

神经网络前沿讲座课程作业

莫梁虹

中山大学 物理学院, 广州 510275

1 引言

黄老师的综述文章^[1]通过介绍自旋玻璃的 SK 平均场模型引入了复本对称破缺这一概念, 并讲述了复本对称破缺在神经网络及优化问题研究方面的影响。这篇文章谈到的两个点让我产生了浓厚兴趣:

1. 复本对称破缺这一概念, 尤其是其物理含义;
2. 复本方法与记忆相之间的联系。

为了讲清楚我个人的看法, 本文会以自旋玻璃的 SK model 为主线, 重新梳理一遍复本方法的引入, 以及复本对称破缺的含义。

2 SK model

SK model 实际上是 Ising Model 的改进版, 我们先以 Ising Model 作为引入, 计算其自由能和序参量, 观察这样一个磁性系统的简单特性。

Ising 模型的哈密顿量如下:

$$H = -J \sum_{(ij) \in B} S_i S_j - h \sum_{i=1}^N S_i \quad (1)$$

我们可以写出这个系统的配分函数 Z :

$$Z = \sum_S e^{-\beta H} = \text{Tre}^{-\beta H} \quad (2)$$

由于配分函数的求和计算量过大, 我们考虑平均场近似简化该处计算, 即引入平均自旋 m :

$$m = \sum_i \langle S_i \rangle / N = \langle S_i \rangle \quad (3)$$

$$\delta S_i = S_i - m \quad (4)$$

将上述等式替换进 Ising 模型的哈密顿量中, 通过简单的求和代数关系, 可以得到如下的哈密顿量形式:

$$H = N_B J m^2 - (J m z + h) \sum_i S_i \quad (5)$$

其中 z 是系统的维数, N_B 是总自旋的数目。

将平均场近似下的哈密顿量代入可以得到如下的配分函数:

$$Z = \text{Tr} \exp[\beta \{-N_B J m^2 + (J m z + h) \sum_i S_i\}] \quad (6)$$

$$= e^{-\beta N_B J m^2} \{2 \cosh \beta (J m z + h)\}^N \quad (7)$$

$$(8)$$

进而可以计算得到自由能和磁化率 m :

$$F = -T \log Z \quad (9)$$

$$= -NT \log \{2 \cosh \beta (J m z + h)\} + N_B J m^2 \quad (10)$$

$$\approx -NT \log 2 + \frac{J z N}{2} (1 - \beta J z) m^2 + \frac{N}{12} (J z m)^4 \beta^3 \quad (11)$$

$$m = \frac{\text{Tr} S_i e^{-\beta H}}{Z} = \tanh \beta (J m z + h) \quad (12)$$

我们可以画出自由能 F 和磁化率 m 的图像2.1.

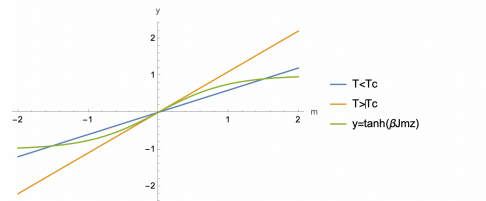


图 1 磁化率 m

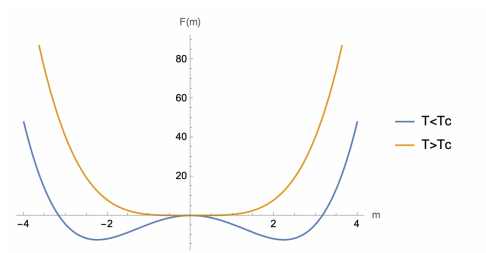


图 2 自由能与磁化率 m 之间的关系

根据上述图像我们可以发现, 当温度低于临界温度时, 这个系统存在两个极小值, 即两个热力

学平衡态，只有当外界的能量足够高时才有可能从其中一个态转移到另外一个状态。能量较低时，系统在任意一个极小值附近的演化，最后也只能落入临近的极小值中。对于更加复杂的系统，比如长程的 Ising 模型，以及两个自旋之间的耦合系数为特定分布的 SK 模型，尽管能量极小值的分布会更加复杂，但是这种极小值附近的热力学态演化的记忆性是仍然保持的。

