

# 计算物理学 教案



## 目 录

目 录 .....	I
第一章 随机变量及其分布 .....	1
1.1 概率论简介 .....	1
1.2 随机变量和随机向量 .....	4
1.3 一些重要的概率分布 .....	7
第二章 蒙特卡洛方法 .....	13
2.1 蒙特卡洛方法的数学基础 .....	13
2.2 随机变量的抽样方法 .....	16
2.2.1 变换抽样法 .....	16
2.2.2 舍选抽样法 .....	17
2.2.3 Metropolis 抽样法 .....	21
2.2.4 迭加原理 .....	22
2.2.5 条件概率密度法 .....	23
2.3 减小蒙特卡洛方差的技巧 .....	24
第三章 蒙特卡洛方法的应用 .....	27
3.1 定积分的掷点法求解 .....	27
3.2 多体末态过程的相空间产生 .....	28
3.3 量子力学路径积分形式的求解 .....	32
3.4 量子力学变分形式的求解 .....	35
3.5 中子输运问题的蒙特卡洛模拟 .....	37

第四章 偏微分方程的求解 .....	43
4.1 有限差分方法 .....	43
4.1.1 微商的差商格式 .....	43
4.1.2 偏微分方程的离散化 .....	45
4.2 有限元素方法 .....	48
4.2.1 基本物理思想 .....	48
4.2.2 二维场的有限元素法 .....	49

## 第一章 随机变量及其分布

本章将介绍三部分的内容：1. 概率论初步知识；2. 随机变量及其分布；3. 一些重要的概率分布。

### 1.1 概率论简介

本节中我们将给出概率的定义，并引伸出条件概率、边缘概率的概念，给出全概率公式及贝叶斯公式。

随机试验：可以在“相同条件”下重复进行的、所有可能结果都已明确知道但每次试验结果无法确定的试验。

随机事件：随机试验中可能出现的各种结果。

#### 1. 概率函数的定义

$P$  是定义在样本空间  $S$  上的函数，如果满足

$$(1) \quad 0 \leq P(A) \leq 1, \quad \forall A \subset S \quad (1.1)$$

$$(2) \quad P(S) = 1$$

$$(3) \quad P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{k=1}^n P(A_k), \quad P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) = \sum_{k=1}^{\infty} P(A_k)$$

其中， $A_k$  ( $k = 1, 2, \dots$ ) 为不相容事件。则函数  $P$  称为概率函数。上述三式分别称为概率函数的正定性、归一性、有限（可列）可加性。

## 2. 概率函数的性质

$$(1) \quad P(A) + P(\bar{A}) = 1 \quad (1.2)$$

$$(2) \quad P(\emptyset) = 0$$

$$(3) \quad P(A) \leq P(B), \quad A \subset B$$

$$(4) \quad \sum_{i=1}^n P(A_i) = 1, \quad A_i \ (i = 1, 2, \dots, n) : \text{partition of } S$$

$$(5) \quad P(A - B) = P(A) - P(B), \quad A \supset B$$

$$(6) \quad P(A \cup B) = P(A) + P(B) - P(AB)$$

$$(7) \quad P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) \\ - P(A_1A_2) - P(A_2A_3) - P(A_3A_1) + P(A_1A_2A_3)$$

## 3. 条件概率

已知  $A, B$  为一随机试验的两个事件, 事件  $A$  的概率为  $P(A)$ , 则在事件  $A$  发生的条件下事件  $B$  发生的概率称为条件概率, 记为  $P(B|A)$ 。由于在事件  $A$  发生的条件下事件  $B$  发生的概率应该是以  $A$  作为样本空间时  $A, B$  事件同时发生的概率, 所以有

$$P(B|A) = \frac{P(AB)}{P(A)} \quad (1.3)$$

由条件概率的定义, 我们可以得到如下的概率乘法定理:

$$P(AB) = P(B|A) \cdot P(A) = P(A|B) \cdot P(B) \quad (1.4)$$

## 4. 事件的独立性

两个事件  $A$  和  $B$  相互独立的定义为

$$P(AB) = P(A) \cdot P(B) \quad (1.5)$$

这相当于

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B) \quad (1.6)$$

三个事件  $A$ 、 $B$  和  $C$  相互独立的定义为

$$\begin{aligned} P(AB) &= P(A) \cdot P(B) \\ P(BC) &= P(B) \cdot P(C) \\ P(CA) &= P(C) \cdot P(A) \\ P(ABC) &= P(A) \cdot P(B) \cdot P(C) \end{aligned} \quad (1.7)$$

由上述的定义可知，彼此独立的三事件间两两独立。但是，该命题的逆命题不成立。

## 5. 边沿概率、全概率公式及贝叶斯公式

$A_i$  ( $i = 1, 2, \dots, m$ ) 和  $B_j$  ( $j = 1, 2, \dots, n$ ) 为样本空间  $S$  的两个分割，事件  $A_i$  和  $B_j$  的边沿概率分别定义为

$$P(A_i) = \sum_{j=1}^n P(A_i \cap B_j) \quad \text{and} \quad P(B_j) = \sum_{i=1}^m P(B_j \cap A_i) \quad (1.8)$$

利用边沿概率及条件概率的定义，我们得到全概率公式：

$$P(A) = \sum_{j=1}^n P(A|B_j)P(B_j) \quad (1.9)$$

贝叶斯公式：

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)} \quad (1.10)$$

## 1.2 随机变量和随机向量

本节中我们将引入随机变量（向量）的概念，并介绍随机变量（向量）的分布函数、概率密度函数及数字特征。

随机变量是定义在样本空间  $S$  上的单值实函数，分为离散型随机变量和连续型随机变量两大类。

### 1. 随机变量的分布函数

$X$  为一随机变量，其取值范围为  $x_{\min} \leq x \leq x_{\max}$ 。随机变量  $X$  取值小于等于  $x$  的概率称为分布函数，用  $F(x)$  表示为

$$F(x) = P(X \leq x) \quad (1.11)$$

分布函数具有如下性质：

$$\begin{aligned} (1) \quad & 0 \leq F(x) \leq 1 \\ (2) \quad & F(x_{\min}) = 0 \quad \text{and} \quad F(x_{\max}) = 1 \\ (3) \quad & F(x_2) - F(x_1) \geq 0, \quad x_2 > x_1 \end{aligned} \quad (1.12)$$

如果  $X$  为一离散型随机变量，其取值为分立的实数  $x_i$  ( $i = 1, 2, \dots$ )，相应的概率为  $P(X = x_i) = p_i$ ，则分布函数为

$$F(x) = \sum_{x_i \leq x} p_i \quad (1.13)$$

### 2. 连续型随机变量的概率密度函数

$X$  为一连续型随机变量，若存在非负连续实函数  $f(x)$ ，对于任意实数  $x \in [x_{\min}, x_{\max}]$  下式成立：

$$F(x) = \int_{x_{\min}}^x f(t) dt \quad (1.14)$$

则  $f(x)$  称为随机变量  $X$  的概率密度函数。



概率密度函数具有如下性质：

$$\begin{aligned}
 (1) \quad & f(x) \geq 0 \\
 (2) \quad & \int_{x_{min}}^{x_{max}} f(x)dx = 1 \\
 (3) \quad & \int_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1), \quad x_2 > x_1 \\
 (4) \quad & f(x) = F'(x)
 \end{aligned} \tag{1.15}$$

若随机变量  $X$  和  $Y$  之间存在一一映射： $y = y(x)$ （或  $x = x(y)$ ），则  $X$  的概率密度函数  $f(x)$  与  $Y$  的概率密度函数  $g(y)$  之间存在如下关系：

$$g(y) = f(x(y)) \left| \frac{dx(y)}{dy} \right| \tag{1.16}$$

### 3. 随机变量的数字特征

随机变量的数字特征包括：期望值、方差、偏度、峰度等各阶原点矩或中心矩。

• 期望值：

$$\mu = \langle X \rangle = \int_{\Omega} x f(x) dx = \sum_i x_i p_i \tag{1.17}$$

• 方差：

$$\sigma^2 = \langle (X - \mu)^2 \rangle \tag{1.18}$$

• 偏度：

$$\gamma_1 = \frac{\langle (X - \mu)^3 \rangle}{\sigma^3} \tag{1.19}$$

• 峰度：

$$\gamma_2 = \frac{\langle (X - \mu)^4 \rangle}{\sigma^4} - 3 \tag{1.20}$$

可以证明，正态分布的偏度系数及峰度系数均为 0。

#### 4. 高维随机向量

一维随机变量的相关概念可以推广定义至高维随机向量。以二维随机向量  $(X, Y)$  为例，分布函数定义为

$$F(x, y) = P(X \leq x, Y \leq y) \quad (1.21)$$

函数  $f(x, y)$  称为二维连续型随机向量  $(X, Y)$  的概率密度函数，如果

$$F(x, y) = \int_{-\infty}^x du \int_{-\infty}^y dv f(u, v) \quad (1.22)$$

显然，

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \quad (1.23)$$

与一维随机变量不同的是：对于高维随机向量，我们有边沿概率和条件概率的概念。

- 随机向量  $(X, Y)$  关于  $X$  和  $Y$  的边沿分布函数：

$$\begin{aligned} F_X(x) &= F(x, y_{\max}) \\ F_Y(y) &= F(x_{\max}, y) \end{aligned} \quad (1.24)$$

- 随机向量  $(X, Y)$  关于  $X$  和  $Y$  的边沿概率密度函数（如果存在）：

$$\begin{aligned} f_X(x) &= \int_{\Omega_y} f(x, y) dy \\ f_Y(y) &= \int_{\Omega_x} f(x, y) dx \end{aligned} \quad (1.25)$$

- 在  $Y = y$  ( $X = x$ ) 下随机变量  $X$  ( $Y$ ) 的条件概率密度函数：

$$\begin{aligned} f(x|y) &= \frac{f(x, y)}{f_Y(y)} \\ f(y|x) &= \frac{f(x, y)}{f_X(x)} \end{aligned} \quad (1.26)$$

由独立性的定义知：

$$\begin{aligned} X, Y : \text{independent} &\iff F(x, y) = F_X(x)F_Y(y) \\ &\iff f(x, y) = f_X(x)f_Y(y) \end{aligned} \quad (1.27)$$

对于二维随机向量  $(X, Y)$ ，定义协方差为

$$\text{cov}(X, Y) = \langle (X - \mu_X)(Y - \mu_Y) \rangle \quad (1.28)$$

$\text{cov}(X, Y) > 0$  称为正相关； $\text{cov}(X, Y) < 0$  称为负相关； $\text{cov}(X, Y) = 0$  称为不相关。可以证明：两随机变量独立时必不相关；反之未必正确。

### 1.3 一些重要的概率分布

本节中我们将介绍一些物理上非常重要的概率分布，包括离散分布和连续分布两大类。

#### 1. 伯努利分布

伯努利试验的结果有两种： $A$  表示“成功”， $\bar{A}$  表示“失败”。 $A$  和  $\bar{A}$  的概率分别为  $P(A) = p$ ， $P(\bar{A}) = 1 - p = q$ 。用离散型随机变量  $X$  表示伯努利试验的结果， $X = 1$  表示成功， $X = 0$  表示失败，则  $X$  的概率分布为

$$B(r; p) = p^r(1 - p)^{1-r} \quad (1.29)$$

此概率分布称为伯努利分布或  $(0, 1)$  分布。期望值和方差分别为  $p$  和  $pq$ 。

说明：伯努利随机变量  $X$  的值为单次伯努利试验中成功的次数。

#### 2. 二项分布

$X$  为  $n$  次彼此独立的同概率伯努利试验中成功的次数，取值为： $0, 1, \dots, n$

$$X = X_1 + X_2 + \dots + X_n \quad (1.30)$$

所满足的概率分布为

$$B(r; n, p) = \binom{n}{r} p^r (1-p)^{n-r} \quad (1.31)$$

期望值和方差分别为  $np$  和  $npq$ 。

说明：考虑随机变量  $Y = X/n$ ，其期望值和方差分别为  $p$  和  $pq/n$ 。这表明，对多次独立的同概率伯努利试验取平均可以减小方差。

### 3. 多项分布

多项分布是二项分布的推广。在二项分布中，单次伯努利试验的结果只有两种。一般地，设单次随机试验的结果有  $l$  种，记为事件  $A_j$  ( $j = 1, 2, \dots, l$ )，每种结果出现的概率为  $P(A_j) = p_j$ 。作  $n$  次独立的随机试验，事件  $A_j$  出现  $r_j$  次的概率分布为

$$M(\vec{r}; n, \vec{p}) = \frac{n!}{r_1! r_2! \cdots r_l!} p_1^{r_1} p_2^{r_2} \cdots p_l^{r_l} \quad (1.32)$$

