 2\pi$ such that the 'uncertainty' in oscillation frequencies gives rise to a phase uncertainty 2π. For large photon numbers, this happens for $t \sim \pi/g$. This is illustrated in the figure for $g = 0.1$. Interestingly, however, this is not the end of the story. As is shown in the inset of the figure, the 'collapse' of the initial state is followed by its 'revival' at larger times. To understand this phenomenon, notice that for modes in the interval $[[\bar{n}-\delta n, \bar{n}+\delta n]$, the spacing between consecutive discrete frequencies is given by $\delta\omega = 2g(\sqrt{n+1} - \sqrt{n}) = \frac{g}{\sqrt{n}} \simeq \frac{g}{\sqrt{\bar{n}}}$. For times $t\Delta\omega = k2\pi \Rightarrow t = k2\pi\sqrt{\bar{n}}g$, all cosines are approximately in phase and may add to a finite fraction of the initial value. The revival phenomenon isn't perfect, because of the spread in frequency spacings which we neglected and which causes residual decoherence.



A recent observation[12] of a collapse/revival pattern in a atom-optics setting is shown in the figure. The take-home message from of this discussion is that the interaction of matter states with quantum light sources may lead to a wide spectrum of phenomena. For coherent irradiation we do not usually observe a straightforward decay of excited states. Rather, one may be met with oscillations, or oscillations superimposed by 'beads'. However, the interaction with 'incoherent' light, i.e. an uncorrelated superposition of modes of different frequency generally causes the decay of excited matter states.

2.3 Summary & Outlook

This concludes our preliminary discussion of applications of the second quantization. In this chapter, we have introduced second quantization as a tool whereby problems of many-body quantum mechanics can be addressed more efficiently than by the traditional language of symmetrized many-body wave functions. We have discussed how the two approaches are related to each other and how the standard operations of quantum mechanics can be performed by second quantized methods.

[12] S. Will, *Time-resolved observation of coherent multi-body interactions in quantum phase revivals*, Nature **465**, 197 (2010).

One may note that, beyond qualitative discussions, the list of concrete applications encountered in this chapter involved problems that either were non-interacting from the outset, or could be reduced to a solvable problem with a small number of degrees of freedom (the James-Cummings system). However, we carefully avoided dealing with interacting problems where no such reductions are possible – the majority by far of the problems encountered in many particle physics. What can be done in situations where interactions, i.e. operator contributions of fourth or higher order, are present and no tricks can be played? Obtaining answers to this question is a key challenge in many areas of modern physics — and one beyond the scope of the present text. However the machinery introduced above leaves you well equipped to delve into the quantum physics of many particle systems.

Chapter 3

Relativistic quantum mechanics

In this chapter we aim to reconcile quantum mechanics with the principles of special relativity. We start with a discussion of the Klein-Gordon equation, the first historic attempt of a relativistic generalization of the Schrödinger equation. We will run into a number of conceptual problems which motivate the search for alternative formulations, and ultimately lead to the Dirac equation. We will present two alternative derivations of the Dirac equation, a short and pragmatic one (Dirac's original derivation), and a more extended derivation which highlights its true conceptual status. Specifically, we will understand that the Dirac equation should not actually be interpreted as a relativistic generalization of the Schrödinger equation. It is far more natural to think of it as a descriptor of a many particle *field*, the field of relativistic fermions, and their anti-particles.

3.1 Synopsis of special relativity

Recapitulate the basics of special relativity, notably the postulates of relativity, the physics of Lorentz transformations, their formal description in terms of the Lorentz group, and the principles of relativistic mechanics.

3.1.1 Covariant notation

The relativistic extension of quantum mechanics to be discussed below will be formulated in the co- and contravariant notation of special relativity. We begin this chapter with a brief reminder of the (largely standard) 'relativistic notation'.

Space-time events are recorded in the **contravariant space time vector** $v = \{v^\mu\} \equiv (v^0, v^i)$. Greek indices $\mu = 0, 1, 2, 3$ generally index time ($\mu = 0$) and space ($\mu = 1, 2, 3$), while latin indices $i = 1, 2, 3$ refer to space. Important examples of contravariant vectors include the **space-time vector** $x = (ct, \mathbf{x})$, and the **four-momentum vector** $p = (E/c, \mathbf{p})$. The corresponding **covariant vector**[1] $\{x_\mu\}$ is defined by

$$x_\mu = g_{\mu\nu} x^\nu,$$

[1] An abuse of language. It would be more appropriate to call $\{x_\mu\}$ a covariant tensor of rank 1. However the general denotation 'vector' for everything that carries four indices has stuck.

where

$$g = \{g_{\mu\nu}\} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix},$$

g is the **Minkovski metric** and an Einstein summation convention ($x_{\ldots\mu}y^{\mu\cdots} \equiv \sum_{\mu=0}^{3} x_{\ldots\mu}y^{\mu\cdots}$) is implied. Similarly,

$$x^\mu = g^{\mu\nu}x_\nu,$$

where

$$g = \{g^{\mu\nu}\} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix},$$

is the inverse of $\{g_{\mu\nu}\}$. (In a Minkovski space-time basis, $\{g_{\mu\nu}\}$ is self-inverse.) Derivatives w.r.t. co- and contravariant vector indices are defined as

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = (\partial_0, \nabla),$$
$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = (\partial_0, -\nabla),$$

in terms of which the **d'Alambert operator** is defined as

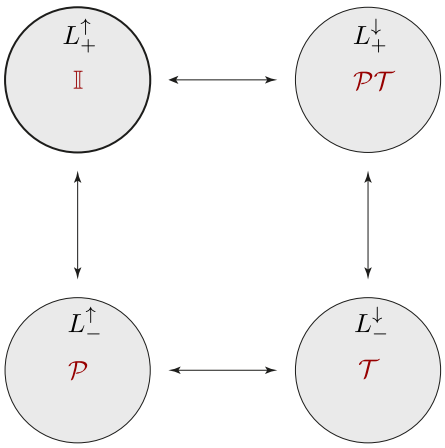
$$\Box = \partial_\mu\partial^\nu = \frac{1}{c^2}\partial_t^2 - \Delta.$$

To simplify the notation, we will set $c = 1$ throughout.

3.1.2 Lorentz group essentials

The relativistic extension of quantum mechanics below will heavily rely on the postulated Lorentz invariance of all relevant equations. For later reference, we here summarize the essential properties of the Lorentz group relevant to the construction of the theory.

The Lorentz group, L, contains all linear transformations $\Lambda : \mathbb{R}^4 \to \mathbb{R}^4, x \mapsto \Lambda x$ of Minkovski space, that leave the scalar product g invariant, $\Lambda^T g \Lambda = g$. In components, this is expressed as $x'_\mu y'^\mu = x_\mu y^\mu$ for Lorentz transformed four-vectors $x'_\mu = \Lambda_\mu{}^\nu x_\nu$ and $y^\mu = \Lambda^\mu{}_\nu y^\nu$ where $\Lambda = \{\Lambda^\mu{}_\nu\}$ and indices are raised and lowered by metric multiplication, $\Lambda^\mu{}_\nu = g^{\mu\mu'}\Lambda_{\mu'}{}^{\nu'}g_{\nu'\nu}$, etc. (Recapitulate why L forms a group and not just a set.)



