Rashba spin-orbit coupling and spin precession in carbon nanotubes

A De Martino and R Egger§

Institut für Theoretische Physik, Heinrich-Heine-Universität, D-40225 Düsseldorf

Abstract. The Rashba spin-orbit coupling in carbon nanotubes and its effect on spin-dependent transport properties are analyzed theoretically. We focus on clean non-interacting nanotubes with tunable number of subbands N. The peculiar band structure is shown to allow in principle for Datta-Das oscillatory behavior in the tunneling magnetoresistance as a function of gate voltage, despite the presence of multiple bands. We discuss the conditions for observing Datta-Das oscillations in carbon nanotubes.

PACS numbers: 72.25.-b, 72.80.Rj, 73.63.Fg

Submitted to: *J. Phys.: Condens. Matter*

§ To whom correspondence should be addressed (egger@thphy.uni-duesseldorf.de)

1. Introduction

Spintronics in molecular conductors is a field attracting more and more attention, both from fundamental physics as well as from application-oriented material science [1]. Here the quantum-mechanical electronic spin is the central object controlling transport properties. For a conductor sandwiched between ferromagnetic leads, a different resistance can be observed depending on the relative orientation of the lead magnetizations. Quite often, the resistance is larger in the antiparallel configuration than in the parallel one, but sometimes also the reverse situation can be observed. It is useful to define the tunnel magnetoresistance (TMR), $\rho_t = (R_{AP} - R_P)/R_P$, as the relative difference between the corresponding resistances.

A particularly interesting material in that context is provided by carbon nanotubes (CNTs), see Refs. [2, 3] for general reviews. Quite a number of experimental studies concerning spin transport through individual multi- (MWNT) or single-walled (SWNT) nanotubes contacted by ferromagnetic leads have been reported over the past few years [4, 5, 6, 7, 8, 9, 10]. In particular, the experiments of the Basel group [9, 10] use thin-film PdNi alloys as ferromagnetic leads in order to contact either SWNTs or MWNTs, where the shape anisotropy and the geometry of the setup allow for the study of the spin-dependence of electrical transport. These experiments have revealed *oscillatory* behavior of the TMR as a function of the external gate voltage. Similar oscillations were predicted as a consequence of the gate-voltage-tunable Rashba spin-orbit (SO) interaction [11, 12] in a classic paper by Datta and Das some time ago [13]. Since Datta-Das oscillations have still not been observed experimentally so far, a thorough theoretical investigation of this effect in nanotubes is called for and provided here. Unfortunately, from our analysis below, we find that the weakness of SO couplings in nanotubes excludes an interpretation of these data in terms of the Datta-Das effect – they can, however, be explained in terms of quantum interference effects [10]. Nevertheless, we show that the presence of multiple bands in CNTs is not detrimental, and under certain circumstances, the effect may be sufficiently enhanced to be observable, e.g., by a tuning of the number of bands via external gates along the lines of Ref. [14]. In the original Datta-Das proposal [13], subband mixing was ignored so that different channels just add up coherently, but subband mixing has later been argued to spoil the effect [15, 16]. In CNTs, the special band structure requires a careful re-examination of the Datta-Das idea in this context, and we shall show that the arguments of Refs. [15, 16] do not necessarily apply here.

Recent theoretical studies of spin-dependent transport in CNTs have mainly focused on the single-channel limit, taking into account electron-electron interactions within the framework of the Luttinger liquid theory [17, 18, 19, 20, 21] (see also [22, 23, 24] for related discussions on interacting quantum wires with Rashba SO coupling). Here we confine ourselves to the noninteracting problem in order to not overly complicate the analysis, but study the many-band case and details of the band structure. Interactions can be taken into account within the Luttinger liquid approach at a later stage, and may enhance the effect of SO couplings [22, 25]. We shall also neglect disorder effects.

Mean free paths in high-quality SWNTs typically exceed 1μm, while in MWNTs this may be a more severe approximation for some samples. However, high-quality MWNTs with ultra-long mean free paths have also been reported recently [26].

The structure of this paper is as follows. In Sec. 2 we derive the Rashba spin-orbit hamiltonian from microscopic considerations. The resulting tight-binding SO hamiltonian will be studied at low energy scales in Sec. 3, where we derive its continuum form. In Sec. 4, the consequences with regard to Datta-Das oscillations in the TMR are analyzed. We shall always consider the zero-temperature limit, and (in most of the paper) put $\hbar = 1$.

2. Rashba spin-orbit coupling in nanotubes

We start by noting that transport effectively proceeds through the outermost shell of a MWNT only, such that we can take a single-shell model even when dealing with a MWNT. Experimentally and theoretically, it is understood that such a model works very well in good-quality MWNTs [2], essentially because only the outermost shell is electrically contacted and tunneling between different shells is largely suppressed [27, 28]. Naturally, a single-shell description is also appropriate for SWNTs, where we assume a sufficiently large radius R such that occupation of multiple subbands can be possible. (For a MWNT, R denotes the radius of the outermost shell.) Depending on the electrochemical potential μ (doping level), we then have to deal with N spin-degenerate bands. We assume full quantum coherence (no dephasing), so that the usual Landauer-Büttiker approach applies, and exclude external magnetic fields or electric field inhomogeneities, say, due to the electrodes. We proceed to derive the Rashba SO interaction, H_{so}, for this problem. Notice that this is different from the intrinsic atomic SO interaction discussed in Refs. [18, 29]. In particular, the SO coupling in Refs. [18, 29] vanishes in the limit of large radius, which is not the case for the Rashba SO coupling we discuss below. Though Ando's SO coupling [18] could straightforwardly be included in our analysis, being gate-voltage independent it could not change our conclusions relative to the gate-voltage dependent oscillations in the magnetoresistance and is neglected in what follows.

