Lecture notes on topological insulators

Ming-Che Chang

Department of Physics, National Taiwan Normal University, Taipei, Taiwan

(Dated: October 10, 2022)

CONTENTS

I. QUANTUM ANOMALOUS HALL EFFECT

A. Qi-Wu-Zhang model

The Hall effect in a metallic ferromagnet is called **anomalous Hall effect** (AHE). In the usual Hall effect, the trajectory of an electron is deflected in an external magnetic field due to the Lorentz force. However, in a ferromagnet that hosts the AHE, exchange interaction would generate an effective magnetic field that is much stronger than an applied magnetic field. Therefore, the band structure, and the electron motion, is strongly influenced by the magnetization of the ferromagnet.

The AHE can be caused by the Berry curvature of the modified Bloch band, which is an intrinsic effect. It can also be caused by scatterings with impurities (that lead to skew scatterings or side jumps). In order for the electron motion to couple with the direction of the majority spin (magnetization), spin-orbit coupling is required in AHE. For a comprehensive review on the AHE, see Nagaosa *et al.*, 2010.

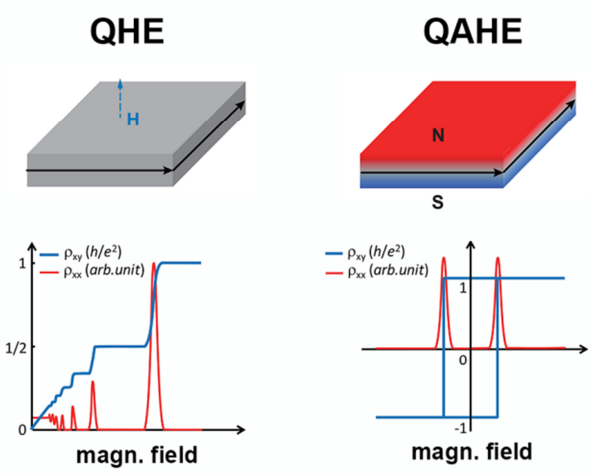

FIG. 1 A comparison between quantum Hall effect and quantum anomalous Hall effect.

Since the *intrinsic* AHE is due to the Berry curvature, we can expect quantized Hall conductivity from a filled band in a 2D system, *if* that band has a non-zero Chern number. To engineer a topological band, one can utilize the surface states of a topological insulator (Yu *et al.*, 2010). Dope the surface states with magnetic elements, then under proper condition, exchange field combined with spin-orbit coupling could generate quantized Hall conductivity. This proposal was first realized in Chang *et al.*, 2013. For subsequent experimental works, see Halperin, 2015, Chang and Li, 2016, or Zhao *et al.*, 2020. Fig. 1 shows a comparison of the Hall plateaus in QHE and QAHE.

In this section, we introduce a simple 2D model proposed by Qi, Wu and Zhang to illustrate the connection between Berry curvature and AHE (Qi *et al.*, 2006). The fermions are living on a 2D square lattice, and each lattice point is allowed to have two degrees of freedom (spin, or quasi-spin). The Qi-Wu-Zhang model has the Hamiltonian,

$$\begin{aligned}
&\mathsf{H}(\mathbf{k}) = \mathsf{H}_0 + \mathsf{H}_m + \mathsf{H}_{so}, \qquad (1.1)\\
&\mathsf{H}_0 = \varepsilon_0(\mathbf{k}) + \\
&t \begin{pmatrix} 2-\cos k_x a - \cos k_y a & 0 \\ 0 & -(2-\cos k_x a - \cos k_y a) \end{pmatrix},\\
&\mathsf{H}_m = m \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\\
&\mathsf{H}_{so} = \lambda \begin{pmatrix} 0 & \sin k_x a - i \sin k_y a \\ \sin k_x a + i \sin k_y a & 0 \end{pmatrix}.
\end{aligned}$$

H_m is a mass term (due to exchange interaction) that opens an energy gap, and H_{so} is a spin-orbit coupling.

