Gapless chiral excitons in thin films of topological insulators

S. A. JAFARI\( ^{(a)} \)

Department of Physics, Sharif University of Technology - Tehran 11155-9161, Iran, Center of Excellence for Complex Systems and Condensed Matter (CSCM), Sharif University of Technology Tehran 14588-9694, Iran and School of Physics, Institute for Research in Fundamental Sciences - Tehran 19395-5531, Iran

received 21 May 2014; accepted in final form 23 June 2014
published online 14 July 2014

PACS 03.65.Vf – Phases: geometric; dynamic or topological
PACS 71.35.Lk – Collective effects (Bose effects, phase space filling, and excitonic phase transitions)
PACS 03.65.Pm – Relativistic wave equations

Abstract – In a nanoscopic thin film of a strong topological insulator (TI) the Coulomb interaction in the channel that exchanges the two electrons with the same chirality in two different planes of the slab takes advantage of the minus sign resulting from such “exchange” and gives rise to a bound state between the positive-energy states in one surface and the negative-energy states in the opposite surface. Therefore particle and hole pairs in the undoped Dirac cone of the TI thin film form an inter-surface spin-singlet state that lies below the continuum of free particle-hole pairs. This mode is similar to the excitons of semiconductors, albeit formed between the electron and hole pairs from two different two-dimensional surfaces. For low momenta the dispersion relation characterizing this collective mode is linear. Experimental comparison of two slabs with different thicknesses can capture the exponential dependence of the present effect on the slab thickness.

Copyright © EPLA, 2014

Introduction. – Strong topological insulators (TIs) are characterized by a bulk gap, while the surface surrounding the material hosts an odd number of gapless Dirac fermions [1,2]. The classification of the topological insulators is based on the non-interacting picture and their basic band structure [3]. The gapless helical surface states are characteristic of topological insulators as long as the many-body interactions can be ignored. Therefore an important question is: What happens when interactions are turned on? The nature of collective excitations of the helical metallic states [4] and possible instabilities [5] are of particular interest. If the interactions are of short-range nature and are strong enough to destroy the metallic state of the surface, a Mott insulating state in the surface is expected, albeit with surface spinon bands [6]. When the interactions are not strong enough to destroy the metallic nature of the surface states, the Coulomb interactions within a surface of a TI doped away from the Dirac point gives rise to plasmonic and zero sound collective excitations in the helical metallic state of the surface [4]. Such collective excitations require a Fermi surface and hence disappear in the undoped surface Dirac cone, \( \mu = 0 \). The Dirac cone of undoped graphene supports a particle-hole bound state — i.e. a split-off state below the free particle-hole continuum (PHC) — which disperses linearly with momentum [7–9]. In contrast to the antibonding nature of plasmon excitations [10], a minus sign required to render the interaction vertex attractive is provided by the exchange interactions [7]. The Dirac cone describing the topological surface states features a spin-momentum locking [1] which makes the physical spin inaccessible for the exchange process. The question is, then: Can there be a mechanism to give rise to a bound state below the PHC of the free Dirac fermions topological insulators? The physical spin of the electrons in helical metals is locked to the momentum and hence the spin-exchange scattering requires their momentum to be reversed. But such a back-scattering is prohibited by the chiral nature of carriers. Therefore unlike graphene [7], the combination of time-reversal symmetry and the chiral nature of helical states prevents from taking advantage of the minus sign (attraction) in the exchange Coulomb vertex in the spin-triplet particle-hole channel. In this work we

\( ^{(a)} \) E-mail: jafari@physics.sharif.edu
show that when the TI is cut into a thin film the two surfaces can be described by a pseudo-spin $\hat{\tau}$ taking on two values $\tau^2(\equiv \tau) = \pm 1$. Then the $\tau$-exchange scattering part of the Coulomb interaction will provide attractive interaction between the positive-energy state from one surface and a negative-energy state from the other surface and hence leads to a bound state in the particle-hole channel which is of singlet nature with respect to spin, but can be considered of triplet nature with respect to pseudo-spin $\hat{\tau}$. This is akin to excitons of a semiconductor, but with particle and hole states coming from two different surfaces $\tau = \pm$ surfaces of the TI slab. It can also be thought of as the “pseudo-spin” version of a collective mode in the Dirac cone of graphene [7,8]. Since the constituent particle and hole pairs come from two different planes, the center of mass of this collective mode will be in a bisecting plane parallel to the surface. Therefore in planes, the center of mass of this collective mode will be different surfaces $\tau$. But with particle and hole states coming from two different surfaces $\tau = \pm$ surfaces of the TI slab. It can also be thought of as the “pseudo-spin” version of a collective mode in the Dirac cone of graphene [7,8]. Since the constituent particle and hole pairs come from two different planes, the center of mass of this collective mode will be in a bisecting plane parallel to the surface. Therefore in planes, the center of mass of this collective mode will be different surfaces $\tau$. But with particle and hole states coming from two different surfaces $\tau = \pm$ surfaces of the TI slab. It can also be thought of as the “pseudo-spin” version of a collective mode in the Dirac cone of graphene [7,8]. Since the constituent particle and hole pairs come from two different planes, the center of mass of this collective mode will be in a bisecting plane parallel to the surface. Therefore in planes, the center of mass of this collective mode will be different surfaces $\tau$.