### 4. 泊松分布

泊松分布是物理中极其重要的一个离散分布，许多的物理量服从泊松分布。泊松随机变量的取值为  $0, 1, 2, \dots$ ，泊松分布为

$$P(r; \mu) = \frac{1}{r!} \mu^r e^{-\mu} \quad (1.33)$$

期望值和方差均为  $\mu$ ，偏度系数  $\gamma_1 = 1/\sqrt{\mu}$ ，峰度系数  $\gamma_2 = 1/\mu$ 。

### 5. 均匀分布

$[a, b]$  区间上均匀分布的概率密度函数为

$$f(x) = \begin{cases} 1/(b-a), & a \leq x \leq b \\ 0, & \text{else} \end{cases} \quad (1.34)$$

其数字特征为

$$\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}, \quad \gamma_1 = 0, \quad \gamma_2 = -1.2 \quad (1.35)$$

## 6. 指数分布

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1.36)$$

## 7. 正态分布

正态分布又称为高斯分布，一维、 $n$  维正态分布的概率密度函数分别为

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = N(\mu, \sigma^2) \\ f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \left( \frac{x_1-\mu_1}{\sigma_1} \right)^2 + \left( \frac{x_2-\mu_2}{\sigma_2} \right)^2 \right. \right. \\ &\quad \left. \left. - 2\rho \left( \frac{x_1-\mu_1}{\sigma_1} \right) \left( \frac{x_2-\mu_2}{\sigma_2} \right) \right] \right\} \\ f(\vec{x}) &= \frac{1}{(2\pi)^{n/2} \sqrt{\det V}} e^{-Q/2} \end{aligned} \quad (1.37)$$

其中， $n$  维正态分布的二次型  $Q$  定义为

$$Q = (\vec{x} - \vec{\mu})^T V^{-1} (\vec{x} - \vec{\mu}) \quad (1.38)$$

而  $V$  则为  $n$  维正态随机向量  $X = (X_1, X_2, \dots, X_n)$  的协方差矩阵。

- 正态分布坐标定理：多维正态分布随机向量的各坐标满足一维正态分布。

$$\int N(\vec{\mu}, V) \prod_{j \neq i} dx_j = N(\mu_i, \sigma_i^2) \quad (1.39)$$

- 正态分布加法定理：相互独立的正态分布的和分布仍是正态分布。

$$\sum_i \alpha_i N(\mu_i, \sigma_i^2) = N(\mu, \sigma^2), \quad \left( \mu = \sum_i \alpha_i \mu_i, \quad \sigma^2 = \sum_i \alpha_i \sigma_i^2 \right) \quad (1.40)$$

### 8. 柯西 - 布雷特 - 维格纳分布

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \quad -\infty < x < +\infty \quad (1.41)$$

柯西 - 布雷特 - 维格纳分布广泛地应用于不稳定粒子的描述中。四动量为  $p^\mu$ ，质量为  $M$ ，衰变宽度为  $\Gamma$  的不稳定粒子为跃迁几率提供了一运动学因子

$$\frac{1}{(p^2 - M^2)^2 + M^2 \Gamma^2} \quad (1.42)$$

显然，这是一未归一化的布雷特 - 维格纳分布。

### 9. $\chi^2$ 分布

$$f(y; n) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}, \quad y \geq 0 \quad (1.43)$$

称为自由度为  $n$  的  $\chi^2$  分布，记为  $\chi^2(n)$ 。期望值和方差分别为  $\mu = n$ ， $\sigma^2 = 2n$ 。

- $\chi^2$  分布加法定理：相互独立的  $\chi^2$  分布的和分布仍是  $\chi^2$  分布。

$$\sum_i \chi^2(n_i) = \chi^2(n), \quad \left( n = \sum_i n_i \right) \quad (1.44)$$

- 与正态分布关系 (1)： $n$  个相互独立的标准正态变量的平方和服从  $\chi^2(n)$  分布。

$$\sum_{i=1}^n N(0, 1)^2 = \chi^2(n) \quad (1.45)$$

推论： $n$  维正态分布的二次型  $Q$  服从  $\chi^2(n)$  分布。

- 与正态分布关系 (2)： $n \rightarrow \infty$  的情形下， $\chi^2$  分布趋向于正态分布。

$$\lim_{n \rightarrow \infty} \frac{\chi^2(n) - n}{\sqrt{2n}} = N(0, 1) \quad (1.46)$$

- 与均匀分布关系： $X_i$  ( $i = 1, 2, \dots, n$ ) 为彼此独立的  $[0, 1]$  区间上均匀分布的随机变量，则

$$Y = -2 \sum_{i=1}^n \ln X_i \quad (1.47)$$

服从自由度为  $2n$  的  $\chi^2$  分布。

### 10. $t$ 分布

$$f(t; n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < +\infty \quad (1.48)$$

称为自由度为  $n$  的  $t$  分布，记为  $t(n)$ 。期望值和方差分别为  $\mu = 0$ ,  $\sigma^2 = \frac{n}{n-2}$ 。

- 与柯西分布关系：自由度为 1 的  $t$  分布为柯西分布。

$$f(t; 1) = \frac{1}{\pi} \frac{1}{1 + t^2} \quad (1.49)$$

- 与正态分布关系：  $n \rightarrow \infty$  的情形下，  $t$  分布趋向于标准正态分布。

$$\lim_{n \rightarrow \infty} t(n) = N(0, 1) \quad (1.50)$$

- $t$  分布意义：随机变量  $X \sim N(0, 1)$ ,  $Y \sim \chi^2(n)$ , 并且  $X$  和  $Y$  相互独立，则

$$t = \frac{X}{\sqrt{Y/n}} \quad (1.51)$$

服从自由度为  $n$  的  $t$  分布。

### 11. $F$ 分布

$$f(y; n_1, n_2) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} y^{\frac{n_1}{2}-1} \left(1 + \frac{n_1}{n_2}y\right)^{-\frac{n_1+n_2}{2}}, \quad y \geq 0 \quad (1.52)$$

称为自由度为  $(n_1, n_2)$  的  $F$  分布，记为  $F(n_1, n_2)$ 。

- $F$  分布意义：随机变量  $U_1 \sim \chi^2(n_1)$ ,  $U_2 \sim \chi^2(n_2)$ , 且  $U_1$  和  $U_2$  相互独立，则

$$Y = \frac{U_1/n_1}{U_2/n_2} \quad (1.53)$$

服从自由度为  $(n_1, n_2)$  的  $F$  分布。





## 第二章 蒙特卡洛方法

本章将介绍三部分的内容：1. 蒙特卡洛方法的数学基础；2. 随机变量（向量）的抽样方法；3. 减小蒙特卡洛方差的技巧。

### 2.1 蒙特卡洛方法的数学基础

通过两个简单的例子，引入蒙特卡洛方法，进而阐述作为蒙特卡洛方法数学基础的两个定理：大数定理和中心极限定理。

#### 1. Buffon 投针试验

试验方案：在平滑桌面上划一组相距为  $s$  的平行线，向此桌面随意地投掷长度  $l = s$  的细针，从针与平行线相交的概率可以得到  $\pi$  的数值。

Buffon 投针试验是伯努利试验，试验结果只有两种：针与平行线相交，针与平行线不相交。针与平行线相交的概率为

$$p = \int_0^\pi \frac{l |\cos \alpha|}{s} \frac{1}{\pi} d\alpha = \frac{2}{\pi} \quad (2.1)$$

$n$  次 Buffon 试验中针与平行线相交的次数  $X$  服从二项分布  $B(r; n, p)$ 。随机变量  $Y = X/n$  的期望值和方差分别为

$$\langle Y \rangle = p, \quad V(Y) = p(1 - p)/n \quad (2.2)$$

由此得到  $p = \langle Y \rangle$  的估计值和标准误差为

$$p \approx \frac{r}{n} \pm \sigma, \quad \sigma = \sqrt{\frac{p(1 - p)}{n}} \quad (2.3)$$

最终得到  $\pi = 2/p$  的估计值和标准误差为

$$\pi \approx \frac{2n}{r} \pm \sigma \quad (2.4)$$

## 2. 定积分计算

如图 2.1 所示，曲线  $y = f(x)$  定义于边长为 1 的正方形区域内，计算该曲线下面积。

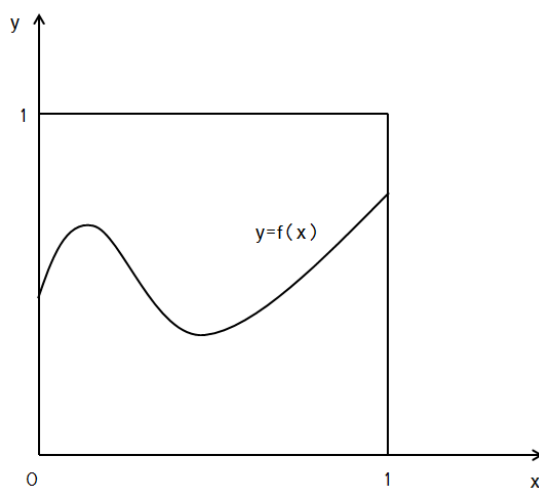


图 2.1 定积分计算示意图。

伯努利试验：正方形区域内随机投点。点落在曲线下方的概率  $p$  为曲线下面积

$$p = I = \int_0^1 f(x) dx \quad (2.5)$$

在正方形区域内均匀投下  $n$  个点，其中，落在曲线下方的点的个数  $X$  服从二项分布  $B(n, p)$ 。由二项分布的期望值公式知

$$I = \langle X/n \rangle \approx \frac{r}{n} \quad (2.6)$$

## 3. 大数定理

● 切比雪夫大数定理： $X_i$  ( $i = 1, 2, \dots$ ) 为相互独立的随机变量序列，它们有有限的期望值和方差，且方差有公共上界，则对于任意的  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n \mu_i \right| < \varepsilon \right) = 1 \quad (2.7)$$

推论：对于  $X_i$  ( $i = 1, 2, \dots$ ) 具有相同的期望值和方差的特殊情形，有

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| < \varepsilon \right) = 1 \quad (2.8)$$

• 辛钦大数定理： $X_i$  ( $i = 1, 2, \dots$ ) 为同分布的相互独立的随机变量序列，并且期望值有限，则对于任意的  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| < \varepsilon \right) = 1 \quad (2.9)$$

大数定理表明，当  $n$  充分大时，具有相同期望值的随机变量  $X_1, X_2, \dots, X_n$  的算术平均接近于它们的期望值  $\mu$ 。

• 定积分的蒙特卡洛估计值（期望值法，平均值法）：

$$\lim_{n \rightarrow \infty} \left[ I_n = \sum_{i=1}^n h(x_i) \right] = \left[ I = \int_a^b h(x)g(x)dx \right] \quad (2.10)$$

