Spin-orbit coupling   V. M. STOJANOVIĆ

In atomic physics, the spin-orbit (SO) interaction originates from a nonrelativistic approximation to the Dirac equation. The resulting Pauli SO term reads

$$H_{SO} = -\frac{\hbar}{(2m_0 c)^2}\,\vec{\sigma}\cdot\vec{p}\times\nabla V_0 \, ,$$

where m_0 is the free-electron mass, $\vec{p}$ is the momentum operator, V_0 is the Coulomb potential of the atomic core, and $\vec{\sigma}\equiv(\sigma_x,\sigma_y,\sigma_z)$ the vector of Pauli spin matrices. It is an established fact that atomic spectra are accurately described only if SO coupling is taken into account.

Now we will explore some of the consequences of SO coupling for solid-state systems. We know that the motion of electrons in a crystalline solid is described by the energy-band dispersions $E_n(\vec{k})$, where n is the band index (describing "orbital motion" of electrons) and $\vec{k}$ the quasimomentum.

As it turns out, SO coupling stronly affects the energy band structure. For instance, in the semiconductor GaAs it leads to a splitting of the highest-lying valence band.

GaAs
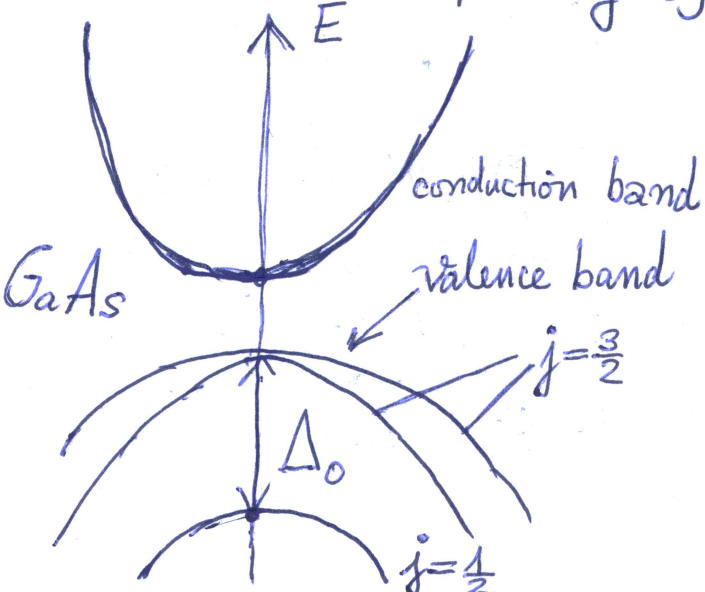
The electron states at the valence-band edge are _p-like_ (orb. angular momentum $l=1$). With SO coupling included, we get states with $j=\frac{3}{2}$ and $j=\frac{1}{2}$ with the splitting Δ_0 (the SO gap).

It was Dresselhaus who first anticipated that the SO coupling may have important consequences for the band structure of semiconductors without the bulk inversion symmetry. The semiconductors such as Si and Ge have diamond structure which is inversion-symmetric. On the other, GaAs is an example of semiconductors with zinc-blende structure which do not have a center of inversion! In such semiconductors we can have a spin splitting of electron and hole states at $\vec{k} \neq 0$ even in the absence of an external magnetic field!

Why there is no SO-induced splitting in the inversion-symmetric case?

The time-reversal symmetry (which is not violated upon inclusion of SO coupling) implies that $E_n(\vec{k}, \uparrow) = E_n(-\vec{k}, \downarrow)$ (Kramers' degeneracy). In systems with inversion symmetry

$$E_n(\vec{k}, \uparrow) = E_n(-\vec{k}, \uparrow) \quad , \quad E_n(\vec{k}, \downarrow) = E_n(-\vec{k}, \downarrow).$$

Combining the last results, we obtain

$$E_n(\vec{k}, \uparrow) = E_n(-\vec{k}, \downarrow) = E_n(\vec{k}, \downarrow).$$

In other words, in inversion-symmetric systems the energy cannot depend on the spin even in the presence of SO coupling. Thus in semiconductors like Si and Ge we have a twofold degeneracy of the Bloch states at each $\vec{k}$.

clearly the spin splitting of the Bloch states in zinc-blende semiconductors must come from SO interaction, otherwise the spin degree of freedom would not know whether it was in an inversion-symmetric or an inversion-asymmetric structure.

