國立台灣大學理學院物理學研究所

博士論文

Department of Physics
College of Science
National Taiwan University
Doctoral Dissertation

二維拓樸絕緣體的電子自旋操控

Spin manipulation in two dimensional topological insulator

陳國進
Kuo-Chin Chen

指導教授: 張慶瑞博士
Advisor: Ching-Ray Chang, Ph.D.

中華民國 105年七月
July, 2016

doi:10.6342/NTU201603562

To my parents and aunt

謹以本論文獻給我的父親、 母親與阿姨

陳木胤先生，謝春貴女士與謝秀姻女士

doi:10.6342/NTU201603562

國立臺灣大學博士學位論文
口試委員會審定書

二維拓樸絕緣體的電子自旋操控

Spin manipulation in two dimensional topological insulator

本論文係陳國進君（F98222059）在國立臺灣大學物理學系、所完成之博士學位論文，於民國 105 年 07 月 20 日承下列考試委員審查通過及口試及格，特此證明

口試委員： 張慶瑞（簽名）
（指導教授）

林育中

胡崇德　　唐毓慧

陳松賢　　朱仲夏

Acknowledgment

就在特別炎熱的2016年暑假，我從博士班畢業了，在台大六加一年的時光裡，我從一個只讀過普通物理和近代物理的門外漢，慢慢地變成一個物理博士，在台大的這七年裡我學到了很多東西，不管是科學方面的知識或者是對人生體認。科學方面我學到的也不僅僅是對於電子自旋傳輸的了解，也學到了如何探索一個未知的領域，勇敢地花數個月的時間在一個不知道是不是好結果是一個很大的挑戰。在一個好的結果中往往有著多到數不清楚的失敗結果，但是不斷地嘗試新的方法來解決目前的問題是我覺得博士班中學到一個很重要的事情。

博士班是個漫長的道路，家人的陪伴是支撐我完成博士的動力，父親默默的付出讓我回家可以像廢人一樣地休息，母親不時的鼓勵並支持我做任何的人生決定，阿姨們無怨無悔地把我當兒子一樣的付出，還有祖母在我小時候的鼓勵則是啓蒙我勇敢追求博士的動力 ，介延這個從小到大的朋友也常常聽我發牢騷當我當心情的垃圾桶。

這篇論文的誕生，首先要感謝張慶瑞老師的指導，張老師不會限制我的想法並且給了我很大的空間，不時地提出建議讓這篇論文變得完整許多。而多次的出國報告機會也讓我從害羞不敢與人攀談到能夠勇敢地與陌生人侃侃而談自己的研究成果。在覺得有點無力的時候張老師也給了我很多的方向以及鼓勵。

doi:10.6342/NTU201603562

除了張老師之外，也要感謝林育中老師的指導，林老師嚴謹的學術態度在組上開會的時候提出許多本來沒想過的問題以及建議都是很精闢的，口試前與張老師多次的提出我報告表達時的缺陷並且提出改進的建議也讓我有機會可以改進自己報告時的缺點。 在博班期間還有許多想感謝的老師，中研院的關肇正老師讓我瞭解了第一原理，NUS的Mansor和 Seng Ghee在我去新加坡交換時也時常和我討論，接觸國際的研究學者讓我科學的眼界大開。

人生的體認則要引用劉彥甫博士的論文致謝“這是一個人生的博士”。在博士班期間慢慢地瞭解自己要追求怎麼樣的人生，也結交了許多好朋友，不管在我求學的時候或者是做一些人生抉擇的時候都給我很大的幫助。 甫哥是個樂於幫助所有人的學長，其豐富的理論背景給了我論文上很多的建議，他更是個保有赤子之心的好友，在我徬徨無助的時候總是義氣相挺的聽我說發一次又一次的牢騷，並且給了我無數多的建議，在我去美國開磁性會議的時候一起在加州闖蕩也給我相當多的幫助，是我在研究和生活上的摯友。

培瑋，思圻與一禎是凝態中心的好朋友，雖然身處不同研究室，不過也給我相當多的建議，不管是我論文的改進或是對我面對一些人生的抉擇都有很大的幫助，在我人生中最低潮的時候不時地伸出援手，一起用歡樂的態度度過博士中最難熬的最後這幾年。不時的騎著腳踏車出遊，在大太陽中辛苦的揮灑汗水如今都是甜美的回憶，墾丁與小琉球的求婚之旅更是讓我充飽了電面對眾多的挑戰，希望你們未來人生的道路順遂，我們的友誼長存，以後還有機會一起出遊。

同研究室的學長欣翰，建良和盛哥，是我博班剛進來研究上的小導師，從什麼都不懂到進入這個領域給我莫大的幫忙，和欣翰去韓國開會時在異鄉闖蕩是個難忘的經驗 。 松賢學長是學術上的好榜樣，任何疑難雜症請問他馬上有豁然開朗的舒暢感。明鑑與景皓則是在我博班的後期給了我相當多的意見，眾多的討論讓

doi:10.6342/NTU201603562

我完成了這最後一個題目。501時代就認識的柏文，黃黃，港仔和斌哥是以前玩橋牌的好戰友，港仔喊出seven no king的事蹟會一輩子記住的。 還有許多沒有提到的朋友也大大的豐富了我的生活，謝謝你們。

KUO-CHIN, CHEN

National Taiwan University
Department of Physics
July, 2016

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

Abstract

This thesis presents a theoretical study of spin manipulation in a two-dimensional topological insulator (2DTI). Non-equilibrium Greens function approach and Landauer Büttiker Formalism cooperates with tight binding band calculation numerically study the electron spin transport properties of two-dimensional electron gas with strong spin-orbit coupling. A topological insulator is a material that a strong intrinsic spin-orbit interaction exists. The non-dispersive edge states make it be a promising material in spintronics application. Adopting non-magnetic field control is predominating in the research field of semiconductor devices, thus, we adopt non-magnetic field to control electron spin in the two-dimensional topological insulator. The thesis provides two ways to manipulate electron spin in a two-dimensional system.

First, quantum interference induced by external electric field is adopted. A persistent quantum resonance device is proposed in an H-shaped 2DTI embedded a non-magnetic impurity at the center. Transmissions between each branch of the H-shaped 2DTI shows two kinds of quantum resonance in this device, Breit-Wigner resonance, and Fano-like resonance. These resonances can be realized in the device through modulating the onsite impurity potential. A phase transition between the Fano-like and the Breit-Wigner resonances through modulating the thickness of the 2DTI leads is also presented.

doi:10.6342/NTU201603562

Second, we present a band study of a 2DTI-normal metal junction. The helical edge state of 2DTI hybridized with the quantum well state of normal metal is well studied. A systematical study of the band in terms of the coupling strength between 2DTI and normal metal shows there are two interesting phenomena in this junction (i) A helical state induced splitting that similar to Rashba field existed in normal metal. The Rashba-like field generates a spin precession that the precession length can be modulated by the coupling strength. The Rashba-like field can be a promising way to create giant Rashba spin orbital via material manipulation. (ii) The band of spin down opens due to the spin down electron of normal metal penetrated to 2DTI and backward moved restrictedly. Energy band gap opens for one spin channel thus a polarized spin current flows into normal metal in the energy region of the band gap. By modulating the Fermi energy, it is possible to convert the quantum spin hall system into a spin filter.

doi:10.6342/NTU201603562

摘要

本論文的研究主題是在二維拓樸絕緣體中操控自旋電子，採用tight binding model配合Non-equilibrium Greens function以及Landauer Büttiker Formalism探討自旋電子的傳輸行爲，並且配合tight binding的能帶計算來研究拓樸絕緣體這一個有強自旋軌道耦合的材料，其非耗散的表面態受到時間反轉對稱守恆的保護，使此材料有很大的機會應用在自旋電子元件上面。利用電場來控制電子的傳輸在半導體工業已經發展得相當成熟。本論文則討論兩種方法用電場來控制在二維拓樸絕緣體中電子自旋。

第一種方法是利用局部的電場產生量子干涉並且進一步地控制穩固非耗散自旋流的開關 。我們用電場調控一個放置在工字形二維拓樸絕緣體中間的雜質能量，電極之間的傳輸係數在不同電極寬度下展現出Fano-like 共振或者Breit-Wigner 共振，而此兩種共振可以藉由調控電極寬度產生相變。

第二種方法是在二維拓樸絕緣體接薄金屬，二維拓樸絕緣體的表面態會和金屬的量子井態混成，有兩個特性是值得被討論的，第一個就是此混成造成自旋的分裂，產生了類似Rashba自旋軌道耦合的效應，第二個則是混成造成能隙的打開，此混成在正（負）的動量時讓自旋向下（上）電子的能帶產生能隙，因此藉由調控能量，可以將這個系統變成自旋閥的元件。

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

List of Publication

[1]. Kuo-Chin Chen and Ching-Ray Chang. Geometrical effect on spin transport. SPIN Vol. 3, No. 3 (2013) 1340006.

[2]. Kuo-Chin Chen, Yu-Hsin Su, Son-Hsien Chen, and Ching-Ray Chang. Non-equilibrium study of spin wave interference in systems with both Rashba and Dresselhaus (001) spin-orbit coupling. Journal of Applied Physics **115**, 17C305 (2014)

[3]. Kuo-Chin Chen, Hsin-Han Lee, and Ching-Ray Chang. Persistent quantum resonance transition in spin Hall transport . Phys. Rev. B **93**, 035405 (2016)

[4]. Kuo-Chin Chen, Ching-Hao Chang, Ming-Chien Hsu, and Ching- Ray Chang. Normal metal and two dimensional topological insulator junction(in progress)

doi:10.6342/NTU201603562

Contents

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

List of Figures

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

doi:10.6342/NTU201603562

1

Introduction

> *Condensed matter physics is the field of physics that deals with the macroscopic and microscopic physical properties of matter. In particular, it is concerned with the condensed phases that appear whenever the number of constituents in a system is extremely large and the interactions between the constituents are strong. The most familiar examples of condensed phases are solids and liquids, which arise from the electromagnetic forces between atoms.*
>
> – Wikipedia

Almost all of the world that we can see is condensed matter. Such as the cause of the different electrical conductivity of conductor and insulator, the magnetization of the ferromagnetic and antiferromagnetic material, and the superconducting behavior in low temperature all can be explained by theory and verified experimentally. In fact, the condensed matter physics does not only satisfy the curiosity of humankind but also made many remarkable achievements for human technologies such as engineer new materials, superconductor, optical physics etc. The most re-

doi:10.6342/NTU201603562

markable achievement in this field is the application in the semiconductor industry, which enabled electronics devices such as computers, monitor, and everything else we often used to go into our lives. However, in the semiconductor industry, the length scale has a limitation. The device in the semiconductor industry is based on complementary metal-oxide- semiconductor (CMOS) [1]. In CMOS, the leakage currents are increasing dramatically in sub-100nm processes. So we need another approach to manipulating the electron. The spin is an intrinsic degree of freedom of an electron that has raised much attention in condensed matter physics. The spin comes out of the application starting from the giant magnetic resistance[2, 3] in 1988 that gives a significant progress of data storage.

1.1 Spintronics in heterojunction

The systems of the thesis are built on heterojunction. In this chapter, we come on give a brief introduction of heterojunction first. And we then give a model to illustrate Rashba spin-orbit interaction [4]. Finally, we will focus on a special spin-orbit interaction that exists in HgTe/CdTe heterojunction is called a two-dimensional topological insulator (2DTI) [5]. 2DTI is a unique material with strong spin-orbit interaction with absorbing spin transport behavior, spin-momentum locking, and topological protection [6].

1.1.1 Heterojunction

A heterojunction is the interface of two different conjunct layers of crystalline semiconductors with alternating band gaps and can be utilized in electrical devices such as solar cells, lasers, and complementary metal-oxide-semiconductor. The behavior of the alignment of energy bands determines the behavior of a heterojunction. There

doi:10.6342/NTU201603562

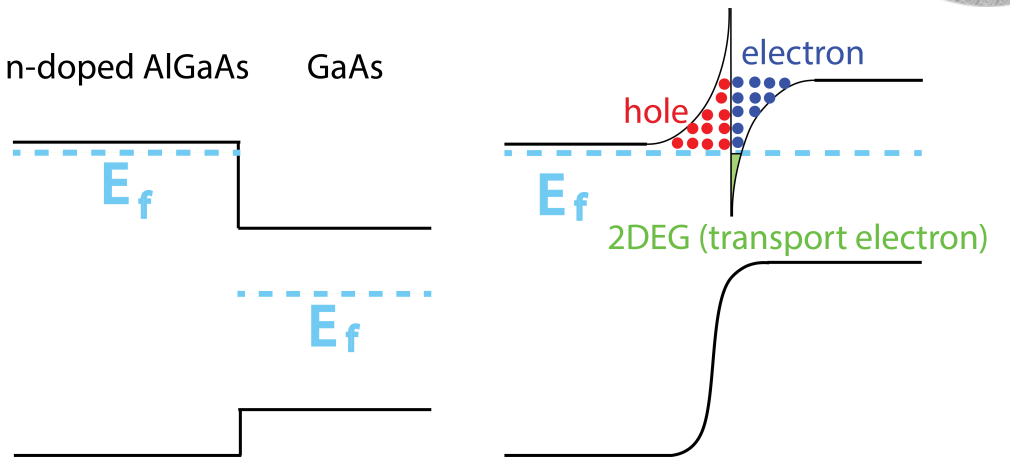

Figure 1.1: (Left) Band diagram for an AlGaAs-GaAs interface in terms of the direction that perpendicular to the interface. Band gap and Fermi energy are different for AlGaAs and GaAs and the energy for AlGaAs and GaAs has not been equilibrium. (Right) a two-dimensional electron gasses induced by redistribution of charge.

are three types of energy bands that define three types of heterojunctions straddling the gap, staggered gap, or broken gap. In the case of the type of staggering gap, the band gap is different from the two semiconductors. The conduction band of the smaller gap semiconductor lies above the conduction band of the larger gap semiconductor. We are going to focus on the staggering gap and take AlGaAs-GaAs as an example and showed in Fig. 1.1. AlGaAs is an n-doped semiconductor that Al give electronic donors, however, GaAs is undoped so that the chemical potential is lower for conduction band that the schematic figure Fig. 1.1 shows the energy band diagram in terms of the position of the perpendicular direction of the heterojunction. This is unstable near the interface due to the difference of the chemical potentials for two semiconductors. The two systems are in a nonequilibrium state and the electron in n-doped AlGaAs will move into GaAs so that the system would be more stable. After the migration of the electron positively, ionized donors generated near

doi:10.6342/NTU201603562

the interface on the side of AlGaAs due to the electron would no longer screen Al atoms. The electron would be trapped on the side of GaAs near the interface thus the charge is negative near GaAs and is positive near AlGaAs.

How does the band become to correspond the migration of the electron? The slope of the potential is the electric fields and the sign of curvature is the sign of charge. The direction of electric field is from n-doped AlGaAs to undoped GaAs due to the charge distribution. The charge is positive in the region of AlGaAs and is negative in the region of GaAs, so the sign of the curvature is positive in the region of AlGaAs and is negative in the region of GaAs. The schematic figure shows the band after the migration and the Fermi energy now below than the conduction band of AlGaAs and lower than the conduction band of GaAs. The interface is a 2 dimensional quantum well with electric field perpendicular to the plane. A two-dimensional electron gasses would be trapped in the heterojunction.

1.1.2 Rashba spin-orbit interaction

Spin-orbit interaction is the core of the thesis and is utilized to control the spin transport [7, 8, 9]. We consider the linear term of spin-orbit interaction in the dissertation that is an additional term of Hamiltonian. A linear combination of product of spin operator and momentum operator can express spin-orbit interaction

$$H_{so} = \sum_{i,j} S_i * P_j \tag{1.1}$$

where S_i are the spin operators and P_j are the momentum operators.

The origin of spin-orbit interaction is complicated that depend on the material. Nevertheless, we can give a brief example to illustrate Rashba spin-orbit coupling that is a special kind of spin-orbit coupling can be found in GaAs/AlGaAs het-

doi:10.6342/NTU201603562

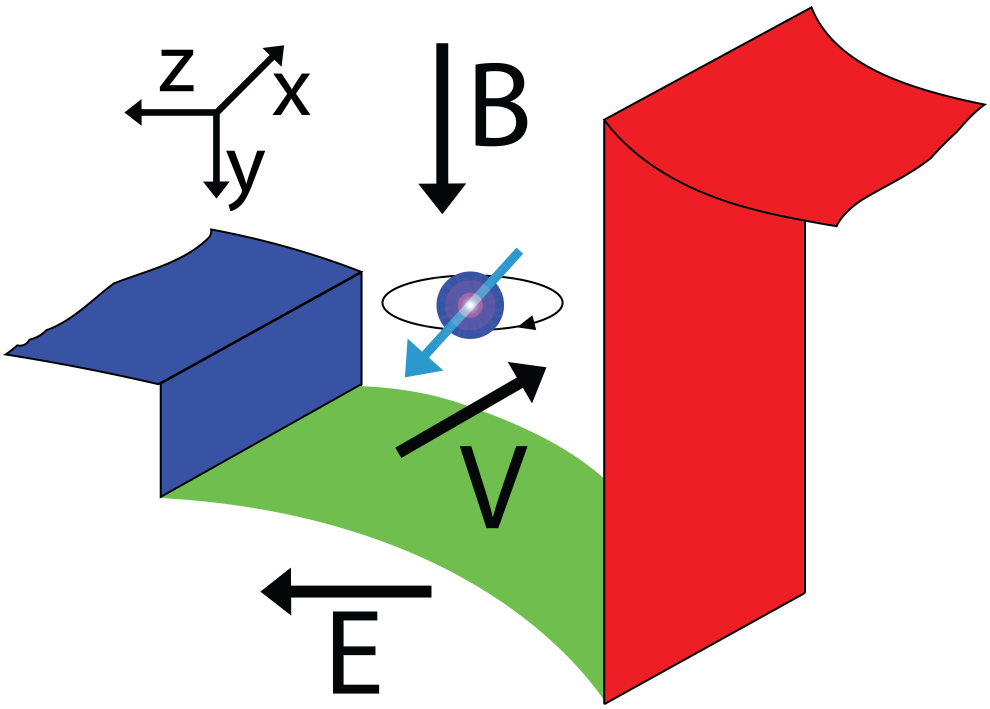

Figure 1.2: Schematic of an electron moving at velocity v in a constrained two dimensional systems. The different chemical energy induces an electric field that is perpendicular to the two dimensional system

doi:10.6342/NTU201603562

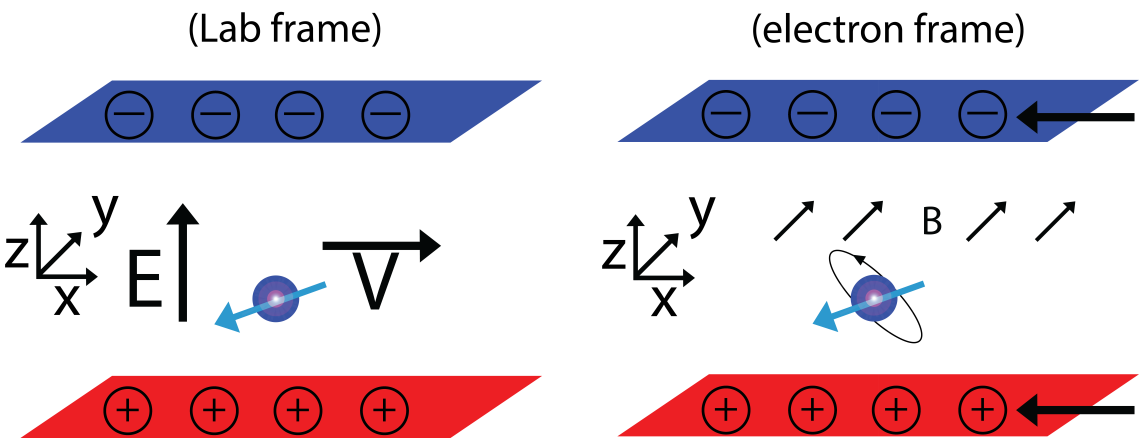

Figure 1.3: (Left) Schematic of an electron moving at a velocity v to a two-dimensional plane that in the between of positively charged plane and negatively charged plane in the lab frame with an electric field that is generated by the charged plane. (Right) In the frame of the electron, positively charged plane and negatively charged plane in lab frame become charged current and induced an effective magnetic field B_{eff}.

doi:10.6342/NTU201603562

erojunction. The Hamiltonian form of linear Rashba spin-orbit interaction can be expressed as

$$H_{so} = P_y * S_x - P_x * S_y \tag{1.2}$$

Now we about illustrate how does the GaAs/AlGaAs heterojunction generate Rashba spin-orbit coupling. A two-dimensional electron gasses can be formed in the GaAs/AlGaAs heterojunction and a positive charge and negative charge distribution on the side of AlGaAs and GaAs respectively as shown in Fig. 1.2. Assume the charge distribution is uniform so that the heterojunction can be regarded as a charged capacitor. An electric field E that is perpendicular to the plane of the two-dimensional electron gas as shown in Fig. 1.3.

$$E = E_z\hat{z} \tag{1.3}$$

The electron in the two-dimensional electron gas is restrictedly moving in the x-y plane within the electric field. If we are located in the frame of moving electron, a hole and electron current is created on the side of AlGaAs and GaAs respectively. We can now regard this system as uniformly distributed positive electric current and negative electric current flow on the top and down of the electron. The magnetic field B is an effective field in the frame of moving electron that can be expressed as

$$B = V \times E \tag{1.4}$$

and the Hamiltonian can be expressed by exchange coupling

$$H_{so} = -S\cdot B = -S\cdot(V\times E) = -S\times(P\cdot E) = E\cdot S\times(P_z) = E(P_y\cdot S_x - P_x\cdot S_y) \tag{1.5}$$

where P is the momentum operator. We have explained the Rashba spin-orbit coupling exist in the heterojunction. Now we are going to illustrate another material, HgTe/CdTe heterojunction, that a distinctive spin-orbit coupling can be found in the material. We will give a brief introduction of this heterojunction in the next subsection.

doi:10.6342/NTU201603562

1.2 Topological insulator

1.2.1 Topology

A topological insulator is a material with non-trivial topological order behaves as an insulator in its bulk, however, conducting surface states will be induced near the surface that is the interface of non-trivial topological order (1) and trivial topological order (0). This term "topology" is a branch of mathematics. That is the field of investigation of properties of space that are preserved under continuous deformations.

> *Topology can be formally defined as "the study of qualitative properties of certain objects (called topological spaces) that are invariant under a certain kind of transformation (called a continuous map), especially those properties that are invariant under a certain kind of transformation (called homeomorphism)."*
>
> – Wikipedia

One famous example is that a continuous deformation of a mug into a doughnut (torus) as shown in Fig. 1.4. The topology of a mug and a doughnut are the same, nevertheless, these shapes are changed apparently. The summation of Gaussian curvature K of the surface of a mug or a doughnut is the same. The Gaussian curvature K is the product of the two principal curvatures k_1 and k_2. The principal curvatures at p denoted k_1 and k_2, are the maximum and minimum values of this curvature as shown in Fig. 1.5.

$$K = k_1 * k_2 \tag{1.6}$$

The summation of K is actually a topological invariance.

doi:10.6342/NTU201603562

Figure 1.4: A continuous deformation of a mug into a doughnut. A flexible doughnut is reshaped to a mug could be reshaped to a coffee cup by creating and enhancing a dimple by degrees, and shrinking the hole into a handle simultaneously.

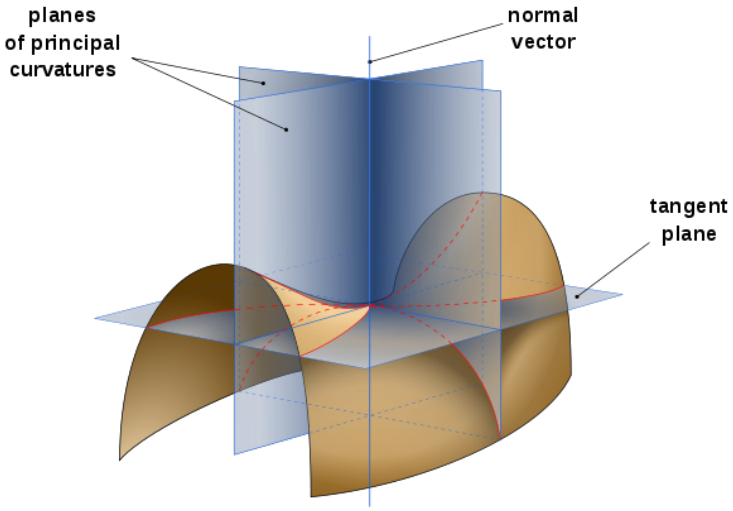

Figure 1.5: a saddle surface with normal planes in directions of the principal curvatures. The curvature at a given point along a given direction is defined as the curvature of the intercepted curve of a plane contains the loop and the surface at that given point. The principal curvatures actually can be viewed as the maximum and minimum values of this curvature of a given point.

doi:10.6342/NTU201603562

What is the application of topology in physics? The summation of a certain quantity in momentum space is topological invariance and deformation (small interaction) cannot change the topology thus preserve the physics properties. The TKNN number in the quantum hall effect [10] and Z_2 number of quantum spin hall effect [11] are the topological invariance number in momentum space which leads to topological protection. We have given a brief introduction of topology and the relation with physics. In the next section, we will introduce a material that has quantum spin hall effect, the topological insulator, and illustrate the important physical properties.

