2012

Specular Andreev Reflection in the Interface of a Two-Dimensional Semiconductor with Rashba Spin-Orbit Coupling and a d-Wave Superconductor

Bo Lv
Peking University

C Zhang
University of Wollongong, czhang@uow.edu.au

Zhongshui Ma
Peking University, Beijing, China, zma@uow.edu.au

http://ro.uow.edu.au/engpapers/4452

Publication Details

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
Specular Andreev Reflection in the Interface of a Two-Dimensional Semiconductor with Rashba Spin-Orbit Coupling and a d-Wave Superconductor

Bo Lv, C. Zhang, and Zhongshui Ma

School of Physics, Peking University, Beijing 100871, China
School of Engineering Physics, University of Wollongong, New South Wales 2522, Australia

We reveal that the recently discovered specular Andreev reflection (SAR) [C. W. J. Beenakker, Phys. Rev. Lett. 97, 067007 (2006)] can occur in semiconductors where the spin-orbit coupling is finite. We demonstrate this finding in the hybrid of a two-dimensional electron gas with Rashba spin-orbit coupling and a superconductor. In the limit of low density or a strong spin-orbit coupling, specular Andreev reflection is finite. We also show that unit electron-hole conversion is possible in a specular Andreev reflection due to the different topological structures of the equal-energy surface between electrons and holes. The SAR in the semiconductor is determined by the relative orientation of wave vector to group velocity, which can be analyzed by ray equations.

Model.—We consider a ballistic R2DEG/S junction where the R2DEG is in the region \( x < 0 \) and the superconductor is in the region \( x > 0 \). Their interface is located at \( x = 0 \) (along the \( y \) axis) and a \( \delta \)-potential barrier of strength \( Z \) separates two materials, \( U(x) = Z\delta(x) \). Transport is along the \( x \) axis. We assume that the Fermi wavelength \( \lambda_F \) is much shorter than the BCS coherence length and the London penetration depth. In general, the relevant dimensions of the NS device are of the same order of magnitude as \( \lambda_F \). For simplicity we also assume that the

Andreev reflection (AR) [1] is an important phenomenon of quantum tunneling in normal metal/superconductor (NS) junctions. It is a two particle process in which an incident electron in the normal metal couples with another electron below Fermi level to form a Cooper pair across the interface into the superconductor [2]. Equivalently, a hole is reflected tracking the opposite path of the oblique incident electron [retro-Andreev-reflection (RAR)] in normal metal and a Cooper pair is created in the superconductor, simultaneously. Applying the scattering wave function method by approximating the insulating barrier at the surface as a \( \delta \) function, Blonder, Tinkham, and Klapwijk (BTK) [3] studied AR in NS junctions. They showed that the conductance increases considerably if the applying voltage is within the gap of the superconductor and the barrier strength is not strong. The AR in the combined structures involving the semiconductor [4], ferromagnets [5], spintronic systems [6], and \( s \)- and \( d \)-wave superconductors [8] have been investigated. The properties of AR involving different normal states and different superconducting states can be found in a recent review [9].

The reflection of holes at the NS interface not along the incident direction, or specular Andreev reflection (SAR), is a rare physical phenomenon. SAR had not been predicted until Beenakker [10,11] discovered the possibility of an unusual electron-hole conversion in the reflection of relativistic electrons in graphene at a superconductor. By combining the Dirac equation with the Bogoliubov–de Gennes (BdG) equation of superconductivity, it was shown that [10,11] AR can be a RAR as well as a SAR. Because of the unique band structure of graphene, an electron above the Dirac point can be reflected into a hole below the Dirac point, behaving as a specular reflected hole. SAR was also studied in a graphene ferromagnet/superconductor junction [12]. Recently, superconducting states in graphene have been realized by proximity effect through contact with superconducting electrodes [13]. This provides a possibility of experimentally observing SAR in the future.

