Moderne Theoretische Physik II

Advanced Quantum Mechanics

Alexander Shnirman

Institute TKM, Karlsruhe Institute of Technology, Karlsruhe, Germany

(Dated: December 10, 2024)

Contents

I. TIME-DEPENDENT PERTURBATION THEORY

A. Evolution operator

Schrödinger equation:

$$i\hbar\frac{d}{dt}\left|\Psi\right\rangle = H(t)\left|\Psi\right\rangle \ , \tag{1}$$

where $H(t)$ is (in general time-dependent) Hamiltonian.

Initial condition: We know the state of the system (wave function) at $t = t_0$. We denote it as $|\Psi(t_0)\rangle$.

We want to know the state at arbitrary time $t > t_0$. Formal solution reads

$$\left|\Psi(t)\right\rangle = U(t,t_0)\left|\Psi(t_0)\right\rangle \ , \tag{2}$$

where $U(t,t_0)$ is the evolution operator.

Evolution operator satisfies the same Schrödinger equation:

$$i\hbar\frac{d}{dt}U(t,t_0) = H(t)U(t,t_0) \ , \tag{3}$$

and the initial condition $U(t_0,t_0) = \hat{1}$ (unity operator). As t_0 is arbitrary we have $U(t,t) = \hat{1}$ for arbitrary t.

Evolution operator is unitary:

$$[U(t,t_0)]^\dagger = [U(t,t_0)]^{-1} = U(t_0,t) \ . \tag{4}$$

Proof: $(i\hbar\dot{U} = HU) \to (-i\hbar\dot{U}^\dagger = U^\dagger H)$ *on the other hand* $(UU^{-1} = 1) \to (\dot{U}U^{-1} = -U\dot{U}^{-1}) \to (\dot{U}^{-1} = -U^{-1}\dot{U}U^{-1} = -U^{-1}HUU^{-1}/(i\hbar)) \to (-i\hbar\dot{U}^{-1} = U^{-1}H)$. *End of proof.*

Since we assumed initially that $t > t_0$, the operator $U(t_0,t)$ should be understood as evolution operator backward in time.

For time-independent Hamiltonian

$$U(t,t_0) = e^{-iH(t-t_0)/\hbar} \ . \tag{5}$$

B. Transitions

By transitions we usually (but not always) understand the following: In the far past

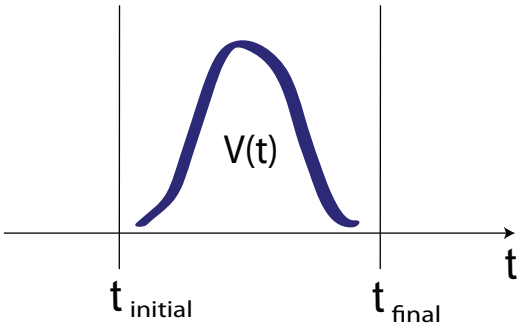

FIG. 1: Schematic picture showing the perturbation $V(t)$ which acts only during a finite time period.

and in the far future the system is described by a time-indepndent Hamiltonian H_0. In the middle a perturbation $V(t)$ is added so that the full Hamiltonian reads

$$H = H_0 + V(t) \ . \tag{6}$$

Example: a pulse of electric field acts on an atom and excites it.

Assume that $V(t < t_{\text{initial}}) = 0$ and $V(t > t_{\text{final}}) = 0$ (Fig. 1).

We introduce the eigenstates $|n\rangle$ of H_0

$$H_0 \, |n\rangle = E_n \, |n\rangle \ , \tag{7}$$

These states evolve in time as

$$|n(t)\rangle = e^{-iE_n t/\hbar} \, |n\rangle \ . \tag{8}$$

(by choice $|n(t=0)\rangle = |n\rangle$).

Assume that before the perturbation $V(t)$ is switched on the system was in one of the eigenstates of H_0. Say $|\Psi(t < t_{\text{initial}})\rangle = |n(t)\rangle$. After the perturbation passes, i.e., for $t > t_{\text{final}}$, the wave function can be expanded in the basis of the eigenstates:

$$|\Psi(t > t_{\text{final}})\rangle = \sum_m a_{n\to m} \, |m(t)\rangle \ . \tag{9}$$

Thus the state is a superposition of eigenstates with amplitudes $a_{n\to m}$. The amplitudes are time-independent, since $V(t > t_{\text{final}}) = 0$. The probability of transition from state $|n(t)\rangle$ to state $|m(t)\rangle$ is given by $P_{n\to m} = |a_{n\to m}|^2$.

C. Interaction picture

Interaction picture is in the middle between Schrödinger and Heisenberg pictures. In Schrödinger picture the wave functions evolve with full Hamiltonian. In Heisenberg picture the wave functions do not evolve at all, instead the operators evolve. Idea of interaction picture: the wave function evolves only due to perturbation $V(t)$.

We introduce new wave functions $|\Psi_I(t)\rangle$ by

$$|\Psi(t)\rangle = e^{-iH_0 t/\hbar} |\Psi_I(t)\rangle \ , \tag{10}$$

where $|\Psi(t)\rangle$ is the Schrödinger wave function.

Equation of motion (Schrödinger equation) in the interaction picture:

$$i\hbar \frac{d}{dt} |\Psi_I\rangle = V_I(t) |\Psi_I\rangle \ , \tag{11}$$

where

$$V_I(t) \equiv e^{iH_0 t/\hbar} V(t) e^{-iH_0 t/\hbar} \tag{12}$$

D. Dyson expansion

Evolution operator in the interaction picture:

$$|\Psi_I(t)\rangle = U_I(t, t_0) |\Psi_I(t_0)\rangle \tag{13}$$

It satisfies the equation of motion:

$$i\hbar \frac{d}{dt} U_I(t, t_0) = V_I(t) U_I(t, t_0) \ , \tag{14}$$

Initial condition: $U_I(t_0, t_0) = \hat{1}$.

Rewrite Eq. (14) as

$$\begin{aligned} U_I(t, t_0) &= U_I(t_0, t_0) + \int_{t_0}^{t} \left(-\frac{i}{\hbar} \right) V_I(t') U_I(t', t_0) dt' \\ &= \hat{1} + \int_{t_0}^{t} \left(-\frac{i}{\hbar} \right) V_I(t') U_I(t', t_0) dt' \ . \end{aligned} \tag{15}$$

Now iterate: as zeroth order approximation take $U_I^{(0)}(t,t_0) = \hat{1}$. Substitute into the RHS of Eq. (15). This gives the first order approximation:

$$U_I^{(1)}(t,t_0) \approx 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\, V_I(t_1)\ . \tag{16}$$

Substitute again into the RHS of Eq. (15). Result: second order approximation:

$$U_I^{(2)}(t,t_0) \approx 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\, V_I(t_1) + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, V_I(t_1)V_I(t_2)\ . \tag{17}$$

And so on:

$$U_I^{(n)}(t,t_0) \approx 1 + ... + \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 ... \int_{t_0}^{t_{n-1}} dt_n\, V_I(t_1)V_I(t_2)....V_I(t_n)\ . \tag{18}$$

Note, that in this integral the times are ordered: $t_1 > t_2 > .... > t_n$. Can be rewritten as

$$U_I^{(n)}(t,t_0) \approx 1 + ... + \frac{1}{n!}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 ... \int_{t_0}^{t} dt_n\, T\left[V_I(t_1)V_I(t_2)....V_I(t_n)\right]\ , \tag{19}$$

where T is the time ordering operator: $T\left[V_I(t_1)V_I(t_2)\right] = V_I(t_1)V_I(t_2)$ if $t_1 > t_2$ and $T\left[V_I(t_1)V_I(t_2)\right] = V_I(t_2)V_I(t_1)$ if $t_2 > t_1$.

A symbolic way of writing:

$$U_I(t,t_0) = T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\, V_I(t')\right) \tag{20}$$

E. 1-st order transition probability

We assume that at $t_0 < t_{\text{initial}}$ the system was prepared in state $|n(t)\rangle$. In the interaction picture that means that $|\Psi_I(t < t_{\text{initial}})\rangle = |n\rangle$ (no time dependence). We want to know the state $|\Psi_I(t)\rangle$ for arbitrary $t > t_{\text{final}}$. It is given by

$$|\Psi_I(t)\rangle = U_I(t,t_0)\ |n\rangle \approx \left(1 - \frac{i}{\hbar}\int_{t_0}^{t} dt'\, V_I(t')\right)\ |n\rangle\ , \tag{21}$$

The probability that the system is in state $|m\rangle \neq |n\rangle$ is given by

$$P_{n\to m} \equiv \left|\langle m\,|\Psi_I(t)\rangle\right|^2 \approx \frac{1}{\hbar^2}\left|\int_{t_0}^{t} dt'\, \langle m|\, V_I(t')\,|n\rangle\right|^2 \tag{22}$$

F. The Golden Rule

Assume that the perturbation is time-independent between t_{initial} and t_{final} (Fig. 2). That

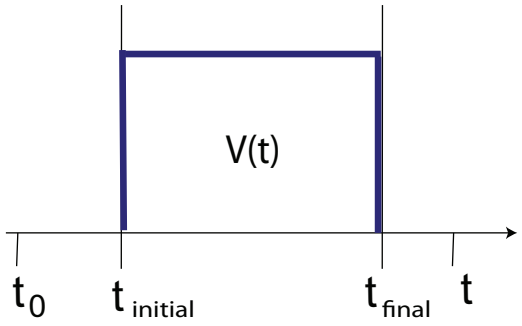

FIG. 2: Time-independent perturbation $V(t)$.

is $V(t_{\text{initial}} < t < t_{\text{final}}) = V$ and $V(t) = 0$ otherwise. Then $\langle m| V_I(t) |n\rangle = V_{mn} e^{-i(E_n - E_m)t/\hbar}$, where $V_{mn} \equiv \langle m| V |n\rangle$. Thus

$$P_{n\to m} \approx \frac{1}{\hbar^2} \left| V_{mn} \int_{t_0}^{t} dt' e^{-i(E_n - E_m)t'/\hbar} \right|^2 = \frac{1}{\hbar^2} |V_{mn}|^2 \frac{\sin^2(\omega_{mn}\Delta t/2)}{(\omega_{mn}/2)^2} , \tag{23}$$

where $\omega_{mn} \equiv (E_m - E_n)/\hbar$ and $\Delta t \equiv t_{\text{final}} - t_{\text{initial}}$.

The function $\frac{\sin^2(\omega\Delta t/2)}{(\omega/2)^2}$ (considered as function of ω with Δt being a parameter) is plotted in Fig. 3. It has a sharp peak near $\omega \approx 0$ of width $\sim 1/(\Delta t)$. Thus transitions happen only to

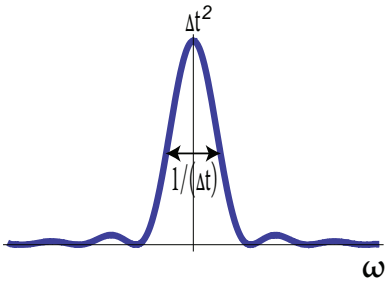

FIG. 3: "Delta function"

the states with energies close to that of the original state E_n. More precisely, $|\omega_{mn}||\Delta t| < 1$ or $|E_m - E_n||\Delta t| < \hbar$. This is called uncertainty relation for energy-time.

It is possible to show that

$$\int d\omega \frac{\sin^2(\omega\Delta t/2)}{(\omega/2)^2} = 2\pi\Delta t . \tag{24}$$

This is why *in certain situations* the following substitution is valid

$$\frac{\sin^2(\omega\Delta t/2)}{(\omega/2)^2} \to 2\pi\delta(\omega)\Delta t \ . \tag{25}$$

Then, rather formally,

$$P_{n\to m} \approx \frac{1}{\hbar^2}\left|V_{mn}\right|^2 \frac{\sin^2(\omega_{mn}\Delta t/2)}{(\omega_{mn}/2)^2} \approx \frac{2\pi}{\hbar}\left|V_{mn}\right|^2 \delta(E_m - E_n)\,\Delta t \ . \tag{26}$$

This gives for the transition rate

$$\Gamma_{n\to m} = \frac{P_{n\to m}}{\Delta t} = \frac{2\pi}{\hbar}\left|V_{mn}\right|^2 \delta(E_m - E_n) \ . \tag{27}$$

The last formula is the most common form of the Golden Rule.

What are the situations when the substitution by the delta function are allowed and how can one use the rate containing a delta function? Mostly, the Golden Rule is applicable if we have to sum over many densely packed levels $|m\rangle$ to which the transition happens, i.e., for transitions to continuum (Fig. 4). That is, if we calculate the probability to leave

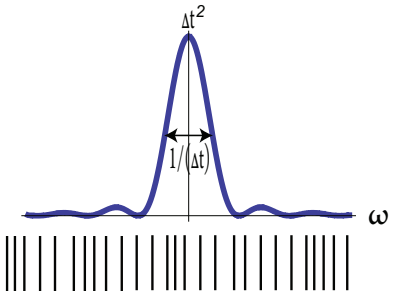

FIG. 4: "Delta function" with dense spectrum of final states $|m\rangle$.

the original state $|n\rangle$ and to go somewhere, it is given by $P_{n\to} = \sum_m P_{n\to m}$. Assuming the matrix element V_{mn} is more or less a constant (more generally V_{mn} depends smoothly on the energy of the final state E_m and does not depend on other quantum numbers of the state $|m\rangle$) we obtain

$$\begin{aligned} P_{n\to} &= \sum_m P_{n\to m} = \sum_m \frac{1}{\hbar^2}\left|V_{mn}\right|^2 \frac{\sin^2(\omega_{mn}\Delta t/2)}{(\omega_{mn}/2)^2} \\ &\approx \frac{1}{\hbar^2}\int dE_m \nu(E_m)\left|V_{mn}\right|^2 2\pi\delta((E_m - E_n)/\hbar)\Delta t \\ &= \frac{2\pi}{\hbar}\left|V(E_n)\right|^2 \nu(E_n)\Delta t \ , \end{aligned} \tag{28}$$

where $\nu(E)$ is the density of states of the final states $|m\rangle$. Then $\nu(E_n)$ is the density of the final states $|m\rangle$ evaluated at $E_m = E_n$. Analogously $V(E_n) \equiv V_{mn}$ with $E_m = E_n$. The probability grows linearly in time. We obtain transition rate $\Gamma = \frac{2\pi}{\hbar} |V(E_n)|^2 \nu(E_n)$. For the dense spectrum one may have a situation when $P_{n\to m} \ll 1$ for all m. Thus the first order perturbation theory is justified. Nevertheless, the total transition probability $P_{n\to} = \sum_m P_{n\to m}$ does not have to be small. Thus the Golden Rule describes the process at all relevant times.