1.2.2 Topological insulator

Topological insulators are materials that the bulk electronic band structure is an ordinary band insulator with the Fermi energy lays in the gap that in between of the conduction and valence bands, however, linear conducting bands exist in the band gap and it is topologically protected states accumulated on their edge [6] . These conducting edge states are due to the spin-orbit interactions and preservation of time-reversal symmetry. The linear surface band can be described as a Dirac electron by modified Dirac equation [12] and nearly massless that preserve excellent transport behavior just like graphene [13]. There are two characters that make topological insulator become a fascinating material. The first one is spin-momentum locking due to unusual spin-orbit coupling. The second one is the surface conducting state is topological protected. These two characters make the topological insulator become a promising material in the application of spintronics.

doi:10.6342/NTU201603562

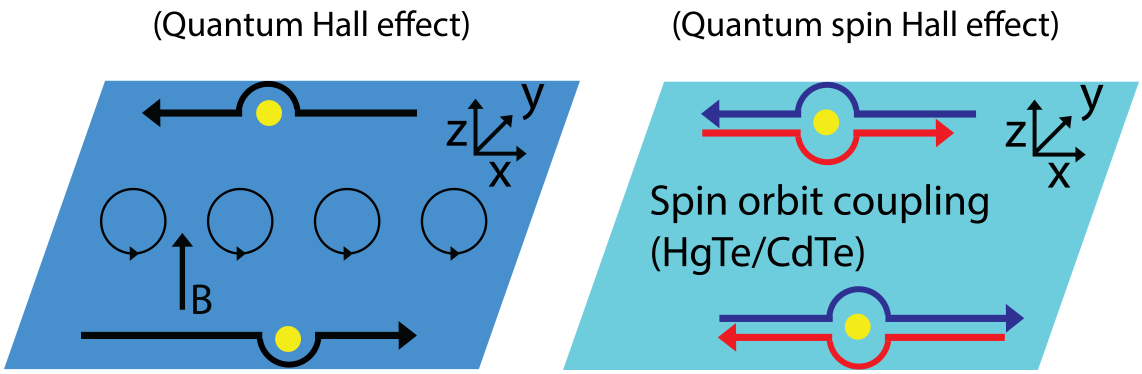

Figure 1.6: (Left) Electron moving in quantum Hall system that can be realized by adding a magnetic field in a two-dimensional system. The motion of an electron is cyclotron so the electron can only move forward at the lower edge. Due to lack backward state at the lower edge. The states are robust such that encountering an impurity without scattering. (Right) an electron moving in quantum spin Hall system that can be realized in a quantum well of HgTe/CdTe that is proposed as a topological insulator. Instead adding a magnetic field, the intrinsic SOC induces 4 basic degrees of freedom that is spatially separated at the edges. Spin up and down electron both can move in the upper and lower edges, however, in the opposite moving direction. The same analogy of quantum Hall system, the lacking of backward state of spin down electron at the lower edge, the edge states are robust against impurity.

doi:10.6342/NTU201603562

1.2.3 Spin momentum locking

Spin momentum locking is that the spin and momentum are perpendicular due to strong spin-orbit interaction. Surface states of a topological insulator are spin-momentum locking so that the electron is spin-polarized when the electron is transported. The band structure shows spin evolves in momentum space. We may take HgTe/CdTe quantum well as an example that is a two-dimensional topological insulator. The orientation of spin is in z axis and the electron restrictedly moves in the x-y plane. For one edge, the spin up electron is in the positive slope of the band and spin down is in the negative slope of the band. That means the direction of electron movement determines the orientation of spin. Furthermore, the spin orientation of opposite movement of the electron is opposite as shown in Fig. 1.6. However, the spin-momentum locking is unique for topological insulators. It can be found in Au(111) surface [14], the cooperate with topological protecting is actually made topological insulator be so special.

1.2.4 Topological protection

Topological protection is announced in the two-dimensional electron gas system with a strong magnetic field in the quantum system. The magnetic field B can be regarded as an electromagnetic vector potential A that changes the phase of the wavefunction and make a cyclotron orbit of the electron. Landau quantization can be derived by solving the Hamiltonian so that the electron can only occupy orbits with discrete energy values, called Landau levels. States of Landau levels are accumulated at the edge of two-dimensional electron gas systems and cannot be scattered back due to the cyclotron motion as showed in Fig. 1.6.

Topological protection in a topological insulator is based on the preservation

doi:10.6342/NTU201603562

of time reversal symmetry (i.e. without a magnetic field) instead. The surface state cannot be removed by surface non-magnetic impurities due to the preservation of time reversal symmetry, so the property makes the application become more feasible. We may be wondering what will go on if an impurity near the surface. The topological protection gives us a simple physical picture as showed in Fig. 1.6.

doi:10.6342/NTU201603562

Bibliography

[1] Kikuchi, Hideo, Haruyoshi Takaoka, and Shigenori Baba. "Complementary metal-oxide semiconductor." U.S. Patent No. 4,288,804. 8 Sep. 1981.

[2] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, Phys. Rev. Lett. **61**, 2472 1988.

[3] G. Binasch, P. Grunberg, F. Saurenbach, and W. Zinn, Phys. Rev. B **39**, 4828 1988 .

[4] Bychkov, Yu A., and Emmanuel I. Rashba. "Oscillatory effects and the magnetic susceptibility of carriers in inversion layers." Journal of physics C: Solid state physics **17**.33 (1984): 6039.

[5] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[6] Qi, Xiao-Liang, and Shou-Cheng Zhang. "Topological insulators and superconductors." Reviews of Modern Physics **83** (2011): 1057.

[7] Datta, S., and Das, B. Applied Physics Letters **56**, 665 (1990).

[8] Liu, Ming-Hao, and Ching-Ray Chang. "Datta-Das transistor: Significance of channel direction, size dependence of source contacts, and boundary effects." Physical Review B **73** (2006): 205301.

doi:10.6342/NTU201603562

[9] Chuang, Pojen, et al. "All-electric all-semiconductor spin field-effect transistors." Nature Nanotechnology **10**, 35-39 (2015)

[10] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).

[11] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005)

[12] Shen, Shun-Qing. Topological Insulators: Dirac Equation in Condensed Matters. Vol. 174. Springer Science and Business Media, (2013).

[13] Neto, AH Castro, et al. "The electronic properties of graphene." Reviews of modern physics **81**, 109 (2009)

[14] LaShell, S., McDougall, B., and Jensen, E. (1996). Spin splitting of an Au (111) surface state band observed with angle resolved photoelectron spectroscopy. Physical review letters, 77, 3419.

doi:10.6342/NTU201603562

2

Method

2.1 Construction of Hamiltonian(Finite difference method)

In this section, we construct the spatial resolved Hamiltonian matrix **H** describing a two-dimensional electron system. The Hamiltonian we described in the thesis is an effective Hamiltonian derived from the first principle calculation or fitting from the experiment. The Hamiltonian is usually represented in momentum space, however, we would like to know the spatial information in the transport aspect. By adopting a spatially resolved basis, position of lattice sites. Allows us to construct a spatial resolved Hamiltonian. One thing that I need in order to mention here is that these sites do not represent the realistic crystal lattice of atoms. The numbering of the sites does not alter the physical properties, nevertheless, the numbering is important in the numerical calculation. Now we are going to illustrate the way to change the basis from momentum space (continuous) to spatial space (discretized).

The finite difference method is a numerical method to approximate differen-

doi:10.6342/NTU201603562

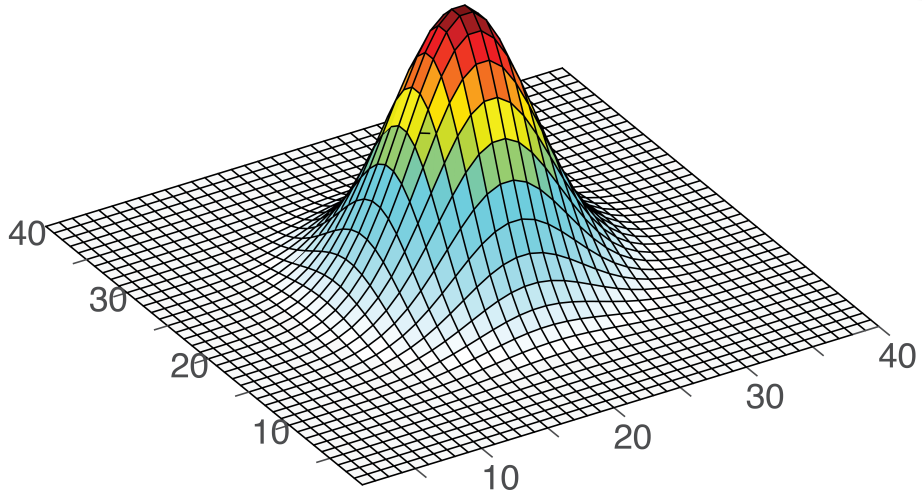

Figure 2.1: Discretization of wave function

tial equation with difference equations. The difference equations allow us to solve differential equation numerically.= We start this section by illustrating the finite difference method by taking one-dimensional free electron gas without the spin degree of freedom as an example. The Hamiltonian describing one dimensional (x-axis) free electron can be given by

$$H\varphi(x) = \frac{p_x^2}{2m}\varphi(x) \tag{2.1}$$

where p_x is the momentum operator along the x-axis that is a spatial derivative operator can be expressed as

$$p_x\varphi(x) = -i\hbar\frac{d}{dx}\varphi(x) = -i\hbar\lim_{a\to 0}\frac{\varphi(x+a/2)-\varphi(x-a/2)}{a} \tag{2.2}$$

The derivative of the wave function can be approximated as the variation of wave function if a is small enough. The approximation is valid actually when a is much less than the Fermi wave vector.

$$p_x\varphi(x) \approx -i\hbar\frac{\varphi(x+a/2)-\varphi(x-a/2)}{a} \tag{2.3}$$

doi:10.6342/NTU201603562

If the momentum operator p_x act on the (2.3) again, we have the wave function on which kinetic term p_x^2 act on

$$p_x^2\varphi(x) \approx -\hbar^2\frac{\varphi(x+a)-2\varphi(x)+\varphi(x-a)}{a^2} \tag{2.4}$$

Above approximation is the finite difference method. We can express the Hamiltonian by local spatial wave function with the help of finite different method as following. The spatial wave function $\varphi(x)$ is a summation of all local wave function $\varphi_n(x)$ given by

$$\varphi(x)=\sum_n \varphi_n(x) \tag{2.5}$$

Where n is the index of the site, so $\varphi_n(x)$ is the wave function at site n. Using the relation above, we can know how does the Hamiltonian operator act on wave function

$$\begin{aligned} H\varphi(x) &= H\sum_n \varphi_n(x) = \frac{p_x^2}{2m}\sum_n \varphi_n(x) \\ &= -\frac{\hbar^2}{2m}\sum_n \frac{\varphi_n(x+a)-2\varphi_n(x)+\varphi_n(x-a)}{a^2} \end{aligned}$$

So now the Hamiltonian can be expressed as the local wave function by matrix that adopts $\varphi_n(x)$ as basis

$$\begin{aligned} H_{m,n} &= \varphi_m^\dagger(x')H\varphi_n(x) \\ &= -\frac{\hbar^2}{2m}\varphi_m^\dagger(x')\frac{\varphi_n(x+a)-2\varphi_n(x)+\varphi_n(x-a)}{a^2} \\ &= -t_0\varphi_m^\dagger(x')\left[\varphi_n(x+a)-2\varphi_n(x)+\varphi_n(x-a)\right] \end{aligned}$$

where $t_0=\frac{\hbar^2}{2ma^2}$, is defined as hoping energy. Assume the wave function is not so extensive in space, only the nearest overlapping have to be accounted (tight binding

doi:10.6342/NTU201603562

approximation), we have

$$H_m = \begin{cases} -t_0 & \mathbf{m} = \mathbf{m} \pm \mathbf{a} \\ 2t_0 & \mathbf{m} = \mathbf{m} \end{cases} \tag{2.6}$$

We can easily extend the 1-D spin independent system to a 2-D Rashba spin orbit system that we have discussed in Chapter1

$$H_{\mathbf{m},\mathbf{m}'} = \begin{cases} -t_0 I_s - i t_r \sigma^y, & \mathbf{m} = \mathbf{m}' + \mathbf{a_0 e_x} \\ -t_0 I_s + i t_r \sigma^x, & \mathbf{m} = \mathbf{m}' + \mathbf{a_0 e_y} \end{cases} \tag{2.7}$$

Now I am going to give another point of view of the discretized Hamiltonian. We use the second quantization to express the Hamiltonian as following. First we would like to define Fermion operators a_i and $a_i^\dagger$. $a_i^\dagger$ creates a fermion in the site i and a_i annihilates a fermion in the site i. The Fermion operators obey the following relations:

$$\begin{aligned} a_i a_j + a_j a_i &= 0 \\ a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger &= 0 \\ a_i a_j^\dagger + a_j^\dagger a_i &= \delta_{i,j} \end{aligned} \tag{2.8}$$

The occupation number operator whose eigenvalues are the number of particles in state i is

$$\begin{aligned} n_i &= a_i^\dagger a_i \\ 1 - n_i &= a_i a_i^\dagger \end{aligned} \tag{2.9}$$

With the help of second quantization, we can express the wave function by the creation and annihilation operator.

$$\psi(r) = \sum_i u_i(r) a_i \tag{2.10}$$

Where $u_i(r)$ is the single-particle wave functions corresponding to site i. The Hamiltonian of the notation of second quantization can be expressed as following

$$H = \sum_n \epsilon_n a_n^\dagger a_n + \sum_{i,j} t_{i,j} a_i^\dagger a_j, \tag{2.11}$$

doi:10.6342/NTU201603562

Where $\epsilon_n = \int u^\dagger H u dr$ and $t_{i,j} = \int u_i^\dagger H u_j dr$. We adopt the tight-binding approximation in the thesis, so $t_{i,j} \neq 0$ as i and j are the nearest neighborhood. The approximation makes sense if the wave function is not extended in real space, we can ignore the second nearest site interaction.

We can now give another viewpoint of the second quantized Hamiltonian. The electron hop from one site to another site via $t_{i,j}$ and onsite energy ϵ_n. If we use the operator to extend Hamiltonian, then the matrix representation of it is

$$H_{i,j} = \epsilon_i \delta_{i,j} + t_{i,j} \tag{2.12}$$

Note the form of second quantized Hamiltonian is similar to the finite difference Hamiltonian that we have obtained.

2.2 Periodical structure (Bloch Theorem)

an electron in a crystal structure is in a periodical field, so the infinite system can be simplified by adopting Bloch theorem. The statement of Bloch theorem can be expressed as following: For an electron constraint in a periodical field $U(r)$ with a period R as showed in Fig. 2.2. The Hamiltonian H acting on a wave function ψ can be expressed

$$H\psi = \left[-\frac{\hbar^2}{2m}\nabla^2 + U(r) \right] \psi \tag{2.13}$$

where $U(r)$ is a periodical potential with the period R

$$U(r) = U(r + R) \tag{2.14}$$

The translational symmetry reduce a degree of freedom so that the eigenstates ψ of the Hamiltonian H can be chosen to have the form

$$\psi_(r + R) = e^{ikR}\psi(r) \tag{2.15}$$

doi:10.6342/NTU201603562

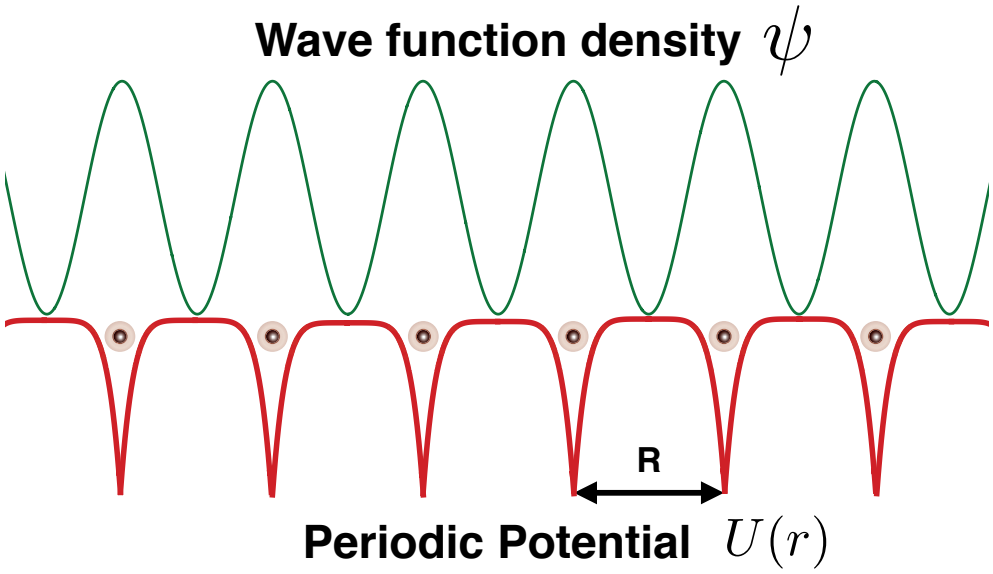

Figure 2.2: (Bloch Theorem) Electron in periodic potential have periodical density distribution

If we adopt the second quantization to express a wave function $\psi(r) = \sum_i u_i(r)a_i$, where i is the degree of freedom of the periodical direction. Then we have the relation

$$u_i a_i = u_{i+1} a_{i+1} e^{ikR} \tag{2.16}$$

2.2.1 Band structure and wave density

Now we are going to apply Bloch theorem in a system which is infinite and periodical along the x-axis, however, finite size in the y-axis as showed in Fig. 2.3. Adopting the tight binding approximation, the Hamiltonian can be written as

$$H = \sum_n \epsilon_n a_n^\dagger a_n + \sum_{<i,j>} t_{i,j} a_i^\dagger a_j \tag{2.17}$$

Where ϵ_n is onsite energy of each sites, and $t_{i,j}$ is the hopping energy between sites and $< i, j >$ represents the nearest neighbor. The periodic system is infinite so it is

doi:10.6342/NTU201603562

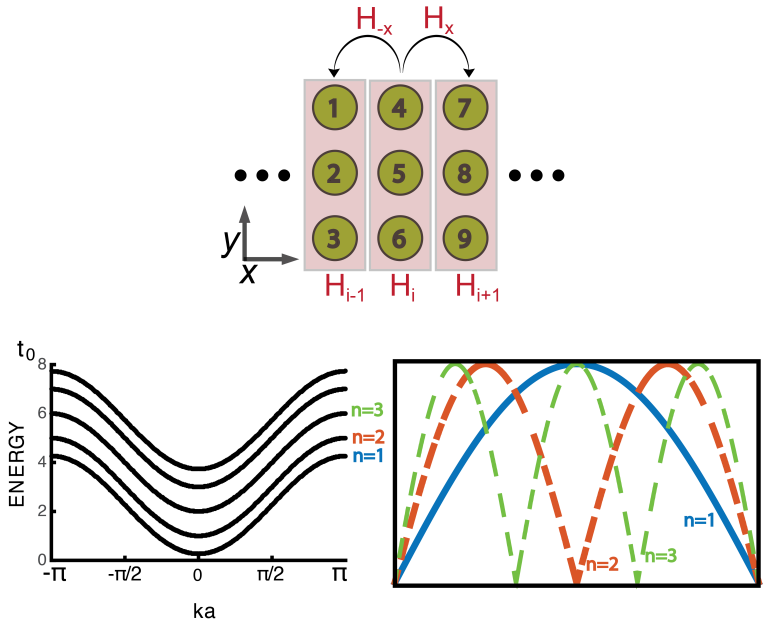

Figure 2.3: Now we can solve the eigenvalue and eigenvector of the Hamiltonian. The eigenvalue is the band structure and eigenvector is the correspondent wave function. Now let us take 2DTI as an example

doi:10.6342/NTU201603562

hard to calculate the eigenenergy of the overall system. Thus we can use the Bloch theorem to reduce the calculation as following. Taking Fig. 2.3 as an example, we can define a unit cell named H_i, are the same for all i, couple with the nearest unit cell via a coupling matrix H_x and H_{-x} that represented the unit cell hop to another unit cell along x and $-x$ respectively. For example H_i couple with $H_{i\pm1}$ via $H_{\pm x}$ in Fig. 2.3. The unit cell H_i are the same as the system is periodical and we can also use Bloch theorem eqrefbloch to proof it trivially. Without loss of generality, we use H_0 to replace H_i and define H_0 is

$$H_0 = \sum_i \epsilon_i a_i^\dagger a_i + \sum_{<i,j>} t_{i,j} a_i^\dagger a_j \tag{2.18}$$

i, j is the label of the site in the unit cell. $H_x^\dagger = H_{-x}$ due to Hamiltonian must be Hermitian operator and can be defined as

$$H_x = \sum_m t_{m,m+1} a_m^\dagger a_{m+1} \tag{2.19}$$

Where m label the site of H_i and $m+1$ label the site of H_{i+1}. m and $m+1$ are the nearest neighbor. We can use Bloch theorem to transform a_{m+1} to a_m by adding a factor e^{ikR}, so we can rewrite H_x

$$H_x = \sum_m t_{m,m+1} a_m^\dagger a_{m+1} = \sum_m t_{m,m+1} a_m^\dagger a_m e^{ikR} \tag{2.20}$$

Now we can represent the overall Hamiltonian by H_0 and H_x by

$$H = H_0 + H_x e^{ikR} + H_{-x} e^{-ikR} \tag{2.21}$$

The unit cell of a periodical stripe can be described by H_0 and the coupling matrix between each unit cell H_x and H_{-x}.

Now we use an example to illustrate how to get the band structure. If the unit cell is 5 site in y-direction and is periodical in x-direction. Without loss of generality,

doi:10.6342/NTU201603562

we can assume the field is isotropy, then the hopping energy of x and y-direction is the same. Then we can express H_0

$$H_0 = t_0 * \begin{bmatrix} 4 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 4 \end{bmatrix} \tag{2.22}$$

where t_0 is the hopping energy and we set the onsite energy to be $4t_0$. The coupling matrix is then can be expressed as

$$H_x e^{ika} + H_{-x} e^{-ika} = t_0 * 2cos(ka) \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{2.23}$$

Now we can solve the eigenvalue and eigenvector of the Hamiltonian.

$$H(k)\phi_i = E(k)_i \phi_i \tag{2.24}$$

We have the eigenvalue to be

$$E(k) = \begin{bmatrix} 4-\sqrt{3} \\ 3 \\ 4 \\ 5 \\ 4+\sqrt{3} \end{bmatrix} - 2 * cos(ka) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{2.25}$$

As showed in Fig. 2.3. The eigenvalue in terms of k is actually the band structure. The norm of the eigenvector is the spatially resolved wave function density. The

doi:10.6342/NTU201603562

behavior of wave function density is like the wave function density of a particle in a one-dimensional hard box due to the confinement in the y-direction. The wave function density of the lowest band, indicated by $n = 1$, is the ground state and the first excited state is the band indicated by $n = 2$ as showed in Fig. 2.3.

Now let us take 2DTI as an example. The tight binding Hamiltonian of 2DTI can be expressed by a four-band model. The Hamiltonian of an HgTe/CdTe quantum well yielding a 2DTI system can be given as follows:

$$\mathbf{H_{TI}} = \begin{bmatrix} \mathbf{h} & 0 \\ 0 & \mathbf{h}^* \end{bmatrix} \tag{2.26}$$

There are four base vectors of this Hamiltonian, where $|s,\uparrow\rangle$ and $|p_x + ip_y,\uparrow\rangle$ are the pseudo-spin-up states for the upper 2×2 block, and $|s,\downarrow\rangle$ and $|-p_x + ip_y,\downarrow\rangle$ are the pseudo-spin-down states for the lower block. The tight-binding Hamiltonian of $\mathbf{h}$ [2, 3]in real space can be represented by

$$\begin{aligned} \mathbf{h} &= \sum_{\mathbf{i}} \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} E_{is} & 0 \\ 0 & E_{ip} \end{bmatrix} \varphi_{\mathbf{i}} \\ &+ \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{ss} & V_{sp} \\ V_{sp}^* & V_{pp} \end{bmatrix} \varphi_{\mathbf{i}+\delta\mathbf{x}} + \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{ss} & iV_{sp} \\ -iV_{sp}^* & V_{pp} \end{bmatrix} \varphi_{\mathbf{i}+\delta\mathbf{y}} \end{aligned}$$

Where

$$\varphi = \left[c_{s,\mathbf{i}}^{\uparrow}, c_{p,\mathbf{i}}^{\uparrow}\right]^T \tag{2.27}$$

are spinors for the spin up components of the s and p orbits, the index $\mathbf{i}$ represents position in real space, and δx and δy are unit vectors of the lattice constant along the x and y directions.

$$H_0 = \sum_{\mathbf{i}} \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} E_{is} & 0 \\ 0 & E_{ip} \end{bmatrix} \varphi_{\mathbf{i}} + \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{ss} & iV_{sp} \\ -iV_{sp}^* & V_{pp} \end{bmatrix} \varphi_{\mathbf{i}+\delta\mathbf{y}} \tag{2.28}$$

doi:10.6342/NTU201603562

$$H_x = \sum_{\mathbf{i}} \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{ss} & V_{sp} \\ V_{sp}^{*} & V_{pp} \end{bmatrix} \varphi_{\mathbf{i}+\delta\mathbf{x}} \tag{2.29}$$

$$H = H_0 + H_x e^{ik\delta x} + H_{-x} e^{-ik\delta x} \tag{2.30}$$

$$H\phi_i = E_i \phi_i \tag{2.31}$$

The wave function density is the norm of the wave function and the summation of the wave function density at each energy point is 1 (normalised).

$$\rho = \phi_i^{\dagger} \phi_i \tag{2.32}$$

As showed in Fig. 2.4, the energy desperation is linear near energy equal to zero, and the wave function density shows the edge state accumulates at one edge. If we focus on the negative k region, the wave function density will show the edge state accumulates at the other edge due to time reversal symmetry. If the energy is going away from the Fermi energy, the wave function density shows a bulk state (orange energy point) and an edge state penetrate to the bulk (green energy point). We will discuss the detail of the band properties in the next chapter.