In many ways, the Lorentz group is similar to the orthogonal group $\mathrm{O}(4)$, i.e. the set of all transformations $O^T\mathbb{I}_4 O = \mathbb{I}_4$ leaving the trivial metric of $\mathbb{R}^4$ invariant. However, there are important differences: first, it is non-compact, i.e. unlike with the rotation 'angles' parameterizing a rotation group, it cannot be spanned by a set of compact parameters (see below.) Second it comes with a richer topology. To understand the latter point, notice that the relation $\Lambda^T g\Lambda = g$ implies $\det(\Lambda)^2 = 1$, or $\det\Lambda = \pm 1$. Second, there is no way to continuously interpolate between a Lorentz matrix with zero-zero element $\Lambda_{00} > 0$ and one with $\Lambda_{00} < 0$. To see this, notice that the zero-zero element of the invariance condition reads $1 = g_{00} = (\Lambda^T g\Lambda)_{00} = (\Lambda_{00})^2 - \sum_i (\Lambda_{0i})^2$. This condition can only be satisfied for non-vanishing Λ_{00}, i.e. Λ_{00} cannot be tuned through zero. We thus conclude that $L = L_+^\uparrow \cup L_+^\downarrow \cup L_-^\uparrow \cup L_-^\downarrow$ decomposes into four disjoint subsets defined by

$$\begin{aligned} L_+^\uparrow &: \quad \det\Lambda = +1, \quad \Lambda_{00} \geq +1, \\ L_+^\downarrow &: \quad \det\Lambda = +1, \quad \Lambda_{00} \leq -1, \\ L_-^\uparrow &: \quad \det\Lambda = -1, \quad \Lambda_{00} > +1, \\ L_-^\downarrow &: \quad \det\Lambda = -1, \quad \Lambda_{00} \leq -1. \end{aligned} \tag{3.1}$$

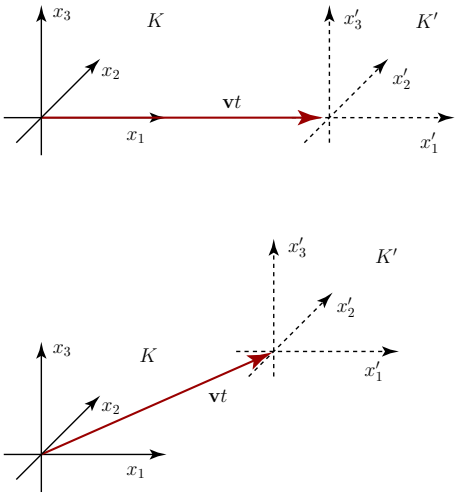
Of these, only $L_+^\uparrow$ forms a sub*group* (why?), in particular it contains the unit element $\mathbb{I}_4 \in L_+^\uparrow \subset L$. The component $L_+^\uparrow$ is called the **restricted Lorentz group**. The remaining three components accomodate Lorentz transformations which cannot be continuously deformed to unity. Specifically, the **time reversal** operation $\mathcal{T} : (t, \mathbf{x}) \mapsto (-t, \mathbf{x})$ belongs to $L_\downarrow^-$, **parity**, $\mathcal{T} : (t, \mathbf{x}) \mapsto (t, -\mathbf{x})$ belongs to $L_\uparrow^-$, and the product of the two, $\mathcal{PT}$ to $L_\downarrow^+$.

In the mathematical literature the Lorentz group is commonly denoted $\mathrm{O}(1,3)$, i.e. the group of 'orthogonal' transformations leaving a metric of signature $(+,-,-,-)$ invariant, the unit-determinant sector $L_+ = L_+^\uparrow \cup L_+^\downarrow$ is called $\mathrm{SO}(1,3)$, and the proper Lorentz group $L_+^\uparrow$ is denoted $\mathrm{SO}^+(1,3)$. Let us parameterize proper Lorentz transformations as $\mathrm{SO}^+(1,3) \ni \Lambda \equiv \exp(T)$, where $T \in \mathrm{so}^+(1,3)$ belongs to the Lie algebra $\mathrm{so}^+(1,3)$. It is straightforward to verify that the condition $\Lambda^T g\Lambda = \gamma$ requires $T^T g = -gT$.

Much like the Lie algebra, $\mathrm{so}(3)$, of the three-dimensional rotation group $\mathrm{SO}(3)$ is spanned by three linearly independent generators, $\hat{J}_{1,2,3}$, of rotations around the three coordinate axes (given explicitly in Eq. (3.11) below), the algebra $\mathrm{so}(1,3)$ has six linearly independent generators $\hat{J}_{1,2,3}$ and $\hat{K}_{1,2,3}$. These are given by

$$\hat{J}_i = \begin{pmatrix} 0 & 0 \\ 0 & \hat{J}_i \end{pmatrix}, \qquad \hat{K}_i = \begin{pmatrix} 0 & \mathbf{e}_i^T \\ \mathbf{e}_i & 0_3 \end{pmatrix}. \tag{3.2}$$

Here, $\hat{J}_i$ generate rotations of three dimensional space, a Lorentz transformation which we again represent by the symbol $\hat{J}_i$.[2] The matrices $\hat{K}_i$ contain a unit vector $\mathbf{e}_i$ in the 3×1 dimensional space-time sector and its transpose in the 1×3 dimensional time-space sector. Physically, these matrices generate **Lorentz boosts** along the i-axis, as can be seen by straightforward exponentiation. E.g.



$$\Lambda = \exp(\phi\hat{K}_1) = \begin{pmatrix} \cosh(\phi) & \sinh(\phi) & 0 & 0 \\ \sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (3.3)$$

represents the Lorentz transformation (cf. the top panel of the figure) $x' = \Lambda x$, with

$$x^{0\prime} = \gamma(x^0 + \beta x^1), \quad x^{1\prime} = \gamma(\beta x^0 + x^1), \quad x^{2\prime} = x^2, \quad x^{3\prime} = x^3,$$

where $\beta \equiv v/c$ and

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \cosh(\phi). \quad (3.4)$$

Lorentz boosts along arbitrary vectors $\mathbf{v} = \{v_i\}$ (bottom panel) are generated by appropriate linear combinations of generators, $v_i\hat{K}_i$.

It is a straightforward exercise to work out the $\mathbf{so}(3,1)$**–commutation relations**

$$\begin{aligned} [\hat{J}_i, \hat{J}_j] &= \epsilon_{ijk}\hat{J}_k, \\ [\hat{K}_i, \hat{K}_j] &= \epsilon_{ijk}\hat{J}_k, \\ [\hat{J}_i, \hat{K}_j] &= -\epsilon_{ijk}\hat{K}_k. \end{aligned} \quad (3.5)$$

Notice that the pure Lorentz transformations (generated by the $\hat{K}$'s) do not form a group. This means that sequences of non-colinear Lorentz boosts necessarily involve a *rotation* of space, an effect that is at the root of the physical phenomenon of **Thomas precession**.

Finally, notice that the **Lorentz group is non-compact**, i.e. unlike with the rotation group it cannot be spanned by a compact (closed and bounded) set of parameters. For example, the variable ϕ used to parameterize the Lorentz boost above runs over the entire real axis.

[2] I.e. the generator $\hat{J}_i$ in (3.2) is a 4×4 matrix containing the three dimensional rotation generator $\hat{J}_i$ as its 3×3 spatial block.

3.2 Klein-Gordon equation

After this preparation, we now set out to 'quantize' relativistic mechanics. To this end, let us recall the algorithm of quantizing the non-relativistic theory: i) write down a 'dispersion relation', $E-H=0$, where $H=\mathbf{p}^2/(2m)$ is kinetic energy, ii) upgrade to operators, $E\to i\partial_t$, $\mathbf{p}\to -i\partial_{\mathbf{x}}$, iii) define wave functions through $(i\partial_t+\frac{1}{2m}\Delta)\psi=0$. Let us now transfer this procedure to the relativistic setting. We start from the relativistic dispersion relation

$$E^2-\mathbf{p}^2=m^2,$$

and upgrade to operators as in ii) above. Denoting the relativistic wave function by $\phi=\phi(x)$, we are led to the equation $(\partial_t^2-\Delta+m^2)\phi=0$, or

$$\boxed{(\Box+m^2)\phi=0.} \tag{3.6}$$

This is the celebrated **Klein-Gordon equation**. The Klein-Gordon equation beautifully embodies the principle of relativistic invariance. However, on second thought, we may wonder if it really is the equation we have been looking for. First, the Klein-Gordon operator $(\Box+m^2)$ differs from the Schrödinger operator in that it carries dimension (energy)^2, and not just energy. It is, thus, not an equation trivially reducing to the Schrödinger equation in the non-relativistic limit $p^2/2m\ll m$.