For a slab of TI with thickness $\ell$, the Coulomb energy scale is be given by $e^2/(\epsilon_d f)$, where $\epsilon_d$ is the relative dielectric constant of the insulating medium (the bulk) separating the two surfaces. The separation $\ell$ serves to tune the strength of interactions for a fixed bandwidth of surface states. For nanosopic separations on the scale of $\ell \sim 10 \, \text{nm}$, on the one hand the strength of the Coulomb interaction will become comparable to the bandwidth of the surface states [9]. On the other hand, such separation is large enough to prevent the direct hybridization of the atomic orbitals on the opposite surfaces, and hence the single-particle part of the spectrum is protected from gap opening and will continue to be a gapless Dirac theory. Increasing the strength of $\tau$-off-diagonal interactions is expected to give rise to fascinating possibilities, including the proposal of topological excitonic condensation [5,12].

Non-interacting Hamiltonian. – To fix the notations for the rest of the paper, let us summarize the non-interacting part of the Hamiltonian of a TI thin film and discuss the relevant matrix elements. The low-energy effective theory describing the gapless metallic states on the surface of a TI is the so-called “helical metal” and is given by [1].

$$H_0 = v_F \sum_k \epsilon_k \hat{c}_{\tau \mu k} \cdot (\mathbf{k} \times \hat{\sigma}) \hat{c}_{\tau \mu k},$$

where the two-dimensional momentum $\mathbf{k}$ (we set $\hbar = 1$) is in the surface characterized by normal vector $\hat{n}$. The momentum is locked to its spin $\hat{\sigma}$. We use the spinor notation $\hat{c}_{\tau \mu k} = (\hat{c}_{\uparrow \mu k}, \hat{c}_{\downarrow \mu k})$ where $\uparrow$ and $\downarrow$ are the $z$-component of the spin of electrons. For upper surface $\hat{n} = +\hat{z}$, while for the lower surface $\hat{n} = -\hat{z}$, amounting to $\hat{n} = \tau \hat{z}$, with $\tau = \pm 1$ for the two surfaces [13]. When two opposite surfaces labeled by the pseudo-spin values $\tau = \pm$ are brought to nano-meter distance $\ell$, as long as the surfaces are not close enough to allow for the overlap of atomic wave functions, the Hamiltonian of the thin film will be given by two copies of the above Dirac model as

$$H_{\ell} = v_F \sum_{\tau \lambda k} c_{\tau \lambda k}^\dagger \tau \sigma \cdot (k \times \hat{\sigma}) c_{\tau \lambda k}. (1)$$

A constant energy configuration within this Hamiltonian is described by a circulating pattern of in-plane spin texture oriented perpendicular to their two-dimensional momentum vector $\mathbf{k}$ [2]. If we denote the azimuthal angle of $\mathbf{k}$ with respect to the $k_x$-axis with $\theta_k$, then in terms of the $\tau/2$ rotated momentum vector $\hat{k}$ whose azimuthal angle is $\varphi_k = \theta_k + \pi/2$, the $2 \times 2$ Hamiltonian matrix in the above equation will be given by

$$H_{\ell} = v_F \tau \sigma \cdot (k \times \hat{\sigma}) c_{\tau \lambda k}^\dagger c_{\tau \lambda k}.$$ (4)

where the subscript $\lambda = \pm$ stands for the helicity and the cone-like dispersion is given by $e_{\tau \lambda k} = \tau \lambda v_F k \equiv \tau \lambda \epsilon_{k}$. The eigen-mode $|\tau \lambda \mathbf{k}\rangle$ of the helical metal is created by

$$a_{\tau \lambda k}^\dagger |\text{vac}\rangle = \frac{\tau}{\sqrt{2}} (\hat{c}_{\tau \downarrow k}^\dagger + \lambda e^{i\varphi_k} \hat{c}_{\tau \uparrow k}^\dagger) |\text{vac}\rangle.$$ (5)

