

凝聚态材料中的拓扑相与拓扑相变

——2016年诺贝尔物理学奖解读

戴 希^{1,2,†}

(1 中国科学院物理研究所 量子物质科学协同创新中心 北京 100190)

(2 中国科学院大学 物理学院 北京 100049)

2016-11-18收到

† email: daix@iphy.ac.cn

DOI: 10.7693/wl20161201

Topological phases and transitions in condensed matter systems

DAI Xi^{1,2,†}

(1 Collaborative Innovation Center of Quantum Matter, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China)

(2 College of Physics, University of Chinese Academy of Sciences, Beijing 100049, China)

摘 要 凝聚态物理中拓扑相变和拓扑物态的发现, 获得了2016年度诺贝尔物理学奖。文章系统介绍了凝聚态物理中拓扑性的起源, 并简要介绍了目前凝聚态物理中发现的主要几类拓扑态: 拓扑绝缘体、量子反常霍尔效应、拓扑晶体绝缘体和拓扑半金属。

关键词 拓扑量子态, 拓扑绝缘体, 拓扑相变

Abstract Two months ago, three physicists won the Nobel physics prize for their discovery of topological phases and transitions. In this paper, we review the origin of the concept of topology in condensed matter physics, then present a brief introduction to the main classes of topological states in solid-state materials, including topological insulators, the quantum anomalous Hall effect, topological crystalline insulators, and topological semimetals.

Keywords topological quantum state, topological insulator, topological phase transition

2016年的诺贝尔物理学奖授予了三位理论物理学家, 分别是美国华盛顿大学的David J. Thouless, 普林斯顿大学的F. Duncan M. Haldane和布朗大学的J. Michael Kosterlitz(图1), 以表彰他们在理论上提出了凝聚物质中的拓扑相变和拓扑相^[1]。作为在这一领域工作多年的研究者, 衷心地为几位前辈的获奖而感到高兴, 同时也深受鼓舞。许多朋友来信, 要求科普一下凝聚态物理学中的拓扑相, 我在这里就向大家简单介绍一下。很多人都觉得拓扑这个数学概念很深奥, 似乎很难跟生活中的物理现象发生联系, 其实不然, 生活中到处都有拓扑现象。比如说涡旋, 就是一种常见的

拓扑现象, 我们每天清晨在洗漱台见到, 在水流湍急的河边偶遇, 在每一个台风过境的夜晚, 我们怀着忐忑的心情等待。本次诺贝尔物理学奖得主Thouless先生和Kosterlitz先生发现的二维超流/超导体中的Kosterlitz—Thouless(KT)相变^[2, 3], 就跟其中的涡旋激发有着密切的关系。因此我们先来讲讲涡旋激发和KT相变。

1 超流/超导体中的KT相变

在介绍超流/超导体中的涡旋之前, 我先向大家介绍一点非常基本的场论, 一个二维标量场,



David J. Thouless F. Duncan M. Haldane J. Michael Kosterlitz

图1 2016年度诺贝尔物理学奖获得者

比如一个绷紧的鼓面，考虑它的横向振动，就是一个典型的二维标量场，它的运动可以由平面坐标 (x, y) 点处的振幅 $Z(x, y)$ 来描写，它的集体运动模式就是我们熟悉的声波振动，如图2所示。

与之类似，二维超流/超导薄膜的集体运动也可以由一个二维标量场来刻画，即平面坐标 (x, y) 点处的波函数相位 $\theta(x, y)$ ，由于超流/超导相变破缺了 $U(1)$ 对称性，当不存在集体激发时，刻画超流/超导体的波函数相位在空间各处是固定在一个值上的，可以假设为 $\theta(x, y)=0$ 。我们可以用图3中每一点的单位矢量与 x 轴的夹角来形象地演示相位场 $\theta(x, y)$ 。接下来讨论系统的集体激发，它可以被分成两类模式，第一类是 $\theta(x, y)$ 在0附近做小幅振动，这种模式跟鼓面上的声波振动是非常类似的，如图3(b)所示。

而另一种模式则是鼓面振动问题中没有的，也就是涡旋激发模式，为什么会存在这种不同呢？这是因为超导/超流问题和鼓面振动中的二维标量场是很不一样的，前者是紧致的(compact)而后者是非紧致的。紧致和非紧致是一个数学上的概念，它的意义很容易理解，通俗地说就是该标量的取值范围是在一个开放的空间(像鼓膜的横向振动)，还是在一个首尾相连的封闭空间(如相位)。相位空间是一个典型的紧致空间，它是一个定义在0到 2π 之间的角度，而零度则严格等价于 2π 。正因为相位场的紧致性，导致了在超流/超导体中另一类激发，拓扑型激发，也就是涡旋的出现。一个典型涡旋的空间构型如图3(c)所示，围绕着涡旋中心，相位场 $\theta(x, y)$ 连续地从0变化到 2π ，注意由于0和 2π 的等价性，相位场的连续

性并未被破坏。从图3(b)和(c)的对比中，可以看出涡旋和普通声波振动的不同，图3(b)中的声波激发构型可以通过对相位场施加连续形变来消除掉，而图3(c)中的涡旋在连续形变下只能移动位置而不能消除，把涡旋消掉的唯一办法就是通过连续形变，将一对手性相反(左旋和右旋)的涡旋挪到同一点上，让它们互相抵消。从拓扑学的角度看，具有单个涡旋激发的相位场构型跟没有涡旋的构型，是不能通过连续形变来互相转化的，属于不同的拓扑类，因此把涡旋激发称为是“拓扑型激发”。1972—1973年，Thouless和Kosterlitz两位先生通过漂亮的理论工作，推导出了在二维超流/超导体中涡旋运动的理论模型，他们发现两种手性的涡旋(左旋和右旋)可以看成是被限制在二维平面内的正负“电荷”，互相之间存在二维“库仑作用”。更有意思的是，他们还发现随着温度的变化，这一体系中还存在着一种特殊的相变。在相变温度以上，存在着自由运动的涡旋，从而破坏了超导/超流的长程相位有序；而在相变温度之下，不同手性的涡旋只能两两配对形成束缚态，这时候存在着波函数相位的“准长程序”，也就是说相位关联随距离按幂函数而不是指数函数衰减。这种相位无序到“准长程序”的特殊相变，是涡旋游离态到束缚态转变所导致的，并不伴随着对称性的破缺，因此不能用朗道相变理论来刻画，后来以两位发现者的姓名命名为Kosterlitz—Thouless相变。在实验上，Koster-

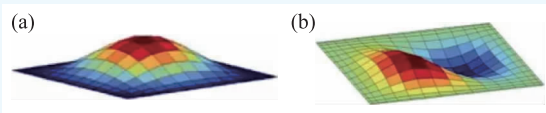


图2 鼓面的不同振动模式由一个标量场描述

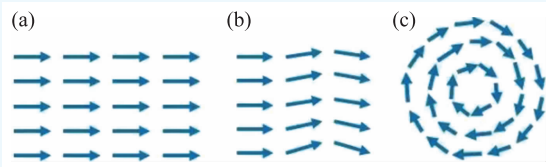


图3 超导/超流体波函数的相位构型 (a)基态构型；(b)普通型激发；(c)涡旋激发

litz—Thouless 相变在许多超导薄膜中被观测到，这种由拓扑元激发导致的相变，受到了广泛的关注。此外，前苏联科学家 Berezinskii 也完全独立地对理解这一概念作出了重大贡献^[4, 5]，因此这一相变也被称为是 Berezinskii—Kosterlitz—Thouless 相变^[2-5]。