INFO One might get the idea to take the **'square root' of the Klein-Gordon equation**, i.e. $E=\pm\sqrt{\mathbf{p}^2+m^2}$, or $(i\partial_t\mp\sqrt{-\Delta+m^2})\psi=0$. Then, however, we need to decide which solution to pick, the one with positive, or with negative energy. We might decide to discard the negative energy solutions 'on physical grounds'. But can we really? Second, the equation is now of infinite order in the differential operator Δ. Only in the non-relativistic limit does the positive energy branch reduce to the Schrödinger (!) equation $i\partial_t\psi\simeq(m-\frac{1}{2m}\Delta)\psi$, where $m=mc^2$ plays the role of the rest frame energy. In general, however, an infinite order equation hardly seems to be a viable option.

The above difficulties notwithstanding, the idea of 'taking the square root' was attractive enough; it formed the historic basis of Dirac's extension of the Klein-Gordon approach, to be discussed below.

Irritatingly, the E^2-problem is not the only difficulty we meet with the Klein-Gordon equation: according to the general principles of quantum mechanics, the correspondence $\hat{p}\leftrightarrow -i\nabla$ leads to the interpretation of $\mathbf{j}\equiv\frac{1}{2mi}\left(\bar{\phi}\nabla\phi-\phi\nabla\bar{\phi}\right)$ as a **current density**. As it is our top priority to stay relativistically covariant, we are then *forced* to introduce $j_0\equiv\frac{1}{2mi}\left(\bar{\phi}\partial_0\phi-\phi\partial_0\bar{\phi}\right)$ as the corresponding 0-component, so that we obtain

$$j_\mu=\frac{1}{2mi}\left(\bar{\phi}\partial_\mu\phi-\phi\partial_\mu\bar{\phi}\right), \tag{3.7}$$

as the **quantum mechanical four current**. In line with the interpretation of the spatial components j_i as currents, we think of j_0 as a component describing particle densities. Indeed,

it is easy to verify that for ϕ a solution of the Klein-Gordon equation j_μ obeys a **continuity equation**,

$$\partial^\mu j_\mu = 0, \tag{3.8}$$

much like the density $\rho = \bar{\psi}\psi$ and current $\mathbf{j} = \frac{1}{2mi}(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi})$ carried by solutions ψ of the non-relativistic Schrödinger equation obey one (recall this point.)

So far everything looks good. The problems start when we realize that unlike with the non-relativistic ρ, the relativistic 'density' j_0 is not positive definite! The Klein-Gordon equation is second order in time, which means that both ϕ and $\partial_t\phi$ can be chosen arbitrarily as initial conditions. Pick positive ϕ and negative $\partial_t\phi$ and you obtain a negative 'density' j_0. How do we interpret this phenomenon? One option follows from the observation that non-trivial solutions ϕ must be complex (for for real ϕ both density and current are trivially zero.) Now recall that in the quantum mechanics of charged particles, gauge transformations $A_\mu \to A_\mu + \partial_\mu\phi$ act on wave functions as $\psi \to \exp(i\phi)\psi$. In other words, charge requires complexity of wave functions, too. This observation suggests an interpretation of the relativistic j_μ as a charge current, as opposed to a mere particle or probability current. Within this interpretation, negative j_0, i.e. negative electric charge density, creates much less of a headache. Still we are left with a feeling of uneasiness: the charge interpretation seems vague, at this point. And even if it holds, it is irritating that the theory *forces* us to introduce negative charge states (for whatever they describe) along with positive ones.
The bottom line of the discussion above is that the relativistic extension comes at the prize of 'negative solutions' – negative in energy, or negative in density – bullying their way into the theory. We will soon understand that these solutions anticipate the physics of anti-particles, i.e. partners negative in energy and charge relative to the original particles. While these states remain hidden in the non-relativistic setting, we have to introduce them once we go relativistic.

3.3 Dirac equation

The Dirac equation is relativistic equation carrying physical dimension (energy)^1. In this respect, it may be thought of as a 'square root' of the Klein-Gordon equation. However, in contradistinction to the naive square root discussed above, it is of first (and not infinite) order in derivatives. Heuristically, this first order-ness embodies the linear dispersion characteristic for theories at the ultra-relativistic limite: the law $E \sim (cp)^2$ (Klein-Gordon) gets replaced by $E \sim (cp)$ (Dirac). But how is the construction of an effective first order equation achieved in practice? There are different routes to this end. The most direct one is Dirac's historic construction, reviewed in the majority of textbooks on the subject, and briefly sketched below for the convenience of the reader. While Dirac's approach is concise and very much to the point, it may feel a bit 'ad hoc' to some. Certainly, it doesn't give one much physical insight. The opposite extreme would be a representation-theory oriented approach, which deduces the structure of the linear evolution equation hierarchically, outgoing from the known transformation behavior of contravariant vectors under Lorentz transformations. This formulation (reviewed in many textbooks of mathematical physics) does provide insight into the structure

of the theory, but this comes at the price of an extended construction. We will here try to follow a 'middle way', which emphasizes symmetry under transformations (the most important principle of relativity, after all), but dispenses with mathematical rigor. Although this approach, too, has its shortcomings, it compromises between efficiency and physical transparency, which may be appropriate for this text.

3.3.1 Rotation invariance: $\mathrm{SU}(2)$–$\mathrm{SO}(3)$ correspondence

Unlike with the Klein-Gordon equation (or the Schrödinger equation for that matter), the formulation of Dirac equation necessitates the presence of spin. It turns out that the intimate relation between linearity of the dispersion, (relativistic) invariance, and spin finds a precursor within the more familiar context of non-relativistic quantum mechanics. We will first discuss this point to then generalize to the relativistic setting.

Consider a fictitious Schrödinger equation governed by the operator $\hat{H} = \hat{H}_0 + x_i\hat{l}_i$, where $\hat{l}_i$, $i = 1, 2, 3$ are the cartesian components of the angular momentum operator, x_i are the three components of a mathematical object whose realization we are to identify, and rotational invariance is required. Under a rotation, $\hat{l} = \{l_i\}$ transforms as an axial vector. Rotational invariance of the theory then requires that the components $x = \{x_i\}$ must transform like an axial vector, too (for then the 'bilinear form' $x_i\hat{l}_i \equiv x^T\hat{l}$ will stay invariant as required.)

A vector-object x showing the desired 'active' transformation behavior is realized in the quantum mechanics of spin: chose $x_i = \sigma_i$ as Pauli matrices, i.e. generators of $\mathrm{SU}(2)$ in a spin $1/2$ representation.[3] Our Hamilton operator now acts in a Hilbert space $\mathcal{H} \otimes \mathbb{C}^2$, where $\mathcal{H}$ is the 'orbital' Hilbert space in which angular momentum acts, and the second factor accommodates spin.

