高等统计物理

陈志

Email： chenzyn@ustc.edu.cn
Web： http://staff.ustc.edu.cn/~chenzyn/Modern_statistical_physics.html

参考书：
1.**量子统计物理学，杨展如。**
2. 统计物理现代教程 （上册）， L.E. Reichl(雷克)。
3. Equilibrium statisitical physics, 3rd edition, by M. Plischke and B. Bergersen (世图有英文影印版， 中文版)
4. 统计力学（英文第二版）， 黄克孙。
5. 统计力学（中文第三版）， 帕斯里亚等。

主要讲述内容

- 概率论基础和随机过程简介，主方程（master equation），参考雷克的统计物理现代教程第五-六章。
- 杨展如的量子统计物理学第一---九章：

1. 量子统计物理学基础：三种系综，统计算符和密度矩阵；
2. 系综的配分函数，热力学函数的奇异性，经典和量子集团展开；
3. 波色系统，波色-爱因斯坦凝聚，超流性；
4. 费米系统，顺磁性和抗磁性，准粒子和元激发；
5. 相变和临界现象的基本概念：序参量，临界指数，标度律，平均场理论，自发对称破缺；
6. 几种典型的晶格统计模型：Ising模型，XY模型，Potts模型等；
7. 重整化群理论（第八章和第九章部分内容）。

概率论基础和主方程

- 热力学和统计物理：

热力学：从若干经验定律出发，研究由大量微观粒子组成的宏观系统，给出物理量平均值之间的相互关系；

统计物理：从单个粒子的力学运动规律出发，加上统计的假设，通过大量微观粒子（自由度很大的系统）的集体表现来描述宏观物理量的行为，宏观量是相应微观物理量的统计平均值。

- 参考雷克的统计物理现代教程第五-六章。
- 讨论随机理论用于非线性系统涨落现象的动力学研究。
- 随机理论对统计物理在其它领域的应用有很重要的价值。

第一章 基本概率理论

1.1 概率

- 随机事件：在一定条件下，如果一个事件可能发生也可能不发生，这事件称为随机事件。

- 随机事件的概率（发生可能性的大小）：当观测次数N趋于无穷时，某一事件A发生的次数N_A与总观测次数的比值将趋于稳定的极限值—概率P(A):

$$P(A) = \lim_{N\to\infty} \frac{N_A}{N}$$

- # 互斥事件概率的加法原理：

互斥事件：在一次观测中不可能同时发生的事件。

加法原理：对两互斥事件A和B，它们中任意一个出现的概率为：

$$P(A+B) = P(A) + P(B)$$

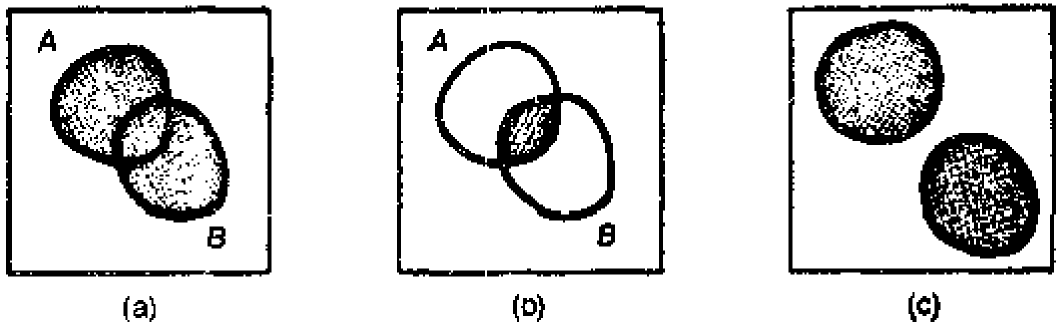
完备性：完备的互斥事件出现的总概率为1：$P(\sum_i A_i) = 1.$