2 永远的 TKNN: 动量空间中的拓扑不变量

上一节介绍了在超导/超流薄膜中的拓扑型元激发——涡旋，以及涡旋激发导致的 BKT 相变。从中大家可以看到拓扑型激发的几个特点：一是全局性，一个涡旋产生以后对距离很远的相位场都产生影响；二是对局部扰动的稳定性，局部扰动只能移动涡旋的空间位置而不能马上消灭它。同时，产生这种拓扑激发的本质原因也很清楚，就是相位空间是紧致的，说穿了就因为它是一个角度，只能在 0 到 2π 之间取值。类似的拓扑型激发在各种不同的凝聚态系统中都存在，比如最近研究很热的斯格明子，就是在铁磁系统中出现的拓扑型激发。

上面介绍的都是实空间的拓扑构型，拓扑激发，如涡旋和斯格明子，无论是描写它们的数学方程，还是显示其空间构型的图片，都非常美妙。但在凝聚态物理中，这只是刚刚开始，更美妙的数学结构存在于动量空间，也就是说在看似复杂的电子态波函数背后居然隐藏着异常丰富和美妙的拓扑结构，目前在这一领域进展非常快，像拓扑绝缘体、量子自旋霍尔效应、量子反常霍尔效应、外尔半金属等新材料、新现象层出不穷，而这一重要领域的开山之作，就是大名鼎鼎的 TKNN，也是这次获奖的工作之一^[6]。这篇经典文章由四位科学家联合署名，其中 T 就是这次获奖的 Thouless 先生，上文介绍的 KT 相变也是他的贡献，从中可以看出 Thouless 对当代凝聚态物理的巨大贡献，这次获奖的确是实至名归。TKNN 要研究的问题，其实属于电子态波函数的拓扑分类，这是一个很大的领域，TKNN 关注的

是其中最简单的一类。首先，这类波函数是无相互作用的，可以由一个 Slater 行列式来描写；其次这类波函数描述一个能带绝缘体。由于无相互作用，多电子波函数可以由一批占据单电子波函数以形成 Fock 态来表示，又由于是能带绝缘体，所有被占据的单电子态将铺满整个布里渊区。这样对多电子态的分类，可以简化为单电子波函数在整个布里渊区内是否具有拓扑结构的问题。目前研究正热的拓扑绝缘体、拓扑半金属等，都是在针对单电子态而言的，本质上都是 TKNN 的工作在广度和深度上的推广。下面就具体介绍 TKNN 指数。

我们来考虑一个二维无相互作用绝缘体的波函数，对布里渊区内的每一个 k 点，被占据的单电子波函数构成了一个希尔伯特空间，而这些单电子波函数 $\psi_k(r)$ 连同承载它们的二维布里渊区就构成了一个叫做纤维丛(fiber bundle)的数学结构。所谓的纤维丛，就是定义在一个几何体(数学上叫底流形)的函数(数学上叫纤维)。比如图 4 的这把梳子，就可以看成是一个纤维丛，底流形是一个柱面，而纤维则是上面的一根根梳齿(线段)。

对于无相互作用绝缘体系统来说，底流形就是我们的二维布里渊区，(注意它是一个轮胎面或者说是当肯甜甜圈，这一点非常重要，后面会详细阐述)，而单电子波函数则是“长”在上面的纤维。对于纤维丛的拓扑分类问题，数学大师陈省身先生曾做过系统的研究，按照微分几何理论，这样的纤维丛可以分成不同的类，每一个拓扑上等



图4 生活中处处可见“纤维丛”

价的类都具有相同的“陈数”，TKNN在二维绝缘体系统中找到的不变量，正是数学上的“陈数”。

要从物理上理解“陈数”这个比较抽象的数学概念其实并不难，首先说明一下为什么二维布里渊区是一个轮胎面，这是因为布里渊区内的动量 \mathbf{k} ，跟上节中的角度一样，是定义在一个单值区间 $[-\pi, \pi]$ 之内的，换句话说也是紧致的，因此在图 5(a) 中布里渊区的上边界跟下边界，左边界跟右边界都是等价的，两边连起来一个平面就成了轮胎面，如图 5(b) 所示。理解了底流形布里渊区的拓扑特性之后，再来看定义在其上的“纤维”——波函数。固体中的电子态，学过一点固体物理的同学们都知道，就是一条条的能带，如图 6 就是一个典型的绝缘体的能带色散关系图，上面的空态是导带，下面的占据态是价带。在 TKNN 以前，人们只对电子态的本征能量

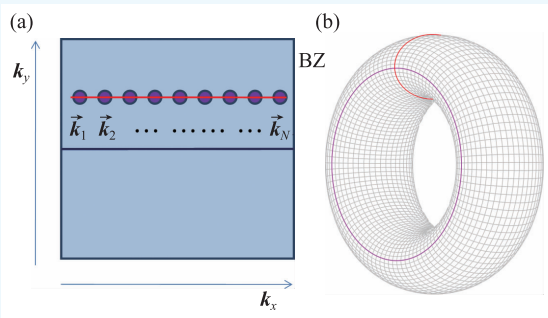


图5 二维布里渊区示意图 (a)沿 k_x, k_y 方向周期性延伸；(b)轮胎面

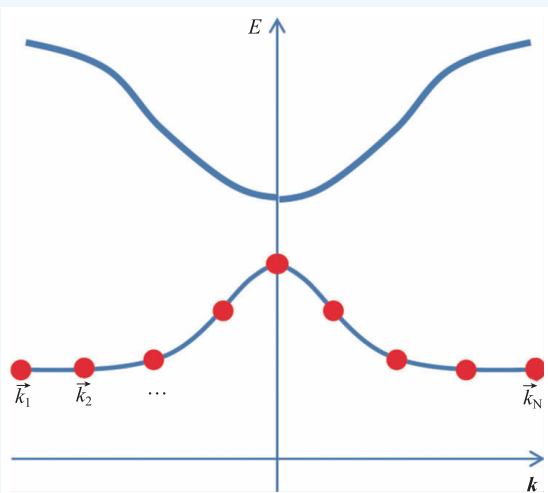


图6 最简单的绝缘体能带图

或色散关系感兴趣，因为它可以决定固体的许多性质，而对具体波函数的性质则不太关心。随着 TKNN 的工作和对 Berry 相因子研究的进一步深入，人们越来越认识到波函数内部结构的重要性。

其实，跟上文介绍的涡旋激发类似，动量空间中波函数的拓扑结构还是隐藏在它的相位当中。但是一个重要的不同点是，现在我们的体系具有 $U(1)$ 规范不变性，也就是说每个波函数 $\psi_k(r)$ 的相位是不定的，可以通过规范变换 $\psi'_k(r) = \psi_k(r)e^{i\theta(k)}$ 来进行改变，而任何有意义的物理可观测量都必须在这样的变换下保持不变。因此，我们要解决的问题就转化为是否能在看似任意变化的波函数相位结构中找到规范不变的量来，TKNN 给出的答案就是陈数。