A rotation of space is mediated by a matrix $\hat{R} \in \mathrm{SO}(3)$, which acts on the angular momentum operator as $\hat{l} \mapsto \hat{R}\hat{l}$, i.e. by matrix multiplication. The Hamilton operator transforms as $\hat{H} \to \hat{H}_{\hat{R}} \equiv \hat{H}_0 + (\hat{R}\hat{l})^T\hat{J}$. The whole point now is that spinor states transform, too, viz. as $|\psi\rangle \to \hat{U}|\psi\rangle$, where $\hat{U} \in \mathrm{SU}(2)$ is a matrix to be identified momentarily. Rotational invariance requires expectation values (i.e. the detectable output of the theory) to be independent of the frame of reference. We thus demand that

$$\langle\psi|\hat{H}|\psi\rangle \stackrel{!}{=} \langle(\hat{U}\psi)|\hat{H}_{\hat{R}}|\hat{U}\psi\rangle = \langle\psi|(\hat{U}^\dagger\hat{H}_{\hat{R}}\hat{U})|\psi\rangle. \tag{3.9}$$

Since this is to hold for any $|\psi\rangle$, we have the condition $U^\dagger\hat{H}_{\hat{R}}\hat{U} = \hat{H}$. Substituting the definition of $\hat{H}_{\hat{R}}$, we obtain the condition

$$\hat{U}^\dagger((\hat{R}\hat{l})^T\sigma)\hat{U} = (\hat{R}\hat{l})^T(\hat{U}^\dagger\sigma\hat{U}) = \hat{l}^T\hat{R}(\hat{U}^\dagger\sigma\hat{U}) \stackrel{!}{=} \hat{l}^T\sigma,$$

which is satisfied if

$$\hat{R}(\hat{U}^\dagger\sigma\hat{U}) = \sigma \Leftrightarrow \hat{U}^\dagger\sigma\hat{U} = \hat{R}\sigma,$$

[3] More generally, we might consider $x_i = \hat{J}_i$, where $\hat{J}_i$ are $\mathrm{SU}(2)$ generators in a general spin S representation. However, to keep the notation simple, we restrict ourselves to $S = 1/2$.

or, in components

$$\hat{U}^\dagger \sigma_i \hat{U} = R_{ij}\sigma_j. \tag{3.10}$$

Now, the 'natural' matrix U to choose in the present context follows from the familiar isomorphy of the Lie algebras $\mathrm{so}(3)$ and $\mathrm{su}(2)$ of the groups $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$, resp. The correspondence between the two algebras is expressed by the table

$$\begin{array}{ccc} \mathrm{su}(2) & & \mathrm{so}(3) \\ \hat{J}_x = \frac{i}{2}\sigma_x = \frac{i}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \leftrightarrow & \hat{J}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\ \hat{J}_y = \frac{i}{2}\sigma_y = \frac{i}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} & \leftrightarrow & \hat{J}_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \\ \hat{J}_z = \frac{i}{2}\sigma_z = \frac{i}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \leftrightarrow & \hat{J}_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \tag{3.11}$$

assigning to $\mathrm{SU}(2)$ group generators their $\mathrm{SO}(3)$ analogs.[4] In our present context, the group $\mathrm{SO}(3)$ acts by rotation of three dimensional real space, while $\mathrm{SU}(2)$ acts through spinor rotation in a spin 1/2 representation. Let us represent a rotation matrix $\hat{R} \in \mathrm{SO}(3)$ as $\hat{R} = \exp(v_i \hat{J}_i)$, where the parameters $v_i \in \mathbb{R}$ define the rotation angles. The algebra isomorphy above then suggests a correspondence

$$\hat{R} = \exp(v_i \hat{J}_i) \leftrightarrow \exp(v_i \hat{J}_i) = \hat{U}, \tag{3.12}$$

where $\hat{U} \in \mathrm{SU}(2)$ and the generators on the right hand side are $\mathrm{su}(2)$ generators. This identification indeed does the job. The easiest way to check this is to consider rotations around a coordinate axis, e.g. $v_i = \theta\delta_{i1}$ a rotation by an angle θ around the x-axis. A straightforward calculation shows

$$\hat{R}_\theta \equiv \exp(\theta \hat{J}_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

and

$$\hat{U}_\theta \equiv \exp(\theta \hat{J}_x) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & i\sin\left(\frac{\theta}{2}\right) \\ i\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}.$$

[4] On a formal level, the Lie algebra correspondence is expressed by the commutation relations $[\hat{J}_i, \hat{J}_j] = \epsilon_{ijk}\hat{J}_k$.

These matrices indeed satisfy the compatibility relation (3.10). We thus conclude that the 'natural' identification of $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ group identification suggested by the structure of Lie algebras leads to the required spinor rotation invariance of the Schrödinger equation.

INFO Notice, however, that in spite of the isomorphy of their Lie algebras the two groups $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ are not globally isomorphic: consider a path $[0, 2\pi] \to \mathrm{SO}(3), t \mapsto \exp(t\hat{J}_z)$ in $\mathrm{SO}(3)$ group space parameterizing a rotation by an angle increasing from 0 to 2π. This path begins and ends at the unit element of the group, $\mathbb{I}_3 = \exp(0\hat{J}_z) = \exp(2\pi\hat{J}_z)$. However, its $\mathrm{SU}(2)$ analog $\exp(t\hat{J}_x) = \exp(\frac{it}{2}\sigma_1)$ begins at $\mathbb{I}_2$ and ends at $\exp(\frac{i2\pi}{2}\sigma_1) = -\mathbb{I}_2$. This mismatch shows that the group $\mathrm{SU}(2)$ is 'bigger' than $\mathrm{SO}(3)$ in that the *two* group elements $\pm\mathbb{I}_2 \in \mathrm{SU}(2)$ correspond to the single $\mathbb{I}_3 \in \mathrm{SO}(3)$. In other words,

The $\mathrm{SU}(2)$–$\mathrm{SO}(3)$ group correspondence is 2–1.

In mathematics, one says that $\mathrm{SU}(2)$ is the **universal covering group** of $\mathrm{SO}(3)$. The above non-uniqueness has important topological consequences. In $\mathrm{SO}(3)$, our reference path begins and ends at the same point $\mathbb{I}_3$. However, it can not be continuously contracted to a 'trivial' constant path $t \mapsto \mathbb{I}_3$ (think why.) The non-contractability of a closed path is a signature of a non-simply connected manifold. (A circle would be another example, i.e. a loop winding once around the circle cannot be contracted.) By contrast, $\mathrm{SU}(2)$ is simply connected. In fact, one may think of it as the smallest group 'covering' $\mathrm{SO}(3)$ to yield a simply connected manifold.

In physics, the 2–1 correspondence means that a 360deg rotation of coordinate space corresponds to a 180deg rotation in spin space. It takes a double 720deg space rotation to fully rotate spin. This phenomenon is detectable in interference experiments.

3.3.2 Relativistic invariance: $\mathrm{SL}(2,\mathbb{C})$–$\mathrm{SO}^+(3,1)$ correspondence

We will construct Dirac theory by generalization of the above construction to a relativistic setting. Recall from our discussion of section 3.2 that we aim for a Hamiltonian linear in the four momentum p. We thus start out from an ansatz,

$$\hat{H} = x^\mu \hat{p}_\mu, \tag{3.13}$$

where $\hat{p}_\mu$ now take over the role of the angular momentum components $\hat{l}_i$ of the previous section. Under a proper Lorentz transformation $\Lambda \in \mathrm{SO}^+(1,3)$ the four momentum transforms as $p_\mu \to \Lambda_\mu{}^\nu p_\nu$, and we seek for a realization of the 'coefficients' x^μ which transforms accordingly, $x^\mu \to \Lambda^\mu{}_\nu x^\nu$, so as to make $\hat{H}$ Lorentz-invariant.

Now, before turning more explicit, let us summarize a number of salient features of the theory:

- ▷ The Lorentz group contains the three-dimensional rotation group as a subgroup. This means that our theory must be derivable by extension of the theory constructed above. Corollary: Dirac theory necessitates the presence of a spin structure!
- ▷ However, unlike with the rotation group, the Lorentz group is non-compact. This means that the role of $\mathrm{SU}(2)$ above must be taken by a larger, and non-compact group (containing $\mathrm{SU}(2)$ as a subgroup.)

▷ Our group must be six-dimensional, with a Lie algebra satisfying the commutation relations (3.5).