The above eigen-states correspond neither to $|\uparrow\rangle$ nor to $|\downarrow\rangle$ spin states. Therefore, while the energy eigenmodes correspond to spin-half ($S^2/2 = 3h^2/4 |\tau \lambda \mathbf{k}\rangle$), their spin has no component perpendicular to the surface, i.e. $\langle \mathbf{k} | S_z | \tau \lambda \mathbf{k}\rangle = 0$. Rather, at each $\mathbf{k}$ point the eigen-state has a definite spin angular momentum in the transverse direction $k$ which is in the plane of the surface and perpendicular to $\mathbf{k}$. Such a locking between the spin angular momentum on the linear momentum roots in the coupling between spin and orbital motion and forces the planar spin of the positive- and negative-energy states in the same plane to be always in opposite directions [2,11]. Therefore low-momentum transfer excitations between negative- and positive-energy states will always correspond to spin-singlet particle-hole pairs. So it is not possible to take advantage of the spin-flip (exchange) processes in order to stabilize a bound state in the particle-hole channel as long as a fixed surface is concerned.

However, when the other surface of the TI is brought closer, the surfaces on the two sides of the slab can be labeled by $\tau = \pm$. The index $\tau$ for an infinitely large plane will be an attribute of the electron motion (constant of motion) and therefore a pseudo-spin $\tau$ emerges.
very naturally. The Coulomb interaction in the channel that “exchanges” the $\tau$ attributes (fig. 1) can take advantage of this pseudo-spin structure. Hence the Coulomb repulsion between electrons translates into attraction between a positive-energy state in one plane (say $\tau = +$) and the negative-energy state in the opposite plane ($\tau = -$) and hence is expected to bind them below the PHC.

Collective excitations in slab of topological insulator. – In the plasmon and zero sound excitations considered in [4] for a single value of $\tau$, only states with a fixed value of $\lambda$ are involved, as they are intra-band excitations. But when the chemical potential is at the Dirac node, $\mu = 0$, both positive- and negative-energy states are expected to play a role in the formation of the particle-hole bound state. The non-interacting part of the Hamiltonian $H^\text{lab}_0$ describes electrons at $|\tau\lambda k\rangle$ states, where the sign of energy is given by the product $\tau\lambda$ of the pseudo-spin value and the helicity. Hence a positive-energy state from one plane and the negative-energy state from the opposite plane must have the same helicity $\lambda$. For an electron of given energy traveling in some direction in the surface $\tau$ whose planar spin is $\sigma$, there will be another state at the same energy corresponding to a propagating electron in the opposite direction in the plane $\bar{\tau}$ with the same planar spin $\sigma$. This can be alternatively thought of as a hole in the $\bar{\tau}$ surface whose spin is $\bar{\sigma}$ and propagates in the same direction as the electron in state corresponding to $|\tau,\sigma\rangle$.

The resulting attraction between the electron and hole propagating in the same direction binds them in the spin-singlet channel.

Now let us put the above idea in a formal setting. To do this, we employ an interaction term of the following form:

$$H^\text{int} = \frac{1}{2A} \sum_{q\tau} \rho_{q\tau} \rho_{-q\bar{\tau}} v(q),$$

which is the standard density-density Coulomb interaction and the density fluctuation $\rho_{q\tau}$ is the Fourier component of the density operator at surface $\tau$ and $A$ is the area of the thin film under consideration. A peculiar feature of the following problem would be due to the interaction $v(q)$: The momentum $q \equiv (q_x, q_y)$ is a two-dimensional momentum and the Coulomb interaction is between two planes separated by $\ell$ along the $z$-direction. Hence $v(q)$ is a partial Fourier transform or the three-dimensional Coulomb potential with respect to planar momentum $q$, which will be given by

$$v(q) = \frac{\epsilon^2}{4\pi\epsilon_d} \exp(-q\ell),$$

where the two surfaces are assumed to be at $z = 0$ and $z = \ell$. In the basis where the $H^\text{lab}_0$ is diagonal, the density operator at surface $\tau$ is represented as

$$\rho_{q\tau} = \sum_{\lambda\lambda'k} a_{q\tau\lambda}^\dagger \langle \tau\lambda'k | \tau\lambda k \rangle a_{\tau\lambda k},$$

where $k' = k + \mathbf{q}$ is understood, and the wave function overlap factors are

$$F_{\tau'\lambda'k',\tau\lambda k} \equiv \langle \tau' \lambda' k' | \tau \lambda k \rangle = \tau\tau' \frac{1 + \lambda' \lambda e^{i\phi_{\lambda'k'} - i\phi_{\lambda k}}}{2}.$$
Note that to obtain the above equation, the interactions in the channel that exchanges the \( \eta = \tau \lambda \) index of the incoming particles are essential. The equations of motions (11) and (12) suggest that the collective operator defined in the last line of eq. (12) satisfies the following collective mode equation:

\[
\frac{1}{v(q)} = \frac{1}{A} \sum_{k} n_{+\bf{k}'} - n_{-\bf{k}} \frac{\omega - (\epsilon_{k} + \epsilon_{k'})}{2} + \frac{1}{e^{\theta_{kk'}}} = \chi_0.
\]