我们可以通过一个叫做“Wilson loop”的方法^[7, 8]，很直观地向大家介绍陈数这个看似深奥的数学概念。首先，在看似随意的波函数相位中找一个规范不变的量，可以把二维布里渊区按照不同的 k_y 切成一条条的横线，如图 5(a) 所示，并且在每一条线上按照一定的间距分成 N 个点，于是把相邻 k 点的占据态波函数求内积再乘起来就是规范不变的，如下式，

$$Ae^{i\theta(k_y)} = \langle \psi_{k_1} | \psi_{k_2} \rangle \langle \psi_{k_2} | \psi_{k_3} \rangle \cdots \langle \psi_{k_{N-1}} | \psi_{k_N} \rangle \langle \psi_{k_N} | \psi_{k_1} \rangle .$$

需要注意的是，这个连乘是从 k_1 开始，最后回到 k_1 ，因为刚才讲过，布里渊区其实是轮胎面，所以 k_N 和 k_1 也是相邻的。很容易证明这个量在规范变换下是不变的，并且当 N 趋于无穷大时，这个复数的模 A 趋于 1，而只留下一个相位 $\theta(k_y)$ ，这个量是 k_y 的函数，因为对每一个固定 k_y 的环我们都可以做同样的计算。然后我们来看相位 $\theta(k_y)$ 随着 k_y 是如何演化的。假设 k_y 从 $-\pi$ 变化到 π ，因为 $-\pi$ 和 π 其实是等价的，相位 $\theta(k_y)$ 必须回到初始值或者变化 2π 的整数倍。

于是随着 k_y 变化一圈，相位角 $\theta(k_y)$ 必须在布里渊区的轮胎面上绕整数 C 圈，如图 7 所示，这个整数 C 就是 TKNN 发现的陈数！大家可以感受一下，绝缘体系统在 \mathbf{k} 空间的相位结构，居然跟一个涡旋系统有着如此的相似性，可见物理体

系背后的数学结构是多么的美妙。

更有意思的是，陈数并不单单只是数学游戏，对于二维绝缘体系统来说，它有明确的物理意义，就是霍尔电导的量子化，不同的陈数对应于不同的量子霍尔效应台阶。1988年本年度诺贝尔物理学奖的另一位获奖者F. D. M. Haldane教授第一次提出，无需外磁场而依靠特殊的能带波函数结构也可以产生非平庸的陈数，从而实现量子反常霍尔效应^[9]。2010年戴希、方忠、张首晟等人第一次提出在磁性掺杂的拓扑绝缘体薄膜 $\text{Bi}_2\text{Se}_3/\text{Bi}_2\text{Te}_3$ 中可能存在陈数为1的量子反常霍尔效应态^[10]，3年后薛其坤领衔的实验团队第一次在上述体系中观测到了量子反常霍尔效应态^[11]，引起了很大的关注，这时离TKNN最初的文章已经过去了31年，离F. D. M. Haldane 1988年的理论工作也过去了25年，可见几位获奖前辈在科学上的超人远见。

3 从TKNN到 Z_2 拓扑绝缘体

TKNN的经典论文发现了TKNN不变量即陈数，利用这一整数，可以对所有的二维绝缘体态进行分类。但是在TKNN那个年代，80年代，人们只是在二维电子气在强磁场下形成的量子霍尔效应系统^[12]——这一人造体系中实现了具有非零陈数的物质态。那么能不能在天然晶体材料中发现类似的拓扑非平庸物质态呢？这是一个很自然的问题，而为了回答这个问题，科学家们又付出了20多年的努力。想要在天然晶体中找到与量子霍尔效应体系类似的拓扑材料，首先要确定搜索的方向。如果这个新的拓扑材料是用陈数来刻画的，那么它必须满足两个条件，(1)必须是二维材料；(2)必须破坏时间反演对称。第一个条件很显然，因为根据我们前面的介绍，陈数本身就是定义在二维体系中的。第二个条件需要进一步解释一下，事实上对凝聚态物质中时间反演对称的思考和研究，直接导致了拓扑绝缘体的发

现。那么时间反演对称是怎么回事呢？我们说体系具有某种对称，是指它在相应的对称操作下保持不变，比如镜面对称，就是指系统在镜面反射之后还保持原来的模样。对于现代物理学来说，对称性是个太重要的概念，在凝聚态体系中各种空间对称已经被研究得很透彻，但对时间反演对称却尚有很大的研究空间。简单地说，时间反演就是时间箭头反向，向前运动的粒子掉头往后，原来往后运动的粒子转而向前。除了粒子平动自由度以外，内禀转动自由度也要反向，原来顺时针转的变为逆时针，逆时针转的变为顺时针，导致的后果就是粒子自旋方向的翻转。数学上很容易证明，就像单个粒子的平动和自旋运动，在时间反演操作下二维绝缘体的陈数必须反号，那么对于任何晶体材料，只要具有时间反演对称，其陈数就必须为零。换言之，具有非零陈数的晶体材料，只能在破缺时间反演对称，也就是具有自发磁性的系统中去找，而绝大多数天然晶体材料是没有磁性的，难道在这些材料中就不会有拓扑非平庸的电子结构吗？能否找到陈数以外的拓扑不变量来刻画具有时间反演对称的绝缘体系统呢？答案是肯定的，近年来这方面的重大突破就是拓扑绝缘体理论。根据这一理论，可以把具有时间反演对称的绝缘体系统，分为拓扑绝缘体和普通绝缘体两类，而用来刻画其拓扑特性的不变量就叫做 Z_2 拓扑不变量。

在介绍 Z_2 拓扑不变量的数学形式之前，让我们首先来回顾一下历史。在保持时间反演不变的体系中寻找拓扑物态的早期尝试，是基于一种非常朴素的想法，既然时间反演操作下绝缘体系统

