热学
第4讲 万能法则和熵

黄志琦

教材：《热学》第二版，赵凯华，罗蔚茵，高等教育出版社
课件下载 https://github.com/zqhuang/SYSU_TD

上一讲内容回顾

- 态是相空间里的离散“小块”（学会相空间划小块的游戏，量子力学的物理图像就掌握一半了）
- 高斯积分是行走江湖的必备技能（反正你现在不掌握将来也是要掌握的，不如现在花点时间）
- 麦克斯韦分布和麦克斯韦-玻尔兹曼分布（考试重点）

数学概念充电：多元变量的概率密度

假设有块长为200米，宽为200米的方形麦田。我们取麦田中心为原点，用坐标(x, y) (单位均为米）来描述麦田里的点。$-100 \le x, y \le 100$。

田鼠随机地选择一个点打洞，我们忽略田鼠窝的大小，把田鼠窝抽象为一个点。

数学概念充电：多元变量的概率密度

田鼠窝在$(50, 50)$的概率是多少?

数学概念充电：多元变量的概率密度

田鼠窝在$(50 \leq x \leq 51, 50 \leq y \leq 51)$这个小方块里的概率是多少?

数学概念充电：多元变量的概率密度

田鼠窝在$(50 \leq x \leq 50 + dx, 50 \leq y \leq 50 + dy)$这个小方块里的概率是多少?

数学概念充电：多元变量的概率密度

以x, y为变量，田鼠窝在$(50, 50)$的**概率密度**是多少?

数学概念充电：多元变量的概率密度

田鼠窝在$50 \leq x \leq 51$这个长条内的概率是多少?

数学概念充电：多元变量的概率密度

田鼠窝在$50 \leq x \leq 50 + dx$ 这个长条内的概率是多少?

数学概念充电：多元变量的概率密度

以x为变量，田鼠窝在$x = 50$处的**概率密度**是多少?

数学概念充电：多元变量的概率密度

以$s \equiv x^2$为变量，田鼠窝在$s = 2500$处的概率密度是多少?

数学概念充电：多元变量的概率密度

以$s_x \equiv x^2$，$s_y \equiv y^2$为变量，田鼠窝在$s_x = 2500, s_y = 2500$处的概率密度是多少?

本讲内容预告

- 能均分定理与热容量
- 熵 (entropy)

能均分定理(Theorem of Equipartition of Energy)

假设某个自由度的能量形式为（广义）动量的二次型$\varepsilon = cp^2$ (例如任何一个自由度的平动动能或转动动能）或者（广义）坐标的二次型 $\varepsilon = cx^2$（任何一个自由度的小振幅振动的势能等）， 我们可以用"万能法则"——概率$\propto e^{-\frac{\varepsilon}{kT}}$以及我们上节课学习的高斯积分公式证明这个自由度贡献了$\frac{1}{2}kT$给分子平均能量。

动量二次型能量的能均分定理证明

$$
\begin{aligned}
\frac{\int_{-\infty}^{\infty} c\mathfrak{p}^2 e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}}{\int_{-\infty}^{\infty} e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}}
&= \frac{-\frac{kT}{2}\int_{-\infty}^{\infty} \mathfrak{p}\, d\left(e^{-\frac{c\mathfrak{p}^2}{kT}}\right)}{\int_{-\infty}^{\infty} e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}} \\
&= \frac{-\frac{kT}{2}\left[\mathfrak{p} e^{-\frac{c\mathfrak{p}^2}{kT}}\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}\right]}{\int_{-\infty}^{\infty} e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}} \\
&= \frac{\frac{kT}{2}\int_{-\infty}^{\infty} e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}}{\int_{-\infty}^{\infty} e^{-\frac{c\mathfrak{p}^2}{kT}}\, d\mathfrak{p}} \\
&= \frac{kT}{2}
\end{aligned}
$$

计算也可以直接用上节课讲的“交换积分和求导次序”的办法来进行。

坐标二次型能量的能均分定理证明

坐标二次型的情况，把上述证明中的p换成x即可。

证明都是利用了相空间（的态）是按坐标和动量均匀划分的。

思考题

如果能量是某自由度动量的四次型$\varepsilon = cp^4$ ($c > 0$为常量)，那么能均分定理还成立吗?

