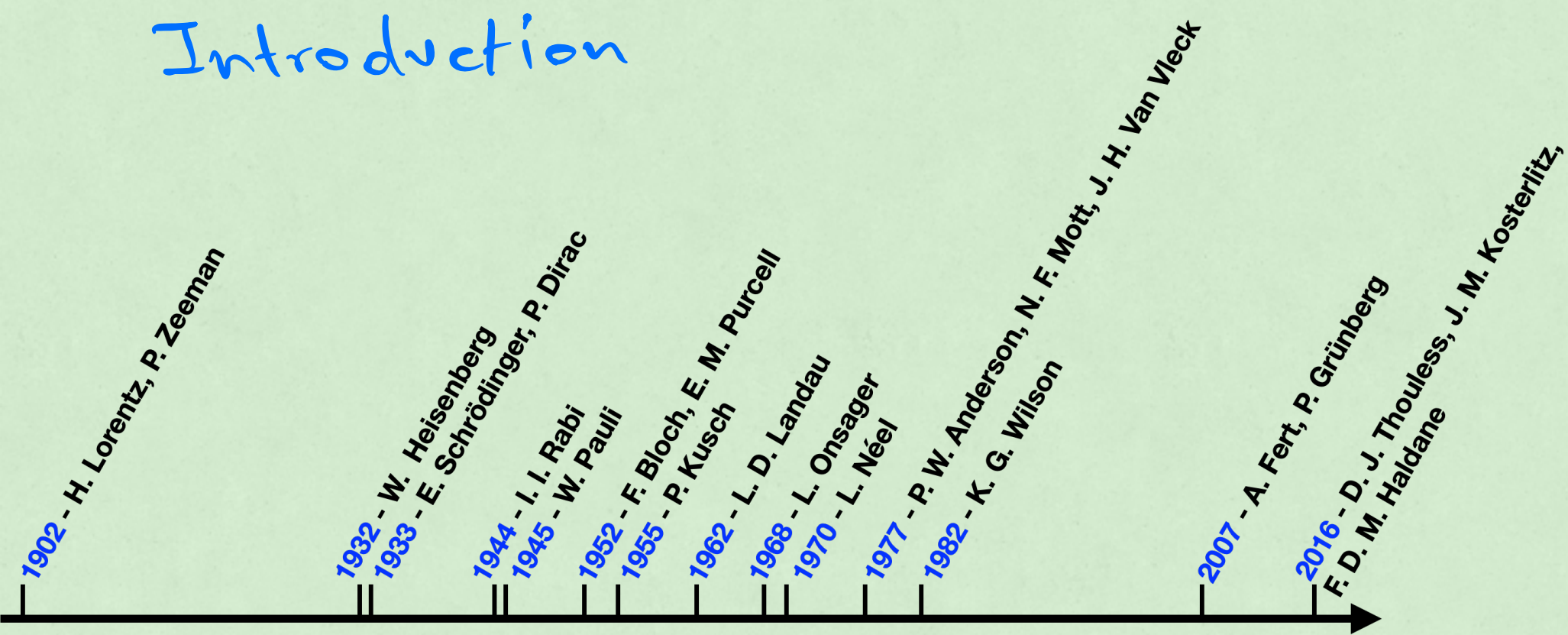
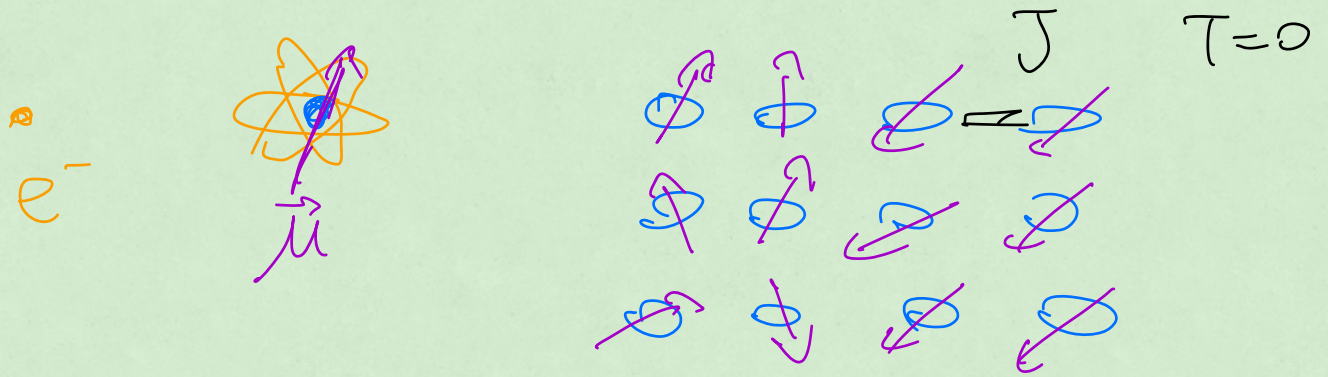