If the matrix elements V_{mn} depend not only of the energy E_m but also on the other quantum numbers, one has to use the standard Golden Rule formula

$$\Gamma_{n\to m} = \frac{2\pi}{\hbar} \, |V_{mn}|^2 \, \delta(E_m - E_n) \; . \tag{29}$$

and substitute it into an integral over the states $|m\rangle$.

G. Harmonic perturbation

$V(t_{\rm initial} < t < t_{\rm final}) = V \, \cos(\omega t)$ and $V(t) = 0$ otherwise.

$$\begin{aligned} P_{n\to m} &\approx \frac{1}{\hbar^2} \left| V_{mn} \int_{t_0}^{t} dt' \cos(\omega t') e^{-i(E_n - E_m)t'/\hbar} \right|^2 \\ &\approx \frac{1}{\hbar^2} \frac{|V_{mn}|^2}{4} \left[\frac{\sin^2\left([\omega_{mn} + \omega]\Delta t/2\right)}{\left([\omega_{mn} + \omega]/2\right)^2} + \frac{\sin^2\left([\omega_{mn} - \omega]\Delta t/2\right)}{\left([\omega_{mn} - \omega]/2\right)^2} \right] \; . \end{aligned} \tag{30}$$

(the mixed terms neglected) We obtain transitions with energy change $\pm\hbar\omega$. The Golden rule result in this case reads

$$\Gamma_{n\to m} = \frac{\pi}{2\hbar} \, |V_{mn}|^2 \, \left(\delta(E_m - E_n + \hbar\omega) + \delta(E_m - E_n - \hbar\omega)\right) . \tag{31}$$

Again it makes sense only if submitted into an integral, i.e., when the states $|m\rangle$ form a kontinuum.

H. Adiabatic approximation

Assume now another situation when a perturbation remains forever (see Fig. 5).

We use again the formula

$$|\Psi_I(t)\rangle = U_I(t, t_0) \; |n\rangle \approx \left(1 - \frac{i}{\hbar} \int_{t_0}^{t} dt' \, V_I(t') \right) \; |n\rangle \;\; , \tag{32}$$

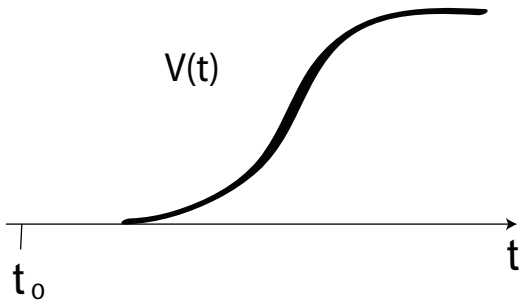

FIG. 5: Perturbation $V(t)$ that does not disappear.

(Recall, $|m\rangle$ are the eigenstates of H_0). The amplitude $a_{n\to m}$ in the expansion of the wave function $|\Psi_I(t)\rangle$ is given by

$$a_{n\to m} = -\frac{i}{\hbar}\int_{t_0}^{t} dt'\, V_{mn}(t')\, e^{i\omega_{mn}t'} \ . \tag{33}$$

We integrate by parts:

$$a_{n\to m} = -\frac{V_{mn}(t')}{\hbar\omega_{mn}}\, e^{i\omega_{mn}t'}\Big|_{t_0}^{t} + \frac{1}{\hbar\omega_{mn}}\int_{t_0}^{t} dt' \left(\frac{\partial V_{mn}(t')}{\partial t'}\right) e^{i\omega_{mn}t'} \ . \tag{34}$$

At t_0 there was no perturbation, $V(t_0) = 0$, thus

$$a_{n\to m} = -\frac{V_{mn}(t)}{\hbar\omega_{mn}}\, e^{i\omega_{mn}t} + \frac{1}{\hbar\omega_{mn}}\int_{t_0}^{t} dt' \left(\frac{\partial V_{mn}(t')}{\partial t'}\right) e^{i\omega_{mn}t'} \ . \tag{35}$$

Let us try to understand the meaning of the first term of (35). Recall the time-independent perturbation theory for Hamiltonian $H = H_0 + V$. The corrected (up to the first order) eigenstate $|\tilde{n}\rangle$ is given by

$$|\tilde{n}\rangle \approx |n\rangle - \sum_{m\neq n} \frac{V_{mn}}{E_m - E_n}\, |m\rangle \tag{36}$$

It is clear that the first term of (35) and the first order corrections in (36) have something to do with each other (are the same). To compare we have to form in both cases the time-dependent Schrödinger wave function. From (35) we obtain the interaction representation wave function $|\Psi_I\rangle = |n\rangle + \sum_{m\neq n} a_{n\to m}(t)\, |m\rangle$, which leads to (neglecting the second term in the RHS of (35))

$$|\Psi_I(t)\rangle = |n\rangle - \sum_{m\neq n} \frac{V_{mn}}{E_m - E_n}\, |m\rangle\, e^{i(E_m - E_n)t/\hbar} \ . \tag{37}$$

In the Schrödinger picture $|\Psi(t)\rangle = e^{-iH_0 t/\hbar}\,|\Psi_I(t)\rangle$ we obtain

$$|\Psi(t)\rangle = |n\rangle\, e^{-iE_n t/\hbar} - \sum_{m\neq n} \frac{V_{mn}}{E_m - E_n}\,|m\rangle\, e^{-iE_n t/\hbar} \tag{38}$$

The same we get from (36) (upon neglecting corrections to the energy E_n).

Thus, the first term of (35) corresponds to the state "adjusting" itself to the new Hamiltonian, i.e., remaining the eigenstate of the corrected Hamiltonian $H = H_0 + V$. The "real" transitions an only be related to the second term of (35).

Idea of the adiabatic approximation: $\partial V_{mn}(t)/\partial t$ is small and moreover multiplied in the integral of (35) by an oscillating exponent. This integral can be neglected. Thus, no real transition will happen, and, the state will remain the eigenstate of a slowly changing Hamiltonian.

I. Instantaneous frame

New idea: follow the eigenstates of the changing Hamiltonian $H(t)$. It is convenient to introduce a vector of parameters $\vec{R}$ upon which the Hamiltonian depends and which change in time. $H(t) = H(\vec{R}(t))$. We no longer consider a situation when only a small perturbation is time-dependent. Diagonalize the Hamiltonian for each $\vec{R}$. Introduce instantaneous eigenstates $\left|n(\vec{R})\right\rangle$, such that $H(\vec{R})\left|n(\vec{R})\right\rangle = E_n(\vec{R})\left|n(\vec{R})\right\rangle$. Since $H(\vec{R})$ changes continuously, it is reasonable to assume that $\left|n(\vec{R})\right\rangle$ do so as well. Introduce a unitary transformation

$$\Omega(t,t_0) \equiv \sum_n \left|n(\vec{R}(t_0))\right\rangle \left\langle n(\vec{R}(t))\right| = \sum_n |n_0\rangle \left\langle n(\vec{R}(t))\right| \ . \tag{39}$$

For brevity $|n_0\rangle \equiv \left|n(\vec{R}(t_0))\right\rangle$. Idea: if $|\Psi(t)\rangle \propto \left|n(\vec{R}(t))\right\rangle$, i.e., follows adiabatically, the new wave function: $|\Phi(t)\rangle = \Omega(t,t_0)\,|\Psi(t)\rangle \propto |n_0\rangle$ does not change at all. In other words in the adiabatic approximation the time evolution operator would be approximated as $U(t,t_0) \approx [\Omega(t,t_0)]^{-1}$. Let's find the Hamiltonian governing the time evolution of $|\Phi(t)\rangle$:

$$i\hbar\left|\dot{\Phi}\right\rangle = i\hbar\Omega\left|\dot{\Psi}\right\rangle + i\hbar\dot{\Omega}\,|\Psi\rangle = \Omega H(t)\,|\Psi\rangle + i\hbar\dot{\Omega}\,|\Psi\rangle = \left[\Omega H\Omega^{-1} + i\hbar\dot{\Omega}\Omega^{-1}\right]|\Phi\rangle \tag{40}$$

Thus the new Hamiltonian is given by

$$\tilde{H} = \Omega H\Omega^{-1} + i\hbar\dot{\Omega}\Omega^{-1} \tag{41}$$

The first term is diagonal. Indeed

$$\Omega(t,t_0)H(t)[\Omega(t,t_0)]^{-1} = \sum_{nm} |n_0\rangle \left\langle n(\vec{R}(t))\right| H(\vec{R}(t)) \left|m(\vec{R}(t))\right\rangle \langle m_0| = \sum_n E_n(\vec{R}(t)) |n_0\rangle \langle n_0| \ . \quad (42)$$

Thus transitions can happen only due to the second term which is proportional to the time derivative of Ω, i.e., it is small for slowly changing Hamiltonian.

J. Geometric Phase

We can treat the Hamiltonian (41) perturbatively. $\tilde{H}(t) = H_0(t) + V(t)$, where

$$H_0(t) = \sum_n E_n(\vec{R}(t)) |n_0\rangle \langle n_0| \ , \quad (43)$$

and

$$V(t) = i\hbar\dot{\Omega}\Omega^{-1} = i\hbar \sum_{n,m} |n_0\rangle \left\langle \dot{n}(\vec{R}(t))\right| m(\vec{R}(t))\rangle \langle m_0| \ . \quad (44)$$

The operator (44) may have diagonal and off-diagonal elements. The latter will be responsible for transitions and will be discussed later. Here we discuss the role of the diagonal elements, e.g.,

$$V_{nn}(t) = \langle n_0| \ i\hbar\dot{\Omega}\Omega^{-1} \ |n_0\rangle = i\hbar \left\langle \dot{n}(\vec{R}(t))\right| n(\vec{R}(t))\rangle = i\hbar \, \dot{\vec{R}} \left\langle \vec{\nabla}_R \ n(\vec{R})\right| n(\vec{R})\rangle \ . \quad (45)$$

Here $\vec{\nabla}_R$ is the gradient in the space of parameters.

This (real) quantity serves as an addition to energy $E_n(\vec{R})$, i.e., $E_n(R(t)) \to E_n(\vec{R}(t)) + V_{nn}(t)$. Thus, state evolves in time as

$$|n_0\rangle \, e^{-(i/\hbar)\int_{t_0}^{t} dt' \left[E_n(\vec{R}(t'))+V_{nn}(t'))\right]} = |n_0\rangle \, e^{-i\phi_{D,n}} e^{-i\phi_{B,n}} \ . \quad (46)$$

The so called dynamical phase is defined as

$$\phi_{D,n} = \frac{1}{\hbar}\int_{t_0}^{t} dt' E_n(\vec{R}(t')) \ . \quad (47)$$

The additional Berry phase or geometric phase is defined as

$$\begin{aligned}\phi_{B,n} &= \frac{1}{\hbar}\int_{t_0}^{t} dt' V_{nn}(t') = i\int_{t_0}^{t} dt'\,\dot{\vec{R}}\cdot\left\langle\vec{\nabla}_R\, n(\vec{R})\right|\left. n(\vec{R})\right\rangle \\ &= i\int_{\vec{R}(t_0)}^{\vec{R}(t)} d\vec{R}\cdot\left\langle\vec{\nabla}_R\, n(\vec{R})\right|\left. n(\vec{R})\right\rangle\ . \end{aligned} \tag{48}$$

The last expression is completely geometric.

The Berry phase is well defined only for closed path, i.e., when the Hamiltonian returns to itself. Indeed the choice of the basis $\left|n(\vec{R})\right\rangle$ is arbitrary up to a phase. Instead of $\left|n(\vec{R})\right\rangle$ we could have chosen $e^{i\chi(\vec{R})}\left|n(\vec{R})\right\rangle$. Instead of (45) we would then obtain

$$\begin{aligned}V_{nn}(t) &= i\hbar\left\langle\dot{n}(\vec{R}(t))\right|\left. n(\vec{R}(t))\right\rangle + \hbar\dot{\chi}(\vec{R}(t)) \\ &= i\hbar\,\dot{\vec{R}}\cdot\left\langle\vec{\nabla}_R\, n(\vec{R})\right|\left. n(\vec{R})\right\rangle + \hbar\dot{\vec{R}}\cdot\vec{\nabla}_R\,\chi(\vec{R}(t))\ . \end{aligned} \tag{49}$$

Instead of (48) we would obtain

$$\phi_{B,n} = \frac{1}{\hbar}\int_{t_0}^{t} dt' V_{nn}(t') = i\int_{\vec{R}(t_0)}^{\vec{R}(t)} d\vec{R}\cdot\left\langle\vec{\nabla}_R\, n(\vec{R})\right|\left. n(\vec{R})\right\rangle + i\int_{\vec{R}(t_0)}^{\vec{R}(t)} d\vec{R}\cdot\vec{\nabla}_R\,\chi(\vec{R})\ . \tag{50}$$

Thus the Berry phase is, in general, not gauge invariant. For a closed path we must choose the basis $\left|n(\vec{R})\right\rangle$ so that it returns to itself. That is $\left|n(\vec{R})\right\rangle$ depends only on the parameters $\vec{R}$ and not on the path along which $\vec{R}$ has been arrived. This means $\chi(\vec{R}(t_0)) = \chi(\vec{R}(t_0+T))$, where T is the traverse time of the closed contour. In this case the integral of $\vec{\nabla}_R\,\chi$ vanishes and we are left with

$$\phi_{B,n} = i\int dt\left\langle\dot{n}(\vec{R}(t))\right|\left. n(\vec{R}(t))\right\rangle = i\int d\vec{R}\left\langle\vec{\nabla}_R\, n\right|\left. n\right\rangle \tag{51}$$

This is Berry's phase. It is also a geometric phase. Physical meaning only for superpositions of different eigenstates.

K. Non-adiabatic processes: transitions

H_0 is diagonal, but time dependent. Interaction representation is simple to generalize:

$$|\Phi(t)\rangle \equiv e^{-(i/\hbar)\int_{t_0}^{t} dt' H_0(t')}\,|\Phi_I(t)\rangle \tag{52}$$

and

$$i\hbar\frac{d}{dt}\left|\Phi_I\right\rangle = V_I(t)\left|\Phi_I\right\rangle \ , \tag{53}$$

where

$$V_I(t) \equiv e^{i\int\limits_{t_0}^{t} dt' H_0(t')/\hbar} V(t) e^{-i\int\limits_{t_0}^{t} dt' H_0(t')/\hbar} \tag{54}$$

($H_0(t)$ commutes with itself at different times).

