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，$\cdots$, 在 t_n 时刻取 y_n 值的联合概率）。

一般性质:

联合概率密度是正的: $P_n \geq 0$;

它们可以被约化:

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

并且是归一化的: $\int P_1(y_1,t_1)\,\mathrm{d}y_1 = 1.$

- 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):

$$\langle y_1(t_1)\, y_2(t_2) \cdots y_n(t_n) \rangle = \int \cdots \int y_1 y_2 \cdots y_n P_n(y_1, t_1; y_2, t_2; \cdots ; y_n, t_n)\, \mathrm{d}y_1 \cdots \mathrm{d}y_n.$$

- 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：

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

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：

$P_{1|1}(y_1, t_1 | y_2, t_2) \equiv$（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）；

它由如下恒等式来定义：

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

- 不同时刻概率密度之间的关系（上式对 y_1积分）：

$$P_1(y_2, t_2) = \int P_1(y_1, t_1)\, P_{1|1}(y_1, t_1 | y_2, t_2)\, \mathrm{d}y_1.$$

- 联合条件概率密度：

$$P_{k|\ell}(y_1,t_1;\cdots;y_k,t_k|y_{k+1},t_{k+1};\cdots;y_{k+\ell},t_{k+\ell})$$

= （固定 $(y_1,t_1;\cdots;y_k,t_k)$ 时，随机变量Y具有值 $(y_{k+1},t_{k+1};\cdots;y_{k+\ell},t_{k+\ell})$ 的联合概率密度）

$$= \frac{P_{k+\ell}(y_1,t_1;\cdots;y_k,t_k;y_{k+1},t_{k+1};\cdots;y_{k+\ell},t_{k+\ell})}{P_k(y_1,t_1;\cdots;y_k,t_k)}.$$

（II）马尔科夫过程(Markov process)

对马尔科夫过程我们有(其中 $t_1 < t_2 < \cdots < t_n$)：

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

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）：

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

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）：

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

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

$P_1(y_2,t)$ 的时间导数为：
$$\frac{\partial P_1(y_2,t)}{\partial t}=\lim_{\tau\to 0}\frac{P_1(y_2,t+\tau)-P_1(y_2,t)}{\tau},$$

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

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的单位时间的转变概率密度（转移率）。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内不转变的概率密度为 $\left[1-\tau\int \mathrm{d}y\,W_{t_1}(y_1,y)\right]\delta(y_1-y_2)$。所以有：

$$P_{1|1}(y_1,t_1|y_2,t_1+\tau)=\tau W_{t_1}(y_1,y_2)+\left[1-\tau\int \mathrm{d}y\,W_{t_1}(y_1,y)\right]\delta(y_1-y_2)+O(\tau^2) \quad (*3)$$

由（*1-3）我们发现：

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

这就是主方程。

(IV) 细致平衡和Monte Carlo模拟

为简单记这里我们考虑离散的情形，这时主方程可写为：

$$\dot{P}_n(t) = \sum_{n'} \left(W_{n'n}(t)P_{n'}(t) - W_{nn'}(t)P_n(t)\right) = \sum_{n'} V_{nn'}P_{n'}(t).$$

这和我们以前学过的统计物理里的刘维尔定理很相似。对统计物理研究的很多系统而言，转移概率一般是不含时的，即与系统是否处于平衡态无关，因此我们一般有：$W_{nn'}(t) = W_{nn'}$ 。对平稳过程，我们有 $\dot{P}_n(t) = 0$, 因此对不同的平衡态有（这里我们略去了时间）：

$$W_{nn'}P_n^{\text{stat}} = W_{n'n}P_{n'}^{\text{stat}},$$

这就是**细致平衡（detailed balance）**。可以证明（见下），即使系统初始处于非平衡态时（这时概率密度函数与时间有关），经过足够长的时间后系统将逐渐进入平衡态，这是我们对系统进行**Monte Carlo模拟**的理论基础。

练习(对离散情形的证明见本章末尾):

考虑相对熵：$S_{\text{rel}}(t) = -k_B \sum_n P_n(t) \ln\left(P_n(t)/P_n^{\text{stat}}\right)$ 。这里 $P_n(t)$是系统处于非平衡态的概率密度函数，P_n^{stat}则是系统处于平衡态的概率密度函数。证明 $S_{\text{rel}}(t) \leq 0$ 和 $\dot{S}_{\text{rel}}(t) \geq 0$，其中等号仅当系统处于平衡态时成立。由此有：

$$\lim_{t\to\infty} P_n(t) = P_n^{\text{stat}}.$$

(V) 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。先做变量代换:

$W(y',y) = \bar{W}(y', y-y') \equiv \bar{W}(y',\xi)$, 类似地 $W(y,y') = \bar{W}(y, y'-y) = \bar{W}(y,-\xi)$,

这里 $\xi = y - y'$ 是跳跃的大小。于是主方程变为

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

由于转移概率 $\bar{W}(y',y)$ 将随ξ的增大而迅速减小，我们把WP₁按ξ的幂次展开：

$$\frac{\partial P_1(y,t)}{\partial t} = \int \mathrm{d}\xi\, [\bar{W}(y,\xi)\, P_1(y,t)] - \int \mathrm{d}\xi\, \xi \frac{\partial}{\partial y}[\bar{W}(y,\xi)\, P_1(y,t)]$$
$$+\frac{1}{2}\int \mathrm{d}\xi\, \xi^2 \frac{\partial^2}{\partial y^2}[\bar{W}(y,\xi)\, P_1(y,t)] + \cdots - P_1(y,t)\int \mathrm{d}\xi\, \bar{W}(y,-\xi).$$

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**：

$$\alpha_n(y) = \int \mathrm{d}\xi\, \xi^n\, \bar{W}(y,\xi).$$

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 $Y=\{y_1,y_2,\ldots,y_\ell\}$，基本时间间隔为1，从t=0到t=1我们有：

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

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

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

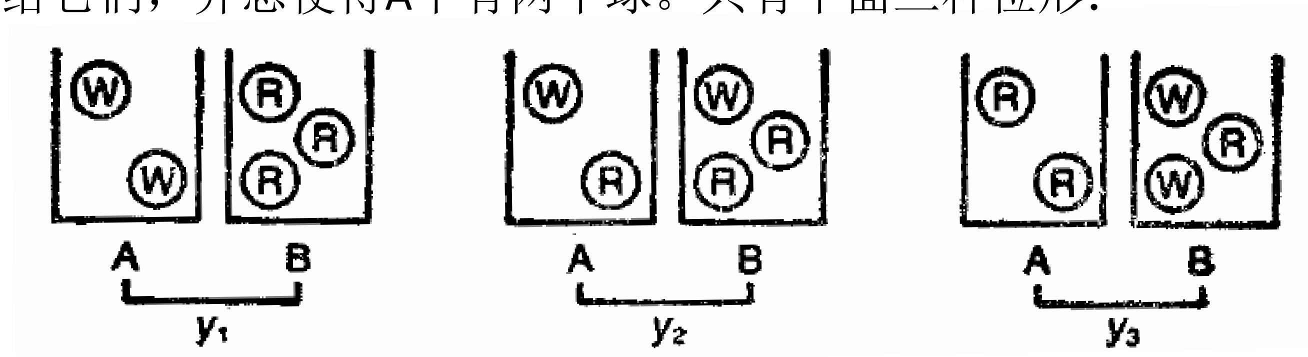
- 在s时刻，我们有：$\boldsymbol{P}(s)=\boldsymbol{P}(0)\mathbf{Q}^s$.
- **P**(s) 在s很大时的行为依赖于转移矩阵的结构。
  若**Q**的某个幂次的全部元素都是正的（正则矩阵），则**P**(s)趋向唯一的确定的与初态无关的定态 $\boldsymbol{\pi}=(\pi_1,\pi_2,\cdots,\pi_\ell)$：

$$\mathbf{Q}^s\to\mathbf{M}=\begin{pmatrix}\pi_1&\pi_2&\cdots&\pi_\ell\\ \pi_1&\pi_2&\cdots&\pi_\ell\\ \cdots&\cdots&\cdots&\cdots\\ \pi_1&\pi_2&\cdots&\pi_\ell\end{pmatrix}.$$

且易证明：

$$\mathbf{M}=\mathbf{QM}=\mathbf{MQ},\ \ \boldsymbol{\pi}=\boldsymbol{P}(0)\mathbf{M}=\boldsymbol{\pi}\mathbf{Q}.$$