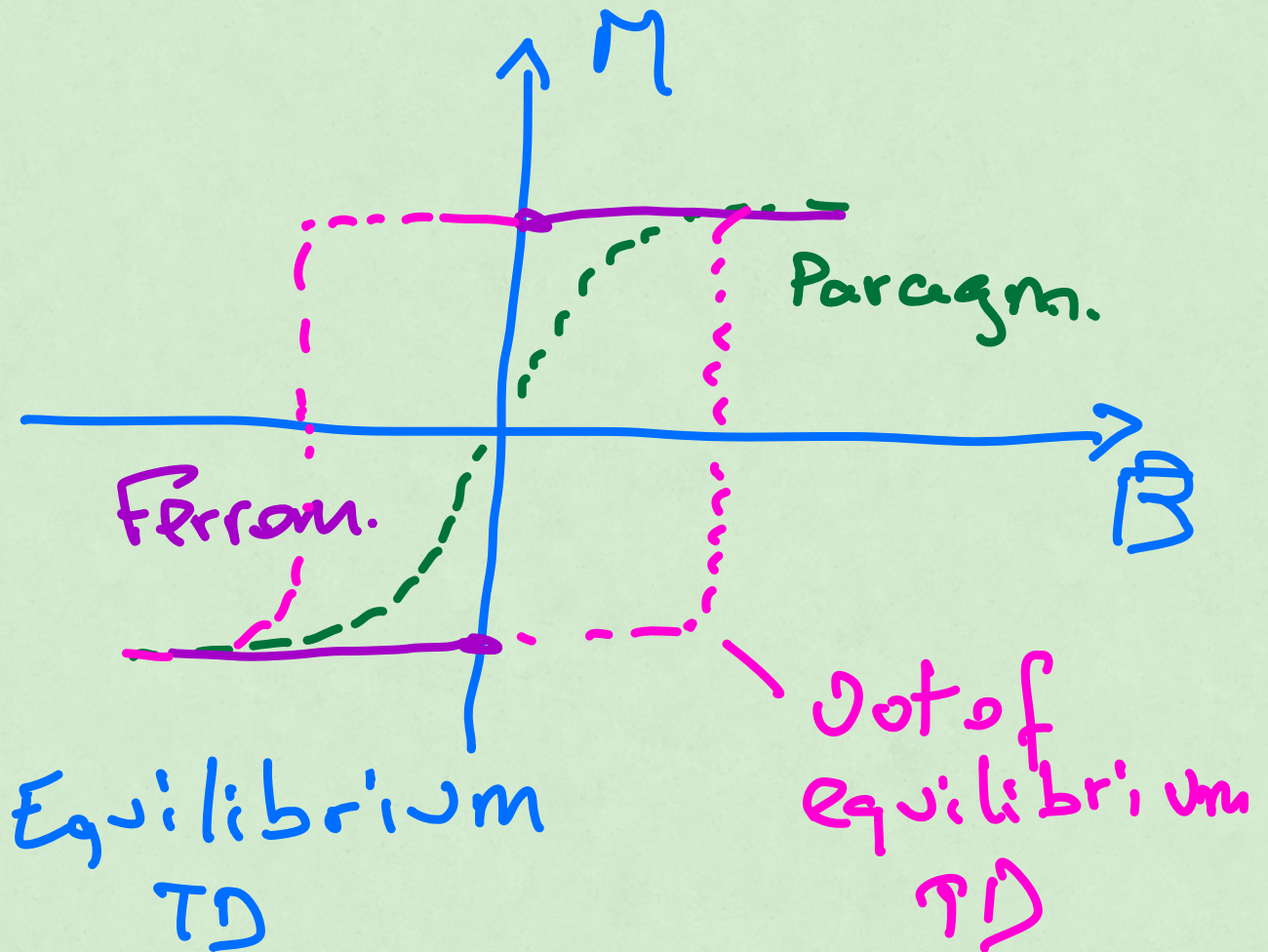
Introduction





$$\vec{M}(B, T) = \frac{\langle \vec{\mu} \rangle}{\Delta V}$$

$T \neq 0$



$$m\ddot{x} = F \quad \text{Newton}$$

invariant w.r.t

$$x' = x - vt$$

t absolute
Galilean tr.

Maxwell eq...

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \vec{E} = 0$$

$$\left(\dots \right) \vec{B} = 0$$

invariant w.r.t
Lorentz tr.

Lorentz
tr

\longrightarrow
 $v \ll c$

Galilean
tr.

$$\sim \quad \sim$$
$$i\hbar \frac{\partial}{\partial t} \psi = H \psi$$

$$i\hbar \frac{\partial}{\partial t} \psi = -\frac{\hbar^2}{2m} \nabla^2 \psi \quad \text{free part.}$$

time and space derivative enter
differently in the Schrödinger eq.

Chapter 1

Magnetism in Atoms

Lecture: 26.09.2022

1.1 Magnetism before quantum physics

The fact that some solids attract or repel other materials by virtue of what nowadays would be called magnetic interaction was already evident to mankind in ancient times. One of the first references to the magnetic properties of magnetite (Fe_3O_4 , lodestone) was reported by the Greek philosopher Thales of Miletus in the 6th century BC. The name “magnet” probably comes from the lodestones found in Magnesia. Though the use of lodestone in compasses for navigation in China dates back to Middle Ages, humans kept perceiving magnetism as a kind of magic till the 19th century, when a clear relationship between this phenomenon and moving electric charges was established^[1]. This observation opened the door to the modern understanding of magnetism in solids in terms of self organization of individual atomic magnetic moments.