2.3 Non-periodic structure (Green's function)

Mesoscopic physics is an important branch of condensed matter physics that handles the system of an intermediate length that ranged in between micrometer and the size of an atom. For the physical system with the size of an atom, the direct solution of n-body Schrödinger equation has to adopt and all of the interaction have to take into account. However, for the physical system with the size of micrometers, a semi-classical approach is adopted and a lot of quantum interactions

doi:10.6342/NTU201603562

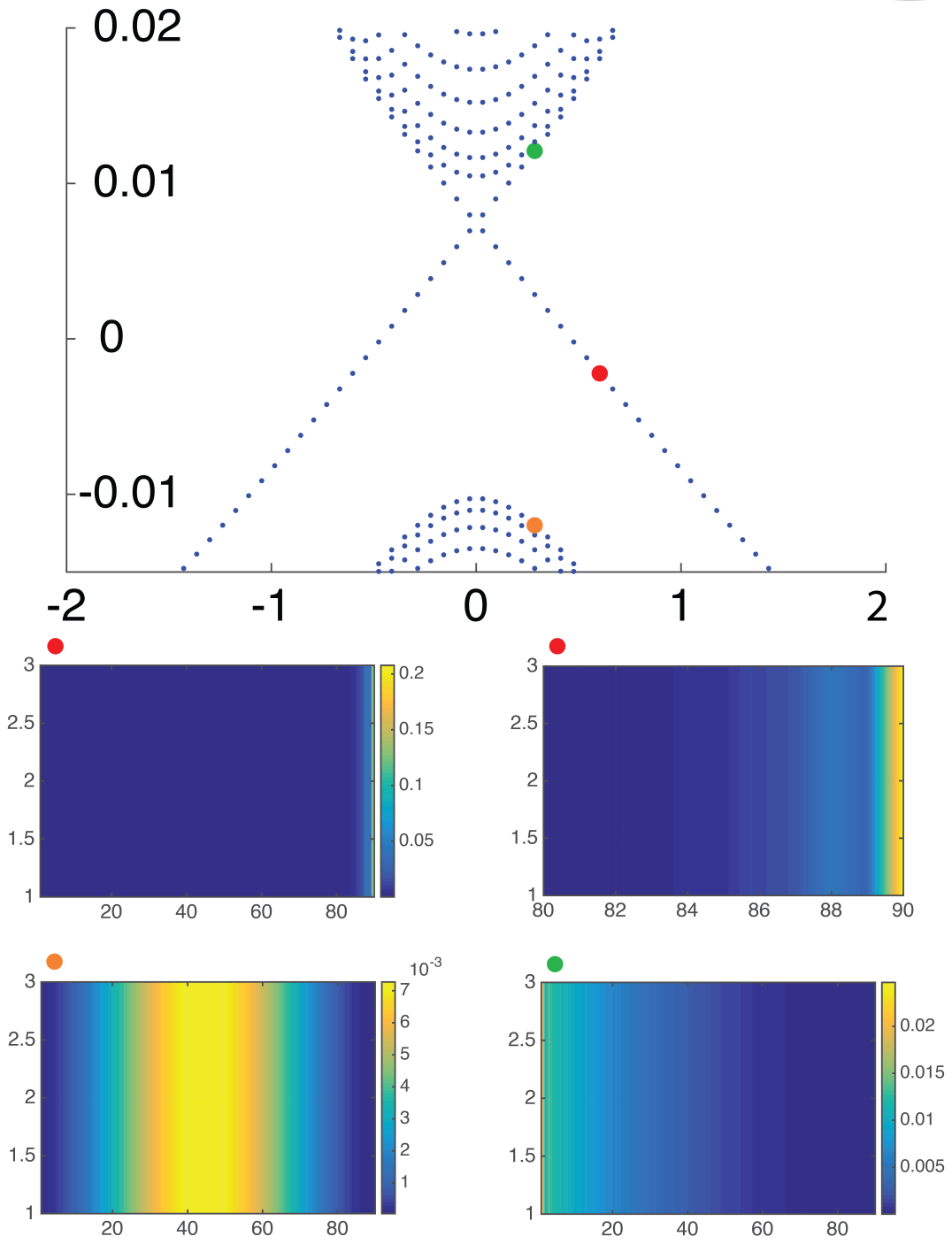

Figure 2.4: The band structure of a stripe of 2DTI near Fermi energy (set to be zero). The wave functions at three point of the eigenenergies.

doi:10.6342/NTU201603562

have to ignore. What if we would like to study the physical system composed of hundreds or thousands of atoms? Both methods are all cannot be adopted due to the complicated computing for n-body Schrödinger equation and ignoring too much information for the semi-classical approach. Thus a tight-binding approach starting from Schrödinger equation and ignoring some complicated interaction is applicable in the mesoscopic system. Interference effects, quantum confinement effects, and charging effects are the three kinds of phenomena in a device of mesoscopic physics.

If we would like to study the basic martial behavior that is periodical in general, the Bloch theory is the method as we have introduced. However, a mesoscopic device is not periodical in general. The device contains leads that play a role of source and drain of an electron and a transport part that connect the leads. The Green's function method with tight-binding model thus takes advantage in the kind of problem. We will present the Green's function and how does the Green's function work in a non-periodical structure. The physical quantities can be expressed in terms of the retarded Green's function G^R. And lesser Green's function $G^<$.

$$\mathbf{G}^r(E) = [(E + i\eta)\,\mathbf{I} - \mathbf{H}_{effect}(E)]^{-1} \tag{2.33}$$

$$G^<_{\mathbf{m},\mathbf{m}'}(E) = G^R(E)\Sigma^<(E)G^{R\dagger}(E) \tag{2.34}$$

The retarded Green's function G^R give us the physical information in equilibrium regime. Thus the local density of state and transmission can be expressed as the retarded Green's function. The spatially resolved spectral function $A(r, r', E)$ can be expressed by retarded Green's function $\mathbf{G}^r(E)$.

$$A(r, r', E) = \frac{-1}{\pi} Im[\mathbf{G}^r(E)] \tag{2.35}$$

The spatially resolved density of states $\mathbf{D}(r, E) = A(r, r, E)$ can be obtained from

doi:10.6342/NTU201603562

the diagonal element of the spectral function. The conductance is obtained by multiplying the conductance unit $\frac{e^2}{h}$ by

$$T_{i,j}(E) = trace[\mathbf{\Gamma}_i(E)\mathbf{G}^{\mathbf{r}}(E)\mathbf{\Gamma}_j(E)\mathbf{G}^{\mathbf{r}\dagger}(E)] \tag{2.36}$$

Where Γ_i is the broadening matrix of lead i that we will defined later. The lesser Green's function $G^<$ contains the physical information of the investigated system in time independent nonequilibrium regime, thus the local spin density and charge density with a small bias (i.e. in linear response regime) can be derived by the lesser Green's function. The local charge density $N_{\mathbf{m}}$ and local spin density $\overrightarrow{S}_{\mathbf{m}}$ can be expressed as

$$\left\langle \overrightarrow{S}_{\mathbf{m}} \right\rangle = \frac{\hbar}{4\pi i}\int_{-\infty}^{\infty} dETr_s[\overrightarrow{\sigma} G^<_{\mathbf{m},\mathbf{m}}(E)] \tag{2.37}$$

$$\langle N_{\mathbf{m}} \rangle = \frac{e}{2\pi i}\int_{-\infty}^{\infty} dETr_s[G^<_{\mathbf{m},\mathbf{m}}(E)] \tag{2.38}$$

We will illustrate how to determine the lesser and retarded Green's function step by step.

2.3.1 Green's function

Green's function is a way to solve differential equations of a linear differential operator $L = L(x)$.

$$Lu(x) = f(x). \tag{2.39}$$

The definition of a Green's function $G(x, s)$ of a linear differential operator $L = L(x)$ acting on a subset of the Euclidean space $\Re n$, at a point s, can be expressed by

$$LG(x, s) = \delta(s - x) \tag{2.40}$$

Where δ is the Dirac delta function. The definition of Green's function give us

$$\int LG(x, s)f(s)\,ds = \int \delta(x - s)f(s)\,ds = f(x). \tag{2.41}$$

doi:10.6342/NTU201603562

$$Lu(x) = \int LG(x,s)f(s)\,ds. \tag{2.42}$$

The operator $L = L(x)$ is linear and acts only on the variable x, thus the operator L can be taken outside of the integration.

$$Lu(x) = L\left(\int G(x,s)f(s)\,ds\right), \tag{2.43}$$

So the solution $u(x)$ of the linear differential equations can be expressed by the integration of the Green's function multiply the source function $f(s)$

$$u(x) = \int G(x,s)f(s)\,ds. \tag{2.44}$$

As a side note, the Green's function as used in physics is usually defined with the opposite sign, instead, that is,

$$LG(x,s) = -\delta(x-s). \tag{2.45}$$

If the operator is translation invariant, then the Green's function can be taken to be a convolution operator

$$G(x,s) = G(x-s). \tag{2.46}$$

In this case, the Green's function is the same as the impulse response of linear time-invariant system theory.

In electromagnetic physics, the Green's function can construct the solution for an arbitrary charge distribution if the linear operator is electronic Gauss's law (i.e. the electric field going out of an arbitrarily spatial space is proportional to the charge inside.)

$$\nabla^2\varphi = \rho/\varepsilon_0. \tag{2.47}$$

doi:10.6342/NTU201603562

Green's functions are also powerful in solving wave equations in quantum mechanics. The Green's function of the Hamiltonian, a linear operator as what we illustrated, is an important connection to the concept of density of states.

2.3.2 Green's function in quantum system

Green's function can be applied in quantum transport system[1]. In this section, we derive the Green's function of a quantum system that can describe the transport in a nano-scale device. I will start from a Schrödinger equation describing the entire system. The entire system contains a device and leads connecting to the device. We use a device connected of a lead as an example so the time independent Schrödinger equation is

$$\begin{bmatrix} H & H_{couple} \\ H_{couple}^{\dagger} & H_L \end{bmatrix} \begin{Bmatrix} \boldsymbol{\Psi} \\ \boldsymbol{\Phi_L} \end{Bmatrix} = E \begin{Bmatrix} \boldsymbol{\Psi} \\ \boldsymbol{\Phi_L} \end{Bmatrix} \tag{2.48}$$

where H and H_L are the Hamiltonian describing the device and the lead respectively. $H_{couple}^{\dagger}$ is the matrix that describes the coupling between the lead and the device. $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi_L}$ are the wave function of the device and the lead respectively, and E is the energy.

The Hamiltonian is a Hermitian operator so the energy is real. By solving the eigenfunction problem, all of the physical properties can be derived. However, there is a difficulty to solve the entire system. The Hamiltonian describing leads are semi-infinite in general, so it is almost impossible to solve the entire system. Hopefully, we are interested in the physical properties of the device only, so we can only calculate the wave function $\boldsymbol{\Psi}$ of the device. We then introduce a concept to simplify the infinite matrix calculation.

We would like to show that the equation can be simplified by eliminating H_L (i.e. the Hamiltonian describing the lead) and the equation describing the device

doi:10.6342/NTU201603562

are obtained by adding a self-energy paying for the eliminating of H_L.

$$E\mathbf{\Psi} = [H + \Sigma]\mathbf{\Psi} + \boldsymbol{S} \tag{2.49}$$

$\boldsymbol{S}$ is a term representing the excitation of the device by the coupling from the lead. Σ could be viewed as a modification of the Hamiltonian that represents the effective field that originates from the lead and act on the device. This equation is describing an open system, instead of a closed system that described by $E\mathbf{\Psi} = H\mathbf{\Psi}$.

We start the derivation form solving the Schrödinger equation of the isolated lead.

$$E\mathbf{\Phi_L} = H_L\mathbf{\Phi_L} \tag{2.50}$$

where $\mathbf{\Phi_L}$ is the wave function of the electrons in the lead.

We now make a bit of revising of the Schrödinger equation by adding a small amount of imaginary number $i\eta$ into energy. So the revised Schrödinger equation can be expressed as

$$(E + i\eta)\mathbf{\Phi_L} = H_L\mathbf{\Phi_L} + \boldsymbol{S_L} \tag{2.51}$$

$i\eta\mathbf{\Phi_L}$ represents the extraction of electrons from the contact and $\boldsymbol{S_L}$ represents the reinjection of electrons from external sources. These two terms are essential that can be viewed as maintaining contact at a constant electrochemical potential.

$i\eta$ is mathematical artifice to make the convergence of a Fourier transform such that ensure numerical calculation convergent. The small imaginary number cannot be dropped in calculating the Green's function numerically due to the convergence of Green's function. As we will discuss later, the surface Green's function can be expressed as

$$G_L \equiv [E - H_L + i\eta]^{-1} \tag{2.52}$$

If we drop $i\eta$, the surface Green's function G_L diverge when the energy is the eigenvalue of the H_L. The revising of Schrödinger equation is actual can be interpreted

doi:10.6342/NTU201603562

in a point of view: The replacement of energy H_L to $H_L - i\eta$ broaden the wave function of lead $\boldsymbol{\Phi_L}$ in energy space. $\boldsymbol{\Phi_L}$ are essentially the eigenfunctions of H_L as the absence of $i\eta$, so $\boldsymbol{\Phi_L}$ are non-zero only as E is the eigenenergies of H_L. As a result, $\boldsymbol{\Phi_L}$ is extremely peak in energy space. Adding $i\eta$, instead,$\boldsymbol{\Phi_L}$ is non-zero when energy "close" to all eigenenergies of H_L, whose sharpness depends on the η. Small η give a sharp $\boldsymbol{\Phi_L}$ in energy space. In numerical calculations, we always used finite-sized leads to saving computation time, so the energy is not continuous, instead, discretized. η have to larger than the level spacing in order to make the lead work well in the calculation. Determining the value of η is actually an ambiguity in our numerical calculation. The Hamiltonian change too much if η is too large, however, we cannot make η too small such the level spacing is too small. In our numerical calculation, we usually set $\eta = 10^{-7}$ in the unit of eV. With this setting, the result is correctly comparing to some analytical result.

Now we couple the device to the lead, the wave function of the lead is not $\boldsymbol{\Phi_L}$, the wave function of isolated lead, so we can define an excite scattered waves χ such that $\boldsymbol{\Phi_L} + \boldsymbol{\chi}$ to be the wave function of the lead in the overall system. $\boldsymbol{\Phi_L} + \boldsymbol{\chi}$ and $\boldsymbol{\Psi}$ have to satisfy to overall revising Schrödinger equation, which we can write in two blocks:

$$\begin{bmatrix} E - H & -H_{couple} \\ -H_{couple}^{\dagger} & E - H_L + i\eta \end{bmatrix} \begin{Bmatrix} \boldsymbol{\Psi} \\ \boldsymbol{\Phi_L} + \boldsymbol{\chi} \end{Bmatrix} = \begin{Bmatrix} 0 \\ \boldsymbol{S_L} \end{Bmatrix} \tag{2.53}$$

Then we expand the matrix and substitute $\boldsymbol{S_L}$ by $[E - H_L + i\eta]\boldsymbol{\Phi_L}$ by assuming $\boldsymbol{S_L}$ remain unchanged and we can have

$$[E - H]\boldsymbol{\Psi} - H_{couple}[\boldsymbol{\Phi_L} + \boldsymbol{\chi}] = 0 \tag{2.54}$$

$$-H_{couple}^{\dagger}\boldsymbol{\Psi} + [E - H_L + i\eta]\boldsymbol{\chi} = 0 \tag{2.55}$$

Now we would like to introduce the surface Green's function of the lead G_L defined

doi:10.6342/NTU201603562

by

$$G_L \equiv [E - H_L + i\eta]^{-1} \tag{2.56}$$

We then can express $\boldsymbol{\chi}$ in in terms of $\boldsymbol{\Psi}$ and the surface Green's function G_L by

$$\boldsymbol{\chi} = G_L H^{\dagger}_{couple} \boldsymbol{\Psi} \tag{2.57}$$

By substituting $\boldsymbol{\chi}$ into (2.54), it is straightforward to defined the self energy Σ

$$\Sigma \equiv H_{couple} G_L H^{\dagger}_{couple} \tag{2.58}$$

and (2.54) become

$$[E - H - \Sigma]\boldsymbol{\Psi} = \boldsymbol{S} \tag{2.59}$$

where $\boldsymbol{S}$ is defined by

$$\boldsymbol{S} \equiv H_{couple} \boldsymbol{\Phi_L} \tag{2.60}$$

Now we have derived the identity and we have the wave function of the device without solving the overall Schrödinger equation, instead, adding an effective field named self energy to the Hamiltonian of the device. We can regard the self energy as a boundary condition of the device that provides a reservoir of electrons. Next, we are going to introduce Green's function and show the physical meaning of these Green's function.

We are going with the express some physical quantity in terms of wave functions, the density of state and the electron density. The density of states of a system is the number of states per interval of energy at each energy level that is available to be occupied. The density of state $D(E)$ can then be defined as

$$D(E) \equiv \sum_i \boldsymbol{\Psi_i} \boldsymbol{\Psi_i^*} \delta(E - \epsilon_i) \tag{2.61}$$

where ϵ_i is a set of eigenenergies of a Hamiltonian and $\boldsymbol{\Psi_i}$ are the eigenfunctions in real space correspond to the eigenenergies ϵ_i. The electron density is the density

doi:10.6342/NTU201603562

of electron occupation in terms of energy. The electron occupies in energy E in a system according to its Fermi-Dirac distribution $F(\epsilon_i - \mu) = 1/[e^{(\epsilon_i-\mu)/kT} - 1]$, so the electron density ρ of the device can be expanded by the eigenfunction and Fermi-Dirac distribution $F(\epsilon_i - \mu)$ in the function of the eigenenergy corresponding the eigenfunction. The relation of the eigenfunction and the electron density can be thus expressed as

$$\rho(r) \equiv \sum_i \mathbf{\Psi_i} F(\epsilon_i - \mu) \mathbf{\Psi_i^*} \tag{2.62}$$

where $F(\epsilon_i - \mu)$ is the Fermi-Dirac distribution function, $\mathbf{\Psi_i}$ is the eigenfunction and ϵ_i is the corresponded eigenenergy, μ is the chemical energy, sometime we also refer it as the Fermi energy. k is the Boltzmann constant, and T is the temperature. In our calculation, we simplify the problem by setting the temperature to be zero so that the Fermi-Dirac distribution function is a step function.

Next, we are going to introduce two types of Green's functions, retarded Green's function and lesser Green's function, and relate the Green's functions with DOS and electron density.

2.3.3 Retarded Green's function

We will illustrate the physics meaning of retarded Green's function in the section. Now let us start from the density of state (DOS) $D(E)$ (2.61) of the overall system.

$$D(E) \equiv \sum_i |\overline{\mathbf{\Psi_i}}|^2 \delta(E - \epsilon_i)$$

where ϵ_i is a set of eigenenergies of the overall Hamiltonian as described eq(2.48) and $\overline{\mathbf{\Psi_i}}$ are the eigenenfunctions in real space correspond to the eigenenergies ϵ_i.

Due to the DOS can be expressed in term of eigenenergies and the correspondence eigenfunctions. We then define overall spatially resolved spectral function

doi:10.6342/NTU201603562

$A_{overall}(r, r', E)$

$$A_{overall}(r, r', E) \equiv 2\pi \sum_i \overline{\mathbf{\Psi}_i}(r)\overline{\mathbf{\Psi}_i^*}(r')\delta(E - \epsilon_i) \tag{2.63}$$

Thus spatially resolved DOS is the diagonal of the spectral function $D(r, E) = A(r, r, E)/2\pi$. The Dirac delta function in (2.63) can be expanded

$$2\pi\delta(E - \epsilon_i) = \left[\frac{2\eta}{(E - \epsilon_i)^2 + \eta^2}\right] = i\left[\frac{1}{(E - \epsilon_i) + i\eta} - \frac{1}{(E - \epsilon_i) - i\eta}\right] \tag{2.64}$$

Where η is a small number as what we have discussed. Spectral function $A_{overall}(r, r', E)$ then can be expanded

$$\begin{aligned} A_{overall}(E) &= 2\pi \sum_n \overline{\mathbf{\Psi}_n}\overline{\mathbf{\Psi}_n^*}\delta(E - \epsilon_n) \\ &= i\sum_n \overline{\mathbf{\Psi}_n\mathbf{\Psi}_n^*}\left[\frac{1}{(E - \epsilon_n) + i\eta} - \frac{1}{(E - \epsilon_n) - i\eta}\right] \\ &= i\sum_n \overline{\mathbf{\Psi}_n}\left[\frac{1}{(E - \epsilon_n) + i\eta}\right]\overline{\mathbf{\Psi}_n^*} \\ &- i\sum_n \overline{\mathbf{\Psi}_n}\left[\frac{1}{(E - \epsilon_n) - i\eta}\right]\overline{\mathbf{\Psi}_n^*} \\ &= i\frac{1}{E - H + i\eta} - i\frac{1}{E - H - i\eta} \end{aligned} \tag{2.65}$$

Define the retarded Green's function $G_{overall}$

$$G_{overall} \equiv \frac{1}{E - H + i\eta} \tag{2.66}$$

We then can express the spectral function $A_{overall}(E)$ in terms of retarded Green's function $G_{overall}$ by

$$A_{overall}(E) = i[G_{overall}(E) - G^{\dagger}_{overall}(E)] \tag{2.67}$$

$A_{overall}(E) = A_{overall}(r, r'E)$ can be expressed with the position representation.

Note the spectral function $A(E)$ describing the device thus can be expressed as

$$A(E) = i[G(E) - G^{\dagger}(E)] \tag{2.68}$$

doi:10.6342/NTU201603562

and $A_L(E)$ describing the lead can be expressed as

$$A_L(E) = i[G_L(E) - G_L^\dagger(E)] \tag{2.69}$$

Now we would like to focus on the self-energy matrix again. We have shown that the self-energy can be expressed as surface Green's function G_L and coupling matrix H_{couple} in eq(2.58). However, the relation of Green's function G of the device and self-energy matrix Σ have not been known yet. If the device is attached to leads, and we do not interest in the local physical properties of the lead then we can use self-energy to simplify the calculation. Now we consider a device that connects with a lead. The overall Hamiltonian can be expressed by

$$\begin{bmatrix} H & H_{couple} \\ H_{couple}^\dagger & H_L \end{bmatrix}$$

where H and H_L are the Hamiltonian describing the device and the lead respectively just like eq(2.48). So the overall retarded Green's function $G_{overall}$ can be easily obtained by the definition of related Green's function eq(2.66)

$$G_{overall} = \begin{bmatrix} G & G_{12} \\ G_{21} & G_L \end{bmatrix} = \begin{bmatrix} E + i\eta - H & -H_{couple} \\ -H_{couple}^\dagger & E + i\eta - H_L \end{bmatrix}^{-1} \tag{2.70}$$

By expanding eq(4.4), the following equation can be easily obtained.

$$[E + i\eta - H]G - H_{couple}G_{21} = 1 \tag{2.71}$$

$$[E + i\eta - H_L]G_{21} - H_{couple}^\dagger G = 0 \tag{2.72}$$

eq(2.72) and the definition of surface Green's function (2.52), G_{21} can be expressed as

$$G_{21} = G_L H_{couple}^\dagger G \tag{2.73}$$

doi:10.6342/NTU201603562

Substituting G_{21} into eq(2.71), we then have the following equation

$$[E + i\eta - H - H_{couple} G_L H_{couple}^{\dagger}]G = 1 \tag{2.74}$$

We can use the definition of self energy eq(2.58), then we have

$$[E + i\eta - H - \Sigma]G = 1 \tag{2.75}$$

Now we can express regarded Green's function of the device by Hamiltonian of the device and self energy by

$$G = [E + i\eta - H - \Sigma]^{-1} \tag{2.76}$$

We always drop $i\eta$ into Σ due to Σ is not a Hermitian matrix so the regarded Green's function of the device is

$$G = [E - H - \Sigma]^{-1} \tag{2.77}$$

From eq(2.59) $[E - H - \Sigma]\boldsymbol{\Psi} = \boldsymbol{S}$ and eq(2.60)$\boldsymbol{S} \equiv H_{couple}\boldsymbol{\Phi_L}$, we can derive an identity that we will use it to derive lessor Green's function and transmission later.

$$\boldsymbol{S} = [E - H - \Sigma]\boldsymbol{\Psi} = G^{-1}\boldsymbol{\Psi} = H_{couple}\boldsymbol{\Phi_L} \tag{2.78}$$

Now we know the power of self energy, we do not need to inverse the overall matrix, instead, only adding a self energy matrix to the Hamiltonian of the device without enlarging the dimension of the Hamiltonian. It is easy to extend the device of one lead to multi-leads leads system. The retarded Green's function is

$$G = \left[E - H - \sum_i \Sigma_i\right]^{-1} \tag{2.79}$$

Where Σ_i is the self energy of i^{th} lead. In the next section, we are going to calculate the self-energy by numerical iteration.

doi:10.6342/NTU201603562

2.3.4 Evaluating the self-energy

There are two ways, analytical method, and numerical method, to evaluate the self-energy. In our calculation, we adopt the numerical method to evaluate the self-energy. The self energy $\Sigma = H_{couple} g_L H_{couple}^{\dagger}$ is determined by the surface Green's function g_L. Now we are going to illustrate how to determine g_L in a numerical way. Generally, the lead is a periodical system, H_0 and H_{01} are the Hamiltonian of a unit cell and the hopping matrix between unit cells as we described in Fig. 2.3 respectively. The Hamiltonian of the lead thus can be expressed as

$$\begin{bmatrix} H_0 & H_{01} & 0 & \dots \\ H_{01}^{\dagger} & H_0 & H_{01} & \dots \\ 0 & H_{01}^{\dagger} & H_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The surface Green's function of n unit cell g_n is then can be constructed iteratively

$$g_2 = [E - H_0 - H_{01}^{\dagger} H_0 H_{01} + i\eta]^{-1} \tag{2.80}$$

$$g_3 = [E - H_0 - H_{01}^{\dagger} g_2 H_{01} + i\eta]^{-1} \tag{2.81}$$

$$g_4 = [E - H_0 - H_{01}^{\dagger} g_3 H_{01} + i\eta]^{-1} \tag{2.82}$$

$$\vdots \tag{2.83}$$

Once the surface Green's function $g_n \sim g_{n+1}$, the surface Green's function do not vary when adding additional unit cell, that means the iteration converges. We then can determine the surface Green's function g_n as a Green's function of semi-infinite lead and $g_n = G_L$.

doi:10.6342/NTU201603562

2.3.5 Lessor Green's function

After we know how to calculate the DOS of a device by using the retarded Green's function, it is also important to express the electron density in terms of Green's function. The lesser Green's function G^n give us the electron density in equilibrium and non-equilibrium regime. We can use it to calculate local spatial electron density, local spatial current density, and transmission (conductance) between leads.