We first define a fixed reference frame $\mathcal{S} = \{\hat{Y}, \hat{Z}, \hat{X}\}$, with unit vector $\hat{X}$ pointing in the axis direction and $\hat{Z}$ perpendicular to the substrate on which the CNT is supposed to be located. Next we introduce a second, local reference frame $\mathcal{S}_i = \{\hat{\rho}_i, \hat{t}_i, \hat{X}\}$ relative to each lattice site $\vec{R}_i$ on the tube surface, where $\hat{\rho}_i$ and $\hat{t}_i$ are unit vectors along the local normal and tangential (around the circumference) directions at $\vec{R}_i$, respectively. Using polar coordinates in the plane transverse to the tube axis, the relation between $\mathcal{S}$ and $\mathcal{S}_i$ is given by

$$\hat{\rho}_i = \cos\varphi_i \hat{Y} + \sin\varphi_i \hat{Z}, \quad \hat{t}_i = -\sin\varphi_i \hat{Y} + \cos\varphi_i \hat{Z}. \tag{1}$$

The position vector of a given carbon atom can then be written as $\vec{R}_i = R\hat{\rho}_i + X_i\hat{X}$. For later convenience, we introduce also another reference frame. For each pair of sites

$\vec{R}_i$ and $\vec{R}_j$, we define

$$\vec{R}_{ij} = \vec{R}_i - \vec{R}_j \equiv X_{ij}\hat{X} + \vec{\rho}_{ij}, \tag{2}$$

and denote the direction perpendicular to $\hat{\rho}_{ij}$ and $\hat{X}$ as $\hat{\rho}^{\perp}_{ij}$. Then $\{\hat{\rho}^{\perp}_{ij}, \hat{\rho}_{ij}, \hat{X}\}$ constitutes a new local frame $\mathcal{S}_{ij}$, and one has

$$\begin{aligned} \hat{\rho}^{\perp}_{ij} &= \cos[(\varphi_i + \varphi_j)/2]\,\hat{Y} + \sin[(\varphi_i + \varphi_j)/2]\,\hat{Z}, \\ \hat{\rho}_{ij} &= -\sin[(\varphi_i + \varphi_j)/2]\,\hat{Y} + \cos[(\varphi_i + \varphi_j)/2]\,\hat{Z}. \end{aligned} \tag{3}$$

The $2p_z$ orbital at position $\vec{R}_i$ can then be represented as

$$\chi_i(\vec{r} - \vec{R}_i) = \alpha(\vec{r} - \vec{R}_i) \cdot \hat{\rho}_i e^{-\beta|\vec{r} - \vec{R}_i|}, \tag{4}$$

where $4\alpha = (2\pi a_0^5)^{-1/2}$, $\beta = (2a_0)^{-1}$, $a_0 = \hbar^2/me^2 = 0.53$Å is the Bohr radius, and m is the electron's mass. We introduce an index i on the orbital in order to keep track of the atom at which it is centered. The wavefunction (4) is expected to be highly accurate for not too small R, where hybridization with the sp^2 orbitals is negligible.

At large distances from the tube, external gates generally produce an electric field perpendicular to the tube axis and the substrate. As it has been shown in detail in previous works [30, 31], polarization effects of the CNT itself due to a transverse field result in a reduction of the externally applied field described by

$$E_0 = \frac{1}{1 + 2\alpha_{0yy}/R^2} E_{ext},$$

where α_{0yy} is the unscreened transverse static polarizability. Since α_{0yy} is approximately proportional to R^2, the factor in front of E_{ext} practically equals a constant, ≈ 0.2 [30]. Then, assuming homogeneity, the electric field due to the gate can be written as

$$\vec{E} = E_0 \hat{Z}, \tag{5}$$

which in turn produces the (first-quantized) Rashba spin-orbit interaction [11, 12]. With standard Pauli matrices $\vec{\sigma}$ acting in spin space,

$$H_{so} = \frac{e\hbar}{4m^2c^2} \vec{E} \cdot (\vec{\sigma} \times \vec{p}). \tag{6}$$

