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)。

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一-八章：

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。

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化学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：

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

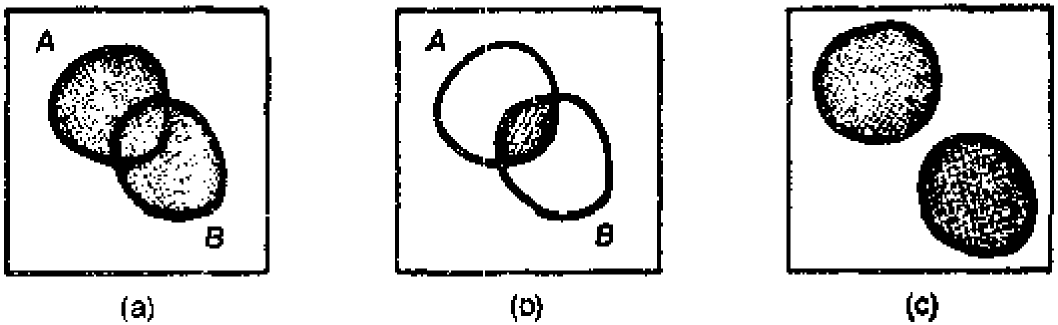
完备性：完备的互斥事件出现的总概率为1：$P(\sum_i A_i)=1.$

- 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 $\sum_i P(A_i) = 1$.
- 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：$P(B \mid A) = \frac{P(A \cap B)}{P(B)}$.
- 由 $P(A \cap B) = P(B \cap A)$ 我们有(Bayes' theorem)：

$$P(A)P(A \mid B) = P(B)P(B \mid A).$$

- 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：

$$f(x_i) \geq 0 \text{和} \sum_i f(x_i) = 1.$$

- 对连续型随机变量，我们有：

$$f_X(x) \geq 0 \text{和} \int f_X(x)dx = 1.$$

- 随机变量的n阶矩(moments):

$$\langle X^n \rangle = \sum_i x_i{}^n f(x_i) \text{或} \langle X^n \rangle = \int dx x^n f_X(x).$$

- 〈X〉为平均值，二阶矩与方差(variance)有关：$Var(x) = \langle x^2 \rangle - \langle x \rangle^2$.
- 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。

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

- 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 $C_1(X) = \langle X\rangle, C_2(X) = \langle X^2\rangle - \langle X\rangle^2$,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？

- 可以很容易计算两个随机变量和$X=X_1+X_2$的密度函数：

x空间:

$$f(x) = \int_{-\infty}^{\infty} dx_1 f_1(x_1) f_2(x - x_1)$$

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！

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：

$$f(x, y) \geq 0, \qquad \iint dxdy f(x, y) = 1.$$

- X和Y的协方差定义为：

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

- 若X和Y独立，我们有：

$f(x, y) = f_X(x) f_Y(y)$和

$\langle XY \rangle = \langle X \rangle \times \langle Y \rangle$ 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此试验后，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：

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

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

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

其中Gaussian分布和Cauchy分布是稳定分布（即N个iid随机变量和的分布与单个随机变量线性函数的分布相同），其中Gaussian分布是有收敛矩的唯一的稳定分布。

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：

$$\Phi(k) = \int e^{ik(y_N - \langle X\rangle)} f_Y(y_N - \langle X\rangle) dy_N$$
$$= \int [e^{i(k/N)((x_1-\langle X\rangle)+(x_2-\langle X\rangle)+\cdots+(x_N-\langle X\rangle))} f_X(x_1)$$
$$\times f_X(x_2)\cdots f_X(x_N) dx_1 dx_2 \cdots dx_N] = [\phi(k/N)]^N$$

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

$$\phi\left(\frac{k}{N}\right) = \int e^{i(k/N)(x_1-\langle X\rangle)} f_X(x_1) dx_1$$
$$= 1 - \frac{1}{2}\frac{k^2}{N^2}\sigma^2 + \cdots.$$

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：

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

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

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

因此，不管$f_X(x)$的形式如何，只要$f(x)$有有限的矩，X的大量独立测量的平均值的分布是中心在〈X〉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<x>的概率趋于零。

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

（注意 $\sigma_X^2 = \int_{-\infty}^{\infty} dx\left(x-\langle x\rangle\right)^2 f_X(x) \geq \int_{-\infty}^{\langle x\rangle-\varepsilon} dx\left(x-\langle x\rangle\right)^2 f_X(x) + \int_{\langle x\rangle+\varepsilon}^{\infty} dx\left(x-\langle x\rangle\right)^2 f_X(x)$ 即得证）

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

$$P(|y_N - \langle x\rangle| \geq \varepsilon) \leq \sigma_{y_N}^2 / \varepsilon^2 = \sigma_X^2 / (N\varepsilon^2)$$

故当N→∞时，若σ_X^2有限，则

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

第二章 主方程(Master equation)