While there exists a mathematical machinery for systematic identification of the relevant group structures, we here simply state that the group doing the job is $\mathrm{SL}(2,\mathbb{C})$, i.e. the group of two-dimensional matrices with complex entries and unit determinant. Indeed, $\mathrm{SU}(2) \subset \mathrm{SL}(2,\mathbb{C})$, as required. The group is non-compact, and six-dimensional (i.e. it is spanned by six *real* parameters.)[5] The Lie algebra, $\mathrm{sl}(2,\mathbb{C})$ is spanned by matrices $\sum_j z_j \frac{1}{2}\sigma_j \equiv \sum_j (a_j \frac{1}{2}\sigma_j + b_j \frac{i}{2}\sigma_j)$, where $z_j \in \mathbb{C}$, and $a_j, b_j \in \mathbb{R}$. Clearly, the sub-algebra $\mathrm{su}(2) \subset \mathrm{sl}(2,\mathbb{C})$ is generated by the parameters b_j, while the parameters a_j must correspond to the generators of Lorentz boosts. Indeed, it is straightforward to compute the commutation relations between the matrices $\frac{1}{2}\sigma_j$ and $\frac{i}{2}\sigma_j$ and to verify the correspondence

$$\begin{array}{ccc} \mathrm{sl}(2,\mathbb{C}) & & \mathrm{so}(1,3) \\ \hat{J}_i = \frac{i}{2}\sigma_i & \leftrightarrow & \hat{J}_i \\ \hat{K}_i = \frac{1}{2}\sigma_i & \leftrightarrow & \hat{K}_i, \end{array} \tag{3.14}$$

which generalizes Eq. (3.11) above, and establishes the Lie algebra isomorphy $\mathrm{sl}(2,\mathbb{C}) \simeq \mathrm{so}(1,3)$.[6] Now, the 'natural' generalization of the construct $\sigma_i \hat{l}_i$ to the present setting reads $\hat{p}_\mu \sigma^\mu$, where $\hat{p}_\mu \equiv -i\frac{\partial}{\partial x^\mu}$ are the components of the momentum operator, and $\{\sigma^\mu\} = (\sigma^0, \sigma^1, \sigma^2, \sigma^3)$ contains the unit matrix, σ^0, and the Pauli matrices σ^i.[7] Now, consider a proper Lorentz transformation

$$\Lambda \equiv \exp(\boldsymbol{\theta}\cdot\hat{\mathbf{J}} + \boldsymbol{\phi}\cdot\hat{\mathbf{K}}) \in \mathrm{SO}^+(1,3)$$

where $\boldsymbol{\theta}\cdot\hat{\mathbf{J}} \equiv \sum_i \theta_i \hat{J}_i$, and the three parameters θ_i (ϕ_i) specify a space rotation (Lorentz boost). Under this transformation, the momentum operator transforms covariantly, $\hat{p}_\mu \to \hat{p}_\nu \Lambda^\nu{}_\mu$, so that the generalization of (3.9) reads

$$\hat{p}_\mu \sigma^\mu \stackrel{!}{=} \hat{p}_\nu \Lambda^\nu{}_\mu (\hat{U}^\dagger \sigma^\mu \hat{U}),$$

where

$$\hat{U} \equiv \exp(\boldsymbol{\theta}\cdot\hat{\mathbf{J}} + \boldsymbol{\phi}\cdot\hat{\mathbf{K}}) \in \mathrm{SL}(2,\mathbb{C}) \tag{3.15}$$

[5] The latter statement follows from the observation that a two dimensional complex matrix $\hat{M}$ is specified by 2×4 real numbers. The (complex) equation $\det(M) = 1$ removes two degrees of freedom, leaves six.

[6] As with the pair $\mathrm{SU}(2) \leftrightarrow \mathrm{SO}(3)$, the correspondence between the Lie algebras does not extend to the full group manifold. Rather, $\mathrm{SL}(2,\mathbb{C})$ turns out to be the universal covering group of the (non-simply connected) group $\mathrm{SO}^+(1,3)$.

[7] We use uppercase indices because $\{\sigma^\mu\}$ will turn out to transform as a contravariant object.

is the 'natural' $\mathrm{SL}(2,\mathbb{C})$ representative of Λ.[8] Indeed is straightforward to verify[9] that

$$\hat{U}^\dagger \sigma^\mu \hat{U} = \Lambda^\mu{}_\nu \sigma^\nu,$$

and this entails the required invariance, $\hat{p}_\nu \Lambda^\nu{}_\mu (\hat{U}^\dagger \sigma^\mu \hat{U}) = \hat{p}_\nu \Lambda^\nu{}_\mu \Lambda^\mu{}_\rho \sigma^\rho = \hat{p}_\mu \sigma^\mu$.

Summarizing, we have confirmed that the operator $\hat{p}_\mu \sigma^\mu$ transforms covariantly, if it acts on two-component states $|\psi\rangle$ carrying an $\mathrm{SL}(2,\mathbb{C})$ spinor representation. However, before exploring the physical properties of these so-called **Weyl spinors**, we need to discuss a complication which gives the relativistic theory structure beyond the generalization $\mathrm{SU}(2) \to \mathrm{SL}(2,\mathbb{C})$.

3.3.3 Left- and right-handed Weyl spinors

In the previous section, we considered the 'natural' linear representation of group elements $\hat{U} \in \mathrm{SL}(2,\mathbb{C})$, i.e. linear transformations $|\psi\rangle \to \hat{U}|\psi\rangle$. Now, $\mathrm{SL}(2,\mathbb{C})$ is richer than its subgroup $\mathrm{SU}(2)$ in that it admits a second, and inequivalent two-dimensional representation. Indeed, we may choose to assign to $\hat{U}$ the transformation $|\psi\rangle \to (\hat{U}^{-1})^\dagger |\psi\rangle$. One can show, that this transformation is unitarily inequivalent to the one above, i.e. there is no *fixed* unitary matrix $\hat{T} \in \mathrm{SU}(2)$, such that $(\hat{U}^{-1})^\dagger = T\hat{U}\hat{T}^{-1}$ for all $\hat{U} \in \mathrm{SL}(2,\mathbb{C})$. Within the boost/rotation parameterization (3.15), we have

$$(\hat{U}^\dagger)^{-1} = \exp(\boldsymbol{\theta} \cdot \hat{\mathbf{J}} - \boldsymbol{\phi} \cdot \hat{\mathbf{K}}), \tag{3.16}$$

i.e. the two representations differ by an inversion of the boost parameter. This shows that for $\hat{U} \in \mathrm{SU}(2) \subset \mathrm{SL}(2,\mathbb{C})$, or $\phi_i = 0$, we have $(\hat{U}^{-1})^\dagger = \hat{U}$, i.e. when restricted to the unitary subgroup the two representations collapse, in line with the fact that, up to unitary transformation, there is only one unitary spin $1/2$ $\mathrm{SU}(2)$-representation.

The alternative representation gives us an option to construct a second covariant formulation of the theory. To see this, let us define the vector of spin matrices

$$\bar{\sigma}^\mu \equiv (\sigma^0, -\sigma^1, -\sigma^2, -\sigma^3).$$

It is then not difficult to verify that

$$(\hat{U})^{-1} \bar{\sigma}^\mu (\hat{U}^{-1})^\dagger = \Lambda^\mu{}_\nu \bar{\sigma}^\mu,$$

again transforms as a contravariant vector. Lacking any a priori physical principles discriminating between the two different representations, we'd better keep an eye on both. For reasons to become clear momentarily, states transforming under $\hat{U}$, and $(\hat{U}^\dagger)^{-1}$, resp., are denoted $|\psi_L\rangle$ and $|\psi_R\rangle$ and called **left- and right-handed Weyl spinors**. Their transformation behavior is summarized by

$$\begin{aligned} |\psi_L\rangle &\xrightarrow{\hat{U}} \hat{U}|\psi_L\rangle, \\ |\psi_R\rangle &\xrightarrow{\hat{U}} (\hat{U}^{-1})^\dagger |\psi_R\rangle. \end{aligned} \tag{3.17}$$

[8] For the appearance of the disclaimer 'natural', see below.

[9] From the previous section we know that the formula holds for rotations. A direct computation shows that it also holds for axes-oriented Lorentz boosts, say $\Lambda = \exp(\phi \hat{K}_x)$. Since a general Lorentz transformation can be obtained by a succession of rotations and boosts, this proves the formula's validity.

where (up to the notorious (2–1) ambiguity) the group element $\hat{U} \in \mathrm{SL}(2,\mathbb{C})$ is determined by the Lorentz transformation under consideration.