摩尔定体热容量C_V^{mol}

摩尔定体热容量$C_V^{\mathrm{mol}} \equiv 1\,\mathrm{mol}$ 固定体积的气体升高$1\,\mathrm{K}$需要吸收的热量

摩尔定体热容量C_V^{mol}

若分子的自由度为s个，由能均分定理，1 mol气体包含内能为

$$U^{\mathrm{mol}} = N_A \frac{skT}{2} = \frac{sRT}{2}$$

固定体积意味着不做功，内能的增长来自于吸收的热量，于是固定体积的摩尔热容为

$$C_V^{\mathrm{mol}} = \frac{s}{2} R$$

思考题

理想情况下，下列物质的摩尔定体热容量为多少?

- 单原子气体
- 双原子气体
- 构成正三角形的三原子分子
- 固体

玻尔兹曼的墓碑(Boltzmann's Grave)

熵 = 微观状态数的对数

而当你翻开热学书 🤦

- 熵=混乱程度(大众理解的熵)

$$S = \text{男生宿舍}$$

- 熵=微观状态数的对数（玻尔兹曼熵）

$$S = k \ln \Omega$$

- 熵=可逆过程的热温比积分(克劳修斯熵)

$$S = \int_{\text{reversible}} \frac{\text{đ}Q}{T}$$

- ...

之所以熵这么多（这么混乱），是因为没有清晰地区分场景。下面我们循序渐进，依次来讨论“概率的熵”，“对象的熵”，“多次重复对象的熵”。

概率的熵

我们假设任何一个概率p $(0 \leq p \leq 1)$都对应一个确定的熵，记为$s(p)$。注意我们并不关心概率描述的事件，而仅仅把熵看成概率的一个函数。(或者，对数学家而言，熵仅仅是某个定义域为$[0, 1]$的函数。)

对这个我们目前还没有给出明确表达式的函数，先来猜想一下它的性质：

- 概率$p = 1$的事件，我们有完整的信息，(猜想）熵为零：$s(1) = 0$。
- 概率$p = 0$的事件，我们有完整的信息，（猜想）熵为零：$s(0) = 0$。
- 概率$0 < p < 1$的事件，我们有部分的信息，（猜想）熵大于零：$s(p)|_{0<p<1} > 0$。

思考题

什么概率的熵最大?

拓宽场景

“一个概率有多混乱”这个问题是非常难回答的，为了得到$s(p)$的合理定义，我们必须拓宽熵的应用场景。

对象(object)和态(state)

- 抛硬币有两种态：正面，反面; 每个态的概率为1/2。

- 掷骰子有六种态：1, 2, 3, 4, 5, 6，每个态的概率为1/6。

- 修热学课有两种态：
  挂科，概率0.6
  不挂科，概率0.4 (概率是我随便瞎编的)

我们把抛硬币，掷骰子，修热学课这些**包含多种态，每种态的概率都已知的事件称为“对象”**。

对象的熵：直接把概率熵加起来就行了

设一个对象有m种态，它处在各个态的概率分别为p_1，p_2，...，p_m。我们把这个对象记作$O(p_1, p_2, \ldots, p_m)$，并定义这个**对象的熵为**

$$S\left[O(p_1, p_2, \ldots, p_m)\right] \equiv \sum_{i=1}^{m} s(p_i)$$

这是再自然不过的定义了。

抛硬币实验$O(\frac{1}{2}, \frac{1}{2})$的熵为$2s(\frac{1}{2})$.

掷骰子实验$O(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$的熵为$6s(\frac{1}{6})$.

修热学课$O(0.6, 0.4)$的熵为$s(0.6) + s(0.4)$.

重复对象的熵

考虑N个相同的对象$O(p_1, p_2, \ldots, p_m)$,并把它记作$O_N(p_1, p_2, \ldots, p_m)$。

自然而然地，定义**重复对象$O_N(p_1, p_2, \ldots, p_m)$的熵为:**

$$S\left[O_N(p_1, p_2, \ldots, p_m)\right] \equiv N \sum_{i=1}^{m} s(p_i)$$

N次抛硬币的熵为$2Ns(\frac{1}{2})$.

N次掷骰子的熵为$6Ns(\frac{1}{6})$.

N次修热学课的熵为$N\left(s(0.6) + s(0.4)\right)$.

玻尔兹曼有点不开心

迄今为止，我们对三种场景定义了熵：概率，包含多个概率的对象，多个重复的对象。按熵可叠加的假设，熵的定义也是自然而然地延拓:

- 每个概率p，不论其含义，都唯一地对应一个熵$s(p)$
- 包含多种可能性的对象的熵 = 每种可能性的概率的熵之和
- N个重复对象的熵 = 单个对象的熵 x N

玻尔兹曼有点不开森：讲了半天，老子的墓碑呢?