其中， $\{x_i | i = 1, 2, \dots, n\}$  为满足概率密度函数  $g(x)$  的随机数序列。

#### 4. 中心极限定理

• 同分布中心极限定理： $X_i$  ( $i = 1, 2, \dots$ ) 为同分布的相互独立的随机变量序列，并且期望值  $\mu$  和方差  $\sigma^2$  有限，则随机变量

$$Y = \frac{(X_1 + X_2 + \dots + X_n) - n\mu}{\sqrt{n}\sigma} \quad (2.11)$$

的分布函数  $F(y)$  对于任意的  $y$  值满足

$$\lim_{n \rightarrow \infty} F(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (2.12)$$

即当  $n \rightarrow \infty$  时，随机变量  $Y$  依概率收敛于标准正态分布  $N(0, 1)$ ，或随机变量  $X_1 + X_2 + \dots + X_n$  依概率收敛于正态分布  $N(n\mu, n\sigma^2)$ 。

• 蒙特卡洛估计值的方差（期望值法，平均值法）：

$$\lim_{n \rightarrow \infty} P \left( -\lambda \sqrt{\frac{V(h)}{n}} \leq I_n - I \leq \lambda \sqrt{\frac{V(h)}{n}} \right) = \int_{-\lambda}^{\lambda} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 1 - \alpha \quad (2.13)$$

## 2.2 随机变量的抽样方法

介绍随机变量的三类抽样方法：变换法，舍选法，Metropolis 法及抽样的迭加原理。

### 2.2.1 变换抽样法

#### 1. 离散型随机变量的直接抽样

$X$  为离散型随机变量，取值为  $x_i$  的概率为  $p_i$  ( $i = 1, 2, \dots$ )，如图 2.2 所示。

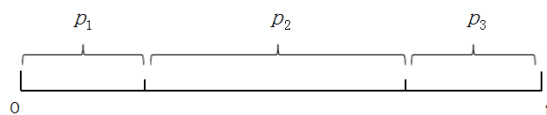


图 2.2 离散型随机变量直接抽样示意图。

- (1) 抽样生成  $[0, 1]$  区间均匀分布的随机数  $\xi$ ;
- (2) 判断  $\xi$  所属的区间段号  $j$ ;
- (3) 选取  $\eta = x_j$  作为抽样值。

#### 2. 连续型随机变量的直接抽样

连续型随机变量  $X$  的概率密度函数为  $f(x)$ ，相应的分布函数为  $F(x)$ 。在映射  $F^{-1}$  下， $[0, 1]$  区间上均匀分布的随机数序列映射为满足概率密度函数  $f(x)$  的随机数序列。

$$F^{-1}: \{\xi_1, \xi_2, \dots, \xi_n\} \longrightarrow \{x_1, x_2, \dots, x_n\} \quad (2.14)$$

例：直接抽样法生成指数分布随机数序列的映射为

$$x = -\frac{1}{\lambda} \ln \xi \quad (2.15)$$

#### 3. 连续型随机变量的变换抽样

随机变量  $X$  和  $Y$  满足的概率密度函数分布为  $f(x)$  和  $g(y)$ ，两者间由一一映射联系，则

$$f(x) = g(y) \left| \frac{dy}{dx} \right| \quad (2.16)$$

二维随机向量  $(X, Y)$  和  $(U, V)$  的概率密度函数分别为  $f(x, y)$  和  $g(u, v)$ ，两者间由一一映射联系，则

$$f(x, y) = g(u, v) \left| \det \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} \right| \quad (2.17)$$

例：  $(U, V)$  为  $[0, 1] \times [0, 1]$  区域上均匀分布的二维随机向量，  $(X, Y)$  和  $(U, V)$  间的映射为

$$\begin{cases} x = \sqrt{-2 \ln u} \cos(2\pi v) \\ y = \sqrt{-2 \ln u} \sin(2\pi v) \end{cases} \quad \text{or} \quad \begin{cases} u = e^{-\frac{1}{2}(x^2+y^2)} \\ v = \frac{1}{2\pi} \arctan \frac{y}{x} \end{cases} \quad (2.18)$$

则随机向量  $(X, Y)$  满足的概率密度函数为

$$f(x, y) = g(u, v) |J| = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} \quad (2.19)$$

即  $(X, Y)$  的两个坐标为相互独立的两个标准正态分布随机变量。

### 2.2.2 舍选抽样法

#### 1. 第一类舍选法

随机变量  $X$  在  $[a, b]$  上的概率密度函数为  $f(x)$ ，在区间  $[a, b]$  上  $f(x)$  的最大值存在，并等于

$$L = \max_{x \in [a, b]} f(x) \quad (2.20)$$

对这类问题采用舍选法的步骤为：

- (1) 抽样产生  $[0, 1]$  区间上两个均匀分布的随机数序列  $\{\xi_1^{(i)} | i = 1, 2, \dots\}$  和  $\{\xi_2^{(i)} | i = 1, 2, \dots\}$ ；
- (2) 利用映射  $\eta = a + (b - a)\xi_1$  生成  $[a, b]$  区间上均匀分布的随机数序列

$\{\eta^{(i)} | i = 1, 2, \dots\}$ ;

(3) 判选条件:

$$\xi_2 \leq \frac{1}{L} f(\eta) \quad (2.21)$$

当上述判选条件成立时, 保留  $\eta$ , 并取  $x = \eta$ ; 否则, 舍弃该点。

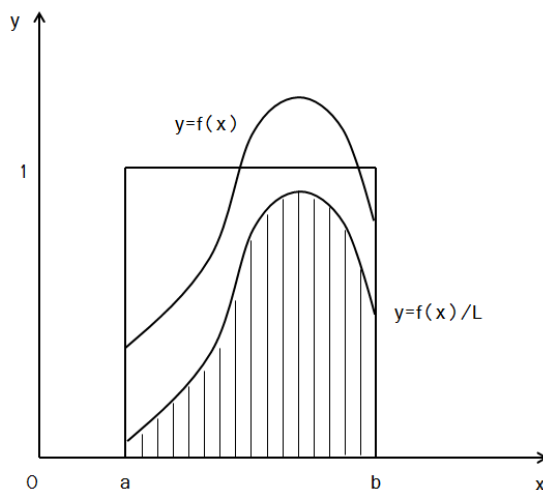


图 2.3 第一类舍选抽样法。

显然, 该舍选抽样法的舍选效率为

$$E = \frac{\int_a^b f(x) dx}{L(b-a)} = \frac{1}{L(b-a)} \quad (2.22)$$

例: 满足概率密度函数

$$f(x) = nx^{n-1}, \quad x \in [0, 1], \quad n = 0, 1, 2, \dots \quad (2.23)$$

的随机变量的抽样方法为

$$x = \max\{\xi_1, \xi_2, \dots, \xi_n\} \quad (2.24)$$

其中,  $\xi_i$  ( $i = 1, 2, \dots, n$ ) 为  $[0, 1]$  区间上的均匀分布的随机数。

## 2. 第二类舍选法

第二类舍选法是第一类舍选法的推广。若随机变量  $X$  在  $[a, b]$  上的概率密

度函数  $f(x)$  可以写为

$$f(x) = Lh(x)g(x) \quad (2.25)$$

其中,  $g(x)$  为概率密度函数;  $|h(x)| \leq 1$ 。相应的舍选法步骤为:

- (1) 抽样产生  $[0, 1]$  区间上均匀分布的随机数序列  $\{\xi^{(i)} | i = 1, 2, \dots\}$ ;
- (2) 按概率密度函数  $g(x)$  抽样产生  $[a, b]$  区间上的随机数序列  $\{\eta^{(i)} | i = 1, 2, \dots\}$ ;
- (3) 判选条件:

$$\xi \leq f(\eta) \quad (2.26)$$

当上述判选条件成立时, 保留  $\eta$ , 并取  $x = \eta$ ; 否则, 舍弃该点。

第二类舍选法的舍选效率为

$$E = \frac{\int_a^b h(x)g(x)dx}{\int_a^b g(x)dx} = \frac{1}{L} \quad (2.27)$$

例: 利用指数分布 ( $\lambda = 1$ ) 随机数序列, 借助第二类舍选法, 可以得到满足标准正态分布的随机数序列。

### 3. 第三类舍选法

如果概率密度函数可以表示成积分形式

$$f(x) = L \int_{-\infty}^{h(x)} g(x, y) dy \quad (2.28)$$

其中,  $g(x, y)$  是二维随机向量  $(X, Y)$  的联合概率密度函数,  $h(x)$  取值在  $y$  的定义域上。常数  $L$  定义为

$$L^{-1} = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{h(x)} dy g(x, y) \quad (2.29)$$

则舍选法步骤为:

- (1) 按联合概率密度函数  $g(x, y)$  抽样产生随机向量序列  $\{(\eta_x^{(i)}, \eta_y^{(i)}) | i = 1, 2, \dots\}$ ;
- (2) 判选条件:

$$\eta_y \leq h(\eta_x) \quad (2.30)$$

当上述判选条件成立时，保留  $\eta_x$ ，并取  $x = \eta_x$ ；否则，舍弃该点。

• 证明：

$$\begin{aligned}
 P(X \leq x) &= P(\eta_x \leq x \mid \eta_y \leq h(\eta_x)) \\
 &= \frac{P(\eta_x \leq x, \eta_y \leq h(\eta_x))}{P(\eta_y \leq h(\eta_x))} \\
 &= \frac{\int_{-\infty}^x dt_1 \int_{-\infty}^{h(t_1)} dt_2 g(t_1, t_2)}{\int_{-\infty}^{+\infty} dt_1 \int_{-\infty}^{h(t_1)} dt_2 g(t_1, t_2)} \\
 &= \int_{-\infty}^x f(t) dt
 \end{aligned} \tag{2.31}$$

• 三种舍选法间关系：

(a) 当  $(X, Y)$  两个分量彼此独立、并且  $Y$  为  $[0, 1]$  区间上均匀分布随机变量时，

$$g(x, y) = g_1(x)g_2(y), \quad g_2(y) = \begin{cases} 1, & y \in [0, 1] \\ 0, & \text{else} \end{cases} \tag{2.32}$$

$$f(x) = Lh(x)g_1(x), \quad 0 \leq h(x) \leq 1 \tag{2.33}$$

(b) 进一步取  $g_1(x) = 1/(b-a)$ ，有

$$f(x) = L(b-a)h(x) \tag{2.34}$$

以上两种情形分别对应于第一、第二类舍选法。

例：各向同性方位角余弦的抽样。

$$f(x) = \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}}, \quad |x| \leq 1 \tag{2.35}$$

第一步：由  $[0, 1] \times [0, 1]$  上均匀分布的随机向量  $(U, V)$  作映射得到概率密度函数为  $g(x, y)$  的二维随机向量  $(X, Y)$ ：

$$\begin{cases} x = (u^2 - v^2)/(u^2 + v^2) \\ y = u^2 + v^2 \end{cases}, \quad g(x, y) = \begin{cases} 1/(4\sqrt{1-x^2}), & |x| \leq 1, 0 \leq y \leq 1 \\ 0, & \text{else} \end{cases} \tag{2.36}$$



第二步：借助第三类舍选法由二维随机向量  $(X, Y)$  给出各向同性方位角余弦：

$$f(x) = \frac{4}{\pi} \int_{-\infty}^1 g(x, y) dy \quad (2.37)$$

抽样步骤：

- (1) 产生  $[0, 1]$  区间上的均匀分布的独立随机数  $u$  和  $v$ ，并计算出  $x$  和  $y$ ；
- (2) 判断  $y \leq 1$  是否成立。如成立，保留  $x$ ，并取方位角余弦  $\cos \theta = x$ ；否则，舍弃该点。