To date, SAR has not been found in any system (non-relativistic) other than graphene (Dirac-like system). In this Letter, we show that SAR can occur in Schrödinger systems such as two-dimensional semiconductors with true spin-orbit coupling (SOC). We shall use a hybrid of a two-dimensional semiconductor with Rashba SOC (R2DEG) [14] and a superconductor to demonstrate this. Herein, the electron-hole conversion can be controlled by the strength of SOC. We will consider both \( s \)- and \( d \)-wave superconductors in order to analyze the influence of interface barrier and quasiparticle spin orientation in superconductors on the SAR. For \( d \)-wave superconductors, the SAR also varies with the anisotropic angle. It is found that the electron reflection characteristics (retro or specular) are solely determined by the relative orientation of the wave vector and group velocity of the reflected hole. The characteristics of reflected holes can be identified by the relative sign among their ray equations [15]. There is another difference between graphene and the present system. In graphene both conduction and valence bands are needed and the SAR reflected hole is in the valence band. In semiconductors, only the conduction band with SOC is required. The valence below the gap is irrelevant. The SAR reflected hole remains in the conduction band.
effective masses in the 2DEG and superconductor are the same. The Fermi levels of them align in line if no bias is applied across the junction. Thus, the physics on the scale of $\Delta_F$ can be described by BdG equations, $H \Psi(r) = e \Psi(r)$, where $H$ is the Hamiltonian describing the R2DEG/S junction, written as

$$H = \left( \begin{array}{cc} h(x) - e_F & i \Delta(r) \sigma_y \Theta(x) \\ -i \Delta(r) \sigma_y \Theta(x) & -h^*(x) + e_F \end{array} \right)$$

with

$$h = \frac{p^2}{2m^*} + \Theta(-x) \hat{e}_z \cdot \left( \frac{\lambda \sigma \times \hat{p}}{\hbar} \right) + \left( Z - \frac{i}{2} \right) \sigma_z \delta(x),$$

where $\sigma$ are Pauli matrices, $m^*$ is the effective mass of electron, $p$ is the momentum operator, $\Delta(r)$ is the pair potential of a quasiparticle, and $\lambda$ is the strength of Rashba SOC. Here $\Theta(x)$ is the Heaviside step function. The problem can be studied by using the BTK formalism, i.e., solving the BdG equations in both sides of the junction subject to the boundary conditions at the interface.

Eigenstates and energy dispersions.—In the R2DEG region $x < 0$, the four eigenstates are given as $\xi^{(e)}_{\pm}(\theta_{k_x}) = 2^{-1/2} (1 \ 0 \ 0 \ i \sigma_z) \tau^T$ and $\xi^{(h)}_{\pm}(\theta_{k_x}) = 2^{-1/2} (0 \ i \sigma_z \ 1 \ 0)^T$, corresponding to spin helicity states $s(= \pm)$. Here $\tau$ represents a transpose, $\theta_{k_x}^{(e/h)} = \arctan[k_{x'}^{(e/h)}/k_{x''}^{(e/h)}]$. The eigenenergies are given as $\epsilon^{(e)}_\pm = \hbar^2(k + sk_{SO})^2/2m^* - \epsilon_{SO}$ and $\epsilon^{(h)}_\pm = -\hbar^2(k + sk_{SO})^2/2m^* + \epsilon_{SO}$, $k_{SO} = m^*/\hbar^2$ is the momentum offset of the annular maximum in the minus branch, and $\epsilon_{SO} = m^*\lambda^2/2\hbar^2$ is the Rashba energy defined as the relative value of minimum energy in the minus branch to the crossing point of two branches, as shown in Fig. 1(a). In the superconductor region $x > 0$, the quasiparticle wave function is a sum of electronlike and holelike quasi-particles. The superconductors can be classified into $s$- and $d$-wave types by the spin dependence of the pair potentials. The Fourier transform of the pair potential is modeled as $\Delta(k)$. For an $s$-wave superconductor $\Delta(k)$ is a real constant $\Delta_0$ because of symmetry and the independence of spin orientation. However, for $d_{x^2-y^2}$ symmetry $\Delta(k)$ depends on the quasiparticle spin orientation in the superconductor, given by $\Delta(k) = \Delta_0 e^{i\alpha} e^{i\varphi}$, where $\varphi = \cos(2\beta \pm \beta)$ as a function of the anisotropy angle $\beta$ (the angle between the normal to the interface and the crystal axis of $d$-wave superconductors), $\alpha$ is the quasiparticle angle in the superconductor $\varphi = \sin^{-1}k_y/|k|$, and $\varphi_\pm$ provides the amplitude of the angle between the electronlike or holelike quasiparticle's wave vector and the interface normal. Solving the BdG equation, we found the steady states in the superconductors.