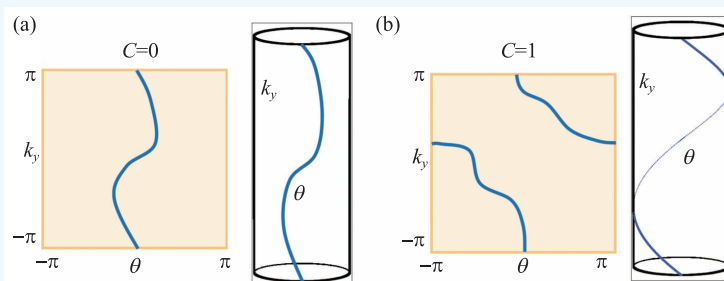


图7 陈数分别为0(a)和1(b)时相位角 $\theta(k_y)$ 的演化示意图

的陈数会变号，那么假设有这样一类体系，它们的电子态由两个互为时间反演态，并且相互之间无耦合的子系统组成，我们不妨分别记为A和B，由于时间反演对称，两个子系统的总陈数为零，但它们各自的陈数可以不等于零，而是符号相反的两个整数，比如(1, -1)或者(2, -2)，这样我们就可以得到一个陈数分类的简单推广，由于对固体电子态而言，最自然的两套子系统，就是电子的两个自旋指向，向上或者向下，于是这种“半个系统”的陈数就被叫做“spin Chern number”，自旋陈数，比如刚才提到的(1, -1)系统，其自旋陈数就等于1。并且当自旋陈数不等于零时，系统就相应地具有自旋霍尔效应。2005年宾州大学的Kane/Mele^[13, 14]和斯坦福大学的Bernevig/张首晟^[15]分别基于石墨烯和特殊应力场下的半导体量子阱，几乎同时提出了这种可以用自旋陈数来描述的自旋霍尔效应体系。然而用自旋陈数来描述电子态的拓扑特性有一个非常重要的前提条件，就是上下自旋的电子态是完全相互独立的，不存在耦合效应，而这一条件在实际材料中是很难实现的，因为自旋轨道耦合效应是普遍存在的。然而他们很快发现，在进一步考虑了自旋轨道耦合效应以后，原先自旋陈数等于1, 3, 5等奇数的系统彼此在拓扑上等价，同时原先自旋陈数等于0, 2, 4等偶数的系统也在拓扑上等价，当然这里还必须满足一个条件，就是加上自旋轨道耦合的过程不能关闭绝缘体的能隙，这样才能保持拓扑学要求的“连续形变”。那么，具有时间反演不变的二维绝缘体就可以分为奇偶两类，数学上叫做 Z_2 分类，其中奇数类具有被时间反演对称保护的狄拉克型边缘态，而偶数类则没有这种受保护的边缘态。在2005年Kane/Mele的文章中^[13]，他们还推导出了这个 Z_2 拓扑不变量具体表达式，从而把TKNN开创的研究领域推进到了新的高度。2006年Bernevig, Hughes和张首晟提出了实验上有可能真正实现的二维拓扑绝缘体系统——HgTe/CdTe量子阱^[16]，次年德国维尔兹堡大学的Molenkamp小组在该体系中第一次观测到了量子自旋霍尔效应的迹象^[17]。此后，傅亮

和Kane等人把拓扑绝缘体的概念推广到了三维体系^[18-20]，2009年我和方忠领导的小组与张首晟小组合作，通过计算提出了目前影响最大的三维拓扑绝缘体材料， Bi_2Se_3 家族^[21]，也对这个领域的研究做出了自己的贡献。2011年我和祁晓亮、Bernevig等人，利用Wilson loop方法把Kane/Mele提出的 Z_2 拓扑不变量写成了更直观的形式，并且可以统一描述系统的陈数和 Z_2 不变量^[7]，下面向大家简要地介绍一下。

跟计算陈数的时候一样，首先我们也还是要把二维布里渊区划分成一条条的平行线，如图5(a)所示。与破坏了时间反演对称的绝缘体不同的是，当具有时间反演对称时，最简单的绝缘体系统将至少具有两条占据能带，这是电子自旋自由度的体现。于是当我们仿照上文的做法构建Wilson loop时，我们发现相邻 k 点的波函数内积不再是一个复数而是一个 2×2 的矩阵，同样我们可以把这些矩阵按照次序乘起来，

$$\hat{W} = \hat{A}_{12} \hat{A}_{23} \hat{A}_{34} \cdots \hat{A}_{N-1, N} \hat{A}_{N1},$$

其中矩阵元 $\hat{A}_{kk', mm} = \langle \phi_{k, n} | \phi_{k', m} \rangle$ 。注意通过这样的Wilson loop得到的矩阵 \hat{W} 不再只是 $U(1)$ 规范不变，而是 $U(2)$ 规范不变，它可以分解成 $U(1)$ 部分和 $SU(2)$ 部分 $\hat{W} = z \hat{W}_0$ ，其中 $U(1)$ 部分 z 就是矩阵 \hat{W} 的行列式，是一个复数，而这个复数的相位角随 k_y 的演化则得到系统的陈数，请参看上节。剩下的 \hat{W}_0 是一个行列式为1的幺正矩阵， Z_2 拓扑不变量就藏在这个矩阵的本征值里。 2×2 幺正矩阵 \hat{W}_0 的本征值是两个模为1且互相共轭的复数对，我们可以提取出它们的相位，记为 θ_1, θ_2 ，(注意，又是相位!)并且考察它们随着 k_y 的演化。由于系统具有时间反演对称，这两个相位的演化必须满足下面两个条件：(1)只有一半的演化过程是独立的，如 $k_y=0$ 到 π ，而另一半($-\pi$ 到 0)只是前一半的简单重复；(2)在 $k_y=0$ 和 π 处，相应的Wilson loop在时间反演操作下是不变的，因此上面提到的两个相位 θ_1, θ_2 必须相等，且满足 $\theta_1 = -\theta_2$ (共轭性的要求)。所以在 $k_y=0$ 到 π 处，它们或者都等于0，或者都等于 π (π 等价于 $-\pi$ ，又是相位紧致性的体现!)。于是当我们考察 $k_y=0$ 到 π 的演化

过程时, 相位 θ_1, θ_2 的演化曲线有两种拓扑上不等价的构型, 分别如图 8(a)和(b)所示。

其中图 8(a)的两条演化线从 0 开始, 最后一起回到 0, 而在图 8(b)中, 则是一起从 0 开始, 最后分别演化成 π 和 $-\pi$ 。由于二维布里渊区的紧致性, 上述演化实际上是在一个开放的柱面上进行的, 理解了这一点就不难弄清楚图 8(a)和(b)之间的本质不同, 在不改变两端取值(时间反演对称的要求)的情况下, 通过在柱面上的连续变形, 永远不可能将图 8(a)的演化线变化成图 8(b)。图 8(b)的情况实际上是两条演化线合起来在柱面上绕了一圈, 所以它跟没有绕圈的图 8(a)在拓扑上是不等价的。下一个问题是, 这样不等价的构型一共有多少个呢? 能不能有绕两圈, 三圈的情况呢? 拓扑绝缘体理论给出的回答是, 在没有额外对称性的情况下, 绕两圈等价于零圈, 而三圈等价于一圈。这一点可以通过图 9(a)和(b)看出来, 如果两条演化线绕柱面两圈(如图 9(a)), 它们必须在中间交叉一次, 假设除了时间反演以外

没有其他晶体对称, 两条演化线之间总可以存在有限的耦合, 从而把交叉点劈裂开, 如图 9(a)所示, 于是两圈就等价于零圈了。同理, 如图 9(b)所示, 三圈也等价于一圈。

综上所述, 我们可以得出结论, 在半个布里渊区的相位角演化只有两种拓扑上不等价的构型, 分别环绕半个布里渊区柱面奇数和偶数圈, 前者对应拓扑绝缘体而后者则对应普通绝缘体。与陈数分类不同, 这种分类在数学上是奇偶类而非整数类, 因此称为 Z_2 拓扑不变量。