- # 独立概率事件的乘法原理：

独立事件：彼此没有任何关联的事件，即一事件的发生与否与别的事件的发生与否毫不相关。

乘法原理：两独立事件同时或依次发生的概率为 $P(A \cdot B) = P(A)\,P(B)$。

一般情形：集合论的描述



(a) 阴影区为A和B的并$P(A\cup B)$
(b) 阴影区为A和B的交$P(A\cap B)$
(c) A和B互相排斥，没有重叠

- 事件A或事件B或二者同时出现的概率$P(A\cup B)$：
$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$
- 互斥事件：$P(A\cup B)=P(A)+P(B).$
- 若事件A_1，A_2,..., A_m是互斥和完备的，则 $P(\sum_i A_i)=1.$
- 事件A和事件B二者同时出现的概率为$P(A\cap B)$。
- 当且仅当$P(A\cap B)=P(A)P(B)$时，事件A和B是独立的。
- 事件A和B互斥时$P(A\cap B)=0$。

条件概率P(B|A)

- 条件概率P(B|A)是在事件B已出现，而又出现事件A的概率（按雷克书中的定义，但很多书与此相反）。
- 我们有： $$P(B|A) = \frac{P(A \cap B)}{P(B)}.$$
- 由 $P(A \cap B) = P(B \cap A)$ 我们有(Bayes' theorem)：
$$P(A)\,P(A|B) = P(B)\,P(B|A).$$
- 若A和B独立，则P(B|A)=P(A).

1.2 随机变量

- 随机变量：如果一变量X以一定的概率在其样品空间S里取各种可能值，则X称为随机变量。随机变量分为离散型和连续型两种。
- 我们通常以X表示随机变量，以 x_i 和 $f(x_i)$ 表示离散型随机变量的可能取值和相应值的概率（分布函数）；以 x和 $f_X(x)$ 表示连续型随机变量的可能取值和相应值的概率密度。
- 对离散型随机变量，我们有：$f(x_i) \geq 0$ 和 $\sum_i f(x_i) = 1.$
- 对连续型随机变量，我们有：$f_X(x) \geq 0$ 和 $\int f_X(x)\mathrm{d}x = 1.$
- 随机变量的n阶矩(moments): $\langle X^n \rangle = \sum_i x_i^n f(x_i)$ 或 $\langle X^n \rangle = \int \mathrm{d}x\, x^n f_X(x).$
- $\langle X \rangle$ 为平均值，二阶矩与方差(variance)有关：$\mathrm{Var}(x) = \langle X^2 \rangle - \langle X \rangle^2.$
- 矩给出了分布函数的范围和形状信息。
- 对连续型随机变量，若所有的矩 $\langle X^n \rangle$都知道，则概率密度函数被完全确定。特殊地，高斯分布被一阶和二阶矩完全确定（反之也成立）。

为什么?

1.3 特征函数

特征函数 $\phi_X(k)$: 概率密度函数 $f_X(x)$ 的傅立叶变换。

$$\phi_X(k) = \langle \mathrm{e}^{ik\cdot x} \rangle = \int \mathrm{d}x \, \mathrm{e}^{ikx} f_X(x) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle X^n \rangle}{n!}.$$

- 在上面的积分里对k（多次）求导再取k=0，很容易就可得到最后的级数表达式。注意级数展开只当高阶矩足够小时才有意义。
- 由上式可见，特征函数和概率密度函数都由各级矩完全确定。
- 另一个相关的量：累积量(cumulants)---其中$C_n(X)$是第n级累积量

$$\phi_X(k) = \exp\left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X)\right].$$

我们有 $C_1(X) = \langle X \rangle, C_2(X) = \langle X^2 \rangle - \langle X \rangle^2, C_3(X) = \langle X^3 \rangle - 3\langle X^2 \rangle \langle X \rangle + 2\langle X \rangle^3,$ 等等。$C_2(X)$是方差，$C_3(X)$可用于判断分布的对称性, $C_4(X)$常用于判断分布与高斯分布的偏离。