现在我们直接进入正题，引入 SK 模型。

SK 模型的哈密顿量为：

$$H = - \sum_{i,j} J_{ij} S_i S_j \quad (13)$$

其中 J_{ij} 的分布可以有很多种，后面讨论 hopfield 模型的时候我们会发现这个耦合系数体现了模型的记忆性。D.Sherrington 和 S.Kerkpatrick 最初取 J_{ij} 为较容易计算的高斯分布。

我们进而可以写出这个系统的配分函数：

$$Z = \sum_{\{S_i\}} \exp(\sum_{ij} \beta J_{ij} S_i S_j) \quad (14)$$

其中 $\sum_{\{S_i\}}$ 是对所有构型的求和。

相应的，自由能可以写为^[2]：

$$f_{av} = -\frac{1}{N} k_B T \prod_{i,j} \int P(J_{ij}) dJ_{ij} \ln \sum_{\{S_i\}} \exp(\sum_{ij} \beta J_{ij} S_i S_j) \quad (15)$$

我们可以发现 J_{ij} 在对数函数中，这个积分很难积，故引入了**副本技巧**，利用

$$\ln z = \lim_{n \rightarrow 0} \frac{z^n - 1}{n} \quad (16)$$

将这个问题转换成了对 z^n 的无序平均。将自由能化简后可以得到：

$$f = \lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nN} \left(\sum_{\{S_i^\alpha\}} \exp\left\{ \beta J_0 \sum_{\alpha=1}^n \left(\sum_i S_i^\alpha \right)^2 + (\beta J)^2 \sum_{\alpha, \beta=1}^n \left(\sum_i S_i^\alpha S_i^\beta \right)^2 \right\} - 1 \right) \quad (17)$$

$$\exp\left\{ \beta J_0 \sum_{\alpha=1}^n \left(\sum_i S_i^\alpha \right)^2 + (\beta J)^2 \sum_{\alpha, \beta=1}^n \left(\sum_i S_i^\alpha S_i^\beta \right)^2 \right\} - 1 \quad (18)$$

根据上式可以引入两个序参量：

$$m^\alpha = \frac{1}{N} \sum_i S_i^\alpha \quad (19)$$

$$q^{\alpha\beta} = \frac{1}{N} \sum_i S_i^\alpha S_i^\beta \quad (20)$$

系统的自由度可以写为上述两个序参量的函数。

从 q 的定义我们可以看出，这个序参量刻画的是 α, β 构型之间的热力学平衡态的重合度，也可以理解为耦合系数。

由于复本是计算过程中我们人为引入的，故很自然地认为，这些辅助计算的复本是完全等价的，或者说，没有任何多余的信息可以表示他们之间存在差异。故自然地我们认为复本对称，则 $q^{\alpha\beta} = q$ 。基于这个条件我们可以计算得到系统的平均自由能：

$$\langle f \rangle \approx -\sqrt{\frac{2}{\pi}} J + \frac{T}{2\pi} \quad (21)$$

由此可以得到基态的熵为 $-\frac{1}{2\pi}$ 。这是一件异常的事情，因此 Parisi 怀疑这是因为前面的复本对称假设出现了问题，并提出了他的复本对称破缺理论。假如我们考虑这个系统存在 3 个平衡态，设他们的耦合系数分别为 q_1, q_2, q_3 ，Parisi 发现，一定会有两个相等，且小于等于第三个。故耦合系数矩阵是一个对称矩阵。当某个热力学平衡态附近存在亚稳态，则可以在其耦合系数矩阵中进行局域的分块，因此，我们可以得到如下不同情况的耦合系数矩阵：

假如系统复本对称性未破缺，则耦合系数为 (n=6) 当复本对称性存在破缺 (eg. $n = 6, m_1 =$

$$\{q_{\alpha\beta}\} = \begin{pmatrix} 0 & & & & & \\ & 0 & & & & q \\ & & 0 & & & \\ & & & 0 & & \\ & & & & 0 & \\ q & & & & & 0 \end{pmatrix}.$$