### 2.2.3 Metropolis 抽样法

Metropolis 方法是通过某种方式的“随机游走”来实现的。只要这个随机游走过程按照一定规则来进行，那末在进行大量的游走并达到平衡之后，所产生点的分布就满足所要求的分布。

● 细致平衡条件：

$$f(x)w(x \rightarrow x') = f(x')w(x' \rightarrow x) \quad (2.38)$$

通常选择过渡概率  $w(x \rightarrow x')$  为

$$w(x \rightarrow x') = \min \left[ 1, \frac{f(x')}{f(x)} \right] \quad (2.39)$$

由  $x_n$  到  $x_{n+1}$  点的游走步骤为：

- (1) 选取试探点  $x_{\text{try}} = x_n + \eta_n$ 。其中， $\eta_n$  为区间  $[-\delta, \delta]$  内均匀分布的随机数；
- (2) 计算  $r = f(x_{\text{try}})/f(x_n)$  的数值；
- (3) 如果  $r \geq 1$ ：取  $x_{n+1} = x_{\text{try}}$ 。返回 (1) 开始  $x_{n+1} \rightarrow x_{n+2}$  的游走试探；
- (4) 如果  $r < 1$ ：产生一个  $[0, 1]$  区间均匀分布的随机数  $\xi$ ；
- (5) 如果  $\xi \leq r$ ：取  $x_{n+1} = x_{\text{try}}$ 。返回 (1) 开始  $x_{n+1} \rightarrow x_{n+2}$  的游走试探；
- (6) 如果  $\xi > r$ ：拒绝游走到  $x_{\text{try}}$  点，即仍留在  $x_n$  点的位置不变；
- (7) 返回 (1) 开始  $x_{n+1} \rightarrow x_{n+2}$  的游走试探。

说明：Metropolis 法的弛豫时间由两个因素决定： $x_0$  and  $\delta$

## 2.2.4 迭加原理

分布密度函数  $f(x)$  为若干个分布密度函数  $g_i(x)$  ( $i = 1, 2, \dots, n$ ) 或  $g(x|y)$  ( $y \in \mathbb{R}$ ) 的加权平均

$$f(x) = \sum_{i=1}^n p_i g_i(x), \quad f(x) = \int_{-\infty}^{+\infty} h(y) g(x|y) dy \quad (2.40)$$

其中,  $p_i$  ( $i = 1, 2, \dots, n$ ) 和  $h(y)$  ( $y \in \mathbb{R}$ ) 为权重因子。

• 迭加原理:

- (1) 按概率  $p_i$  ( $i = 1, 2, \dots, n$ ) 或概率密度  $h(y)$  ( $y \in \mathbb{R}$ ) 抽样确定  $i$  或  $y$  的值;
- (2) 按概率密度函数  $g_i(x)$  或  $g(x|y)$  抽样。

例: 均匀分布的球壳抽样。

$$f(r) = \frac{3r^2}{R_1^3 - R_0^3} \quad (2.41)$$

作仿射变换:  $r = R_0 + (R_1 - R_0)x$ , 则随机变量  $x \in [0, 1]$  的概率密度函数为

$$f(x) = \frac{(R_1 - R_0)^2}{K} \cdot 3x^2 + \frac{3R_0(R_1 - R_0)}{K} \cdot 2x + \frac{3R_0^2}{K} \cdot 1 \quad (2.42)$$

利用舍选法, 结合迭加原理, 即可实现概率密度函数  $f(x)$  的抽样。

• 近似抽样法:

$$f(x) = m f_1(x) + (1 - m) L h(x) f_2(x) \quad (2.43)$$

$f_1(x)$  是  $f(x)$  的一个近似概率密度函数, 且易于抽样;  $0 \leq h(x) \leq 1$ ;  $m$  定义为

$$m = \min_{f_1(x) \neq 0} \frac{f(x)}{f_1(x)} \quad (2.44)$$

利用迭加原理及第二类舍选法,  $f(x)$  的抽样转化为  $f_1(x)$  或者  $f_2(x)$  的抽样。由于  $f_1(x)$  是  $f(x)$  的近似, 有

$$m \rightarrow 1, \quad (1 - m) \rightarrow 0 \quad (2.45)$$

所以,  $f(x)$  的抽样很大程度上是  $f_1(x)$  的抽样, 而  $f_2(x)$  的抽样出现在修正中。

## 2.2.5 条件概率密度法

(以三维随机向量为例,) 三维随机向量的联合概率密度函数写为

$$f(x_1, x_2, x_3) = f(x_1) \cdot f(x_2|x_1) \cdot f(x_3|x_1, x_2) \quad (2.46)$$

其中, 条件概率密度的定义为

$$f(x_2|x_1) = f(x_1, x_2)/f(x_1), \quad f(x_3|x_1, x_2) = f(x_1, x_2, x_3)/f(x_1, x_2) \quad (2.47)$$

边沿概率密度的定义为

$$f(x_1) = \int f(x_1, x_2, x_3) dx_2 dx_3, \quad f(x_1, x_2) = \int f(x_1, x_2, x_3) dx_3 \quad (2.48)$$

抽样步骤:

- (1) 按概率密度函数  $f(x_1)$  抽样, 确定  $x_1$  的值;
- (2) 按条件概率密度函数  $f(x_2|x_1)$  抽样, 确定  $x_2$  的值;
- (3) 按条件概率密度函数  $f(x_3|x_1, x_2)$  抽样, 确定  $x_3$  的值。

例: 中子入射角  $(\varphi, \theta)$  余弦抽样。

$$f(\varphi, \theta) = \frac{1}{K} \left( 1 + \sqrt{3} \sin \varphi \sin \theta \right) \sin \varphi \sin^2 \theta, \quad \varphi, \theta \in [0, \pi/2] \quad (2.49)$$

作余弦变换:  $x = \cos \varphi$ ,  $y = \cos \theta$ , 则随机向量  $(x, y)$  的联合概率密度函数为

$$f(x, y) = \frac{1}{K} \left[ 1 + \sqrt{3} \sqrt{(1-x^2)(1-y^2)} \right] \sqrt{1-y^2}, \quad x, y \in [0, 1] \quad (2.50)$$

引入如下的边沿概率密度及条件概率密度

$$\begin{aligned} f(x) &= \frac{1}{K} \left[ \frac{\pi}{4} + \frac{2\sqrt{3}}{3} \sqrt{1-x^2} \right] \\ f(y|x) &= \frac{1}{\left[ \frac{\pi}{4} + \frac{2\sqrt{3}}{3} \sqrt{1-x^2} \right]} \left[ 1 + \sqrt{3} \sqrt{(1-x^2)(1-y^2)} \right] \sqrt{1-y^2} \end{aligned} \quad (2.51)$$

利用迭加原理, 边沿概率密度  $f(x)$  可以写为

$$f(x) = p_1 \cdot 1 + p_2 \cdot \frac{4}{\pi} \sqrt{1-x^2} \quad (2.52)$$

条件概率密度可以写为

$$f(y|x) = q_1 \cdot \frac{4}{\pi} \sqrt{1-y^2} \cdot 1 + q_2 \cdot \frac{3}{2} (1-y^2) \quad (2.53)$$

其中，权重因子分别为

$$\begin{aligned} p_1 &\sim 1, & p_2 &\sim \frac{2\sqrt{3}}{3} \\ q_1 &\sim \frac{\pi}{4}, & q_2 &\sim \frac{2\sqrt{3}}{3} \sqrt{1-x^2} \end{aligned} \quad (2.54)$$

对概率密度函数  $\frac{4}{\pi}\sqrt{1-x^2}$  和  $\frac{3}{2}(1-x^2)$  的抽样则由舍选法实现。

## 2.3 减小蒙特卡洛方差的技巧

介绍减小蒙特卡洛方差的一些基本方法。

由中心极限定理知，期望值法给出的蒙特卡洛估计值的方差为

$$\sigma^2 = \frac{V(h)}{n} \quad (2.55)$$

由此式我们可以得到减小蒙特卡洛方差的两种基本方法。

### 1. 增加投点数

$$\lim_{n \rightarrow \infty} \sigma^2 = 0, \quad \frac{\partial \sigma^2}{\partial n} < 0, \quad \lim_{n \rightarrow \infty} \frac{\partial \sigma^2}{\partial n} = 0 \quad (2.56)$$

随着投点数  $n$  的增加，蒙特卡洛方差  $\sigma^2$  减小，但减小的速率趋近于 0。

### 2. 重要抽样法

$$I = \int_a^b h(x)g(x)dx = \langle h \rangle \quad (2.57)$$

被积函数为  $f(x) = h(x)g(x)$ 。选择概率密度函数  $g(x)$ ，使得

$$\frac{f(x)}{g(x)} = h(x) \sim \text{const.} \quad (2.58)$$

除了上述两种方法外，以黎曼积分可加性为基础的分层抽样法也是减小蒙特卡洛方差的方法之一。

### 3. 分层抽样法

$$\begin{aligned}
 I &= \int_a^b h(x)g(x)dx = \langle h \rangle \\
 &\parallel \\
 \sum_{j=1}^J p_j I_j &= \sum_{j=1}^J p_j \int_{x_{j-1}}^{x_j} h(x)g_j(x)dx = \sum_{j=1}^J p_j \langle h^{(j)} \rangle
 \end{aligned} \tag{2.59}$$

其中，

$$g_j(x) = \begin{cases} g(x)/p_j, & x_{j-1} \leq x < x_j \\ 0, & \text{else} \end{cases}, \quad p_j = \int_{x_{j-1}}^{x_j} g(x)dx \tag{2.60}$$

积分  $I$  和  $I_j$  ( $j = 1, 2, \dots, J$ ) 的蒙特卡洛方差分别为

$$\sigma^2 = \frac{V(h)}{N}, \quad \sigma_j^2 = \frac{V(h^{(j)})}{n_j} \tag{2.61}$$

则分层抽样法的蒙特卡洛方差为

$$\sigma'^2 = \sum_{j=1}^J p_j^2 \sigma_j^2 \tag{2.62}$$

推导：

$$\begin{aligned}
 V(h) &= \int_a^b [h(x) - I]^2 g(x)dx = \sum_{j=1}^J p_j \int_{x_{j-1}}^{x_j} [h(x) - I]^2 g_j(x)dx \\
 &= \sum_{j=1}^J p_j \int_{x_{j-1}}^{x_j} [h(x) - I_j + I_j - I]^2 g_j(x)dx \\
 &= \sum_{j=1}^J p_j \int_{x_{j-1}}^{x_j} [(h(x) - I_j)^2 + (I_j - I)^2 + 2(I_j - I)(h(x) - I_j)] g_j(x)dx \\
 &= \sum_{j=1}^J p_j V(h^{(j)}) + \sum_{j=1}^J p_j (I_j - I)^2
 \end{aligned} \tag{2.63}$$

所以，两种蒙特卡洛估计值的方差的差为

$$\begin{aligned}\sigma^2 - \sigma'^2 &= \frac{1}{N} \left[ \sum_{j=1}^J p_j V(h^{(j)}) + \sum_{j=1}^J p_j (I_j - I)^2 \right] - \sum_{j=1}^J \frac{p_j^2}{n_j} V(h^{(j)}) \\ &= \sum_{j=1}^J p_j \left( \frac{1}{N} - \frac{p_j}{n_j} \right) V(h^{(j)}) + \frac{1}{N} \sum_{j=1}^J p_j (I_j - I)^2\end{aligned}\quad (2.64)$$

若取

$$\frac{p_j}{n_j} = \frac{1}{N} \quad \longrightarrow \quad n_j = N p_j \quad (2.65)$$