3.3.4 From Weyl– to Dirac–spinors

Our so-far considerations were primarily based on criteria of transformation-invariance, conditions imposed by the mathematical consistency of the theory. We will now add to this kinematic (means, physical) considerations to upgrade to the full Dirac theory. It may be worth mentioning that Dirac theory, too, can be derived solely on the basis of mathematical structures (cf. our remarks on Clifford algebras below). This is manifestation of the strong anchorage of (quantum) relativity in geometric principles.

Assume that the quantum state of a particle of mass m in its rest frame is described by the two-component spinor ψ. Denote the state transformed to a frame moving relative to the rest frame with velocity $\mathbf{v}$ by $\psi(p)$, where $p = \{p_\mu\} = (\gamma m, -\gamma m\mathbf{v})$ is the covariant four momentum. The transformation to the moving frame is described by non-vanishing boost parameters ϕ_i. Now, Eqs. (3.15) and (3.16) tell us that the transformation matrices of generating left– and right–handed Weyl states $\psi_L(p) = \hat{U}\psi(0)$ and $\psi_R(p) = (\hat{U}^{-1})^\dagger\psi(0)$ differ in the sign of the boost parameter, i.e. they are generally different. However, in the rest frame, we have no reason to discriminate between 'left' and 'right' state, $\psi_L(0) = \psi_R(0) \equiv \psi(0)$. This consideration signals that 'left-' and 'right-handedness' must be a property linked to the kinematic motion of particles.

To explore this point, we use that the boost parameter $\boldsymbol{\phi} = \mathbf{n}\phi$, where $\mathbf{n}$ is the unit vector in $\mathbf{p}$-direction, and ϕ is given in (3.4). We then have

$$\begin{aligned}\psi_{L/R}(p) = \exp\left(\pm\frac{1}{2}\boldsymbol{\phi}\cdot\hat{\mathbf{K}}\right)\psi(0) = \left(\cosh(\frac{\phi}{2}) \pm \sinh(\frac{\phi}{2})\mathbf{n}\cdot\boldsymbol{\sigma}\right)\psi(0) = \\ = \left(\left(\frac{\gamma+1}{2}\right)^{1/2} \pm \left(\frac{\gamma-1}{2}\right)^{1/2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)\psi(0) \qquad (3.18)\end{aligned}$$

where the second equality relies on some straightforward algebra based on (3.4). Multiplying the equation by $\left(\frac{\gamma+1}{2}\right)^{1/2} \mp \left(\frac{\gamma-1}{2}\right)^{1/2}\mathbf{n}\cdot\boldsymbol{\sigma}$, we obtain

$$\left(\left(\frac{\gamma+1}{2}\right)^{1/2} \mp \left(\frac{\gamma-1}{2}\right)^{1/2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)\psi_{L/R}(p) = \left(\left(\frac{\gamma+1}{2}\right)^{1/2} \pm \left(\frac{\gamma-1}{2}\right)^{1/2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)\psi_{R/L}(p),$$

and one more multiplication by $\left(\frac{\gamma+1}{2}\right)^{1/2} \pm \left(\frac{\gamma-1}{2}\right)^{1/2}\mathbf{n}\cdot\boldsymbol{\sigma}$ gets us to

$$\psi_{L/R}(p) = \left(\gamma \pm (\gamma^2-1)^{1/2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)\psi_{R/L}(p) = \frac{1}{m}\left(E \pm p\mathbf{n}\cdot\boldsymbol{\sigma}\right)\psi_{R/L}(p).$$

where we used that the particle energy $E^2 = p^2 + m^2$ is related to the parameter γ through $\gamma = E/m$ (show it.)

Using that $E = p_0$, we may formulate this equation as

$$\begin{pmatrix} -m & p_0 + p\cdot\boldsymbol{\sigma} \\ p_0 - p\cdot\boldsymbol{\sigma} & m \end{pmatrix} \begin{pmatrix} \psi_R(p) \\ \psi_L(p) \end{pmatrix} = 0.$$

Introducing four-component spinor

$$\psi(p) = \begin{pmatrix} \psi_R(p) \\ \psi_L(p) \end{pmatrix}, \tag{3.19}$$

and the 4×4 matrices

$$\gamma^0 = \begin{pmatrix} & \mathbb{I} \\ \mathbb{I} & \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} & -\sigma^i \\ \sigma^i & \end{pmatrix}, \tag{3.20}$$

this equation assumes the concise form

$$\boxed{(\gamma^\mu p_\mu - m)\,\psi(p) = 0.} \tag{3.21}$$

Eq. (3.21) is the (momentum representation) of the celebrated **Dirac equation**. In a coordinate representation, $\hat{p}_\mu \to i\frac{\partial}{\partial x^\mu}$, the equation assumes the form

$$(i\gamma^\mu \partial_\mu - m)\,\psi(x) = 0.$$

Finally, interpreting p_μ as eigenvalues of linear operators $\hat{p}_\mu$ it may be written as (3.22). The historical derivation of this equation, different from the one discussed here, is reviewed in the info block below.

INFO The construction above is different from the historical derivation of the Dirac equation. In fact, **Dirac's original derivation** is much shorter, and rather elegant at that. However the price to be payed is that, unlike with the derivation given above, the physical principles behind the mathematical structure of the equation remain opaque. Still, it is rewarding to recapitulate Dirac's original line of arguments: For the reasons outlined in section 3.2, Dirac was after an equation linear in the four-momentum operator. He thus started from an ansatz

$$(\gamma^\mu \hat{p}_\mu - m)|\psi\rangle = 0, \tag{3.22}$$

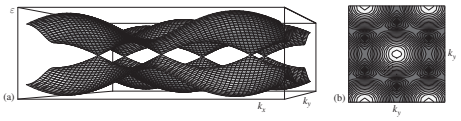
with as yet undetermined 'coefficients' γ^μ. Multiplication with $\gamma^\nu \hat{p}_\nu$ gives

$$\gamma^\nu \hat{p}_\nu(\gamma^\mu \hat{p}_\mu - m)|\psi\rangle = (\gamma^\mu\gamma^\nu \hat{p}_\nu \hat{p}_\mu - m^2)|\psi\rangle = \left(\frac{1}{2}[\gamma^\mu, \gamma^\nu]_+ \hat{p}_\nu \hat{p}_\mu - m^2\right)|\psi\rangle 0,$$

where in the second equality we used the defining equation once more, and the third equality is based on the commutativity of the components of the momentum operator. Now, in all what we are doing, we must stay compatible with the relativistic dispersion relation expressed by the Klein Gordon equation $(\hat{p}^\mu \hat{p}_\mu - m^2)|\psi\rangle = (g^{\mu\nu}\hat{p}_\nu \hat{p}_\mu - m^2)|\psi\rangle = 0$. We satisfy this relation if

$$\boxed{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}.} \tag{3.23}$$

Dirac observed that this relation can not be realized with real, or complex valued coefficients. The simplest mathematical object doing the job turned out to be the γ-matrices (3.20), and this is how these objects where originally introduced.
Formally, one may think of (3.23) as the defining relations of an *algebra* generated by the matrices γ^μ. Indeed, the matrices γ^μ are elements of a vector space (they can be added and multiplied by numbers), endowed with a product (the matrix product, constraint by the relation (3.23).) The set formed by taking all linear combinations of products $\gamma^\mu, \gamma^\mu\gamma^\nu, \gamma^\mu\gamma^\nu\gamma^\rho, \ldots$ thus forms an algebra, the so-called **Clifford algebra**. The algebraic approach to Dirac theory turned out to be eminently rich, both from a physics and a mathematics point of view: our previous discussion has been limited to four-dimensional space, endowed with a metric of signature $(+,-,-,-)$. [10] However, Clifford algebras can be defined for spaces of arbitrary dimension and metric. The representation theory of the general Clifford algebra provides a handle for the construction of Dirac operators in much more general settings, and this includes spaces with metrics of non-trivial curvature, as relevant to general relativity.
Dirac operators of dimension 1,2, and 3 find physical realizations in various condensed matter systems. Arguably the most prominent example are two dimensional sheets of sp^2-hybridized carbon, a material called **graphene**. The dispersion relation of graphene (cf. the figure) contains characteristic cusps, which signals that the low energy physics of the compound is governed by a two-dimensional variant of the Dirac operator. The experimental discovery of graphene in 2005 ($\rightarrow$ Nobel prize 2010) sparked a wave of research on relativistic quantum mechanics in a setting off the beaten tracks of $3+1$-dimensional particle physics.