The first important observations toward a rationalization of magnetism and its origin came with the experiments of Ampère and Oersted, in the early decades of the 19th century. They demonstrated that i) a current is able to influence a magnetic needle (Oersted) and ii) a mechanical force exists between two wires in which an electric current circulates (Ampère). Later, Faraday completed our knowledge of magnetic field by discovering that time dependent magnetic fields can create an electromotive force. The current

¹ *The Theory of Magnetism*, D. C. Mattis – Harper’s physics series – Harper & Row (New York, Evanston, and London, 1965).

description of electromagnetic fields was provided by J. C. Maxwell through the well-known set of equations named after him. The origin of magnetism in matter – precisely of magnetic moments – was more debated: Ampère postulated that magnetism in atoms originates from the existence of closed atomic-sized currents. Poisson and later Maxwell, instead, ascribed the origin of atomic magnetic moments to magnetic charges that appear always coupled into dipoles. Today we know that part of the magnetism of atoms is due to the motion of electrons about the nuclei. This orbital contribution is in line with the intuition of elementary currents put forward by Ampère. Another contribution is given by the *intrinsic* magnetic moment of electrons (spin contribution).

We will start our course showing that electrons *must* have an intrinsic magnetic moment in order that their quantum-mechanical description fulfills the same invariance properties as the Maxwell equations. Moving further in the course, you will appreciate how magnetic materials have been used across the years as benchmark to test our understanding of fundamental achievements in theoretical physics: from quantum mechanics to the theory of critical phenomena.

1.2 The Dirac equation and the electron spin

Lorentz covariance

While Newton's laws of classical mechanics are invariant under Galilean transformations, Maxwell equations are not. The latter are rather invariant (covariant) w.r.t. Lorentz transformations². If \vec{r} and t are the position vector and the time measured in a certain reference frame, in a second reference frame that moves with velocity

$$\vec{v} = c \tanh \alpha \hat{n} \quad (1.1)$$

w.r.t. the first one they transform as

$$\begin{cases} \vec{r}' = \vec{r} + [(\cosh \alpha - 1) (\vec{r} \cdot \hat{n}) - \sinh \alpha (ct)] \hat{n} \\ ct' = \cosh \alpha (ct) - \sinh \alpha (\vec{r} \cdot \hat{n}) \end{cases} \quad (1.2)$$

with c being the light velocity and $\alpha = \operatorname{atanh}(v/c)$. The important take of the Lorentz transformation of coordinates in Eq. (1.2) is that it mixes up

²The Galilean group encompasses Galilean transformations plus rotations and translations; the Poincaré group encompasses Lorentz transformations (boosts) plus rotations and translations.

check that for $v/c \ll 1$ (1.3) $\rightarrow \begin{cases} x' = x - vt \\ t' = t \end{cases}$ (like Galilean tr.)
Hint: in that limit $\frac{v}{c} = \tanh(\alpha) \approx \alpha \approx \sinh(\alpha)$

CHAPTER 1. MAGNETISM IN ATOMS

3

time and space coordinates. For simplicity we limit ourselves to a Lorentz boost along the x direction in which case Eq. (1.2) simplifies as

$$\begin{cases} x' = \cosh \alpha x - \sinh \alpha (ct) \\ ct' = \cosh \alpha (ct) - \sinh \alpha x. \end{cases} \quad (1.3)$$

Q limit $v/c \ll 1$

It is straightforward to verify that the difference of squared spacetime intervals is the same in both reference frames

$$(c\Delta t)^2 - (\Delta x)^2 = (c\Delta t')^2 - (\Delta x')^2. \quad (1.4)$$

In general, spacetime intervals

$$(\Delta s)^2 = (c\Delta t)^2 - (\Delta \vec{r})^2 \quad (1.5)$$

are conserved across frames related by a Lorentz transformation. This leads to the introduction of the 4-vector formalism

$$x^\mu = (ct, \vec{r}) \quad (1.6)$$

with the metric tensor

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (1.7)$$

so that

$$(\Delta s)^2 = g^{\mu\nu} \Delta x^\mu \Delta x^\nu = \Delta x^\mu \Delta x_\mu. \quad (1.8)$$

The invariant spacetime interval Δs is equal to the product of the speed of light by the *proper time* τ , which fulfills the property

$$\Delta \tau = \gamma^{-1} \Delta t \quad \text{with} \quad \gamma^{-2} = 1 - \frac{v^2}{c^2}. \quad (1.9)$$

With these few concepts at hand, one can proceed defining the 4-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{dt}{d\tau} (c, \vec{v}) = \gamma (c, \vec{v}) \quad (1.10)$$

and from this the 4-momentum

$$p^\mu = m u^\mu = m \gamma (c, \vec{v}) \quad (1.11)$$

Q12.