- 这里我们研究概率分布随时间的演化。
- 随机过程：与时间有关的随机变量(time-dependent random variable)
- 我们只考虑仅有短程记忆的过程 – 马尔科夫过程(Markov process)，该过程的时间演化方程就是主方程。
- 主方程是统计物理里最重要的方程之一，它几乎是普遍适用的，并广泛地被应用于化学，生物学，人口动力学，布朗运动，流体，半导体，金融等问题。

2.1 主方程的推导

（I）一般情形

对于随机变量Y的概率密度，将采用以下的记号来表示：

$P_1(y_1, t) \equiv$ (随机变量Y在t_1时刻取y_1值的概率);

$P_2(y_1, t_1; y_2, t_2) \equiv$ (随机变量Y在t_1时刻取y_1值, 在t_2时刻取y_2值的联合概率);

$P_n(y_1, t_1; y_2, t_2; \cdots; y_n, t_n) \equiv$ (随机变量Y在t_1时刻取y_1值, 在t_2时刻取y_2值, $\cdots$, t_n时刻取y_n值的联合概率)。

联合概率密度是正的： $P_n \geq 0$;

它们可以被约化：

$$\int P_n(y_1, t_1; y_2, t_2; \cdots; y_n, t_n)dy_n = P_{n-1}(y_1, t_1; y_2, t_2; \cdots; y_{n-1}, t_{n-1});$$

并且是归一化的： $\int P_1(y_1, t_1)dy_1 = 1.$

- 随机变量与时间有关的矩(表征随机变量在不同时刻的值之间的相关):

$$\langle y_1(t_1)y_2(t_2)\cdots y_n(t_n)\rangle$$
$$= \int\cdots\int y_1y_2\cdots y_nP_n(y_1,t_1;y_2,t_2;\cdots;y_n,t_n)dy_1\cdots dy_n.$$

- 平稳过程：如果一个过程对一切n与τ都有：

$$P_n(y_1,t_1;y_2,t_2;\cdots;y_n,t_n) = P_n(y_1,t_1+\tau;y_2,t_2+\tau;\cdots;y_n,t_n+\tau)$$

在平衡时，所有物理过程都是平稳的。对一个平稳过程，有：

$$P_1(y_1,t)=P_1(y_1),\qquad P_2(y_1,t_1;y_2,t_2) = P_2(y_1,0;y_2,t_2-t_1),$$

而 $\langle y_1(t_1)y_2(t_2)\rangle$ 只依赖于 $|t_1-t_2|$ - 时间差的绝对值。

- 条件概率：

$P_{1|1}(y_1,t_1 \mid y_2,t_2) \equiv$ (在 t_1时刻取y_1值的随机变量Y

在 t_2时刻取y_2值的概率密度)

它由如下恒等式来定义:

$$P_1(y_1,t_1)P_{1|1}(y_1,t_1 \mid y_2,t_2) = P_2(y_1,t_1;y_2,t_2)$$

- 不同时刻概率密度之间的关系（上式对y_1积分）：

$$P_1(y_2,t_2) = \int P_1(y_1,t_1)P_{1|1}(y_1,t_1 \mid y_2,t_2)dy_1$$

- 联合条件概率密度：$P_{k|l}(y_1,t_1;\cdots;y_k,t_k \mid y_{k+1},t_{k+1};\cdots;y_{k+l},t_{k+l})$
$=$ (固定$(y_1,t_1;\cdots;y_k,t_k)$时，随机变量Y具有值$(y_{k+1},t_{k+1};\cdots;y_{k+l},t_{k+l})$的联合概率密度).
$$= \frac{P_{k+l}(y_1,t_1;\cdots;y_k,t_k;y_{k+1},t_{k+1};\cdots;y_{k+l},t_{k+l})}{P_k(y_1,t_1;\cdots;y_k,t_k)}.$$

（II）马尔科夫过程(Markov process)

对马尔科夫过程我们有(其中$t_1<t_2<\ldots<t_n$)：

$$P_{n-1|1}(y_1,t_1;y_2,t_2;\cdots;y_{n-1},t_{n-1} \mid y_n,t_n) = P_{1|1}(y_{n-1},t_{n-1} \mid y_n,t_n).$$

即t_n时刻取y_n的条件概率完全由t_{n-1}时刻y_{n-1}的值确定。马尔科夫过程完全由 $P_1(y,t)$ 和 $P_{1|1}(y_1,t_1 | y_2,t_2)$ [转移概率]两个函数确定。例如：

$$P_3(y_1,t_1;y_2,t_2;y_3,t_3) = P_2(y_1,t_1;y_2,t_2)P_{2|1}(y_1,t_1;y_2,t_2 | y_3,t_3)$$
$$= P_1(y_1,t_1)P_{1|1}(y_1,t_1 | y_2,t_2)P_{1|1}(y_2,t_2 | y_3,t_3)$$