Fermi surface and the different topological structures for the electrons and the holes.—We consider the case that the Fermi energy lies at the crossing point of two branches of the electron spectrum. For a certain energy $\epsilon$, which lies slightly above the band crossing, there are two spin-splitting bands for the electrons, corresponding respectively to two different spin helicity states. The equal-energy surface crossing energy bands $\epsilon^{(e)}_\pm$ identify two concentric Fermi discs of opposed spin helicities, red (medium gray) and yellow (light gray) discs in Fig. 1(b). The radii of Fermi discs are given as $k^{(e)}_\pm = [2m^*(\epsilon + \epsilon_{SO})/\hbar^2]^{1/2} \pm k_{SO}$. This same energy $\epsilon$ intersects the hole band below the band crossing, or only intersects the $\ldots$ branch $\epsilon^{(h)}_\pm$. The equal-energy surface is, then, constituted by two concentric circles of radii $k^{(h)}_\pm$ and $k^{(h)}_\pm = [2m^*(\epsilon_{SO} - \epsilon)/\hbar^2]^{1/2}$ for the outer circle [point $D$ in Fig. 1(b)] and $k^{(h)}_\pm = [2m^*(\epsilon_{SO} - \epsilon)/\hbar^2]^{1/2}$ for the inner circle [point $F$ in Fig. 1(b)]

The density of state (DOS) for two electronic branches $n^{(e)}_\pm$ are $D_+^{(e)}(\epsilon) = (D_0/2)[1 - F^{2}(\epsilon)]$ and $D_-^{(e)}(\epsilon) = (D_0/2)[1 + F^{2}(\epsilon)]$, respectively, where $F(\epsilon) = \epsilon_{SO}/(\epsilon + \epsilon_{SO})$ and $D_{\pm} = m^*/\pi\hbar^2$. The total DOS for the electrons is $D(\epsilon) = (D_0/2)$ (2D characteristics). However, the holes are only in the $\epsilon^{(h)}_\pm$ states. The DOS for the holes is $D_+^{(h)}(\epsilon) = D_0[F(\epsilon)]^{1/2}$, which nonvanishes if $0 < \epsilon < \epsilon_{SO}$ and shows a $(\epsilon + \epsilon_{SO})^{-1/2}$ van Hove singularity behavior representing one-dimensional characteristics.

From these, it is found that the equal-energy surfaces for the electron and hole bands have different topological structures, i.e., two concentric Fermi discs for the electrons and an annular torus for the holes. Correspondingly, the DOS for the electron and hole are very different. The
former has 2D characteristics while the latter has 1D characteristics.

Ray equations for the reflection modes.—We now use the ray analysis[15] to demonstrate SAR in R2DEG. We provide a concrete analysis of SAR by using ray equations and explicitly show that the occurrence of SAR relates to the sign reversal of ray equations in the scattering process. At incident energy $e$, there are two electron states [the wave vectors $A$ and $A'$ in Fig. 1(a) and 1(b)]. Below we analyze the reflection of the $A$ electron. The reflection of the $A'$ electron can be analyzed in a similar manner.

At the point $A$ the wave vector $k = (k_x, k_y)$ and the group velocity $v_g = (v_{g1}, v_{g2})$ are given as $k^{(A)} = k^{(e)}(\cos \theta^{(e)}_k, \sin \theta^{(e)}_k)$ and $v_{g}^{(A)} = v^{(e)}(\cos \theta^{(e)}_k, \sin \theta^{(e)}_k)$. Here $v^{(e)} = [2(e + \epsilon_{SO})/m^*]^{1/2}$. It follows that $k^{(A)} \cdot v^{(A)} > 0$.