引入特征函数有何好处？

- 可以很容易计算任意两个独立随机变量和$X=X_1+X_2$的密度函数：

x空间:

$$f(x) = \int_{-\infty}^{\infty} \mathrm{d}x_1 f_1(x_1) f_2(x-x_1)$$

k空间：

$$\phi(k) = \phi_1(k)\,\phi_2(k)$$

避免了积分！

- 把对概率密度函数的求导转为乘积（可应用于主方程的求解），大大简化计算！

$$\begin{aligned}
\phi(k) &= \int \mathrm{e}^{ikx} f(x)\,\mathrm{d}x = \iint \mathrm{e}^{ikx} f_1(x_1) f_2(x-x_1)\,\mathrm{d}x_1\mathrm{d}x \\
&= \int f_1(x_1)\,\mathrm{d}x_1 \mathrm{e}^{ikx_1} \int \mathrm{e}^{ik(x-x_1)} f_2(x-x_1)\,\mathrm{d}(x-x_1) \\
&= \phi_1(k)\,\phi_2(k).
\end{aligned}$$

1.4 高维情形

- 联合概率密度：

对两随机变量X和Y，其联合概率密度f(x,y)有：

$$f(x,y) \geq 0, \quad \iint \mathrm{d}x\,\mathrm{d}y\, f(x,y) = 1.$$

- X和Y的协方差定义为：

$$\mathrm{cov}(X,Y) = \iint \mathrm{d}x\,\mathrm{d}y\,(x-\langle X\rangle)(y-\langle Y\rangle)\, f(x,y) = \langle XY\rangle - \langle X\rangle \times \langle Y\rangle.$$

- 若X和Y独立，我们有：

$f(x,y) = f_X(x)\, f_Y(y)$ 和

$$\langle XY\rangle = \langle X\rangle \times \langle Y\rangle \text{ 及 } \mathrm{cov}(X,Y) = 0.$$

其中最后两式的逆命题不一定成立。

1.5 几种常见分布

- 二项式分布：设每次试验有两种结果1和-1，概率分别为p和q。则N次独立试验后，n_1次结果为1的概率为：

$$P_N(n_1) \equiv \frac{N!}{n_1!n_2!} p^{n_1} q^{n_2}.$$

- 高斯(Gaussian)分布：在大N和大pN极限下，二项式分布趋于高斯分布：

$$P_N(n_1) = \frac{1}{\sigma_N \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{(n_1 - \langle n_1 \rangle)^2}{\sigma_N^2}\right\}.$$

其中 $\sigma_N = \sqrt{Npq}$ 。

- 泊松(Poisson)分布：在大N和小p极限下，如$Np = a, a$是一个常数，二项分布趋于泊松分布：

$$P_N(n_1) = \frac{a^{n_1} \mathrm{e}^{-a}}{n_1!}.$$

- Lorentz或Cauchy分布（有时也叫Breit-Wigner分布）：所有的矩都发散。

$$P(x) = \frac{1}{\pi} \frac{\gamma}{(x-a)^2 + \gamma^2}, \quad -\infty < x < \infty.$$

其中Gaussian分布和Cauchy分布是稳定分布。

Levy稳定随机变量和稳定分布

这里我们考察N个iid（independent and identically distributed）随机变量 X_i 的和的分布:

$$S_N = X_1 + X_2 + \cdots + X_N.$$

若 S_N 的分布与单个随机变量 X_i 的线性函数的分布相同，即 $P_{S_N} = P_{c_N X + d_N}$，则称这个分布是稳定（stable）的。特别地，若 $d_N = 0$ 则分布是严格稳定的（strictly stable）。上面提到的 Gaussian分布和Cauchy分布就是稳定分布的例子。

稳定分布的重要性：不论初始随机变量 X_i 的分布是什么，当N趋于无穷大时，X_i的和或平均值满足的分布一般将趋近于稳定分布，而后者正是我们在统计物理的研究中经常观测的量。

对稳定分布，我们有：$c_N = N^{1/\alpha},\ 0 < \alpha \leq 2.$

- 对任意满足α≠1的稳定分布P(x)，我们总能找到一个常数μ使得P(x-μ)是严格稳定的；
- 当α=2 我们发现分布为Gaussian分布（对Cauchy分布我们有α=1），若α<2则方差和更高阶的矩发散（若α<1平均值也发散）！ Gaussian分布是唯一有有限方差的稳定分布。