具有时间反演不变的二维拓扑绝缘体可以按照 Z_2 不变量分类, 从本质上讲, 起源于对绝缘体电子态“连续

变形”所施加的限制。在不施加任何限制的情况下, 二维绝缘体只能按照陈数来分类, 而如果限制电子态连续变形过程不能破坏时间反演对称, 则二维绝缘体又可以分为 Z_2 不变量为奇或偶两类, 虽然它们的陈数都等于零。按照这个思路, 考虑越来越多的对称性, 如空间反演、镜面反射、旋转对称等, 系统电子态的拓扑构型也将越来越丰富, 受此启发科学家们此后又提出了拓扑晶体绝缘体等概念。2012 年陈谐、顾正澄、刘正鑫和文小刚总结了前面的工作, 用群论方法对固体中受对称性保护的各种拓扑态进行了分类研究, 提出了 SPT 的概念(symmetry protected topological phase), 又把这方面的研究推进了一大步^[22]。

4 百花齐放的拓扑态: 拓扑晶体绝缘体和拓扑半金属

TKNN 指数可以用来刻画任意二维绝缘体系

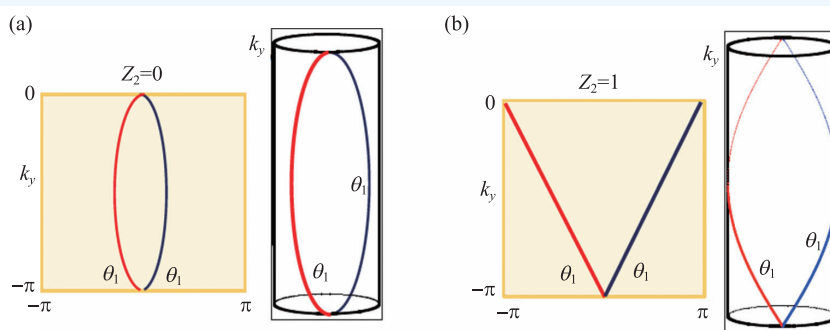


图 8 时间反演对称能带相位演化的两种不同拓扑构型 (a) $Z_2=0$; (b) $Z_2=1$

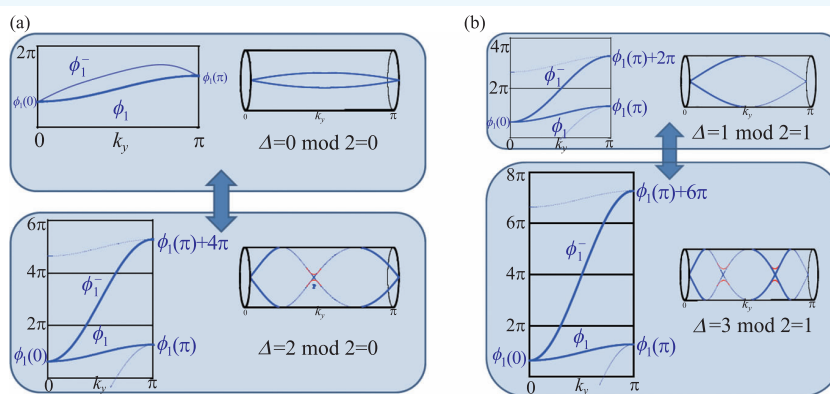


图 9 能带相位演化的两种不同拓扑构型 (a) $Z_2=0$; (b) $Z_2=1$

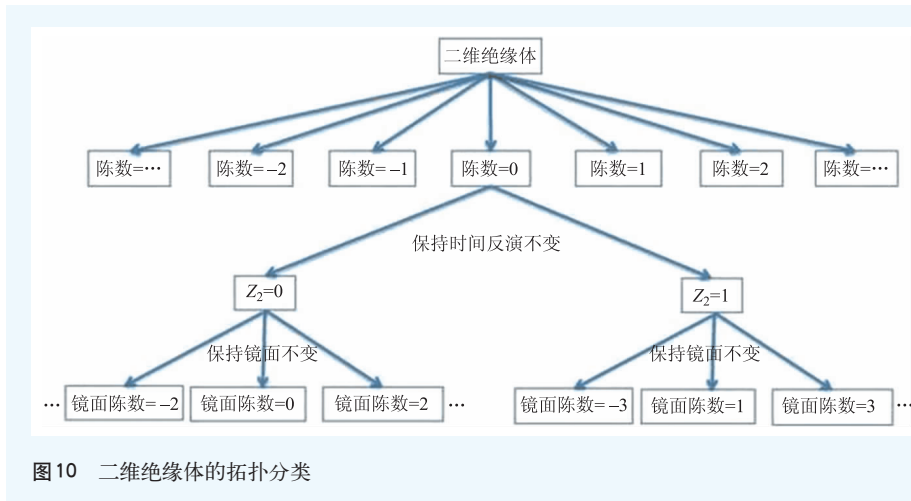


图10 二维绝缘体的拓扑分类

统的拓扑特性， Z_2 指数可以用来刻画具有时间反演对称的二维绝缘体的拓扑特性。这里面蕴含了深刻的对称性和拓扑分类之间的关系，如果不对体系的对称性做任何限制，那么二维绝缘体只能按照TKNN指数分类，也就是每一个整数 C 代表一类，其中所有具有时间反演对称的二维绝缘体都被统一分在TKNN指数 $C=0$ 的类里面。按照拓扑学的基本原理，被分在同一类中的个体(这里指体系的波函数)都在连续变形下彼此等价。然而如果利用对称性对系统的连续变形做进一步的限制，比如要求连续变形过程限制在满足时间反演不变的子空间里，则有些原本彼此等价个体，由于连续变形的受限而变得不再等价，于是这些TKNN指数等于零的二维绝缘体又可以进一步按照 Z_2 指数分为奇偶两类。同理，我们还可以进一步加入各种晶体对称性，对连续变形做更高的对称性要求，从而获得更加丰富的拓扑晶体绝缘体，如镜面陈数(mirror chen number)绝缘体^[23]和Hour Glass绝缘体等^[24, 25]。从中我们可以看到，体系受到的对称性限制越多，则拓扑分类越为丰富，如图10所示。沿着这个方向，用拓扑学的观点来重新审视各类材料的电子结构，是目前发展非常迅速的领域，结出了累累硕果。

前面介绍的是绝缘体系统的拓扑分类问题，对于绝缘体来说，其布里渊区被整数条能带所占据，从微分几何的观点看，是一个定义在轮胎面上的纤维丛，所有的拓扑分类研究都缘此而来。

那么能不能把拓扑分类研究推广到金属体系呢？简单的推广马上就碰到问题，在金属中波函数在布里渊区的分布是不连续的，在布里渊区的某些区域占据着 N 条能带，而在另一些区域则可能占据着 $N+1$ 或 $N-1$ 条能带，这些能带上出现的“断层”，则构成金属中最重要的元素

——费米面。显然对于金属体系而言，我们不能把整个布里渊区及其上的占据波函数当作分类对象。那么对于金属来说什么是合适的分类对象呢？费米面及其上占据的波函数，似乎是一个比较合理的选择。目前针对金属体系的普遍拓扑分类问题，还没有一个明确的答案，但对于一类特殊的金属体系——半金属，我们已经取得了突破，这就是大家都听说过的拓扑半金属^[26]，它可以分为外尔半金属^[27-33]、狄拉克半金属^[33-37]、节线半金属^[38-41]和多重点半金属^[42-45]等等，其中外尔半金属是最基本的，本文的介绍也主要围绕它展开。