3.3.5 Things to be learned from the structure of the Dirac equation

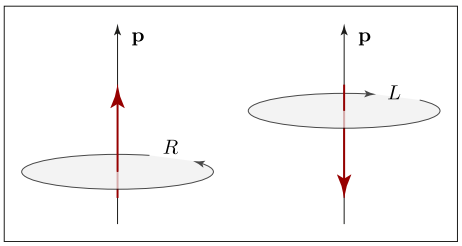
A great deal more was hidden in the Dirac equation than the author had expected when he wrote it down in 1928. Dirac himself remarked in one of his talks that his equation was more intelligent than its author. It should be added, however, that it was Dirac who found most of the additional insights. Weisskopf on Dirac

Above we have seen how the 'kinematics' of Lorentz boosts establishes a correlation between left- and right-Weyl spinors, which finds its algebraic expression in the Dirac equation. In this section, we will take a closer look at the (matrix–)structure of the equation. Specifically, we will discover that the form of the γ-matrices contains important information on the relativistic dynamics of quantum particles which can be extracted without elaborate calculation.

Helicity

[10] The **signature** of a metric is defined by the number of negative eigenvalues. While the magnitude and ordering of eigenvalues can be changed by scaling and unitary transformation, the signature is a robust invariant.

Let us first take a look at the **Dirac equation of massless particles**, $m = 0$. In this limit, the left and the right sector decouple, and the Dirac equation reduces to



$$(p_0 \mp \boldsymbol{\sigma} \cdot \mathbf{p})\psi_{R/L}(p) = 0. \tag{3.24}$$

Now, for $m = 0$ we have $p_0 = |bp|$, which means that

$$\boldsymbol{\sigma} \cdot \mathbf{n}\psi_{R/L}(p) = \pm\psi_{R/L}(p), \tag{3.25}$$

where $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$ as before. These equations explain the origin of the denotation left- and right- spinor. The Weyl spinors are eigenstates of the **helicity operator** $\boldsymbol{\sigma}\cdot\mathbf{n}$, which measures the spin component in the direction of travel, $\mathbf{n}$. The equation tells us that massless quantum particles carry helicity ± 1, i.e. a unit spin which is directed in the direction of travel (R) or opposite to it (L). Eq. (3.25) explains the origin of the denotation 'left-' and 'right-handed' Weyl spinors.

Parity and time reversal

Remember that we introduced the Weyl spinors as a means to incorporate *proper* Lorentz transformations into quantum theory. But what about Lorentz transformations outside the proper Lorentz group $\mathrm{SO}^+(1,3)$, notably parity and time reversal? **Parity** acts on the four momentum as $\mathcal{P} : p = (p_0, \mathbf{p}) \mapsto (p_0, -\mathbf{p}) \equiv \Lambda_{\mathcal{P}} p$. Following the general logics of our approach, we thence need to find a Dirac spinor transformation $|\psi\rangle \to \hat{U}_{\mathcal{P}}|\psi\rangle$, such that

$$\hat{U}_{\mathcal{P}}\gamma^\mu\hat{U}_{\mathcal{P}}^{-1} \stackrel{!}{=} (\Lambda_{\mathcal{P}})^\mu{}_\nu\gamma^\nu$$

and the Dirac bilinear $p_\mu\Lambda^\mu \to p_\nu(\Lambda_{\mathcal{P}})^\nu{}_\mu(\Lambda_{\mathcal{P}})^\mu{}_\rho\gamma^\rho = p_\mu\gamma^\mu$ remains invariant. Using the commutation relations (3.23) of the γ-matrices, it is evident that $\hat{U}_{\mathcal{P}} = \gamma^0$ does the job. Indeed,

$$(\gamma^0)^{-1}(\gamma^0, \gamma^1, \gamma^2, \gamma^3)\gamma^0 = (\gamma^0, -\gamma^1, -\gamma^2, -\gamma^3)$$

as required. Under parity the four-component Dirac spinor thus transforms

$$|\psi\rangle \xrightarrow{\mathcal{P}} \gamma^0|\psi\rangle. \tag{3.26}$$

We observe that parity exchanges the left- and the right-handed components of the spinor. This was to be expected, as parity does not change spin (on account of spin representing an axial vector), but does change momentum. Helicity, thus, will be inverted.

Turning to **time reversal**, our discussion of section 1.4.2 tells us that this transformation is best explored in the coordinate representation of the Dirac equation,

$$(i\partial_\mu\gamma^\mu - m)\psi(x) = 0.$$

Under time reversal, $\mathcal{T} : x = (x^0, \mathbf{x}) \to (-x^0, \mathbf{x}) \equiv \Lambda_{\mathcal{T}} x$. From our previous experience we also know that we will need an anti–unitary transformation, i.e. a product of complex

conjugation, $\hat{K}$, and a unitary transformation, $|\psi\rangle \to \hat{U}_{\mathcal{T}}\hat{K}|\psi^*\rangle$ to implement time reversal. Setting $\hat{U}_{\mathcal{T}} \equiv \gamma_1\gamma_3$, we indeed find

$$U_{\mathcal{T}}\hat{K}(\gamma^0,\gamma^1,\gamma^2,\gamma^3)\hat{K}^{-1}\hat{U}_{\mathcal{T}}^{-1} = U_{\mathcal{T}}(\gamma^0,\gamma^1,-\gamma^2,\gamma^3)\hat{U}_{\mathcal{T}}^{-1} = (\gamma^0,-\gamma^1,-\gamma^2,-\gamma^3) = -\Lambda^\mu{}_\nu\gamma^\nu.$$

From this relation we readily deduce that the time reversed spinor

$$|\psi\rangle \xrightarrow{\mathcal{T}} \gamma^1\gamma^3\hat{K}|\psi\rangle \tag{3.27}$$

solves the Dirac equation for the time reversed derivative operator $\frac{\partial}{\partial(\Lambda_{\mathcal{T}}x)^\mu} = (-\partial_0,\partial_1,\partial_2,\partial_3)$. Notice that

$$\gamma^1\gamma^3 = i\sigma^2 \otimes \mathbb{I},$$

where $\mathbb{I}$ acts in LR-space. We learn that time reversal acts on a Dirac spinor as $\psi(x) \to (i\sigma^2)\psi(x)^*$,[11]. Unlike with non-relativistic quantum mechanics, where the the behavior of spinful states under time reversal (cf. Eq. (1.72)) had to be postulated, it here follows from the criterion of Lorentz invariance. This is one manifestation of a more general principle to be discussed in section xx below: many physical phenomena pertaining to the spin (even of 'slow' particles) are of relativistic origin.

3.4 Many particle interpretation

3.4.1 Particle current

In section 3.2 we have seen that negative energy states and negative particle 'densities' where phenomena accompanying the Klein-Gordon equation. Let us explore how the Dirac equation performs in this compartment. We begin by taking the hermitian conjugate of the Dirac equation,

$$\psi(x)^\dagger\left(-i(\gamma^\mu)^\dagger\overleftarrow{\partial}_\mu - m\right) = 0,$$

where the arrow indicates that the derivative acts on the spinor to the left. Now, $(\gamma^i)\dagger = -\gamma^i$, while $(\gamma^0)^\dagger = \gamma^0$, which tells us that the hermitian adjoint has messed up the covariant structure of the equation. We can repair this defect by multiplying the equation from the right by γ^0. Using the commutation relations (3.23) and defining the **conjugate Dirac spinor**

$$\bar{\psi} \equiv \psi^\dagger\gamma^0,$$

we obtain the equation

$$\bar{\psi}(x)\left(-i\gamma^\mu\overleftarrow{\partial}_\mu - m\right) = 0.$$

[11] Rather than writing $\bar{z}$ for complex conjugation, we here use z^*. In Dirac theory the overbar is commonly reserved for another operation, see section 3.4 below.