此种分层抽样可以减小蒙特卡洛方差。

均匀分层抽样：

$$g(x) = \text{const.}, \quad x_j - x_{j-1} = \frac{b-a}{J}, \quad n_j = \frac{N}{J} \quad (2.66)$$

#### 4. 控制变量法

$$\int_a^b f(x) dx = \int_a^b [f(x) - g(x)] dx + \int_a^b g(x) dx \quad (2.67)$$

其中， $g(x)$  在区间  $[a, b]$  上的积分可以解析求出，且

$$f(x) - g(x) \sim \text{const.} \quad (2.68)$$

则  $f(x)$  的蒙特卡洛积分问题转化成为  $f(x) - g(x)$  的蒙特卡洛积分问题。与  $f(x)$  相比较而言， $f(x) - g(x)$  更为平坦；因此，蒙特卡洛积分的方差也更小。

## 第三章 蒙特卡洛方法的应用

本章将介绍蒙特卡洛方法在数学及物理（特别是粒子物理）中的应用。

### 3.1 定积分的掷点法求解

介绍求解定积分的掷点法，并给出该方法的蒙特卡洛解释。

讨论如图 2.1 所示的曲线  $y = f(x)$  下的面积（积分） $I$ 。掷点法给出的积分估计值为

$$I_n = \frac{r}{n} \quad (3.1)$$

其中， $n$  为总的均匀掷点数目， $r$  为曲线下点的数目。

• 解释一：

$$I = \langle X \rangle \quad (3.2)$$

其中， $X$  服从概率  $p = I$  的伯努利分布  $B(r; p)$ 。按照该解释，蒙特卡洛方差为

$$\sigma^2 = \frac{1}{n} V(X) = \frac{p(1-p)}{n} = \frac{I - I^2}{n} \quad (3.3)$$

• 解释二：

$$I = \langle h(X, Y) \rangle \quad (3.4)$$

其中， $(X, Y)$  为  $[0, 1] \times [0, 1]$  上均匀分布的随机向量；函数  $h(x, y)$  的定义为

$$h(x, y) = \begin{cases} 1, & y \leq f(x) \\ 0, & y > f(x) \end{cases} \quad (3.5)$$

按照该解释，蒙特卡洛方差为

$$\sigma^2 = \frac{1}{n} V(h) = \frac{1}{n} \left[ \langle h^2 \rangle - \langle h \rangle^2 \right] = \frac{I - I^2}{n} \quad (3.6)$$

显然，两种解释的结论是一致的。

• 与  $I = \langle f \rangle$  给出的蒙特卡洛估计值的比较：

该估计值的蒙特卡洛方差为

$$\sigma^2 = \frac{1}{n} V(f) = \frac{1}{n} \left[ \langle f^2 \rangle - \langle f \rangle^2 \right] = \frac{\langle f^2 \rangle - I^2}{n} \quad (3.7)$$

掷点法给出的蒙特卡洛估计值的方差与  $I = \langle f \rangle$  给出的蒙特卡洛估计值的方差之差为

$$\left( \frac{I - I^2}{n} \right) - \left( \frac{\langle f^2 \rangle - I^2}{n} \right) = \int_0^1 f(x) [1 - f(x)] dx \geq 0 \quad (3.8)$$

### 3.2 多体末态过程的相空间产生

介绍基本粒子物理研究中多体末态过程的相空间产生的因子化方法。

基本粒子物理中，描述两粒子对撞过程及单粒子衰变过程的反应几率的物理量分别为衰变宽度及对撞截面。定义分别为

$$\begin{aligned} d\Gamma &= \frac{1}{2M} |\mathcal{M}|^2 d\Phi_n(p; p_1, \dots, p_n) \\ d\sigma &= \frac{1}{4|\vec{p}_{a,\text{cm}}|\sqrt{s}} |\mathcal{M}|^2 d\Phi_n(p_a + p_b; p_1, \dots, p_n) \end{aligned} \quad (3.9)$$

其中， $d\Phi_n$  为  $n$  体末态相空间体积元

$$\begin{aligned} d\Phi_n(p; p_1, \dots, p_n) &= (2\pi)^4 \delta^{(4)}\left(p - \sum_{i=1}^n p_i\right) \prod_{i=1}^n \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i} \\ &= (2\pi)^4 \delta^{(4)}\left(p - \sum_{i=1}^n p_i\right) \prod_{i=1}^n \frac{d^4p_i}{(2\pi)^3} \theta(p_i^0) \delta(p_i^2 - m_i^2) \end{aligned} \quad (3.10)$$



相空间体积元可按如下公式因子化

$$d\Phi_n(p; p_1, \dots, p_n) = d\Phi_j(q; p_1, \dots, p_j) d \times \Phi_{n-j+1}(p; q, p_{j+1}, \dots, p_n) \times \frac{1}{2\pi} dM^2 \quad (3.11)$$

其中,  $M = \sqrt{q^2}$  为四动量  $q$  的不变质量。反复利用因子化公式可以得到

$$d\Phi_n(p; p_1, \dots, p_n) = \frac{1}{(2\pi)^{n-2}} \prod_{j=2}^{n-1} dM_j^2 \prod_{j=2}^n d\Phi_2(j) \quad (3.12)$$

两体末态相空间体积元  $d\Phi_2(j)$  为

$$d\Phi_2(j) = d\Phi_2(q_j; q_{j-1}, p_j), \quad M_j = \sqrt{q_j^2} \quad (3.13)$$

由因子化公式可以看到,  $n$  体末态相空间可以迭代地由两体末态相空间得到。

• 两体末态相空间:

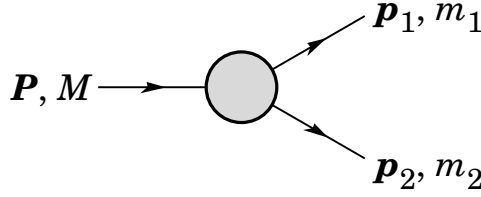


图 3.1 两体末态过程的相空间。

如图 3.1 所示, 初态的四动量和不变质量分别为  $P$  和  $M$ ; 末态粒子的四动量及质量分别为  $p_1, m_1$  和  $p_2, m_2$ 。在初态质心系中, 末态相空间体积元为

$$\begin{aligned} d\Phi_2(P; p_1, p_2) &= (2\pi)^4 \delta^{(4)}(p_1 + p_2 - P) \frac{d^3\vec{p}_1}{(2\pi)^3 2E_1} \frac{d^3\vec{p}_2}{(2\pi)^3 2E_2} \\ &= \frac{1}{(2\pi)^2} \frac{1}{4E_1 E_2} \delta(E_1 + E_2 - M) |\vec{p}_1|^2 d|\vec{p}_1| d\Omega_1 \\ &= \frac{1}{(2\pi)^2} \frac{|\vec{p}_1|^2}{4E_1 E_2} \left[ 1 \left/ \frac{\partial(E_1 + E_2)}{\partial|\vec{p}_1|} \right. \right] d\Omega_1 \end{aligned} \quad (3.14)$$

考虑到

$$\frac{\partial(E_1 + E_2)}{\partial|\vec{p}_1|} = \frac{|\vec{p}_1|}{E_1} + \frac{|\vec{p}_1|}{E_2} \quad (3.15)$$

初态质心系下两体末态相空间体积元可以写为

$$d\Phi_2(P; p_1, p_2) = \frac{1}{(2\pi)^2} \frac{|\vec{p}_1|}{4M} d\Omega_1 \quad (3.16)$$

其中,  $d\Omega_1 = d\varphi_1 d(\cos\theta_1)$  为末态粒子 1 的立体角微元,  $\varphi \in [0, 2\pi]$ ,  $\theta \in [0, \pi]$ ; 末态粒子 1 和 2 的动量大小为

$$|\vec{p}_1| = |\vec{p}_2| = \frac{1}{2M} \sqrt{\lambda(M^2, m_1^2, m_2^2)} \quad (3.17)$$

函数  $\lambda(x, y, z)$  的定义为

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2zx \quad (3.18)$$

• 两体末态相空间  $d\Phi_2(q_j; q_{j-1}, p_j)$  的产生步骤:

(1) 确定到  $q_j$  静止系的顺时正 Lorentz 变换  $\Lambda_j$ :

$$\Lambda_j : \begin{cases} q_j \mapsto (M_j, 0, 0, 0) \\ q_{j-1} \mapsto q'_{j-1} \\ p_j \mapsto p'_j \end{cases} \quad (3.19)$$

(2) 抽样产生  $\vec{p}_j$  的方位角和散射角:

$$\varphi_j = 2\pi\xi_{j,1}, \quad \cos\theta_j = -1 + 2\xi_{j,2} \quad (3.20)$$

(3) 抽样产生  $q'_{j-1}(q_{j-1})$  的不变质量  $M_{j-1}$ :

$$M_{j-1} = \sum_{i=1}^{j-1} m_i + \left( M - \sum_{i=1}^j m_i \right) \xi_{j,3} \quad (3.21)$$

(4) 确定  $\vec{p}_j$  的大小:

$$|\vec{p}_j| = \frac{1}{2M_j} \sqrt{\lambda(M_j^2, M_{j-1}^2, m_j^2)} \quad (3.22)$$

(5) 确定  $p'_j$  和  $q'_{j-1}$ :

$$p'_j = \left( \sqrt{|\vec{p}_j|^2 + m_j^2}, \vec{p}_j \right), \quad q'_{j-1} = \left( \sqrt{|\vec{p}_j|^2 + M_{j-1}^2}, -\vec{p}_j \right) \quad (3.23)$$

(6) 作 Lorentz 变换  $\Lambda_j^{-1}$ :

$$\Lambda_j^{-1} : \begin{cases} q'_{j-1} \mapsto q_{j-1} \\ p'_j \mapsto p_j \end{cases} \quad (3.24)$$

说明：上述两体末态相空间产生步骤正确性的基础是相空间体积元的顺时正 Lorentz 不变性。

•  $n$  体末态相空间  $d\Phi_n(p; p_1, \dots, p_n)$  的产生:

$$j = n \rightarrow (n-1) \rightarrow \dots \rightarrow 3 \rightarrow 2 \quad (3.25)$$

按上述顺序递归地进行两体末态相空间产生，即可得到  $n$  体末态相空间。其中

$$q_n = p, \quad M_n = \sqrt{p^2}, \quad q_1 = p_1, \quad M_1 = m_1 \quad (3.26)$$

该方法产生的相空间点的权重因子为

$$\begin{aligned} w &= \frac{1}{(2\pi)^{n-2}} \prod_{j=2}^{n-1} 2M_j \prod_{j=2}^n \frac{1}{(2\pi)^2} \frac{\sqrt{\lambda(M_j^2, M_{j-1}^2, m_j^2)}}{8M_j^2} \\ &= (2\pi)^{4-3n} 2^{1-2n} \frac{1}{M_n} \prod_{j=2}^n \frac{\sqrt{\lambda(M_j^2, M_{j-1}^2, m_j^2)}}{M_j} \end{aligned} \quad (3.27)$$