In short,

$$\mathsf{H}(\mathbf{k}) = \varepsilon_0(\mathbf{k}) + \mathbf{h}(\mathbf{k}) \cdot \boldsymbol{\sigma}, \qquad (1.2)$$

where

$$\mathbf{h}(\mathbf{k}) = \left(\lambda \sin k_x a, \lambda \sin k_y a, m + t \sum_{j=1}^{2} (1 - \cos k_j a) \right). \qquad (1.3)$$

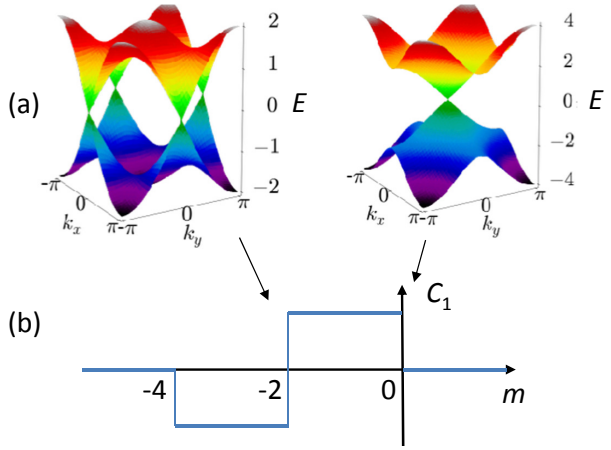

FIG. 2 (a) Dispersion of energy bands at $m = 0$ and $m = -2$. (b) The first Chern number is changed when the energy gap is closed at critical values of m.

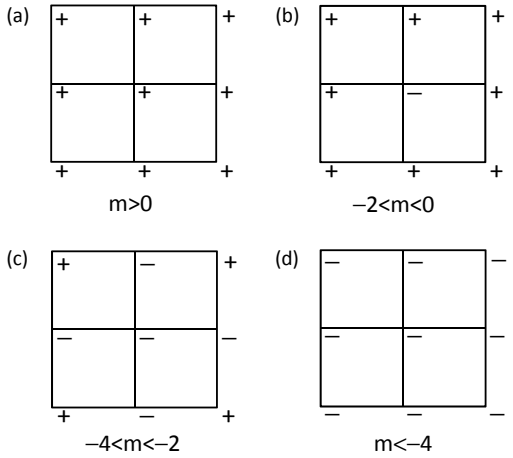

FIG. 3 Distribution of $\mathbf{h}(\mathbf{k})$ vectors in the BZ. Only the signs of $h_z(\mathbf{k})$ are shown.

The eigen-energies are,

$$\varepsilon_\pm(\mathbf{k}) = \varepsilon_0(\mathbf{k}) \pm |\mathbf{h}(\mathbf{k})|. \tag{1.4}$$

For simplification, we will set $t, a = 1$. It is not difficult to see that,

$$\begin{aligned} \mathbf{k}_0 &= 0 \to \varepsilon_\pm(\mathbf{k}_0) = \varepsilon_0 \pm m, && (1.5)\\ \mathbf{k}_0 &= (\pi,0),(0,\pi) \to \varepsilon_\pm(\mathbf{k}_0) = \varepsilon_0 \pm |m+2|, && (1.6)\\ \mathbf{k}_0 &= (\pi,\pi) \to \varepsilon_\pm(\mathbf{k}_0) = \varepsilon_0 \pm |m+4|. && (1.7) \end{aligned}$$

The energy gap closes at $m = 0, -2$, and -4. We will see that the topology of the Bloch states changes at these critical points, when the energy gap closes (Fig. 2).