For the transition probability (of the first order) this gives

$$\begin{aligned} P_{n\to m} &\approx \frac{1}{\hbar^2}\left| i\hbar\int\limits_{t_0}^{t} dt' \left\langle \dot{m}(\vec{R}(t'))\right| n(\vec{R}(t'))\rangle\, e^{-i\int\limits_{t_0}^{t'} dt''\left[E_n(\vec{R}(t''))-E_m(\vec{R}(t''))\right]/\hbar}\right|^2 \\ &= \left|\int\limits_{t_0}^{t} dt' \left\langle \dot{m}(\vec{R}(t'))\right| n(\vec{R}(t'))\rangle\, e^{-i\int\limits_{t_0}^{t'} dt''\left[E_n(\vec{R}(t''))-E_m(\vec{R}(t''))\right]/\hbar}\right|^2 . \end{aligned} \tag{55}$$

L. Landau-Zener transition

$$H(t) = -\frac{1}{2}\left(\epsilon(t)\sigma_z + \Delta\,\sigma_x\right) \tag{56}$$

Parameter $R(t) = \epsilon(t) = \alpha t$.

The eigenstates (dependent on R)

$$\begin{aligned} \left|0(R=\epsilon)\right\rangle &= \cos(\eta/2)\left|\uparrow\right\rangle + \sin(\eta/2)\left|\downarrow\right\rangle \\ \left|1(R=\epsilon)\right\rangle &= -\sin(\eta/2)\left|\uparrow\right\rangle + \cos(\eta/2)\left|\downarrow\right\rangle \ , \end{aligned} \tag{57}$$

where $\tan\eta \equiv \Delta/\epsilon$. We find also the eigenenergies (dependent on $R = \epsilon$) $E_{0/1} = \mp(1/2)\sqrt{\epsilon^2+\Delta^2}$.

Assume that at $t \to -\infty$ the system was in the state $\left|0(R=\epsilon=-\infty)\right\rangle$. What is the probability that at $t\to\infty$ a transition will happen to the state $\left|1(R=\epsilon=\infty)\right\rangle$?

From Eq. (55) we obtain

$$P_{0\to 1} \approx \left|\int\limits_{-\infty}^{\infty} dt' \left\langle \dot{1}\right| 0\rangle\, e^{(i/\hbar)\int\limits_{0}^{t'} dt''\sqrt{\Delta^2+\alpha^2 t''^2}}\right|^2 . \tag{58}$$

(The lower limit of integration in the exponent was changed to 0, this just adds a constant phase which does not change the probability).

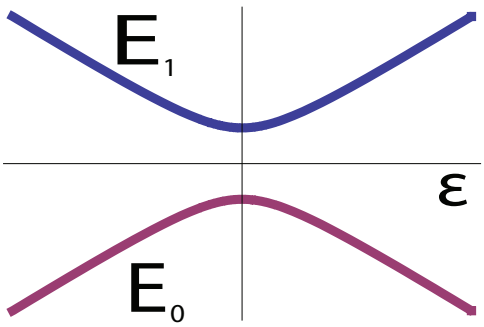

FIG. 6: Energy levels $E_0(\epsilon)$ and $E_1(\epsilon)$.

We obtain $\langle \dot{1} | 0 \rangle = -(1/2)\dot{\eta}$. From $\cot\eta = (\alpha t/\Delta)$ we obtain $\dot{\eta} = -\sin^2\eta(\alpha/\Delta) = -\frac{\Delta\alpha}{\Delta^2+\alpha^2t^2}$. For the transition probability this gives

$$P_{0\to 1} \approx \left| \frac{1}{2} \int_{-\infty}^{\infty} dt' \frac{\Delta\alpha}{\Delta^2+\alpha^2 t'^2} \, e^{(i/\hbar)\int_0^{t'} dt''\sqrt{\Delta^2+\alpha^2 t''^2}} \right|^2 . \tag{59}$$

Introducing a new (dimensionless) time $\tau \equiv \alpha t/\Delta$ we obtain

$$P_{0\to 1} \approx \left| \frac{1}{2} \int_{-\infty}^{\infty} d\tau' \frac{1}{1+\tau'^2} \, e^{i\gamma\int_0^{\tau'} d\tau''\sqrt{1+\tau''^2}} \right|^2 , \tag{60}$$

where $\gamma \equiv \Delta^2/(\hbar\alpha)$. We see that the result depends only on γ.

M. Berry has calculated this integral and got

$$P_{0\to 1} \approx \left| (\pi/3)\, e^{-\frac{\pi\gamma}{4}} \right|^2 = (\pi^2/9)\, e^{-\frac{\pi\gamma}{2}} . \tag{61}$$

Problem: the result is non-analytic in α which characterizes the slowness of the change of the Hamiltonian. Thus the logic of our perturbative expansion does not work. The result is actually exponentially small for small enough α and not only small because it is proportional to α. It cannot be excluded that higher order contributions will give also exponentially small results.

The exact result is known (Landau 1932, Zener 1932, Stückelberg 1932, Majorana 1932). It reads: $P_{0\to 1} = e^{-\frac{\pi\gamma}{2}}$ for arbitrary γ.

II. RELATIVISTIC QUANTUM MECHANICS: DIRAC EQUATION

A. Relativistic mechanics

1. 4-vectors

Space time $x^\mu = (x^0, x^1, x^2, x^3) = (ct, \vec{r})$ (contravariant vector). Metric:

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{62}$$

Covariant vector: $x_\mu = g_{\mu\nu}x^\nu = (ct, -\vec{r})$. The same with arbitrary vector $a^\mu = (a^0, \vec{a})$. Scalar product: $a_\mu b^\mu = g_{\mu\nu}a^\nu b^\mu = a^0 b^0 - \vec{a}\cdot\vec{b}$. Norm of a vector $a_\mu a^\mu$.

If $a_\mu a^\mu < 0$ - space-like (raumartiger) vector

If $a_\mu a^\mu > 0$ - time-like (zeitartiger) vector

If $a_\mu a^\mu = 0$ - vector on the light-cone (Lichtkegel).

The time-like vectors can point into the future: $a^0 > 0$ or into the past $a^0 < 0$.

Gradient operators $\partial_\mu \equiv \partial/\partial x^\mu = ((1/c)\partial/\partial t, \vec{\nabla})$. As the index suggests ∂_μ is a covariant vector. Indeed $\partial_\mu x^\mu = 4$, i.e., a scalar. Respectively $\partial^\mu \equiv \partial/\partial x_\mu = ((1/c)\partial/\partial t, -\vec{\nabla})$.

2. E-m field

$$A^\mu = (A^0, \vec{A}) = (\Phi, \vec{A}) \quad ; \quad A_\mu = (A^0, -\vec{A}) = (\Phi, -\vec{A}) \tag{63}$$

Tensor of e-m field $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{64}$$

Covariant derivative: $D_\mu \equiv \partial_\mu + \frac{iq}{\hbar c}A_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t} + \frac{iq}{\hbar c}\Phi, \vec{\nabla} - \frac{iq}{\hbar c}\vec{A}\right)$. Here q is the charge of the particle. If it is an electron, then $q = -e$, where $e > 0$.

3. Lorentz Group, Lorentz Transformations

Transformations which conserve the norm (interval) $x^\mu x_\mu$. Homogeneous transformation (no translation):

$$x'^\mu = \Omega^\mu{}_\nu x^\nu \ . \tag{65}$$

Elements of $\Omega^\mu{}_\nu$ must be real (coordinates are real). Using $x'_\mu = g_{\mu\alpha}x'^\alpha$ we can obtain the transformation law for the covariant components

$$x'_\mu = g_{\mu\alpha}x'^\alpha = g_{\mu\alpha}\Omega^\alpha{}_\beta x^\beta = g_{\mu\alpha}\Omega^\alpha{}_\beta g^{\beta\nu}x_\nu \ . \tag{66}$$

This equation can be also written as

$$x'_\mu = \Omega_\mu{}^\nu x_\nu \quad , \quad \Omega_\mu{}^\nu \equiv g_{\mu\alpha}\Omega^\alpha{}_\beta g^{\beta\nu} \ . \tag{67}$$

From $x'^\mu x'_\mu = x^\mu x_\mu$ (Lorentz transformations conserve the norm) we obtain

$$x'^\mu x'_\mu = \Omega^\mu{}_\gamma x^\gamma g_{\mu\alpha}\Omega^\alpha{}_\beta g^{\beta\nu}x_\nu \tag{68}$$

Thus the conservation of the norm requires

$$\Omega^\mu{}_\gamma g_{\mu\alpha}\Omega^\alpha{}_\beta g^{\beta\nu} = \delta_\gamma{}^\nu \ . \tag{69}$$

In matrix notation this gives $\Omega^T g\Omega = g$, where we have multiplied by g from the right. This is the condition Ω must satisfy. In particular this means that $\mathrm{Det}\,[\Omega] = \pm 1$.

By Ω we mean $\Omega^\mu{}_\nu$ and not $\Omega_\mu{}^\nu$, which is a different object. For clarity we will denote $\Omega_\mu{}^\nu$ by $\tilde{\Omega}$. In matrix notation $\tilde{\Omega} = g\Omega g^{-1}$. Then Eq. (69) reads $\Omega^\mu{}_\gamma\Omega_\mu{}^\nu = \delta_\gamma{}^\nu$, i.e., in the matrix notation $\Omega^T\tilde{\Omega} = 1$. Thus we obtain also $\tilde{\Omega} = (\Omega^T)^{-1} = (\Omega^{-1})^T$. This allows us to invert the basic relation $x'^\mu = \Omega^\mu{}_\nu x^\nu$ as follows. We write (with no indices) $x' = \Omega x$. Then $x = \Omega^{-1}x'$. Then $x^T = x'(\Omega^{-1})^T = x'(\Omega^T)^{-1} = x'\tilde{\Omega}$. With indices this reads

$$x^\mu = x'^\nu\Omega_\nu{}^\mu \ . \tag{70}$$

Lorentz group - all such transformations.

There are certain important sub-groups:

$\Omega^0{}_0 > 0$ (no change of direction of time for time-like vectors) - orthochronous (orthochrone Lorentz-Gruppe).

$\mathrm{Det}\left[\Omega^\mu{}_\nu\right] = 1$ - proper L.G. (eigentliche L.G.).

$\Omega^0{}_0 > 0$ and Det $\left[\Omega^\mu{}_\nu\right] = 1$ - proper orthochronous L.G. (eigentliche orthochrone Lorentz-Gruppe).

Examples:

Lorentz Boost characterised by velocity $v < c$ of one coordinate system relative to the other:

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} , \tag{71}$$

where $\tanh\eta = v/c$. There are 3 such boosts.

Rotation by angle θ:

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} , \tag{72}$$

There are 3 such rotations.

4. Energy-momentum

In relativistic mechanics the energy and the momentum form a 4-vector. Namely: $p^\mu = (E/c, \vec{p})$, where $E = \sqrt{m^2c^4 + c^2(\vec{p})^2}$. The norm is given by $p^\mu p_\mu = (mc)^2$.

The Hamiltonian of a charged particle in a given e-m field is given by

$$H = c\sqrt{(mc)^2 + \left(\vec{p} - \frac{q}{c}\vec{A}\right)^2} + q\Phi \ . \tag{73}$$

B. Klein-Gordon equation

As in non-relativistic mechanics we substitute $\vec{p} \to -i\hbar\vec{\nabla}$ and $E \to i\hbar\partial/\partial t$. In 4-vector "language" this means $p^\mu = (E/c, \vec{p}) \to i\hbar\partial^\mu = ((i\hbar/c)\partial/\partial t, -i\hbar\vec{\nabla})$.

Using $(E - q\Phi)^2 - c^2\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2 = (mc^2)^2$ we obtain the Klein-Gordon equation

$$\left[\left(i\hbar\frac{\partial}{\partial t} - q\Phi\right)^2 - c^2\left(-i\hbar\vec{\nabla} - \frac{q}{c}\vec{A}\right)^2\right]\Psi = (mc^2)^2\,\Psi \ , \tag{74}$$

where Ψ is the scalar (one component) wave function. In terms of the covariant derivatives D_μ this can be rewritten as

$$\left[\hbar^2 D_\mu D^\mu + (mc)^2\right]\Psi = 0 \tag{75}$$

The Lorentz invariance of this equation is then self-evident.

This is the 2-order differential equation. To specify the state of the system at all times we have to know Ψ and $\partial\Psi/\partial t$ at initial time. Thus Ψ alone does not describe the state of the system: BAD.

We investigate the case of zero e-m field. Then the equation reads

$$\left[\hbar^2 \partial_\mu \partial^\mu + (mc)^2\right]\Psi = 0 \tag{76}$$

Solutions: plane waves $\Psi \propto e^{-(i/\hbar)(Et-\vec{p}\vec{r})}$ with $E = \pm c\sqrt{(mc)^2 + (\vec{p})^2}$. There are negative energy solutions: BAD but we will have the same in Dirac theory and it will at the end be its triumph.

1. Density and current

The complex conjugated wave function Ψ^* also satisfies

$$\left[\hbar^2 \partial_\mu \partial^\mu + (mc)^2\right]\Psi^* = 0 \tag{77}$$

We then obtain

$$\Psi^*(\partial_\mu \partial^\mu \Psi) - (\partial_\mu \partial^\mu \Psi^*)\Psi = 0 \tag{78}$$

and

$$\partial_\mu \left[\Psi^*(\partial^\mu \Psi) - (\partial^\mu \Psi^*)\Psi\right] = 0 \tag{79}$$

Thus if we introduce the 4-current

$$j^\mu = \frac{i\hbar}{2m}\left[\Psi^*(\partial^\mu \Psi) - (\partial^\mu \Psi^*)\Psi\right] \tag{80}$$

it satisfies the continuity equation $\partial_\mu j^\mu = 0$.

The components of j^μ are given by $j^\mu = (c\rho, \vec{j})$, where

$$c\rho = \frac{i\hbar}{2mc}\left[\Psi^* \frac{\partial \Psi}{\partial t} - \frac{\partial \Psi^*}{\partial t}\Psi\right] \tag{81}$$

and

$$\vec{j} = -\frac{i\hbar}{2m}\left[\Psi^*(\vec{\nabla}\Psi) - (\vec{\nabla}\Psi^*)\Psi\right] \tag{82}$$

The continuity equation then reads $\partial\rho/\partial t + \vec{\nabla}\vec{j} = 0$. The continuity equation is GOOD but the density ρ can be negative - BAD. We need another equation.