Now we consider the electron density of the lead ρ_L and of the device ρ, just as what we have defined (2.62) . The electron density of a lead is $\rho_L = \mathbf{\Phi_L}\mathbf{\Phi_L^\dagger}$ where $\mathbf{\Phi_{Li}}$ is the eigenfunction of the Hamiltonian of the lead

$$\begin{aligned}
\rho_L &= \sum_i \mathbf{\Phi_{Li}} F(\epsilon_i - \mu)\mathbf{\Phi_{Li}^*} \\
&= \int dE * F(E-\mu) \sum_i \mathbf{\Phi_{Li}}\delta(\epsilon_i - E)\mathbf{\Phi_{Li}^*} \\
&= \int dE * F(E-\mu) A_L/2\pi
\end{aligned} \tag{2.84}$$

We have used the definition of the spectral function just as eqref17, the definition of spectral function of the overall system. Thus, the spectral function of the lead can be expressed as

$$A_L = 2\pi \sum_i \mathbf{\Phi_{Li}}\delta(\epsilon_i - E)\mathbf{\Phi_{Li}^*} \tag{2.85}$$

Next we can derive the electron density of the device.

$$\begin{aligned}
\rho &= \sum_i \mathbf{\Psi_i} F(\epsilon_i - \mu)\mathbf{\Psi_i^*} \\
&= \int dE * F(E-\mu) \sum_i \mathbf{\Psi_i}\delta(\epsilon_i - E)\mathbf{\Psi_i^*} \\
&= \int dE * F(E-\mu) G H_{couple} [\sum_i \mathbf{\Phi_{Li}}\delta(\epsilon_i - E)\mathbf{\Phi_{Li}^*}] H_{couple}^\dagger G^\dagger \\
&= \int \frac{dE}{2\pi} * F(E-\mu) G H_{couple} A_L H_{couple}^\dagger G^\dagger
\end{aligned} \tag{2.86}$$

doi:10.6342/NTU201603562

We have used eq(2.78) $G\mathbf{\Psi} = H_{couple}\mathbf{\Phi_L}$ to derive $G\mathbf{\Psi_i} = H_{couple}\mathbf{\Phi_{Li}}$. Now we can express the electron density ρ of the device in terms of the spectral function of the lead and the retarded Green's function of the device.

We can thus define G^n

$$G^n \equiv F(E-\mu)GH_{couple}A_LH_{couple}^{\dagger}G^{\dagger} \tag{2.87}$$

Such that electron density $\rho(r)$ then can be expressed as a integration of lesser Green's function

$$\rho(r) = \int dEG^n/2\pi \tag{2.88}$$

So far we can know the electron density and density of states by using the retarded Green's function and lesser Green's function.

Now we would like to introduce some identity that will be useful in deriving the transmission between leads in the next section. From eq(2.58) we know

$$\begin{aligned} \Sigma - \Sigma^{\dagger} &= H_{couple}G_LH_{couple}^{\dagger} - [H_{couple}G_LH_{couple}^{\dagger}]^{\dagger} \\ &= H_{couple}[G_L - G_L^{\dagger}]H_{couple}^{\dagger} \\ &= H_{couple}[A_L/i]H_{couple}^{\dagger} \end{aligned} \tag{2.89}$$

We can define gamma function Γ

$$\Gamma \equiv H_{couple}A_LH_{couple}^{\dagger} = i[\Sigma - \Sigma^{\dagger}] \tag{2.90}$$

So G_n is also expressed as

$$G_n = F(E-\mu)G\Gamma G^{\dagger} \tag{2.91}$$

It is easy to extend the device of one lead to multi-leads leads system. The rerated Green's function is

$$G = \left[E - H - \sum_i \Sigma_i\right] \tag{2.92}$$

doi:10.6342/NTU201603562

and the lessor Green's function can be expressed as

$$G_n = \sum_i F(E - \mu_i) G \Gamma_i G^\dagger \tag{2.93}$$

where Σ_i is the the self energy of i^{th} lead, Γ_i is the gamma function of i^{th} lead, and μ_i is the chemical potential of i^{th} lead.

So far we have introduced a device connected with multi-leads that indeed is a non-equilibrium system. Next, we would like to derive a transport property, the transmission between leads, that is an important physical quantity discussed in the thesis.

2.3.6 Transmission

We have to obtain an expression for the current between each of the leads. First, let us consider the system we have discussed again. We want to derive the current flowing between the device and the lead. The current is actually the time rate of change in the electron density inside the device, so the current can be defined by the time derivative of electron density $\rho = \mathbf{\Psi}\mathbf{\Psi}^\dagger = \sum_i \mathbf{\Psi_i} F(\epsilon_i - \mu)\mathbf{\Psi_i^*}$.

$$I \equiv \frac{d}{dt}\mathbf{\Psi}\mathbf{\Psi}^\dagger = \frac{d}{dt} trace([\mathbf{\Psi}\mathbf{\Psi}^\dagger) = \frac{d}{dt} trace([\mathbf{\Psi}^\dagger\mathbf{\Psi}) \tag{2.94}$$

due to $\mathbf{\Psi}\mathbf{\Psi}^\dagger = trace([\mathbf{\Psi}\mathbf{\Psi}^\dagger)$, its a scale, and the multiplication of matrix is commute if we take a trace ($Trace(AB) = Trace(BA)$).

The derivation of wave function is determined by the time dependent Schrödinger equation

$$\begin{bmatrix} H & H_{couple} \\ H_{couple}^\dagger & H_L - i\eta \end{bmatrix} \begin{Bmatrix} \mathbf{\Psi} \\ \mathbf{\Phi} \end{Bmatrix} = i\hbar \frac{d}{dt} \begin{Bmatrix} \mathbf{\Psi} \\ \mathbf{\Phi} \end{Bmatrix} \tag{2.95}$$

So the time derivation of the wave function in the device can be expressed as

$$i\hbar \frac{d}{dt}\mathbf{\Psi} = H\mathbf{\Psi} + H_{couple}\mathbf{\Phi} \tag{2.96}$$

doi:10.6342/NTU201603562

Substituting (2.96) into (2.94), we have the expression of current

$$I = \frac{Trace\left[\boldsymbol{\Psi}^{\dagger} H_{couple} \boldsymbol{\Phi} - \boldsymbol{\Phi}^{\dagger} H_{couple}^{\dagger} \boldsymbol{\Psi}\right]}{i\hbar} \tag{2.97}$$

The $\boldsymbol{\Phi}$ is actually $\boldsymbol{\Phi_L} + \boldsymbol{\chi}$ as we described in (2.53). Replacing $\boldsymbol{\Phi}$ by $\boldsymbol{\Phi_L} + \boldsymbol{\chi}$ we have

$$I = \frac{Trace\left[\boldsymbol{\Psi}^{\dagger} H_{couple} [\boldsymbol{\Phi_L} + \boldsymbol{\chi}] - [\boldsymbol{\Phi_L} + \boldsymbol{\chi}]^{\dagger} H_{couple}^{\dagger} \boldsymbol{\Psi}\right]}{i\hbar} \tag{2.98}$$

We can divide the current into two parts

$$I = \frac{Trace\left[\boldsymbol{\Psi}^{\dagger} H_{couple} \boldsymbol{\Phi_L} - \boldsymbol{\Phi_L}^{\dagger} H_{couple}^{\dagger} \boldsymbol{\Psi}\right]}{i\hbar} - \frac{Trace\left[\boldsymbol{\chi}^{\dagger} H_{couple}^{\dagger} \boldsymbol{\Psi} - \boldsymbol{\Psi}^{\dagger} H_{couple} \boldsymbol{\chi}\right]}{i\hbar} \tag{2.99}$$

The first term is represented inflow current I_{inflow} due to its proportional to the incident wave $\boldsymbol{\Phi_L}$, and the second term is proportional to scattered wave $\boldsymbol{\chi}$ that represent the outflow current $I_{outflow}$.

There is some equation used to express the inflow current. $\boldsymbol{\Psi} = G\boldsymbol{S}$ and $\boldsymbol{S} = H_{couple}\boldsymbol{\Phi_L}$ in (2.78). The charge density of the lead $\rho_L = \boldsymbol{\Phi_L}\boldsymbol{\Phi_L}^{\dagger} = \int dE * F(E - \mu)A_L/2\pi$ in (2.84). Spectral function of the device $A(E) = i[G(E) - G^{\dagger}(E)]$ in (2.68). Gamma function $\Gamma = i[\Sigma - \Sigma^{\dagger}] = H_{couple}A_L H_{couple}^{\dagger}$ in (2.90). With the above identity, we can derive the inflow current to be a simpler form as following

doi:10.6342/NTU201603562

$$
\begin{aligned}
I_{inflow} &= \frac{Trace\left[\boldsymbol{\Psi}^{\dagger} H_{couple} \boldsymbol{\Phi_L} - \boldsymbol{\Phi}_{\boldsymbol{L}}^{\dagger} H_{couple}^{\dagger} \boldsymbol{\Psi}\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{S}^{\dagger} G^{\dagger} \boldsymbol{S} - \boldsymbol{S}^{\dagger} G \boldsymbol{S}\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{S}^{\dagger} \boldsymbol{S} G^{\dagger} - \boldsymbol{S}^{\dagger} \boldsymbol{S} G\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{S}^{\dagger} \boldsymbol{S} (G^{\dagger} - G)\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{S}^{\dagger} \boldsymbol{S} A\right]}{\hbar} \\
&= \frac{Trace\left[\boldsymbol{\Phi}_{\boldsymbol{L}}^{\dagger} H_{couple}^{\dagger} H_{couple} \boldsymbol{\Phi_L} A\right]}{\hbar} \\
&= \frac{Trace\left[H_{couple} \boldsymbol{\Phi_L} \boldsymbol{\Phi}_{\boldsymbol{L}}^{\dagger} H_{couple}^{\dagger} A\right]}{\hbar} \\
&= \frac{Trace\left[H_{couple} [\int dE * F(E-\mu) A_L / 2\pi] H_{couple}^{\dagger} A\right]}{\hbar} \\
&= \frac{Trace\left[\int dE * F(E-\mu) \Gamma A\right]}{2\pi\hbar}
\end{aligned}
\tag{2.100}
$$

There is some equation used to express the outflow current. $\boldsymbol{\chi} = G_L H_{couple}^{\dagger} \boldsymbol{\Psi}$ in (2.57). Spectral function of the lead $i[G_L^{\dagger} - G_L] = A_L$ in (2.69). The charge density of the device $\rho(r) = \int dE G^n / 2\pi$ in (2.88). To have the expression of the outflow current.

doi:10.6342/NTU201603562

$$
\begin{aligned}
I_{outflow} &= \frac{Trace\left[\boldsymbol{\chi}^{\dagger}H_{couple}^{\dagger}\boldsymbol{\Psi}-\boldsymbol{\Psi}^{\dagger}H_{couple}\boldsymbol{\chi}\right]}{i\hbar} \\
&= \frac{Trace\left[[G_{L}H_{couple}^{\dagger}\boldsymbol{\Psi}]^{\dagger}H_{couple}^{\dagger}\boldsymbol{\Psi}-\boldsymbol{\Psi}^{\dagger}H_{couple}G_{L}H_{couple}^{\dagger}\boldsymbol{\Psi}\right]}{i\hbar} \\
&= \frac{Trace\left[[\boldsymbol{\Psi}^{\dagger}H_{couple}G_{L}^{\dagger}]H_{couple}^{\dagger}\boldsymbol{\Psi}-\boldsymbol{\Psi}^{\dagger}H_{couple}G_{L}H_{couple}^{\dagger}\boldsymbol{\Psi}\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{\Psi}^{\dagger}[H_{couple}G_{L}^{\dagger}H_{couple}^{\dagger}-H_{couple}G_{L}H_{couple}^{\dagger}]\boldsymbol{\Psi}\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{\Psi}^{\dagger}[H_{couple}[G_{L}^{\dagger}-G_{L}]H_{couple}^{\dagger}]\boldsymbol{\Psi}\right]}{i\hbar} \\
&= \frac{Trace\left[\boldsymbol{\Psi}^{\dagger}[H_{couple}A_{L}H_{couple}^{\dagger}]\boldsymbol{\Psi}\right]}{\hbar} \\
&= \frac{Trace\left[\boldsymbol{\Psi}^{\dagger}\Gamma\boldsymbol{\Psi}\right]}{\hbar} \\
&= \frac{Trace\left[\boldsymbol{\Psi}^{\dagger}\boldsymbol{\Psi}\Gamma\right]}{\hbar} \\
&= \frac{Trace\left[\int dE*G^{n}\Gamma\right]}{2\pi\hbar}
\end{aligned}
\tag{2.101}
$$

It is easy to see that the inflow and outflow are equal so the net current is zero in this system due to $G_n = AF$. The current can be easily extended to a with

$$
I_i = \int dE * trace\left[\Gamma_i AF_i - \Gamma_i G^n\right] \tag{2.102}
$$

Now we can derive the transmission as a function of energy by the Landauer formula. Considering a muti leads system, we want to calculate current between two leads $I_{i,j}$ form i^{th}lead to j^{th}lead. Here we have to derive some identity to simplify $I_{i,j}$. First,

$$
G^{-1} - G^{\dagger -1} = \left[E - H - \sum_j \Sigma_j\right] - \left[E - H - \sum_j \Sigma_j^{\dagger}\right] = i\sum_j \Gamma_j \tag{2.103}
$$

doi:10.6342/NTU201603562

So we have a identity: $A = i(G - G^\dagger) = \sum_j G\Gamma_j G^\dagger$. Second, $G^n = \sum_j G\Gamma_j F_j G^\dagger$ as indicated in (2.93) for muti lead system.

$$\begin{aligned} I_i &= \int dE * trace\left[\Gamma_i\left[\sum_j G\Gamma_j G^\dagger\right]F_i - \Gamma_i\left[\sum_j G\Gamma_j F_j G^\dagger\right]\right] \\ &= \int dE * trace\left[\sum_j \Gamma_i G\Gamma_j G^\dagger F_i - \Gamma_i G\Gamma_j G^\dagger F_j\right] \\ &= \int dE * trace\left[\sum_j \Gamma_i G\Gamma_j G^\dagger (F_i - F_j)\right] \\ &\approx (\mu_i - \mu_j)\int dE * trace\left[\sum_j \Gamma_i G\Gamma_j G^\dagger\right] \\ &\approx (\delta\mu_{i,j})T_{i,j}(\bar{\mu}) \end{aligned} \tag{2.104}$$

where

$$T_{i,j}(\bar{\mu}) = \int_{\mu_j}^{\mu_i} dE * trace\left[\Gamma_i G\Gamma_j G^\dagger\right] \tag{2.105}$$

If all of the chemical potential of leads are equal, (i.e. $\delta\mu_{i,j} = 0$), then we have the zero current. With a bias difference between lead, we have the current obey the above equation. This is the Landauer-Büttiker formula that describes the current between lead. Landauer-Büttiker is a linear response formula so the formula work when bias difference between lead is small.

We give an example to illustrate the relation of transmission and the band structure. We consider a stipe of normal metal that electron can move freely in the normal metal. As we have introduced in the section of the periodical structure, the band structure can be calculated as Fig. 2.3. We can also calculate the transmission of the stipe. The relation is shown in Fig. 2.5, the transmission is quantized and increasing as energy 2reaches a band and decreasing as energy leaves a band.

doi:10.6342/NTU201603562

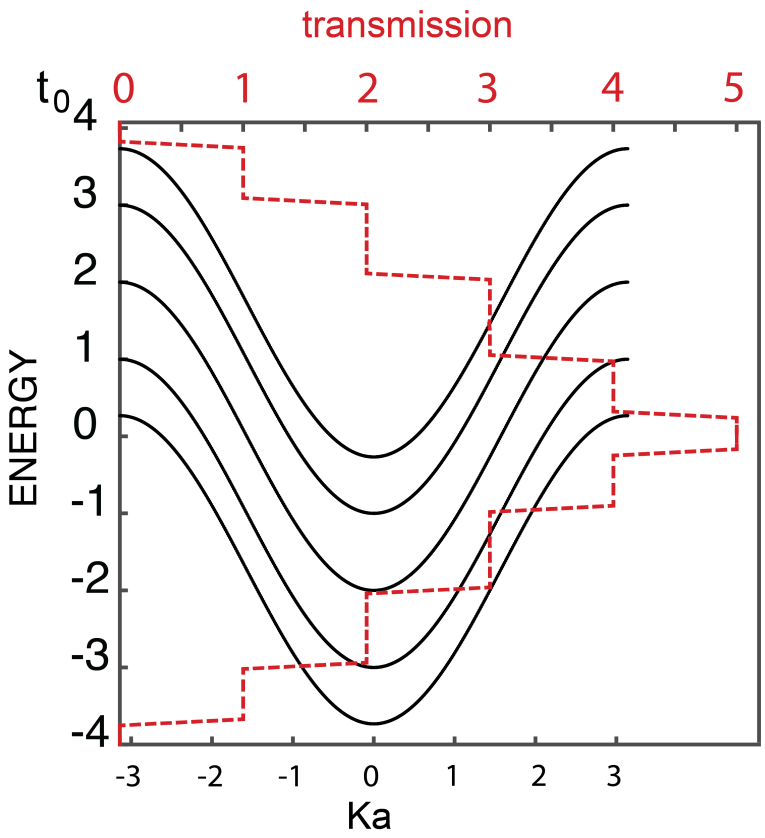

Figure 2.5: The transmission and band structure of a strip of normal metal, the energy unit is the hopping energy t_0.

doi:10.6342/NTU201603562

Bibliography

[1] S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, Cambridge, 2005).

[2] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[3] C. Xu and J.E. Moore, Phys. Rev. B **80**, 165316 (2009).

doi:10.6342/NTU201603562

3

Persistent quantum resonance transition in spin Hall transport

3.1 Overview

A quantum well composed of a CdTe/HgTe/CdTe sandwich structure was predicted to be 2DTI[1], and experiments demonstrated that a strong intrinsic spin-orbit interaction exists in this structure[2]. In this material, the spin direction and momentum are locked together in quantum spin Hall edge states, and the electrons of spin-up and -down propagate along the edge in opposite directions. By virtue of time reversal symmetry conservation, the edge states cannot be scattered by nonmagnetic defects[3, 4] and weak interactions[5, 6]. However, the quantized edge states cannot be conserved completely due to the finite size effect[7] based on the overlapping of wave functions at the two edges, which opens a band gap in energy dissipation near the Fermi energy and that makes the transmissions decrease to zero.

Fano resonances[8] and Breit-Wigner (B-W) resonances are quantum resonances,

doi:10.6342/NTU201603562

and can be realized in one-dimensional transport with a quantum dot[9]. Fano resonances, the coexistence of continuum of propagating modes, and discrete states interfere with each other and cause the transmission to be asymmetrical on line profile. However, the B-W resonances have no propagating modes, and electrons from a source cannot tunnel to a drain except through a quantum dot in a two-leads quantum dot system. Fano resonances cannot be found in a stripe of 2DTI with defect[10] for the reason that time reversal symmetry prevents the backward scattering of the edge states. Recently, an experimental result indicated that a 3DTI embedding impurity (replacing Bi with In) reduces the spin-orbit interaction and increases the overlap of surface state and bulk phonons. The 3DTI thus undergoes a phase transition to the non-TI state. The increase of the Drude background enhances the asymmetry character of the Fano resonance, and reaches its maximum near the phase transition point from TI to a non-TI phase[11]. However, the reported Fano resonance exists due to an overlap of bulk phonons and Dirac electrons and a quantum resonance within the Dirac region does not exist.

In this section, a local gated H-shaped 2DTI with four-terminals is proposed as a quantum resonance device as shown in Fig. 3.1. Local gating is introduced as a non-magnetic impurity which induces localized bound states at appropriate potential values in the 2DTI[12], and breaks the particle-hole symmetry. The H-shaped device provides the electrons within the helical edge states transverse through an impurity and crosswalk to other branches, and destroys the quantization of spin Hall conductance when the two edges reside sufficiently close so that Fano-like resonance can be realized in 2DTI. We propose a picture describing the interference of edge states and bound states. The transmission from leads show the Fano or the B-W resonances by modulating the gate potential in a distinct region of the width of the leads. Furthermore, this novel quantum resonance in Dirac electrons is persistent

doi:10.6342/NTU201603562

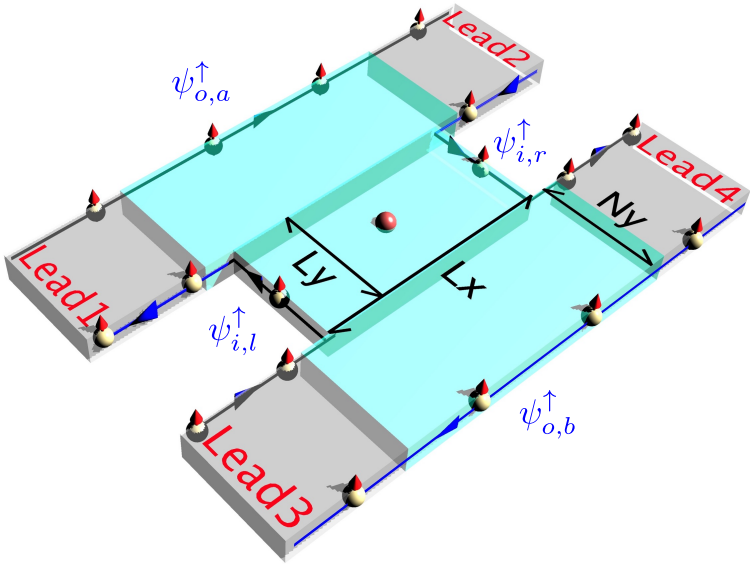

Figure 3.1: A schematic of the H-shaped structure HgTe/CdTe quantum well with an impurity embedded in the central crossover region.

since no backscattering mechanism is mandated by time reversal symmetry conservation. The persistent quantum resonance is very different from any previously reported quantum resonances because of the helical state has no backscattering that is protected by time reversal symmetry. A phase diagram of transition from the Fano resonances to the B-W resonances is also presented. Quantum resonance is critical in a quantum device[13], and it is, therefore, important to find controllable quantum resonances within a topological insulator with potential applications.

3.2 Finite size effect in stripe of 2DTI

Finite size effect plays an important role in the quantum spin hall transport region in 2DTI. An analytical study of the effective 4-band model for the HgTe/CdTe quantum well with a finite strip geometry is presented[7]. Edge states of 2DTI are

doi:10.6342/NTU201603562

localized at edges and penetrate into the bulk, decaying exponentially away from the edges. Then edge states can couple together and create a gap in energy dispersion. The is an attracting phenomenon that is the difference from the quantum Hall edge states. If the edges are close enough, less than about $250nm$, the edge states on the two sides can couple together without breaking time-reversal symmetry as what we will illustrate.

The 4-band model Hamiltonian of an HgTe/CdTe quantum well yielding a 2DTI system can be given as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h} & 0 \\ 0 & \mathbf{h}^* \end{bmatrix} \tag{3.1}$$

There are four base vectors of this Hamiltonian, where $|s,\uparrow\rangle$ and $|p_x + ip_y,\uparrow\rangle$ are the pseudo-spin-up states for the upper 2×2 block, and $|s,\downarrow\rangle$ and $|-p_x + ip_y,\downarrow\rangle$ are the pseudo-spin-down states for the lower block. The tight-binding Hamiltonian of $\mathbf{h}$ in real space can be represented by [1, 14]

$$\begin{aligned} \mathbf{h} = \; &= \sum_{\mathbf{i}} \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} E_{is} & 0 \\ 0 & E_{ip} \end{bmatrix} \varphi_{\mathbf{i}} \\ &+ \sum_{\mathbf{i}} \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{ss} & V_{sp} \\ -V_{sp}^* & V_{pp} \end{bmatrix} \varphi_{\mathbf{i}+\delta\mathbf{x}} + \sum_{\mathbf{i}} \varphi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{ss} & iV_{sp} \\ iV_{sp}^* & V_{pp} \end{bmatrix} \varphi_{\mathbf{i}+\delta\mathbf{y}} \end{aligned} \tag{3.2}$$

where $\varphi = \left[c_{s,\mathbf{i}}^{\uparrow}, c_{p,\mathbf{i}}^{\uparrow}\right]^T$ are spinors for the spin up components of the s and p orbits, the index $\mathbf{i}$ represents position in real space, and δx and δy are unit vectors of the lattice constant along the x and y directions. These parameters are to be fitted with experimental data[2].

A finite strip geometry of width L with the periodic boundary conditions in the x-direction and open boundary conditions in the y direction. We calculate the Bloch band structure in terms of k_x, the good quantum number in this system, and

doi:10.6342/NTU201603562

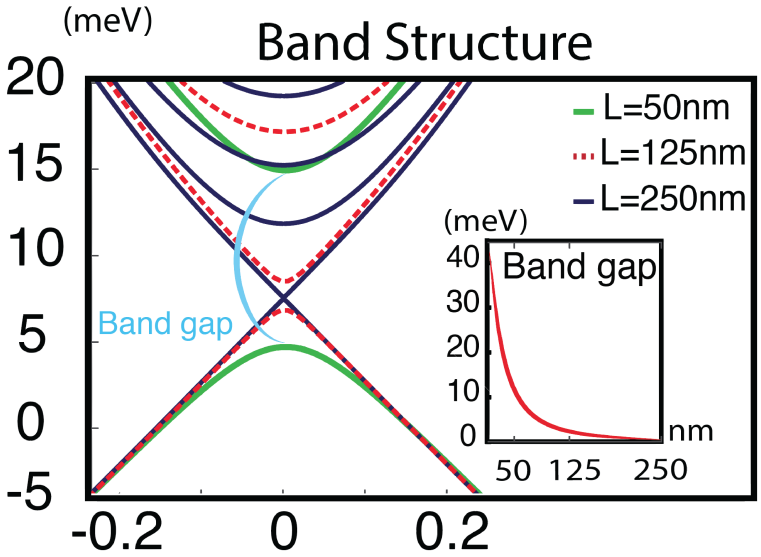

Figure 3.2: Band structure of a stripe of 2DTI with different width L near the Fermi energy. Subfigure is the band gap created by the finite size effect in terms of the width of the stripe L varying from $15nm$ to $250nm$.

show the result in Fig. 3.2. There is a band gap near the Fermi energy, the region of quantum spin Hall transport. The states near the Fermi energy show quantum spin Hall behavior and accumulate at the edges and decrease exponentially in the bulk. However, the finite size enables the states to overlap, couple with each other, and create a gap near the Fermi energy. The finite size effect can be eliminated by increasing the distance between the edges. The overlapping of the edge states is then small, so the coupling between edge is small as well and prevents gap opening. As shown in subfigure of Fig. 3.2, the band gap decreases sharply with increasing the width of the strip[7]. By the calculation presented above, the band gap is less than 0.2 meV if the width of the stripe is 250nm.