图 3 复本对称的耦合系数矩阵

6, $m_2 = 3$)

根据该矩阵可以验证 q 的矩阵元的一般规律，在 n 趋近于 0 的极限下可以将该矩阵写成一个连续函数，此时计算自由能和热力学熵均与蒙特卡罗得到的结果相吻合。

3 复本对称性破缺的物理含义

$q^{\alpha\beta}$ 的一个著名的物理特性是超度量性，但是此处我想关注的是复本对称性破缺的物理含义。复本的引入本身是处理积分时用到的某种数学技巧，

图 4 复本对称破缺的耦合系数矩阵

我们很自然地要求复本是相同的，或者说，耦合系数 q 是一样的。但是基于这个基本假设的计算结果会导致系统的基态熵为负值，而 Parisi 引入复本对称性破缺这一想法之后，成功解释了实验的蒙特卡洛模拟结果。这种操作看上去很强行，甚至让人觉得会不会这种方法一开始就存在问题？但是，我们进一步思考，也许复本的对称性破缺是存在深刻的物理意义的。

自旋玻璃模型中，由于系统哈密顿量的自旋耦合系数不是一个常数，而是可以取随机或者特定分布的，导致系统的热力学平衡态数目远远多于一般的系统。我们自然会想，如果不同的复本处在不同的热力学平衡态，那么复本对称性破缺就是一件自然的事情。我们甚至可以做出一个大胆的猜想，基于复本对称性破缺这个思想，宇宙在诞生之初也伴随产生了众多复本，而不同的复本之间的对称性破缺，可能导致了平行宇宙的出现。如果我们可以找到不同复本之间的存在差异的那一个维度，也许就能找到不同的平行宇宙之间的联通方式。虽然本人的这个猜想还没有经过严格的推导，但是这无疑是一个有意思的解读方式。

4 记忆性

通过上面的分析，我们不难发现，Parisi 的解法神奇之处在于不同复本之间的耦合矩阵是一种无限划分的黑塞矩阵，每一次分块都引进了新的序参量。这个时候我们可以猜测，这些序参量包含着这个系统热力学平衡态的信息，反过来，我们也可以根据这些热力学平衡态的信息，构建耦合矩阵，从而推断出某个热力学态所对应的热力学平衡态，或者说，最有可能演化到的热力学平衡态。

换个角度来说，如果不同的热力学平衡态是

人脑中的一个记忆相，则我们可以通过构建合适的耦合矩阵，从一个非记忆相 (非平衡态) 推算出平衡态。

与此相关的一个经典模型就是 Hopfield 模型，这个模型的哈密顿量可以写为：

$$H = - \sum_{i,j} J_{ij} S_i S_j \quad (22)$$

其中 $J_{ij} = \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$, μ 遍历了所有的副本。