There are four reflection modes with the $k$ vector at $B$ (reflected electronic state “+”), $C$ (reflected electronic state “−”), $D$ (reflected hole with the vector at outer circle), and $F$ (reflected hole with the vector at inner circle) points. The top half of the diagram in Fig. 1(c) illustrates these scattering processes. It is straightforward to show that for reflected electrons, $k^{(B)} \cdot v^{(B)} > 0$ and $k^{(C)} \cdot v^{(C)} > 0$, and for reflected holes, $k^{(D)} \cdot v^{(D)} < 0$ and $k^{(F)} \cdot v^{(F)} < 0$. The magnitude of the group velocity for a hole is $v^{(h)} = [2(e - \epsilon_{SO})/m^*]^{1/2}$. Comparing these ray equations with those of incident electrons, it is interesting to find that the $k^{(h)} \cdot v^{(h)}$ at $D$ point has the opposite sign to that of the electron, while the $k^{(h)} \cdot v^{(h)}$ at $F$ point is the same sign as that of the electron. Therefore the reflection with the wave vector at $D$ point is retroreflection and that with the wave vector at $F$ point is specular reflection, or SAR. By the same ray analysis, it is found that for electrons incident from $A'$ point following that $k^{(A')} \cdot v^{(A')} > 0$, the reflection with the wave vector ending at $D$ point is RAR and that with the wave vector ending at $F$ point is SAR, as depicted in the bottom half of the diagram in Fig. 1(c).

Critical incident angles.—For electrons incident from the state with the $k$ vector at the $A'$ point, there is no restriction on the incident angle $\theta^{(e)}_{k+}$ and the electrons can be reflected to any reflection states. However, if the incident electron is from $A$ point (the negative helicity branch $e^{(e)}$), there exist three critical angles beyond which certain type reflections are forbidden. From the conservation of energy and momentum in the scattering processes, we define three critical angles, $\theta^{(e)}_{c+}, \theta^{(e)}_{c-}$, and $\theta^{(h)}_{c-}$ for the incident $\theta^{(e)}_{k+}$. They are the maximum incident angles for the electron being reflected to the electron state (C point), the hole state at the outer edge of equal-energy circle $e^{(h)}$ (D point), and the hole state at the inner edge of the energy circle (F point). There is a simple relation, $\theta^{(h)}_{c-} > \theta^{(h)}_{c+} > \theta^{(e)}_{c+}$. We are interested in SAR, which corresponds to the reflected hole in the inner circle of the state “−” (F point). The critical angle for SAR is found as,

$$\theta^{(h)}_{c-} = \arcsin[1 - (1 - e/\epsilon_{SO})^{1/2}]/[1 + (1 + e/\epsilon_{SO})^{1/2}].$$

SAR can only occur if the incident angle is smaller than this critical angle. For SAR to dominate, the maximum allowed angle should be large, or close to $\pi/2$. It can be seen that the maximum allowed angle is small if the incident electron is from the “+” branch. On the other hand, the maximum allowed angle extends to $\pi/2$ if the incident electron is from the “−” branch. Therefore our result is most applicable to systems with low densities[16].