The wave density function can also be obtained as a result of calculating the norm of the eigenfunction as we have mentioned in chapter2. We now are going to

doi:10.6342/NTU201603562

focus on a point in energy dispersion that is the linear band with the energy region from $7meV$ to $0meV$ in Fig. 3.2, and the momentum point $k_x a = 0.2513$. The wave density functions at the point are shown in Fig. 3.3. For width equal to 40, the spin up electron wave density function is accumulated at one edge. The slope of the point on this linear band is negative, so the spin up electron move along the negative x-axis that obeys helical property of 2DTI as we mentioned in chapter1(add the helical statement in chapter1). The spin up electron is occupied at the edge and penetrates to the bulk o 2DTI, however, the penetrating electron is decayed so fast that cannot crosstalk the other edge. Halting the width of the stripe, equal to 20, makes the penetrating electron closer to the other edge. And further halting makes the penetrating electron reach to the other edge and open a band gap as showed in Fig. 3.2. As the width of the stripe shrink to be 5, the bulk state dominate this band and the energy dispersion is not so linear, instead, more parabolic than the case of $L = 50nm$ in Fig. 3.2

The transmission in response to the finite size effect is well-studied [7, 10] The transmission of the linear band of 2DTI is 1 for one spin channel in the energy region near Fermi energy, the quantum spin hall region. However, the finite size effect open a band gap of the linear band, and it makes the transmission be zero in the band gap as showed in the Fig. 3.4. The origin of the zero transmission can be regarded as the crosswalk of state from one edge to the other edge. A forward electron from one edge will penetrate to the other edge of a strip of 2DTI and go backward, so the transmission is zero. The transmission in terms of width as the Fermi energy, fixed at $EF = 8.47meV$ in the section, represents a step function. A sharp transition, the transmission goes from zero to one, can be found as width equal to approximately $120nm$ as showed in the upper figure of Fig. 3.4. It is the result of band gap opening due to the finite size effect. The finite size effect opens a

doi:10.6342/NTU201603562

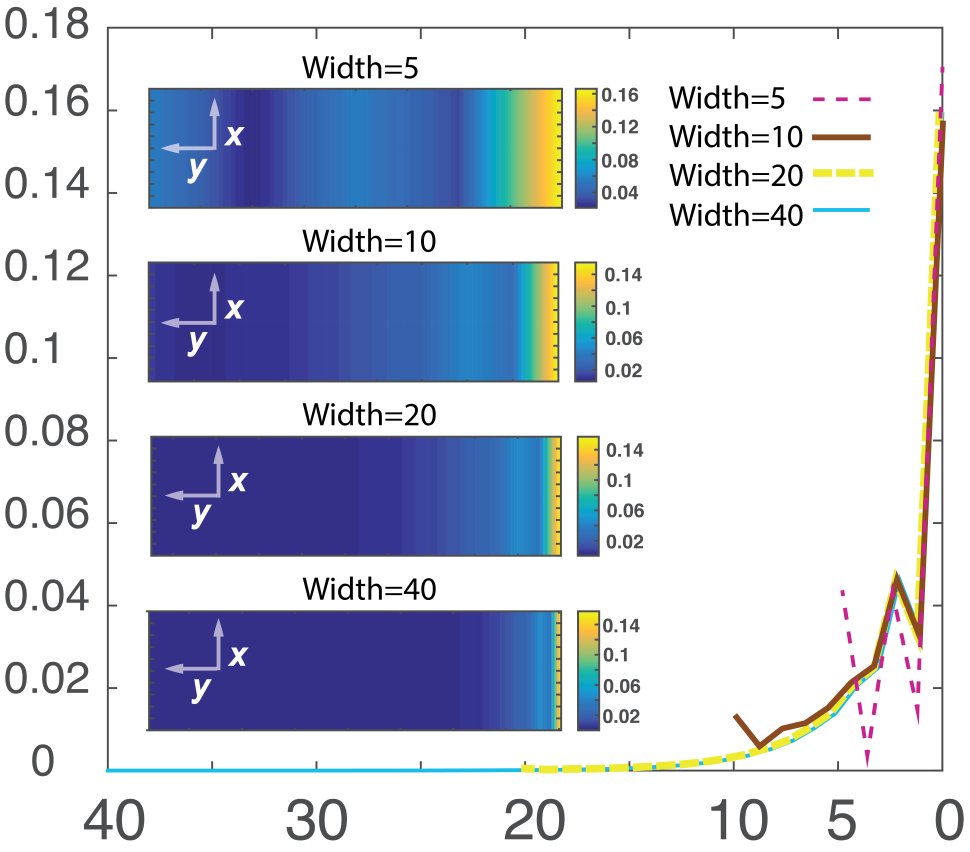

Figure 3.3: The spatial resolved wave function density in terms of the distance away from the edge of the right-hand side in a stripe of 2DTI with the difference in width. The stripe of 2DTI is in periodic boundary conditions in the x-direction and open boundary conditions in the y direction. The overall wave function density in the unit cell is illustrated in the subfigure. The wave function is the eigenvector corresponding to the momentum point $k_x a = 0.2513$ in the region of quantum spin Hall effect.

doi:10.6342/NTU201603562

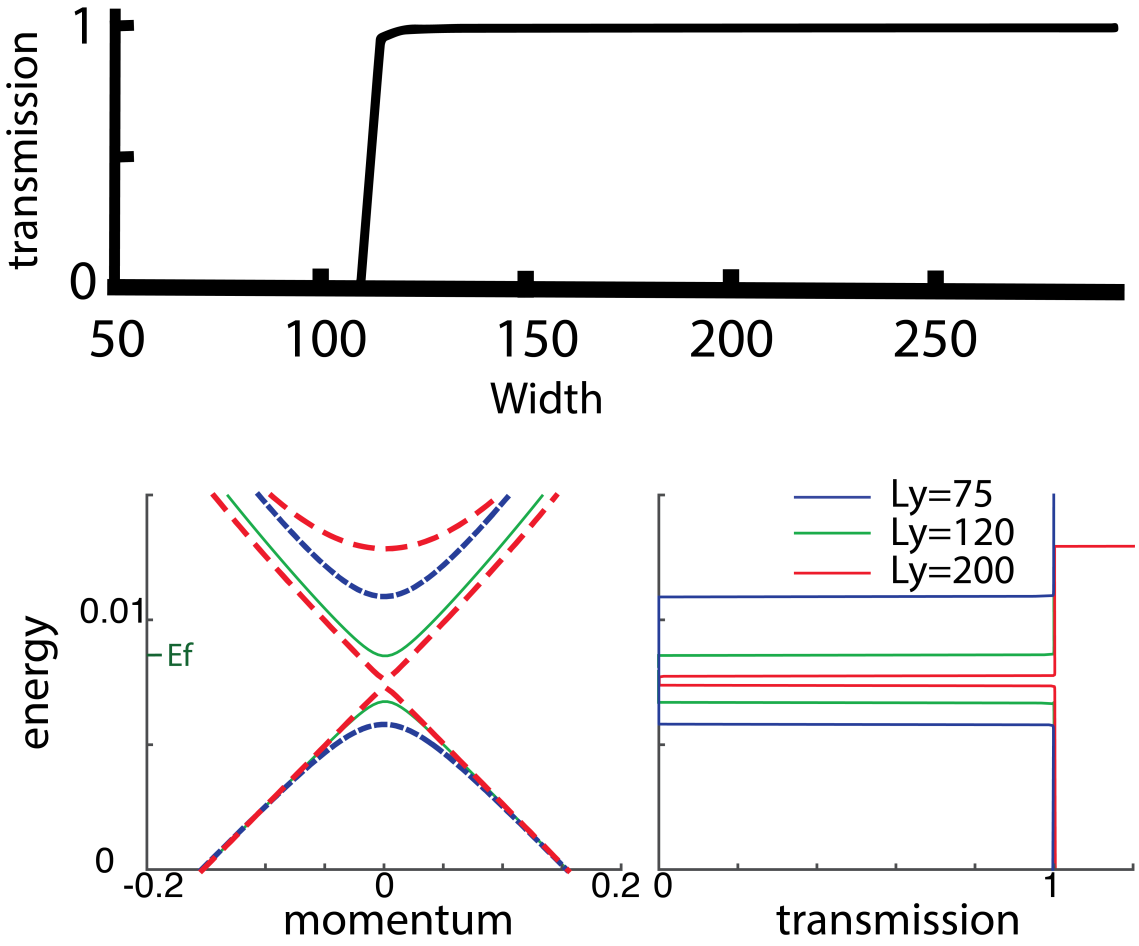

Figure 3.4: (Upper) The transmission coefficient as a function of the sample width for the Fermi energy $E_f = 8.74meV$ (Figure adopted from [10]). (Lower) The band dispersion and transmission of a stripe of 2DTI with different widths L_y. The blue line is the width of $L_y = 75nm$, the green line is the width of $L_y = 120nm$ and the Blue line is the width of $L_y = 200nm$. The band and transmission are sharing the same energy axis (y-axis).

doi:10.6342/NTU201603562

band gap of the linear band, and it makes the transmission be zero in the band gap as showed in the lower Fig. 3.4. As the width of the stripe is 75, the band gap is large and the Fermi energy lay in the gap, so the transmission would be zero. In the case of setting Fermi energy equal to $8.47meV$, finite size vanishes as width larger than $120nm$. However, presenting of impurity would not hold the criteria and we will talk it over later.

The H-shaped 2DTI without impurity, with the width of the leads set to be $330nm$, is analyzed by calculating the transmission from LEAD3 to LEAD4 in terms of the size of the crossover region as shown in Fig. 3.1. The length of crossover region L_x is small enough that two inner edge states $\psi^{\uparrow}_{i,l}$ and $\psi^{\uparrow}_{i,r}$ couple with each other. Increasing L_x reduces the finite size effect, the edge states on the left and the right of the crossover region penetrate into the bulk and exponentially decay away from the edges, so the overlap of the two inner edge states $\psi^{\uparrow}_{i,l}$ and $\psi^{\uparrow}_{i,r}$ is lessened, and the transmission is decreased. However, increasing L_y enhances the finite size effect, increases the unit of overlap of the inner edge states. The overlap increases linearly with L_y, thus increasing the transmission.

3.3 Impurity influence in a stripe of 2DTI

Bound states induced by non-magnetic impurity in TI are well studied[15] In-gap bound states induced by non-magnetic impurities in 2DTI can be derived from a modified Dirac model that depicts the same physical mathematical structure of 2DTI. The model can transform from topological trivial to topological nontrivial by varying a parameter. It can also derive the existence of topologically protected boundary states at the edges of 2DTI due to topological invariance change at the boundary of 2DTI and vacuum[16][17]. A standard Gaussian impurity potential

doi:10.6342/NTU201603562

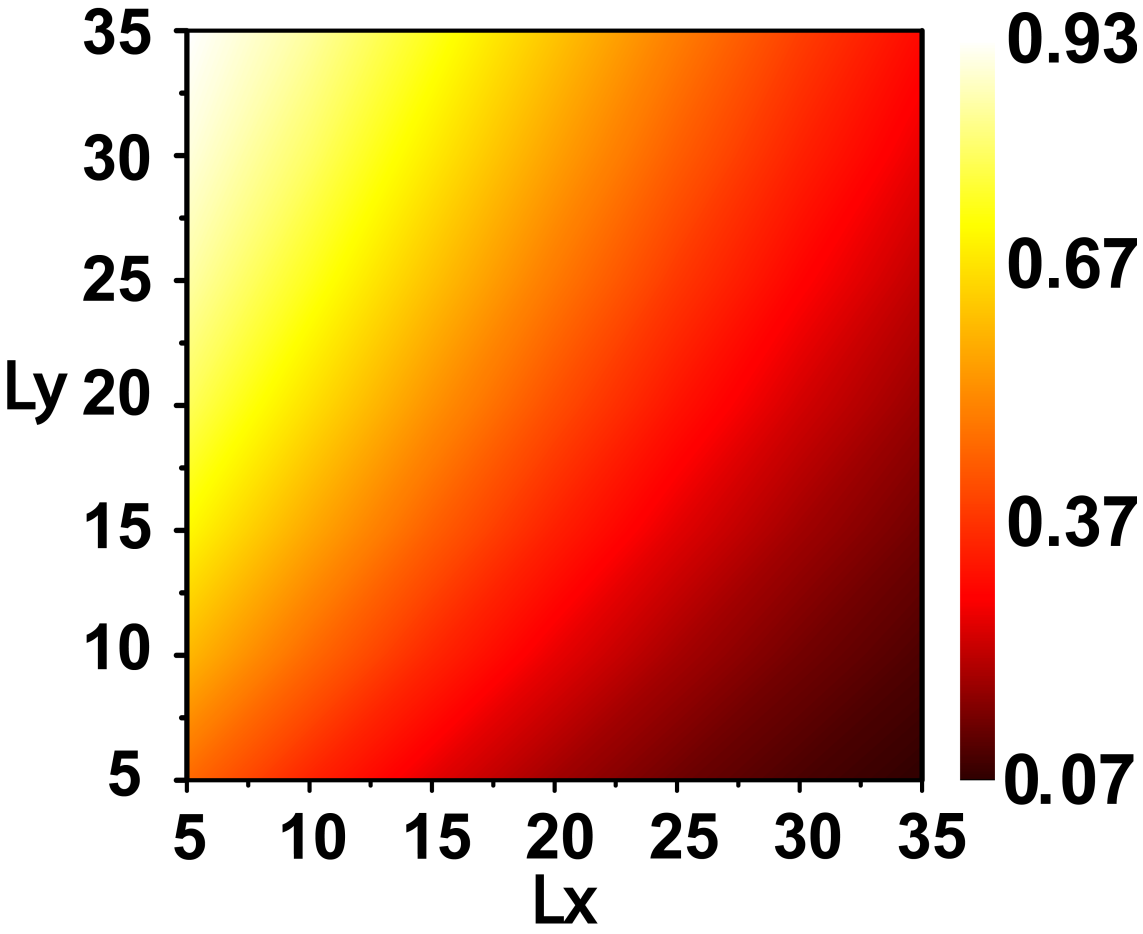

Figure 3.5: The transmission from LEAD3 to LEAD4 in a clean H-shaped 2DTI in terms of size of the crossover region (x-axis is L_x from $5nm$ to $35nm$ and y-axis is L_y from $5nm$ to $35nm$)

doi:10.6342/NTU201603562

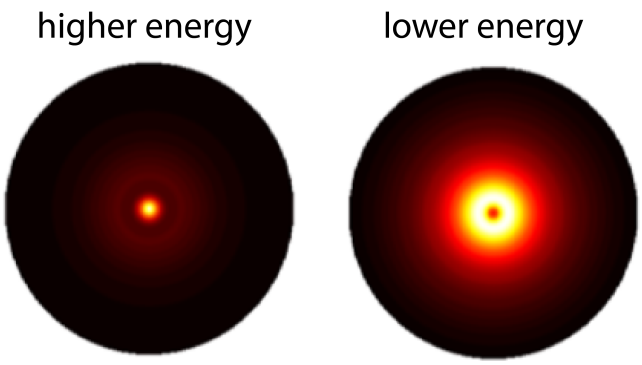

Figure 3.6: In-gap bound states induced by non-magnetic impurities in 2DTI. (Figure adopted from [15])

is introduced in the modified Dirac model to represent an isotropic non-magnetic impurity. This potential breaks particle-hole symmetry, natural behavior of the helical state of 2DTI, and create two pairs of in-gap bound states. The wave function probability distributions of the higher energy one are accumulating in the center of the impurity, however, the wave function probability distributions of the lower energy one are accumulating around the impurity as showed in Fig. 3.6.

Now we are going to study the periodical impurity in the 2DTI with the four-band model as we have adopted. The band dispersion and wave density function can be calculated by adopting Bloch theory. Fig. 3.7 shows the band dispersion of 9 layers in the x-direction, 41 layers in the y-direction as indicated without impurity. The spin up edge state is accumulated at the edge of the right-hand side as what we have mentioned in the previous section. The other edge also accumulates electron due to the wave density function in the negative momentum region. In negative momentum region, the slope of the impurity band is positive so that the spin up electron move in the positive x-direction. The spin up electron at opposite edge

doi:10.6342/NTU201603562

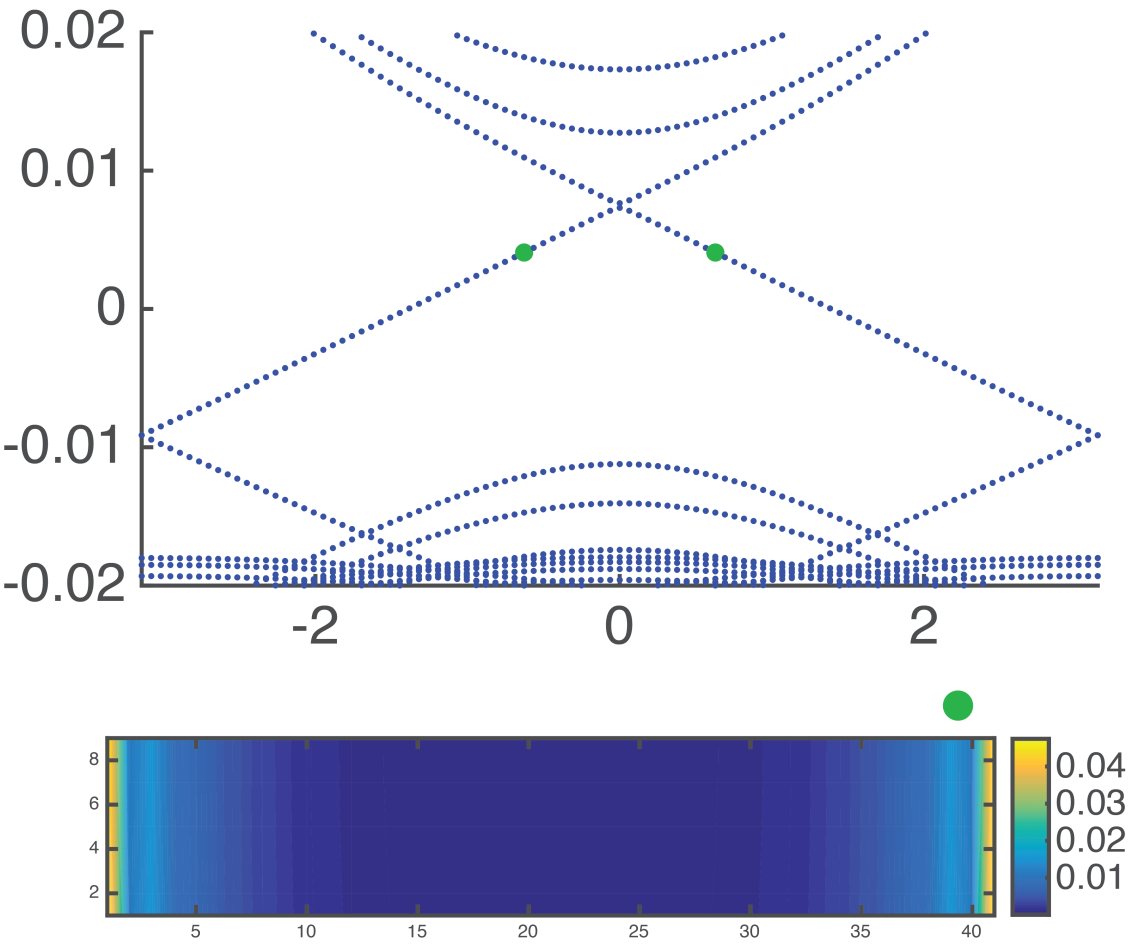

Figure 3.7: (Upper) Band dispersion of a stripe of 2DTI with width 205nm near the Fermi energy. (Lower) Wave function density of the energy indicated as the green point in upper figure. The open boundary conditions in the y direction is set 9 lattice sites (three times of Fig. 3.4).

doi:10.6342/NTU201603562

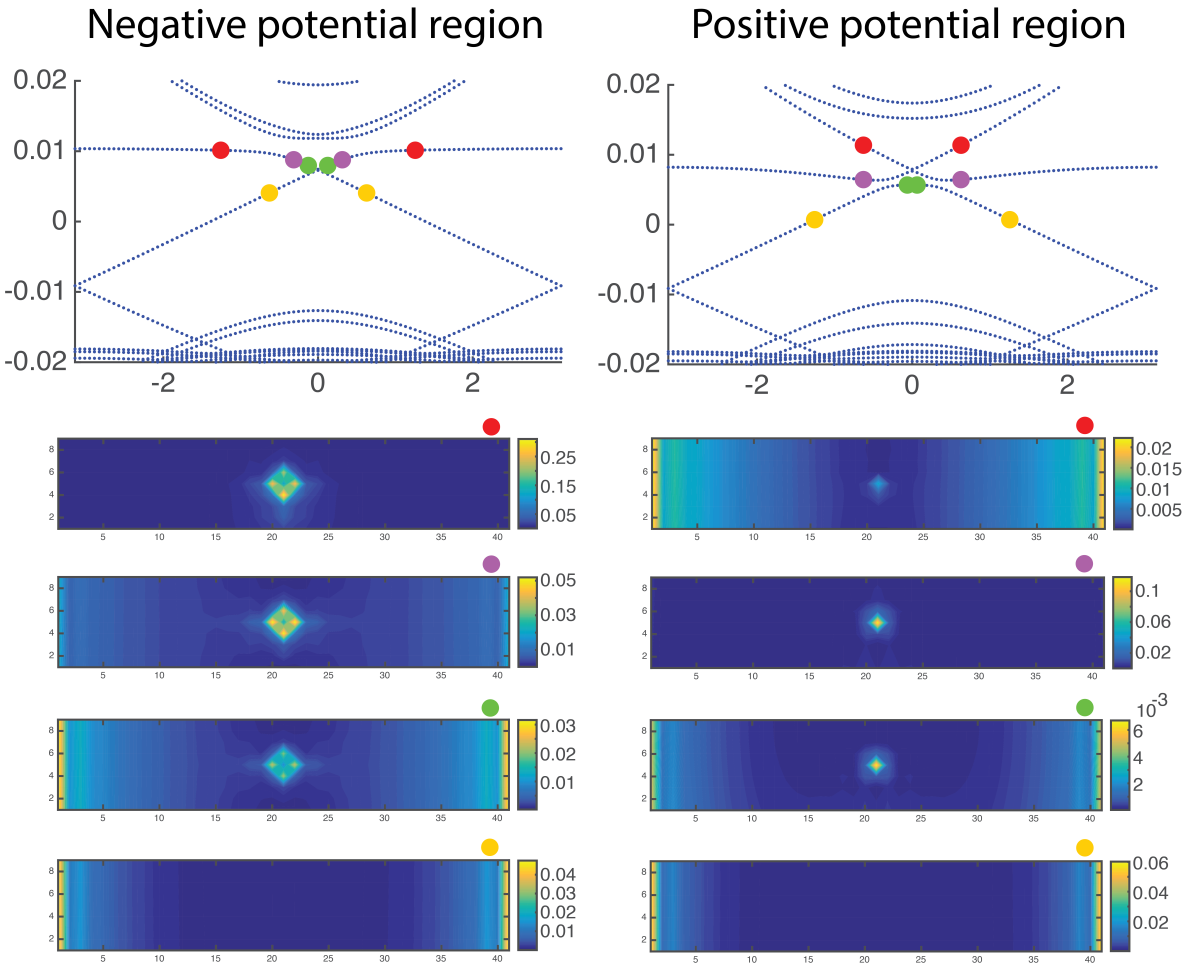

Figure 3.8: (Upper) The same stripe of 2DTI as inducted in Fig. 3.7 that embedded an impurity in the center of the unit cell. Potential energy is $0.068eV$ for the left and $-0.4eV$ for the right figure. (Lower) Spin up wave function density in different energy. Each subfigure referred to the color dots is corresponding to the energy as indicated in the upper figure.

doi:10.6342/NTU201603562

move in the opposite direction due to the helical property of 2DTI.

Now we back to the case at present non-magnetic impurity. A non-magnetic impurity is centered at a unit cell of the stripe of 2DTI as the case in the previous paragraph. The left (right) hand side of Fig. 3.8 is the band dispersion and the wave density function of positive (negative) impurity potential. The case of positive(negative) shows hole-like(electron) band dispersion as the curvature is positive(negative). From the band dispersion, the particle-hole symmetry breaking as stated by considering the modified Dirac model[15]

The wave density function in terms of energy can be obtained from different band and momentum point as labeled by colors. In the case of positive impurity potential, the red and green points show the edge state and impurity bound state co-occurrences. The purple point, the impurity bound state localized in the center of impurity without edge accumulation. As the energy going down, the yellow point shows pure edge states without impurity induced bound state as if no impurity present. In the case of negative impurity potential, the red points show impurity bound state accumulating around the impurity. Lowing down of energy makes the wave density function transit from bound state accumulation to edge states as the case of positive potential. In the yellow point of both cases, the wave density function becomes accumulating at the edges as the non-presence of impurity.

The topological protection gives us a hint that any non-magnetic impurity cannot break the transport. Indeed, it is the mathematical structure that is the most basic of this theory and cannot be broken forever. However, a finite width of 2DTI gives a chance for a crosswalk of edge states as we have mentioned and that will break quantum spin hall effect without breaking time-reversal symmetry.