一个例子（雷克书P.173）：考虑两个罐子A和B，有三个红球和两个白球分配给它们，并总使得A中有两个球。共有下面三种位形：



位形间的转移为：无规则地从A和B中各取一个球进行交换。

转移矩阵为

$$\mathbf{Q}=\begin{pmatrix}0 & 1 & 0\\ 1/6 & 1/2 & 1/3\\ 0 & 2/3 & 1/3\end{pmatrix}$$

且易知 $\mathbf{Q}^2$是正则的:

$$\mathbf{Q}^2=\begin{pmatrix}1/6 & 1/2 & 1/3\\ 1/12 & 23/36 & 10/36\\ 1/9 & 20/36 & 1/3\end{pmatrix}$$

令 $\boldsymbol{\pi}=(x,y,z)$表示定态，由方程：$\boldsymbol{\pi}=\boldsymbol{\pi}\mathbf{Q}$ 可解出定态，结果为：
$\boldsymbol{\pi}=(1/10,6/10,3/10)$，与初态无关。

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：

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

若粒子向左向右运动的概率均为1/2，则

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

原方程可简化为：

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

把上式写为求导的形式，我们有：

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

令 $x = n\ell, t = s\tau,$ 并在 $D \equiv \ell^2/(2\tau)$ 为有限的条件下取极限 $\ell \to 0, \tau \to 0,$ 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P₁(x,t)对x的傅里叶变换（特征函数），扩散方程可变为：

$$\frac{\partial \tilde{P}_1(k,t)}{\partial t}=-Dk^2\tilde{P}_1(k,t), \tilde{P}_1(k,0)=1.$$

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

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

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。

一阶矩和二阶矩：

一阶矩：把扩散方程两边乘以 x 并对位置积分后，我们发现：$\frac{\mathrm{d}\langle x(t)\rangle}{\mathrm{d}t}=0.$ 因此粒子的平均距离不随时间改变；

二阶矩：把扩散方程两边乘以 x^2 并对位置积分后，我们发现：

$$\frac{\mathrm{d}\langle x^2(t)\rangle}{\mathrm{d}t}=2D \implies \langle x^2(t)\rangle=2Dt.$$

这正是扩散过程的特征。

2.4 生灭过程，主方程的求解

- 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:

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

再假定生和灭的概率正比于现存细菌数（线性假设），则有 $\lambda_m = m\lambda$ 和 $\mu_m = m\mu$，上式两边同乘以 $P_1(m, t)$ 并对m求和再化简，即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（characteristic function）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：

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

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

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：

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

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

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

主方程左右两侧同乘z^n并对所有n求和，再把以上表达式带入到主方程中我们有：

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

因此方程（**）和主方程是等价的，我们只需解方程（**）。

容易发现（**）可由方程组 $\mathrm{d}F = \frac{\partial F}{\partial z}\mathrm{d}z + \frac{\partial F}{\partial t}\mathrm{d}t = 0$ 及 $\frac{\mathrm{d}t}{1} = \frac{-\mathrm{d}z}{(z-1)(\lambda z-\mu)}$ 得到。

从dF=0我们发现$F(z,t)$=C_2，由 $\frac{\mathrm{d}t}{1} = \frac{-\mathrm{d}z}{(z-1)(\lambda z-\mu)} \Longrightarrow C_1 = \frac{z-1}{\lambda z-\mu}\,\mathrm{e}^{-(\mu-\lambda)t},$

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

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

故若 $u = (z-1)^{-1}(\lambda z-\mu)$, 则 $\tilde{F}(u) = \left(\frac{\mu-u}{\lambda-u}\right)^m$，代入到（***），

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

2.5 离散平稳马尔科夫过程的普遍解

- 对平稳马尔科夫过程，Chapman-Kolmogorov方程变为：

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

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：

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

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

- 由概率和条件概率的定义我们还有：$P_1(n)=\sum_m P_1(m)P_{1|1}(m|n;s)$，即 $\mathbf{P}=\mathbf{P}\mathbf{Q}^s, s\geq 0$

和 $\sum_n P_{1|1}(m|n;s)=1$ （$\mathbf{Q}^s$ 的归一性）。

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：

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

右本征矢量问题可写为：

$$\lambda_j\, y_{mj} = \sum_{n=1}^{\ell} Q_{mn} y_{nj},$$

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：$Q_{mn} = \sum_i \lambda_i\, y_{mi}\, x_{in}.$ 因此我们有
  $$P_{1|1}(m|n;s) = \sum_i \lambda_i^s\, y_{mi}\, x_{in}.$$
- **Q**至少有一本征值为1，且$|\lambda_i| \le 1.$