所谓外尔半金属，就是两条不简并的能带在三维空间相交在一个 k 点上，这个 k 点就是外尔点。可以从多种角度出发去理解外尔半金属，首先它可以被看成是一种特殊的金属，其特殊性在于它的费米面被缩成了三维空间中的一个点——外尔点；同样外尔半金属也可以被看成是一种特殊的绝缘体，除了几个孤立的外尔点以外，它在布里渊区的其他区域内都具有有限的能隙。在固体的能带结构中，这样的外尔点是由能级的“偶然简并”造成的，由于外尔点的位置并不在布里渊区的任何高对称点和线上，这种简并与晶体对称性导致的简并是完全不同的。这种看似“偶然”形成的简并点，在传统的固体物理中并不受重视，因为它是不稳定的，在哈密顿量中加入一个小的微扰项就能在原先的简并点处打开能隙。然而，对于三维体系，加入微扰项消除原有简并

点的同时，在邻近区域又会产生新的简并点。其原因很简单，两条能级发生简(交)并(叉)的独立条件有三个，对应于三个彼此独立的2*2厄米矩阵(泡利矩阵)，而在三维体系中，有三个独立的动量 k_x 、 k_y 和 k_z ，因而由三个独立条件可以确定三个动量参数。因此，上述数学问题在一定参数范围内总是有解的，对哈密顿量进行连续变形，只能移动外尔点的位置而不能马上消除它。那么如何能真正消掉一个外尔点呢？唯一的方法是通过连(施)续(加)变(微)形(扰)，把两个“手性相反”的外尔点移动到一起，然后它们就能像粒子物理中的正反粒子一样湮灭了。这么看来，外尔点很像我们第一节介绍的涡旋，只有正反涡旋相遇才能相消。事实上，外尔点正是出现在动量空间中的奇点。

在某个特定的外尔点附近，在对空间坐标重新标度以后，体系的低能有效哈密顿量正是大名鼎鼎的外尔方程，那是ETH的外尔教授在上世纪20年代提出来描写无质量费米子的运动方程。外尔方程有一个非常重要的特性，就是它的解具有特定的手性，也就是粒子的自旋运动总是严格绕着质心平动方向进行，因此可以分成左手(图11(a))和右手(图11(b))两类，其相应的自旋指向分别反平行和平行于动量方向 \mathbf{k} 。这就是外尔点手性的起源，当然在凝聚态物理中，外尔点的“自旋”并非电子真正的自旋，而是所谓“赝自旋”，这一点需要反复强调。把外尔点附近的元激发看成是一种无质量的费米子，是一种典型的粒子物理的视角。而从凝聚态物理的视角则可以把它看作是贝里曲率的奇点，如果说贝里曲率可以被看作是动量空间的磁场，那么外尔点就是这种磁场对应的磁单极子^[8, 46, 47]。关于贝里相位和贝里曲率^[46, 47]，这里补充说明一下。对于固体中的能带电子态而言，贝里相位 $\vec{A}(\mathbf{k})$ 和曲率 $\vec{B}(\mathbf{k})$ 分别由下面两个公式来定义：

$$\vec{A}(\mathbf{k}) = i \sum_{\text{occ}} \langle n, \mathbf{k} | \nabla_{\mathbf{k}} | nk \rangle, \quad \vec{B}(\mathbf{k}) = \nabla_{\mathbf{k}} \times \vec{A}(\mathbf{k}),$$

其中 $|nk\rangle$ 代表占据态波函数。从中可以看出，贝里相位反映的是邻近 \mathbf{k} 点的占据态波函数之间的联络关系，因此它也叫做贝里联络。上述定义的

贝里相位可以看作一个规范场，而贝里曲率则是这个规范场所对应的场强。需要强调的是，贝里相位和曲率是对于固体中电子波函数特性的最简洁描述，而由电子波函数特性导致的物理规律，是传统的固体物理研究中遭到忽视的课题，也是本次诺奖得主 Thouless 先生长期耕耘的领域，其中他的华人高徒牛谦教授也在其中做出了卓越的贡献^[48]。个人认为在今后的本科固体物理教学中，除了介绍能带色散导致的各种物理效应以外，也应该更多地加入贝里相位和曲率这些波函数特性导致的物理效应，以免在求学期间形成“能带色散关系决定一切物理特性”的错误观念。这种观念一旦在头脑中固化，将贻害无穷。

从贝里曲率的观点看，TKNN指数就是它在二维布里渊区内的面积分，或者说是流过整个二维布里渊区的“磁通量”。利用微分几何中的斯托克斯定理可以证明这个积分必须是 2π 的整数倍，也就是TKNN指数乘以 2π 。利用贝里曲率还可以非常简洁地证明外尔半金属的一个重要特性，即固体能带中的外尔点都是手性一正一反成对出现的。这其实又是紧致性的一个体现，这次体现在布里渊区本身，大家知道在布里渊区中的晶体动量相差一个倒格矢是完全等价的，就跟相位空间相差 2π 一样。然后让我们来考察图12所示的三维布里渊区，在其中取 k_z 为常数的平面。这样的平面可以看作是某一个二维体系的布里渊区，上面提过外尔半金属也可以看成是某种特殊的绝缘体，其能隙只在有限的几个外尔点关闭，那么对于某个特定的 k_z 方向截面而言，只要其中不包含外尔点，那么这个截面就是一个二维绝缘体系统的布里渊区，可以定义其TKNN指数(陈数)，记为 $C(k_z)$ 。接下来要证明的