假如我们有如下两个模式5,6,我们可以根据构建得到的系数矩阵,利用复本方法得到初始状态为8演化得到的热力学平衡态5(找到了其对应的记忆相)。

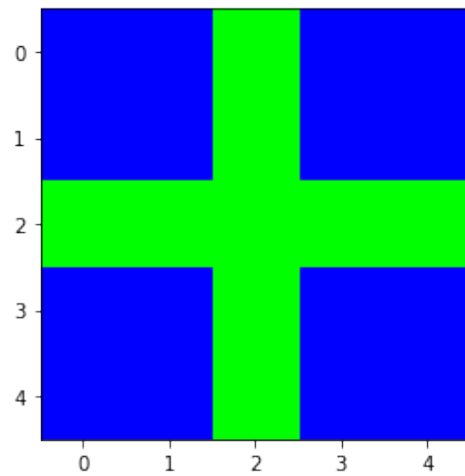


图 5 模式 1

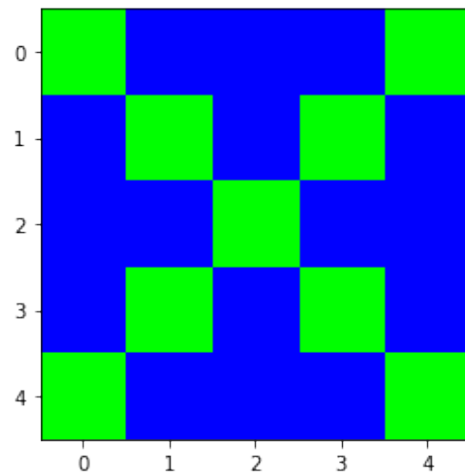


图 6 模式 2

根据上述过程我们可以发现，自旋玻璃模型的这一记忆性可以广泛地应用在其他领域，尤其是量子纠错之中，因为当系统发生局部的错误时，

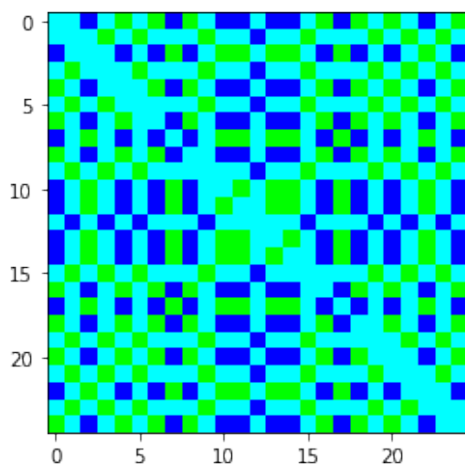


图 7 系数矩阵

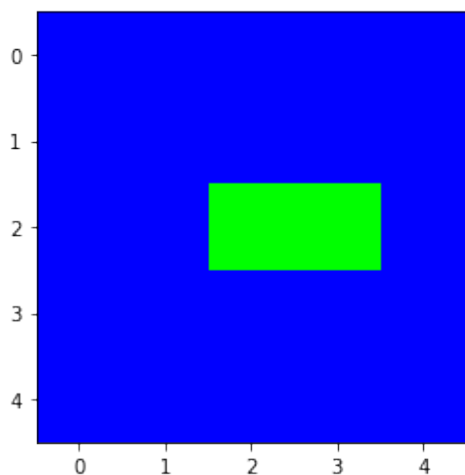


图 8 初始化

其状态类似于偏离某一平衡态的非平衡态，如果我们能很好地利用系统的记忆性，则可实现系统的纠错，但是要注意的是，这样的系统的哈密顿量及耦合系数应该使得势阱尽可能地深，这样产生错误后，系统也不易从一个平衡态跃变到另外一个平衡态，仍然能够保持其记忆性。

5 卷积

卷积神经网络的卷积与纯数学中定义的并不完全一致。

在数学课上，我们学到的卷积定义为：

$$S(t) = \int x(a)\omega(t-a)da \quad (23)$$

离散形式可以写为：

$$S(t) = (x * \omega)(t) = \sum_{a=-\infty}^{\infty} x(a)\omega(t-a) \quad (24)$$

在卷积神经网络的术语中，卷积的第一个参数(x)即输入，第二个参数(ω)即核函数，输出被称为特征映射。

神经网络中的卷积指的是离散卷积。

考虑两个矩阵，第一个矩阵是输入矩阵，第二个矩阵类似于一个滤波器，是一个特征筛选矩阵，其维度小于输入矩阵，具体过程如图11所示。卷积是指滤波矩阵与输入矩阵进行局域内积，当滤波矩阵以步长为1对输入矩阵进行扫描和内积后，可以得到卷积矩阵。我们可以通过 matlab 计算一个

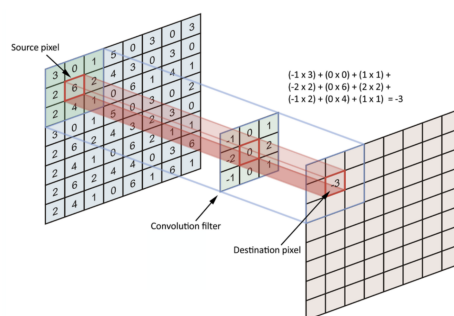


图 9 卷积 (<https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks>)