The distribution of the $\mathbf{h}(\mathbf{k})$ vectors in the BZ changes when an energy gap closes, see Fig. 3. Depending on the value of m, there are four different regimes:
1) $m > 0 : h_z(\mathbf{k}) > 0$ over the whole BZ.
2) $-2 < m < 0 : h_z(\mathbf{k}) < 0$ near $\mathbf{k} = 0$.
3) $-4 < m < -2 : h_z(\mathbf{k}) > 0$ near $\mathbf{k} = (\pi, \pi)$ (and its equivalent points).
4) $m < -4 : H_z(\mathbf{k}) < 0$ over the whole BZ.

The $\mathbf{h}$ in Eq. (1.2) is the "magnetic field" for the quasi-spin. In case (1), the magnetic field only sweeps over the northern hemisphere when $\mathbf{k}$ scans over the BZ. According to the analysis in **??**, we need only one gauge $\mathbf{A}^N(\mathbf{k})$ for the whole BZ. Therefore, the topology is expected to be trivial and the Hall conductivity $\sigma_H = 0$.

In cases (2) and (3), h_z changes sign, so that $\mathbf{h}$ sweeps through the whole sphere, and two gauges, $\mathbf{A}^N$ and $\mathbf{A}^S$ are required to avoid the singularity. Therefore the topology is non-trivial and $\sigma_H \neq 0$.

Case (4) is similar to Case (1), but $\mathbf{h}$ sweeps over the southern hemisphere only. The topology is again trivial and $\sigma_H = 0$.

From such an analysis, we can see that both the magnetization m, which causes band inversion, and the spin-orbit coupling λ, are essential to the appearance of the QAHE. Without band inversion or SOC, one cannot have the nontrivial topology.

The simple picture presented above can be verified by actual calculation of σ_H. First, we show that (Nomura, 2013)

$$F_z^\pm(\mathbf{k}) = \mp \frac{1}{2h^3}\mathbf{h}\cdot\frac{\partial \mathbf{h}}{\partial k_x}\times\frac{\partial \mathbf{h}}{\partial k_y}. \tag{1.8}$$

Pf: The Berry connections in $\mathbf{k}$-space are,

$$\begin{aligned} A_\ell^\pm(\mathbf{k}) &= i\langle \mathbf{h},\pm|\frac{\partial}{\partial k_\ell}|\mathbf{h},\pm\rangle && (1.9)\\ &= \frac{\partial h_\alpha}{\partial k_\ell} i\langle \mathbf{h},\pm|\frac{\partial}{\partial h_\alpha}|\mathbf{h},\pm\rangle && (1.10)\\ &= \frac{\partial h_\alpha}{\partial k_\ell} a_\alpha^\pm(\mathbf{h}), && (1.11) \end{aligned}$$

where $\mathbf{a}^\pm$ are the Berry connections in $\mathbf{h}$-space. Therefore, the Berry curvatures in $\mathbf{k}$-space are,

$$\begin{aligned} F_z^\pm(\mathbf{k}) &= \frac{\partial A_y^\pm}{\partial k_x} - \frac{\partial A_x^\pm}{\partial k_y} && (1.12)\\ &= \frac{\partial}{\partial k_x}\left(\frac{\partial h_\beta}{\partial k_y}a_\beta^\pm\right) - \frac{\partial}{\partial k_y}\left(\frac{\partial h_\alpha}{\partial k_x}a_\alpha^\pm\right) && (1.13)\\ &= \frac{\partial h_\alpha}{\partial k_x}\frac{\partial h_\beta}{\partial k_y}\left(\frac{\partial a_\beta^\pm}{\partial h_\alpha} - \frac{\partial a_\alpha^\pm}{\partial h_\beta}\right) && (1.14)\\ &= \frac{\partial h_\alpha}{\partial k_x}\frac{\partial h_\beta}{\partial k_y}\varepsilon_{\alpha\beta\gamma}f_\gamma^\pm && (1.15)\\ &= \mp\frac{1}{2h^3}\mathbf{h}\cdot\frac{\partial \mathbf{h}}{\partial k_x}\times\frac{\partial \mathbf{h}}{\partial k_y}, && (1.16) \end{aligned}$$

in which $f_\gamma^\pm = \mp h_\gamma/2h^3$ are the Berry curvatures in $\mathbf{h}$-space. End of proof.