Exercise: show that with the e-m field the 4-current density satisfying the continuity equation $\partial_\mu j^\mu = 0$ is given by

$$j^\mu = \frac{i\hbar}{2m}\left[\Psi^*(D^\mu\Psi) - (D^\mu\Psi)^*\Psi\right] = \frac{i\hbar}{2m}\left[\Psi^*(\partial^\mu\Psi) - (\partial^\mu\Psi^*)\Psi\right] - \frac{q}{mc}A^\mu\Psi^*\Psi \tag{83}$$

C. Dirac equation

The idea is to have a first order equation which could be written as Schrödinger equation $i\hbar\partial\Psi/\partial t = H_D\Psi$. For this equation to be Lorentz-invariant H_D must be linear in $\vec{p}$. Thus

$$H_D = c\vec{\alpha}\cdot\vec{p} + \beta mc^2 \ , \tag{84}$$

where $\vec{\alpha}$ and β are some objects independent of time and cordinates. The Schrödinger equation then reads

$$(p_0 - \vec{\alpha}\cdot\vec{p} - \beta mc)\Psi = 0 \ , \tag{85}$$

where $p_0 = p^0 = E/c = (i\hbar/c)\partial/\partial t$.

On the other hand we still want Ψ to satisfy the Klein-Gordon equation

$$(p_0^2 - (\vec{p})^2 - (mc)^2)\Psi = 0 \ , \tag{86}$$

since it just expresses the basic relativistic relation. We multiply (85) by $(p_0 + \vec{\alpha}\cdot\vec{p} + \beta mc)$ from the left

$$(p_0 + \vec{\alpha}\cdot\vec{p} + \beta mc)(p_0 - \vec{\alpha}\cdot\vec{p} - \beta mc)\Psi = 0 \ , \tag{87}$$

and want this to be equivalent to the Klein-Gordon equation. In other words we want to factorize the operator $p_0^2 - (\vec{p})^2 - (mc)^2$ as

$$p_0^2 - (\vec{p})^2 - (mc)^2 = (p_0 + \vec{\alpha}\cdot\vec{p} + \beta mc)(p_0 - \vec{\alpha}\cdot\vec{p} - \beta mc) \tag{88}$$

Now we should determine what $\vec{\alpha}$ and β should be. The operators p_0 and $\vec{p}$ commute with each other and the components of $\vec{p}$ commute as well. If also $\vec{\alpha}$ and β would be commuting numbers we would not get the Klein-Gordon equation. Let's assume they do not commute

(they are matrices). Then

$$
\begin{aligned}
&(p_0 + \vec{\alpha}\cdot\vec{p} + \beta mc)(p_0 - \vec{\alpha}\cdot\vec{p} - \beta mc) \\
&= \big[p_0^2 - \sum_k (\alpha_k)^2 p_k^2 - \beta^2 m^2 c^2 \\
&- \sum_{k<l} (\alpha_k\alpha_l + \alpha_l\alpha_k) p_k p_l - \sum_k (\alpha_k\beta + \beta\alpha_k) mc p_k\big]
\end{aligned} \tag{89}
$$

Thus to obtain the Klein-Gordon equation we must require that

$$(\alpha_k)^2 = 1 \quad , \quad \beta^2 = 1 \ , \tag{90}$$

$$\alpha_k\alpha_l + \alpha_l\alpha_k = 0 \quad (\text{for } k \neq l) \ , \tag{91}$$

$$\alpha_k\beta + \beta\alpha_k = 0 \tag{92}$$

Dirac found that $\vec{\alpha}$ and β must be at least 4×4 matrices. He suggested

$$\vec{\alpha} = \begin{pmatrix} \hat{0} & \vec{\sigma} \\ \vec{\sigma} & \hat{0} \end{pmatrix} \qquad \beta = \begin{pmatrix} \hat{1} & \hat{0} \\ \hat{0} & -\hat{1} \end{pmatrix} \tag{93}$$

where $\vec{\sigma}$ is the vector of Pauli matrices:

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{94}$$

This is not the only possibility. We will not consider the other ones, since they are equivalent.

This means that the wave function Ψ has 4 components. That is

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix} = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \ , \tag{95}$$

where

$$\varphi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} \quad , \quad \chi = \begin{pmatrix} \Psi_3 \\ \Psi_4 \end{pmatrix} \tag{96}$$

!!!! Note, the Ψ is not a 4-vector.

The Dirac equation reads

$$i\hbar\frac{\partial\Psi}{\partial t} = c(\vec{\alpha}\cdot\vec{p})\Psi + mc^2\beta\Psi = -i\hbar c(\vec{\alpha}\cdot\vec{\nabla})\Psi + mc^2\beta\Psi \ , \tag{97}$$

or

$$i\hbar\frac{\partial\varphi}{\partial t} = c(\vec{\sigma}\cdot\vec{p})\chi + mc^2\varphi \ ,$$
$$i\hbar\frac{\partial\chi}{\partial t} = c(\vec{\sigma}\cdot\vec{p})\varphi - mc^2\chi \ . \tag{98}$$

This form is convenient for the non-relativistic limit.

1. Probability density, continuity equation

Let us choose $\rho = \Psi^\dagger\Psi = |\Psi_1|^2 + |\Psi_2|^2 + |\Psi_3|^2 + |\Psi_4|^2$, where $\Psi^\dagger = (\Psi_1^*, \Psi_2^*, \Psi_3^*, \Psi_4^*)$.

Then

$$i\hbar\frac{\partial\rho}{\partial t} = i\hbar\left(\frac{\partial\Psi^\dagger}{\partial t}\,\Psi + \Psi^\dagger\,\frac{\partial\Psi}{\partial t}\right) \tag{99}$$

The conjugated Dirac equation reads (note that $\vec{\alpha}$ are hermitian matrices)

$$-i\hbar\frac{\partial\Psi^\dagger}{\partial t} = i\hbar c(\vec{\nabla}\Psi^\dagger\cdot\vec{\alpha}) + mc^2\Psi^\dagger\beta \ , \tag{100}$$

This gives

$$\begin{aligned} i\hbar\frac{\partial\rho}{\partial t} &= i\hbar c\left(-(\vec{\nabla}\Psi^\dagger\cdot\vec{\alpha})\Psi - \Psi^\dagger(\vec{\alpha}\cdot\vec{\nabla}\Psi)\right) \\ &= -i\hbar c\vec{\nabla}\cdot(\Psi^\dagger\vec{\alpha}\Psi) \end{aligned} \tag{101}$$

From here we conclude

$$\vec{j} = c\Psi^\dagger\vec{\alpha}\Psi \tag{102}$$

No derivatives in $\vec{j}$ and most importantly ρ is positive. GOOD.

4-current is then given by

$$j^\mu = (c\rho, \vec{j}) = (c\Psi^\dagger\Psi, c\Psi^\dagger\vec{\alpha}\Psi) \tag{103}$$

D. Dirac equation with e-m field

$$i\hbar\frac{\partial\Psi}{\partial t} - q\Phi\Psi = c\vec{\alpha}\cdot\left(\vec{p} - \frac{q}{c}\vec{A}\right)\Psi + mc^2\beta\Psi \ , \tag{104}$$

Exercise: show that even with e-m field $\vec{j} = c\Psi^\dagger\vec{\alpha}\Psi$.

E. Covariant form

We multiply the Dirac equation by β from the left. This gives

$$i\hbar\beta\frac{\partial\Psi}{\partial t} = c\beta(\vec{\alpha}\cdot\vec{p})\Psi + mc^2\Psi \ , \tag{105}$$

We introduce: $\gamma^0 = \beta$, $\vec{\gamma} = \beta\vec{\alpha}$, and $\gamma^\mu = (\gamma^0, \vec{\gamma})$

$$\vec{\gamma} = \beta\vec{\alpha} = \begin{pmatrix} \hat{0} & \vec{\sigma} \\ -\vec{\sigma} & \hat{0} \end{pmatrix} \qquad \gamma^0 = \beta = \begin{pmatrix} \hat{1} & \hat{0} \\ \hat{0} & -\hat{1} \end{pmatrix} \tag{106}$$

Since $p_0 = (i\hbar/c)\partial/\partial t$ and $p_\mu = (p_0, -\vec{p})$, the Dirac equation reads

$$(\gamma^\mu p_\mu - mc)\Psi = (i\hbar\gamma^\mu\partial_\mu - mc)\Psi = 0 \tag{107}$$

With the e-m field it is

$$\left(\gamma^\mu\left(p_\mu - \frac{q}{c}A_\mu\right) - mc\right)\Psi = (i\hbar\gamma^\mu D_\mu - mc)\,\Psi = 0 \tag{108}$$

F. Properties of γ^μ

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \tag{109}$$

$$(\gamma^0)^\dagger = \gamma^0 \quad , \quad (\gamma^k)^\dagger = -\gamma^k \tag{110}$$

1. 4-current

4-current is given by

$$j^\mu = (c\rho, \vec{j}) = (c\Psi^\dagger\Psi, c\Psi^\dagger\vec{\alpha}\Psi) = (c\Psi^\dagger\gamma^0\gamma^0\Psi, c\Psi^\dagger\gamma^0\gamma^0\vec{\alpha}\Psi) = c\bar{\Psi}\gamma^\mu\Psi \ , \tag{111}$$

where $\bar{\Psi} \equiv \Psi^\dagger\gamma^0 = (\Psi_1^*, \Psi_2^*, -\Psi_3^*, -\Psi_4^*)$.

G. Form-invariance of Dirac equation under Lorentz transformations

We introduced γ^μ as if it was a 4-vector (of matrices). Here we clarify the issue. Consider a Lorentz transformation of the system of coordinates

$$x'^\mu = \Omega^\mu{}_\nu x^\nu \quad , \quad x^\mu = x'^\nu \Omega_\nu{}^\mu \tag{112}$$

We will also use a shorter form $x' = \Omega x$, and $x = \Omega^{-1} x'$.

The covariant derivative transforms as

$$D_\mu(x) = D'_\nu(x')\Omega^\nu{}_\mu \tag{113}$$

which actually means the following $\partial_\mu = \partial'_\nu \Omega^\nu{}_\mu$ and $A_\mu(x) = A'_\nu(x')\Omega^\nu{}_\mu$.

Substituting these all into the Dirac equation

$$\left(i\hbar\gamma^\mu D_\mu(x) - mc\right)\Psi(x) = 0 \tag{114}$$

we obtain

$$\left(i\hbar\gamma^\mu D'_\nu(x')\Omega^\nu{}_\mu - mc\right)\Psi(\Omega^{-1}x') = 0 \tag{115}$$

or

$$\left(i\hbar\gamma'^\nu D'_\nu(x') - mc\right)\Psi(\Omega^{-1}x') = 0 \tag{116}$$

where $\gamma'^\nu \equiv \Omega^\nu{}_\mu \gamma^\mu$.

One possibility would be to define $\Psi'(x') = \Psi(\Omega^{-1}x')$, i.e., to say that Ψ is a set of 4 scalars. Then we would have the equation

$$\left(i\hbar\gamma'^\nu D'_\nu(x') - mc\right)\Psi'(x') = 0 \tag{117}$$

It is not a good choice since γ'^μ are no longer unitary and look completely different, even though they still satisfy $\gamma'^\mu\gamma'^\nu + \gamma'^\nu\gamma'^\mu = 2g^{\mu\nu}$. Moreover the equation written in components would have a different form. We do something different instead:

We look for a matrix $S(\Omega)$ such that

$$\gamma'^\nu = \Omega^\nu{}_\mu \gamma^\mu = S^{-1}\gamma^\nu S \tag{118}$$

That is for every Lorentz transformation Ω we want to find a corresponding $S(\Omega)$.

Then, defining $\Psi'(x') = S(\Omega)\Psi(\Omega^{-1}x')$ and multiplying the Dirac equation by $S(\Omega)$ from the left we obtain

$$\left(i\hbar\gamma^\nu D'_\nu(x') - mc\right)\Psi'(x') = 0 \tag{119}$$

That is the equation is of the same form.

Matrix $S(\Omega)$, thus, defines the transformation of the wave function. It gives a different representation of the Lorentz group. More precisely it is a two-fold representation. Indeed, from Eq. (118) we observe that is $S(\Omega)$ solves (118) then also $-S(\Omega)$ does. This is a special and important property of the spin representation of the Lorentz group. The wave function is not a 4-vector but a "spinor" or "4-spinor" to differentiate from the non-relativistic "2-spinors".

1. Examples

For a boost in the 01 "plane" with velocity $v < c$, i.e.

$$\Omega = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{120}$$

where $\tanh\eta = v/c$, one finds

$$S(\Omega) = \exp\left[-\frac{1}{2}\alpha_1\eta\right] . \tag{121}$$

Analogously for other boosts.

For a rotation around the axis 3, i.e., in 12 plane

$$\Omega = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{122}$$

$$S(\Omega) = \exp\left[\frac{i}{2}\Sigma_3\theta\right] , \tag{123}$$

where

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & \hat{0} \\ \hat{0} & \vec{\sigma} \end{pmatrix} \tag{124}$$

2. Infinitesimal transformations

The correspondence between the Lorentz transformations of the 4-vectors and their spin representation is most convenient to discuss for infinitesimal transformations. An infinitesimal Lorentz transformation can be written as $\Omega = 1 + (i/2)\theta_{\mu\nu}\, J^{\mu\nu}$ or with full indexes

$$\Omega^{\alpha}{}_{\beta} = 1 + \frac{i}{2}\,\theta_{\mu\nu}\,[J^{\mu\nu}]^{\alpha}{}_{\beta} \ . \tag{125}$$

Here $\theta_{\mu\nu} = -\theta_{\nu\mu}$, $|\theta_{\mu\nu}| \ll 1$ is an antisymmetric tensor of infinitesimal rotation angles.