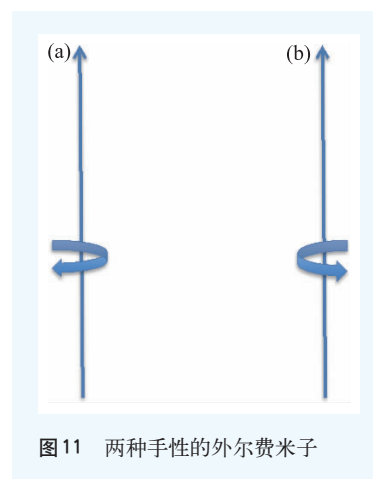


图11 两种手性的外尔费米子

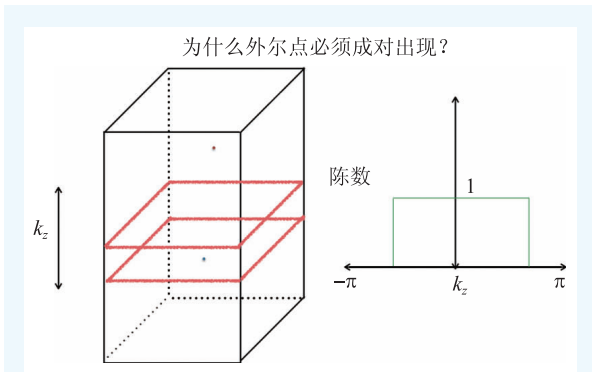


图12 总是成对出现的外尔点

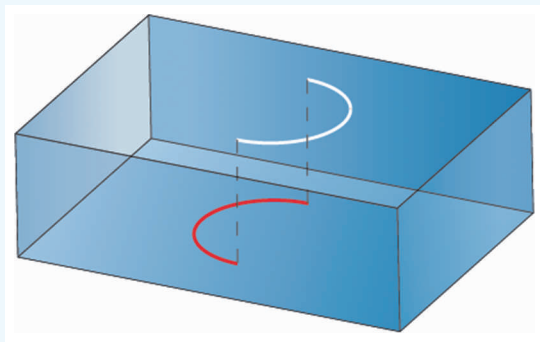


图13 外尔半金属表面的费米弧

是，当 k_z 方向截面穿过某一个外尔点时，按照该外尔点的手性，其TKNN指数(陈数)将跳变+1或-1。为证明这一点，让我们来考察图12中由跳变前后的两个 k_z 截面(红色截面)以及四个侧面所组成的长方体。显然这个长方体是一个三维空间中的封闭曲面，并且包围着一个孤立的外尔点。前面讲过，从贝里曲率的观点看，外尔点是贝里曲率或者说是动量空间磁场所对应的“磁单极子”，满足高斯定律，即流出任何一个封闭曲面的“磁通量”积分等于其中包围的“磁单极子”也就是外尔点的总“磁荷”。而对于图12定义的长方体而言，很容易证明通过四个侧表面流出的“磁通量”两两互相抵消，原因是由于布里渊区的紧致性，左边侧面等价于右边侧面，所以左边流出的“磁通量”严格等于右边流入的“磁通量”，同理前后侧面流出的“磁通量”也将互相抵消。于是高斯定理告诉我们上下表面流过的“磁通量”之差，必须等于被夹在其中的外尔点的“磁(手)荷(性)”，而前面说过上下表面可以看

成是二维布里渊区，其上流过的总“磁通量”积分就等于TKNN指数(陈数)，所以这就证明了当 k_z 方向截面穿过某一个外尔点时，其TKNN指数(陈数)将发生由该外尔点手性所决定的+1或-1的跳变。然后再让我们来考察当 k_z 从整个三维布里渊区的下底面 $k_z=-\pi$ 演化到上底面 $k_z=\pi$ 时，定义在相应 k_z 截面上的TKNN指数的变化，又由于布里渊区的紧致性上下底面是严格等价的，也就是说它们的TKNN指数是一样的，于是从 $k_z=-\pi$ 到 π ，跨越整个布里渊区的过程中遇到的所有外尔点的“磁(手)荷(性)”相加必须等于零。这就证明了在任何格点体系中，外尔点必须是一正一反成对出现的，以保证整个布里渊区内的总“磁(手)荷(性)”为零。

表面上的费米弧是外尔半金属中出现的另一个奇异物性。所谓费米弧，指的是开放费米面。在金属中，费米面的定义就是准粒子能量等于化学势的等能曲面。在普通二维金属中，由于能带色散是连续函数，所以它的等能线通常是闭合的。而外尔半金属的表面态则很特殊，它虽然也是一个二维体系却不能脱离三维体材料而单独存在。本质上，外尔半金属表面态形成的费米面仍然是封闭的，只是另半边费米弧通过体内的外尔点连到了下表面，如图13所示。这种特殊的费米弧其实也不难理解，让我们来考察图14(a)中包围着一个外尔点的圆柱体，再啰嗦一句，这个圆柱体其实是一个轮胎面，因为上下底面等价，所以它也可以被看作是某个二维体系的布里渊区。于是再一次利用高斯定理，我们可以得到，穿过这个圆(轮)柱(胎)体(面)的总“磁通量”积分，也就是该二维体系的TKNN指数，等于被围于其中的外尔“磁荷”总数。具体到图14(a)的情况，假设被围的外尔点手性为+1，那么这个二维体系的TKNN指数就是+1。然后，我们再来考察在 z 方向形成的表面，图14(a)中的圆(轮)柱(胎)体(面)投影到该表面形成一个圆环，围着外尔点的表面投影点，如图14(b)中红线所示。这个圆环就是上述TKNN指数为+1的二维体系的边缘布里渊区，而我们知道这样的二维体系会有一个单向色散的边

缘态，就像量子霍尔效应或量子反常霍尔效应体系一样。于是在上述红色圆环上就只有一个费米点，记住这一圆柱体的半径是可以改变的，只要它围着同一个外尔点，于是在表面布里渊区内每一个包围同一个外尔投影点的圆环上都有且只有一个费米点，把它们连起来就形成了一段开放的弧线，不难证明这段弧线只能终止于另一个手性相反的外尔点投影，如图 14(b)所示。对于绝大多数外尔半金属材料而言，费米能都没有那么巧，正好穿过外尔点，总是有一点偏差，由此形成有限大小的包围着外尔点的三维费米面，投影到表面就形成图 14(b)中的蓝色区域，盖住了相应的外尔点投影。利用刚才介绍的取圆柱体的分析方法，不难证明，只要覆盖不同手性外尔点的三维费米面投影(如图 14(b)中的两块蓝色区域)，并不彼此相连，那么虽然设在蓝色区域之内的表面态不再受到拓扑保护，但在蓝色区域之外必然存在一段受到拓扑保护的费米弧存在。这一点请从事角分辨光电子能谱研究的同事切记。

金属的最大特点是会导电，体现于它在外加电磁场下的响应，而外尔半金属的响应则跟普通金属非常不同。在文献中，大家经常可以读到两个互相联系而又有所区别的效应，即手性反常(Chiral anomaly)和手性磁效应(Chiral magnetic effect)。手性反常指的是具有特定手性的外尔费米子数目，在平行磁场和电场的作用下不守恒，由于在实际固体材料中外尔点是成对出现的，所以这种不守恒不会导致整体电子数不守恒，而是把一种手性的电子转变成另一种手性。而手性磁效应则是指在静磁场下，某个特定的外尔点附近的电子态会贡献一种特殊的电流，根据外尔点手性的不同，这种电流平行或反平行于磁场方向。手性磁效应最奇特的地方在于这种手性电流严格由

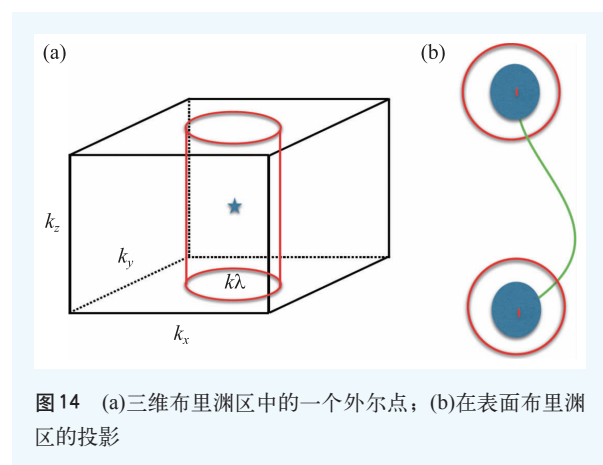


图 14 (a)三维布里渊区中的一个外尔点；(b)在表面布里渊区的投影

磁场方向和外尔点的手性决定，与任何晶体方向或其他材料细节无关，是一种典型的演生现象。由于固体中的外尔点总是成对出现，在平衡态下，不同手性的外尔点贡献的电流严格抵消，体系的宏观电流为零。而一旦在外场驱动下，形成非平衡态，造成不同手性的外尔点处的化学势不相等，则手性磁效应就会导致上述手性电流，比如在平行磁场和电场的作用下，手性反常效应造成不同手性的电子之间出现化学势差，而进一步由手性磁效应形成平行于磁场的电流，在实验上体现出来就是当电场平行于磁场时出现负磁阻。这是手性磁效应在直流输运中的体现，当然这种不同手性外尔点之间的不平衡也可以由其他外场来驱动，如最近我们提出的利用光场和声子场驱动的手性磁效应^[49]。