这部分解释了为什么在自然界里我们发现很多（平均值）的分布是Gaussian分布。

稳定分布的特征函数一般可以写为:

$$\chi(t)=\langle e^{itX}\rangle=\exp\left(ita-|bt|^{\alpha}\left[1-i\beta\operatorname{sgn}(t)u_{\alpha}(t)\right]\right),\ t\in R,\ \text{where } u_{\alpha}(t)=\begin{cases}\tan\left(\frac{\pi\alpha}{2}\right), & \alpha\neq 1\\ -\frac{2}{\pi}\ln(|t|), & \alpha=1\end{cases}$$

这里shift parameter a可以是任意实数，$b\in(0,\infty)$, $\alpha\in(0,2]$, $\beta\in[-1,1]$.
b是scale parameter, β是skewness parameter.

练习：验证两个稳定分布的和的分布仍是稳定分布，并有：

$$\begin{aligned} a &= a_1+a_2,\\ |b| &= (|b_1|^{\alpha}+|b_2|^{\alpha})^{1/\alpha}\\ \beta &= \frac{\beta_1|b_1|^{\alpha}+\beta_2|b_2|^{\alpha}}{|b|^{\alpha}} \end{aligned}$$

熵和相对熵

- 对任一个密度函数为ρ_X 的随机变量X我们可以定义其熵为:

$$S[\rho_X] = -k \int \mathrm{d}x \, \rho_X(x) \ln\left(\frac{\rho_X(x)}{\rho_0}\right).$$

这里我们引入参数 ρ_0 使得 $S \geq 0$ 能够成立。以后可以看到这里定义的熵和物理中的熵密切相关。

- 对任两个密度函数 *p*和*q*我们可以引入相对熵的概念：

$$S[p||q] = -k \int \mathrm{d}x \, p(x) \ln\left(\frac{p(x)}{q(x)}\right).$$

利用不等式：$\ln z \leq z - 1$可以很容易证明（作为练习）： $S[p||q] \leq 0.$

- 熵的广延性（练习）：

对任两个随机变量X_1和X_2，设它们的密度函数分别为 $\rho_{X_1}(x)$ 和 $\rho_{X_2}(x)$ ，两个随机变量的联合概率密度函数为 $\rho_{X_1,X_2}(x_1,x_2)$，利用相对熵$S[\rho_{X_1,X_2}||\rho_{X_1}\rho_{X_2}]$ 可以证明：$S_{12} \leq S_1 + S_2$. 等号当且仅当随机变量X_1和X_2独立时成立。

- 最大熵原理：

如果密度函数$\rho(x)$有最大熵，且满足额外的限制条件：

$$\langle g_i(X)\rangle \equiv \int \mathrm{d}x\, g_i(x)\rho(x) = \eta_i,\ i=1,...,n.$$

这里 $g_i(x), \eta_i$ 均已知，那么$\rho(x)$的形式必为： $\rho(x) = \frac{1}{A}\exp\left[-\sum_{i=1}^{n}\lambda_i g_i(x)\right],$

这个性质可利用相对熵加以证明，具体证明作为练习。

统计物理中的对应：平衡态时系统的熵最大，如正则系综，巨正则系综等。

熵的定义是唯一的吗？ ---- 不是的

通常定义的熵对任两个独立随机变量A和B具有广延性(见前页)，但我们也可以定义不具有广延性的熵。比如Tsallis最近定义的熵(当 $q=1$ 时即回到前面的定义)：

$$S_q = \frac{1-\int \mathrm{d}\Omega\,[p(x)]^q}{q-1},\quad \text{where} \int p(x)\,\mathrm{d}\Omega = 1, \text{ and very often } 0\le q\le 2.$$

这里$\mathrm{d}\Omega$是体积元。可以发现对这个熵有：$S_q(A+B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B).$