There are 6 different angles. The 6 generators $J^{\mu\nu}$ are given by

$$[J^{\mu\nu}]^{\alpha}{}_{\beta} = i\left[g^{\mu\alpha}\delta^{\nu}{}_{\beta} - g^{\nu\alpha}\delta^{\mu}{}_{\beta}\right] \ . \tag{126}$$

There are only 6 different generators because again $J^{\mu\nu} = -J^{\nu\mu}$. The 6 generators are 3 boost generators $K_n \equiv J^{0n}$ $(n = 1, 2, 3)$ and 3 generators of rotations $L_m \equiv (1/2)\epsilon_{mlk}J^{lk}$ $(m, l, k = 1, 2, 3)$. If we have just one non-zero angle, e.g., $\theta_{01} = -\theta_{10} = \eta$ and the rest vanish, the infinitesimal Lorentz transformation is the expansion of $\Omega = e^{i\eta J^{01}} = e^{i\eta K_1}$. It is easy to check that this matrix is exactly the matrix (120).

The algebra of the 6 generators is described by the following commutators:

$$\begin{aligned}
[L_n, L_m] &= i\epsilon_{nmk}L_k \ , \\
[K_n, K_m] &= -i\epsilon_{nmk}L_k \ , \\
[L_n, K_m] &= i\epsilon_{nmk}K_k \ .
\end{aligned} \tag{127}$$

The corresponding spin transformation reads

$$S(\Omega) = 1 + \frac{i}{2}\,\theta_{\nu\mu}\,\tilde{J}^{\mu\nu} \ , \tag{128}$$

where

$$\tilde{J}^{\mu\nu} = \frac{1}{2}\,\sigma^{\mu\nu} = \frac{i}{4}\,[\gamma^{\mu}, \gamma^{\nu}] \ . \tag{129}$$

(We use here also $\sigma^{\mu\nu}$ since these objects are frequently used.) Thus we only replace the generators $J^{\mu\nu} \to \tilde{J}^{\mu\nu}$. We can again define 3 boost generators $\tilde{K}_n \equiv \tilde{J}^{0n}$ $(n = 1, 2, 3)$ and 3 generators of rotations $\tilde{L}_m \equiv (1/2)\epsilon_{mlk}\tilde{J}^{lk}$ $(m, l, k = 1, 2, 3)$. These satisfy the same algebra (127). If again the only non-vanishing angle is $\theta_{01} = -\theta_{10} = \eta$ we obtain $S = e^{i\eta\tilde{J}^{01}} = e^{i\eta\tilde{K}_1}$. With $\sigma^{01} = i\,\gamma^0\gamma^1 = i\,\alpha_1$ we obtain $S = \exp[-(1/2)\alpha_1\eta]$.

Analogously, for a rotation around the z axis $\theta_{12} = -\theta_{21} = \theta$ (and the rest vanish) we obtain $\Omega = e^{i\theta J_{12}} = e^{i\theta L_3}$ and $S(\Omega) = e^{i\theta \tilde{J}^{12}} = e^{i\theta \tilde{L}_3}$. With

$$\tilde{L}_3 = \tilde{J}^{12} = (i/2)\gamma^1\gamma^2 = (i/2)\begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} = \frac{1}{2}\Sigma_3 \tag{130}$$

we obtain $S(\Omega) = e^{(i/2)\theta\Sigma_3}$. The important feature of the spin representation of the Lorentz group is that for $\theta = 2\pi$ we obtain $\Omega = 1$, but $S(\Omega) = -1$.

H. Spin operator, total angular momentum.

Orbital angular momentum

$$\vec{L} = \vec{r} \times \vec{p} = -i\hbar \vec{r} \times \vec{\nabla} \tag{131}$$

does not commute with the Hamiltonian $H_D = c\vec{\alpha} \cdot \vec{p} + \beta mc^2$.

$$\begin{aligned} [\vec{L}, H_D] &= (\vec{r} \times \vec{p})(c\vec{\alpha} \cdot \vec{p} + \beta mc^2) - (c\vec{\alpha} \cdot \vec{p} + \beta mc^2)(\vec{r} \times \vec{p}) \\ &= (\vec{r} \times \vec{p})(c\vec{\alpha} \cdot \vec{p}) - (c\vec{\alpha} \cdot \vec{p})(\vec{r} \times \vec{p}) \\ &= c\epsilon_{ijk} r_j p_k \alpha_m p_m - c\alpha_m p_m \epsilon_{ijk} r_j p_k \end{aligned} \tag{132}$$

In the first term we use $r_j p_m - p_m r_j = i\hbar\delta_{jm}$ and obtain

$$[\vec{L}, H_D] = i\hbar c \epsilon_{ijk} p_k \alpha_j = i\hbar c \vec{\alpha} \times \vec{p} \tag{133}$$

We introduce the spin operator.

$$\vec{S} = \frac{1}{2}\hbar\vec{\Sigma} = \frac{1}{2}\hbar \begin{pmatrix} \vec{\sigma} & \hat{0} \\ \hat{0} & \vec{\sigma} \end{pmatrix} \tag{134}$$

Since $\vec{S}$ obviously commutes with β we get

$$\begin{aligned} [\vec{S}, H_D] &= \frac{1}{2}\hbar c \left[\begin{pmatrix} \vec{\sigma} & \hat{0} \\ \hat{0} & \vec{\sigma} \end{pmatrix} \begin{pmatrix} \hat{0} & \vec{p} \cdot \vec{\sigma} \\ \vec{p} \cdot \vec{\sigma} & \hat{0} \end{pmatrix} - \begin{pmatrix} \hat{0} & \vec{p} \cdot \vec{\sigma} \\ \vec{p} \cdot \vec{\sigma} & \hat{0} \end{pmatrix} \begin{pmatrix} \vec{\sigma} & \hat{0} \\ \hat{0} & \vec{\sigma} \end{pmatrix} \right] \\ &= \frac{1}{2}\hbar c \begin{pmatrix} \hat{0} & [\vec{\sigma}, \vec{p} \cdot \vec{\sigma}] \\ [\vec{\sigma}, \vec{p} \cdot \vec{\sigma}] & \hat{0} \end{pmatrix} = \frac{1}{2}\hbar c \begin{pmatrix} \hat{0} & 2i\vec{p} \times \vec{\sigma} \\ 2i\vec{p} \times \vec{\sigma} & \hat{0} \end{pmatrix} \\ &= -i\hbar c \vec{\alpha} \times \vec{p} \end{aligned} \tag{135}$$

Thus, only the total angular momentum

$$\vec{J} = \vec{L} + \vec{S} \tag{136}$$

is conserved, i.e., $[\vec{J}, H_D] = 0$.

I. Plane waves

Look for solutions $\varphi = \varphi_p e^{-ip_\mu x^\mu/\hbar}$, $\chi = \chi_p e^{-ip_\mu x^\mu/\hbar}$ (recall that $p_\mu = (E/c, -\vec{p})$). Thus

$$(E - mc^2)\varphi_p = c(\vec{\sigma}\cdot\vec{p})\chi_p \ ,$$
$$(E + mc^2)\chi_p = c(\vec{\sigma}\cdot\vec{p})\varphi_p \ . \tag{137}$$

We see that, e.g., φ_p can be chosen arbitrarily. Then from the second equation $\chi_p = c(\vec{\sigma}\cdot\vec{p})\varphi_p/(E+mc^2)$. Substituting this into the first equation we observe that it is satisfied if $E^2 = (mc^2)^2 + c^2(\vec{p})^2$. Thus the general solution looks like

$$\Psi = Ne^{-ip_\mu x^\mu/\hbar}\begin{pmatrix} \varphi_p \\ \\ \frac{c(\vec{\sigma}\cdot\vec{p})}{(E+mc^2)}\,\varphi_p \end{pmatrix} \ , \tag{138}$$

where N is the normalisation constant. This representation is convenient for positive energy, i.e. for $E = +\sqrt{(mc^2)^2 + c^2(\vec{p})^2}$, as for $|\vec{p}| \ll mc$ the lower part is much smaller than the upper. However it can, in principle, be also used for $E = -\sqrt{(mc^2)^2 + c^2(\vec{p})^2}$. If for the normalisation one would require the density of particles to be equal $1/V$, i.e., one particle in volume V, one would use $\rho = \Psi^\dagger\Psi = 1/V$. A simple calculation, assuming ϕ_p is normalised, i.e., $\phi_p^\dagger\phi_p = 1$, would give then $N = \frac{1}{\sqrt{V}}\sqrt{\frac{E+mc^2}{2E}}$. However this is not a Lorentz invariant choice. The correct choice for $E > 0$ is $\bar{\Psi}\Psi = \Psi^\dagger\gamma^0\Psi = 1/V$. This gives $N = \frac{1}{\sqrt{V}}\sqrt{\frac{mc^2+E}{2mc^2}}$.

Alternatively, if we choose χ_p to represent a solution it will be

$$\Psi = Ne^{-ip_\mu x^\mu/\hbar}\begin{pmatrix} \frac{c(\vec{\sigma}\cdot\vec{p})}{(E-mc^2)}\,\chi_p \\ \\ \chi_p \end{pmatrix} \tag{139}$$

with $N = \frac{1}{\sqrt{V}}\sqrt{\frac{mc^2+|E|}{2mc^2}}$. This representation is convenient for negative energies.

It is convenient to choose the basis states to be eigenstates of the helicity $\vec{p}\cdot\vec{\Sigma}$ (this operator commutes with H_D). This means to choose, e.g., φ_p so that $(\vec{\sigma}\cdot\vec{p})\varphi_p = \pm|\vec{p}|\varphi_p$. These are NOT eigenstates of the spin as $\vec{S}$ does not commute with the Hamiltonian.

J. Prediction of antiparticles

Problem of negative energies. Dirac sea. Positrons as holes.

K. Charge conjugation

Assume we have the wave function $\Psi(x) = \Psi(t, \vec{r})$, which solves the Dirac equation in a given field $A_\mu(x)$:

$$\left(\gamma^\mu \left(p_\mu - \frac{q}{c} A_\mu\right) - mc\right) \Psi = 0 \tag{140}$$

We want to find a wave function Ψ^c which satisfies the same equation but with $q \to -q$. Complex conjugating the Dirac equation we obtain

$$\left((\gamma^\mu)^* \left(-p_\mu - \frac{q}{c} A_\mu\right) - mc\right) \Psi^* = \left(-(\gamma^\mu)^* \left(p_\mu + \frac{q}{c} A_\mu\right) - mc\right) \Psi^* = 0 \ . \tag{141}$$

Thus we have the minimal coupling combination $\left(p_\mu + \frac{q}{c} A_\mu\right)$ with the opposite charge $q \to -q$. Assume the desired wave function is given by $\Psi^c = U_C \Psi^*$ with U_C being a (unitary) matrix. Then we obtain

$$\left(-(\gamma^\mu)^* \left(p_\mu + \frac{q}{c} A_\mu\right) - mc\right) U_C^{-1} \Psi^c = 0 \ . \tag{142}$$

We multiply by U_C from the left:

$$\left(-U_C (\gamma^\mu)^* U_C^{-1} \left(p_\mu + \frac{q}{c} A_\mu\right) - mc\right) \Psi^c = 0 \ . \tag{143}$$

Thus our aim would be achieved if we would find U_C such that

$$U_C (\gamma^\mu)^* U_C^{-1} = -\gamma^\mu \ . \tag{144}$$

It is easy to see that $U_C = U_C^{-1} = i\gamma^2$ satisfies this condition. This has to do with the fact that γ^0, γ^1, and γ^3 are real whereas γ^2 is imaginary. This is specific for the Dirac representation of the Dirac matrices. Thus we obtain

$$\Psi^c = i\gamma^2 \Psi^* \ . \tag{145}$$

Consider now a stationary solution $H_D \Psi = E\Psi$, where $H_D = c\vec{\alpha} \cdot \vec{p} + \beta mc^2$, i.e., no field. It is easy to show that Ψ^c corresponds to the opposite energy, i.e., $H_D \Psi^c = -E\Psi^c$. This follows from the fact that $U_C^{-1} H_D U_C = -H_D^*$.

L. Non-relativistic limit, Pauli Hamiltonian

We start with

$$\begin{aligned} i\hbar \frac{\partial \varphi}{\partial t} &= c(\vec{\sigma} \cdot \vec{\pi})\chi + q\Phi\varphi + mc^2\varphi \ , \\ i\hbar \frac{\partial \chi}{\partial t} &= c(\vec{\sigma} \cdot \vec{\pi})\varphi + q\Phi\chi - mc^2\chi \ , \end{aligned} \tag{146}$$

where $\vec{\pi} \equiv \vec{p} - (q/c)\vec{A}$. Consider stationary case $i\hbar\partial/\partial t \to E$ and rewrite as

$$(E - mc^2 - q\Phi)\varphi = c(\vec{\sigma}\cdot\vec{\pi})\chi \ ,$$
$$(E + mc^2 - q\Phi)\chi = c(\vec{\sigma}\cdot\vec{\pi})\varphi \ . \tag{147}$$

Consider the limit of small positive energies: $E + mc^2 - q\Phi \approx 2mc^2$. That is $E - q\Phi - mc^2 \ll mc^2$. Since, classically, $E = q\Phi + c\sqrt{(mc)^2 + (\vec{\pi})^2}$, this limit corresponds to $|\vec{\pi}| \ll mc$ or $v \ll c$. Then from the second equation we obtain

$$\chi \approx \frac{(\vec{\sigma}\cdot\vec{\pi})\varphi}{2mc} \tag{148}$$

Substituting to the first equation we obtain

$$(E - mc^2 - q\Phi)\varphi = \frac{(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi})}{2m}\,\varphi \tag{149}$$

We introduce $\tilde{E} \equiv E - mc^2$ and obtain

$$\tilde{E}\varphi = \frac{(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi})}{2m}\,\varphi + q\Phi\varphi \tag{150}$$

Substituting $\vec{\pi} \equiv \vec{p} - (q/c)\vec{A}$ we obtain the Hamiltonian

$$H = \frac{\left(\vec{\sigma}\cdot\left(\vec{p} - \frac{q}{c}\vec{A}\right)\right)\left(\vec{\sigma}\cdot\left(\vec{p} - \frac{q}{c}\vec{A}\right)\right)}{2m} + q\Phi \ . \tag{151}$$

We use $(\vec{\sigma}\cdot\vec{C})(\vec{\sigma}\cdot\vec{D}) = \vec{C}\cdot\vec{D} + i\vec{\sigma}\cdot(\vec{C}\times\vec{D})$ where $\vec{C}$ and $\vec{D}$ can be orbital operators. This gives

$$H = \frac{\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2}{2m} + \frac{i}{2m}\vec{\sigma}\cdot\left[\left(\vec{p} - \frac{q}{c}\vec{A}\right)\times\left(\vec{p} - \frac{q}{c}\vec{A}\right)\right] + q\Phi \ . \tag{152}$$

With $\vec{p} = -i\hbar\vec{\nabla}$ we obtain the Pauli Hamiltonian

$$H = \frac{\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2}{2m} - \frac{\hbar q}{2mc}\vec{B}\cdot\vec{\sigma} + q\Phi \ . \tag{153}$$

with the correct giro-magnetic ratio.