对y_2积分，容易得到（Chapman-Kolmogorov方程）：

$$P_{1|1}(y_1,t_1 | y_3,t_3) = \int dy_2 P_{1|1}(y_1,t_1 | y_2,t_2)P_{1|1}(y_2,t_2 | y_3,t_3)$$

*Chapman-Kolmogorov*方程的重要性：告诉我们对马尔科夫过程来说两个相继步骤的转移概率是两个单个步骤转移概率的**乘积**，而且相继的步骤是统计**独立**的。

(III) 主方程(Master equation)

在时刻$t_1+\tau$（ τ 是一个非常小的正数），由定义我们有（以连续变量为例）：

$$P_1(y_2, t_1 + \tau) = \int P_1(y_1, t_1) P_{1|1}(y_1, t_1 \mid y_2, t_1 + \tau) dy_1 \qquad (*1)$$

当$\tau=0$时，由（*1）得： $P_{1|1}(y_1, t_1 \mid y_2, t_1) = \delta(y_1 - y_2)$ （*2）

$P_1(y, t)$ 的时间导数为： $\dfrac{dP_1(y,t)}{dt} = \lim\limits_{\tau \to 0} \dfrac{P_1(y, t+\tau) - P_1(y,t)}{\tau}$,

计算（*1）的时间导数我们必须考虑：$\lim\limits_{\tau \to 0} \dfrac{P_{1|1}(y_1, t \mid y_2, t+\tau) - P_{1|1}(y_1, t \mid y_2, t)}{\tau}$

这里我们定义 $W_{t_1}(y_1, y_2)$是系统在时间间隔 $t_1 \to t_1 + \tau$ 内，从态y_1变到态y_2的单位时间的条件（或转变）概率密度。因此

在时间 $t_1 \to t_1 + \tau$内，从态y_1转变到态y_2的概率密度为 $\tau W_{t_1}(y_1, y_2)$；
在时间τ内不转变的概率为 $\left(1 - \tau \int dy W_{t_1}(y_1, y)\right)\delta(y_1 - y_2)$。所以有：

$$P_{1|1}(y_1, t_1 \mid y_2, t_1 + \tau) = \tau W_{t_1}(y_1, y_2) + \left(1 - \tau \int dy W_{t_1}(y_1, y)\right)\delta(y_1 - y_2) \qquad (*3)$$

由（*1-3）我们发现：

$$\frac{\partial P_1(y_2, t)}{\partial t} = \int dy_1 \{W(y_1, y_2) P_1(y_1, t) - W(y_2, y_1) P_1(y_2, t)\}. \qquad (*4)$$

这就是主方程。

(IV) 福克-普朗克(Fokker-Planck)方程

当y是一个**连续变量**，而且y的改变以**小跳跃**的方式发生时，我们可导出 $P_1(y,t)$ 的偏微分方程---福克-普朗克方程。这时，我们有

$$W(y';y)=W(y';y-y')\equiv W(y';\xi),$$

其中$\xi=y-y'$是跳跃的大小。于是主方程变为

$$\frac{\partial P_1(y,t)}{\partial t}=\int d\xi W(y-\xi;\xi)P_1(y-\xi,t)-P_1(y,t)\int d\xi W(y;-\xi).$$

由于转移概率 $W(y',y)$ 将随ξ的增大而迅速减小，我们把W按ξ的幂次展开：

$$\frac{\partial P_1(y,t)}{\partial t}=\int d\xi W(y;\xi)P_1(y,t)-\int d\xi\xi\frac{\partial}{\partial y}W(y;\xi)P_1(y,t)$$

$$+\frac{1}{2}\int d\xi\xi^2\frac{\partial^2}{\partial y^2}W(y;\xi)P_1(y,t)+\cdots-P_1(y,t)\int d\xi W(y;-\xi).$$

上式右边第一项和最后一项可消去，因此得到：

$$\frac{\partial P_1(y,t)}{\partial t}==-\frac{\partial}{\partial y}[\alpha_1(y)P_1(y,t)]+\frac{1}{2}\frac{\partial^2}{\partial y^2}[\alpha_2(y)P_1(y,t)],\qquad (+)$$

这就是福克-普朗克(Fokker-Planck)方程。

其中$\alpha_n(y)$是第n级跃变矩：

$$\alpha_n(y)=\int d\xi\xi^n W(y;\xi).$$

2.2 马尔科夫链(Markov chain)

- 马尔科夫链：是马尔科夫过程的一个例子，是在离散时刻出现的离散随机变量Y取值之间的转移。
- 设Y可取值Y={y_1,y_2,..,y_l}，基本时间间隔为1，从t=0到t=1我们有：

$$P_1(y_i,1) = \sum_{j=1}^{l} P_1(y_j,0) P_{1|1}(y_j,0 \mid y_i,1),$$

- 引入$Q_{ji} \equiv P_{1|1}(y_j,0 \mid y_i,1)$，我们可把上式改写为矩阵方程：

$$\mathrm{P}(1) = \mathrm{P}(0)\mathbf{Q}.$$

- 在s时刻，我们有：$\mathrm{P}(s) = \mathrm{P}(0)\mathbf{Q}^s.$
- **P**(s) 在s很大时的行为依赖于转移矩阵的结构。若**Q**的某个幂次的全部元素都是正的（正则矩阵），则**P**(s)趋向唯一的确定的定态。