Having discussed these general implications of SO coupling for the "orbital" motion of electrons in semiconductors, can we say something about its effect on the spin degree of freedom? We know that in an external magnetic field the spins undergo Zeeman splitting, which is given by $g\mu_B B$ (μ_B - Bohr magneton; B - external field). Here g is the so-called g-factor which normally is equal two: $g_0 = 2$ is the free-electron g-factor. As it turns out, in the presence of SO coupling this is not the case and the g-factor can be parameterized using the SO gap Δ_0!

Bulk band structure of semiconductors within the framework of the $\vec{\kappa} \cdot \vec{p}$ method

The simplest model for the band structure of a direct bandgap semiconductor close to the fundamental gap is an _isotropic parabolic dispersion_ for both the conduction (c) and the valence (v) band

$$E_{c,v}(\vec{k}) = \pm\left(\frac{E_g}{2} + \frac{\hbar^2 k^2}{2m^*_{c,v}}\right),$$

where E_g is the energy gap,
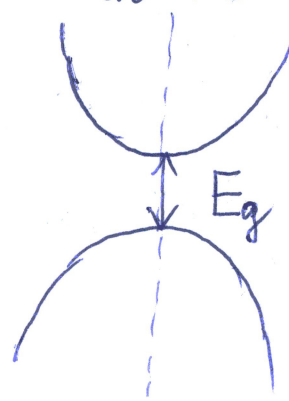
and m_c^*, m_v^* are the effective masses for the conduction and valence bands, respectively. In a more realistic model, one has to take into account the nonparabolicity, anisotropy, and the coupling between the heavy holes and the light holes.

The development of the $\vec{k} \cdot \vec{p}$ method starts by recasting the Schrödinger equation for Bloch electrons

$$\left(\frac{p^2}{2m_0} + V(\vec{r}) \right) \Psi_{v\vec{k}}(\vec{r}) = E_v(\vec{k}) \Psi_{v\vec{k}}(\vec{r})$$

into the form that involves their lattice-periodic part $U_{v\vec{k}}(\vec{r})$ of the Bloch functions [recall that $\Psi_{v\vec{k}}(\vec{r}) \equiv e^{i\vec{k}\cdot\vec{r}} U_{v\vec{k}}(\vec{r})$; we will here use "$v$" to denote orbital motion of electrons, instead of the usual Bloch-band index n] :

$$\left[\frac{p^2}{2m_0} + V(\vec{r}) + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} \vec{k} \cdot \vec{p} \right] |v\vec{k}\rangle = E_v(\vec{k}) |v\vec{k}\rangle ,$$

where $U_{v\vec{k}}(\vec{r}) \equiv \langle \vec{r} | v\vec{k} \rangle$ (i.e., the last equation was written in a representation-independent form).

In the presence of SO coupling, the lattice-periodic parts of the Bloch functions become spinors $|n\vec{k}\rangle$ with two components $|n\vec{k}\rangle = \begin{pmatrix} |\vec{k},\uparrow\rangle \\ |\vec{k},\downarrow\rangle \end{pmatrix}$ and the last Schrödinger equation becomes :

$$\left[\frac{p^2}{2M_0} + V(\vec{r}) + \frac{\hbar^2 K^2}{2M_0} + \frac{\hbar}{M_0} \vec{k} \cdot \vec{\pi} + \frac{\hbar}{(2M_0 C)^2} \vec{p} \cdot \vec{\delta} \times \nabla V \right] |n\vec{k}\rangle$$

$$= E_n(\vec{k}) |n\vec{k}\rangle, \quad \text{where} \quad \vec{\pi} \equiv \vec{p} + \frac{\hbar}{(2M_0 C)^2} \vec{\delta} \times \nabla V.$$

Note that in the presence of SO coupling the spin quantum number δ is not a good quantum number!