(13)

where \( \bf{k}' = \bf{k} + \bf{q} \) is understood. Moreover \( \theta_{kk'} = \varphi_{k} - \varphi_{k'} \), and \( n_{+\bf{k}} \) is the expectation value \( \left\langle a^\dagger_\bf{k} a_{-\bf{k}} \right\rangle \). Due to the symmetry under \( \bf{q} \rightarrow -\bf{q} \) in the above equation, the \( \sin(\theta_{kk'}) \) does not contribute. The remaining integral is identical to those appearing in the polarization of undoped graphene which has been calculated by many authors [15], and gives rise to the following simple form:

\[
\chi_0 = \frac{q^2}{16\sqrt{q^2v_F^2} - \omega^2}.
\]

(14)

This bubble is real and positive inside the gap below the particle-hole continuum \( \omega < qv_F \). Hence the equation for the dispersion relation of the inter-surface spin-singlet exciton will become,

\[
\omega = qv_F \left( 1 - \tilde{\alpha}^2 e^{-2\eta q}\right)^{1/2},
\]

(15)

where \( \tilde{\alpha} = \frac{\tilde{\alpha}}{\sqrt{q}v_F} \) plays a role similar to fine structure constant. Indeed it can be written as \( \tilde{\alpha} = \frac{\alpha}{\sqrt{q} v_F} \), where \( \alpha \approx 1/137 \) is the fine structure constant. For Fermi velocities in the scale of \( 10^4 \sim 10^5 \) m/s in the Bismuth family [16] one expects the range of values \( \tilde{\alpha} \sim 10^{-2} \sim 10^{-1} \). According to this expression, an instability in the \( \bf{q} \rightarrow 0 \) limit can be triggered when \( \tilde{\alpha} \) becomes large which is only possible by finding materials with smaller \( v_F \). For the Bismuth family TIs, given the very small value of \( \tilde{\alpha} \), such instability does not seem feasible. The smallness of \( \tilde{\alpha} \) allows us to Taylor expand the right-hand side of the dispersion relation to get

\[
\omega \approx qv_F \left( 1 - \frac{1}{2} \tilde{\alpha}^2 e^{-2\eta q}\right),
\]

(16)

which defines the binding energy of the bosonic mode with respect to the lower edge of the free particle-hole continuum as

\[
E_B = qv_F \tilde{\alpha}^2 \frac{1}{2} e^{-2\eta q}.
\]

(17)

If the thickness \( \ell \) of the slab becomes very large, the exponential factor will diminish the binding energy and the bosonic mode will be absorbed to the lower edge of the continuum of free particle-hole pairs. However when the \( \ell \) is not very large, the dispersion relation (16) defines a relativistic branch of bosons. The boson velocity acquires the following form:

\[
\frac{v_B}{v_F} = 1 + \left( \eta q - \frac{1}{2}\right) \ell \tilde{\alpha}^2 e^{-2\eta q}.
\]

(18)

The maximum of the binding energy is obtained for \( 2\eta q = 1 \) and is \( E_B^{\text{max}} = \tilde{\alpha}^2 v_F/(4\ell) \). Below this point \( v_B < v_F \) and above it \( v_B > v_F \). At this point itself, \( v_B \) becomes equal to \( v_F \). Therefore for practical purposes in the low-momentum region defined by \( 2\eta q \lesssim 1 \), one may assume that the bosonic branch is almost degenerate with fermionic branch and has the approximate dispersion relation

\[
\omega = qv_B, \quad v_B \lesssim v_F.
\]

(19)

For \( \ell = 1 \) nm, we expect \( E_B^{\text{max}} \sim 10^{-2}\) meV. Although the direct observation of the above excitonic mode is difficult, but it may have indirect consequences on the spectral properties of the fermions themselves.