目前，对于包括外尔半金属在内的各种拓扑半金属的研究方兴未艾^[26, 41]，除了特殊的表面态以外，对其输运和光学性质的研究是其中最迷人的部分，特别是在磁场下的行为，与传统的金属和半导体都截然不同，是大有可为的创新领域。

参考文献

- [1] http://www.nobelprize.org/nobel_prizes/physics/laureates/2016/
 [2] Kosterlitz J M, Thouless D J. *J. Phys. C: Solid State Phys.*, 1973, 5 (11): 124
 [3] Kosterlitz J M, Thouless D J. *J. Phys. C: Solid State Phys.*, 1973,

- 6 (7): 1181
 [4] Berezinskii V L. *Sov. Phys. JETP*, 1971, 32 (3): 493
 [5] Berezinskii V L. *Sov. Phys. JETP*, 1971, 34 (3): 610
 [6] Thouless D J, Kohmoto M, Nightingale M P *et al.* *Phys. Rev.*



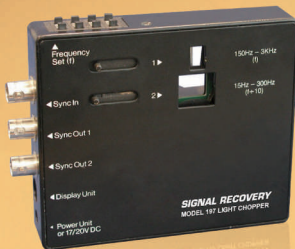
微弱信号检测 半个世纪的骄傲

Model 7210
多通道锁相放大器

全球唯一
通道之最



Model 197光学斩波器



生产商: 阿美特克商贸(上海)有限公司北京分公司
电话: 010-85262111-10 传真: 010-85262141-10
Email: info@ametech.cn
网址: www.signalrecovery.com.cn

中国代理商: 北京三尼阳光科技发展有限公司
电话: 010-65202180/81 传真: 010-65202182
Email: sales@sunnytek.net
网址: www.sunnytek.net

Lett., 1982, 49:405

- [7] Yu R, Qi X L, Bernevig A *et al.* Phys. Rev. B, 2011, 84 (7): 75119
- [8] Weng H M, Yu R, Hu X *et al.* Advances in Physics, 2015, 64 (3): 227
- [9] Haldane F D M. Phys. Rev. Lett., 1988, 61: 2015
- [10] Yu R, Zhang W, Zhang H J *et al.* Science, 2010, 329 (5987): 61
- [11] Chang C Z, Zhang J S, Feng X *et al.* Science, 2013, 340 (6129): 167
- [12] Klitzing K V, Dorda G, Pepper M. Phys. Rev. Lett., 1980, 45 (6): 494
- [13] Kane C L, Mele E J. Phys. Rev. Lett., 2005, 95: 146802
- [14] Kane C L, Mele E J. Phys. Rev. Lett., 2005, 95: 226801
- [15] Bernevig B A, Zhang S C. Phys. Rev. Lett., 2006, 96: 106802
- [16] Bernevig B A, Hughes T L, Zhang S C. Science, 2006, 314 (5806): 1757
- [17] König M, Wiedmann S, Brüne C *et al.* Science, 2007, 318 (5851): 766
- [18] Fu L, Kane C, Mele E. Phys. Rev. Lett., 2007, 98 (10): 106803
- [19] Moore J E, Balents L. Phys. Rev. B: Condens. Matter, 2007, 75: 121306
- [20] Roy R. Phys. Rev. B: Condens. Matter, 2009, 79 (19): 195322
- [21] Zhang H, Liu C X, Qi X L *et al.* Nat. Phys., 2009, 5 (6): 438
- [22] Chen X, Gu Z C, Liu Z X *et al.* Phys. Rev. B, 2013, 87: 155114
- [23] Hsieh T H, Lin H, Liu J W *et al.* Nature Communications, 2012, 3: 982
- [24] Wang Z J, Alexandradinata A, Cava R J *et al.* Nature, 2016, 532: 189
- [25] Ma J Z, Yi C J, Lv B Q *et al.* Experimental Discovery of the First Nonsymmorphic Topological Insulator KHgSb. 2016, arXiv: 1605.06824
- [26] Weng H M, Dai X, Fang Z. J. Phys.: Condens. Matter, 2016, 28: 303001
- [27] Wan X, Turner A M, Vishwanath A *et al.* Phys. Rev. B, 2011, 83: 205101
- [28] Xu G, Weng H M, Wang Z *et al.* Phys. Rev. Lett., 2011, 107: 186806
- [29] Weng H M, Fang C, Fang Z *et al.* Phys. Rev. X, 2015, 5: 011029
- [30] Huang S M, Xu S Y, Belopolski I *et al.* Nature Communications, 2014, 6: 7373
- [31] Lv B Q, Weng H M, Fu B B *et al.* Phys. Rev. X, 2015, 5: 031013
- [32] Lv B Q, Xu N, Weng H M *et al.* Nature Physics, 2015, 11: 724
- [33] Xu S Y, Belopolski I, Alidoust N *et al.* Science, 2015, 349: 613
- [34] Wang Z, Sun Y, Chen X Q *et al.* Phys. Rev. B, 2012, 85: 195320
- [35] Wang Z, Weng H M, Wu Q *et al.* Phys. Rev. B, 2013, 88: 125427
- [36] Liu Z K, Zhou B, Zhang Y *et al.* Science, 2014, 343: 864
- [37] Liu Z K, Jiang J, Zhou B *et al.* Nat. Mater., 2014, 13: 677
- [38] Burkov A A, Hook M D, Balents L. Phys. Rev. B, 2011, 84: 235126
- [39] Weng H M, Liang Y, Xu Q *et al.* Phys. Rev. B, 2015, 92: 045108
- [40] Fang C, Chen Y, Kee H Y *et al.* Phys. Rev. B, 2015, 92: 081201(R)
- [41] Fang C, Weng H M, Dai X *et al.* Chin. Phys. B, 2016, 25 (11): 117106
- [42] Bradlyn B, Cano J, Wang Z J *et al.* Science, 2016, 353: 558
- [43] Weng H M, Fang C, Fang Z *et al.* Phys. Rev. B, 2016, 93: 241202(R)
- [44] Zhu Z M, Winkler G W, Wu Q S *et al.* Phys. Rev. X, 2016, 6: 031003
- [45] Weng H M, Fang C, Fang Z *et al.* Phys. Rev. B, 2016, 94: 165201
- [46] Berry M V. Proc. R. Soc. Lond. A, 1984, 392(1802): 45
- [47] Fang Z, Nagaosa N, Takahashi K S *et al.* Science, 2003, 302(5642): 92
- [48] Niu Q, Thouless D J, Wu Y S. Phys. Rev. B, 1985, 31: 3372
- [49] Song Z D, Zhao J M, Fang Z *et al.* Detecting Chiral Magnetic Effect by Lattice Dynamics, arXiv: 1609.05442