Suppose the lower band is completely filled, and the

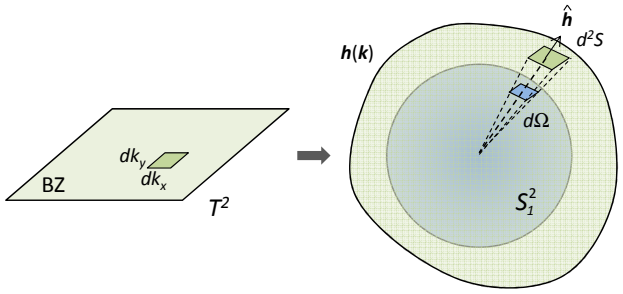

FIG. 4 Mapping a small area d^2k in the 2D BZ to a small area $d^2\mathbf{S}$ on the surface of $\mathbf{h}(\mathbf{k})$. Its solid angle $d\Omega$ is equal to the area of $\hat{h} \cdot d^2\mathbf{S}$ projected on a unit sphere S_1^2.

upper band is empty, then

$$\sigma_H = \frac{e^2}{h}\frac{1}{2\pi}\int_{BZ} d^2k F_z^-(\mathbf{k}) \tag{1.17}$$

$$= \frac{e^2}{h}\frac{1}{4\pi}\int_{BZ} d^2k \frac{1}{h^3}\mathbf{h}\cdot\frac{\partial \mathbf{h}}{\partial k_x}\times\frac{\partial \mathbf{h}}{\partial k_y}. \tag{1.18}$$

In the integrand, $\hat{\mathbf{h}}\cdot\left(\frac{\partial \mathbf{h}}{\partial k_x}dk_x\right)\times\left(\frac{\partial \mathbf{h}}{\partial k_y}dk_y\right)$ is actually the area $\hat{h}\cdot d^2\mathbf{S}$ on the $\mathbf{h}$-surface in Fig. 4. After being divided by h^2, it becomes the solid angle $d\Omega$ extended by that area. Since the BZ is a closed surface (a 2D torus, or T^2), under a continuous mapping, it would map to a closed surface in $\mathbf{h}$-space (see Fig. 5). The integral in Eq. (1.18) gives the total solid angle extended by that $\mathbf{h}$-surface. For a closed surface, it must be an integer multiple of 4π, thus

$$\sigma_H = w\frac{e^2}{h}, w \in Z. \tag{1.19}$$

The integer w, which is equal to the first Chern number C_1, is the number of times the $\mathbf{h}$-surface wraps over a unit sphere S_1^2. It characterizes the topology of the mapping (and the Bloch states) and is called the **winding number** (or wrapping number). We emphasize that w is an integer only if the base space is a closed surface (in this case, T^2), which requires the valence band to be completely filled (that is, an insulator).

In addition, one could apply an external magnetic field to the QAH insulator (aka **Chern insulator**). Then there will be orbital quantization and Landau levels in energy spectrum. Interested readers can consult p. 103 of Bernevig and Hughes, 2013 for more details.

B. Edge state in Qi-Wu-Zhang model

Topological materials (insulators) have an important property: their surface states are stable against perturbations. They can be destroyed only if the energy gap of the bulk bands closes so that the topology of the electronic states is trivialized. In general, *the interface between two materials with different topologies would have robust interface states.* A heuristic explanation is as follows: to go from one material to another, the spatial-dependent energy gap (in the sense of the Thomas-Fermi approximation) needs to close near the interface, otherwise it is simply impossible for the topology to change. This gapless region is where the surface states reside.