We proceed to derive the second-quantized spin-orbit hamiltonian within the tight-binding approximation. For that purpose, we need the matrix element of the momentum operator between two $2p_z$ orbitals $\vec{p}_{ij} = \langle \chi_i | \vec{p} | \chi_j \rangle$, from which we get the following form for the SO hamiltonian:

$$H_{so} = g \sum_{ij} c_i^{\dagger} \left[(\vec{\sigma} \times \vec{p}_{ij}) \cdot \hat{Z} \right] c_j, \tag{7}$$

where the fermionic operator $c_{i\sigma}$ destroys an electron with spin $\sigma = \uparrow, \downarrow$ in the $2p_z$ orbital centered at $\vec{R}_i$, and $g = E_0/4m^2c^2$. For calculational convenience, the matrix element $\vec{p}_{ij}$ can be written as $g\vec{p}_{ij} = i(\vec{v}_{ij} + \vec{u}_{ij})$, where the spin-orbit vectors $\vec{v}_{ij}$ and $\vec{u}_{ij}$ are defined as

$$\vec{v}_{ij} = -g\alpha \int d^3\vec{r}\, \chi_i(\vec{r} - \vec{R}_i) \hat{\rho}_j e^{-\beta|\vec{r} - \vec{R}_j|}, \tag{8}$$

$$\vec{u}_{ij} = g\beta \int d^3\vec{r}\, \chi_i(\vec{r} - \vec{R}_i) \frac{\vec{r} - \vec{R}_j}{|\vec{r} - \vec{R}_j|} \chi_j(\vec{r} - \vec{R}_j), \tag{9}$$

Note that the modulus of $\vec{v}_{ij}$ and $\vec{u}_{ij}$ has dimension of energy, and their sum (but not necessarily each term separately) is antisymmetric under exchange of i and j.

We first observe that the spin-orbit vectors connecting a site with itself clearly vanish, since $\langle\chi_i|\vec{p}|\chi_i\rangle = 0$. Let us then discuss spin-orbit vectors connecting different sites. Since the orbitals (4) decay exponentially, it is sufficient to consider only the case of nearest neighbors. We start with $\vec{v}_{ij}$. Shifting $\vec{r} \to \vec{s} + \vec{R}_i$ in Eq. (8) and using Eq. (4), we obtain

$$\vec{v}_{ij} = -g\alpha^2 \hat{\rho}_j \int d^3\vec{s}\,(\vec{s}\cdot\hat{\rho}_i) e^{-\beta s} e^{-\beta|\vec{s}+\vec{R}_{ij}|}.$$

Using $\vec{s} = s_\parallel \hat{R}_{ij} + \vec{s}_\perp$, we then rewrite the above integral as

$$\int d^3\vec{s}\,(s_\parallel \hat{R}_{ij} + \vec{s}_\perp)\cdot\hat{\rho}_i\, e^{-\beta s} e^{-\beta\sqrt{(s_\parallel+d)^2+s_\perp^2}},$$

where we use $|\vec{R}_{ij}| = d$, with the nearest-neighbor distance among carbon atoms in graphene $d = 1.42$ Å. Note that $\beta d = 1.34$. The second term in the brackets is odd in $\vec{s}_\perp$ and thus vanishes, and we obtain

$$\vec{v}_{ij} = -g\alpha^2 \hat{\rho}_j \frac{2R}{d} \sin^2\!\left(\frac{\varphi_i - \varphi_j}{2}\right) d^4\gamma_0, \tag{10}$$

where we have used $\hat{R}_{ij}\cdot\hat{\rho}_i = \frac{2R}{d}\sin^2(\frac{\varphi_i-\varphi_j}{2})$ and the dimensionless numerical factor γ_0 :

$$\gamma_0 = \int dx\,dy\,dz\,x\, e^{-\beta d\sqrt{x^2+y^2+z^2}} e^{-\beta d\sqrt{(x+d)^2+y^2+z^2}}.$$

For $\vec{v}_{ji}$, we find

$$\vec{v}_{ji} = g\alpha^2 \hat{\rho}_i \frac{2R}{d} \sin^2\!\left(\frac{\varphi_i - \varphi_j}{2}\right) d^4\gamma_0.$$

Notice that, up to higher orders in d/R, the unit vectors $\hat{\rho}_{i,j}$ can be replaced by $\hat{\rho}_{ij}^{\perp}$, which makes clear that $\vec{v}_{ij}$ is normal to the tube surface. Now $|\sin[(\varphi_i - \varphi_j)/2]|$ varies between zero (when the two sites are aligned in the axis direction) and $d/2R \ll 1$ (when the two sites are aligned in the circumferential direction). Thus, to zeroth order in d/R, $\vec{v}_{ij}$ vanishes: it is a pure curvature effect, peculiar of nanotubes, which does not exist in graphene. In practice, $\vec{v}_{ij}$ is tiny and certainly subleading to $\vec{u}_{ij}$, which turns out to be of order $(d/R)^0$. We shall therefore neglect it in what follows.