  由**P=PQ**及λ$_i$**X**$_i$=**X**$_i$**Q**和λ$_i$**Y**$_i$=**QY**$_i$，易得 $PY_i = P\mathbf{Q}Y_i = \lambda_i PY_i \Longrightarrow (\lambda_i - 1)PY_i = 0.$
  因Y$_i$构成完备本征矢，若所有$\lambda_i \neq 1$则对所有i均有PY$_i$=0，这不可能。故存在i使得 $\lambda_i = 1.$
  用以下两种办法可以证明$|\lambda_i| \le 1$：
  (a) 左本征矢方程两边取绝对值：$|\lambda_i| \cdot |x_{im}| \le \sum_n |x_{in}|\, Q_{nm},$对m求和并考虑Q的归一 性即得：$|\lambda_i| \le 1;$
  (b) 右本征矢方程两边取绝对值有：$|\lambda_j| \cdot |y_{mj}| \le \sum_n Q_{mn}\, |y_{nj}|.$ 设对所有m, $\max(|y_{mj}|) = |y_{\tilde{m}j}| = C,$
  对 $\tilde{m}$我们有：$|\lambda_j| \cdot |y_{\tilde{m}j}| = |\lambda_j| C \le \sum_n Q_{\tilde{m}n}\, |y_{nj}| \le C \sum_n Q_{\tilde{m}n} = C \Rightarrow |\lambda_j| \le 1.$

- 对正则转移矩阵，若Q只有一个本征值 $\lambda_1 = 1$，则

$$\lim_{s\to\infty} P_{1|1}(m|n;s) = \lim_{s\to\infty} \left(y_{m1}\, x_{1n} + \sum_{i\neq 1} \lambda_i^s\, y_{mi}\, x_{in} \right) = y_{m1}\, x_{1n} = P_1.$$

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$\rho(x,t)=P_1(n,t)$,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：

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

- 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：

$$\frac{\partial \langle n^2(t)\rangle}{\partial t} = \sum_{m,n} \left(m^2-n^2\right) W(n,m)P_1(n,t) = \langle a_2(n)\rangle + 2\langle n\, a_1(n)\rangle.$$

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：

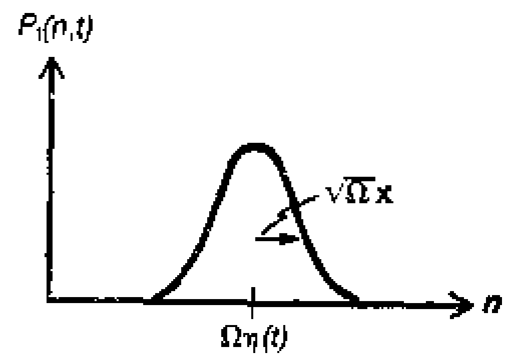
$$W(m,n) = f(\Omega)\left[\omega_0\left(\frac{m}{\Omega},\Delta n\right) + \frac{1}{\Omega}\omega_1\left(\frac{m}{\Omega},\Delta n\right) + \cdots\right],$$

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$\sigma_n$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$\sqrt{\Omega}\,x$ 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）：

$$P_1(n,t)\Delta n \equiv \pi(x,t)\Delta x,$$

其中 $\Delta n = \sqrt{\Omega}\Delta x$，$\pi(x,t) \equiv \sqrt{\Omega}\,P_1(\Omega\,\eta(t) + \sqrt{\Omega}\,x,t)$.