2.3无规行走和扩散方程

考虑一个粒子在x轴上运动，且各步行走是统计独立的。设步长为l，步间时间为τ，n=0，±1，±2,…为粒子的绝对位置，则有：

$$P_1(n_2 l, s\tau) = \sum_{n_1} P_1(n_1 l, (s-1)\tau) P_{1|1}(n_1 l, (s-1)\tau \mid n_2 l, s\tau)$$

若粒子向左向右运动的概率均为1/2，则

$$P_{1|1}(n_1 l, (s-1)\tau \mid n_2 l, s\tau) = \frac{1}{2}\delta_{n_2, n_1+1} + \frac{1}{2}\delta_{n_2, n_1-1}$$

原方程可简化为：

$$P_1(nl, s\tau) = \frac{1}{2} P_1((n-1)l, (s-1)\tau) + \frac{1}{2} P_1((n+1)l, (s-1)\tau).$$

把上式写为求导的形式，我们有：

$$\frac{P_1(nl, s\tau) - P_1(nl, (s-1)\tau)}{\tau} = \frac{l^2}{2\tau} \frac{P_1((n-1)l, (s-1)\tau) + P_1((n+1)l, (s-1)\tau) - 2P_1(nl, (s-1)\tau)}{l^2}.$$

令$x = nl, t = s\tau$, 并在$D \equiv l^2/2\tau$为有限的条件下取极限$l \to 0, \tau \to 0$, 便可得到扩散方程：

$$\frac{\partial P_1(x,t)}{\partial t} = D\frac{\partial^2 P_1(x,t)}{\partial x^2}$$

假定初始时刻$P_1(x,0) = \delta(x)$, 并引入$P_1(x,t)$的傅里叶变换（特征函数），扩散方程可变为：

$$\frac{\partial \tilde{P}_1(k,t)}{\partial t} = -Dk^2 \tilde{P}_1(k,t).$$

该方程的解为：$\tilde{P}_1(k,t) = e^{-Dk^2 t}$，再取逆变换，可得：

$$P_1(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}.$$

这是粒子在t=0从x=0出发，到t时刻于x点找到它的概率。

2.4 生灭过程

- 生灭过程：在一个时刻只能进行一步转移。
- 我们这里处理一个可用生成函数严格求解的情形。

考虑t时刻有m个细菌的一个群体：

(i)在时间$t \to t+\Delta t$内死亡一个细菌的概率为$\mu_m \Delta t$;

(ii)在时间$t \to t+\Delta t$内出生一个细菌的概率为$\lambda_m \Delta t$;

(iii)在时间$t \to t+\Delta t$内细菌数目不变的概率为$(1-\lambda_m \Delta t-\mu_m \Delta t)$;

(iv)在时间$t \to t+\Delta t$内出生或死亡数超过1的概率为零。

于是有

$$P_{1|1}(m,t \mid n,t+\Delta t)=(1-\lambda_m \Delta t-\mu_m \Delta t)\delta_{n,m}$$
$$+(\lambda_m \delta_{n,m+1}+\mu_m \delta_{n,m-1})\Delta t+\cdots.$$

再假定生和灭的概率正比于现存细菌数，则有 $\lambda_m = m\lambda$ 和 $\mu_m = m\mu$，按（2.1）同样的步骤，我们可得线性生灭过程的主方程：

$$\frac{\partial P_1(n,t)}{\partial t}=(n-1)\lambda P_1(n-1,t)+(n+1)\mu P_1(n+1,t)-(n\lambda+n\mu)P_1(n,t).$$

对依赖于离散随机变量的主方程的求解：生成函数法

生成函数可写为：$F(z,t)=\sum_{n=-\infty}^{\infty}P(n,t)z^{n}$

对z求导后令z--->1,可得到随机变量n的各阶距：

$$<n>=\lim_{z\to 1}\frac{\partial}{\partial z}F(z,t)=\sum_{n=-\infty}^{\infty}nP(n,t)$$

$$<n^{2}>-<n>=\lim_{z\to 1}\left(\frac{\partial}{\partial z}\right)^{2}F(z,t)=\sum_{n=-\infty}^{\infty}(n^{2}-n)P(n,t)$$

一般地，我们有：$\left(z\frac{\partial}{\partial z}\right)^{k}F(z,t)=\sum_{n=-\infty}^{\infty}P(n,t)z^{n}n^{k},$

因此对一般的多项式函数r(n), 我们有：

$$r\left(z\frac{\partial}{\partial z}\right)F(z,t)=\sum_{n=-\infty}^{\infty}P(n,t)z^{n}r(n),$$