Let us now turn to $\vec{u}_{ij}$. We shift $\vec{r} \to \vec{s} + (\vec{R}_i + \vec{R}_j)/2$ in Eq. (9), and rewrite $\vec{u}_{ij}$ as the sum of two terms:

$$\vec{u}_{ij}^{(1)} = g\beta \int d^3\vec{s}\,\chi_i(\vec{s} - \vec{R}_{ij}/2)\,\chi_j(\vec{s} + \vec{R}_{ij}/2) \frac{\vec{s}}{|\vec{s} + \vec{R}_{ij}/2|}, \tag{11}$$

$$\vec{u}_{ij}^{(2)} = \frac{g\beta}{2}\vec{R}_{ij} \int d^3\vec{s}\,\chi_i(\vec{s} - \vec{R}_{ij}/2)\,\chi_j(\vec{s} + \vec{R}_{ij}/2) \frac{1}{|\vec{s} + \vec{R}_{ij}/2|}. \tag{12}$$

Writing again $\vec{s} = s_\parallel \hat{R}_{ij} + \vec{s}_\perp$, the computation of the above integrals leads, to the lowest non-vanishing order in d/R, to the following expressions:

$$\vec{u}_{ij}^{(1)} = g\beta\alpha^2 \vec{R}_{ij} d^4\gamma_1 \equiv u_1 \vec{R}_{ij}, \tag{13}$$

$$\vec{u}_{ij}^{(2)} = \frac{g\beta}{2}\alpha^2 \vec{R}_{ij} d^4\gamma_2 \equiv u_2 \vec{R}_{ij}, \tag{14}$$

with the dimensionless numerical factors

$$\gamma_1 = \int dx\, dy\, dz \frac{x z^2 e^{-\beta d\sqrt{(x-1/2)^2+y^2+z^2}} e^{-\beta d\sqrt{(x+1/2)^2+y^2+z^2}}}{\sqrt{(x+1/2)^2+y^2+z^2}} \simeq -0.0375, \tag{15}$$

and

$$\gamma_2 = \int dx\, dy\, dz \frac{z^2 e^{-\beta d\sqrt{(x-1/2)^2+y^2+z^2}} e^{-\beta d\sqrt{(x+1/2)^2+y^2+z^2}}}{\sqrt{(x+1/2)^2+y^2+z^2}} \simeq 0.3748. \tag{16}$$

To lowest order in d/R, it does not make a difference whether we take the tangent unit vector at $\vec{R}_j$, $\vec{R}_i$, or at $(\vec{R}_i + \vec{R}_j)/2$. Hence we may write $\hat{\rho}_{ij} \to \hat{e}_\varphi$, where $\hat{e}_\varphi$ is the unit tangent vector at $(\vec{R}_i + R_j)/2$. We then get SO couplings along the axial and along the circumferential direction,

$$\vec{u}_{ij} = u\left[(\vec{R}_{ij}\cdot\hat{X})\hat{X} + (\vec{R}_{ij}\cdot\hat{e}_\varphi)\hat{e}_\varphi\right], \tag{17}$$

with $u = u_1 + u_2$. Note that we have neglected a tiny component of $\vec{R}_{ij}$ normal to the tube surface. The above discussion then results in the tight-binding hamiltonian $H = H_0 + H_{so}$, where

$$H_0 = -t\sum_{\vec{r},a} c^\dagger_{B,\vec{r}+\vec{\delta}_a} c_{A,\vec{r}} + h.c.,$$

with $t \approx 2.7$ eV [3]. Here the $\vec{r}$ denote all sublattice-A tight-binding sites of the lattice. Furthermore, the $\vec{\delta}_{a=1,2,3}$ are vectors connecting $\vec{r}$ with the three nearest-neighbor sites which are all located on sublattice B [3]. Since we consider the limit $d/R \ll 1$, the $\vec{\delta}_a$ at each site effectively lie in the tangent plane to the tube surface at that site. The Rashba spin-orbit hamiltonian then reads

$$H_{so} = iu\sum_{\vec{r},a} c^\dagger_{B,\vec{r}+\vec{\delta}_a}\left[\left(\vec{\sigma}\times[(\vec{\delta}_a\cdot\hat{X})\hat{X} + (\vec{\delta}_a\cdot\hat{e}_\varphi)\hat{e}_\varphi]\right)\cdot\hat{Z}\right]c_{A,\vec{r}} + h.c. \tag{18}$$

3. Continuum limit

Since we are interested in the low-energy long-wavelength properties, we now expand the electron operator around the Fermi points K, K' in terms of Bloch waves [3],

$$\frac{c_{p\vec{r}}}{\sqrt{S}} = e^{i\vec{K}\cdot\vec{r}}F_{1p}(\vec{r}) + e^{-i\vec{K}\cdot\vec{r}}F_{2p}(\vec{r}), \tag{19}$$

where $S = \sqrt{3}a^2/2$ is the area of the unit cell, $a = \sqrt{3}d$, and $p = A/B$ is the sublattice index. The $F_{\alpha p}$ are slowly varying electron field operators, and we choose the Fermi points at $\vec{K} = (4\pi/3a, 0)$ and $\vec{K}' = -\vec{K}$ [3]. We then expand $F(\vec{r}+\vec{\delta}) \simeq F(\vec{r}) + \vec{\delta}\cdot\nabla F(\vec{r})$ and use the bond vectors

$$\vec{\delta}_1 = \frac{a}{\sqrt{3}}(0,-1), \qquad \vec{\delta}_2 = \frac{a}{2}(1, 1/\sqrt{3}), \qquad \vec{\delta}_3 = \frac{a}{2}(-1, 1/\sqrt{3}). \tag{20}$$