具体例子：

```
1 % 滤波矩阵
2 S=[1,0,-1;1,0,-1;1,0,-1];
3 % 输入矩阵
```

```

4 I=[3,0,1,2,7,4;1,5,8,9,3,1;2,7,2,5,1,3;
5   0,1,3,1,7,8;
6   4,2,1,6,2,8;
7   2,4,5,2,3,9];
8 Con=zeros(4,4);
9 for m=1:4
10  for n=1:4
11      A=I(m:m+2,n:n+2);
12      Con(m,n)=sum(sum(A.*S)) ;
13  end
14 end
15 Con

```

我们得到的输出矩阵为10

```

Con = 4x4
    -5    -4     0     8
   -10    -2     2     3
     0     -2    -4    -7
    -3     -2    -3   -16

```

图 10 卷积输出矩阵

6 池化

池化过程类似于卷积，即用 $n \times n$ 的滤波器矩阵，进行步长为 n 的扫描，得到这个 $n \times n$ 区域的最大值或者平均值，若是取出极大值，该滤波过程称为最大池化，若是计算平均值，则该过程称为平均池化。池化的过程与卷积类似，只需要将代码中

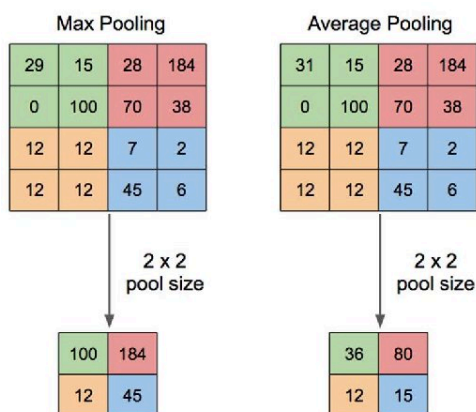


图 11 两种池化方式^[3]

循环部分的 Con 改为：

```

1 Pool(m,n)=max(max(A)) ;

```

我们得到的输出矩阵为12

```

Pool = 4x4
     8     9     9     9
     8     9     9     9
     7     7     7     8
     5     6     7     9

```

图 12 池化输出矩阵

7 复习笔记 Note

7.1 1. 光学量子计算

7.1.1 Hong-Ou-mandel

当光子进入分束器时，有两种可能性：它要么被反射，要么被透射。透射和反射的相对概率由分束器的反射率决定。在这里，我们假设一个 1:1 的分束器，其中一个光子被反射和透射的概率相等。

接下来，考虑两个光子，一个在 1:1 分束器的每个输入模式中。关于光子的行为有四种可能性：

1. 从上面进来的光子被反射，从下面进来的光子被透射。
2. 两个光子都被传输。
3. 两个光子都被反射。
4. 从上面进来的光子被透射，从下面进来的光子被反射。

我们现在假设这两个光子在它们的物理特性（即偏振、时空模式结构和频率）上是相同的。此外，来自分束器底部的反射引入了 π 的相对相移，对应于叠加中相关项中的因子 -1。这是分束器的可逆性（或量子演化的统一性）所要求的。由于这两个光子是相同的，我们无法区分可能性 2 和 3 的输出状态，它们的相对负号确保这两项抵消。

因此，当两个相同的光子进入 1:1 分束器时，它们将始终以相同（但随机）的输出模式离开分束器。即，最后的态为 $|02\rangle - |20\rangle$

7.1.2 玻色取样

“玻色取样”基本概念：当 n 个全同玻色子经过一个干涉仪（线性变换器）之后，求特定分布的输出概率。例如，在一个 7 进 7 出干涉仪的 1、2、3 口同时输出 3 个全同玻色子，求 3 个光子在 2、3、5 口各输出一个光子的分布概率。