The transmission is 1 for the spin up if the impurity is not presenting. The transmission of impurity in a stripe of 2DTI can be studied by considering connecting

doi:10.6342/NTU201603562

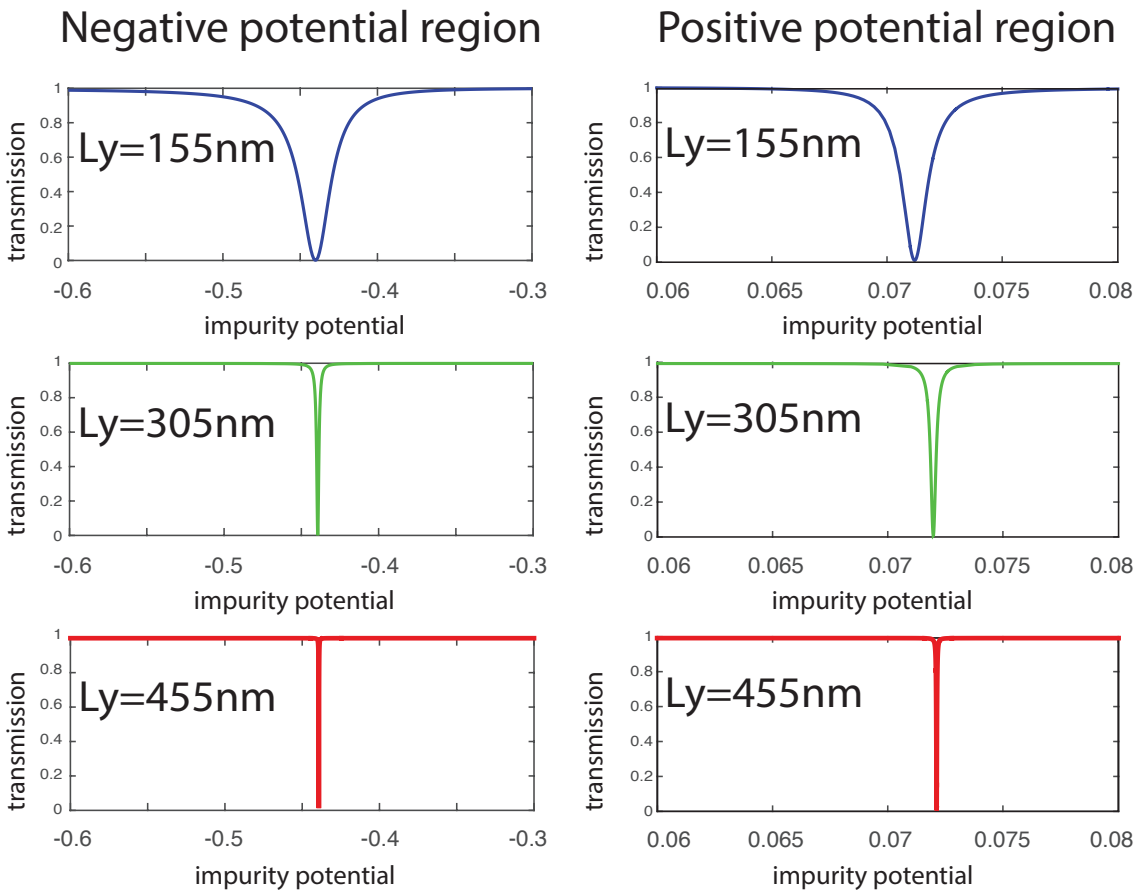

Figure 3.9: Transmission of a stripe of 2DTI that embedded an impurity at the center with the different width (155nm for the upper one, 305nm for the mid one and455nm for the lower one) as a function of impurity potential. The figures on the left-hand side are the transmission in the positive potential region and positive potential region for the left-hand side.

doi:10.6342/NTU201603562

two 2DTI lead the unit cell of 2DTI centered a non-magnetic impurity. Now we are going to study the finite size effect in the presence of a non-magnetic impurity.

The presence of impurity gives a resonance near two energy regions. As showed in Fig. 3.9, the transmission as a function of impurity potential indicate that the resonance occurs in positive (near 0.07eV) and negative (near -0.45eV) impurity potential region. The transmission drops down until zero that represent the breaking of the transmission. For the case of $Ly = 31$, the distance between edge state and impurity is 15, and this distance will show conspicuous finite size effect as we have mentioned. The resonance actually is due to the crosswalk of the edge state through the bulk via impurity so the transmission goes to zero as the impurity create a bound state that can be verified by the density of states as the potential in this region. Increasing the width of 2DTI reduces overlapping of wave function thus less crosswalk from edge state to the other edge. That result shrinks the width of resonance and we could also find the result in H-shaped 2DTI. If the width of 2DTI is large enough, then the contribution of impurity can be ignored as showed[10].

3.4 Path of H-shaped 2DTI and Formalism

In the H-shaped structure of 2DTI with impurity, there are several possible paths which electrons can follow. Three different kinds of paths with preserved time-reversal symmetry are shown in Fig. 3.10. The first kind of path is denoted by blue lines followed by electrons which stay within the same edge. It is similar to the conventional helical edge states of a two-terminal topological insulator system, the four edge states including two outer ($\psi^{\uparrow}_{o,l}$, $\psi^{\uparrow}_{o,r}$), and two inner ($\psi^{\uparrow}_{i,l}$, $\psi^{\uparrow}_{i,r}$) states. Another kind of path is denoted by purple dots denoting electrons which tunnel by the finite size effect from one edge state directly into the other edge state following a

doi:10.6342/NTU201603562

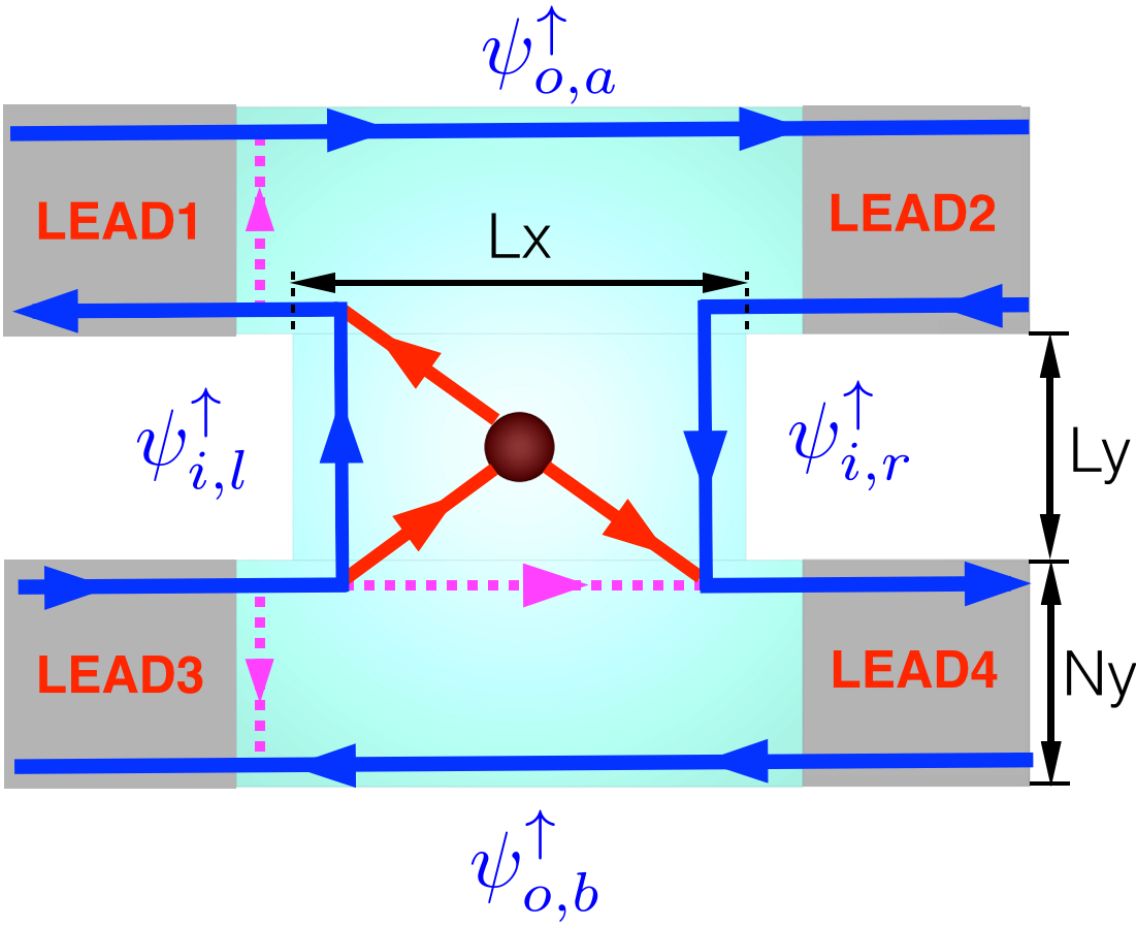

Figure 3.10: A schematic of the H-shaped structure HgTe/CdTe quantum well, consisting of two inner edge states, two outer edge states, and localized bound states. The impurity is centered in an H-shaped HgTe/CdTe quantum well consisting of a gap of width L_y, a crossover region of width L_x and four TI leads with the width of N_y. The blue lines represent the helical edge states, the purple dotted lines are finite-size induced paths, and the red lines are bound states induced paths.

doi:10.6342/NTU201603562

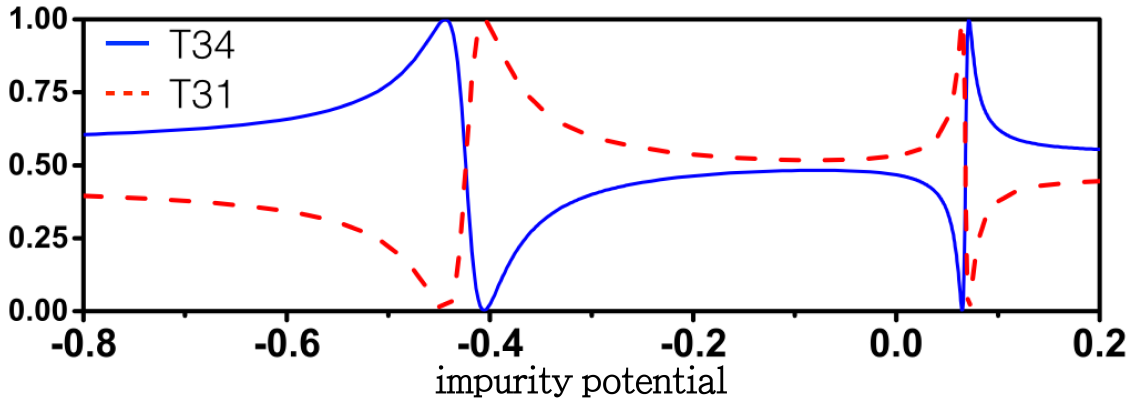

Figure 3.11: (Fano-like resonance) Transmission from LEAD3 to LEAD4 (blue solid line) and LEAD3 to LEAD1 (red dash line) in terms of impurity potential in the unit of eV (electron volt) for $L_x = 45$nm. The Fermi energy is 8.7403 meV, width of leads is $N_y = 330$nm, and width of crossover region L_y is 5nm.

path which does not involve impurity. The last kind of path is denoted by red lines followed by electrons which tunnel with impurity and show resonance properties. The electrons from the inner edge states can tunnel in and out of the impurity, however, the time-reversal symmetry requires that electrons of the helical edge states either move backward to the same side or move forward to the other side.

3.5 Fano and Breit-Wigner resonance in 2DTI

We include a non-magnetic impurity within an H-shaped 2DTI in the topological insulator states in order to study the interplay of the impurity bound states and the edge states. The transmissions are shown in Fig. 3.11 and Fig. 3.12 in terms of modulations of the impurity potential. Our result demonstrates that the transmissions from LEAD3 to LEAD1 and from LEAD3 to LEAD4 are complementary, and this result indicates that the electrons do not scatter back from the outer edge to the

doi:10.6342/NTU201603562

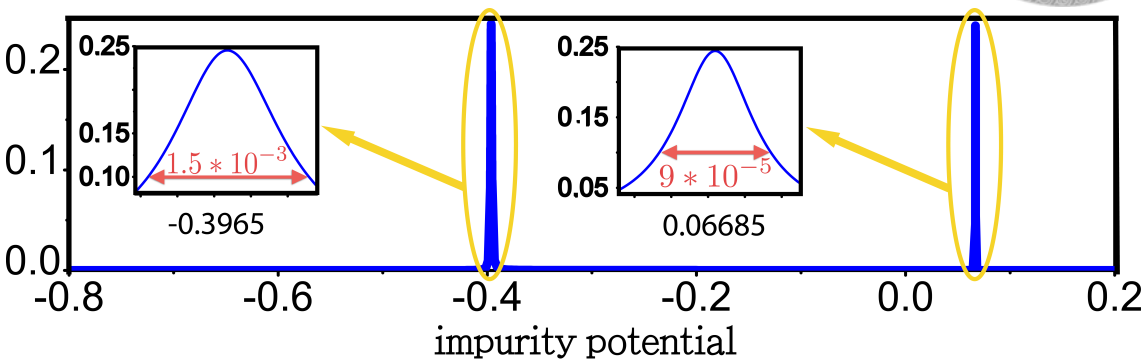

Figure 3.12: (B-W resonance) Transmission from LEAD3 to LEAD4, modulates the width of leads N_y to be 50nm.

source lead by the finite size effect. An interesting resonance can appear in two impurity potential regions where the transmission from LEAD3 to LEAD4 fluctuates significantly. This phenomenon corresponds to the Fano-like resonance representing quantum interference of discretized and continuum states. These Fano-like resonances are slightly distinct from the Fano resonances reported for a quantum-dot system[9], the interference between background states and localized bound states with backward scattering; however, backward scattering does not exist because of time reversal symmetry conservation and spin-momentum locking in 2DTI. The Fano-like resonance within the H-shaped four-terminal 2DTI is due to interference between two amplitudes of electron waves, one due to the background process of tunnelling of the edge states $\psi^{\uparrow}_{i,l}$ (blue line in crossover region in Fig. 3.10) and tunnelling states from LEAD3 to LEAD4 (purple dots in crossover region in Fig. 3.10), and the second due to a resonant process of excitation of discrete impurity states (red line in Fig. 3.10). Modulating the onsite impurity potential, the electrons of one edge state can proceed by resonant tunneling into the other edge states which control the switching from LEAD3 to LEAD4 and LEAD1. Indeed, the interesting spectrum of transmission is a representative of Fano-like resonance owing to

doi:10.6342/NTU201603562

the competition between the discrete states (i.e., the impurity bound states) and the background states. For impurity potential far from the resonant regions, the tunneling process between edge states dominates, while near the resonant energy, the amplitude of the electrons changes both in magnitude and phase abruptly, thus inducing the asymmetric profile of transmission.

If N_y is reduced, then the finite size effect of leads of 2DTI cannot be ignored. The inner edge states crosswalk to the outer edge states, the band gap of a stripe of 2DTI increases, and the transmission goes to zero[10]. Once this channel is blocked, the only way for electrons that passing from LEAD3 to LEAD4 is tunneling with the help of resonance by an impurity. Modulating the impurity potential, it is possible for electrons tunneling from LEAD3 to LEAD4 through the bound states that are created by the impurity in two impurity potential regions. B-W resonances result if we include an impurity in the center of the crossover region. As shown in Fig. 3.12, the transmission from LEAD3 to LEAD4 behaves as a symmetrical B-W resonance, differs significantly from the Fano-like resonances whose the shape is asymmetrical. Comparing with Fano-like resonances, the asymmetrical behavior comes from the interference of the resonant states and background states. The B-W resonances do not have background states, so the interference disappears, and the spectrum of transmission is symmetrical.

3.6 Phase transition between B-W resonances and Fano-like resonances

Increasing the width of leads brings about a transition from B-W resonances to Fano-like resonances. The phase diagram Fig. 3.13, derived by calculating the trans-

doi:10.6342/NTU201603562

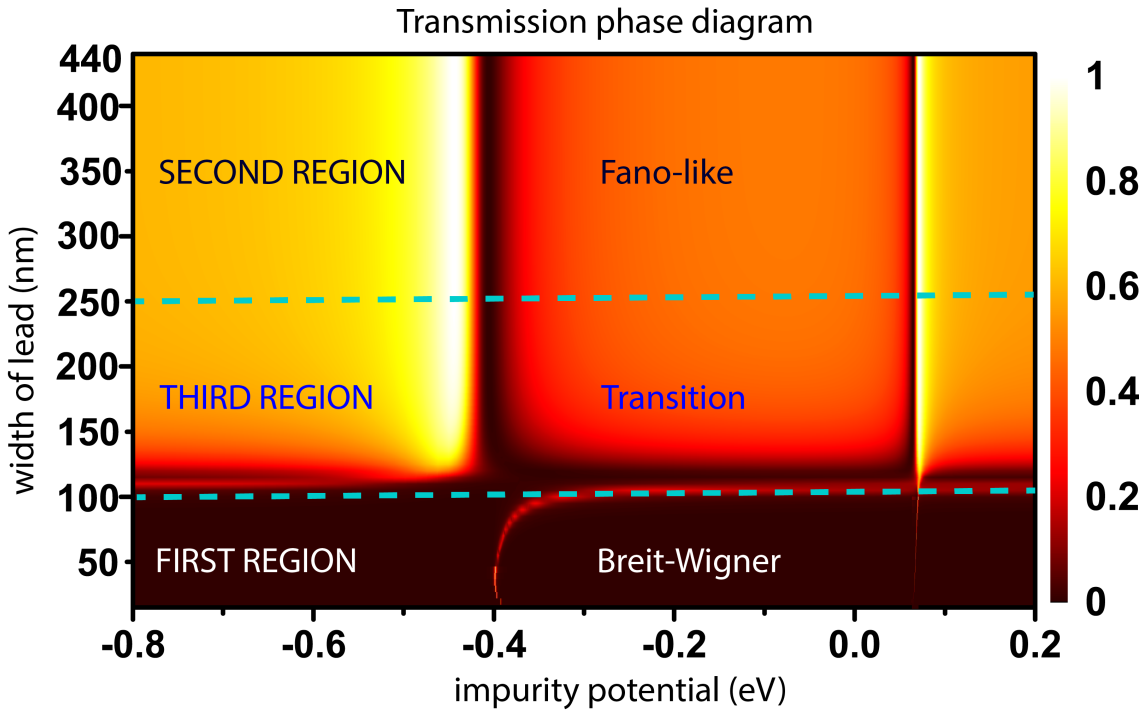

Figure 3.13: Phase diagram of transmission with N_y varying from 15nm to 440nm. The size of the crossover region is fixed. The longitudinal direction is the length of leads (in units of a nanometer), and the transverse direction is the impurity potential (in units of eV).

doi:10.6342/NTU201603562

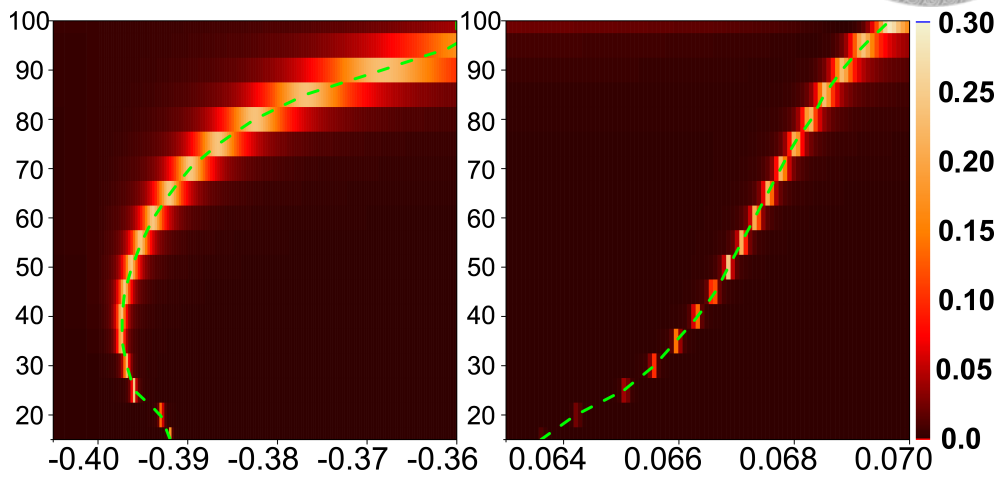

Figure 3.14: The region of B-W resonance, N_y varies from 15nm to 100nm, and the impurity potential is near the two resonance energy regions. The green dashed lines show the peak of traversal lines.

mission from LEAD3 to LEAD4 while modulating the width of the leads N_y and of the impurity potential, can be divided into three regions: The first one belongs to short-width of leads ($N_y < 100$nm), the second to long-width of leads ($N_y > 250$nm), and the third region is due to mid-width of leads ($250\text{nm} > N_y > 100\text{nm}$).

The B-W resonances can be observed in the first region as shown in Fig. 3.14. Along transversal lines (i.e., modulating impurity potential and fixing the width of leads), the oscillation of transmission can be found, and there are two bright points correspond to two peaks of transmission which represent two resonances. The width of B-W resonance in a negative impurity potential is larger than in the positive one. We will give a more detailed discussion of the width of resonance in the next section.

The Fano-like resonances can be found in the second region as shown in Fig. 3.13. The oscillation of the Fano-like resonances corresponds to the change of bright longitudinal lines for constructive interference and dark longitudinal lines for destructive interference. The widths of oscillation near the two resonances region show also the

doi:10.6342/NTU201603562

same features as in the case of the B-W resonances which are larger in the negative potential region. The Fano-like resonances are stable with increasing N_y due to the absence of finite size effect of the leads, which can be reduced if the width of leads is increased.

In the third region, the transition from the B-W resonance to the Fano-like resonance, the channel from LEAD3 to LEAD4 is partially blocked. An electron from LEAD3 can transport through any of the paths described in Fig. 3.10. On the vertical axis in the absence of impurity (i.e., the impurity potential is zero), the transmissions between leads are calculated: The transmission from LEAD3 to LEAD2 is increased, cross-walking of $\psi^{\uparrow}_{i,l}$ to $\psi^{\uparrow}_{o,a}$ (i.e. the electrons near LEAD1 tunnel from the inner edge to the outer edge), and reaches its maximum at $N_y =$ 120nm, and then decreases as the width of leads widened. The transmission from LEAD3 to LEAD4 is increased, however, and Fig. 3.13 shows that it fluctuates near $N_y = 115$nm. It is the interference with the electron from LEAD2 due to which the strength of the electron wave from LEAD3 varies in response to the finite size effect, which makes the fluctuation of transmission a function of the width of the leads. We have found a similar fluctuation of the transmission from LEAD3 to LEAD2, due to the interference with the electron from LEAD1. In this region, the resonance shows Fano-like behavior, asymmetry of oscillation, and the transmission cannot reach unit of conductance until the finite size effect reduces. The width of the stripe of the 2DTI is about 250nm as calculated in the section III the finite size effect.

3.7 Widths of resonances

We will study the width of resonance in this section. The widths of resonance are related to the coupling of local bound states and edge states (i.e. resonant states).

doi:10.6342/NTU201603562

Enhancing the coupling between the edge states and the local bound states broadens the width of the resonance.

In the Fano-like region, the width of the Fano-like resonance can be varied by varying L_x. The width of the Fano-like resonance increases in response to enhanced overlapping of the edge states and the local bound state achieved by reducing L_x. Reducing L_x leads to an enhancement of transmission in a clean H-shaped 2DTI as indicated in the section on the finite size effect. In an H-shaped 2DTI with embedded impurity, a reduction of L_x leads to an enhancement of the overlapping of the edge states and the bound state, so the resonant states are enhanced and the width of the Fano-like resonance broadens.

In the B-W region, the width of the resonance broadens with increasing N_y as shown in Fig. 3.14 and simultaneously the finite size effect of the leads decreases and two inner edge states (i.e. $\psi^{\uparrow}_{i,l}$, $\psi^{\uparrow}_{i,r}$ indicated in Fig. 3.10) are enhanced. This effect also enhances the overlapping of the edge states, and of the impurity bound state, so the coupling between edge and impurity is also enhanced. A simple model in 1D-quantum dot system[18] shows that the enhanced coupling yields wider the width of resonance, in correspondence with our result. The width of B-W resonance in a negative impurity potential is larger than in the positive one due to the wave properties of the bound states as indicated in Fig. 3.15. The local densities of states (DOS) near resonance region are presented in Fig. 3.16. The impurity DOS in a negative potential is more extended spatially than in a positive one. The greater spatial extension of the DOS implies more coupling between the edge states and the bound states (blue curve in Fig. 3.15).

The coupling of the edge state and the local bound state can be verified by the spectral function[1]. The spectral function gives the number of states, and the peak means there is a state or there are several degenerate states there. The more

doi:10.6342/NTU201603562

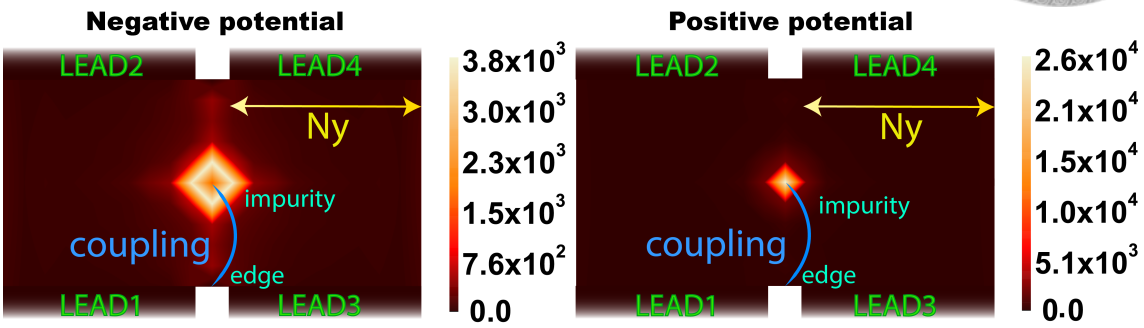

Figure 3.15: DOS (in the units of inverse eV) of negative (left) and positive (right) potential resonance region. The width of the leads is 50nm. DOS of positive potential accumulates in the center of impurity, but DOS of negative potential accumulates around the impurity. The coupling between the impurity bound state and the edge state is indicated by the blue curve. The spectral functions between those states are shown in Fig. 3.16.

extensive is the spectral function, the stronger is the coupling between two sites. The spectral function of the impurity site and the edge as a function of Fermi energy is presented in Fig. 3.16. We set the impurity potential to be -0.3919 eV which is the resonant potential of $N_y = 15$nm at the fixed Fermi energy (8.74 meV) as shown in Fig. 3.16. There is a sharp peak for $N_y = 15$nm at the Fermi energy equal 8.74 meV, which means that the coupling between the bound state and the edge state exists near a small region of the Fermi energy and that the coupling is weak. The peaks of the spectral functions are the resonant points as a function of the Fermi energy in terms of the width of the leads N_y. The width of the leads affects the resonant state in the crossover region because of the finite size effect. Increasing N_y leads to an enhancement of the correlation of impurity and helical states, so the resonant states are enhanced.