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

主方程随之变为：

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

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

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

并重新标定时间 f(Ω)t=Ωτ 后，主方程最终变为：

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

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：

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

要满足上式，只须取

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

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：

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

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 $s(\tau)=\ln\left[a_1'(\eta'(0))/a_1'(\eta'(\tau))\right]$，及 $x=y\,\mathrm{e}^{-s}$ 和 $\pi'(x,s)=\mathrm{e}^{s}Q(y,s)$，上述Fokker-Planck方程可变为一个广义扩散方程：

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

- 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。

2.7 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：

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

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 $(m-1)/\Omega$ 贡献，这里Ω是系统的大小。这样转移率变为：

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

- 由于在时间τ内不转变的概率为 $\left[1-\tau\int \mathrm{d}y\, W_{t_1}(y_1,y)\right]\delta(y_1-y_2)$，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：

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

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：

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

此方程的解为

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

特点：

- 当$\lambda-\mu<0$，则$\eta(t)\to 0$，人口消亡；
- 当$\lambda-\mu>0$，则$\eta(t)\xrightarrow[t\to\infty]{}\frac{\lambda-\mu}{\gamma}$，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”！

2.8 马尔科夫链的相对熵增加定理

在这一节里我们要证明马尔科夫链的相对熵的一个重要性质，即下面的定理：

定理：设 μ_n 和 μ'_n 分别为同一个有限态马尔科夫链在时刻*n*的两个概率密度函数，那么它们的相对熵 $S[\mu_n\|\mu'_n] = -k\sum\mu_n \ln(\mu_n/\mu'_n) \leq 0$ 是时刻*n*的一个单调递增函数。特别地，如果*μ*是唯一的一个平稳(stationary)分布，我们有$S[\mu_n\|\mu] \xrightarrow{n\to\infty} 0.$

这里我们为了和条件概率中的"|"区分，用了"||"来分开相对熵中的两个概率密度函数。

- 首先，对任两个联合概率密度函数*p*(*x*,*y*)和*q*(*x*,*y*)，相对熵有一个链式规则（条件概率仍沿用雷克书的写法）：

$$S[p(x,y)\|q(x,y)] = S[p(x)\|q(x)] + S[p(x|y)\|q(x|y)].$$

这由条件概率的定义即可得证（作为练习）。注意我们有：

$$p(x,y)/q(x,y) = p(x)p(x|y)/[q(x)q(x|y)] \quad \text{and} \quad \sum_y p(x|y) = 1,$$

$$S[p(x|y)\|q(x|y)] \equiv -k\sum_x p(x)\sum_y p(x|y)\ln\frac{p(x|y)}{q(x|y)} = \sum_x p(x)S[p(X=x|Y)\|q(X=x|Y)] \leq 0.$$

上面不等式中的等号仅当 $p(x|y) = q(x|y)$时才成立。

- 我们考虑相邻时刻*n*和*n*+1的联合概率密度函数$p(x_n,x_{n+1})=p(x_n)\,r(x_n|x_{n+1})=\mu_n\,r(x_n|x_{n+1})$

和 $q(x_n,x_{n+1})=q(x_n)\,r(x_n|x_{n+1})=\mu'_n\,r(x_n|x_{n+1})$, 这里 $r(x_n|x_{n+1})$ 是同一个马尔科夫链的不同时刻的条件概率，易知对密度函数*p*和*q*均相同。于是利用链式规则有：

$$\begin{aligned} S[p(x_n,x_{n+1})\|q(x_n,x_{n+1})] &= S[p(x_n)\|q(x_n)]+S[p(x_n|x_{n+1})\|q(x_n|x_{n+1})] \\ &= S[p(x_{n+1})\|q(x_{n+1})]+S[p(x_{n+1}|x_n)\|q(x_{n+1}|x_n)]. \end{aligned}$$

在上面第一个等式右边，由于 $p(x_n|x_{n+1})=q(x_n|x_{n+1})=r(x_n|x_{n+1})$, 故它们的相对熵为零（已定义时刻 *n*+1 在时刻 *n* 之后）。再利用相对熵的非正性，由第二个等式我们即有：

$$S[p(x_n)\|q(x_n)]\le S[p(x_{n+1})\|q(x_{n+1})] \text{ 或 } S[\mu_n\|\mu'_n]\le S[\mu_{n+1}\|\mu'_{n+1}].$$

等号仅当 $p(x_{n+1}|x_n)$和 $q(x_{n+1}|x_n)$ 相等时成立。

特别地,如 μ'_n 是平稳概率密度函数，则有: $\mu'_n=\mu'_{n+1}=\mu$, 于是 $S[\mu_n\|\mu]\le S[\mu_{n+1}\|\mu]$, $S[\mu_n\|\mu]$ 是上界为零的单调递增函数。若平稳概率密度函数只有一个，则有：

$$S[\mu_n\|\mu]\xrightarrow{n\to\infty}0.$$