in standard 3D space, rotations conserve $(\Delta \vec{r})^2$
 Here Lorentz transform. conserve $(\Delta s)^2 = c^2 \Delta t^2 - (\Delta \vec{r})^2$

which has conserved length $p^\mu p_\mu = m^2 c^2$ (using the metric tensor in Eq. (1.7)). Since energy and momentum in restricted relativity are defined as

$$\begin{aligned}\vec{p} &= \gamma m \vec{v} \\ E &= \gamma m c^2\end{aligned}\tag{1.12}$$

the 4-momentum can also be written as follows

$$p^\mu = \left(\frac{E}{c}, \vec{p} \right)\tag{1.13}$$

which leads to the famous energy-momentum relation

$$p^\mu p_\mu = \frac{E^2}{c^2} - (\vec{p})^2 = m^2 c^2 \quad \Rightarrow \quad E^2 = (c\vec{p})^2 + m^2 c^4.\tag{1.14}$$

Using the 4-vector formalism Maxwell equations can be written in an elegant and compact way, in which they are manifestly covariant, i.e., invariant under Lorentz transformations. We will limit ourselves to the few ideas exposed so far because they suffice to describe the early attempts to develop a quantum-mechanical theory of the atom compatible with restricted relativity.

Early attempts

Starting from the quantum Hamiltonian of a non-relativistic free particle

$$\mathcal{H} = \frac{\hat{\mathbf{p}}^2}{2m}\tag{1.15}$$

the mapping of \mathcal{H} and the momentum to operators

$$\begin{aligned}\mathcal{H} &\rightarrow i\hbar \frac{\partial}{\partial t} \\ \hat{\mathbf{p}} &\rightarrow -i\hbar \vec{\nabla}\end{aligned}\tag{1.16}$$

leads to the Schrödinger equation for a free particle

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi.\tag{1.17}$$

Identifying E in Eq. (1.14) with \mathcal{H} , it seems natural to assume

$$\mathcal{H} = \sqrt{(c\vec{p})^2 + m^2 c^4}\tag{1.18}$$

as the Hamiltonian of a relativistic free particle and accordingly

$$i\hbar \frac{\partial \psi}{\partial t} = \sqrt{-\hbar^2 c^2 \nabla^2 + m^2 c^4} \psi \quad (1.19)$$

Problem:
what is the
sqrt of
operators?

as the relativistic equivalent of the Schrödinger equation (1.17). The reader is immediately faced with the problem of handling the square-root operator, as scientists in the 1920s were. One way to overcome this problem is to square operators on both sides of the equation, which yields the so-called Klein-Gordon equation

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = (-\hbar^2 c^2 \nabla^2 + m^2 c^4) \psi. \quad (1.20)$$

Plane waves

$$\psi = N e^{i\vec{x}\cdot\vec{p}/\hbar} \quad (1.21)$$

are solutions to the Klein-Gordon equation with eigenvalues

$$E = \pm \sqrt{(cp)^2 + m^2 c^4} \quad (1.22)$$

Such eigenvalues indeed fulfill the energy-momentum relation in Eq. (1.14). However, about a century ago it was not obvious how to interpret negative-energy solutions. Nowadays we know that they are associated with antiparticles. The final blow for this approach to a relativistic formulation of quantum-mechanics is the fact that the *probability density* (modulus of the wave function) obtained from the solutions to the Klein-Gordon equation is not positive defined.

The Dirac approach

Following the historical path taken by Paul F. Dirac in 1928, we list the requirements that a relativistically covariant version of the Schrödinger equation should fulfill:

- for the free-particle case its eigenvalues should be the same as the Klein-Gordon Eq. (1.20);
- time and space derivatives should enter that equation similarly;
- its wave function should correspond to a positive-defined probability density.

K-G $\frac{\partial^2}{\partial t^2}$ ∇^2

Dirac $\frac{\partial}{\partial t}$ ∇ $\frac{\partial}{\partial x}$

The first requirement is trivially fulfilled by the Klein-Gordon equation; the second one as well because just second derivatives of time and space coordinates appear; the third requirement is not fulfilled. To circumvent this problem – without discarding the positive aspects of the Klein-Gordon equation – Dirac proposed to replace the Schrödinger equation with a set of N -coupled equations *linear* both in time and space coordinates. Formally, he represented the wave function with a vector (of functions!)

$$\Psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{bmatrix} \quad (1.23)$$

which is solution of the equation

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{bmatrix} = -i\hbar \left(\underbrace{\alpha_x}_{\partial_x} \hat{p}_x + \underbrace{\alpha_y}_{\partial_y} \hat{p}_y + \underbrace{\alpha_z}_{\partial_z} \hat{p}_z + mc^2 \beta \right) \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{bmatrix} \quad (1.24)$$

with $\alpha_x, \alpha_y, \alpha_z, \beta$ being $N \times N$ matrices. The requirement that the spectrum of the eigenvalues of the Dirac equation be the same as for the Klein-Gordon equation implies that these four matrices must obey the following algebra:

$$\begin{aligned} \alpha_h \alpha_k + \alpha_k \alpha_h &= 2\mathbb{I} \text{ (if } k = h), \quad \mathbf{0} \text{ (otherwise)} \\ \alpha_k \beta + \beta \alpha_k &= \mathbf{0} \\ \alpha^2 = \beta^2 &= \mathbb{I} \end{aligned} \quad (1.25)$$