We can deduce three features from the spectral function in Fig. 3.16. First, the

doi:10.6342/NTU201603562

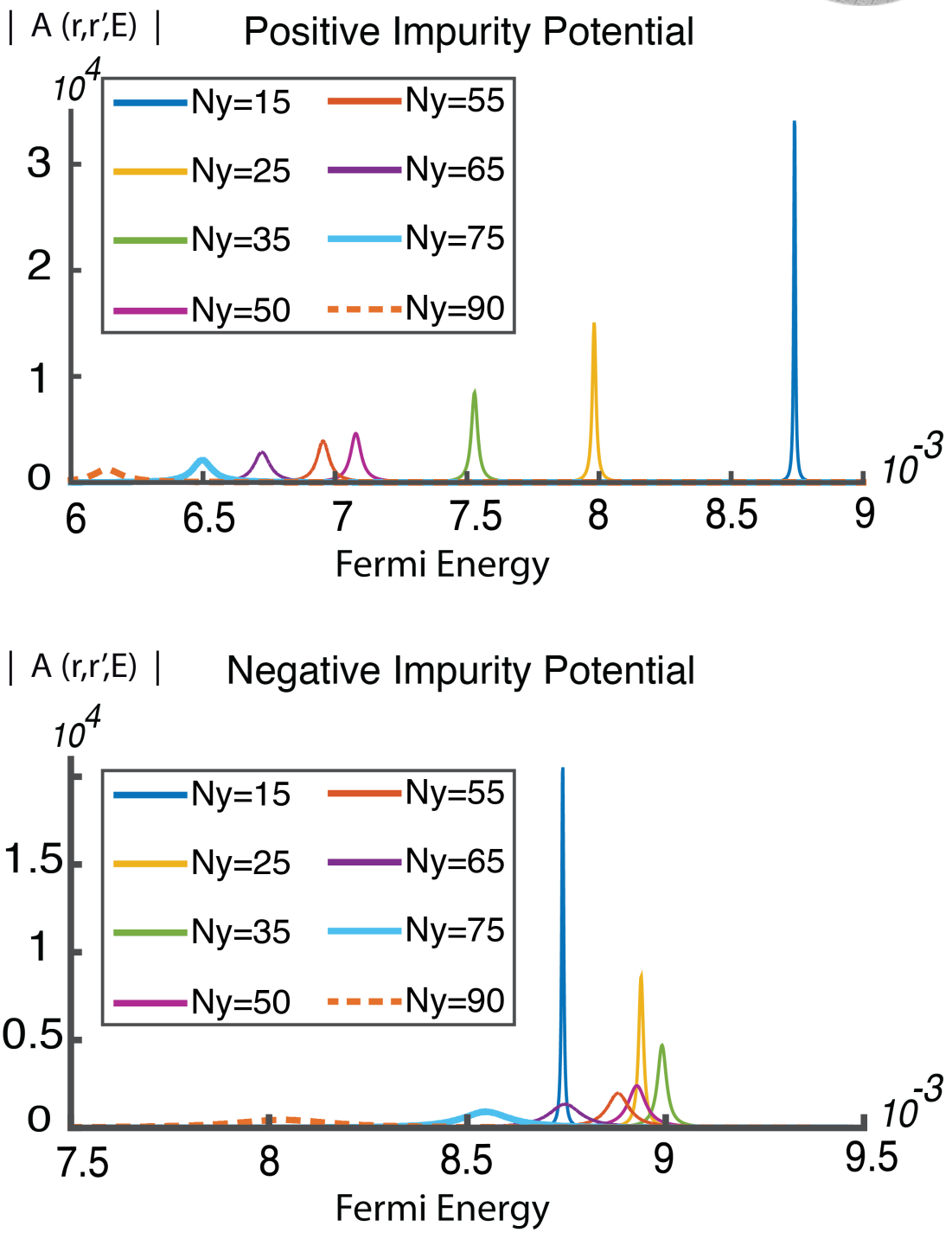

Figure 3.16: (b) Spatially resolved spectral function $\mid A(r, r', E) \mid$ of the impurity bound state and the edge state for a negative potential (upper figure) and a positive potential (lower figure) in terms of the Fermi energy (in units of eV) for different sizes of leads (in units of nm). The fixed impurity potential is set to be -0.392eV for the negative potential and 0.0636eV for the positive potential.

doi:10.6342/NTU201603562

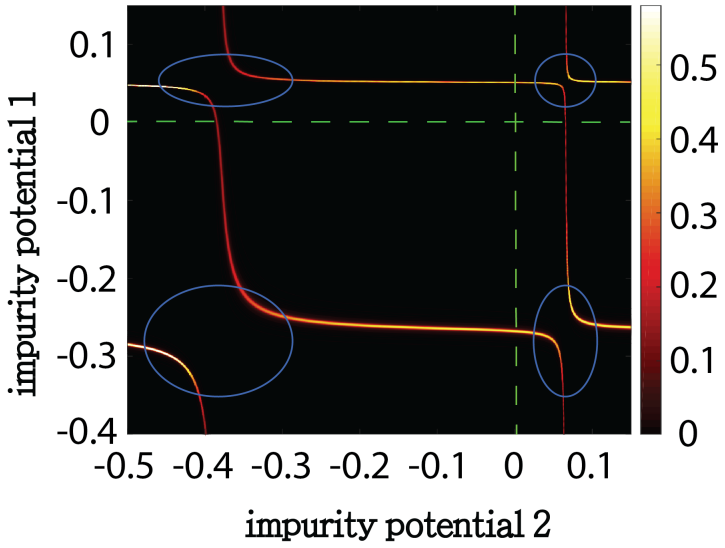

Figure 3.17: Transmission of two impurities in the crossover region in terms of an impurity potential (from -0.4 to 0.15 eV) located near LEAD3 (vertical axis) and an impurity potential (from -0.5 to 0.15 eV) located near LEAD4 (horizontal axis).

spectral functions become extensive with increasing width of the leads. This implies a stronger correlation of impurity and helical edge states so that the resonance of width increases. Second, the peaks of spectral functions, that is, the resonant energies, are distinct for each N_y. As in the first feature, the resonant state varies in terms of N_y, thus changing the resonant energy. Thirdly, the spectral functions are more extensive for a negative potential at each N_y. This signifies that the correlation between impurity and helical edge states in a negative potential is stronger than in a positive one for a fixed N_y. This also affects the width of the resonance and consists of the discussion of DOS as shown in Fig. 3.16.

doi:10.6342/NTU201603562

3.8 Two impurities interplay

A single impurity can be manipulated with a gate voltage and thus it is quite realistic for real experiments. However, defect or non-magnetic interaction may also create an impurity potential so that multiple scattering may not be avoided. In order to illustrate this point, an additional impurity is embedded in the crossover region. Varying the additional impurity potential in the crossover region creates two additional resonances in the region of positive and negative values of the potential respectively. The transmissions from LEAD3 to LEAD4 of two impurities in the crossover region in terms of an impurity potential (from -0.4 to 0.15 eV) located near LEAD3 (vertical axis) and an impurity potential (from -0.5 to 0.15 eV) located near LEAD4 (horizontal axis) are shown in Fig. 3.17. Two green dashed lines represent an impurity with zero potential, and only the other impurity gives rise to resonances. The position of the resonant energy peak in the B-W region depends on the position of the impurity thus the resonant energies of impurities are distinct. If one impurity potential is fixed, there is only one pair of resonances as in the case of one impurity. The influence of an additional impurity is changing the resonant energy peak in the four blue circles as shown in Fig. 3.17, indicating the region where the two impurities both resonate. However, out of the blue circle region, the resonance of an impurity is barely affected by another impurity as shown. If the potential of the additional impurity is fixed, then there are only two resonances which can be found, and all general physics of persistent quantum resonance can thus be clearly demonstrated on a single impurity. It is therefore not necessary to consider multiple scattering.

doi:10.6342/NTU201603562

3.9 Summary

In summary, we have observed a phase transition of Fano-like resonances and B-W resonances in a 2DTI with H-shaped geometry. We have explained the phenomenon by bound states created by an impurity and crosswalk states generated by a finite size effect. The H-shape geometry allows another path of edge states, so Fano-like resonances can be observed. The transmission shows B-W resonances as the edge channels are all blocked. From the application point of view, the system is potentially an efficient switch of spintronics devices, blocking the channel of the same spin orientation completely, and maintain high transport performance for the other spin orientation. The transmission can be efficiently modulated by the impurity potential, and the impurity potential can be controlled by a local gate voltage. Additional impurity in the crossover region creates two additional resonances in the region of positive and negative values of the potential respectively. However, additional impurity barely effects the original transport properties except additional impurity is in its resonance region.

doi:10.6342/NTU201603562

Bibliography

[1] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[2] M. König *et al.*, Science **318**, 766 (2007).

[3] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[4] X.-L. Qi and S.-C. Zhang, Physics Today **63**, 33 (2010).

[5] C. Xu and J.E. Moore, Phys. Rev. B **73**, 045322 (2006).

[6] C. Wu, B. A. Bernevig, and S.-C. Zhang, Phys. Rev. Lett. **96**, 106401 (2006).

[7] Bin Zhou, Hai-Zhou Lu, Rui-Lin Chu, Shun-Qing Shen, and Qian Niu, Phys. Rev. Lett. 101, 246807 (2008)

[8] A. E. Miroshnichenko, S. Flach, and Y. S. Kivshar, Rev. Mod. Phys. **82**, 2257 (2010).

[9] A. C. Johnson, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **93**, 106803 (2004).

[10] H. H. Lee, J. Y. Liu, C. R. Chang, and S. Q. Shen, Phys. Rev. B **88**, 195149 (2013).

doi:10.6342/NTU201603562

[11] S. Sim, N. Koirala, M. Brahlek, J. H. Sung, J. Park, S. Cha, M. H. Jo, S. Oh, and H. Choi, Phys. Rev. B 91, 235438 (2015)

[12] W.-Y. Shan, J. Lu, H.-Z. Lu, and S.-Q. Shen, Phys. Rev. B **84**, 035307 (2011).

[13] B. luk'yanchuk *et al.*, Nature Mater. **9**, 707 (2010).

[14] C. Xu and J.E. Moore, Phys. Rev. B **80**, 165316 (2009).

[15] Jie Lu, Wen-Yu Shan, Hai-Zhou Lu and Shun-Qing Shen, New Journal of Physics **13** (2011) 103016

[16] Liang Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007)

[17] Xiao-Liang Qi, Yong-Shi Wu, and Shou-Cheng Zhang, Phys. Rev. B **74**, 045125 (2006)

[18] A. Goker, Phys. Status Solidi B **247**, 129 (2010).

doi:10.6342/NTU201603562

4

Normal metal two-dimensional topological insulator junction

4.1 Overview

The generation and control of spin polarization that can filter spin and produce spin current are essential in spintronics. Several methods were proposed to realize spin polarization or spin filter. Some approaches make use of the magnetic field to let different spins filtered or differentiable by magnetoresistance response [1, 2] or the Aharonov-Bohm interference [3]. Spin filtering is also possible by increasing concentrations of spin-filtering defects [4]. It is natural to take advantage of magnetic materials by passing carriers through adjacent ferromagnetic (or ferrimagnetic) layers [5, 6, 7]. Some realizations directly use ferromagnetic materials as a quantum well state to influence spin polarization [8, 9, 10, 11]. However, using ferromagnetic materials gets another way to filter spin from the material point of view. Spin-orbit coupling (SOC) is actually another good choice to change spins

doi:10.6342/NTU201603562

in carrier transport through materials, especially on the basis of spintronics integrated on semiconductor applications. For example, it was considered to be crucial that intrinsic spin-polarized current can be generated in the system with substantial Rashba spin-orbit coupling [12]. Spin filter or selectivity can also be achieved by asymmetric interband tunneling exploiting large band SOC [13]. In strongly curved materials like helical DNA, spin selectivity can even be induced by a dynamical evolution of orbit angular momentum through SOC [14].

It has a great advantage to find giant SOC materials that can realize substantial spin polarization to allow us to control spin related phenomena electrically without using a magnetic field. Therefore, it has long been part of the key issues to realize large Rashba interaction or sizable spin splitting. One way to change Rashba strength in surface alloys is by electron doping or adsorption of suitable atoms [15, 16, 17]. However, this is equivalent to change material nature chemically. Other methods like quantum well state (QW) or monolayer deposited on a substrate can enhance Rashba splitting by the interplay of atomic and interface potential gradients [18, 19]. It is found that giant Rashba splitting can be achieved even at quantum well state with light atoms on substrates [20, 21, 22]. This is due to the coupling or hybridization of QW state with the surface states. Spin polarization and spin splitting can be manipulated by the interplay of quantum confinement and surface spin-split states from SOC [23, 24, 25]. Therefore, quantum confined states in thin films or QWs coupled to other materials or substrate has proven to be an effective way to enhance or decrease Rashba splitting by ways such as varying the thickness of thin film.

On the other hand, materials that are utilized to hybridize with a quantum well state in order to manipulate Rashba splitting can be varied. People usually choose materials with strong enough SOC to provide surface spin-split states. However,

doi:10.6342/NTU201603562

a topological insulator (TI) which also originates from SOC but provides more restricted and interesting surface states offer another choice to couple with the thin films. The TI has nontrivial topological quantity Z_2 invariant [26, 27]. When TI is connected to other trivial materials like a normal bulk insulator, the interface will have surface or edge states due to the bulk-edge correspondence [28, 29, 30]. The surface or edge states of TI is spin-momentum locked and hence has natural spin selectivity. At one interface, the edge spin up state flows in one direction and the spin down state flows in the opposite direction. It implies that in each chosen direction, one of the spins is blocked and therefore the edge states behave like spin-split states. When the TI is adjacent to the quantum well states, the spin selectivity provided by the TI helps the QW state produce giant Rashba-like spin splitting by hybridization, as presented in this paper. As a demonstration, we build the two-dimensional TI/normal metal (NM) QW mixture as transport channels along the interface direction. When connected to different leads for the NM and TI separately in the drain end, the TI/NM junction can divert the spin transport efficiently, a key feature required by spintronics.

4.2 Spin splitting induced by helical state

TI/NM junction can be investigated by considering the system that is comprised of one layer of NM attaching HgTe quantum well which is proposed as a 2DTI [31]. Energy band studying for the junction with varying the coupling strength between (see method) NM and HgTe quantum well is presented in this article. The band dispersion is calculated in terms of k_x that is the good quantum number in this system by considering a finite strip geometry of one site of NM attaching HgTe quantum well with the periodic boundary conditions in the x-direction and open

doi:10.6342/NTU201603562

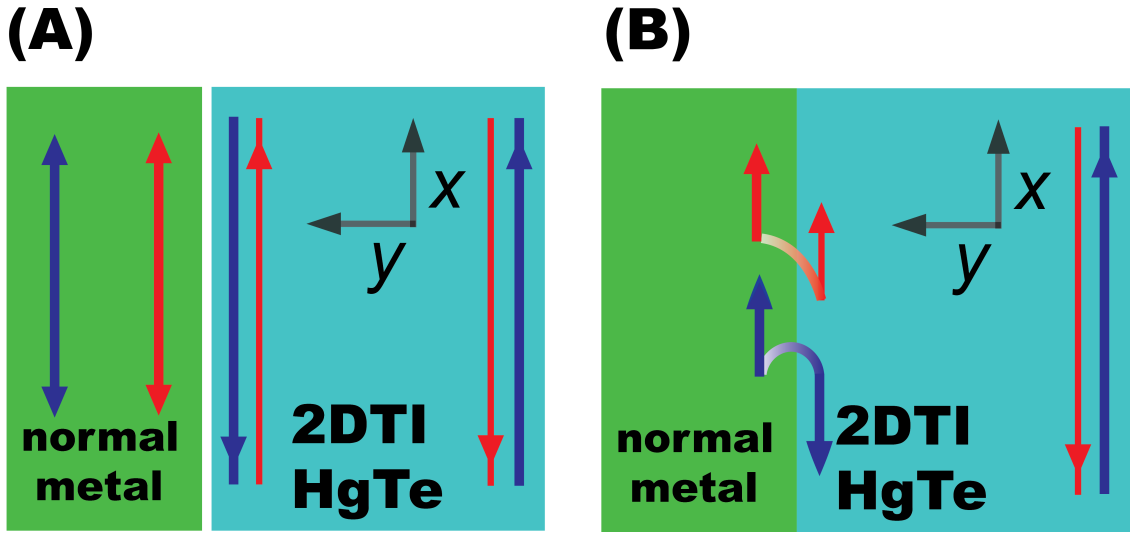

Figure 4.1: (left) Schemetic figure of path of isolated NM and 2DTI. (right) Schemetic figure of path of 2DTI/NM junciton. The red (blue) color represents the path of spin up (down) electron and the array is the direction of movement of electron.

doi:10.6342/NTU201603562

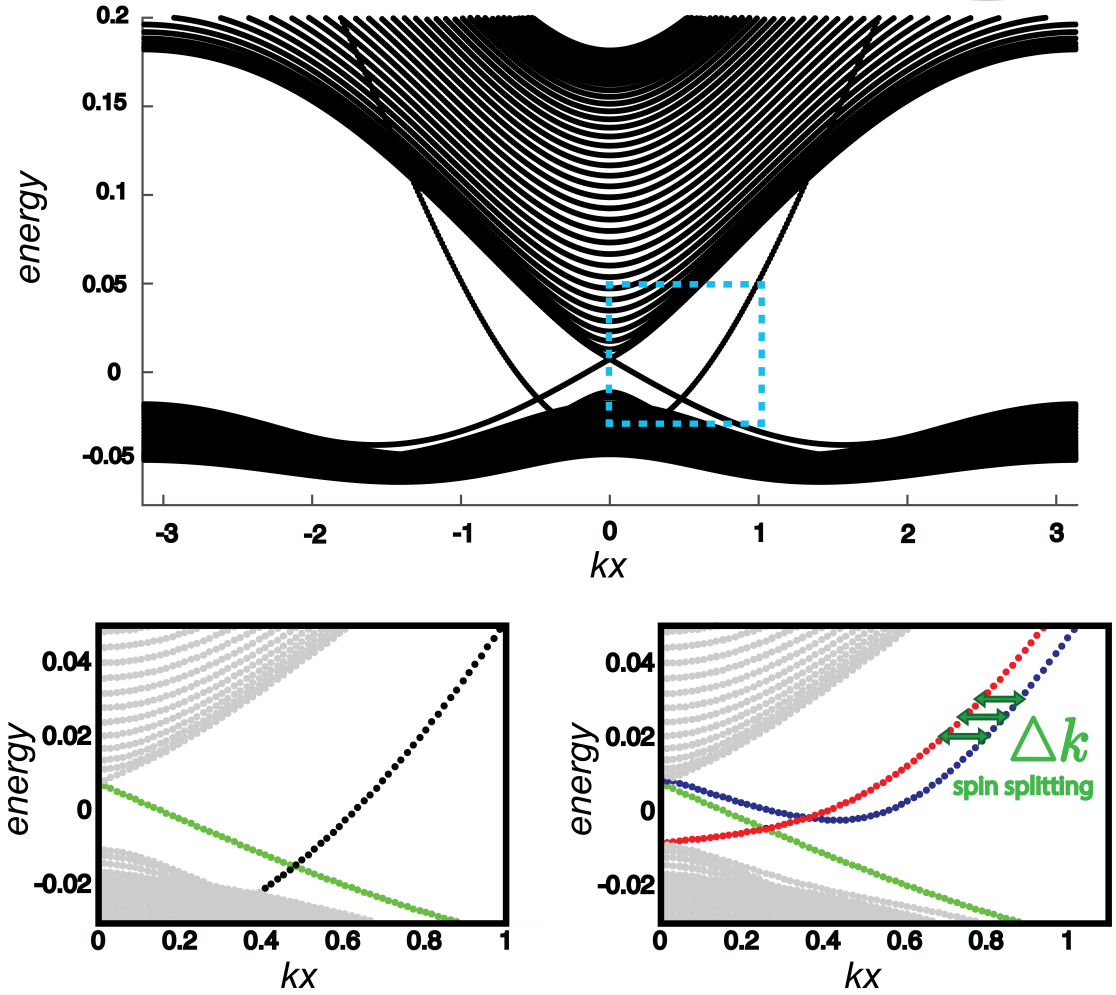

Figure 4.2: (upper) The band structure of normal metal and HgTe quantum well. The band of normal metal is parabolic and the band of 2DTI is linear near Fermi energy (energy equal to zero). (left) The band structure of normal metal and 2DTI near Fermi energy. The green dots are the edge states of 2DTI and black dots are the quantum well state of normal metal. The edge states are linear dispersive, spin up and down are degenerated and localized at different edges, and show Dirac electron behavior that can not be scattered back by impurities. The quantum well is parabolic dispersive and spin up and down degenerated. (right) The junction of normal metal and HgTe quantum well. The quantum well state of normal metal hybridizes with one of the edge states of 2DTI (the edge state near the interface between 2DTI and NM) and open a band gap, however, the other spin hybridizes with the bulk state. Thus the spin of normal metal split and form Rashba-like band and the difference of momentum for spin up and down denoted by $\triangle k$.

doi:10.6342/NTU201603562

boundary conditions in the y-direction. The 2DTI is a strong spin-orbit coupling system, so we need to take the degrees of freedom of spin into the band calculation. The band dispersion of the junction is symmetry along the gamma point, however, the orientation of spin in positive k_x region and negative k_x region is in opposite direction. We can thus focus on the band dispersion of the positive k_x region and the negative k_x region is just an inversion of spin orientation along the gamma point. As showed in the left figure in fig. 4.2, the band of isolated NM and isolated 2DTI are spin degenerated and behave parabolic and linear respectively near the Fermi energy that is placed at zero in tight-binding model approach. As indicated in the section of "method", the spin and orbit degree of freedom in the Hamiltonian of NM are decouple, however, those are coupled for the Hamiltonian of 2DTI. The coupled spin and orbit degree of freedom make the transport behavior of electron spin in NM and 2DTI are different. The electron spin can travel freely in any direction in NM, however, the spin-momentum locking in 2DTI enforces the spin up and down edged electron move in the opposite direction near the Fermi energy[32, 33].

The spin up electrons move clockwise in 2DTI and spin down electrons move counterclockwise along the edge of 2DTI. The edged electron spin is actually the helical states that behave like Dirac electron due to the band is linear dispersive near Fermi energy defined as quantum spin hall energy region[31, 32]. The helical states are robust against impurity [34, 35] if the preservation of time reversal symmetry stays on the total system. If there are unpolarized currents flow into NM, there is no local magnetization induced, however, the moving electron accumulates at edges with opposite orientation of spin in 2DTI.

Once the NM and 2DTI do connect, the electron hops through the junction from helical states of 2DTI to the quantum well states of NM and vice versa. The helical state at the edge that does not connect with the NM (green dot in the right figure

doi:10.6342/NTU201603562

of fig. 4.2) is remaining the linear desperation without hybridization due to the distance is too far away from the interface of NM and 2DTI. However, the helical state at the edges that connect with the NM hybridizes with the quantum well states of NM. Take the hopping is spin independent so that there is no flipping mechanism in this penetration. We can predict how do bands hybridize in the junction base on two characters. First, the spin up(down) quantum well states and the state of NM can only hybridize with the spin up(down) helical state of 2DTI due to spin independent hopping in the quantum spin hall region. Second, The electron moving in opposite direction is in opposite spin orientation due to spin-momentum locking, thus the helical electron move in positive x direction is spin up and the spin down electron move in negative x direction at the edge that connects with the NM as shown in fig. 4.1. Combining these two characters, we can conclude that the spin up electron moving in positive x direction in NM can remain the transport direction as hopping to the helical state of 2DTI, the spin down electron which moves in positive x direction in NM have to change the transport direction as hopping to the helical state of 2DTI as shown in fig. 4.1. The linear band in the region of positive k_x is the negative slope in terms of k_x. Thus this band describes the helical electron moving in the negative x direction. So spin down linear band hybridizes with quantum well state and the spin up quantum well state lack some state in response to electron penetration in the region of positive k_x as showed in the right figure in fig. 4.2. The origin of the asymmetry hybridization is actually brought about by the unique property of topological insulator, spin-momentum locking. There are two points we can point out in the band dispersive as showed in fig. 4.2. First, the band of spin up and down in NM is splitting and we denote the splitting by helical state induced splitting (HSIS). Second, the band of spin down(up) opens a band gap in positive(negative) k_x region that makes the junction be a good spin filter and we

doi:10.6342/NTU201603562

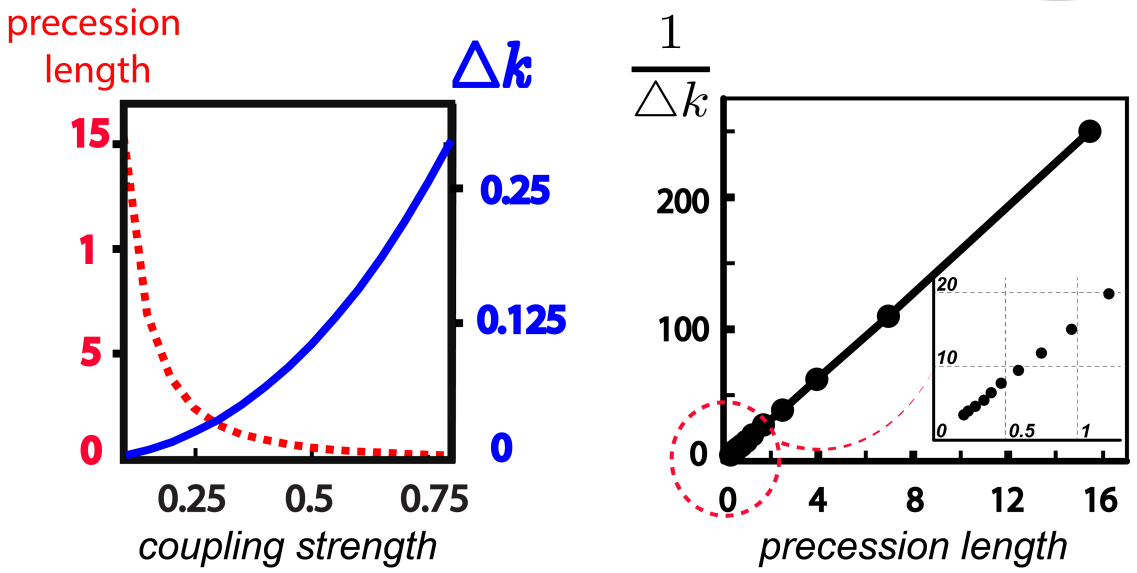

Figure 4.3: (left) Spin precession length (red dots and in the unit of 100 nm) and momentum difference of spin up and down $\triangle k$ (blue line and in the unit of 1/5 (1/nm)) as defined in Fig. 4.2 in terms of coupling strength (in the unit of hopping energy of s orbit in 2DTI) between 2DTI and normal metal (in the unit of hopping energy of s orbit of 2DTI). Spin precession length decreased, however, $\triangle k$ increased in response to increasing coupling strength. (right) The relation between precession length and inverse of $\triangle k$ shows linear behavior. This relation indicates that the precession behavior is the same as Rashba field. The subfigure is located in the strong coupling region that coupling strength larger than 0.35 and less than 0.8.

will give a more detailed discussion in the section (spin filter).