Coefficient of SAR.—In the R2DEG region ($x < 0$), for an incident electron with the energy slightly higher than the Fermi energy, there are four possible reflections, two of them are reflected electrons at the wave vectors $k^{(h)}$ and $k^{(C)}$, and the other two are the reflected holes at $k^{(h)}$ and $k^{(F)}$. According to the analysis given above, the hole state with the $k$ vector at $D$ is RAR but the state with the $k$ vector at $F$ is SAR. The total wave function for the $s$ state on the incident side can be written as $\Psi(x, y) = \Phi^{(L)} e^{ik_yy}$ with $\Phi^{(L)}(x) = \frac{e^{i\theta^{(e)}_k}}{\sin \theta^{(e)}_k} e^{ik_xx} + \sum_{\alpha=\pm} c_{\alpha} e^{i\theta^{(h)}_e} e^{ik_{\alpha} x} + \sum_{\alpha=\pm} c_{\alpha} e^{i\theta^{(h)}_e} e^{ik_{\alpha} x}$, where $\alpha = \pm$ corresponds to the two possible incident states $\xi^{\pm}(\theta_k)$, $k_{\pm} = k^{(e)} \cos \theta_k \mp k^{(h)} \sin \theta_k$. The coefficients $r^{(e)}_{\alpha}$ and $r^{(h)}_{\alpha}$ are the amplitudes of reflected probabilities corresponding to the reflected electron states in wave vectors $k^{(e)}_x = -k^{(e)} \cos \theta_k$ and $k^{(h)}_x = \pm k^{(h)} \cos \theta_k$, and the reflected hole states in wave vectors $k^{(h)}_x = k^{(h)} \cos \theta_k$ and $k^{(F)}_x = -k^{(F)} \cos \theta_k$. In the region $x > 0$, the transmission wave functions are a superposition of electron- and holelike states of quasiparticles. The reflection coefficients $r^{(e)}_\alpha$ and $r^{(h)}_\alpha$ are determined by the boundary conditions at the interface, (i) the wave function is continuous, and (ii) the first derivative of the wave function is discontinuous with a step $m^* (2Z - i\lambda \sigma_{\psi}) \Phi(\beta, 0)$.

Specular reflected fluxes and the SAR conductance.—From the reflection coefficients $r^{(e)}_\alpha$ and $r^{(h)}_\alpha$, we can calculate the flux for each reflection mode by the formula $J_\alpha = (1/m^*) \Im(\psi_{\alpha}^\dagger \partial_y \psi_{\alpha}) + \lambda \Theta(x) \psi_{\alpha}^\dagger (e_x \times \sigma_y) \psi_{\alpha}$, and analyze the ratio of a given reflection flux to the incident flux. The angular dependence of the SAR coefficients (defined as the ratio of SAR fluxes to the incident fluxes) is shown in Fig. 2 for s-wave superconductors and in Fig. 3 for d-wave superconductors. The results of angle distributions of SAR fluxes reveal some interesting physics.

For $s$-wave superconductors, we found that, for either incident state ($A$ or $A'$), there is a nonzero amplitude of the specular reflected hole ($F$). This specular reflected hole state can become dominant if the incident electron is from the “+” branch, the “$A$” state. Unit conversion of an electron to a SAR hole is possible near normal incidence from the “+” branch, i.e., from the $A'$ state to the $F$ state. The change of wave vector between $A'$-$F$ is much smaller compared to that between $A$-$F$, suggesting a better phase
matching as the electron is converted to a hole from $A'$ to $F$. Furthermore, the critical angle of SAR for the incident from $A'$ extends to $\pi/2$. Over a large energy range, the dominant SAR contribution occurs at $\varepsilon = 1.0\Delta$. This large SAR is apparently not affected by the interface potential strength. As a result, there is a cusplike peak in the SAR conductance at $\varepsilon = 1.0\Delta$ (shown in Fig. 4). For other energies, the SAR coefficient decreases with the interface potential strength.

The dependence of SAR on incident energy and on interface potential in $d$-wave superconductors is very similar to that in $s$-wave superconductors; i.e., SAR is dominated by $A'$-$F$ reflection. In addition, SAR is also strongly dependent on the anisotropic angle $\beta$. While unit conversion remains for small $\beta$ at $\varepsilon = 1.0\Delta$, the probability of this process decreases rapidly with $\beta$. As $\beta$ increases, the total SAR amplitude decreases. Under zero interface potential (insets of Fig. 3), a large SAR amplitude occurs under a small anisotropic angle. This suggests that in the absence of any interface potentials, a large superconducting gap enhances the SAR. The interface potential has very little effect on the SAR at $\Delta\varepsilon = \Delta$. However, the interface potential can change the SAR amplitude for an incident electron with energy $\Delta\varepsilon \neq \Delta$. The SAR amplitude decreases with $Z$ for $\beta = 0$ and increases with $Z$ for $\beta = \pi/8$. Further increasing the $\beta$ to $\pi/4$ completely removes the SAR. For $d$-wave superconductors, the gap vanishes as $\beta \to \pi/4$. Therefore, the gap remaining finite is a necessary requirement for SAR to occur in semiconductor/superconductor junctions.