Indeed, introducing the Bohr magneton $\mu_B \equiv \frac{e\hbar}{2mc}$ and taking $q = -e$ (we use $e > 0$) we obtain the Zeeman term as

$$H_Z = \mu_B\vec{B}\cdot\vec{\sigma} = g\frac{\mu_B}{\hbar}\vec{B}\cdot\vec{S} \ , \tag{154}$$

where $\vec{S} \equiv \hbar\vec{\sigma}/2$ and $g = 2$. From the electrodynamics we know that the interaction of a magnetic moment $\vec{\mu}$ with the magnetic field reads $H = -\vec{\mu} \cdot \vec{B}$. Thus, for the magnetic moment of the electron we get $\vec{\mu}_e = -(g/\hbar)\mu_B \vec{S}$.

To compare with the coupling of the magnetic field to the orbital angular momentum $\vec{L} \equiv \vec{r} \times \vec{p}$ we consider the Pauli Hamiltonian (153) in the regime of weak constant in space and time magnetic field $\vec{B}$. For such a magnetic field a convenient choice of the vector potential is

$$\vec{A}(\vec{r}) = \frac{1}{2}\,\vec{B} \times \vec{r}\,. \tag{155}$$

Then (since $\vec{\nabla} \cdot \vec{A} = 0$)

$$\frac{\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2}{2m} = \frac{(\vec{p})^2}{2m} - \frac{q}{mc}\vec{A} \cdot \vec{p} + \frac{q^2}{2mc^2}(\vec{A})^2\,. \tag{156}$$

The middle term can be rewritten using

$$\vec{A} \cdot \vec{p} = \frac{1}{2}\left(\vec{B} \times \vec{r}\right) \cdot \vec{p} = \frac{1}{2}\,(\vec{r} \times \vec{p}) \cdot \vec{B} = \frac{1}{2}\vec{B} \cdot \vec{L}\,. \tag{157}$$

Thus, for the Pauli Hamiltonian we obtain

$$H = \frac{(\vec{p})^2}{2m} - \frac{q}{2mc}\vec{B} \cdot (\vec{L} + 2\vec{S}) + \frac{q^2}{8mc^2}\left(\vec{B} \times \vec{r}\right)^2\,. \tag{158}$$

For a sufficiently weak magnetic field we can neglect the last term. We see that the spin coupes to the magnetic field twice stronger than the orbital angular momentum.

1. Next order terms

The further expansion is frequently performed with the help of Foldy-Wouthuysen transformation. The result reads

$$\begin{aligned} H \;=\; & \frac{\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2}{2m} - \frac{\hbar q}{2mc}\vec{B} \cdot \vec{\sigma} + q\Phi \\ & - \frac{(\vec{p})^4}{8m^3c^2} - \frac{q\hbar^2}{8m^2c^2}\,(\vec{\nabla} \cdot \vec{E}) + \frac{\hbar q}{8m^2c^2}\,\vec{\sigma} \cdot (\vec{p} \times \vec{E} - \vec{E} \times \vec{p})\,. \end{aligned} \tag{159}$$

The three last terms are 1) the relativistic correction to the kinetic energy; 2) the Darwin term; 3) the spin-orbit coupling.

The spin-orbit term can be further analysed for the case of a central potential $V(r) = q\Phi(r)$. For the electric field we obtain

$$q\vec{E} = -q\vec{\nabla}\Phi = -q\frac{\vec{r}}{r}\frac{d\Phi}{dr} = -\frac{\vec{r}}{r}\frac{dV}{dr}. \tag{160}$$

Then

$$q\left(\vec{p}\times\vec{E}\right) = -\left(\vec{p}\times\vec{r}\right)\frac{1}{r}\frac{dV}{dr} = \vec{L}\frac{1}{r}\frac{dV}{dr} = \frac{1}{r}\frac{dV}{dr}\vec{L} = -q\left(\vec{E}\times\vec{p}\right). \tag{161}$$

Thus, for the spin-orbit term we obtain

$$H_{SO} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\vec{S}\cdot\vec{L}\,. \tag{162}$$

2. Density and current density in the non-relativistic limit

Investigate now the particle density and the current density . The density is given by $\rho = \Psi^\dagger\Psi = \varphi^\dagger\varphi + \chi^\dagger\chi$. As we see from Eq. (148) $\frac{\chi^\dagger\chi}{\varphi^\dagger\varphi} \sim v^2/c^2$. Thus we obtain

$$\rho \approx \varphi^\dagger\varphi \tag{163}$$

For the current density we get

$$\vec{j} = c\Psi^\dagger\vec{\alpha}\Psi = c(\varphi^\dagger\vec{\sigma}\chi + \chi^\dagger\vec{\sigma}\varphi) \tag{164}$$

We substitute

$$\chi = \frac{\vec{\sigma}\cdot\left[(-i\hbar\vec{\nabla} - (q/c)\vec{A})\varphi\right]}{2mc}\,, \tag{165}$$

and

$$\chi^\dagger = \frac{\left[(i\hbar\vec{\nabla} - (q/c)\vec{A})\varphi^\dagger\right]\cdot\vec{\sigma}}{2mc}\,. \tag{166}$$

Here, $\vec{\nabla}$ remained on the left side of $\varphi^\dagger$. Thus it was not transposed and the sign in front of $i\hbar\vec{\nabla}$ had to change.

We use $\vec{\sigma}\cdot(\vec{\sigma}\cdot\vec{a}) = \vec{a} - i(\vec{\sigma}\times\vec{a})$ and $(\vec{a}\cdot\vec{\sigma})\cdot\vec{\sigma} = \vec{a} - i(\vec{a}\times\vec{\sigma})$. This results in

$$\vec{j} = -\frac{i\hbar}{2m}(\varphi^\dagger(\vec{\nabla}\varphi) - (\vec{\nabla}\varphi^\dagger)\varphi) - \frac{q}{mc}\vec{A}\varphi^\dagger\varphi + \frac{\hbar}{2m}\left[\vec{\nabla}\times(\varphi^\dagger\vec{\sigma}\varphi)\right] \tag{167}$$

The charge density and the electric current density are obtained as $\rho_q = q\,\rho$, $\vec{j}_q = q\,\vec{j}$.

The last term of (167) is the spin related current, i.e., as if the electron was spinning around. It does not contribute to the charge flow since $\vec{\nabla}\cdot(\vec{\nabla}\times ...) = 0$. The last term of

(167) can be related to the spin-related magnetisation. Indeed, from the electrodynamics we know that the magnetisation density $\vec{M}(t,\vec{r})$ and charge current density (of bound charges) are related via $\vec{j}_q = c(\vec{\nabla}\times\vec{M})$. Thus we can identify the spin related magnetisation density $\vec{M} = -(g/\hbar)\mu_B\,\varphi^\dagger\vec{S}\varphi = -\mu_B\,\varphi^\dagger\vec{\sigma}\varphi$. Everything becomes consistent with $q=-e$.

III. IDENTICAL PARTICLES

Particles are usually characterised by their position $\vec{r}$ and their spin projection (say on z-direction) m_s. If we are talking about particles of spin s, e.g., $s=0$, $s=1/2$, $s=1$, the quantum number m_s takes values $m_s = -s, -s+1, \ldots, s-1, s$. So a wave function of a single particle can be written as $\psi(\vec{r}, m_s)$. We can also think of it as a spinor with $2s+1$ components $[\psi(\vec{r}, m_s = s), \psi(\vec{r}, m_s = s-1), \ldots, \psi(\vec{r}, m_s = -s)]^T$, but here it would be more convenient to think of it as a function of two arguments $\vec{r}$ and m_s. A wave function of two particles of the same kind would be then in general $\Psi(\vec{r}_1, m_{s1}, \vec{r}_2, m_{s2})$. One frequently uses an abbreviation

$$\Psi(1,2) \equiv \Psi(\vec{r}_1, m_{s1}, \vec{r}_2, m_{s2})\ . \tag{168}$$

For N particles we would have in general

$$\Psi(1,2,\ldots,N) \equiv \Psi(\vec{r}_1, m_{s1}, \vec{r}_2, m_{s2}, \ldots, \vec{r}_N, m_{sN})\ . \tag{169}$$

The question is what happens with such a wave function if two particles are exchanged. Formally we introduce a permutation operator P_{12} such that

$$P_{12}\Psi(1,2) = P_{12}\Psi(\vec{r}_1, m_{s1}, \vec{r}_2, m_{s2}) \equiv \Psi(2,1) = \Psi(\vec{r}_2, m_{s2}, \vec{r}_1, m_{s1})\ . \tag{170}$$

Analogously one can define P_{mn} echanging the particles m and n in a wave function with N particles. Note, we exchange both the positions $\vec{r}_1 \leftrightarrow \vec{r}_2$ and the spin projections $m_{s1} \leftrightarrow m_{s2}$.

It is argued that if the particles are identical and indistinguishable after exchange we should get the same physical state, i.e.,

$$P_{12}\Psi(1,2) = \Psi(2,1) = \lambda\Psi(1,2)\ . \tag{171}$$

Here $|\lambda| = 1$ so that the normalisation is kept. In 3D it is further argued that the only two possibilities are $\lambda = \pm 1$. This is because exchanging two particles twice is equivalent to doing nothing, $P_{12}^2 = 1$, i.e. $P_{12}P_{12}\Psi(1,2) = \Psi(1,2)$. Interestingly, in 2D the situation

is not so simple, but we will not touch this issue here. The particles, for which $\lambda = +1$ are called bosons (Satyendra Nath Bose, 1894-1974), the particles, for which $\lambda = -1$ are called fermions (Enrico Fermi, 1901-1954).

In relativistic quantum field theory a very important theorem has been proven, the **spin-statistics** theorem. It states that particles with integer spin are bosons and particles with half-integer spin are fermions. Thus electrons (spin 1/2) are fermions.

A Hamiltonian describing many identical particles must commute with the permutation (exchange) operator P_{mn}. If it does not, it would distinguish between the particles. So, for example two interacting particles would be in general governed by a Hamiltonian of the type

$$H = \frac{(\vec{p}_1)^2}{2m} + \frac{(\vec{p}_2)^2}{2m} + V(\vec{r}_1) + V(\vec{r}_2) + U(|\vec{r}_1 - \vec{r}_2|) \ . \tag{172}$$

Clearly P_{12} commutes with H. A legitimate Hamiltonian could also have spin-dependence, e.g.,

$$H = \frac{(\vec{p}_1)^2}{2m} + \frac{(\vec{p}_2)^2}{2m} + V(\vec{r}_1) + V(\vec{r}_2) + U(|\vec{r}_1 - \vec{r}_2|) + J\vec{S}_1 \cdot \vec{S}_2 \ . \tag{173}$$

A. Non-interacting particles

A special and very important case is that of non-interacting particles. Let us for simplicity consider two particles. The Hamiltonian has the form

$$H = H^{(1)}(1) + H^{(1)}(2) = H^{(1)}(\vec{r}_1, \vec{p}_1, \vec{S}_1) + H^{(1)}(\vec{r}_2, \vec{p}_2, \vec{S}_2) \ . \tag{174}$$

The superscript (1) in $H^{(1)}$ means a single particle Hamiltonian. For example

$$H^{(1)}(\vec{r}, \vec{p}) = (\vec{p})^2/2m + V(\vec{r}) \ . \tag{175}$$

Assume $H^{(1)}$ has a set of eigenstates ψ_n with energies E_n, i.e., $H^{(1)}\psi_n = E_n\psi_n$. What are the two-particle states? Let's take first a pair of states ψ_1 and ψ_2. For distinguishable particles we could construct four different states: 1) $\psi_1(1)\psi_1(2)$ with energy $E = 2E_1$, 2) $\psi_2(1)\psi_2(2)$ with energy $E = 2E_2$, 3) $\psi_1(1)\psi_2(2)$ with energy $E = E_1 + E_2$, 4) $\psi_2(1)\psi_1(2)$ with energy $E = E_1 + E_2$.

For indistinguishable particles the situation is different. For fermions the only possible eigenstate is

$$\Psi(1,2) = \frac{1}{\sqrt{2}} \left(\psi_1(1)\psi_2(2) - \psi_1(2)\psi_2(1)\right) \quad , \quad E = E_1 + E_2 \ . \tag{176}$$

For bosons the following possibilities are available

$$\Psi(1,2) = \psi_1(1)\psi_1(2) \quad , \quad E = 2E_1 \ ,$$

or

$$\Psi(1,2) = \psi_2(1)\psi_2(2) \quad , \quad E = 2E_2 \ ,$$

or

$$\Psi(1,2) = \frac{1}{\sqrt{2}}\left(\psi_1(1)\psi_2(2) + \psi_2(2)\psi_1(1)\right) \quad , \quad E = E_1 + E_2 \ . \tag{177}$$

Of course the pairs of states are arbitrary. We clearly observe the Pauli principle: two fermions cannot be in the same state.

Consider now N particles. The situation is again very different for bosons and for fermions. For bosons we can construct a state in which N_1 particles are in state ψ_1, N_2 particles are in state ψ_2 etc.. This state reads

$$|N_1, N_2, ...\rangle = \left(\frac{N_1! N_2! \ldots}{N!}\right)^{1/2} \sum_P \psi_{P_1}(1)\psi_{P_2}(2)\ldots\psi_{P_N}(N) \ , \tag{178}$$

where $N = N_1 + N_2 + \ldots$ is the total number of particles. The permutation P counts all different arrangements of N_1 numbers 1, N_2 numbers 2 etc.