完美N次测量

如果N个重复对象的测量结果完全符合统计预期，即任意第i个态都精确地出现了$p_i N$次$(i = 1, 2, \ldots, m)$，就称为对象$O(p_1, p_2, \ldots, p_m)$的**完美N次测量**。

如果修100次热学课，挂了60次，过了40次，就是修热学课的完美100次测量。

实现方法数Ω

例如抛硬币的完美4次测量，必须2次正面，2次反面。实现方法有：

正正反反
正反正反
正反反正
反正正反
反正反正
反反正正

总共6种。我们记$\Omega = 6$。

实现方法数Ω （续）

例如修热学课的完美5次测量，必须3次挂，2次过。实现方法有：

挂挂挂过过
挂挂过挂过
挂挂过过挂
挂过挂挂过
挂过挂过挂
挂过过挂挂
过挂挂挂过
过挂挂过挂
过挂过挂过
过过挂挂挂

总共10种，记$\Omega = 10$

实现方法数Ω （续）

一般性地，对象$O(p_1, p_2, \ldots, p_m)$的完美N次测量的实现方法数为

$$\Omega = \frac{N!}{(Np_1)!(Np_2)!\ldots(Np_m)!}$$

（不会证的请复习初中排列组合内容🤔）

例如上面的例子$O(0.4, 0.6)$的完美5次测量的实现方法数为：

$$\Omega = \frac{5!}{(5 \times 0.4)!(5 \times 0.6)!} = \frac{5!}{2!3!} = 10$$

万事俱备，召唤墓碑

记对象O的完美N次测量的实现方法数为Ω，为了安抚玻尔兹曼，我们希望

$$S(O) = \lim_{N\to\infty} \frac{\ln\Omega}{N}$$

当然，我们暂时抛开了无关紧要的常数k (它只是重新定义了熵的单位而已）。
代入Ω的表达式，即有

$$\sum_{i=1}^{m} s(p_i) = \lim_{N\to\infty} \frac{\ln\left(\frac{N!}{(Np_1)!(Np_2)!\ldots(Np_m)!}\right)}{N}$$

Stirling公式

当N很大时，忽略小于$O(N)$的项:

$$\begin{aligned}
\ln(N!) &= \sum_{i=1}^{N} \ln i \\
&\approx \int_1^N \ln x \, dx \qquad \text{(请无视数学老师幽怨的眼神)} \\
&= (x\ln x - x)\big|_1^N \\
&\approx N\ln N - N
\end{aligned}$$

于是我们得到

$$\sum_{i=1}^{m} s(p_i) = \lim_{N\to\infty} \frac{N\ln N - N - \sum_{i=1}^{m}\left[Np_i\ln(Np_i) - Np_i\right]}{N}$$

概率熵的表达式

最后，我们利用 $\sum_{i=1}^{m} p_i = 1$，化简得到

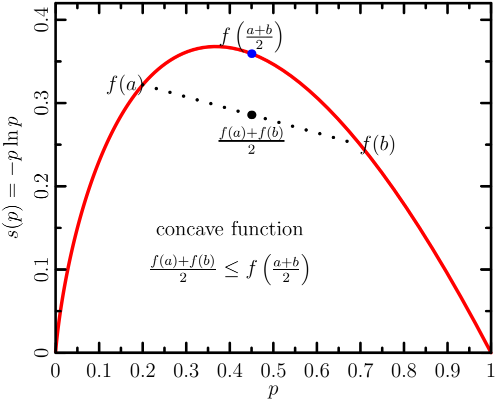
$$\sum_{i=1}^{m} s(p_i) = -\sum_{i=1}^{m} p_i \ln(p_i)$$

很显然，唯一能令玻尔兹曼开心的办法是**定义概率熵为：**

$$s(p) = -p \ln p$$

注意：当$p = 0$时，$s(p) = \lim_{p \to 0^+}(-p \ln p) = 0$.
我们不仅可以验证之前关于$s(p)$的若干猜想，还可以算出$e^{-1} \approx 0.368$是“最混乱”的概率。

熵函数是个凸函数(concave function)



在某区间内部二阶导数恒为负数的函数是该区间内的凸函数。

对$0 < p < 1$

$$s''(p) = -\frac{1}{p} < 0$$

琴生不等式(Jensen's Inequality)