where \mathbb{I} is the $N \times N$ identity matrix and $\mathbf{0}$ the $N \times N$ null matrix. From this algebra it follows that eigenvalues of the α_h , ($h = x, y, z$) and β matrices can only be ± 1 and their trace has to be zero. Both requirements can only be fulfilled if N is even. For $N = 2$ we know that only 3 independent matrices with vanishing trace can be built, i.e., the Pauli matrices. The smallest dimension in which the algebra of Eqs. (1.25) can be realized is thus $N = 4$. A possible 4×4 representation of the α_h , ($h = x, y, z$) and β matrices reads

$$\alpha_h = \begin{pmatrix} \mathbf{0} & \hat{\sigma}_h \\ \hat{\sigma}_h & \mathbf{0} \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & -\mathbb{I} \end{pmatrix} \quad (1.26)$$

where $\hat{\sigma}_h$ are the (2×2) Pauli matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.27)$$

$$\text{Tr}(\underline{\alpha}_k) = 0$$

$$\underline{\alpha}_k^2 = \mathbb{1}$$

\Rightarrow

$$N = \cancel{2}, 4, 6, 8, \dots$$

$$\text{eig.}(\underline{\alpha}_k) = \pm 1$$

$$\underline{\alpha}_k = \begin{pmatrix} \ominus & & & \hat{\sigma}_k \\ & \ddots & & \\ & & \oplus & \\ \sigma_k & & & \ominus \\ \hline & & & \\ & & & \\ & & & \\ & & & \end{pmatrix} \quad \underline{\beta}_k = \begin{pmatrix} \hat{\sigma}_k & & & \\ & \mathbb{1}_{(2 \times 2)} & & \\ & & \mathbb{1}_{(2 \times 2)} & \\ & & & -\mathbb{1}_{(2 \times 2)} \end{pmatrix}$$

Electron at rest

$$\mathcal{H} \psi^1 = e^{-i} m_e c^2 \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathcal{H} \psi^1 = \underbrace{m_e c^2}_{> 0 \text{ electron}} \psi^1 \quad \text{same for } \psi^2$$

$$\mathcal{H} \psi^4 = e^{+i} m_e c^2 \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} =$$

$$\mathcal{H} \psi^4 = \underbrace{-m_e c^2}_{< 0} \psi^4 \quad \text{same for } \psi^3$$

positron

For an electron at rest with mass m_e the Dirac equation (1.24) reduces to

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} = m_e c^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} \quad (1.28)$$

whose solutions are given by

$$\Psi^1 = e^{-\frac{i m_e c^2 t}{\hbar}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \Psi^2 = e^{-\frac{i m_e c^2 t}{\hbar}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (1.29)$$

$$\Psi^3 = e^{\frac{i m_e c^2 t}{\hbar}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \Psi^4 = e^{\frac{i m_e c^2 t}{\hbar}} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (1.30)$$

The first two eigenfunctions correspond to positive energy (electron), while the second two to negative energy (positron).

The emergence of spin in the low-energy limit

We consider now the behavior of an electron with charge q_e and mass m_e in a static electromagnetic field at low energy, namely much smaller than the particle rest energy $m_e c^2$. The non-relativistic Hamiltonian is obtained adding the electrostatic energy term to the free-particle Hamiltonian and replacing the momentum according to the minimal coupling prescription

$$\hat{\mathbf{p}} \rightarrow \hat{\mathbf{p}} - q_e \vec{A}, \quad (1.31)$$

which yields

$$\mathcal{H} = \frac{1}{2m_e} \left(\hat{\mathbf{p}} - q_e \vec{A} \right)^2 + q_e \phi \quad (1.32)$$

The Dirac equation acquires similar terms. As we are interested in the non-relativistic limit, it is convenient to split the four components of the wave function into two components ($\tilde{\varphi}_1, \tilde{\varphi}_2$) associated with the positive-energy solutions (electron) and two components ($\tilde{\chi}_1, \tilde{\chi}_2$) associated with the negative-energy solutions (positron). The corresponding Dirac equation reads

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \\ \tilde{\chi}_1 \\ \tilde{\chi}_2 \end{bmatrix} = c \hat{\boldsymbol{\sigma}} \cdot \left(\hat{\mathbf{p}} - q_e \vec{A} \right) \begin{bmatrix} \tilde{\chi}_1 \\ \tilde{\chi}_2 \\ \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \end{bmatrix} + q_e \phi \begin{bmatrix} \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \\ \tilde{\chi}_1 \\ \tilde{\chi}_2 \end{bmatrix} + m_e c^2 \begin{bmatrix} \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \\ -\tilde{\chi}_1 \\ -\tilde{\chi}_2 \end{bmatrix} \quad (1.33)$$

Note that in this notation we have already applied the α_h , ($h = x, y, z$) and β matrices: the first ones swap the position of the $(\tilde{\varphi}_1, \tilde{\varphi}_2)$ and $(\tilde{\chi}_1, \tilde{\chi}_2)$ components and bring the $\hat{\sigma}$ operators; β just brings a minus sign in front of the $(\tilde{\chi}_1, \tilde{\chi}_2)$ components.