(n is here a common index for the orbital and the)
(spin degrees of freedom)

For any fixed quasimomentum $\vec{k}_0$, the sets of lattice periodic functions $\{|v\vec{k}_0\rangle\}$ and $\{|n\vec{k}_0\rangle\}$ provide complete orthonormal bases in the space of single-particle states. We will intentionally make use of the basis $\{|v, \vec{k}_0 = 0\rangle\}$ even in the presence of SO coupling, since in most cases it is convenient to treat this coupling as a perturbation! Then we have an expansion like

$$|n\vec{k}\rangle = \sum_{v', \delta' = \uparrow\downarrow} C_{nv'\delta'} |v'\delta'\rangle,$$

where $|v'\delta'\rangle \equiv |v', \vec{k} = 0\rangle \otimes |\delta'\rangle.$

The Schrödinger equation involving the states $|n\vec{k}\rangle$ can then be recast as:

$$\sum_{\nu'\delta'} \left\{ \left[E_{\nu'}(0) + \frac{\hbar^2 k^2}{2 m_0} \right] \delta_{\nu\nu'} \delta_{\delta\delta'} + \frac{\hbar}{m} \vec{k} \cdot \vec{P}_{\delta\delta'}^{\nu\nu'} \right.$$

$$\left. + \Delta_{\delta\delta'}^{\nu\nu'} \right\} C_{n\nu'\delta'}(\vec{k}) = E_n(\vec{k}) \, C_{n\nu\delta}(\vec{k}),$$

where $\quad P_{\delta\delta'}^{\nu\nu'} \equiv \langle \nu\delta \,|\, \vec{\pi} \,|\, \nu'\delta' \rangle \quad$ and

$$\Delta_{\delta\delta'}^{\nu\nu'} \equiv \frac{\hbar}{(2 m_0 c)^2} \langle \nu\delta \,|\, \vec{p} \cdot \vec{\delta} \times \nabla V \,|\, \nu'\delta' \rangle.$$

Note that in the spin-free case $P_{\delta\delta'}^{\nu\nu'} = P_{\nu\nu'} \delta_{\delta\delta'}$,

where $\quad P_{\nu\nu'} \equiv \langle \nu, \vec{k}=0 \,|\, \vec{p} \,|\, \nu', \vec{k}=0 \rangle.$

By solving the last eigenvalue problem one obtains
the dispersion $E_n(\vec{k})$ of the relevant bands. While
the last problem is in principle infinite-dimensional,
one is usually interested only in a handful of
bands (and often only in their dispersions in the
vicinity of $\vec{k}=0$ point!).

Let us consider some consequences for the band
structure of semiconductors in the spin-free case.
Using 2nd order perturbation theory one obtains

$$E_\nu(\vec{k}) = E_\nu(0) + \frac{\hbar^2 k^2}{2 m_\nu^*} \quad \left(\begin{array}{l} m_\nu^* - \text{effective mass} \\ \quad \text{in band } \nu \end{array} \right),$$

where $\dfrac{m_0}{m_\nu^*} = 1 + \dfrac{2}{m_0} \sum_{\nu'} \dfrac{P_{\nu\nu'}^2}{E_\nu(0) - E_{\nu'}(0)}.$

Assuming that the dominant contribution to the last sum for the conduction band comes from the valence band, and vice versa, we obtain

$$\boxed{\frac{M_0}{M_{c,v}^*} \approx \frac{2}{M_0} \frac{P_{cv}^2}{E_g}} \;,\; i.e., \; \boxed{M_{c,v}^* \propto E_g}.$$

Thus, to a good approximation the effective masses M_c^* and m_v^* are proportional to the fundamental band gap.

Two-dimensional electron gas (2DEG)
(read Bruus & Flensberg, Ch. 2.3.2)

Ability to confine electrons in 2D is important not only for studying fundamental physics, but also for electronic applications, since one can obtain extremely long mean free paths of $\sim 10 \mu m$.

What is nowadays used for this purpose are interfaces of lattice-matched materials such as GaAs and $Ga_x Al_{1-x} As$. The Fermi energy in wide-gap GaAlAs is higher than in GaAs. Therefore, electrons will migrate towards GaAs layer and leave behind positively charged ions (donors). This leads to a formation of a space-charge distribution and the ensuing electrostatic potential brings about a band-bending.