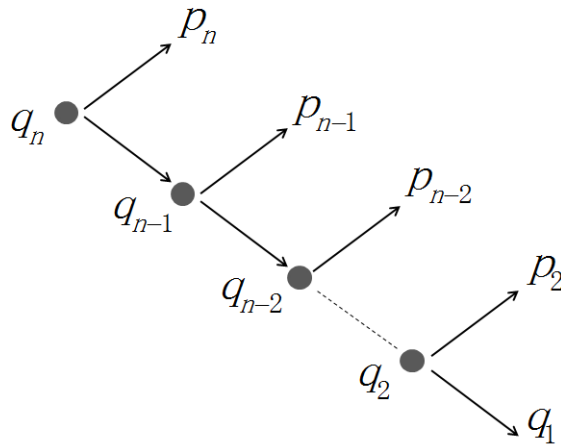


图 3.2  $n$  体末态相空间的因子化产生方法。

### 3.3 量子力学路径积分形式的求解

介绍 Feynman 传播子的路径积分形式及随机向量抽样的 Metropolis 方法，并以路径积分为例给出 Metropolis 方法的物理图像解释。

量子力学的正则形式可以在 Schrodinger 和 Heisenberg 两种图像下表达：

$$\begin{aligned} \text{Schrodinger picture :} \quad i \frac{\partial}{\partial t} |t\rangle &= \hat{H} |t\rangle \\ \frac{\partial}{\partial t} \hat{O} &= 0 \end{aligned} \quad (3.28)$$

$$\begin{aligned} \text{Heisenberg picture :} \quad i \frac{\partial}{\partial t} |t\rangle_H &= 0 \\ \frac{\partial}{\partial t} \hat{O}_H(t) &= i [\hat{H}, \hat{O}_H(t)] \end{aligned} \quad (3.29)$$

两种图像间的关系为

$$|t\rangle_H = e^{i\hat{H}t} |t\rangle, \quad \hat{O}_H(t) = e^{i\hat{H}t} \hat{O} e^{-i\hat{H}t} \quad (3.30)$$

利用 Schrodinger 方程得到两种图像下的时间演化算符分别为

$$\hat{U}(t, t_0) = e^{-i\hat{H}(t-t_0)}, \quad \hat{U}_H(t, t_0) = 1 \quad (3.31)$$

两者在坐标算符本征态下的矩阵元相同，称为 Feynman 传播子。

$$D_F(\vec{r}, t; \vec{r}_0, t_0) = \langle \vec{r} | \hat{U}(t, t_0) | \vec{r}_0 \rangle = \langle \vec{r}, t | \vec{r}_0, t_0 \rangle_H \quad (3.32)$$

• Feynman 传播子  $D_F(x, t; x_0, 0)$  的路径积分形式：

$$D_F(x, t; x_0, 0) = \int \prod_{j=1}^{N-1} dx_j \prod_{j=1}^N \langle x_j, t_j | x_{j-1}, t_{j-1} \rangle_H \quad (3.33)$$

这里，我们已将时间区间  $[0, t]$  按步长  $\varepsilon = t/N$  等间距分割为

$$0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = t \quad (3.34)$$

并在各时间节点上利用了坐标本征态的完备性。

$$\begin{aligned}
 \langle x_j, t_j | x_{j-1}, t_{j-1} \rangle_H &= \langle x_j | e^{-i\varepsilon \hat{H}} | x_{j-1} \rangle \\
 &= \langle x_j | 1 - i\varepsilon \left( \frac{\hat{p}^2}{2m} + V(\hat{x}) \right) | x_{j-1} \rangle + \mathcal{O}(\varepsilon^2) \\
 &= \int \frac{dp}{2\pi} \left[ 1 - i\varepsilon \left( \frac{p^2}{2m} + V(x_j) \right) \right] e^{ip(x_j - x_{j-1})} + \mathcal{O}(\varepsilon^2) \\
 &= \int \frac{dp}{2\pi} \exp \left\{ -i\varepsilon \left( \frac{p^2}{2m} + V(x_j) \right) + ip(x_j - x_{j-1}) \right\} + \mathcal{O}(\varepsilon^2) \\
 &= \sqrt{\frac{m}{2\pi i\varepsilon}} \exp \left\{ i\varepsilon \left[ \frac{1}{2}m \left( \frac{x_j - x_{j-1}}{\varepsilon} \right)^2 - V(x_j) \right] \right\} + \mathcal{O}(\varepsilon^2)
 \end{aligned} \tag{3.35}$$

记  $x = x_N$ ，则有

$$\begin{aligned}
 D_F(x, t; x_0, 0) &= \lim_{N \rightarrow \infty} \left( \frac{m}{2\pi i\varepsilon} \right)^{\frac{N}{2}} \int \prod_{j=1}^{N-1} dx_j \exp \left\{ i\varepsilon \sum_{j=1}^N \left[ \frac{1}{2}m \left( \frac{x_j - x_{j-1}}{\varepsilon} \right)^2 - V(x_j) \right] \right\} \\
 &= \int \mathcal{D}[x(t)] \exp \left( i \int_0^t dt L(t) \right)
 \end{aligned} \tag{3.36}$$

• 积分路径的 Metropolis 抽样：

$$\begin{aligned}
 D_F(x, -it; x_0, 0) &= \int \mathcal{D}[x(t)] \exp \left( - \int_0^t dt H(t) \right) \\
 &= \lim_{N \rightarrow \infty} \left( \frac{m}{2\pi i\varepsilon} \right)^{\frac{N}{2}} \int \prod_{j=1}^{N-1} dx_j \exp \left\{ -\varepsilon \sum_{j=1}^N \left[ \frac{1}{2}m \left( \frac{x_j - x_{j-1}}{\varepsilon} \right)^2 + V(x_j) \right] \right\}
 \end{aligned} \tag{3.37}$$

通常用到的是 Feynman 传播子在虚时下的渐进展开。例如：基态能量  $E_0$  和基态波函数  $\psi_0(x)$  分别可以表达为

$$\begin{aligned}
 \int D_F(x, -it; x, 0) dx &\longrightarrow e^{-E_0 t}, \quad (t \rightarrow +\infty) \\
 |\psi_0(x)|^2 &= \lim_{t \rightarrow +\infty} \frac{D_F(x, -it; x, 0)}{e^{-E_0 t}}
 \end{aligned} \tag{3.38}$$

这些积分的计算需要采用重要抽样法，概率密度函数为

$$f(x_0, x_1, \dots, x_N) \sim \exp \left( -\varepsilon E(x_0, x_1, \dots, x_N) \right) \tag{3.39}$$

其中, 函数  $E(x_0, x_1, \dots, x_N)$  的定义为

$$E(x_0, x_1, \dots, x_N) = \sum_{j=1}^N \left[ \frac{1}{2} m \left( \frac{x_j - x_{j-1}}{\varepsilon} \right)^2 + V(x_j) \right] \quad (3.40)$$

概率密度函数  $f(x_0, x_1, \dots, x_N)$  描述的随机向量为以下几种情形之一:

- (1)  $(N+1)$  维随机向量  $(x_0, x_1, \dots, x_N)$ ;
- (2)  $(N-1)$  维随机向量  $(x_1, x_2, \dots, x_{N-1})$ ;
- (3)  $N$  维随机向量  $(x_0, x_1, \dots, x_N)$ , 其中  $x_0 = x_N$ 。

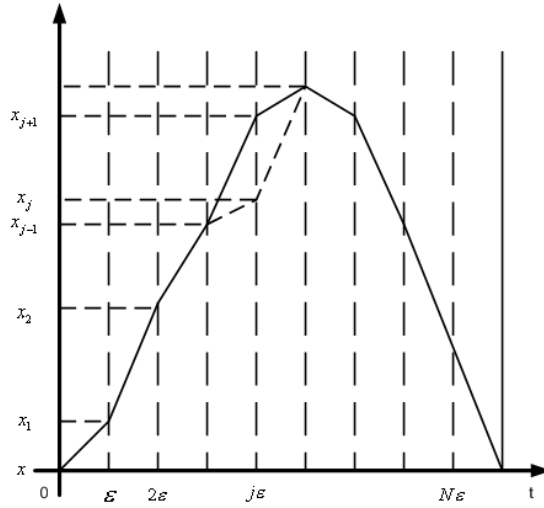


图 3.3 积分路径的 Metropolis 抽样。

如图 3.3 所示, 向量  $(x_0, x_1, \dots, x_N)$  可以等价地表述为路径  $L$ 。借助 Metropolis 方法, 可以实现随机向量  $(x_0, x_1, \dots, x_N)$  (路径  $L$ ) 的抽样, 得到随机向量 (路径) 序列

$$\left\{ L_n = \left( x_0^{(n)}, x_1^{(n)}, \dots, x_N^{(n)} \right) \mid n = 1, 2, \dots \right\} \quad (3.41)$$

由路径  $L_n$  到路径  $L_{n+1}$  点的游走步骤为:

- (1) 选取试探路径:

$$L_{\text{try}} = L_n + \left( 0, \dots, 0, x_j^{(n)} + \eta^{(n)}, 0, \dots, 0 \right) \quad (3.42)$$

其中,  $j \in \{0, 1, \dots, N\}$ ,  $\eta^{(n)} \in [-\delta, \delta]$  为均匀分布的随机数;

(2) 计算过渡概率:

$$w(L_n \rightarrow L_{\text{try}}) = \min \left[ 1, e^{-\varepsilon(E_{\text{try}} - E_n)} \right] \quad (3.43)$$

其中,  $E_n$  和  $E_{\text{try}}$  为函数  $E(x_0, x_1, \dots, x_N)$  在路径  $L_n$  和  $L_{\text{try}}$  上的值;

(3) 依过渡概率  $w(L_n \rightarrow L_{\text{try}})$  接受  $L_{\text{try}}$  为  $L_{n+1}$ , 即:

$$P(L_{n+1} = L_{\text{try}}) = w(L_n \rightarrow L_{\text{try}}) \quad (3.44)$$

• Metropolis 方法的物理含义:

$$w(L \rightarrow L') = \min \left[ 1, e^{-\varepsilon(E_{L'} - E_L)} \right] \quad (3.45)$$

由路径  $L$  经游走得到的路径  $L'$  被接受的概率为

$$P(L' \text{ is accepted}) = \begin{cases} 1, & E_{L'} < E_L \\ e^{-\varepsilon \Delta E}, & E_{L'} > E_L \end{cases} \quad (3.46)$$

其中,  $\Delta E = E_{L'} - E_L$ 。显然, 上式反映的正是 (1). 能量最低原理; (2). Boltzman 分布律。

### 3.4 量子力学变分形式的求解

利用极值原理, 给出基态能量及基态波函数的变分形式, 并用蒙特卡洛方法解决之。

在 Hilbert 空间扫描所有可能的物理状态, 基态是使得

$$E = \frac{\int \psi^*(\vec{r}) \hat{H} \psi(\vec{r}) d^3 \vec{r}}{\int |\psi(\vec{r})|^2 d^3 \vec{r}} \quad (3.47)$$

取极小值的状态。我们首先选择一个试探波函数  $\psi_0(x)$ , 然后计算在此试探波函数下的变分能量, 从而寻找基态波函数和基态能量。

引入随机向量  $\vec{r}$ ，其满足的概率密度函数为

$$g(\vec{r}) \sim |\psi(\vec{r})|^2 \quad (3.48)$$

$\epsilon$  为随机向量  $\vec{r}$  的函数，定义为

$$\epsilon(\vec{r}) = \psi^{-1}(\vec{r}) \hat{H} \psi(\vec{r}) \quad (3.49)$$

按照大数定理，期望值  $E$  的蒙特卡洛估计值为

$$E = \langle \epsilon(\vec{r}) \rangle \approx \frac{1}{N} \sum_{n=1}^N \epsilon(\vec{r}_n) \quad (3.50)$$

其中，序列  $\{\vec{r}_n | n = 1, 2, \dots, N\}$  服从概率密度函数  $g(\vec{r})$ 。

• 选择一个物理上合理的近似基态波函数  $\psi_0(\vec{r})$  作为试探波函数，开始如下的基态波函数试探步骤  $\psi_i(\vec{r}) \rightarrow \psi_{i+1}(\vec{r})$ ：