In the low-energy limit we are considering here – in which $m_e c^2$ is much larger than any other energy scale in the problem – the wave function can be factorized into a part varying slowly with time and in a fast-varying (with frequency $m_e c^2 / \hbar$) part:

$$\begin{bmatrix} \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \\ \tilde{\chi}_1 \\ \tilde{\chi}_2 \end{bmatrix} = e^{-\frac{i m_e c^2 t}{\hbar}} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \chi_1 \\ \chi_2 \end{bmatrix} \quad (1.34)$$

Defining

$$\hat{\pi} = \hat{\mathbf{p}} - q_e \vec{A} \quad (1.35)$$

the slow-varying parts of the wave function should fulfill the equation

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \chi_1 \\ \chi_2 \end{bmatrix} = c \hat{\sigma} \cdot \hat{\pi} \begin{bmatrix} \chi_1 \\ \chi_2 \\ \varphi_1 \\ \varphi_2 \end{bmatrix} + q_e \phi \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \chi_1 \\ \chi_2 \end{bmatrix} - 2m_e c^2 \begin{bmatrix} 0 \\ 0 \\ \chi_1 \\ \chi_2 \end{bmatrix} \quad (1.36)$$

For kinetic energies and field-interaction energies much smaller than $m_e c^2$ each equation for the two χ components can be approximated as

Approx.

$$0 = c \hat{\sigma} \cdot \hat{\pi} \varphi_r - 2 m_e c^2 \chi_r \Rightarrow \chi_r = \frac{\hat{\sigma} \cdot \hat{\pi}}{2 m_e c} \varphi_r \quad (1.37)$$

with $r = 1, 2$. Note that this equation implies that the amplitude of each χ_r be roughly v/c times smaller than the corresponding φ_r component. Equation (1.37) can be substituted into the equation of the φ_r components to obtain

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \left[\frac{(\hat{\sigma} \cdot \hat{\pi})(\hat{\sigma} \cdot \hat{\pi})}{2m_e} + q_e \phi \right] \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} \quad (1.38)$$

Using the properties of the Pauli matrices the first term on the r.h.s. can be

written in a more familiar way³

$$\begin{aligned}
 (\hat{\sigma} \cdot \hat{\pi})(\hat{\sigma} \cdot \hat{\pi}) &= \hat{\pi}^2 + i\hat{\sigma} \cdot \hat{\pi} \times \hat{\pi} = \\
 &= (\hat{\mathbf{p}} - q_e \vec{A})^2 + i\hat{\sigma} \cdot (\hat{\mathbf{p}} - q_e \vec{A}) \times (\hat{\mathbf{p}} - q_e \vec{A}) \\
 &= (\hat{\mathbf{p}} - q_e \vec{A})^2 + i^2 q_e \hbar (\vec{\nabla} \times \vec{A} + \vec{A} \times \vec{\nabla}) \\
 &= (\hat{\mathbf{p}} - q_e \vec{A})^2 - q_e \hbar \hat{\sigma} \cdot \vec{B}
 \end{aligned} \tag{1.39}$$

it actually means
 $\vec{B} \varphi_r = (\vec{\nabla} \times \vec{A}) \varphi_r$

to obtain the famous Pauli equation

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \left[\frac{(\hat{\mathbf{p}} - q_e \vec{A})^2}{2m_e} - \frac{q_e \hbar}{2m_e} \hat{\sigma} \cdot \vec{B} + q_e \phi \right] \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} \tag{1.40}$$

For a uniform magnetic field the vector potential can be written as⁴ $\vec{A} = (\vec{B} \times \mathbf{r})/2$. After some elementary passages the Pauli equation takes the form

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \left[\frac{\hat{\mathbf{p}}^2}{2m_e} - \frac{q_e \hbar}{2m_e} (\hat{\mathbf{l}} + 2\hat{\mathbf{s}}) \cdot \vec{B} + \frac{q_e^2}{8m_e} (\vec{B} \times \hat{\mathbf{r}})^2 + q_e \phi \right] \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} \tag{1.41}$$

Kin Zeeman es pot
 diamagn

where $\hat{\mathbf{l}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}/\hbar$ is the orbital momentum and $\hat{\mathbf{s}} = \hat{\sigma}/2$ is the particle spin. We found it convenient to indicate the position operator with $\hat{\mathbf{r}}$, for coherence of notations throughout the lecture notes. Moreover, we adopt the convention of expressing angular momenta in \hbar units.