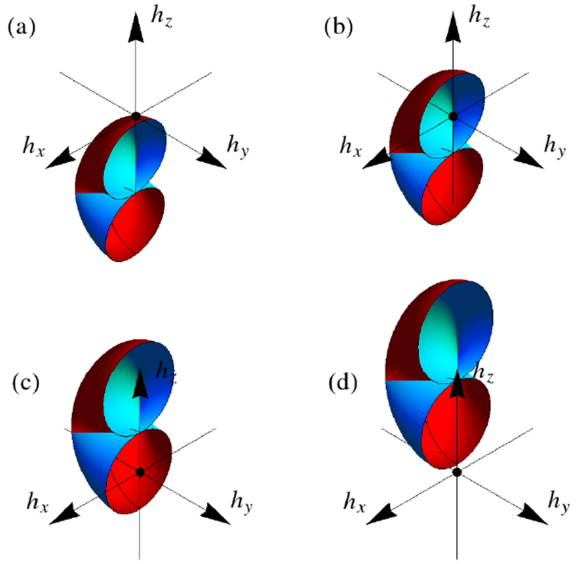

FIG. 5 The line segment $k_x \in [-\pi, \pi]$ in a BZ would map to a closed loop in $\mathbf{h}$-space. As one sweeps the line segment from $k_y = -\pi$ to $k_y = 0$, the loop in $\mathbf{h}$-space sweeps out a torus-like structure, as shown in the figures. Further advancement of k_y from 0 to π would map out another half of the torus (not shown for the sake of clarity). The winding number w depends on whether the origin is enclosed in the torus-like structure or not. In (a) and (d) (for $m < -4$ and $m > 0$), the origin is outside of the surface, so w=0. In (b) and (c) (for $m = -3$ and $m = -1$), the winding numbers are -1 and $+1$ respectively (not easily seen, though). [These figures are from Asbóth *et al.*, 2013]

Take the 2D QWZ model as an example. Divide the space to two parts where

$$m(x)\begin{cases} > 0 \text{ for } x > 0, \\ < 0 \text{ for } x < 0, \end{cases} \tag{1.20}$$

so that there is a 1D boundary along the y-axis (see Fig. 6(a)). For simplicity, consider only the small k limit,

$$\mathsf{H}(\mathbf{k}) = \varepsilon_0 + \begin{pmatrix} m & \lambda(k_x - ik_y) \\ \lambda(k_x + ik_y) & -m \end{pmatrix} + O(k^2). \tag{1.21}$$

The exact profile of $m(x)$ does not matter, as long as it is monotonic and smooth (compared to the electron wavelength). To solve for the surface states, one needs to re-quantize the Hamiltonian with the substitution $\mathbf{k} \to$

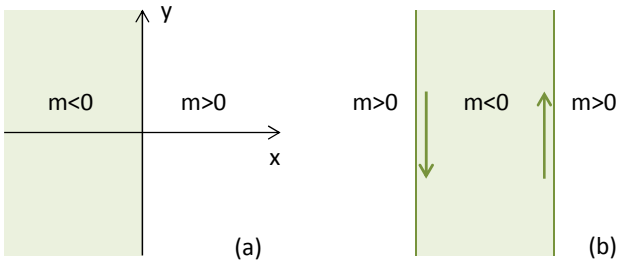

FIG. 6 (a) A topological phase occupies the left side of the space. (b) A finite sample with two boundaries. The edge states move in a definite direction along an edge.

$\frac{1}{i}\frac{\partial}{\partial \mathbf{r}}$, such that

$$\mathsf{H}(\mathbf{p}) = \varepsilon_0 + \begin{pmatrix} m(x) & \lambda\left(\frac{1}{i}\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) \\ \lambda\left(\frac{1}{i}\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) & -m(x) \end{pmatrix}. \quad (1.22)$$

The x-directions extend to infinity on both ends, and PBC is imposed along the y-direction.