而且：$\sum_{n=-\infty}^{\infty}\left(r(n+1)P(n+1,t)-r(n)P(n,t)\right)z^{n}=\left(\frac{1}{z}-1\right)r\left(z\frac{\partial}{\partial z}\right)F(z,t).$

把以上表达式带入到主方程中并注意 $\sum_n r(n+k)P(n+k,t) = \sum_n r(n)P(n,t)$，
我们有：

$$\frac{\partial F}{\partial t} = (z-1)(\lambda z-\mu)\frac{\partial F}{\partial z}, \qquad (**)$$

因此方程(**)和主方程是等价的，我们只需解方程(**)。

容易发现(**)和俩方程 $dF = \frac{\partial F}{\partial z}dz + \frac{\partial F}{\partial t}dt = 0$ 及 $\frac{dt}{1} = \frac{-dz}{(z-1)(\lambda z-\mu)}$ 等价。

从dF=0我们发现F(z,t)=C2，由 $\frac{dt}{1} = \frac{-dz}{(z-1)(\lambda z-\mu)} \Rightarrow C_1 = \frac{(z-1)}{(\lambda z-\mu)}e^{-(\mu-\lambda)t}$,

故一般解为 $F(z,t) = F(1/C_1) = F\left(\frac{(\lambda z-\mu)}{(z-1)}e^{(\mu-\lambda)t}\right) = C_2.$

设t=0时，细菌数目为m，则 $F\left(\frac{(\lambda z-\mu)}{(z-1)}\right) = F(z,0) = \sum_n P_1(n,0)z^n = z^m.$

故若 $u = (z-1)^{-1}(\lambda z-\mu)$, 则

$$F(u) = \left(\frac{\mu-u}{\lambda-u}\right)^m,$$

结果我们求得

$$F(z,t) = \left[\frac{(\mu z-\mu)e^{(\lambda-\mu)t} - \lambda z+\mu}{(\lambda z-\lambda)e^{(\lambda-\mu)t} - \lambda z+\mu}\right]^m.$$

2.5 离散平稳马尔科夫过程的普遍解

- 对离散平稳马尔科夫过程，Chapman - Kolmogorov 方程变为：

  $$P_{1|1}(y_1 \mid y_3, t) = \int P_{1|1}(y_1 \mid y, t_0) P_{1|1}(y, t_0 \mid y_3, t - t_0) dy,$$

  这里 $t = t_3 - t_1$ 和 $t_0 = t_2 - t_1$.

- 考虑离散随机变量 $y = n\Delta$ 和离散时间 $t = s\tau$, 其中 n 和 τ 是整数，这时我们得到了一个马尔科夫链，我们有

  $$P_{1|1}(m \mid n; s) = \sum_k P_{1|1}(m \mid k; s-1) P_{1|1}(k \mid n; 1).$$

  这里 $P_{1|1}(k \mid n; 1)$ 是系统处于 k 态时下一步跳到 n 态的条件概率，它包包含了系统转移机制的一切必要信息。$P_{1|1}(k \mid n; 1)$ 组成了矩阵 $\mathbf{Q}$ 的分量：$P_{1|1}(k \mid n; 1) \equiv Q_{kn}$，由（2.2）节我们并有

  $$P_{1|1}(k \mid n; s) \equiv (\mathbf{Q}^s)_{kn}.$$

- 由概率和条件概率的定义我们还有：

  $$P_1(n) = \sum_m P_1(m) P_{1|1}(m \mid n; s),$$ 即 $\mathbf{P} = \mathbf{PQ}$, 和

  $$\sum_n P_{1|1}(m \mid n; s) = 1.$$

转移矩阵Q

lxl转移矩阵Q一般不是对称阵，因而其左，右本征矢量不同。其左本征矢量问题可写为：

$$\lambda_i x_{im} = \sum_{n=1}^{l} x_{in} Q_{nm},$$

右本征矢量问题可写为：
$$\lambda_j y_{mj} = \sum_{n=1}^{l} Q_{mn} y_{nj},$$

其中**λ**是方程 det|Q-Iλ|=0的解。由以上两式可以证明：

- 正交归一性：即 $\sum_m x_{im} y_{mj} = \sum_m y_{im} x_{mj} = \delta_{ij}.$
- Q可以用其左，右本征矢展开：$Q_{nm} = \sum_i \lambda_i y_{ni} x_{im}$. 因此我们有
  $$P_{1|1}(m \mid n; s) = \sum_i \lambda_i^s y_{mi} x_{in}.$$
- **Q**至少有一本征值为1，且 $|\lambda_i| \le 1.$
  若所有 $|\lambda_i| \le \varepsilon < 1$，由上可知 $P_{1|1}(m \mid n; s) \le \varepsilon^s \sum_i y_{mi} x_{in} \le \varepsilon^s,$
  则对足够大的s方程 $\sum_n P_{1|1}(m \mid n; s) = 1$ 不成立，这不可能。