The reader will recognize in the term linear in \vec{B} the *Zeeman* interaction. This interaction consists of an orbital and a spin contribution. The first contribution does not vanish for an electron that has a finite component of the angular momentum along the applied field, assumed as z direction, e.g., an electron occupying an atomic orbital with $l = 1$ and $m = -1$ (see later). Due to the negative charge of the electrons, the orbital part of the Zeeman interaction favors quantum states for which \hat{l}^z has the minimal expectation value. In a semiclassical picture one would say that the Zeeman interaction favors antiparallel alignment between the angular momentum \vec{l} and \vec{B} . As for the spin contribution to the Zeeman energy, we would like to remark that it

³The reader has to interpret this operator as applied to each component of the wavefunction φ_r (with $r = 1, 2$). In particular, one has $\vec{\nabla} \times (\vec{A}\varphi_r) = \varphi_r(\vec{\nabla} \times \vec{A}) + (\vec{\nabla}\varphi_r) \times \vec{A} = \varphi_r(\vec{\nabla} \times \vec{A}) - \vec{A} \times (\vec{\nabla}\varphi_r)$ and therefore $(\vec{\nabla} \times \vec{A} + \vec{A} \times \vec{\nabla}) \varphi_r = \varphi_r(\vec{\nabla} \times \vec{A}) - \vec{A} \times (\vec{\nabla}\varphi_r) + \vec{A} \times (\vec{\nabla}\varphi_r) = \varphi_r(\vec{\nabla} \times \vec{A}) = \varphi_r \vec{B}$.

⁴This choice of \vec{A} corresponds to the Coulomb gauge $\vec{\nabla} \cdot \vec{A} = 0$.

See Assignment for the Classical equivalent

requirement of having a version Sch. eq. compatible with restricted relativity

emerged spontaneously from the Dirac equation taking the low-energy limit. Note that the number 2 in front of the spin operator $\hat{\mathbf{s}}$ correctly accounts for the gyromagnetic factor of the spin being twice the gyromagnetic factor of the orbital momentum (up to 0.1%). Henceforth, we will express the *Zeeman* interaction of an electron (with electric charge $q_e = -e$) as

$$\mathcal{H}_Z = \mu_B (\hat{\mathbf{l}} + 2\hat{\mathbf{s}}) \cdot \vec{B} = -\hat{\boldsymbol{\mu}} \cdot \vec{B}, \quad (1.42)$$

where $\mu_B = e\hbar/(2m_e) = 9.274 \dots \times 10^{-24}$ J/T = $5.788 \dots \times 10^{-2}$ meV/T is the Bohr magneton and $\hat{\boldsymbol{\mu}} = -\mu_B (\hat{\mathbf{l}} + 2\hat{\mathbf{s}})$ is the magnetic-moment operator. As anticipated in the first section, this magnetic moment consists of an orbital and an intrinsic (spin) contribution.

The term in the Pauli equation (1.41) proportional to B^2 is much weaker than the Zeeman interaction and it is only observable if the total angular momentum of electrons in one atom exactly vanishes. This term is responsible for the so-called *diamagnetic* contribution⁵.

The comparison between the Schrödinger equation and the Pauli equation (1.41) highlights that the latter includes the spin contribution explicitly. However, as spatial and spin d.o.f. are not directly coupled, the solutions of the Pauli equation can be written as

One can understand

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \text{ as } \begin{pmatrix} \varphi^\uparrow \\ \varphi^\downarrow \end{pmatrix} \quad \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \psi_{\text{Schr.}}(\underline{r}) |spin\rangle \quad (1.43)$$

with $|spin\rangle$ indicating the spin part of the wave function. The spatial part $\psi_{\text{Schr.}}(\underline{r})$ is a solution of the Schrödinger equation.

In the next section we will see that the *spin-orbit* interaction, instead, directly couples spatial and spin d.o.f.. Therefore, factorizing spatial and spin parts in the wave function is not accurate when spin-orbit interaction is taken into account.

Spin-orbit coupling

The calculation reproduced in the previous section is not the only possible way to take the low-energy limit of the Dirac equation for an electron. Foldy and Wouthuysen⁶ developed a systematic procedure to decouple the positive-

⁵From the expression for the diamagnetic energy one can define an effective diamagnetic moment per atom as $\mu_d = -\partial E_d/\partial B$ amounting to $\sim -10^{-11}$ eV/T. Thus, the diamagnetic moment per electron points antiparallel to the applied magnetic field and is far smaller than the paramagnetic moment.

⁶Leslie L. Foldy and Siegfried A. Wouthuysen, Phys. Rev. **78**, 29 (1950).

QED
Kuch

and negative-energy solutions by expanding the Hamiltonian in powers of (energy terms)/ $(m_e c^2)$ along with the successive application of canonical transformations. The resulting low-energy Hamiltonian for the positive-energy solutions (electrons) reads:

$$\mathcal{H} = m_e c^2 + \frac{\hat{\mathbf{p}}^2}{2m_e} - \frac{q_e \hbar}{2m_e} (\hat{\mathbf{1}} + 2\hat{\mathbf{s}}) \cdot \vec{B} + \frac{q_e^2}{8m_e} (\vec{B} \times \hat{\mathbf{r}})^2 + q_e \phi$$

rest. kin Zeeman diam. es pot.

$$- i \frac{q_e \hbar^2}{4m_e^2 c^2} \hat{\mathbf{s}} \cdot \vec{\nabla} \times \vec{E} - \frac{q_e \hbar}{2m_e^2 c^2} \hat{\mathbf{s}} \cdot \vec{E} \times \hat{\mathbf{p}} \quad \text{SO coupling (1.44)}$$

will be neglected throughout the course

$$- \frac{\hat{\mathbf{p}}^4}{8m_e^3 c^2} - \frac{q_e \hbar^2}{8m_e^2 c^2} \vec{\nabla} \cdot \vec{E}$$