**Summary and conclusions.** — Using the random phase approximation implemented by the equations of motion we explicitly constructed spin-singlet \( \tau \)-triplet collective operators for a nano-meter thick film of topological insulator. The \( \tau \)-exchange part of the Coulomb interaction gave rise to an attraction between the positive-energy states of one surface and the negative-energy states of the other surface. Such attraction results in a binding bosonic mode below the continuum of free particle-hole excitations. The binding energy has a characteristic \( \exp(-\eta q) \) dependence which vanishes when the thickness \( \ell \) of the slab becomes very large and hence reduces to proper limit. Although the binding energy of the present inter-surface excitonic mode is a fraction of meV, but given its chiral nature, it is expected to leave Kerr and Faraday signatures in optical absorption [17]. Compared to the excitons in semiconductors where, within the same formulation, the spin-triplet exciton bound state obtains, here, the pseudo-spin \( \tau \) that emerges quite naturally, implies the exciton formation in spin-singlet channel. The latter is more accessible to optical experiments than the former. The self-energy corrections from such inter-surface excitons will indirectly show up in angular resolved photo-emission spectroscopy and scanning tunneling spectroscopy. Such effects can become more pronounced by approaching the excitonic instability [18].

Finally let us briefly discuss the finite-size effects. The analysis of the present study that assumes infinitely large surfaces is valid as long as the bound electron and hole pairs from the opposite surfaces do not hit the ends of the sample to wind around in the other surface. The characteristic energy scale in the present problem is 0.01 meV. This translates into time scales of the order \( 10^{-12} \) s. Given typical Fermi velocities of \( 10^5 \) m/s in Bismuth families or \( 10^6 \) m/s in typical metallic states of matter, one obtains a length scale of \( 10^2 \sim 10^3 \) nm. Therefore the present calculations makes sense in finite-size samples as long as the linear dimension of the surface is at least several micrometers.

***

This research was completed while the author was visiting Yukawa Institute for Theoretical Physics by...
the fellowship S13135 from Japan Society for Promotion of Science. Useful discussions with T. Tohyama, N. Kawakami, H. Fukuyama, M. Ogata, Y. Fuseya, A. Mishchenko, Bohm-Jung Yang and M. S. Bahramy are acknowledged. We thank W. Koshiba for hospitality during a short visit to RIKEN.

REFERENCES

[1] Zhang H. et al., Nat. Phys., 5 (2009) 439.
[2] Hasan M. Z. and Kane C. L., Rev. Mod. Phys., 82 (2010) 3045; Hasan M. Z., Su-Yang Xu, Hsieh D., Wary L. A. and Xia Y., arXiv:1401.0848v1 (2014).
[3] Shun-Qing Shen, Topological Insulators: Dirac Equation in Condensed Matters (Springer) 2013.
[4] Raghu S., Suk Bum Chung, Xiao-Liang Qi and Shou-Cheng Zhang, Phys. Rev. Lett., 104 (2010) 116401.
[5] Seradjeh B., Moore J. E. and Franz M., Phys. Rev. Lett., 103 (2009) 066402; Seradjeh B., Phys. Rev. B, 85 (2012) 235146; Efimkin D. K., Lozovik Yu. E. and Sokolik A. A., Phys. Rev. B, 86 (2012) 115436.
[6] Dmytro Pesin and Leon Balents, Nat. Phys., 6 (2010) 376.
[7] Baskaran G. and Jafari S. A., Phys. Rev. Lett., 89 (2002) 016402; Baskaran G. and Jafari S. A., Phys. Rev. Lett., 92 (2004) 19702.
[8] Jafari S. A. and Baskaran G., J. Phys. Condens. Matter, 24 (2012) 096600.
[9] Tsuchiya S., Ganesh R. and Nikuni T., Phys. Rev. B, 88 (2013) 014527.
[10] Hwang E. H. and Das Sarma S., Phys. Rev. B, 75 (2007) 205418.
[11] Hsieh D. et al., Nature, 460 (2009) 1101.
[12] Moon E. K. and Xu C., EPL, 97 (2012) 66008.
[13] Burkov A. A. and Balents L., Phys. Rev. Lett., 107 (2011) 127205.
[14] Demler E. and Shou-Cheng Zhang, Phys. Rev. Lett., 75 (1995) 4126.
[15] Son D. T., Phys. Rev. B, 75 (2007) 235423.
[16] Liu Y. and Allen R. E., Phys. Rev. B, 52 (1995) 1566.
[17] Efimkin D. K. and Lozovik Yu. E., Phys. Rev. B, 87 (2013) 245416.
[18] Rist Stefan, Varlamov A. A., MacDonald A. H., Fazio Rosario and Polini Marco, Phys. Rev. B, 87 (2013) 075418.