These vectors are given in a fixed reference frame for a 2D graphene sheet, and we then must perform a rotation to longitudinal and circumferential directions via the chiral

angle. This rotation results in fixed phases that can be absorbed in the definition of $F_{\alpha p}$ and do not appear in final results. This is of course expected from the $U(1)$ symmetry emerging at low energies in the dispersion relation of graphene [3]. After some algebra, the usual Dirac hamiltonian for the kinetic term follows,

$$H_0 = v \int d^2\vec{r}\, F^\dagger \left[(T_0 \otimes \tau_2 \otimes \sigma_0)(-i\partial_x) + (T_3 \otimes \tau_1 \otimes \sigma_0)(-i\partial_y) \right] F, \qquad (21)$$

where $v = \sqrt{3}at/2 \simeq 8 \times 10^5$ m/sec is the Fermi velocity, and x, y are longitudinal and circumferential coordinates, respectively, with $0 < y \leq 2\pi R$. Finally, T_i and τ_i are also Pauli matrices that now act in the space of Fermi (K, K') points and sublattice space (A, B), respectively. For $i = 0$, these are defined as 2×2 unit matrices.

The low-energy limit of the SO term (18) can be obtained in the following way. First we observe that

$$\left[(\vec{\delta}_a \cdot \hat{X})\hat{X} + (\vec{\delta}_a \cdot \hat{e}_\varphi)\hat{e}_\varphi \right] \times \hat{Z} = -(\vec{\delta}_a \cdot \hat{X})\hat{Y} - \sin(y/R)(\vec{\delta}_a \cdot \hat{e}_\varphi)\hat{X}.$$

Here the only approximation is the assumption that the bond vectors lie in the plane tangent to the nanotube surface at $\vec{r}$. Second, by using the bond vectors (20) and taking into account the chiral angle η between the fixed direction on the graphite sheet and the circumferential direction on the nanotube, one obtains

$$\sum_a c^\dagger_{\vec{r}+\vec{\delta}_a} \left(\vec{\delta}_a \cdot \hat{X} \right) c_{\vec{r}} \approx \frac{-3d}{2} \left(F^\dagger_{B1} e^{-i\eta} F_{A1} + F^\dagger_{B2} e^{i\eta} F_{A2} \right),$$

$$\sum_a c^\dagger_{\vec{r}+\vec{\delta}_a} \left(\vec{\delta}_a \cdot \hat{e}_\varphi \right) c_{\vec{r}} \approx \frac{-3d}{2} \left(iF^\dagger_{B1} e^{-i\eta} F_{A1} - iF^\dagger_{B2} e^{i\eta} F_{A2} \right).$$

Notice that we take into account exactly the relative orientation of the bond vectors with respect to the directions $\hat{X}$ and $\hat{e}_\varphi$ for a generic nanotube, which is encoded in the chiral angle η. The constant phases $e^{\pm i\eta}$ can be absorbed by appropriately redefining the operators as $F_{A2} \to e^{-i\eta} F_{A2}$ and $F_{B1} \to e^{-i\eta} F_{B1}$, and the final result can be written down in the form

$$H_{so} = \int d^2\vec{r}\, F^\dagger \left[u_\parallel T_0 \otimes \tau_1 \otimes \sigma_2 + u_\perp \sin(y/R) T_3 \otimes \tau_2 \otimes \sigma_1 \right] F, \qquad (22)$$

with $u_\parallel = u_\perp = 3du/2$. For the sake of generality, we continue to use different coupling constants $u_\perp$ and $u_\parallel$. It is worthwhile to mention that the leading term for the Rashba spin-orbit coupling in a CNT, Eq. (22), does not depend on longitudinal momentum. This is due to the peculiar band structure of graphene with its isolated Fermi (K) points. In the above derivation, we also find terms that are linear in momentum, i.e., contain spatial derivatives of the electron operators. Such terms only produce tiny renormalizations of the velocities and will be neglected here. The second term in Eq. (22) allows for spin flips and mixes transverse subbands.

From now on, for simplicity, we consider just a single Fermi point, say, K. After the global $SU(2)$ rotation $\sigma_1 \to \sigma_2 \to \sigma_3$ in spin space, we get in compact notation

$$\mathcal{H}_0 = v \left[-i\tau_1 \partial_y - i\tau_2 \partial_x \right]. \qquad (23)$$

$$\mathcal{H}_{so} = u_\parallel \tau_1 \sigma_3 + u_\perp \sin(y/R) \tau_2 \sigma_2. \qquad (24)$$

Note that the exact spectrum of $\mathcal{H}_0+\mathcal{H}_{so}$ with $u_\perp = 0$ can be obtained straightforwardly. In general, however, due to the smallness of the SO coupling (see below), it is enough to treat $\mathcal{H}_{so}$ perturbatively. The following detailed derivation is then necessary to correctly evaluate the effect of the SO coupling, and moreover it is interesting and important for the generalization to the interacting case, and for the analysis of features involving the electron wavefunction (as for instance electron-phonon interactions).