The first row of the equation above comprises the terms already encountered in the Pauli equation plus the rest-mass energy. The two terms in the second row correspond to the spin-orbit coupling. For static electromagnetic fields one has $\vec{\nabla} \times \vec{E} = 0$ and, accounting for the sign of the electron charge $q_e = -e$, the spin-orbit interaction takes the form

$$\mathcal{H}_{so} = \frac{e\hbar}{2m_e^2 c^2} \hat{\mathbf{s}} \cdot \vec{E} \times \hat{\mathbf{p}}. \quad (1.45)$$

In the next sections we will specialize the field \vec{E} to the electron-nucleus Coulomb interaction, possibly corrected with the screening effect of the other electrons in the atom (see Eq. (??)). However, the validity of Eq. (1.45) goes beyond the atom and it applies, for instance, to the spin of electrons in a semiconductor that experience an electric field generated by an interface or applied externally (Rashba spin-orbit coupling). One can interpret the spin-orbit Hamiltonian (1.45) as a Zeeman interaction of the spin magnetic moment $\hat{\boldsymbol{\mu}} = -2\mu_B \hat{\mathbf{s}}$ with an effective magnetic field $\vec{B} = (\vec{E} \times \vec{v})/(2c^2)$ experienced by the electron moving with velocity $\vec{v} = \vec{p}/m_e$ in a field \vec{E} . Classically, this effective magnetic field would be two times larger. The discrepancy between experimentally observed splitting of energy levels (fine structure) and the result of the classical (non-relativistic) calculation animated a long debate at the beginning of the twentieth century.

When the electric field is the gradient of a *spherically symmetric* electrostatic potential $\phi(r)$, like for electron-nucleus Coulomb interaction, it takes the form:

$$\vec{E} = -\nabla\phi(r) = -\frac{1}{r} \frac{\partial\phi(r)}{\partial r} \hat{\mathbf{r}}. \quad (1.46)$$

Replacing this term into Hamiltonian (1.45) the familiar form of the spin-orbit interaction is obtained

$$\mathcal{H}_{so} = -\frac{e\hbar}{2m_e^2 c^2} \frac{1}{r} \frac{\partial\phi(r)}{\partial r} \hat{\mathbf{s}} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{p}}) = -\frac{e}{2} \left(\frac{\hbar}{m_e c} \right)^2 \frac{1}{r} \frac{\partial\phi(r)}{\partial r} \hat{\mathbf{s}} \cdot \hat{\mathbf{l}} = \xi_{so} \hat{\mathbf{s}} \cdot \hat{\mathbf{l}} \quad (1.47)$$

QM. $\hbar \hat{\mathbf{l}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$
 cl. Phys $\vec{\mathbf{l}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$



for the electron: $\xi_{so} > 0$



QM.

$$\hat{j} = \hat{l} + \hat{s}$$

$$\hat{j}^2 = (\hat{l} + \hat{s})^2 = \hat{l}^2 + \hat{s}^2 + 2\hat{s} \cdot \hat{l}$$

$$\hat{s} \cdot \hat{l} = \frac{1}{2} (\hat{j}^2 - \hat{l}^2 - \hat{s}^2)$$

$$\text{H}_{so} |j m_j\rangle = \frac{\hbar}{2} \xi_{so} [j(j+1) - l(l+1) - s(s+1)]$$

for (s, l) given so coupling favors multiplet with smaller j

where we used that $\underline{\hat{r}} \times \hat{\mathbf{p}} = \hbar \hat{\mathbf{l}}$. Since $\partial\phi(\underline{r})/\partial r$ is negative for electrons, the quantity

$$\xi_{\text{so}} = -\frac{e}{2} \left(\frac{\hbar}{m_e c} \right)^2 \frac{1}{r} \frac{\partial\phi(\underline{r})}{\partial r}, \quad (1.48)$$

averaged over the spatial wave function, is positive. Therefore, the spin-orbit interaction favors antiparallel alignment between the spin and the orbital angular momenta of an electron. This can be seen more formally defining the sum of the spin and orbital momentum as $\hat{\mathbf{j}} = \hat{\mathbf{s}} + \hat{\mathbf{l}}$ and squaring both sides of the equation. The scalar product $\hat{\mathbf{s}} \cdot \hat{\mathbf{l}}$ can then be expressed in terms of the squares of the three angular momenta:

$$\hat{\mathbf{s}} \cdot \hat{\mathbf{l}} = \frac{1}{2}(\hat{\mathbf{j}}^2 - \hat{\mathbf{s}}^2 - \hat{\mathbf{l}}^2), \quad (1.49)$$

Inserting this result into Eq. (1.47) gives

$$\xi_{\text{so}} = \frac{1}{2} \xi_{\text{so}} (\hat{\mathbf{j}}^2 - \hat{\mathbf{s}}^2 - \hat{\mathbf{l}}^2). \quad (1.50)$$

which manifestly favors minimal values of the total angular momentum j .

The terms in the last row of Eq. (1.44) – proportional to $\hat{\mathbf{p}}^4$ and to $\vec{\nabla} \cdot \vec{E}$ (the so-called Darwin term) – are very small and will be neglected henceforth.