(1) 改变试探波函数中的变分参数值，得到新的试探波函数  $\psi_{i+1}(\vec{r})$ ，要求

$$|\psi_{i+1}(\vec{r}) - \psi_i(\vec{r})| \leq \delta \quad (3.51)$$

(2) 采用 Metropolis 方法抽样生成满足概率密度函数  $|\psi_{i+1}(\vec{r})|^2$  的随机向量序列

$$\{\vec{r}_n | n = 1, 2, \dots, N\} \quad (3.52)$$

(3) 计算能量  $E$  的蒙特卡洛估计值

$$E_{\text{try}}^{(i+1)} = \frac{1}{N} \sum_{n=1}^N \epsilon(\vec{r}_n), \quad \Delta E_{i+1} = E_{\text{try}}^{(i+1)} - E_{\text{try}}^{(i)} \quad (3.53)$$

(4) 试探波函数  $\psi_{i+1}(\vec{r})$  被接受，如果

$$\Delta E_{i+1} \leq 0 \quad (3.54)$$

说明：循环往复进行如上的试探过程，直至能量估计值不再有明显的改变，则最后一次试探得到的波函数  $\psi_f(\vec{r})$  及能量估计值  $E_f$  为基态波的一组蒙特卡洛估计。该方法的蒙特卡洛方差强烈依赖于试探波函数的具体形式。



### 3.5 中子输运问题的蒙特卡洛模拟

介绍中子输运过程的物理图像，给出了中子透射率的三种不同的蒙特卡洛估计，并比较了它们的方差。

中子在介质中的传输是个典型的随机过程，蒙特卡洛方法是解决中子输运这类本身就具有统计特性的问题的最直接的方法。

#### 1. 直接模拟法

输运过程是典型的随机统计过程，物理观测量是大量粒子的行为共同贡献的统计平均结果。

- (1) 模拟粒子在介质中随机运动的历史，记录其对感兴趣的物理模拟量的贡献；
- (2) 大量的重复模拟之后，对物理模拟量进行统计平均。

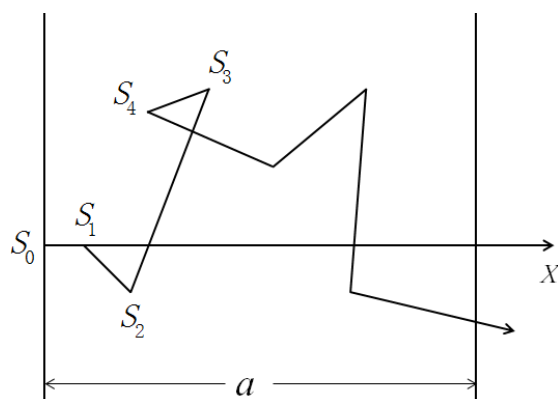


图 3.4 中子在介质中的输运模拟。

- 中子与介质的相互作用的截面描述：

$$\begin{aligned}
 \text{scattering cross section : } & \sigma_s^m(E) \\
 \text{absorbing cross section : } & \sigma_a^m(E) \\
 \text{full cross section : } & \sigma_t^m(E)
 \end{aligned} \tag{3.55}$$

- 模拟跟踪中子的位形：

$$S = (x, E, \cos \theta) \tag{3.56}$$

中子的初始位形为

$$S_0 = (x_0 = 0, E_0, \cos \theta_0 = 1) \quad (3.57)$$

•  $S_{i-1} \rightarrow S_i$  的模拟中依序完成如下随机变量的抽样：

(1) 中子自由程：

$$f(y) = \sigma_T(E_{i-1}) \exp[-\sigma_T(E_{i-1})y] \quad (3.58)$$

其中，自由程（指数分布随机变量） $y$  和中子作用于单位体积介质的总截面  $\sigma_T(E)$  分别为

$$y = \frac{x_i - x_{i-1}}{\cos \theta_{i-1}}, \quad \sigma_T(E) = \sum_m \rho_m \sigma_t^m(E) \quad (3.59)$$

(2) 与中子碰撞的介质中核子种类：

$$P(\text{nucleon} = m) = p_m = \frac{\sigma_T^m(E_{i-1})}{\sigma_T(E_{i-1})} \quad (3.60)$$

其中，中子与介质单位体积中第  $m$  种核子反应的总截面  $\sigma_T^m(E)$  定义为

$$\sigma_T^m(E) = \rho_m \sigma_t^m(E) \quad (3.61)$$

(3) 中子与第  $m$  种核子的反应类型：

$$P(\text{reaction}) = \begin{cases} p_s, & \text{reaction} = \text{scattering} \\ p_a, & \text{reaction} = \text{absorption} \end{cases} \quad (3.62)$$

其中，散射及吸收概率分别为

$$p_s = \frac{\sigma_s^m(E_{i-1})}{\sigma_t^m(E_{i-1})}, \quad p_a = \frac{\sigma_a^m(E_{i-1})}{\sigma_t^m(E_{i-1})}, \quad p_s + p_a = 1 \quad (3.63)$$

(4) 中子与第  $m$  种核子散射的散射角：

$$f(\cos \theta_c) = \frac{d\sigma_s^m(E_{i-1})}{d\cos \theta_c} \bigg/ \int_{-1}^1 \frac{d\sigma_s^m(E_{i-1})}{d\cos \theta_c} d\cos \theta_c \quad (3.64)$$

其中， $\theta_c$  为 neutron-nucleon 质心系中的散射角。

(5) 中子与第  $m$  种核子散射的方位角:

$$f(\varphi) = \frac{1}{2\pi} \quad (3.65)$$

• 位形  $S_i$  的确定:

$$S_i = \left( x_{i-1} + y \cos \theta_{i-1}, \frac{1}{2} E_{i-1} \left[ (1+r) + (1-r) \cos \theta_c \right], \cos \theta_{i-1} \cos \theta_L + \sin \theta_{i-1} \sin \theta_L \cos \varphi \right)$$

$$\begin{array}{ccc} \Downarrow & \Downarrow & \Downarrow \\ x_i & E_i & \cos \theta_i \end{array}$$

其中, 实验室系中的散射角  $\theta_L$  为

$$\cos \theta_L = (1 + A \cos \theta_c) / \sqrt{1 + A^2 + 2A \cos \theta_c} \quad (3.66)$$

参数  $r$  和  $A$  的定义为

$$r = \left( \frac{A-1}{A+1} \right)^2, \quad A = \frac{m_{\text{nucleon}}}{m_{\text{neutron}}} \quad (3.67)$$

• 中子透射率:

单个中子透射试验为伯努利试验。中子透射次数  $X = r \in \{0, 1\}$  为随机变量, 满足概率 (透射率) 为  $p$  的伯努利分布  $B(r; p)$ 。由伯努利分布知

$$\langle X \rangle = p, \quad V(X) = p(1-p) \quad (3.68)$$

中子输运过程的蒙特卡洛模拟给出随机变量  $X$  的等价描述 (即: 随机变量  $X$  的抽样方法) 为

$$X = r = \begin{cases} 1, & x_M \geq a \\ 0, & \text{else} \end{cases} \quad (3.69)$$

由  $p = \langle X \rangle$  诱导出的透射率  $p$  的蒙特卡洛估计值及其方差分别为

$$\bar{P} = \frac{1}{N} \sum_{n=1}^N r_n, \quad \sigma_{\bar{P}}^2 = \frac{1}{N} p(1-p) \quad (3.70)$$

其中,  $\{r_n | n = 1, 2, \dots, N\}$  为满足伯努利分布的随机数序列。  $N$  次反复独立地模

拟中子在介质中的输运，即可得到该随机数序列。

$$\begin{array}{ccc}
 N & & N_1 = \sum_{n=1}^N r_n \\
 \Downarrow & \text{---} \times P \text{---} & \Downarrow \\
 \text{incoming} & & \text{outgoing}
 \end{array} \quad (3.71)$$

## 2. 权重法

采用权重法时，引入散射的概率权重  $w$ 。此时中子的状态描写为

$$S = (x, E, \cos \theta, w) \quad (3.72)$$

$S_{i-1} \rightarrow S_i$  的模拟中， $w_i$  由下面公式确定

$$w_i = w_{i-1} \frac{\sigma_s^m(E_{i-1})}{\sigma_t^m(E_{i-1})}, \quad w_0 = 1 \quad (3.73)$$

### • 中子透射率：

我们知道，随机变量的抽样方法等价地描述了随机变量的概率密度函数。引入连续型随机变量  $Y = y \in [0, 1]$ ，其抽样方法为

$$Y = y = \begin{cases} w_{M-1}, & x_M \geq a \\ 0, & \text{else} \end{cases} \quad (3.74)$$

则该随机变量的期望值和方差分别为

$$\langle Y \rangle = p, \quad V(Y) = \langle Y^2 \rangle - \langle Y \rangle^2 \quad (3.75)$$

由  $p = \langle Y \rangle$  诱导出的透射率  $p$  的蒙特卡洛估计值及其方差分别为

$$\bar{P}' = \frac{1}{N} \sum_{n=1}^N y_n, \quad \sigma_{\bar{P}'}^2 = \frac{1}{N} \left( \langle Y^2 \rangle - p^2 \right) \quad (3.76)$$

### • 权重法与直接模拟法方差的比较：

从本质上说，直接模拟法等价于求期望值积分的“掷点法”，而权重法则相当于“平均值法”。我们知道，平均值法比掷点法的方差要小；因此，权重法的方差

应小于直接模拟法的方差。证明如下

$$\because w_{M-1} < 1, \quad w_{M-1}^2 < w_{M-1} \quad \therefore \langle Y^2 \rangle < p, \quad \sigma_{\bar{P}'}^2 < \sigma_{\bar{P}}^2 \quad (3.77)$$

### 3. 统计估计法

- 中子透射率:

引入连续型随机变量  $Z = z \in [0, +\infty)$ , 其抽样方法为

$$Z = z = \sum_{i=0}^{M-1} p^{(i)} \quad (3.78)$$

其中,  $p^{(i)}$  是在碰撞点  $i$  处的透射概率, 定义为

$$p^{(i)} = \begin{cases} w_i \exp \left\{ -\sigma_T(E_i) \frac{a - x_i}{\cos \theta_i} \right\}, & \cos \theta_i > 0 \\ 0, & \text{else} \end{cases} \quad (3.79)$$

随机变量  $Z$  的期望值和方差分别为

$$\langle Z \rangle = p, \quad V(Z) = \langle Z^2 \rangle - \langle Z \rangle^2 \quad (3.80)$$

由  $p = \langle Z \rangle$  诱导出的透射率  $p$  的蒙特卡洛估计值为

$$\bar{P}'' = \frac{1}{N} \sum_{n=1}^N z_n = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{M-1} p_n^{(i)} \quad (3.81)$$

相应的蒙特卡洛方差为

$$\sigma_{\bar{P}''}^2 = \frac{1}{N} \left( \langle Z^2 \rangle - p^2 \right) \quad (3.82)$$

- 方差评论:

随机变量  $X, Y, Z$  的定义域不同, 因此二阶原点矩的性质不同:

$$\begin{array}{ccc} r \in \{0, 1\} & y \in [0, 1] & z \in [0, +\infty) \\ \Downarrow & \Downarrow & \Downarrow \\ \langle X^2 \rangle = \langle X \rangle & \langle Y^2 \rangle < \langle Y \rangle & \langle Z^2 \rangle = ? \end{array} \quad (3.83)$$