  再由**P=PQ**及λi**X**i=**X**i**Q**和λi**Y**i=**QY**i，易得 $PY_i = PQY_i = \lambda_i PY_i \Rightarrow (\lambda_i - 1) PY_i = 0.$
  因Yi构成完备本征矢，若所有λi≠1则对所有i均有PYi=0，这不可能。故存在i使得 $\lambda_i = 1.$
  $|\lambda_i| > 1$ 的情形不成立可由Perron–Frobenius theorem得出，雷克书中的简单证明是错误的，该定理的描述可见http://en.wikipedia.org/wiki/Perron-Frobenius_theorem。

- 对正则转移矩阵，若Q只有一个本征值 $\lambda_1 = 1$, 则

$$\lim_{s\to\infty} P_{1|1}(m \mid n; s) = \lim_{s\to\infty}\left(y_{m1} x_{1n} + \sum_{i\neq 1} \lambda_i^s y_{mi} x_{in} \right) = y_{m1} x_{1n}.$$

2.6 近似方法--- Ω展开
（I）简单例子：一维无规行走

考虑一个有边界条件的一维无规行走，其主方程为：

$$\partial P_1(n,t)/\partial t = \alpha P_1(n+1,t) + \beta P_1(n-1,t) - (\alpha+\beta)P_1(n,t).$$

这里-L≤n≤L而且L>>1，因而系统大小Ω=2L+1>>1。我们引入：x=n/L,
并记ρ(x,t)=P_1(n,t),主方程可改写为：

$$\partial\rho(x,t)/\partial t = \alpha\rho(x+1/L,t) + \beta\rho(x-1/L,t) - (\alpha+\beta)\rho(x,t).$$

由于1/L是小量，我们可以把 $\rho(x \pm 1/L, t)$ 对1/L展开：

$$\frac{\partial\rho(x,t)}{\partial t} = (\alpha-\beta)\frac{1}{L}\frac{\partial}{\partial x}\rho(x,t) + \frac{(\alpha+\beta)}{2}\frac{1}{L^2}\frac{\partial^2}{\partial x^2}\rho(x,t) + O(\frac{1}{L^3}).$$

- 情形1：α=β: 这时上式右边第一项1/L项消失。为简单记我们令α=1并记 $\tau = t/L^2$. 这样重新标度后我们有：

$$\frac{\partial\rho(x,\tau)}{\partial\tau} = \frac{\partial^2}{\partial x^2}\rho(x,\tau)$$

  这是扩散方程(Fokker-Planck方程)。
- 情形2：α≠β: 这时只用考虑主方程右边第一项。令τ=t/L我们有：

$$\frac{\partial\rho(x,\tau)}{\partial\tau} = (\alpha-\beta)\frac{\partial}{\partial x}\rho(x,\tau).$$

  这是一个有向无规行走且x(τ)满足： $\dot{x}(\tau) = -(\alpha-\beta).$

(II) 一般情形

这里考虑连续时间和离散随机变量的主方程并假定转移率W与时间无关，这样主方程可写为：

$$\frac{\partial P_1(n,t)}{\partial t} = \sum_m \{W(m,n)P_1(m,t) - W(n,m)P_1(n,t)\}.$$

类似于连续随机变量的情形我们可以定义跃变矩：

$$a_p(n) \equiv \sum_m (m-n)^p W(n,m).$$

跃变矩是随机变量n的方程。我们一般感兴趣的是随机变量n及其各级矩的运动方程，这些已知的话系统的性质就基本清楚了。

- $\langle n(t)\rangle$的运动方程：在主方程两边乘以n并对n求和，在对右边第一项作交换n<->m后我们获得：

$$\frac{\partial \langle n(t)\rangle}{\partial t} = \sum_{m,n} (m-n)W(n,m)P_1(n,t) = \langle a_1(n)\rangle.$$

- $\langle n^2(t)\rangle$的运动方程：在主方程两边乘以n^2并对n求和，在对右边第一项作交换n<->m后我们获得：

$$\frac{\partial \langle n^2(t)\rangle}{\partial t} = \sum_{m,n} (m^2-n^2)W(n,m)P_1(n,t) = \langle a_2(n)\rangle + 2\langle na_1(n)\rangle.$$

因此不用解主方程，通过转移率W(n,m)我们就可以得到系统的大量信息。

近似：W对系统参量Ω的展开

对大系统，我们可以把W对表征系统大小的参量Ω做展开（因1/ Ω 是一个小量），并将其带入到主方程中，获得一个近似的主方程，这个方程的解可能对系统的性质做出较好的描述。

在转移率中重要的参量是密度m/ Ω和步长Δn=n-m。因此我们把W(m,n)展开为：

$$W(m,n) = f(\Omega)\left[\omega_0\left(\frac{m}{\Omega},\Delta n\right) + \frac{1}{\Omega}\omega_1\left(\frac{m}{\Omega},\Delta n\right) + \cdots\right],$$