We now solve the differential equation,

$$\mathsf{H}(\mathbf{p})\psi(x,y) = \varepsilon\psi(x,y). \quad (1.23)$$

Use the method of separation of variables and write

$$\psi(x,y) = \phi_1(x)\phi_2(y). \quad (1.24)$$

Since the y-direction is invariant under translation, a trivial solution is $\phi_2(y) = e^{ik_y y}$, a plane wave. Therefore, the equation for $\phi_1(x) = (f(x), g(x))^T$ is,

$$\begin{pmatrix} m(x) & \frac{\lambda}{i}\left(\frac{d}{dx} + k_y\right) \\ \frac{\lambda}{i}\left(\frac{d}{dx} - k_y\right) & -m(x) \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} = \varepsilon_e(k_y) \begin{pmatrix} f \\ g \end{pmatrix}. \quad (1.25)$$

Eliminate g from the coupled equations to get

$$-\lambda^2\left(\frac{d^2}{dx^2} - k_y^2\right) f = \left(\varepsilon_e^2 - m^2\right) f. \quad (1.26)$$

We have assumed that $m(x)$ varies slowly so that dm/dx can be neglected. It follows that the eigenvalue $\varepsilon_e(k_y) = \lambda k_y$, and the eigenstate

$$\phi_1(x) = e^{-\frac{1}{\lambda}\int_0^x dx' m(x')} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad (1.27)$$

which is localized near $x = 0$.

On the other hand, if

$$m(x) \begin{cases} > 0 \text{ for } x < 0, \\ < 0 \text{ for } x > 0, \end{cases} \quad (1.28)$$

then

$$\phi_1(x) = e^{\frac{1}{\lambda}\int_0^x dx' m(x')} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (1.29)$$

is a localized eigenstate with $\varepsilon_e(k_y) = -\lambda k_y$.

Therefore, in a sample with finite width (see Fig. 6(b)), the electrons on the right edge move with velocity $\frac{1}{\hbar}\frac{\partial \varepsilon_e}{\partial k_y} = \lambda/\hbar$; the ones on the left move with velocity $-\lambda/\hbar$. They are called **chiral edge states**. The two edges can be treated as independent only if the strip is wide enough (compared to the decay length of the edge state) so that the edge states on two sides do not couple with each other. In the small k_y limit, the energy dispersion $\varepsilon_e(k_y)$ of the edge state is linear in k_y. This is not so for larger k_y's, where numerical calculation is required.

Exercise

1. The effective two-band Hamiltonian for graphene with *gapped* Dirac points is

$$\mathsf{H}_0 = \hbar v_F\left(\pm k_x\tau_x + k_y\tau_y\right) + m v_F^2\tau_z, \text{ near } \pm K, \quad (1.30)$$

where v_F is the Fermi velocity, $\boldsymbol{\tau}$ is the quasi-spin for orbital degrees of freedom (i.e., conduction and valence bands), and $\pm K$ are the indices for two Dirac valleys.
(a) Find out the eigen-energies $\varepsilon_\pm(\mathbf{k})$ near the two Dirac points and plot their energy dispersions. How large is the energy gap?
(b) Find out the Berry curvatures $F_z^{\pm K}$ for the lower energy bands ε_- near $+K$ and $-K$.
(c) What would the Berry curvatures become if the Dirac points are gapless (i.e. $m = 0$)?
2. In Prob. 3 of Chap 2, we derived the effective Hamiltonian of an electron moving in a non-uniform magnetic field $\mathbf{B}(\mathbf{r},t) = B_0\hat{m}(\mathbf{r},t)$,

$$\left[\frac{1}{2m}\left(\mathbf{p} - \hbar\mathbf{A}_n\right)^2 + \hbar V_n + \varepsilon_n\right]\psi_n = i\hbar\frac{\partial\psi_n}{\partial t}, \quad (1.31)$$

where V_n and $\mathbf{A}_n$ can be found there.
(a) Show that an electron with spin up/down feels an effective electromagnetic field,

$$\tilde{E}_\alpha = \mp\frac{1}{2}\hat{m}\cdot\frac{\partial\hat{m}}{\partial r_\alpha}\times\frac{\partial\hat{m}}{\partial t}, \quad (1.32)$$

$$\tilde{B}_\gamma = \mp\frac{1}{4}\epsilon_{\alpha\beta\gamma}\hat{m}\cdot\frac{\partial\hat{m}}{\partial r_\alpha}\times\frac{\partial\hat{m}}{\partial r_\beta}. \quad (1.33)$$