F_F

E_V E_F Z
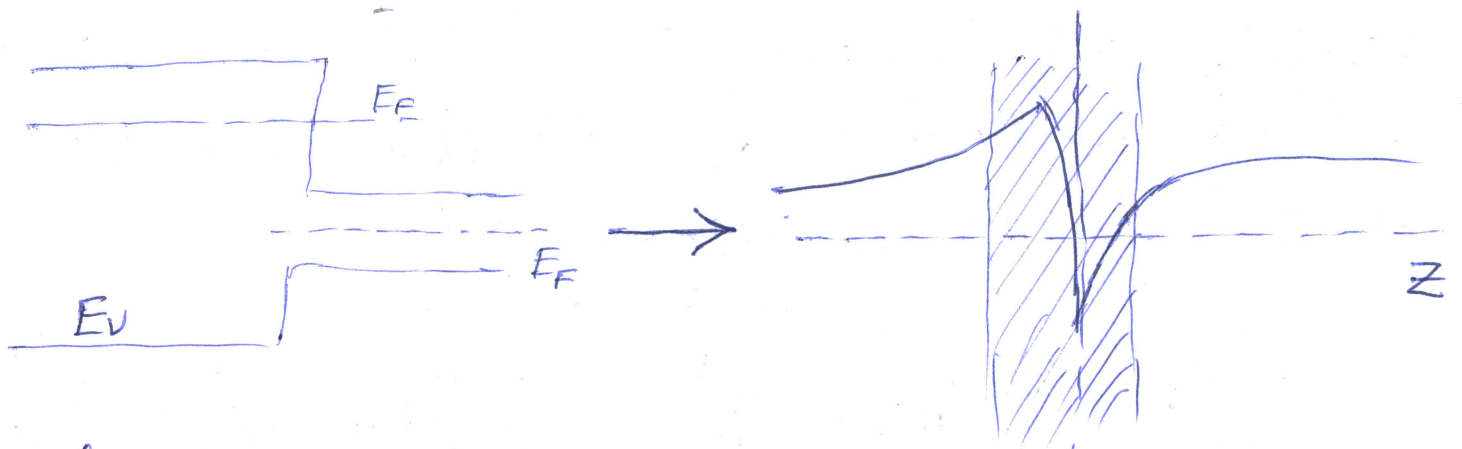
The carrier concentrations in the 2DEG (the x-y plane here) are typically between $2 \times 10^{11} cm^{-2}$ and $2 \times 10^{12} cm^{-2}$.

2DEG form in layers of thickness $\sim 100 \mathring{A}$

SO coupling in the 2DEG

The effective SO coupling is often written as

$$H_{SO} = \frac{\hbar}{M_*} (\vec{K}_{SO} \cdot \vec{\sigma}),$$ where here $\vec{K}_{SO}$ is some linear combination of electron-momentum components p_x and p_y; M_* is an effective mass, which is usually much smaller than the bare electron mass M_0.

There are two main types of SO interaction in a 2DEG.

A confining potential which is asymmetric under space inversion gives rise to Rashba's SO coupling

$$\boxed{H_R = \frac{\hbar k_R}{m} (P_y \sigma_x - P_x \sigma_y) \equiv \frac{\alpha}{\hbar} (P_y \sigma_x - P_x \sigma_y)}.$$

κ_R typically depends strongly on the electric field $\vec{E} = -\nabla V$ in the z-direction! Therefore the Rashba SO coupling can be modulated using an external electric field.

Another type of SO coupling (which historically was studied first) is <u>Dresselhaus-type</u> SO coupling. While it is normally cubic in the electron momentum, on the 2DEG it adopts the form

$$H_D = \frac{\hbar \kappa_D}{m} \left(P_x \hat{\sigma}_x - P_y \hat{\sigma}_y \right) = \frac{\beta}{\hbar} \left(P_x \hat{\sigma}_x - P_y \hat{\sigma}_y \right).$$

This type of SO coupling is due to inversion asymmetry of the bulk ~~band~~ semiconductor structure.

Strictly speaking, the third type of SO coupling $\propto \frac{\gamma}{\hbar} \left(\hat{\sigma}_x P_x P_y^2 - \hat{\sigma}_y P_y P_x^2 \right)$ also exists, but this so-called trilinear term is usually small and can be safely neglected.