多个概率的熵满足下面的不等式（概率的熵的平均不大于平均概率的熵）：

$$\frac{1}{m}\sum_{i=1}^{m} s(p_i) \leq s\left(\frac{1}{m}\sum_{i=1}^{m} p_i\right)$$

或者简写为

$$\overline{s(p)} \leq s(\bar{p})$$

等号当且仅当所有p_i均相等时取到。

这个不等式其实对所有凸函数都成立 （20秒还没有想出怎么证明的要好好补习下高数了）

等概率原理

考虑对象$O(p_1, p_2, \ldots, p_m)$，由于总概率为1，平均概率是固定的$\bar{p} = \frac{1}{m}$。于是由琴生不等式：

$$S = m\,\overline{s(p)} \leq ms(\bar{p}) = m\,s\left(\frac{1}{m}\right) = \ln m$$

等号当且仅当所有p_i相等时才能取到。

态的个数固定的对象，如果没有额外条件的约束，则当所有态的概率相等时，熵达到最大。这时我们拥有的信息量为零（不知道哪种可能性更容易发生）。

例子：抛硬币游戏的熵

抛硬币，如果正面向上概率为x，反面向上概率为$1-x$，则熵为

$$S = -x\ln x - (1-x)\ln(1-x)$$

通过求导易求出$x = 1/2$时熵最大，为$S_{\max} = \ln 2$。
和等概率原理的结论一致。

如果我们知道硬币是动过手脚的（有额外约束条件），$x \neq 1/2$，那么熵就要小一点了。

抛骰子游戏的熵

掷骰子，根据等概率原理，当且仅当所有p_i为1/6时熵取到最大值$S_{\max} = \ln 6$.

如果我们知道谁在掷骰子（有额外约束条件），就不好说了。

掷骰子——比小——出老千

现在我们去趟澳门，跟赌王掷骰子比小。

江湖险恶，十赌九骗，赌王用的是一个高科技骰子。

如果结果为6，则会触发骰子里的老千系统，使下一次出现1的概率为1/2，出现2至6中任何一个概率都是1/10。
如果连续多次地投掷骰子，每次掷骰子的熵变为多少了呢?

显然这时由于老千系统（额外约束条件），熵不再是$S = \ln 6$了。

掷骰子——比小——出老千（续）

这个老千系统比较简单，很容易列出方程：

$$
\begin{aligned}
p_1 &= \frac{1}{2}p_6 + \frac{1}{6}(1-p_6) \\
p_2 &= \frac{1}{10}p_6 + \frac{1}{6}(1-p_6) \\
p_3 &= \frac{1}{10}p_6 + \frac{1}{6}(1-p_6) \\
p_4 &= \frac{1}{10}p_6 + \frac{1}{6}(1-p_6) \\
p_5 &= \frac{1}{10}p_6 + \frac{1}{6}(1-p_6) \\
p_6 &= \frac{1}{10}p_6 + \frac{1}{6}(1-p_6)
\end{aligned}
$$

解出$p_2 = p_3 = p_4 = p_5 = p_6 = 5/32$，$p_1 = 7/32$.
熵 $S = 1.783 < \ln 6 = 1.792$. 这个老千系统只让熵变小了0.01.

别小看老千系统

但是，可别小看这0.01的熵差。如果你跟赌王掷了10000次骰子，熵差就变为100。

如果你还记得玻尔兹曼的墓碑，你就明白赌王通过出老千，得到了一个看起来概率约为e^{-100}的“小概率事件”！

上帝的老千

现在让我们回到物理世界，考虑在相空间有多个可能态的微观粒子。

虽然粒子之间可以发生“碰撞”使态发生变化，但“碰撞”受到自然法则的约束：任何碰撞过程能量守恒。

能量守恒——这就是上帝出的老千。

下面让我们通过研究一个粒子数守恒的封闭系统来揭秘上帝的老千的工作原理。

揭秘上帝的老千

- 设处于a,b两态的粒子可以发生碰撞，成为a'，b'两态的粒子。单位时间内$a+b\to a'+b'$发生概率为$\Gamma(a,b;a',b')$。
- 假设微观物理是时间反演对称的，则逆过程$a'+b'\to a+b$的发生概率相等$\Gamma(a',b';a,b)=\Gamma(a,b;a',b')$。
- 设共有N个粒子，考虑粒子在a态出现的概率p_a随时间的演化：

$$\frac{d(Np_a)}{dt}=N^2\sum_{b;a',b'}(p_{a'}p_{b'}-p_ap_b)\Gamma(a,b;a',b')$$

- 把多粒子系统看成“N个重复对象”，其熵随时间的演化：

$$\frac{dS}{dt}=-N\sum_a\frac{dp_a}{dt}(\ln p_a+1)=-N\sum_a\frac{dp_a}{dt}\ln p_a$$