For fermions the wave functions must be antisymmetric. This means that the occupation numbers are either 0 or 1, $N_i = 0, 1$. This is called Pauli principle. The wave function for N particles occupying states $i = 1, ..., N$ is given by

$$|1_1, 1_2, ..., 1_N\rangle = \left(\frac{1}{N!}\right)^{1/2} \sum_P (-1)^P \psi_{P_1}(1)\psi_{P_2}(2)\ldots\psi_{P_N}(N) \ , \tag{179}$$

To fix the signs one chooses a certain order of states. For example, the permutation with $P_1 < P_2 < \cdots < P_N$ will be assigned the positive sign. This wave function can be conveniently presented as a determinant of a matrix (Slater determinant). For example for 3 particles is 3 different states ψ_1, ψ_2, ψ_3 the wave function reads

$$|1_1, 1_2, 1_3\rangle = \sqrt{\frac{1}{3!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \psi_3(1) \\ \psi_1(2) & \psi_2(2) & \psi_3(2) \\ \psi_1(3) & \psi_2(3) & \psi_3(3) \end{vmatrix} \tag{180}$$

B. Helium Atom

We now consider a helium atom with the nucleus with charge $+Ze$ ($Z = 2$, $e > 0$) and two electrons. The Hamiltonian of the two electrons reads

$$H = \frac{(\vec{p}_1)^2}{2m} + \frac{(\vec{p}_2)^2}{2m} - \frac{Ze^2}{|\vec{r}_1|} - \frac{Ze^2}{|\vec{r}_2|} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \ . \tag{181}$$

Let us first disregard the interaction between the electrons. More precisely we write $H = H_0 + V$, where

$$H_0 = \frac{(\vec{p}_1)^2}{2m} + \frac{(\vec{p}_2)^2}{2m} - \frac{Ze^2}{|\vec{r}_1|} - \frac{Ze^2}{|\vec{r}_2|} \tag{182}$$

and

$$V = \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \ . \tag{183}$$

The eigenstates of H_0 can be constructed from the eigenstates of the single particle Hamiltonian (hydrogen atom):

$$H^{(1)} = \frac{(\vec{p})^2}{2m} - \frac{Ze^2}{|\vec{r}|} \ . \tag{184}$$

These states are

$$\psi_{k,\sigma}(\vec{r}, m_s) = \phi_k(\vec{r})\chi_\sigma(m_s) \ . \tag{185}$$

Here $k = (n, l, m)$ and $\sigma = \uparrow, \downarrow$. Choosing 2 different states k_1, σ_1 and k_2, σ_2 we can construct the corresponding Slater determinant.

1. Ground state

The ground state of H_0 corresponds to $k_1 = k_2 = (1, 0, 0)$, $\sigma_1 = \uparrow$, $\sigma_2 = \downarrow$. The ground state wave function then reads

$$\Psi_{gs}(1, 2) = \phi_{1,0,0}(\vec{r}_1)\phi_{1,0,0}(\vec{r}_2) \cdot \frac{1}{\sqrt{2}} \left[\chi_\uparrow(m_{s1})\chi_\downarrow(m_{s2}) - \chi_\downarrow(m_{s1})\chi_\uparrow(m_{s2})\right] \ . \tag{186}$$

This state has the form of the Slater determinant, since we have 2 particles and 2 single-particle states are involved. We also observe that this state has s symmetric orbital part $\phi_{1,0,0}(\vec{r}_1)\phi_{1,0,0}(\vec{r}_2)$ and an anti-symmetric spin part $\frac{1}{\sqrt{2}} \left[\chi_\uparrow(m_{s1})\chi_\downarrow(m_{s2}) - \chi_\downarrow(m_{s1})\chi_\uparrow(m_{s2})\right]$. The spin part corresponds to a singlet state with total spin zero. The energy of the ground state is given by

$$E_{gs0} = 2 \times \left(-\frac{Z^2 R_y}{1^2}\right) = -8R_y \ . \tag{187}$$

Here $R_y \equiv \frac{\hbar^2}{2ma_0^2}$ is the Rydberg constant and $a_0 = \frac{\hbar^2}{me^2}$ is the Bohr radius. (In SI $a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2}$).

We estimate the ground state energy of the full Hamiltonian H (181) using the first order perturbation theory in V. This is not really justified, as V is not small. One can, first, analyse a hypothetical situation in which V is multiplied by a small number $\eta \ll 1$, i.e., $V = \eta e^2/|\vec{r}_1 - \vec{r}_2|$. Then, the perturbation treatment would probably be justified. Then we put $\eta = 1$ and hope that the results remain correct at least qualitatively.

Thus we need

$$\delta E_{gs} = \langle \Psi_{gs} | \, V \, | \Psi_{gs} \rangle \ . \tag{188}$$

Since the perturbation V does not act on the spin we obtain

$$\begin{aligned} \delta E_{gs} \ &= \ \int d^3r_1 d^3r_2 \, \phi^*_{100}(\vec{r}_1)\phi^*_{100}(\vec{r}_2) \, \frac{\eta e^2}{|\vec{r}_1 - \vec{r}_2|} \, \phi_{100}(\vec{r}_1)\phi_{100}(\vec{r}_2) \\ &= \ \int d^3r_1 d^3r_2 \, |\phi_{100}(\vec{r}_1)|^2 |\phi_{100}(\vec{r}_2)|^2 \, \frac{\eta e^2}{|\vec{r}_1 - \vec{r}_2|} \ . \end{aligned} \tag{189}$$

A straightforward integration with

$$\phi_{100}(\vec{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{3/2} e^{-Z|\vec{r}|/a_0} \tag{190}$$

gives

$$\delta E_{gs} = 2Z \times \frac{5}{8} \, \eta R_y = \frac{5}{2} \, \eta R_y \ . \tag{191}$$

Thus we obtain

$$E_{gs} = \left(-8 + \frac{5}{2}\eta \right) R_y \ . \tag{192}$$

For $\eta = 1$ the correction is not really small, but it turns out to be rather close to the experimental observations (we get $E_{gs} \approx -75$eV, the experimental value $E_{gs}^{\rm exp} \approx -79$eV). We will not try to improve here (variational method introducing a wave function (190) with $Z \to Z_{eff}$).

2. Excited states

To analyse the excited states it is important to notice that the continuum (ionisation) starts already at $E = -4R_y$, when one electron is ionised (promoted to E=0) and the second remains in the ground state. If we excite both electrons to $n = 2$, e.g., consider a state with $k_1 = k_2 = (2, 0, 0)$, its energy $E = 2 \times (-Z^2 R_y/2^2) = -2R_y$ is already

above the threshold. Thus, one considers only the states, where $k_1 = (1,0,0)$ and $k_2 = (n,l,m)$. The energy of all these states (with different choices of l, m, and spin projections) is given by $E = -4(1+1/n^2)R_y$. The degeneracy is $4n^2$: the usual n^2 for all the possible choices of l and m and another 4 for the spins. Indeed, if we fix l and m we still have 4 spin states. For distinguishable particles the $4n^2$ states would be $\psi_{k_1,\sigma_1}(1)\psi_{k_2,\sigma_2}(2) = \phi_{100}(\vec{r}_1)\chi_{\sigma_1}(m_{s1})\phi_{nlm}(\vec{r}_2)\chi_{\sigma_2}(m_{s2})$, with $\sigma_1 = \uparrow / \downarrow$ and $\sigma_2 = \uparrow / \downarrow$. Since $n \neq 0$, i.e., we have no situation of both particles occupying the same single-particle state, each of these $4n^2$ states can be anti-symmetrised, producing $4n^2$ Slater determinants:

$$\begin{aligned}\Psi_{k_1,\sigma_1;k_2,\sigma_2}(1,2) &= \frac{1}{\sqrt{2}}\left[\psi_{k_1,\sigma_1}(1)\psi_{k_2,\sigma_2}(2) - \psi_{k_1,\sigma_1}(2)\psi_{k_2,\sigma_2}(1)\right] \\ &= \frac{1}{\sqrt{2}}\left[\phi_{100}(\vec{r}_1)\chi_{\sigma_1}(m_{s1})\phi_{nlm}(\vec{r}_2)\chi_{\sigma_2}(m_{s2}) - \phi_{100}(\vec{r}_2)\chi_{\sigma_1}(m_{s2})\phi_{nlm}(\vec{r}_1)\chi_{\sigma_2}(m_{s1})\right] \ .\end{aligned} \tag{193}$$

Formally, in order to take into account the interaction term V we could use the degenerate perturbation theory and diagonalise a $4n^2 \times 4n^2$ matrix. Fortunately, there are conserved quantities which make the task much easier. Indeed, the whole Hamiltonian (181) commutes with the total orbital angular momentum $\vec{L} = \vec{L}_1 + \vec{L}_2 = \vec{r}_1 \times \vec{p}_1 + \vec{r}_2 \times \vec{p}_2$. Since $k_1 = (1,0,0)$ and $k_2 = (n,l,m)$, we have $\vec{L}^2\Psi_{k_1,\sigma_1;k_2,\sigma_2} = \hbar^2 L(L+1)\Psi_{k_1,\sigma_1;k_2,\sigma_2}$ with $L = l$ and $L_z\Psi_{k_1,\sigma_1;k_2,\sigma_2} = \hbar M\Psi_{k_1,\sigma_1;k_2,\sigma_2}$ with $M = m$. This is so simple because one of the electrons has zero angular momentum. Thus the perturbation does not mix states with different values of l and m. For each choice of l and m we have to diagonalise a 4×4 matrix:

$$\langle \Psi_{k_1,\sigma_1;k_2,\sigma_2}|\ V\ \left|\Psi_{k_1,\sigma_1';k_2,\sigma_2'}\right\rangle \ . \tag{194}$$

A simple inspection shows that this matrix is not diagonal.

A better strategy is to realise that we have two spins 1/2 that can be added to either spin $S = 1$ (triplet) or spin $S = 0$ (singlet). The perturbation would not mix different spin states. Thus we consider the total spin $\vec{S} = \vec{S}_1 + \vec{S}_2$ and construct the eigenstates of $\vec{S}^2$ (quantum number S) and S_z (quantum number M_s). These are

$$\begin{aligned}|S=0,M_s=0\rangle &= \frac{1}{\sqrt{2}}\left[\chi_\uparrow(m_{s1})\chi_\downarrow(m_{s2}) - \chi_\downarrow(m_{s1})\chi_\uparrow(m_{s2})\right] = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \ , \\ |S=1,M_s=1\rangle &= \chi_\uparrow(m_{s1})\chi_\uparrow(m_{s2}) = |\uparrow\uparrow\rangle \ , \\ |S=1,M_s=0\rangle &= \frac{1}{\sqrt{2}}\left[\chi_\uparrow(m_{s1})\chi_\downarrow(m_{s2}) + \chi_\downarrow(m_{s1})\chi_\uparrow(m_{s2})\right] = \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \ , \\ |S=1,M_s=-1\rangle &= \chi_\downarrow(m_{s1})\chi_\downarrow(m_{s2}) = |\downarrow\downarrow\rangle \ .\end{aligned} \tag{195}$$

The first state (singlet) is anti-symmetric, whereas the three triplet states are symmetric. Since the total wave function must be antisymmetric we should combine these spin states with either symmetrized and anti-symmetrized orbital wave functions. We introduce

$$\Phi^A_{k1,k2}(\vec{r}_1,\vec{r}_2) = \frac{1}{\sqrt{2}}\left[\phi_{k1}(\vec{r}_1)\,\phi_{k2}(\vec{r}_2) - \phi_{k1}(\vec{r}_2)\,\phi_{k2}(\vec{r}_1)\right] \ , \tag{196}$$

and

$$\Phi^S_{k1,k2}(\vec{r}_1,\vec{r}_2) = \frac{1}{\sqrt{2}}\left[\phi_{k1}(\vec{r}_1)\,\phi_{k2}(\vec{r}_2) + \phi_{k1}(\vec{r}_2)\,\phi_{k2}(\vec{r}_1)\right] \ . \tag{197}$$

The proper 4 wave functions are then $|k1,k2,S,M_s\rangle$:

$$\begin{aligned}
|k_1,k_2,0,0\rangle &= \Phi^S_{k1,k2}(\vec{r}_1,\vec{r}_2)\,\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \ , \\
|k_1,k_2,1,1\rangle &= \Phi^A_{k1,k2}(\vec{r}_1,\vec{r}_2)\ |\uparrow\uparrow\rangle \ , \\
|k_1,k_2,1,0\rangle &= \Phi^A_{k1,k2}(\vec{r}_1,\vec{r}_2)\,\frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \ , \\
|k_1,k_2,1,-1\rangle &= \Phi^A_{k1,k2}(\vec{r}_1,\vec{r}_2)\ |\downarrow\downarrow\rangle \ .
\end{aligned} \tag{198}$$

States with $S=0$ are called "parahelium", states with $S=1$ - "orthohelium".

The perturbation V is diagonal in the basis $|k1,k2,S,M_s\rangle$. For $S=0$ states one obtains the energy correction

$$\delta E_{S=0} = I_{k1,k2} + J_{k1,k2} \ . \tag{199}$$

For $S=1$ one obtains

$$\delta E_{S=1} = I_{k1,k2} - J_{k1,k2} \ . \tag{200}$$

Here

$$\begin{aligned}
I_{k1,k2} &\equiv \int d^3r_1 d^3r_2\,\phi^*_{k_1}(\vec{r}_1)\phi^*_{k_2}(\vec{r}_2)\,\frac{e^2}{|\vec{r}_1-\vec{r}_2|}\,\phi_{k_1}(\vec{r}_1)\phi_{k_2}(\vec{r}_2) \\
&= \int d^3r_1 d^3r_2\,|\phi_{k_1}(\vec{r}_1)|^2\,|\phi_{k_2}(\vec{r}_2)|^2\,\frac{e^2}{|\vec{r}_1-\vec{r}_2|} \ ,
\end{aligned} \tag{201}$$

and

$$J_{k1,k2} \equiv \int d^3r_1 d^3r_2\,\phi^*_{k_2}(\vec{r}_1)\phi^*_{k_1}(\vec{r}_2)\,\frac{e^2}{|\vec{r}_1-\vec{r}_2|}\,\phi_{k_1}(\vec{r}_1)\phi_{k_2}(\vec{r}_2) \ . \tag{202}$$

The integral $I_{k1,k2}$ is called *direct* integral. The integral $J_{k1,k2}$ is called *exchange* integral. If $J_{k1,k2} > 0$ (usual case) the triplet states have lower energy than the singlet ones. In our case ($k_1 = (1,0,0)$, $k_2 = (n,l,m)$) both integrals depend only on n and l, not on m. Thus, $2l+1$ orbital degeneracy remains in addition to the 3-fold degeneracy of the $S=1$ states. The good quantum numbers are $N=n, L=l, S$.