The spin up and down in NM split and the splitting difference of band structure in k_x space is defined by $\triangle k$ as Fig. 4.2. The splitting of spin is the same as a Rashba field, a SOC field that can be induced when an electron moves restrictedly in a two-dimensional plane with an electric field perpendicular to the two-dimensional plane.

The Rashba SOC splits the spin degree of freedom of a parabolic band with the

doi:10.6342/NTU201603562

relation $E = \frac{k^2\hbar^2}{2m} + \sigma\alpha k$[36], where σ is the denoted by 1 for the band of the spin up and -1 for spin down, α is the strength of Rashba SOC, and k is the momentum of electron. The relation gives us a band shifting along positive k_x for the spin up and along negative k_x for the spin down, and the shifting is a constant for all energy. This spin splitting induces the spin precession has demonstrated in spin FET[37] and unified analytical study[38] and non-equilibrium numerical study[39]. Above studies show the spin precession length is inversely proportional to the Rashba strength and satisfy the relation, the spin precession length multiplies the difference of momentum displacement is equal to $2 * \pi$.

Now let us go back to the HSIS of the junction, a spin precession is also induced by HSIS and the spin precession length and the momentum difference of spin up and down $\triangle k$ can be studied by varying the strength of coupling between NM and 2DTI as shown in Fig. 4.3. The momentum differences $\triangle k$ is increasing in response to increasing coupling between NM and 2DTI. The spin precession length is, however, decrease as increasing the coupling strength due to the increasing of $\triangle k$. The relation of the precession length and the momentum difference $\triangle k$ of HSIS is shown in Fig. 4.3 and the spin precession length multiply the difference of momentum $\triangle k$ is also equal to $2 * \pi$ that is the same as Rashba effect. However, there are two different features between the HSIS proposed in the paper and the Rashba splitting that induced by an electric field. First, the spin density is getting reduced as energy reach to the lowest point of band of spin down as shown in Fig. 4.4 due to the penetration of electron from NM to 2DTI as indicated in Fig. 4.1.Second, the difference in momentum $\triangle k$ is varied in terms of energy, especially in the low energy region in Fig. 4.4.

doi:10.6342/NTU201603562

4.3 Spin filter

A spin density dependence band dispersion can be calculated by the projection of the spin density of NM into the band dispersion, and the dispersion shows how the occupation of the electron spin in NM in terms of energy and momentum. The band of spin down to open a band gap because of the spin-momentum locking of helical edge state of 2DTI as what we have discussed and the spin density is increasing with increasing k_x. The band of spin up, instead of opening a band gap, only reduce the density as increasing k_x. In high momentum region, k_x is larger than 0.6, the density of spin up and down are closed and show HSIS. In the low momentum region, k_x is less than 0.6, the band gap of spin down can be used as a spin filter.

The spin filter suggests in the TI/NM junction of with three terminals. TI/NM junction connects to a lead composed of NM and a lead composed of 2DTI as showed in Fig. 4.4. LEAD1 connects to the main system composing a 6 layers of TI/NM junction,is also NM 2DTI junction. LEAD2(LEAD3), the same with of the NM(2DTI) in the junction, connects to the side of NM(2DTI) of the main system. The transmission from LEAD1 to LEAD2 or LEAD3 can be calculated by Landauer formalism in terms of energy as showed in Fig. 4.4. The transmission from LEAD1 to LEAD2 (NM lead), in the same energy region of band structure, shows the spin down will be blocked as the energy below energy point of band-gap opening, approximately $-5 * 0^{-3}eV$ as coupling strength equal to 0.4, and the transmission of spin up electron decrease due to the hybridization near the interface. As energy below energy point of band-gap opening, the polarization is positive polarized, however, the polarization becomes less as increasing the energy and reach to zero as the energy beyond approximately $5 * 10^{-}3eV$. In the energy region, the system can be a spin flitter by varying energy of the system.

doi:10.6342/NTU201603562

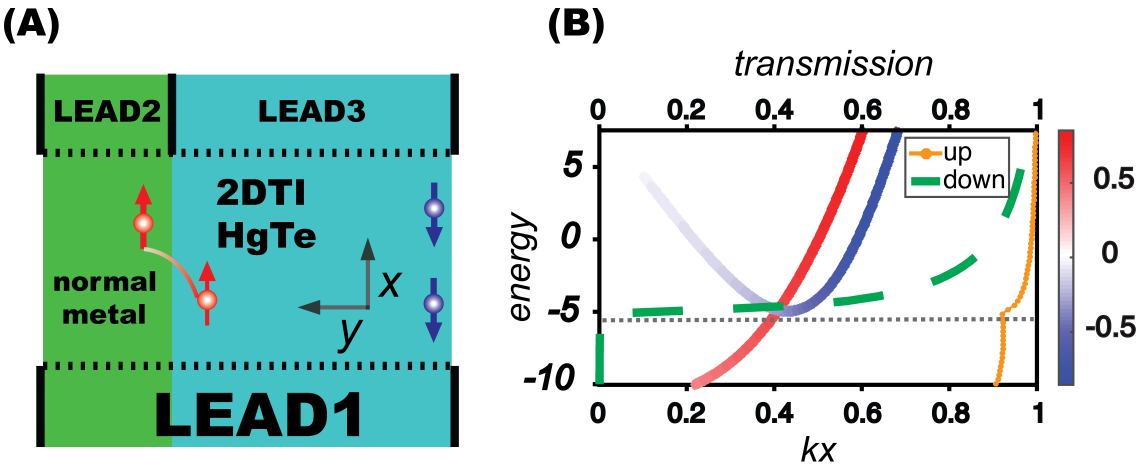

Figure 4.4: (A) Schematic of the spin filter which composed of a junction of 2DTI and normal metal (LEAD1 and the central part), a lead made of normal metal (LEAD2) and a lead made of 2DTI (LEAD3). The transport properties of the central part, sandwiched in the between of leads, are studied by calculating the transmission between leads. The transport direction is along the positive x that defined in the schematic. (B) The band structure and transmission between LEAD1 and LEAD2 (i.e. the transmission from the lead composed by the junction of the normal metal lead) in terms of energy (in the unit of meV).The coupling strength t set to be 0.4. The dashed green line is the transmission of spin down and orange line with a star is the transmission of spin down. The red and blue line with gradient is the band structure near Fermi energy (set to be near zero) with the momentum resolved state density of the normal metal. The dashed gray line indicates the band bottom of spin down the channel and the transmission also drop down to zero.

doi:10.6342/NTU201603562

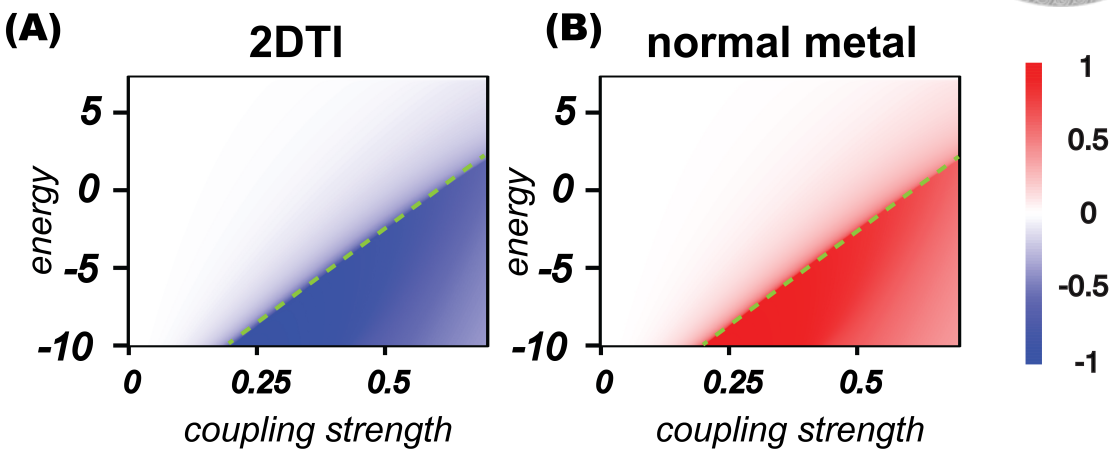

Figure 4.5: The polarization between LEAD1 and LEAD2 (A) and LEAD3 (B) in terms of coupling strength. They share the same color bar that indicted the currents flow into LEAD2 is spin up and into LEAD3 is spin down.

The transmission from LEAD1 to LEAD3 (2DTI lead) can be readily deduced as following. The transmission from LEAD1 to LEAD3 is equal to one for each spin channel if the NM and 2DTI are not connected. Once connected, the transmission of spin down electron in the device is maintained 1 due to it is on the other side of the interface of NM and 2DTI, so the channel would not interact with the NM due to absence finite size effect[40, 41, 42] The polarization, transmission of spin up electron minus transmission of spin up electron, is the same magnitude, however, the opposite direction of the polarization of NM lead. This is the result of conservation of time reversal symmetry, the unpolarized input electron will be unpolarized in the output, so the polarization from LEAD1 to LEAD2 plus the polarization from LEAD1 to LEAD3 should be zero. Under the above condition, we can derive the transmission of spin up electron is equivalent to 1− polarization of NM.

We have proposed a device that divides an unpolarized electron into two sorts of current with highly polarized in the opposite direction. In Fig. 4.5, we also

doi:10.6342/NTU201603562

demonstrate how does the polarization react the coupling strength between NM and 2DTI. The shade of the color is the polarization of the transmission, and the red and blue color indicate the spin up and down respectively. White is the region as the polarization of currents are zero, and become red (blue) is the energy lower down at each coupling strength as showed in Fig. 4.5. The points of change of color are in fact the energy point of band gap opening. The pattern of polarization shows the maximum of polarization occurs near the energy point of band gap opening. The increasing of coupling strength makes the polarization pattern sprayed owing to an enhancement of hybridization. Enhancement of hybridization result from the electron jump to 2DTI more easily and reduced the transport channel that reduces the transmission of electron spin up in NM under the energy point of band gap opening.

4.4 Spin flip in the normal metal 2DTI junction

In the section, we discuss the impact of the spin-flipping mechanism in the TI/NM junction. If the electron penetrates between the interface change the spin orientation, the spin of an electron is not preserved as the penetration. The effect may happen in the difference of fermi energy of NM and 2DTI. The difference of fermi energy induces an electric field in the y-direction. The electric field provides an effective magnetic field as electron move in the x-direction. The effective magnetic field is actually the conventional Rashba field as we have discussed. The field makes the spin flip due to spin precession and the flipping strength is positively associated with the electric field in the y-direction. We examine the flipping field by adopting the coupling matrix of the NM and 2DTI is not block diagonal. The off-diagonal term is a parameter ”flip” multiply the hopping energy of s orbit of 2DTI. By varying

doi:10.6342/NTU201603562

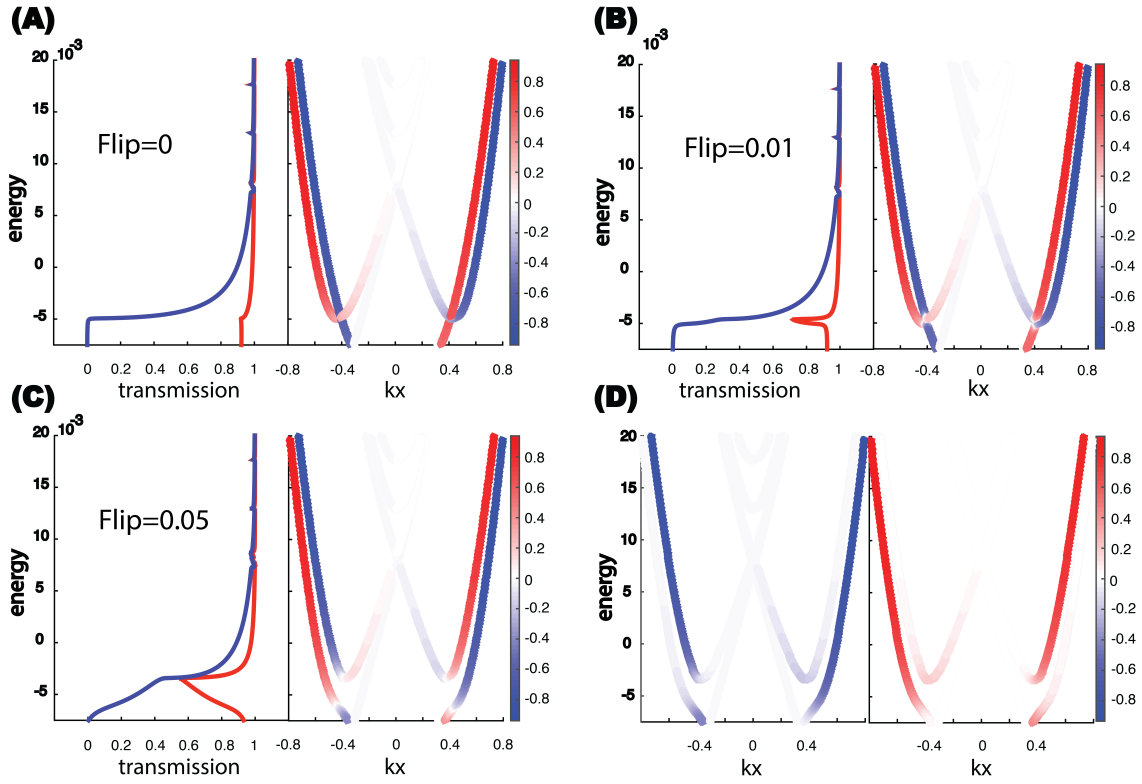

Figure 4.6: (A) (B) (C) The band structure and polarization between LEAD1 and LEAD2 (i.e. the transmission from the lead composed by the junction to the normal metal lead) in terms of energy. (D) The separate band of spin up and down for flip= 0.05, the strong flipping.

doi:10.6342/NTU201603562

the flipping parameter and calculate the band structure and transmission as showed in Fig. 4.6, we can know how does the mechanism influent the system. As flip equal to zero, the hybridization is what we have in Fig. 4.4. The spin up of normal metal hybridizes with the spin up of 2DTI. The splitting makes the HSIS and spin are highly polarized below the energy point of band gap opening. As the flip equal to 0.01, the flipping mechanism is small and the band structure does not vary much even if the presence of the flipping mechanism. However, we can find that there is a slight fluctuation of the transmission near the energy point of band gap opening. So the spin-flipping can be found near the energy point of band-gap opening, nevertheless, the small amount of flipping cannot make a major change of the transport. As the flip increase to 0.05, there is a great change of the band spectrum. The spin up of normal metal hybridizes to 2DTI so that the spin down of normal metal cannot open a band gap in the positive momentum region. The transmission also shows the non-vanished spin down channel. So we can conclude that the property of spin filter is destroyed if the flipping mechanism is robust enough. If we separate the band by spin as showed in Fig. 4.6, the spin up and down is all occupied in the energy region we are focusing. That indicates that the transmission for the spin up and down is all alive. Nevertheless, the HSIS would not be damaged by the flipping mechanism.

Methods

The Hamiltonian of an HgTe/CdTe quantum well yielding a 2DTI system can be given as follows: $\mathbf{H} = \begin{bmatrix} \mathbf{h} & 0 \\ 0 & \mathbf{h}^* \end{bmatrix}$. There are four base vectors of this Hamiltonian, where $|s,\uparrow\rangle$ and $|p_x + ip_y,\uparrow\rangle$ are the pseudo-spin-up states for the upper 2×2 block, and $|s,\downarrow\rangle$ and $|-p_x + ip_y,\downarrow\rangle$ are the pseudo-spin-down states for the lower block.

doi:10.6342/NTU201603562

The tight-binding Hamiltonian of **H** in real space can be represented by

$$\begin{aligned}
\mathbf{h_{2DTI}} &= \sum_{\mathbf{j}} \varphi_{\mathbf{j}}^{\dagger} \begin{bmatrix} E_s & 0 & 0 & 0 \\ 0 & E_p & 0 & 0 \\ 0 & 0 & E_s & 0 \\ 0 & 0 & 0 & E_p \end{bmatrix} \varphi_{\mathbf{j}} \\
&+ \sum_{\mathbf{j}} \varphi_{\mathbf{j}}^{\dagger} \begin{bmatrix} V_{ss} & V_{sp} & 0 & 0 \\ -V_{sp}^{*} & V_{pp} & 0 & 0 \\ 0 & 0 & V_{ss}^{*} & V_{sp}^{*} \\ 0 & 0 & -V_{sp} & V_{pp}^{*} \end{bmatrix} \varphi_{\mathbf{j}+\delta\mathbf{x}} \\
&+ \sum_{\mathbf{j}} \varphi_{\mathbf{j}}^{\dagger} \begin{bmatrix} V_{ss} & iV_{sp} & 0 & 0 \\ iV_{sp}^{*} & V_{pp} & 0 & 0 \\ 0 & 0 & V_{ss}^{*} & -iV_{sp}^{*} \\ 0 & 0 & -iV_{sp} & V_{pp}^{*} \end{bmatrix} \varphi_{\mathbf{j}+\delta\mathbf{y}}
\end{aligned}$$

, where $\varphi = \left[c_{s,\mathbf{j}}^{\uparrow}, c_{p,\mathbf{j}}^{\uparrow}, c_{s,\mathbf{j}}^{\downarrow}, c_{p,\mathbf{j}}^{\downarrow}\right]^{T}$ are spinors for the spin up and down components of the s and p orbits, the index **j** represents position in real space, and δx and δy are unit vectors of the lattice constant along the x and y directions. These parameters are to be fitted with experimental data.

The Hamiltonian of an NM can be given as follows:

$$\mathbf{h_{NM}} = \sum_{\mathbf{i}} \psi_{\mathbf{i}}^{\dagger} \begin{bmatrix} E_{up} & 0 \\ 0 & E_{down} \end{bmatrix} \psi_{\mathbf{i}} + \sum_{\mathbf{i}} \psi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{up} & 0 \\ 0 & V_{down} \end{bmatrix} \psi_{\mathbf{i}+\delta\mathbf{x}} + \sum_{\mathbf{i}} \psi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_{up} & 0 \\ 0 & V_{down} \end{bmatrix} \psi_{\mathbf{i}+\delta\mathbf{y}}$$

, where $\psi_{=} \left[a_{\mathbf{i}}^{\uparrow}, a_{\mathbf{i}}^{\downarrow},\right]^{T}$ are spinors for the spin up and down components, the index **i** represents position in real space, and δx and δy are unit vectors of the lattice constant along the x and y directions. Hopping energy for spin up and down are

doi:10.6342/NTU201603562

equal by setting $V_{up} = V_{down} = 2V_{ss} = -0.09584eV$. Onsite energy, are equal for spin up and down in normal metal, determine the band shifting and we set $E_{up} = E_{down} = 0.155eV$.

The coupling between NM and 2DTI can be expressed as

$$\mathbf{h_{couple}} = \sum_{i,j} \psi_{\mathbf{i}}^{\dagger} \begin{bmatrix} V_c & V_c & V_f & V_f \\ V_f & V_f & V_c & V_c \end{bmatrix} \varphi_{\mathbf{j}}$$

, where $V_c = t * V_{ss}$ and t is the coupling strength as what we have discussed in the article. The index **i** represents position in real space of NM and **j** represents position that connected to site **i** of 2DTI (site **i** connect to site **j** via V_c). The hopping energy from NM to s and p orbit of 2DTI are the same is our assumption. V_f =flip$*V_{ss}$ is a parameter determine the possibility for spin flipping mechanism in the last section. The wave density resolved band dispersion can be calculated by these tight-binding Hamiltonian by Bloch theorem.

Landauer formalism is suitable for a realistic non-periodical system[1]. The conductance is obtained by multiplying the conductance unit $\frac{e^2}{h}$ by $T_{n,m}(E) = trace[\Gamma_n(E)\mathbf{G^r}(E)\Gamma_m(E)\mathbf{G^{r\dagger}}(E)]$. The transmission from lead n to lead m, can be calculated from the Landauer-Buttiker formalism, the trace of the matrix of matrix multiplication of the retarded Green's function $\mathbf{G^r}(E)$ and broadening matrix $\Gamma_n(E) = i[\Sigma_n - \Sigma_n^{\dagger}]$, where Σ_n are self-energies of lead n. The retarded Green's function $\mathbf{G}^r(E)$ can be defined as $\mathbf{G}^r(E) = [(E + i\eta)\,\mathbf{I} - \mathbf{H}_{effect}(E)]^{-1}$, where $\mathbf{I}$ is the identity matrix, E is the Fermi energy, $i\eta$ is a small complex number, and $\mathbf{H}_{effect} = \mathbf{H}_{device} + \mathbf{H}_{leads}$. $\mathbf{H}_{device}$ is the spatial representation of the Hamiltonian of the H-shaped device, and $\mathbf{H}_{leads}$ is the Hamiltonian of the leads.

doi:10.6342/NTU201603562

4.5 Summary

HSIS can be discovered in NM/2DTI junction and can be varied by modulating the coupling strength of NM and 2DTI. The helical edge state of 2DTI hybridized with the quantum well state of normal metal is well explored in this article. The helical edge state near contacted edge penetrates to NM via the spin independent coupling that preserving time reversal symmetry and flows into the NM lead while the isolated edge state flows into 2DTI lead. Energy gap will be open for one spin thus reducing the spin current for NM channel and the preserving of time reversal symmetry makes the whole spin channels to be complementary. By modulating the energy for gate control, it is possible to convert the quantum spin hall system into a spin filter.

doi:10.6342/NTU201603562

Bibliography

[1] S. Sahoo, *et al.*, Nat. Phys. **1**, 99 (2005).

[2] R. Jansen, B.C. Min and S. P. Dash, Nat. Mat. **9**, 133 (2010).

[3] S. Matityahu, A. Aharony, O. Entin-Wohlman, and S. Tarucha, New J. Phys. **15**, 125017 (2013).

[4] Y. Puttisong, X. J. Wang, I. A. Buyanova, H. Carrére, F. Zhao *et al.*, Appl. Phys. Lett. **96**, 052104 (2010).

[5] R. C. Myers, A. C. Gossard, and D. D. Awschalom, Phys. Rev. B **69**, 161305 (2004).

[6] E. Wada, K. Watanabe, Y. Shirahata, M. Itoh, M. Yamaguchi and T. Taniyama, Appl. Phys. Lett. **96**, 102510 (2010).

[7] X. Li, O. E. Tereshchenko, S. Majee, G. Lampel, Y. Lassailly, D. Paget and J. Peretti, Appl. Phys. Lett. **105**, 052402 (2014).

[8] K. N. Altmann, *et al.*, J. Elec. Spec. Rel. Phen. **101**, 367 (1999).

[9] V. Renken, D. H. Yu, M. Donath, Surf. Sci. **601**, 5770 (2007).

[10] Florent Perez, Phys. Rev. B **79**, 045306 (2009).

doi:10.6342/NTU201603562

[11] J. S. Park, A. Quesada, Y. Meng, *et al.*, Phys. Rev. B 83, 113405 (2011).

[12] J. Sinova, D. Culcer, Q. Niu, N. A. Sinitsyn, T. Jungwirth, and A. H. MacDonald, Phys. Rev. Lett. **92**, 126603 (2004).

[13] D. Z.-Y. Ting and X. Cartoixà, Appl. Phys. Lett. **81**, 4198 (2002).

[14] J. Gersten, K. Kaasbjerg, and A. Nitzan, J. Chem. Phys. **139**, 114111 (2013).

[15] H. Bentmann, *et al.*, Europhys. Lett. **87**, 37003, (2009).

[16] H. Bentmann and F. Reinert, New J. Phys. **15**, 115011 (2013).

[17] L. V. Bondarenko *et al.*, Sci. Rep. **3**, 1826 (2013).

[18] A. M. Shikin, *et al.*, Phys. Sol. Sta., **52**, 1515 (2010).

[19] A. M. Shikin, *et al.*, New J. Phys. **15**, 095005 (2013).

[20] A. G. Rybkin, A. M. Shikin, V. K. Adamchuk, *et al.*, Phys. Rev. B, **82**, 233403 (2010).

[21] A. G. Rybkin, A. M. Shikin, D. Marchenko, A. Varykhalov, and O. Rade, Phys. Rev. B, **85**, 045425 (2012).

[22] A. G. Rybkin, *et al.*, Phys. Rev. B, **86**, 035117 (2012).

[23] K. He, T. Hirahara, T. Okuda, S. Hasegawa, *et al.*, Phys. Rev. Lett **101**, 107604 (2008).

[24] E. Frantzeskakis, S. Pons, H. Mirhosseini, *et al.*, Phys. Rev. Lett. **101**, 196805 (2008).

doi:10.6342/NTU201603562

[25] G. Bian, L. Zhang, Y. Liu, T. Miller, and T.-C. Chiang, Phys. Rev. Lett. **108**, 186403 (2012).

[26] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).

[27] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).

[28] X. L. Qi, Y. S. Wu, and S. C. Zhang, Phys. Rev. B **74**, 045125 (2006).

[29] G. M. Graf & M. Porta, Commun. Math. Phys., **324**, 851 (2013).

[30] M. Z. Hasan & C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[31] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[32] Hsieh, D., et al., 2009a, Nature (London) **460**, 1101.

[33] Markus König et al. Science **318**, 766 (2007).

[34] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[35] X.-L. Qi and S.-C. Zhang, Physics Today **63**, 33 (2010).

[36] Laurens W. Molenkamp, Georg Schmidt, and Gerrit E. W. Bauer, Phys. Rev. B **64**, 121202 (2001)

[37] Takaaki Koga, Junsaku Nitta, Hideaki Takayanagi, and Supriyo Datta, Phys. Rev. Lett. **88**88, 126601 (2002)

[38] Ming-Hao Liu, Ching-Ray Chang, and Son-Hsien Chen, Phys. Rev. B **71**, 153305 (2005)

[39] Kuo-Chin Chen, Yu-Hsin Su, Son-Hsien Chen, and Ching-Ray Chang, Journal of Applied Physics **115**, 17C305 (2014)

doi:10.6342/NTU201603562

[40] Bin Zhou, Hai-Zhou Lu, Rui-Lin Chu, Shun-Qing Shen, and Qian Niu, Phys. Rev. Lett. 101, 246807 (2008)

[41] H. H. Lee, J. Y. Liu, C. R. Chang, and S. Q. Shen, Phys. Rev. B **88**, 195149 (2013).

[42] Kuo-Chin Chen, Hsin-Han Lee, and Ching-Ray Chang, Phys. Rev. B **93**, 035405 (2016)

doi:10.6342/NTU201603562