Aaronson 和 Arkhipov 研究发现, n 光子“玻色取样”的分布概率正比于 n 维矩阵积和式 (Permanent) 的模式

玻色取样的本质就是生成模拟信号, 是蒙特卡洛模拟。在数字计算机中实现的蒙特卡洛模拟, 是通过软件来实现的模拟, 而玻色取样则是通过硬件来实现的蒙特卡洛模拟。

7.2 中微子和暗物质

7.2.1 中微子简史

beta 衰变中子变成质子 + 电子 + 反中微子

7.2.2 中微子的手征性

中微子是左旋的, 反中微子是右旋的;

7.2.3 中微子振荡

中微子存在质量本征态和味道本征态, 希尔伯特空间 $2d$ 。

7.3 生物物理

单分子荧光技术单分子荧光共振能量转移 FRET FIONA 研究肌球蛋白光镊操纵材料光镊就是用一束激光俘获微笑束缚在光束的焦点位置

单分子荧光成像技术可以分为两类: 一类是在外力作用下研究单分子活动, 通常通过原子力显微镜 (AFM)、光镊 (OT) 或磁镊 (MT) 将力施加到单个分子上。另一类就是用荧光显微成像观察生物系统中单分子活动。

荧光显微成像是生命科学领域观察生物体结构的经典方法。这其中, 用荧光探针标记, 检测和分析单个分子的单分子荧光成像技术, 能够帮助科学家们在破坏生命体正常生理状态的情况下, 清晰地观察到单个分子的活动。

7.4 强子与夸克

7.4.1 核力的特点与电荷无关

核子间交换 π 介子产生核力场

7.5 拓扑物理

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 和斯坦福大学的 Bernevig / 张首晟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”。这样, 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。

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”。这样，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。

2011 年 zhang 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

这时候你会发现在费米子链的两端存在两个孤立的马约拉纳，如果链的长度足够长，使得两个马约拉纳的波函数不发生重叠的话，这两个马约拉纳将会稳定的存在于链的两端。

所有的马约拉纳结合成一个费米子都需要能量，但是这个哈密顿量不包含链两端的两个马约拉纳，这意味着，如果我们让链两端的马约拉纳结合其实并不需要消耗能量！为了更清楚的看到这两个马约拉纳零模，一个简单又方便的方法，就是用上面的哈密顿量构造一条有限长的链，通过数值严格对角化后

可以看到对于后一种情况，存在两个零能态孤子解，并且波函数分布在链两端，而对于一般的情况零能态并没有被占据，波函数也主要分布在链条内部。

从能谱不难看出，内外边界的传输方向是相反的，其次该体系不存在 0 模，但是斜率相反的两个传输通道具有手性，并且参与输运的通道只有半个自由度，因此被称为马约拉纳手征模 (chiral Majorana modes)！现在如果我们在中心空白区域通入半个磁通，这个体系会发生一个显著的变化：波函数绕磁通一周后会获得半个磁通贡献的相位，从而引入一个新的负号。这意味着波函数现在具有了周期性边界条件而非半周期！！周期性边界条

件意味着，现在的 n 可以取整数了（可以为零了）！！

有带隙的能带结构都能在拓扑性质上被分类，分类方法即是：如果一个能带结构能通过不闭合带隙（即能带不相互交叉，或不增加交叉的数量）的条件下，连续地变换到另一个能带结构，则可以把这两个能带结构对应的布洛赫哈密顿量 [公式] 视为一个等价类（关于等价类知乎上的回答已经足以解释此部分的需要，可自行搜索）。这些等价类可以通过一个拓扑不变量 [公式]（[公式] 是整数集合）来区分，这个拓扑不变量称为陈不变量 (Chern invariant)。也就是说，如果两个能带能够相互连续变换，且不闭合能带，他们就有着相同的陈不变量。

参考文献

- [1] 黄海平. 统计物理、无序系统和神经网络[J]. 科学, 2022, 74(005): 40-44.
- [2] BROUT R. Statistical mechanical theory of a random ferromagnetic system[J]. Physical Review, 1959, 115(4): 824.
- [3] YANI M, et al. Application of transfer learning using convolutional neural network method for early detection of terry' s nail[C]//Journal of Physics: Conference Series: volume 1201. IOP Publishing, 2019: 012052.