The eigenvalues of $\mathcal{H}_0$ are given by

$$\epsilon_{an\sigma}(q) = av\sqrt{k_\perp^2(n) + q^2} \equiv a\epsilon_n(q), \tag{25}$$

where $k_\perp(n) = (n+n_0)/R$ denotes the transverse momentum, q the longitudinal one, $a = \pm$ labels the conduction/valence band, and $\sigma = \pm$ the spin. Here $n_0 = 0$ for intrinsically metallic shells, but generally it can be taken as $0 \le n_0 \le 1/2$ to take into account chirality gaps or orbital magnetic fields along $\hat{X}$. The transverse subbands are labeled by integer values $n = 0, 1, 2, \ldots, \mathcal{N}-1$, where $\mathcal{N} = 2(N^2+M^2+NM)/\gcd(2M+N, 2N+M)$ for (N, M) tubes [3]. $\mathcal{N}$ is typically much larger than the actual number $N = [k_F R]$ of occupied subbands, where we define $k_F = \mu/v$ with the doping level μ that we assume positive here. The velocity v_n for electrons in subband n at the Fermi level (in the absence of H_{so}) and the corresponding Fermi momentum q_n are then given by

$$v_n = v\sqrt{1 - [(n+n_0)/(k_F R)]^2}, \quad q_n = k_F v_n/v. \tag{26}$$

The eigenvalues (25) are spin-independent and thus doubly degenerate. The corresponding eigenstates are denoted $|nqa\sigma\rangle$, where $|n\rangle$ and $|q\rangle$ are respectively plane waves in circumferential and longitudinal direction. In coordinate representation they read

$$\psi_{nqa\sigma}(x, y) \equiv \langle x, y|nqa\sigma\rangle = \frac{e^{ik_\perp(n)y}}{\sqrt{2\pi R}} e^{iqx} \xi_{na}(q) \otimes \chi_\sigma, \tag{27}$$

with the bispinor (in sublattice space)

$$\xi_{n,a=\pm}(q) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\theta_n(q)/2} \\ \pm e^{-i\theta_n(q)/2} \end{pmatrix}, \quad e^{i\theta_n(q)} = \frac{v(k_\perp(n) - iq)}{\epsilon_n(q)}. \tag{28}$$

A different, and here more convenient basis is given by the sublattice states $|nqp\sigma\rangle$. Their coordinate representation is

$$\psi_{nqp\sigma}(x, y) = \frac{e^{ik_\perp(n)y}}{\sqrt{2\pi R}} e^{iqx} \xi_p \otimes \chi_\sigma, \tag{29}$$

where $p = A, B$ and

$$\xi_A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \xi_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Their usefulness stems from the fact that the $|nqp\sigma\rangle$ can be factorized as

$$|nqp\sigma\rangle = |n\rangle|q\rangle \otimes |p\sigma\rangle, \quad |p\sigma\rangle = \xi_p \otimes \chi_\sigma, \tag{30}$$

where $|p\sigma\rangle$ is independent of n and q. Using this basis, we can expand the field operator $F(\vec{r})$ on the tube surface as

$$F(\vec{r}) = \sum_{n,p,\sigma} \int \frac{dq}{2\pi} \psi_{nqp\sigma}(x,y) c_{np\sigma}(q) = \sum_n F_n(x) \langle y|n\rangle, \tag{31}$$

where the operator $c_{np\sigma}(q)$ destroys an electron in the state $|nqp\sigma\rangle$, and we introduce the 1D field operators $F_n(x)$. Alternatively, using the basis of eigenstates of H_0, $F(\vec{r})$ can be expanded as

$$F(\vec{r}) = \sum_{n,a,\sigma} \int \frac{dq}{2\pi} \psi_{nqa\sigma}(x,y) c_{na\sigma}(q), \tag{32}$$

where the operators $c_{na\sigma}(q)$ destroy conduction ($a = +$) or valence ($a = -$) electrons with spin σ in subband n. Notice that in what follows the spin index is left implicit. The relation between the operators c_{na} and c_{np} is easily found to be

$$\begin{pmatrix} c_{n+}(q) \\ c_{n-}(q) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\theta_n(q)/2} & e^{i\theta_n(q)/2} \\ e^{-i\theta_n(q)/2} & -e^{i\theta_n(q)/2} \end{pmatrix} \begin{pmatrix} c_{nA}(q) \\ c_{nB}(q) \end{pmatrix}. \tag{33}$$

We now proceed by treating the spin-orbit hamiltonian using perturbation theory. First, we diagonalize $H_0 - \mu N$ for a fixed transverse subband n,

$$\begin{aligned} H_0^{(n)} - \mu N^{(n)} &= v \int dx\, F_n^\dagger [k_\perp(n)\tau_1 + (-i\partial_x)\tau_2 - \mu] F_n \\ &= \sum_{a=\pm} \int \frac{dq}{2\pi} [a\epsilon_n(q) - \mu] c_{na}^\dagger c_{na}. \end{aligned}$$