What the Rashba and Dresselhaus-type SO couplings have in common is that they can cause a splitting of otherwise degenerate energy levels at each $\vec{k}$!

For instance, by diagonalizing the 2DEG Hamiltonian with Rashba SO coupling

$$H = \frac{p^2}{2M} + \frac{\alpha}{\hbar}(p_y \sigma_x - p_x \sigma_y) \quad \text{one}$$

obtains dispersions that look like this

free-electron gas

2DEG with Rashba spin-orbit coupling
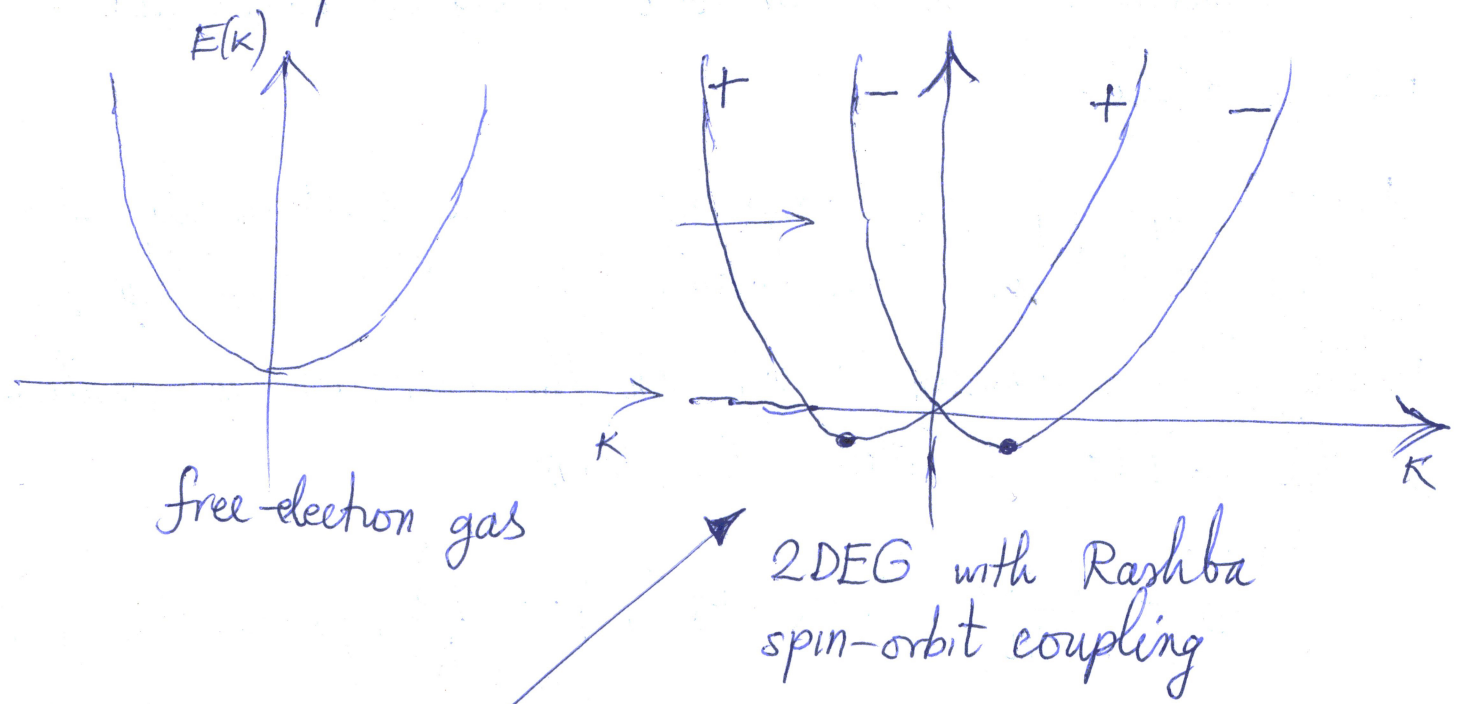
The "+" and "−" used to denote different branches underscore the fact that the spin is not a good quantum number!