The contribution of an angle averaged conductance from SAR at zero temperature can be obtained as

$$\frac{\sigma_{\text{SAR}}}{\sigma_0} = \text{Re} \sum_{s = \pm 1} \int -\pi/2 \langle D_s(e_F)\gamma_s^{(\text{SAR})}(V, \theta_s^{(e)})d\theta_s^{(e)},$$

where the sum is over two incident states, $D_s(e)$ is the electron DOS for the spin-splitting states of incidence, and $\gamma_s^{(\text{SAR})}(eV, \theta_s^{(e)}) = \gamma_s^{(i)}|F_s^{(h)}|^2 \cos\theta_s^{(h)}$ is the specular reflected hole flux for the incident $s$-electron state, and $\sigma_0 = e^2/(2\pi^2)$. We would like to point out that while the SAR proposed here is an intrinsic phenomenon, it is an experimental challenge to directly quantify it via standard total conductance measurement. Since the total conductance is a macroscopic property and contains contributions from all possible transport channels, it is not straightforward to separate contributions from different processes such as the SAR. In Fig. 4, we show explicitly the SAR conductance of a semiconductor/superconductor junction for different interface potentials. The SAR conductance for

FIG. 2 (color online). SAR coefficients in $s$-wave superconductors as a function of the incident angle for different values of $Z$, $\lambda_{SO} = 0.4$.

FIG. 3 (color online). SAR coefficients in $d$-wave superconductors as a function of the incident angle for different values of $Z$, $\lambda_{SO} = 0.4$. The insets are SAR coefficients for the interface potential $Z = 0$.

FIG. 4 (color online). SAR conductance versus bias for different interface potentials. (a) For $s$-wave and (b) for $d$-wave superconductors with $\beta = \pi/8$. The insets show the derivative of the SAR conductance.
$d$-wave superconductors with $\beta = 0$ is very similar to that of $s$-wave superconductors. Increasing $\beta$ reduces the resonance height and total SAR conductance. It is found that the SAR contribution to the total conductance is mostly concentrated around the $eV = \Delta$. Since SAR conductance has a cusplike resonance around $eV = \Delta$, we suggest that the most sensitive way to experimentally identify the SAR is by fine measuring $d\sigma_{\text{SAR}}/d(eV)$ around the resonance (the inset in Fig. 4).

In conclusion, we revealed SAR in a semiconductor/superconductor junction. This is yet another example of an entirely new phenomenology recently discovered in graphene by Beenakker [10,11]. In the present case, it is the interplay of true SOC and the superconductivity which results in SAR. What is more interesting in the present case is the tunability of the Rashba SOC parameter under a gate voltage. This makes the retro to SAR tunable by an external means. From a more fundamental point of view, the SAR may indicate an intrinsic connection of relativistic dynamics in graphene and the relativistic effect behind that of the SOC in semiconductors, as the relativistic Hamiltonian of graphene is naturally described by the electron-pseudo spin coupling and SOC follows a non-relativistic approximation of relativistic dynamics. In predicting SAR in graphene, Beenakker pointed out [10,11] that the practical significance of his prediction rests on the fabrication of high-quality contacts between a superconductor and graphene. We note that the semiconductor/s-wave superconductor junction can be controlled experimentally. Therefore the system proposed here may be an alternative structure for an experimental test of SAR.

B. L. and Z. S. M. acknowledge the financial support from the NNSFC Grant (No. 91021017) and NBRP of China (2012CB921300). C. Z. acknowledges support from the Australian Research Council Grant (No. DP0879151).