

熵之为熵(1)：缘起

已有 5778 次阅读 2016-6-3 00:18 | 个人分类:科研笔记 | 系统分类:科研笔记

【特别申明：本系列博文涉及学术原创，其他写作者在引用其中思想时，敬请注明博文的网址；转载时，也请注明来源和作者。另：本系列博文的发表，并不影响作者在其他学术刊物上以论文的形式再次发表。】

(1)

大贤曹则贤兄说过：“内涵就在表面”。读者诸君，若我有内涵而不言，则无表面表现，也就没有了内涵。没有了内涵，则我就不算是我了。

熵之于我，内涵之内涵也。经年累月而思，不在心头，就在眉头，到了不敢不言，不能不言的境地。

有人说了，若要谈熵，何不寻一正规的学术杂志发表？你不知道科学网上都是些民科么？就是官科，也主要是插科打诨的，当不得真。

我来说说我的烦恼之处。按如今学术文章之泛滥，而相对读者数量之稀少，随便寻一低档杂志发表了，无非是又舔几张擦屁股纸而已；而要寻高档的杂志发表，则文章之写法、内容之时髦、编辑之好恶等等，都变成了枷锁，也许文章还没出门，就被打死在家里了。

抛开枷锁看世界，只要不被首发位置，档次高低等等概念迷惑，科学网确是个发表的好地方。此处水草肥美，粉丝众多，而呼朋唤友，砥砺切磋，把酒笑谈，正经扯淡，真正再好不过了。

当然，前有张树凤、黄秀清、张志东等，外带文小刚，在科学网上叫唤而折了名声，让我不得不小心谨慎。但是，“不疯魔，不成活”，有话要讲，这名声就顾不得了。大不了，哪天去斯德哥尔摩领奖，便改名叫“邢志忠”罢了。

(2)

言归正传。

本系列所谈，乃是香农熵、热力学熵和Boltzmann-Gibbs熵之间的缝隙，我希望像泥瓦匠一般将这些缝隙勾勒弥补起来。

关于这些熵的历史，我会在相关的章节提到。

现在概说这些熵之间的缝隙。

香农熵的缘起，来源于对通信情况的描述。简言之，若一个系统有 $i = 1, 2, \dots, N$ 种状态，系统取其中第 i 种状态的概率为 p_i ，则此系统的香农熵为：

$$S = - \sum_{i=1}^N p_i \ln p_i \quad (1-1)$$

如果有两个完全独立的系统, 则我们很容易证明这两个系统的熵-如果它们分别为 S_1 和 S_2 的话-加和到一起就变成了将两个系统合起来看成一个系统的熵, 即:

$$S = S_1 + S_2 = S_1 + S_2 \quad (1-2)$$

若设一个系统出现各种状态的出现概率都是一样大小的, 即 $p_i = 1/N$, 那么我们很容易计算:

$$S = -\ln N = -\ln N \quad (1-3)$$

以上三个公式, 我们在统计力学中也常常会见到他们的身影: 第一个公式往往是在我们知道系统按能级的统计分布-比如波尔兹曼分布-的情况下, 用来计算体系的熵的; 第二个公式是用来证明熵是一个广延量的; 第三个公式, 则是对于孤立系算熵的基本公式, 也是Boltzmann-Gibbs熵的出发点, 这个公式也暗含了等概率假说-孤立系的各种状态出现的概率是相同的。

但是, 细心的读者会发现, 这里有一道Boltzmann-Gibbs熵和Shannon熵的缝隙。刻在波尔兹曼墓碑上的公式是:

$$S = k \ln W = k \ln W \quad (1-4)$$

不错, 这里 W 和公式(3)中的 N 含义相同, 为体系的状态数。所以(1-3)和(1-4)非常相似-除了一个差别, 系数 k , 即波尔兹曼常数。 k 是有量纲的, 为 $[J K^{-1}]$ 。正是这个量纲, 将熵和一个统计力学处理对象的能量和温度联系起来。

最初的 k 值仅仅个带量纲的比例常数罢了, 但是到了Plank处理黑体辐射, 搞出光量子的時候, 这个常数也被定了下来。(https://en.wikipedia.org/wiki/Boltzmann_constant) 今天, 其精确值为

$$1.38064852 \times 10^{-23} J K^{-1}$$

这个常数的确定, 给我们带来了第二道缝隙。

比如, 以常用的二维Ising铁磁模型为例, 一对指向向上的自旋和一对指向向下的自旋之间相互作用按能量计, 是相同的。在相变温度以下, 体系会自发的取某种指向, 这意味着取相反指向的自旋相当少, 从吉布斯的正则系综出发, 我们应该把两种体系指向的状态都考虑在内。当我们计算体系的热容的时候, 只要考虑自旋相互作用就可以了, 而不用考虑两种不同的指向, 这时, 体系似乎有了两种不同的统计力学熵值-一种是考虑两个不同的体系指向的, 一种不考虑, 这两种结果在 k 固定的情况下, 会有 $k \ln 2$ 的差异。本来这个差异不重要, 因为整个体系的熵完全可以看做是个势函数, 零点取在哪里都可以。但是, k 的固定背后, 意味着相空间划分成相格的大小也固定了, 同时也从统计上指明了绝对零度时体系的熵的计算方式, 熵也不再是个相对大小的概念, 而应该是个绝对的数值。那么我们到底该怎么计算系统的熵呢?