In the 4-dimensional subspace k_1, k_2 we obtain an effective spin Hamiltonian (exchange interaction)

$$H_{ex} = -\frac{J_{k1,k2}}{2}\left(1 + \frac{4}{\hbar^2}\,\vec{S}_1 \cdot \vec{S}_2\right) \ . \tag{203}$$

Indeed $2\vec{S}_1 \cdot \vec{S}_2 = (\vec{S}_1 + \vec{S}_2)^2 - \vec{S}_1^2 - \vec{S}_2^2 = \vec{S}^2 - \vec{S}_1^2 - \vec{S}_2^2 = \hbar^2(S(S+1) - 3/2)$. We get $\vec{S}_1 \cdot \vec{S}_2 = -\hbar^2/4$ for $S = 0$ and $\vec{S}_1 \cdot \vec{S}_2 = 3\hbar^2/4$ for $S = 1$.

IV. EXTRA READING

A. Parity operator in Dirac theory

In non-relativistic quantum mechanics parity operator is trivial:

$$P\Psi(\vec{r}, t) = \Psi(-\vec{r}, t) \tag{204}$$

States with orbital angular momentum l are eigenvectors of P and

$$PY_{lm}(\vec{r}) = (-1)^l Y_{lm}(\vec{r}) \tag{205}$$

This is the orbital *parity*. The spin parity, i.e., how spin transforms under spatial inversion $\vec{r} \to -\vec{r}$ could be chosen arbitrarily. One could even choose $P^2 = \pm 1$ (consistent with half-integer s). This choice would change nothing.

For relativistic quantum mechanics parity operator is more important (along with time inversion T and charge conjugation C). The unitary linear operator P should possess the following properties: $P\vec{r}P^{-1} = -\vec{r}$ and $P\vec{p}P^{-1} = -\vec{p}$. For the Dirac equation to remain form-invariant we should then choose $P\vec{\gamma}P^{-1} = -\vec{\gamma}$ and $P\gamma^0 P^{-1} = \gamma^0$ (or $P\vec{\alpha}P^{-1} = -\vec{\alpha}$ and $P\beta P^{-1} = \beta$). One of the possible choices is

$$P = \gamma^0 P_{\rm orb} = \beta P_{\rm orb} \ , \tag{206}$$

where $P_{\rm orb}\psi(\vec{r}) = \psi(-\vec{r})$. Thus 4-spinors transform less trivially than in the non-relativistic theory. The possible eigenvalues of P are ± 1.

B. Motion in spherically symmetric potential, hydrogen atom in Dirac theory

1. *Eigenstates of J^2, J_z, and P*

We now study the case of spherically symmetric potential $V(|\vec{r}|) = V(r)$. The Dirac Hamiltonian then reads

$$H_D = c\vec{\alpha}\cdot\vec{p} + \beta mc^2 + V(r) \ , \tag{207}$$

We determine conserved quantities. These are the full angular momentum $\vec{J}$ and P. That is $[H_D, \vec{J}] = 0$ and $[H_D, P] = 0$.

Let us take the non-relativistic spherical spinors. These are eigenstates of J^2 and J_z obtained by adding an orbital momentum l and spin $1/2$. There are two possible values of j then, namely $j = l \pm 1/2$. Except when $l = 0$ there is only one possibility $j = l + 1/2 = 1/2$. The states are

$$\Omega_{j=l\pm1/2,m}(\theta,\phi) = \begin{pmatrix} \pm\sqrt{\frac{l\pm m+(1/2)}{2l+1}}\, Y_{l,m-1/2}(\theta,\phi) \\ \\ \sqrt{\frac{l\mp m+(1/2)}{2l+1}}\, Y_{l,m+1/2}(\theta,\phi) \end{pmatrix} \tag{208}$$

We invert the logic and say that we obtain j from either $l = j - 1/2$ or from $l = j + 1/2$ and denote these spinors as $\Omega_{j,m;l=j\pm1/2}$. Then the eigenvectors of J^2, J_z, and P read

$$\begin{pmatrix} \varphi(r,\theta,\phi) \\ \chi(r,\theta,\phi) \end{pmatrix} = \begin{pmatrix} f(r)\cdot\Omega_{j,m;l=j-\kappa/2} \\ g(r)\cdot\Omega_{j,m;l=j+\kappa/2} \end{pmatrix} \ , \tag{209}$$

where κ can take two values $\kappa = \pm 1$. Indeed in this case we obtain

$$J^2 \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \begin{pmatrix} \hbar^2 j(j+1)\varphi \\ \hbar^2 j(j+1)\chi \end{pmatrix} \tag{210}$$

$$J_z \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \begin{pmatrix} \hbar m\varphi \\ \hbar m\chi \end{pmatrix} \tag{211}$$

and

$$P \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \beta P_{\rm orb} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \begin{pmatrix} P_{\rm orb}\varphi \\ -P_{\rm orb}\chi \end{pmatrix} = \begin{pmatrix} (-1)^{j-\kappa/2}\,\varphi \\ -(-1)^{j+\kappa/2}\,\chi \end{pmatrix} = \pm \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \tag{212}$$

2. Dirac equation for a given j

The Dirac equation reads

$$(E - mc^2 - V)\varphi = c(\vec{\sigma}\cdot\vec{p})\chi \ ,$$
$$(E + mc^2 - V)\chi = c(\vec{\sigma}\cdot\vec{p})\varphi \ . \tag{213}$$

We need, e.g., $(\vec{\sigma}\cdot\vec{p})\chi$. We obtain

$$(\vec{\sigma}\cdot\vec{p})\chi = r^{-2}(\vec{\sigma}\cdot\vec{r})(\vec{\sigma}\cdot\vec{r})(\vec{\sigma}\cdot\vec{p})\chi = r^{-2}(\vec{\sigma}\cdot\vec{r})(\vec{r}\cdot\vec{p} + i\vec{\sigma}\cdot\vec{L})\chi \tag{214}$$

We use $\hbar\vec{\sigma}\cdot\vec{L} = (\vec{L} + (\hbar/2)\vec{\sigma})^2 - L^2 - 3\hbar^2/4 = \hbar^2[j(j+1) - l(l+1) - 3/4]$. Since for $\chi = g(r)\cdot\Omega_{j,m;l=j+\kappa/2}$ we obtain

$$(\vec{\sigma}\cdot\vec{L})\chi = -\hbar[\kappa(j+1/2)+1]g(r)\cdot\Omega_{j,m;l=j+\kappa/2} \tag{215}$$

On the other hand the operator $\vec{r}\vec{p}$ acts only on $g(r)$ as

$$(\vec{r}\vec{p})g(r) = -i\hbar r\frac{\partial g}{\partial r} \tag{216}$$

Finally (without proof) $r^{-1}(\vec{\sigma}\cdot\vec{r})\Omega_{j,m;l=j+\kappa/2} = -\Omega_{j,m;l=j-\kappa/2}$ (for $\kappa = \pm 1$, Landau-Lifschitz get here an extra i due to a different definition of Y_{lm}).

Thus

$$(\vec{\sigma}\cdot\vec{p})\chi = i\hbar r^{-1}\left\{r\frac{\partial g}{\partial r} + [\kappa(j+1/2)+1]g(r)\right\}\cdot\Omega_{j,m;l=j-\kappa/2} \tag{217}$$

and

$$(\vec{\sigma}\cdot\vec{p})\varphi = i\hbar r^{-1}\left\{r\frac{\partial f}{\partial r} - [\kappa(j+1/2)-1]f(r)\right\}\cdot\Omega_{j,m;l=j+\kappa/2} \tag{218}$$

The Dirac equation now reads

$$(E - mc^2 - V)f = ic\hbar\left\{\frac{\partial g}{\partial r} + \frac{1}{r}[\kappa(j+1/2)+1]g\right\} \ ,$$
$$(E + mc^2 - V)g = ic\hbar\left\{\frac{\partial f}{\partial r} - \frac{1}{r}[\kappa(j+1/2)-1]f\right\} \ . \tag{219}$$

for $G \equiv -irg$ and $F \equiv rf$ this becomes

$$(E - mc^2 - V)F = -c\hbar\left\{\frac{\partial G}{\partial r} + \frac{\kappa(j+1/2)G}{r}\right\} \ ,$$
$$(E + mc^2 - V)G = \ \ c\hbar\left\{\frac{\partial F}{\partial r} - \frac{\kappa(j+1/2)F}{r}\right\} \ . \tag{220}$$

3. Hydrogen atom

For $V = -Ze^2/r$ we obtain

$$\begin{aligned}(E - mc^2)F &= -c\hbar \left\{ \frac{\partial G}{\partial r} + \frac{\kappa(j+1/2)G + Z\alpha F}{r} \right\} , \\ (E + mc^2)G &= \quad c\hbar \left\{ \frac{\partial F}{\partial r} - \frac{\kappa(j+1/2)F + Z\alpha G}{r} \right\} ,\end{aligned} \tag{221}$$

where $\alpha \equiv e^2/(\hbar c) \approx 1/137$ is the fine structure constant (in SI $\alpha \equiv e^2/(4\pi\epsilon_0\hbar c)$).

We are looking for bound states, i.e., $E < mc^2$. Thus we introduce $\rho \equiv r\sqrt{m^2c^4 - E^2}/(\hbar c)$. Then the equations read

$$\begin{aligned}\sqrt{\frac{mc^2 - E}{mc^2 + E}}\, F &= \left\{ \frac{\partial G}{\partial \rho} + \frac{Z\alpha F + \kappa(j+1/2)G}{\rho} \right\} , \\ \sqrt{\frac{mc^2 + E}{mc^2 - E}}\, G &= \left\{ \frac{\partial F}{\partial \rho} - \frac{Z\alpha G + \kappa(j+1/2)F}{\rho} \right\} .\end{aligned} \tag{222}$$

Introduce $A \equiv \sqrt{\frac{mc^2-E}{mc^2+E}}$, $B \equiv Z\alpha$, $C \equiv \kappa(j+1/2)$. Then

$$\begin{aligned}A\, F &= \left\{ \frac{\partial G}{\partial \rho} + \frac{BF + CG}{\rho} \right\} , \\ \frac{1}{A}\, G &= \left\{ \frac{\partial F}{\partial \rho} - \frac{BG + CF}{\rho} \right\} .\end{aligned} \tag{223}$$

We now use an ansatz:

$$F \;=\; \rho^s e^{-\rho} \sum_{m=0}^{\infty} F_m \rho^m \,, \tag{224}$$

$$G \;=\; \rho^s e^{-\rho} \sum_{m=0}^{\infty} G_m \rho^m \,. \tag{225}$$

Substituting we obtain the following recurrence relations for $m \geq 0$:

$$\begin{aligned}AF_{m-1} &= (s + m + C)G_m - G_{m-1} + BF_m \,, \\ A^{-1}G_{m-1} &= (s + m - C)F_m - F_{m-1} - BG_m \,.\end{aligned} \tag{226}$$

For $m = 0$ this gives

$$(s + C)G_0 + BF_0 = 0 \quad \textit{and} \quad (s - C)F_0 - BG_0 = 0 \tag{227}$$

which means

$$s = \pm\sqrt{C^2 - B^2} = \pm\sqrt{(j+1/2)^2 - Z^2\alpha^2} \tag{228}$$

The minimum value of j is $1/2$. Thus everything OK if $Z\alpha < 1$, i.e., if $Z < 137$. Then we take the positive root so that the wave functions vanish at $r = 0$.

$$s = \sqrt{(j+1/2)^2 - Z^2\alpha^2} \tag{229}$$

For $Z > 137$ interesting things happen but we will not consider this case.

The eigenvectors are found from the condition that for some value of $m = n'$ $F_{n'+1} = G_{n'+1} = 0$, whereas $F_{n'} \neq 0$ and $G_{n'} \neq 0$. Then both of Eq. (226) give for $m = n' + 1$

$$AF_{n'} = -G_{n'} \, . \tag{230}$$

Now we use Eq. (226) give for $m = n'$. Eliminating $G_{n'-1}$ and $F_{n'-1}$ and using (230) we obtain

$$2(s + n') = -B(A - A^{-1}) \tag{231}$$

This equation gives

$$E = \frac{mc^2}{\sqrt{1 + \frac{Z^2\alpha^2}{(s+n')^2}}} \tag{232}$$

The allowed values of n' are $0, 1, ...$ For $n' = 1, 2...$ one can choose arbitrary j (positive and half-integer) and arbitrary parity (κ). For $n' = 0$ there is an extra condition. As one can conclude from (227) and (230) we must have $C > 0$, i.e., $\kappa = +1$.

4. The results

Principal quantum numbers: $n = n' + j + 1/2$.

For a given n possible values of j are $j = 1/2, 3/2, ..., n - 1/2$. As one can see $j = n - 1/2$ corresponds to $n' = 0$. Thus for all values of $j < n - 1/2$ the states are double-degenerate, but for $j = n - 1/2$ only $\kappa = +1$ is allowed. For each $j < n - 1/2$ there are two values of κ. One names the states by the l value of φ, i.e., by $l = j - \kappa/2$. For $j = n - 1/2$ only one value of l is allowed $l = j - 1/2$, i.e., $\kappa = 1$.

The states are named as nl_j, where $l = j \pm 1/2$, except for $j = n - 1/2$ when $l = j - 1/2$. The first states are $1s_{1/2}$, $2s_{1/2}$, $2p_{1/2}$, $2p_{3/2}$ etc.

The energy is given by

$$E_{nj} = mc^2 \left[1 + \frac{Z^2\alpha^2}{(n-\epsilon_j)^2} \right]^{-\frac{1}{2}} , \tag{233}$$

$$\epsilon_j = j + \frac{1}{2} - \sqrt{\left(j + \frac{1}{2} \right)^2 - Z^2\alpha^2} \, . \tag{234}$$

For given n the energy increases with j. The only degeneracy left is between $l = j \pm 1/2$. So the states $2s_{1/2}$ and $2p_{1/2}$ are degenerate. This degeneracy is lifted by the radiative corrections (Lamb shift).

V. LITERATURE

1) L.D. Landau & E.M. Lifschitz, Lehrbuch der theoretischen Physik,
Band 3: Quantenmechanik,
Band 4: Quantenelektrodynamik.

2) T. Fließbach, Lehrbuch zur Theoretischen Physik,
Band III: Quantenmechanik.

3) A. Messiah, Quantum Mechanics.

4) J. D. Bjorken, S. D. Drell, Relativistic Quantum Mechanics.

5) J. J. Sakurai, J. Napolitano, Modern Quantum Mechanics.

6) C. Cohen-Tannoudji, B. Diu, F. Laloë, Quantenmechanik: Band 2, Band 3.