Next we expand around the Fermi points $\pm q_n$ defined in Eq. (26), which introduces right- and left-movers, $r = \pm = R/L$, as the relevant low-energy degrees of freedom. For small deviations k from $\pm q_n$, Taylor expansion yields $\epsilon_n(\pm q_n + k) \simeq \mu \pm v_n k$, where v_n is given in Eq. (26). Since we assumed $\mu > 0$, we may now restrict ourselves to the conduction band, $a = +$. For the hamiltonian, we then obtain

$$\begin{aligned} H_0^{(n)} - \mu N^{(n)} &= \sum_{r=\pm} v_n \int \frac{dk}{2\pi} (rk) c_{nr}^\dagger(k) c_{nr}(k) \\ &= \sum_{r=\pm} v_n \int dx\, \psi_{nr}^\dagger (-ir\partial_x) \psi_{nr}, \end{aligned}$$

where $c_{nr}(k) \equiv c_{n+}(rq_n + k)$ and $\psi_{nr}(x) = \int \frac{dk}{2\pi} e^{ikx} c_{nr}(k)$. This introduces R/L-moving 1D fermion operators for each subband n (and spin σ). The relation of these 1D fermions with the original operator $F_n(x)$ is given by

$$\begin{aligned} F_n(x) = e^{iq_n x} \int \frac{dk}{2\pi} \frac{e^{ikx}}{\sqrt{2}} \begin{pmatrix} e^{i\theta_n(q_n)/2} \\ e^{-i\theta_n(q_n)/2} \end{pmatrix} c_{nR}(k) \\ + e^{-iq_n x} \int \frac{dk}{2\pi} \frac{e^{ikx}}{\sqrt{2}} \begin{pmatrix} e^{-i\theta_n(q_n)/2} \\ e^{i\theta_n(q_n)/2} \end{pmatrix} c_{nL}(k). \end{aligned} \tag{34}$$

Notice that, while in general the unitary transformation from sublattice space to the conduction/valence band description depends on longitudinal momentum, in the continuum limit, one can use the transformation directly at the Fermi momenta. This

is consistent with the neglect of band curvature effects implicit in the linearization of the dispersion relation, which is unproblematic away from van Hove singularities associated with the onset of new subbands [32]. At these points, the concept of R/L-movers breaks down, and some of our conclusions below may change.

Next we express the Rashba hamiltonian (24) in terms of R/L-movers. The first term results in

$$H_{so}^{\parallel} = \frac{u_{\parallel} v}{\mu} \sum_{nr} k_{\perp}(n) \int \frac{dk}{2\pi} c_{nr}^{\dagger}(k) \sigma_3 c_{nr}(k). \tag{35}$$

The presence of the factor $k_{\perp}(n)$ results from a careful treatment of the phases in Eq. (34). In Eq. (35) we omit an additional term mixing right- and left-movers. This term contains a rapidly oscillating factor $e^{\pm 2iq_n x}$ and therefore is strongly suppressed by momentum conservation. The second term in Eq. (24) again contains the oscillating phase factor $e^{\pm i(q_n \pm q_{n+1})x}$, which leads to a drastic suppression of $H_{so}^{\perp}$ at low energies and long wavelengths. Of course, this argument relies in an essential way on the smallness of the coupling $u_{\perp}$, as one expands around the hamiltonian H_0. We conclude that away from van Hove singularities, the only important Rashba term is given by $H_{so}^{\parallel}$ in Eq. (35). This term has the appearance of a static homogeneous but *channel-dependent magnetic field*.

4. Oscillatory TMR effects in nanotubes

In this section we will analyze the consequences of our findings regarding spin-orbit couplings in CNTs, see Eqs. (35), for the observability of spin precession effects encoded in the Datta-Das oscillations of the TMR. Based on our expressions, it is possible to estimate the order of magnitude of this effect.

For a concrete estimate, let us put $E_0 = 0.2eV_G/(\kappa D)$, where D is the gate-tube distance, V_G the gate voltage, and κ denotes the dielectric constant of the substrate. For a given channel n, the Rashba-induced energy splitting is then easily estimated as

$$\frac{\Delta E_n}{eV_G} = (\gamma_1 + \gamma_2) \frac{0.6dv}{\mu} \frac{|n + n_0|}{R} \frac{\alpha^2 \beta d^4 \lambda_c^2}{4\kappa D},$$

where $\lambda_c = \hbar/mc = 3.86 \times 10^{-13}$ m is the Compton length. Plugging in the definition of α, β, we get

$$\frac{\Delta E_n}{eV_G} = \frac{0.6(\gamma_1 + \gamma_2)}{256\pi\kappa} (d/a_0)^5 \frac{\lambda_c^2}{D a_0} \frac{|n + n_0|}{k_F R}. \tag{36}$$

Bands with small n are only weakly split, and hence do not contribute to oscillatory TMR behavior. This argument suggests that Datta-Das oscillations in principle could survive in a CNT, even when there are many channels. The major contribution will come just from the few bands with the largest n.

To estimate the accumulated phase difference due to the different precession length of the two split eigenstates, let us put $(n + n_0)/(k_F R) \to 1$, which represents the

dominant contribution, and set $\kappa = 1$. Then Eq. (36) gives as order-of-magnitude estimate

$$\Delta E/(eV_G) \approx 2 \times 10^{-6}\ a_0/D. \tag{37}$$

Even when assuming a very close-by gate, this gives only a tiny splitting, in retrospect justifying perturbation theory. This splitting now translates into a momentum splitting $\Delta k_n = \Delta E/v_n$, and hence into a precession phase mismatch along the CNT of length L [13]. For the nth band, this phase difference is

$$\Delta\phi_n = \Delta k_n L \approx 2 \times 10^{-6} \frac{L}{D} \frac{eV_G}{\hbar v_n/a_0}. \tag{38}$$

This phase difference should be of order 2π to allow for the observation of Datta-Das oscillatory TMR effects [13].