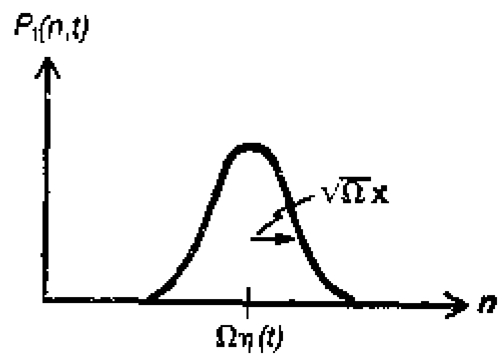
这里f(Ω)是Ω的任意函数。对大Ω我们略去上式中的高阶项并带入到主方程中，得：

$$\frac{\partial P_1(n,t)}{\partial t} = f(\Omega)\sum_{\Delta n}\left\{\omega_0\left(\frac{n-\Delta n}{\Omega},\Delta n\right)P_1(n-\Delta n,t) - \omega_0\left(\frac{n}{\Omega},-\Delta n\right)P_1(n,t)\right\}.$$

对大量独立客体的行为，根据中心极限定理我们知道$\langle n(t)\rangle \propto \Omega$，宽度σn正比于$\sqrt{\Omega}$. 于是我们可以把n在其平均值附近展开：

$$n(t) = \Omega\eta(t) + \sqrt{\Omega}x(t)$$

其中 $\sqrt{\Omega}x(t)$ 是n对其平均值$\Omega\eta(t)$的偏移。



我们可以把主方程用x来表示，在n取值n->n +Δn内，我们定义（这样η(t)显式地依赖于t）：

$$P_1(n,t)\Delta n \equiv \pi(x,t)\Delta x, \text{其中}$$

$$\pi(x,t) \equiv \sqrt{\Omega}P_1(\Omega\eta(t) + \sqrt{\Omega}x, t)$$

于是我们有： $\frac{1}{\sqrt{\Omega}}\frac{\partial \pi}{\partial x} = \sqrt{\Omega}\frac{\partial P_1}{\partial n}$，及 $\frac{\partial \pi}{\partial t} = \sqrt{\Omega}\left\{\Omega\frac{d\eta}{dt}\frac{\partial P_1}{\partial n} + \frac{\partial P_1}{\partial t}\right\}$.

主方程随之变为：

$$\frac{\partial \pi}{\partial t} - \sqrt{\Omega}\frac{d\eta}{dt}\frac{d\pi}{dx} =$$

$$f(\Omega)\sum_{\Delta n}\left\{\omega_0\left(\eta(t) + \frac{x}{\sqrt{\Omega}} - \frac{\Delta n}{\Omega};\Delta n\right)\pi(x - \frac{\Delta n}{\sqrt{\Omega}}, t) - \omega_0\left(\eta(t) + \frac{x}{\sqrt{\Omega}};-\Delta n\right)\pi(x,t)\right\}.$$

把上式右边第一项在 $\Delta n / \sqrt{\Omega} = 0$ 附近作泰勒展开

$$\omega_0\left(\eta(t) + \frac{x}{\sqrt{\Omega}} - \frac{\Delta n}{\Omega};\Delta n\right)\pi(x - \frac{\Delta n}{\sqrt{\Omega}}, t)$$

$$= \left(1 - \frac{\Delta n}{\sqrt{\Omega}}\frac{\partial}{\partial x} + \frac{1}{2}\left(\frac{\Delta n}{\sqrt{\Omega}}\right)^2\frac{\partial^2}{\partial x^2} + \cdots\right)\times\omega_0\left(\eta(t) + \frac{x}{\sqrt{\Omega}};\Delta n\right)\pi(x,t),$$

并重新标定时间f(Ω)t= Ωτ后，主方程最终变为：

$$\frac{\partial \pi'}{\partial \tau} - \sqrt{\Omega}\frac{d\eta'}{d\tau}\frac{d\pi'}{dx} = \sum_{\Delta n}\Omega\left[-\left(\frac{\Delta n}{\sqrt{\Omega}}\right)\frac{\partial}{\partial x} + \frac{1}{2}\left(\frac{\Delta n}{\sqrt{\Omega}}\right)^2\frac{\partial^2}{\partial x^2} + \cdots\right]\omega_0\left(\eta'(\tau) + \frac{x}{\sqrt{\Omega}};\Delta n\right)\pi'(x,\tau).$$

其中 $\pi'(x,\tau) = \pi(x, \Omega\tau / f(\Omega))$， $\eta'(\tau) = \eta(\Omega\tau / f(\Omega))$ 和

$$\omega_0\left(\eta'(\tau) + \frac{x}{\sqrt{\Omega}};\Delta n\right) = \left(1 + \frac{x}{\sqrt{\Omega}}\frac{\partial}{\partial \eta'} + \cdots\right)\omega_0(\eta'(\tau), \Delta n).$$