揭秘上帝的老千(续)

- 于是我们得到

$$\frac{dS}{dt} = N^2 \sum_{a,b;a',b'} \Gamma(a,b;a',b')\,(p_a p_b - p_{a'} p_{b'})\ln p_a$$

- 我们当然可以把求和指标a, b互换，则上式变成

$$\frac{dS}{dt} = N^2 \sum_{a,b;a',b'} \Gamma(a,b;a',b')\,(p_a p_b - p_{a'} p_{b'})\ln p_b$$

- 上面两式相加得到

$$\frac{dS}{dt} = \frac{N^2}{2} \sum_{a,b;a',b'} \Gamma(a,b;a',b')\,(p_a p_b - p_{a'} p_{b'})\ln(p_a p_b)$$

- 我们当然可以把a, b和a', b'互换得到：

$$\frac{dS}{dt} = \frac{N^2}{2} \sum_{a,b;a',b'} \Gamma(a',b';a,b)\,(p_{a'} p_{b'} - p_a p_b)\ln(p_{a'} p_{b'})$$

细致平衡

- 把前页的最后两式相加，得到

$$\frac{dS}{dt} = \frac{N^2}{4} \sum_{a,b;a',b'} \Gamma(a,b;a',b')\,(p_a p_b - p_{a'} p_{b'})\ln(\frac{p_a p_b}{p_{a'} p_{b'}})$$

显然，只要任何一个$p_a p_b \neq p_{a'} p_{b'}$上式右边就是正的，熵会持续增大，直到所有的

$$p_a p_b = p_{a'} p_{b'}$$

这称为细致平衡条件。
当细致平衡条件满足时，我们认为该孤立系统达到了热平衡，它的熵达到了最大并不再变化。

热力学温度

热平衡时，所有“碰撞”过程$a + b \to a' + b'$都需要满足细致平衡条件：

$$\ln p_a + \ln p_b = \ln p_{a'} + \ln p_{b'}$$

在微观上，这是一条碰撞过程的守恒律。显然，把一大堆粒子放在一起并不会产生新的微观守恒律，那么它必然是一条已有的守恒律。

热力学温度（续）

$$\ln p_a + \ln p_b = \ln p_{a'} + \ln p_{b'}$$

在微观上有能量守恒定律和三个方向上的动量守恒定律，看来让$\ln p_a$和能量或者动量的任何一个分量成线性关系即可满足条件?

问题是，无论让$\ln p_a$和哪个方向上的动量成线性关系都将破坏微观规律的空间对称性。

热力学温度(续)

$$\ln p_a + \ln p_b = \ln p_{a'} + \ln p_{b'}$$

至此，我们别无选择，只能要求对所有的态a，$\ln p_a$和能量ε_a成线性关系。另外，我们知道态的能量趋向无穷时，在该态出现的概率就趋向于零，所以$\ln p_a$和ε_a之间的是负线性相关。不妨设

$$\ln p_a = -\frac{\varepsilon_a - \mu}{T}$$

$T > 0$和μ为能量量纲的常量。很快，你会发现这样定义的T的数值往往太大，需要重新取一个合适的单位，于是就有了把T换成kT以及之后的热力学温度的一系列故事。

最后，照顾一下克劳修斯的感受

考虑一个单纯加热（无做功）的微小可逆过程:
由

$$dS = -N\sum_a dp_a \ln p_a = \frac{N}{T}\sum(\varepsilon_a - \mu)dp_a = \frac{N}{T}\sum \varepsilon_a dp_a$$

即得

$$đQ = N\sum_a \epsilon_a dp_a = TdS$$

其他更复杂的情况不再深入讨论。

熵的总结

我们以后会学到，能量动量守恒定律是时间和空间平移对称性的结果。那么上面的所有的推导，除了随机初条件（即开始时没有额外限制条件）之外，我们只用到了时间和空间的各种对称性。

正是因为熵的单向性，覆水难收，我们无法抹平记忆，回到过去，于是才会有了时间的方向。

随机初始条件 + 微观世界的时间和空间对称性 $\Rightarrow$ 宏观世界时间的方向

是不是不可思议！

第四周作业(序号接第三周)

- 教材习题2-7
- 教材习题2-22
- 教材习题2-23