With this definition, a good candidate expression for a **four-particle current** reads

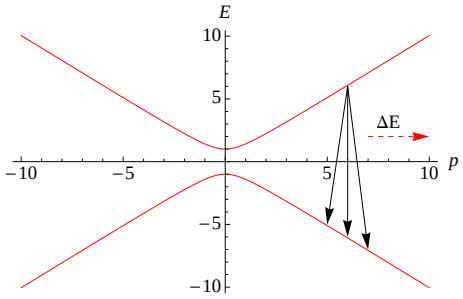
$$j^\mu \equiv \bar{\psi}\gamma^\mu\psi. \tag{3.28}$$

From our previous discussion, we know that $j \equiv \{j^\mu\}$ transforms as a contravariant vector, as it should. Further, j obeys a continuity equation:

$$\partial_\mu j^\mu = \bar{\psi}\left(\overleftarrow{\partial}_\mu\,\gamma^\mu + \partial_\mu\gamma^\mu + im - im\right)\psi = 0,$$

where in the second equality, we added $0 = im - im$. Unlike with the Klein-Gordon equation, the Dirac density $j^0 = \bar{\psi}\gamma^0\psi = \psi^\dagger\psi$ is positive, i.e. our problem of negativ particle densities is gone!

Our enthusiasm about the recovered positivity of density may be tempered by the follow-up observation that the second problem of the Klein-Gordon equation, the existence of negative energy states, persists. To see this, multiply the Dirac equation $(E\gamma^0 - \mathbf{p}\cdot\boldsymbol{\gamma} - m)\psi(p)$ by γ^0 to obtain $(E - \gamma^0(\mathbf{p}\cdot\boldsymbol{\gamma} - m))\psi(p) = 0$. It is straightforward to compute the eigenvalues (try to do it in 'elegant' terms, i.e. without elaborate calculation) as



$$E = \pm\sqrt{|\mathbf{p}|^2 + m^2}.$$

As expected, the energies obtain by taking the square root of the relativistic expression $E^2 = |\mathbf{p}|^2 + m^2$, and this includes both the positive and the negative branch.

Notice the consequences of this observation: it looks like a Dirac particle prepared in an arbitrary initial state particle may cascade down to negative energies, thence releasing an infinite amount of transition energy ΔE. This doesn't make physical sense.

3.4.2 Dirac's way out

Confronted with these problems, Dirac proposed a revolutionary re-interpretation of his equation. Within the new picture, the peculiar structure of the Dirac spectrum not only ceased to be problematic, it actually became the source of physical phenomena crucial to the foundations of modern particle physics. In essence, **Dirac's postulates** can be summarized as follows:

- ▷ The particles described by the Dirac equation necessarily have to be **fermions**. The double occupancy of Dirac eigenstates is, thence, forbidden.

- ▷ (At zero temperature), all negative energy eigenstates are occupied. The union of all filled states forms the so-called **Dirac sea**. Together with the continuum of empty states, the Dirac sea defines the **Dirac vacuum**

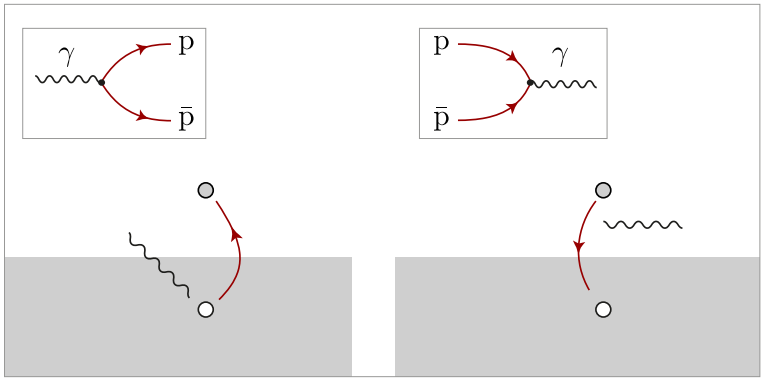


Figure 3.1: Left: Vacuum polarization. High energy photons create particle/anti-particle pairs by excitation of filled states into the empty positive energy continuum. Right: Particle/anti-particle annihilation. The reverse process. A created photon carries the excess energy away.

Comments: (i) the first postulate amounts to a linkage of two of the most fundamental properties of quantum particles, spin and exchange statistics. The Dirac equation requires spin $1/2$ and, according to the postulate, fermionic exchange statistics. (ii) The first postulate effectively states that a meaningful interpretation of Dirac theory requires a **many particle system**. (iii) The filling of the negative eigenstates prevents the constituents of this system from cascading down in energy. However, if the system is exposed to an energy source providing energy $\Delta E > 2m$ larger than the gap between the Dirac sea and the empty states, negative energy particles may be kicked out into the continuum of empty states (Fig. 3.1, left.) What we are left with, then, is a 'hole' in the Dirac sea, plus a positive energy particle. Interpreting the hole as an **anti-particle**, we conclude that sources of high energy radiation, such as high energy photons, γ, may 'polarize' the vacuum by creating particle/anti-particle pairs, $\gamma \to \mathrm{p} + \overline{\mathrm{p}}$. Here p may be any particle described by the Dirac equation, electrons, protons, neutrinos, ... Conversely, a positive energy particle may occupy a hole in the Dirac sea, thereby releasing energy (Fig. 3.1, right.) According to the interpretation above, we should think of this as a particle/anti-particle annihilation process, $\mathrm{p} + \overline{\mathrm{p}} \to \gamma$. (Notice that the minimum energy released in the process is twice the rest mass energy. For a proton pair, this will be about $2\mathrm{GeV}$ — matter/anti-matter reactions make for a violent source of energy!) (iv) Dirac's interpretation thus predicts the existence of anti-particles for all particles. Dirac's 1928 prediction of antiparticles, followed by the 1929 discovery of the positron $\overline{e}$, marked a cornerstone in the advent of particle physics.

3.5 Dirac many particle theory

The important bottom line to be drawn from our discussion of section 3.5 is that Dirac theory has to be interpreted as a many particle theory. Before, we had attempted to construct the Dirac equation as a generalization of the Schrödinger equation, i.e. we were firmly operating within the realm of single particle quantum mechanics. The question then arises as to how

our previous consideration might be integrated into a many body framework.

With the machinery of second quantization in place, it is actually not difficult to solve this problem. When we formulated the Dirac equation (3.22), we met the (prejudiced) choice to interpret the Dirac spinor $|\psi\rangle$ as a single particle 'wave function', an approach we saw does not pass the test of physical consistency. Now, from many body theory, we know that to each Schrödinger like equation $(i\partial_t - \hat{H})|\psi\rangle = 0$, we have an associated Heisenberg equation $(i\partial_t + [\hat{H},\ ])\hat{A} = 0$ for operators. Specifically, in a single particle basis characterized by indices $\{l\}$, the equation $(i\delta_{lk}\partial_t - H_{lk})\psi_k$ corresponds to a second quantized representation $\hat{H} = a_j^\dagger H_{jk} a_k$. Applied to the operator a_l, the Heisenberg equation assumes the form $(i\delta_{lk}d_t - H_{lk})a_k$, equal in form to the Schrödinger equation, only that the object we act upon is an operator.

This observation suggests a **reinterpretation of the Dirac equation**,

$$(i\gamma^\mu \partial_m - m)\psi(x) = 0, \tag{3.29}$$

where $\psi(x)$ now is a four-component *operator valued* spinor. Being fermion operators, the components of ψ obey the **equal time commutation relations**,

$$x^0 = x^{0\prime}: \qquad [\psi^a(x), \psi^{a'}(x')]_+ = \delta^{aa'}\delta(\mathbf{x} - \mathbf{x}'). \tag{3.30}$$