而论及统计力学熵与热力学熵的缝隙，这问题就比较多了。

第三道缝隙，来自热力学自身的历史。众所周知，在考虑封闭系时，仅考虑压强的作用，热力学的基本关系式为：

$$dU = TdS - PdV \quad (1-5)$$

从这个关系出发，再结合热力学第0、第1和第2定律，我们发现，S在体系趋于绝对零度时，可以指定为任意值。但是，在热力学的发展史上，出现了热力学第三定律

(https://en.wikipedia.org/wiki/Third_law_of_thermodynamics)。而根据热力学第三定律，绝对零度处，体系的熵应该为零。(https://en.wikipedia.org/wiki/Thermodynamic_equations,

https://en.wikipedia.org/wiki/Fundamental_thermodynamic_relation) 但是，一般认为，从统计力学来看，在绝对零度处，有些体系的熵是应该有非零值的，为用基态简并度 (ground state degeneracy) 作为体系状态数，然后用公式 (1-4) 计算的结果。换言之，一般认为，这第三道缝隙，由统计力学来弥补了。

第四道缝隙，非常著名，即Gibbs佯谬 (https://en.wikipedia.org/wiki/Gibbs_paradox)。Gibbs假定，有个箱子，用无限薄、无限光滑的隔板 (当然质量就为零了) 隔成两个大小一样的空间。这两个空间都装满了理想气体，压强也一致。这时将隔板抽走，那么两边空间的气体就会混合。从统计上看，对于两边的每个气体分子，状态数量都会增加一倍，结果整个体系的熵平均到每个气体分子会增加 $k \ln 2$ 。但是从热力学的角度看这个问题，系统在抽动质量为零又无限光滑的隔板时，体系和环境之间没有功交换；而抽板前后，体系中气体的状态从宏观上看，也没有变化，因此系统和环境之间也没有热交换。那么根据 $dS = dQ/T$ ，则体系熵增为零。那么，到底熵是针对每个气体分子增加了 $k \ln 2$ ，还是没有增加呢？

我看过种种处理，从量子力学到经典理论扩展，似乎都与真实的问题擦肩而过。先说量子统计的处理。通过使用粒子的全同性概念-像气体分子这样自由运动的粒子，其所对应的波函数的非定域性，使得粒子间彼此无法通过空间位置来区分。-体系可以将这增加的熵去除。但是，对于具有定域性的粒子，比如绝缘固体内的电子，这个全同性的概念则无法起作用。当然，你可以说，绝缘固体内的电子位置固定，不会像气体那样混合，也就没有这重顾虑了。但是，这样通过全同性的方式处理问题，虽然精巧，难免牵强。贝克莱可以说：“存在即被感知”，我们这里的全同性处理，就完全可以说“存在即被分得出位置”。对我这样一个粗陋的工程师而言，是难以接受的。

第五个缝隙，则是统计力学处理问题的方式和热力学处理实际问题之间的差异。在统计力学里，我们总是会让体系处于系统无限大的所谓热力学极限情况下来讨论问题，又会引入系综平均等于时间平均来分析体系的行为。所谓无限大体系，往往会和如今处理问题的对象相矛盾。比如纳米尺度的颗粒，我们该如何分析呢？而有的体系行为，我们几乎没有办法使用时间平均来代替系综平均。比如低温下的铁磁系统，从概率上讲，只要不是绝对零度，这个体系总是有机会从一种spin的总体取向变成另外一个相反的取向，或者说，只要经过足够长的时间，体系总是可以从一种取向变成另一种。但是，这个时间是多长呢？估计体系稍微大点，我们就可以让这个时间和宇宙诞生到现在的时间相媲美。当然，这个实验是没法做了。但是，我们很清楚，这些实验，我们依然可以考虑其温度，分析其热容，讨论其热力学熵。这个缝隙使得我们不得不重新来审视热力学熵的统计学含义，而不敢再简单地引用波尔兹曼的最初假定，然后给出个简洁的结论了。

(3)

在这个系列中，我将以信息论的视角，来讨论以上提到的缝隙，以便使缝隙明晰。并且，在尽可能的情况下，我会提一些弥补这些缝隙的方案。

在后面的章节里，我将在第一部分讲述需要用到的信息熵知识和统计学知识，重点是各种信息熵的基本概念和马尔科夫链；按照我的爱好，我自然也会提到相关的历史。而第二部分，我则要讨论热力学熵的统计含义。第三部分则是讨论热力学熵和统计力学熵之间的缝隙以及相应的处理手段的。

在本博文最后，我要向曹则贤兄致谢。因为正是曹则贤兄，给我传来了Palmer的文章，坚定了我诉说的勇气。而在博客上讨论专业的学术问题，曹兄也是最终的动因。我觉得他仿佛对我说：“要爱，就要勇敢说出来。”