- 在主方程里保留到 $\sqrt{\Omega}$，可得：

$$\frac{d\eta'}{d\tau}\frac{d\pi'}{dx} = \sum_{\Delta n}\Delta n\frac{\partial}{\partial x}\omega_0\left(\eta'(\tau),\Delta n\right)\pi'(x,\tau) = \sum_{\Delta n}\Delta n\omega_0\left(\eta'(\tau),\Delta n\right)\frac{\partial\pi'}{\partial x}.$$

要满足上式，只须取

$$\frac{d\eta'}{d\tau} = \sum_{\Delta n}\Delta n\omega_0\left(\eta'(\tau),\Delta n\right) = a_1'(\eta'(\tau)).$$

这里 $f(\Omega)a_p'(n/\Omega) = a_p(n)$.

- 在主方程里保留到Ω的零级项，可得关于概率密度 $\pi'(x,\tau)$ 的Fokker-Planck方程：

$$\frac{\partial\pi'}{\partial\tau} = -(\partial_{\eta'}a_1')\frac{\partial}{\partial x}x\pi' + \frac{1}{2}a_2'\frac{\partial^2\pi'}{\partial x^2}.$$

由上式即可得扰动x的平均值和矩等的运动方程。对平稳过程，上述方程右端的系数与时间无关。如在τ=0有 $\pi'(x,0) = \delta(x-x_0)$，并定义 $s(\tau) = \ln[a_1'(\eta'(0))/a_1'(\eta'(\tau))]$，及 $x = ye^{-s}$ 和 $\pi'(x,s) = e^sQ(y,s)$，上述Fokker-Planck方程可变为一个广义扩散方程：

$$\frac{\partial Q}{\partial s} = -\frac{a_2'}{2(\partial_{\eta'}a_1')}e^{2s}\frac{\partial^2 Q}{\partial y^2}.$$

- 主方程展开到Ω的 $1/\sqrt{\Omega}$ 级项时，W的泰勒展开的第二项ω1将开始有贡献。

2.4 非线性生灭过程---马尔萨斯方程

- 对线性生灭过程：

考虑 t时刻有m个人的一个社会，我们有：

$$P_{1|1}(m,t \mid n,t+\Delta t) = (1 - m\lambda\Delta t - m\mu\Delta t)\delta_{n,m} + (m\lambda\delta_{n,m+1} + m\mu\delta_{n,m-1})\Delta t + \cdots.$$

这里我们容易发现转移率为：

$$W(m,n) = m\mu\delta_{n,m-1} + m\lambda\delta_{n,m+1}.$$

- 对非线性生灭过程：我们假设社会成员间的竞争使得死亡率加大，因此死亡率中还有一个正比于其它个体密度的项$(n-1)/\Omega$贡献，这里Ω是系统的大小。这样转移率变为：

$$W(m,n) = (m\mu + \gamma\frac{m(m-1)}{\Omega})\delta_{n,m-1} + m\lambda\delta_{n,m+1}.$$

- 由于在时间τ内不转变的概率为 $\left(1 - \tau\int dyW_{t_1}(y_1,y)\right)\delta(y_1 - y_2)$，我们发现主方程可写为：

$$\frac{\partial P_1(n,t)}{\partial t} = (n-1)\lambda P_1(n-1,t) + \left[(n+1)\mu + \gamma\frac{n(n+1)}{\Omega}\right]P_1(n+1,t) - \left[n\lambda + n\mu + \gamma\frac{n(n-1)}{\Omega}\right]P_1(n,t).$$

在t时刻个体平均数的方程为：

$$\frac{d\langle n(t)\rangle}{dt}=(\lambda-\mu)\langle n(t)\rangle-\frac{\gamma}{\Omega}\langle n^2(t)\rangle+\frac{\gamma}{\Omega}\langle n(t)\rangle,$$

这个方程中一级矩的演化依赖于二级矩。为此把n在其平均值附近展开：

$$n(t)=\Omega\eta(t)+\sqrt{\Omega}x(t)$$

并保留到Ω级的项，我们发现：

$$\frac{d\eta(t)}{dt}=(\lambda-\mu)\eta(t)-\frac{\gamma}{\Omega}\eta^2(t),$$

此方程的解为

$$\eta(t)=\frac{\eta(0)e^{(\lambda-\mu)t}}{1+\eta(0)\dfrac{\gamma}{\lambda-\mu}\left\{e^{(\lambda-\mu)t}-1\right\}}.$$

特点：

- 当$\lambda-\mu<0$, 则$\eta(t)\underset{t\to\infty}{\to}0$, 人口消亡；
- 当$\lambda-\mu>0$, 则$\eta(t)\underset{t\to\infty}{\to}\dfrac{\lambda-\mu}{\gamma}$, 人口趋于一个稳定值($\gamma=0$时趋于无穷大)。

即当转移率改变时，存在一个从一个状态到另一状态的“相变”！