As a result, an electron with velocity $\mathbf{v}$ is subject to a force $\hbar(\tilde{\mathbf{E}} + \mathbf{v}\times\tilde{\mathbf{B}})$.
(b) A magnetic skyrmion is a topological spin texture in magnetic materials. Because of the exchange interaction, the spin texture has an effective magnetic field that can be identified with the $\mathbf{B}(\mathbf{r},t)$-field above. Show that a skyrmion moving rigidly (without change of shape) with velocity $\mathbf{v}_s$ generates an effective electric field that is transverse to the direction of motion,

$$\tilde{\mathbf{E}} = -\mathbf{v}_s\times\tilde{\mathbf{B}}, \quad (1.34)$$

where $\tilde{\mathbf{B}}$ is the effective magnetic field of a static skyrmion.

REFERENCES

Asbóth, János K, László Oroszlány, and András Pályi (2013), "Topological insulators," Unpublished.

Bernevig, B Andrei, and Taylor L. Hughes (2013), *Topological Insulators and Topological Superconductors* (Princeton University Press).

Chang, Cui-Zu, and Mingda Li (2016), "Quantum anomalous hall effect in time-reversal-symmetry breaking topological insulators," Journal of Physics: Condensed Matter **28**, 123002.

Chang, Cui-Zu, Jinsong Zhang, Xiao Feng, Jie Shen, Zuocheng Zhang, Minghua Guo, Kang Li, Yunbo Ou, Pang Wei, Li-Li Wang, Zhong-Qing Ji, Yang Feng, Shuaihua Ji, Xi Chen, Jinfeng Jia, Xi Dai, Zhong Fang, Shou-Cheng Zhang, Ke He, Yayu Wang, Li Lu, Xu-Cun Ma, and Qi-Kun Xue (2013), "Experimental observation of the quantum anomalous hall effect in a magnetic topological insulator," Science **340** (6129), 167–170.

Halperin, Bertrand (2015), "Experimental advances in the quantum anomalous hall effect," Journal Club for Condensed Matter Physics: www.condmatjclub.org/?p=2620.

Nagaosa, Naoto, Jairo Sinova, Shigeki Onoda, A. H. MacDonald, and N. P. Ong (2010), "Anomalous hall effect," Rev. Mod. Phys. **82**, 1539–1592.

Nomura, K (2013), "Fundamental theory of topological insulator," Unpublished, written in Japanese.

Qi, Xiao-Liang, Yong-Shi Wu, and Shou-Cheng Zhang (2006), "Topological quantization of the spin hall effect in two-dimensional paramagnetic semiconductors," Phys. Rev. B **74**, 085308.

Yu, Rui, Wei Zhang, Hai-Jun Zhang, Shou-Cheng Zhang, Xi Dai, and Zhong Fang (2010), "Quantized anomalous hall effect in magnetic topological insulators," Science **329** (5987), 61–64.

Zhao, Yi-Fan, Ruoxi Zhang, Ruobing Mei, Ling-Jie Zhou, Hemian Yi, Ya-Qi Zhang, Jiabin Yu, Run Xiao, Ke Wang, Nitin Samarth, Moses H. W. Chan, Chao-Xing Liu, and Cui-Zu Chang (2020), "Tuning the chern number in quantum anomalous hall insulators," Nature **588** (7838), 419–423.