当$q\neq 1$ 时广延性消失。这个熵当系统有长程相互作用时比较有用。

1.6 中心极限定理和大数定律

- 考虑随机变量X的N次独立测量平均的随机变量Y，其概率密度函数为 $f_Y(y_N-\langle X\rangle)$。$f_Y(y_N-\langle X\rangle)$ 可由X的密度函数获得但表达式很复杂，为简单记我们考虑其特征函数：

$$\begin{aligned}\Phi(k) &= \int \mathrm{e}^{ik(y_N-\langle X\rangle)} f_Y(y_N-\langle X\rangle)\mathrm{d}y_N \\ &= \int \Big[\mathrm{e}^{i(k/N)((x_1-\langle X\rangle)+(x_2-\langle X\rangle)+\cdots+(x_N-\langle X\rangle))} f_{X-\langle X\rangle}(x_1-\langle X\rangle) \\ &\quad \times f_{X-\langle X\rangle}(x_2-\langle X\rangle)\cdots f_{X-\langle X\rangle}(x_N-\langle X\rangle)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_N\Big] = [\phi(k/N)]^N .\end{aligned}$$

- 记 $\sigma^2=\langle X^2\rangle-\langle X\rangle^2$, 则

$$\begin{aligned}\phi\left(\frac{k}{N}\right) &= \int \mathrm{e}^{i(k/N)(x_1-\langle X\rangle)} f_{X-\langle X\rangle}(x_1-\langle X\rangle)\mathrm{d}x_1 \\ &= 1-\frac{1}{2}\frac{k^2}{N^2}\sigma^2+\cdots .\end{aligned}$$

若当$x_1 \to \infty$时函数$f_X(x_1)$向零趋近得够快，则矩是有限的，并有：

$$\Phi(k) = \left[1 - \frac{1}{2}\frac{k^2}{N^2}\sigma^2 + o\left(\frac{k^3}{N^3}\right)\right]^N \xrightarrow[N\to\infty]{} \mathrm{e}^{-k^2\sigma^2/(2N)}.$$

于是概率密度$f_Y(y_N - \langle X\rangle)$变为：

$$\begin{aligned} f_Y(y_N - \langle X\rangle) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{-ik(y_N - \langle X\rangle)}\Phi(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{-ik(y_N - \langle X\rangle)}\mathrm{e}^{-k^2\sigma^2/(2N)} \\ &= \sqrt{\frac{N}{2\pi}}\frac{1}{\sigma}\,\mathrm{e}^{-N(y_N - \langle X\rangle)^2/(2\sigma^2)}. \end{aligned}$$

因此，不管$f_X(x)$的形式如何，只要它有有限的矩，X的大量独立测量的平均值的分布是中心在$\langle X\rangle$的高斯分布，其方差是X的概率密度的方差的1/*N*。这就是中心极限定理。

大数定律

大数定律：当N→∞时，y_N偏离$\langle X\rangle$的概率趋于零。

我们利用Chebyshev's inequality: $P(|x-\langle X\rangle| \geq \varepsilon) \leq \dfrac{\sigma_X^2}{\varepsilon^2}$.

（由 $\sigma_X^2 = \int_{-\infty}^{\infty} \mathrm{d}x\,(x-\langle X\rangle)^2 f_X(x) \geq \left[\int_{-\infty}^{\langle X\rangle-\varepsilon} + \int_{\langle X\rangle+\varepsilon}^{\infty}\right] \mathrm{d}x\,(x-\langle X\rangle)^2 f_X(x)$ 易得）

及y_N和X的方差的关系及 $\langle Y_N\rangle = \langle X\rangle$ 很容易发现：

$$P(|y_N - \langle X\rangle| \geq \varepsilon) \leq \frac{\sigma_{Y_N}^2}{\varepsilon^2} = \frac{\sigma_X^2}{N\varepsilon^2}.$$

故当N→∞时，若 σ_X^2 有限，则

$$\lim_{N\to\infty} P(|y_N - \langle X\rangle| \geq \varepsilon) \to 0.$$