Away from a van Hove singularity, Eq. (38) predicts that oscillations appear on a gate voltage scale of the order of 10^6 to 10^7 V for $L \approx D$, which would make Datta-Das oscillations unobservable. This argument also shows that this interpretation can be ruled out for the parameters relevant for the Basel experiment [10]. From Eq. (38), we can then suggest several ways to improve the situation. First, one should use very long CNTs, while at the same time keeping the gate very close, and second, an enhancement can be expected close to van Hove singularities. Of course, very close to a van Hove singularity, some of our arguments above break down, but the general tendency can nevertheless be read off from Eq. (38). Furthermore, electron-electron interactions can also enhance spin-orbit effects [22, 25].

To conclude, we have presented a detailed microscopic derivation of Rashba spin-orbit coupling in carbon nanotubes. It turns out that the Rashba SO coupling is small, and therefore the prospects for observing spin-precession effects like Datta-Das oscillations in the tunneling magnetoresistance are not too favorable. However, for very long CNTs, close-by gates, and in the vicinity of a van Hove singularity, the requirements for observability of these effects could be met in practice.

Acknowledgments

We thank T. Kontos and C. Schönenberger for motivating this study. Support by the DFG (Gerhard-Hess program) and by the EU (DIENOW network) is acknowledged.

References

[1] Zutic I, Fabian J and Das Darma S 2004 *Rev. Mod. Phys.* **76** 323
[2] Forró L and Schönenberger C 2001 *Topics in Appl. Phys.* **80** 1
[3] Ando T 2005 *J. Phys. Soc. Jpn.* **74** 777
[4] Tsukagoshi K, Alphenaar B W and Ago H 1999 *Nature* **401** 572
[5] Zhao B, Mönch I, Vinzelberg H, Mühl T and Schneider C M 2002 *Appl. Phys. Lett.* **80** 3144
[6] Kim J R, So H M, Kim J J and Kim J 2002 *Phys. Rev B* **66** 233401
[7] Chakraborty S, Walsh K M, Alphenaar B W, Liu L and Tsukagoshi K 2003 *Appl. Phys. Lett.* **83** 1008

[8] Hoffer X, Klinke C, Bonard B M, Gravier L and Wegrowe J E 2004 *Europhys. Lett.* **67** 103
[9] Sahoo S, Kontos T, Schönenberger C and Sürgers S 2005 *Appl. Phys. Lett.* **86** 112109
[10] Sahoo S, Kontos T, Furer J, Hoffmann C and Schönenberger C 2005 (unpublished preprint)
[11] Rashba E I 1960 *Fiz. Tverd. Tela* **2** 1224 [*Sov. Phys. Solid State* **2** 1109]
[12] Bychkov Y and Rashba E I 1984 *J. Phys. C* **17** 6039
[13] Datta S and Das B 1990 *Appl. Phys. Lett.* **56** 665
[14] Stojetz B, Miko C, Forró L and Strunk C 2005 *Phys. Rev. Lett.* **94** 186802
[15] Mireles F and Kirczenow G 2001 *Phys. Rev. B* **64** 024426
[16] Valin-Rodriguez M, Puente A and Serra L 2003 *Eur. Phys. J. B* **34** 359
[17] Balents L and Egger R 2001 *Phys. Rev. B* **64** 035310
[18] Ando T 2000 *J. Phys. Soc. Jpn.* **69** 1757
[19] De Martino A, Egger R, Hallberg K and Balseiro C A 2002 *Phys. Rev. Lett.* **88** 206402
[20] De Martino A, Egger R, Murphy-Armando F and Hallberg K 2004 *J. Phys.: Cond. Matt.* **16**, S1437
[21] Häusler W 2004 *Phys. Rev. B* **70**, 115313
[22] Häusler W 2001, *Phys. Rev. B* **63** 121310
[23] Governale M and Zülicke U 2002 *Phys. Rev. B* **66** 073311
[24] Gritsev V, Japaridze G, Pletyukhov M and Baeriswyl D 2005 *Phys. Rev. Lett.* **94** 137207
[25] Chen G H and Raikh M E 1999 *Phys. Rev. B* **60** 4826
[26] Urbina A, Echeverría I, Pérez-Garrido A, Díaz-Sánchez A and Abellán J 2003 *Phys. Rev. Lett.* **90** 106603
[27] Bachtold A, Strunk C, Salvetat J P, Bonard J M, Forró L, Nussbaumer T and Schönenberger C 1999 *Nature* **397** 673
[28] Schönenberger C, Bachtold A, Strunk C, Salvetat J P and Forró L 1999 *Appl. Phys. A* **69** 283
[29] Chico L, López M P and Muñoz M C 2004 *Phys. Rev. Lett.* **93** 176402
[30] Benedict L X, Louie S G and Cohen M L 1995 *Phys. Rev. B* **52** 8541
[31] Krčmar M, Saslow W and Zangwill A 2003 *J. Appl. Phys.* **93** 3495
[32] Hügle S and Egger R 2002 *Phys. Rev. B* **66** 193311