透射率的三种估计值方差间的等级关系为

$$\left[ \sigma_P^2 > \sigma_{P'}^2 \right] \sim \sigma_{P''}^2 \quad (3.84)$$

## 第四章 偏微分方程的求解

本章将介绍数值求解偏微分方程的两种方法：有限差分法和有限元素法。

### 4.1 有限差分方法

介绍一阶、二阶微商的各种常用差商格式，区域的网格划分格式，以及边界条件的离散化方法。

#### 4.1.1 微商的差商格式

以一元函数为例，函数  $f(x)$  可以做如下形式的 Talor 展开

$$\begin{cases} f(x+h_1) = f(x) + h_1 f'(x) + \frac{1}{2} h_1^2 f''(x) + \frac{1}{3!} h_1^3 f'''(x) + \cdots \\ f(x-h_2) = f(x) - h_2 f'(x) + \frac{1}{2} h_2^2 f''(x) - \frac{1}{3!} h_2^3 f'''(x) + \cdots \end{cases} \quad (4.1)$$

由此得到

$$\begin{aligned} & \alpha [f(x+h_1) - f(x)] + \beta [f(x-h_2) - f(x)] \\ &= (\alpha h_1 - \beta h_2) f'(x) + \frac{1}{2} (\alpha h_1^2 + \beta h_2^2) f''(x) + \frac{1}{3!} (\alpha h_1^3 - \beta h_2^3) f'''(x) + \cdots \end{aligned} \quad (4.2)$$

#### 1. 一阶微商的差商格式

- 向前差商格式：

$$\begin{aligned} f'(x) &= \frac{1}{h} [f(x+h) - f(x)] + \mathcal{O}(h) \\ \text{truncation error : } \mathcal{O}(h) &\approx -\frac{1}{2} h f''(x) \end{aligned} \quad (4.3)$$

- 向后差商格式:

$$f'(x) = \frac{1}{h} [f(h) - f(x-h)] + \mathcal{O}(h) \quad (4.4)$$

$$\text{truncation error : } \mathcal{O}(h) \approx \frac{1}{2} h f''(x)$$

- 中心差商格式:

$$\alpha h_1^2 + \beta h_2^2 = 0 \quad (4.5)$$

$$\downarrow$$

$$f'(x) = \frac{1}{h_1 + h_2} \left[ \frac{h_2}{h_1} f(x+h_1) - \frac{h_1}{h_2} f(x-h_2) \right] + \frac{h_1 - h_2}{h_1 h_2} f(x) + \mathcal{O}(h^2)$$

$$\text{truncation error : } \mathcal{O}(h^2) \approx -\frac{1}{3!} h_1 h_2 f'''(x)$$

可以看到, 向前、向后差商格式的截断误差为  $\mathcal{O}(h)$ , 而中心差商格式的截断误差为  $\mathcal{O}(h^2)$ 。

## 2. 二阶微商的差商格式

$$\alpha h_1 - \beta h_2 = 0 \quad (4.6)$$

$$\downarrow$$

$$f''(x) = \frac{2}{h_1 h_2 (h_1 + h_2)} \left[ h_2 f(x+h_1) + h_1 f(x-h_2) - (h_1 + h_2) f(x) \right] + \mathcal{O}(h)$$

$$\text{truncation error : } \mathcal{O}(h) \approx -\frac{1}{3} (h_1 - h_2) f'''(x)$$

一般情况下, 二阶微商的差商格式的截断误差为  $\mathcal{O}(h)$ ; 当  $h_1 = h_2$  时, 截断误差降低为  $\mathcal{O}(h^2)$ 。

## 3. 二维 Laplace 算子的五点差商格式

记向右、向上、向左、向下的差分步长分别为  $h_1, h_2, h_3$  和  $h_4$ , 则

$$(\nabla^2 \phi)_0 \approx 2 \left[ \frac{h_3 (\phi_1 - \phi_0) + h_1 (\phi_3 - \phi_0)}{h_1 h_3 (h_1 + h_3)} + \frac{h_4 (\phi_2 - \phi_0) + h_2 (\phi_4 - \phi_0)}{h_2 h_4 (h_2 + h_4)} \right] \quad (4.7)$$

其中,  $\phi_{0,1,2,3,4}$  分别为函数  $\phi$  在  $(x, y), (x+h, y), (x, y+h), (x-h, y), (x, y-h)$



处的值； $(\nabla^2\phi)_0$  为  $\nabla^2\phi$  在  $(x, y)$  处的值。

$$h_1 = h_3 = h_x, \quad h_2 = h_4 = h_y \quad (4.8)$$

$$\downarrow$$

$$(\nabla^2\phi)_0 \approx \frac{\phi_1 + \phi_3 - 2\phi_0}{h_x^2} + \frac{\phi_2 + \phi_4 - 2\phi_0}{h_y^2}$$

$$h_x = h_y = h \quad (4.9)$$

$$\downarrow$$

$$(\nabla^2\phi)_0 \approx \frac{1}{h^2} (\phi_1 + \phi_2 + \phi_3 + \phi_4 - 4\phi_0)$$

#### 4.1.2 偏微分方程的离散化

基于物理考虑，研究对象为如下形式的线性偏微分方程边值问题。

$$\begin{cases} L\phi = q \\ \left( g_1\phi + g_2 \frac{\partial\phi}{\partial\hat{n}} \right) \Big|_{\partial\mathcal{D}} = g \end{cases} \quad (4.10)$$

其中，偏微分方程成立于区域  $\mathcal{D}$  上， $\partial\mathcal{D}$  为区域  $\mathcal{D}$  的边界， $\hat{n}$  为边界  $\partial\mathcal{D}$  的外法向单位矢量， $L$  为偏微商算子，其形式为

$$L = \nabla(p\nabla) + f \quad (\text{Sturm} - \text{Liouville operator}) \quad (4.11)$$

有限差分法求解边值问题的步骤：

- (1) 通过坐标变换，尽可能使得区域  $\mathcal{D}$  为矩形或正方形；
- (2) 选取笛卡尔标架的等值面进行网格划分，尽可能为等间距正方形网格；
- (3) 依据微商的差商格式，将线性偏微分方程差分为代数线性方程组；
- (4) 将边界条件离散化为代数线性方程组；
- (5) 求解代数线性方程组。

#### 1. 第一类边界条件

$$\phi \Big|_{\partial\mathcal{D}} = g \quad (\text{Dirichlet boundary}) \quad (4.12)$$

图 4.1 所示为等间距  $h$  的二维网格，点 0 到 1, 2 的距离分别  $h_1, h_2$ 。

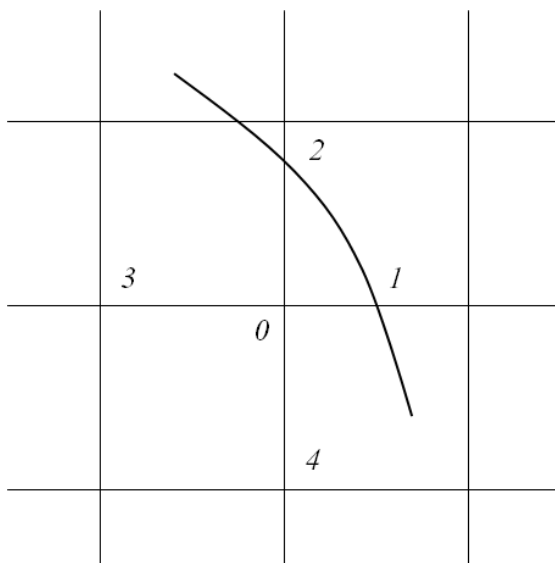


图 4.1 第一类边界条件。

• 直接转移法:

$$\phi_0 = \phi_1 \quad (h_1 < h_2) \quad \text{or} \quad \phi_0 = \phi_2 \quad (h_2 < h_1) \quad (4.13)$$

• 线性插值法:

$$\phi_0 = \frac{h\phi_1 + h_1\phi_3}{h + h_1} \quad (h_1 < h_2) \quad \text{or} \quad \phi_0 = \frac{h\phi_2 + h_2\phi_4}{h + h_2} \quad (h_2 < h_1) \quad (4.14)$$

• 差商法:

利用节点 0, 3, 4 和边界点 1, 2 给出 0 点处的微商算子的差商格式。例如

$$(\nabla^2 \phi)_0 = \frac{2}{h^2} \left[ \frac{1}{\alpha(1+\alpha)} \phi_1 + \frac{1}{\beta(1+\beta)} \phi_2 + \frac{1}{1+\alpha} \phi_3 + \frac{1}{1+\beta} \phi_4 - \left( \frac{1}{\alpha} + \frac{1}{\beta} \right) \phi_0 \right] \quad (4.15)$$

其中,  $\alpha = h_1/h$ ,  $\beta = h_2/h$ 。

## 2. 第二、三类边界条件

$$\left( \frac{\partial \phi}{\partial \hat{n}} + \alpha \phi \right) \Big|_{\partial \mathcal{D}} = g \quad (\text{Neumann/mixed boundary}) \quad (4.16)$$

图 4.2 所示为等间距  $h$  的二维网格， $OQ$  垂直于边界交  $VR$  于  $P$  点， $OP$ ,  $PR$ ,

$PV$  的长度分别为  $ah$ ,  $bh$ ,  $ch$ 。

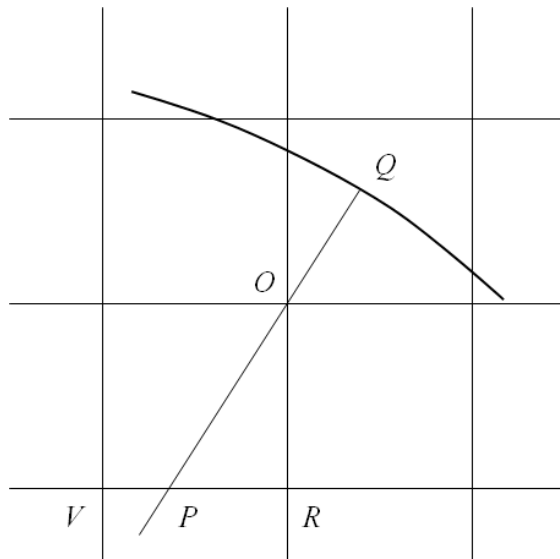


图 4.2 第二、三类边界条件。

将边界条件应用于  $Q$  点，有

$$\left(\frac{\partial \phi}{\partial \hat{n}}\right)_Q + \alpha(Q)\phi_Q = g(Q) \quad (4.17)$$

虑及

$$\left. \begin{aligned} \left(\frac{\partial \phi}{\partial \hat{n}}\right)_Q &= \left(\frac{\partial \phi}{\partial \hat{n}}\right)_O + \mathcal{O}(h) \\ \left(\frac{\partial \phi}{\partial \hat{n}}\right)_O &= \frac{\phi_O - \phi_P}{ah} + \mathcal{O}(h) \\ \phi_P &= b\phi_V + c\phi_R + \mathcal{O}(h^2) \end{aligned} \right\} \Rightarrow \left(\frac{\partial \phi}{\partial \hat{n}}\right)_Q \approx \frac{\phi_O - b\phi_V - c\phi_R}{ah} \quad (4.18)$$

$$\phi_Q = \phi_O + \mathcal{O}(h) \quad (4.19)$$

得到离散化的第二、三类边界条件为

$$\frac{1}{ah}(\phi_O - b\phi_V - c\phi_R) + \alpha(Q)\phi_O = g(Q) \quad (4.20)$$

说明：至此，偏微分方程及边界条件经离散化处理后成为代数线性方程组。

## 4.2 有限元素方法

介绍有限元素法的基本物理思想，并以二维场为例给出作用量泛函离散化为作用量函数的具体方法。

### 4.2.1 基本物理思想

标准的 Sturm-Liouville 边值问题为

$$\begin{cases} -\nabla(p\nabla\phi) + f\phi = \rho & (\text{Sturm - Liouville equation}) \\ \frac{\partial\phi}{\partial\hat{n}}\Big|_{\partial V-S} = 0 & (\text{Dirichlet boundary}) \\ \left(\frac{\partial\phi}{\partial\hat{n}} + \alpha\phi\right)\Big|_S = g & (\text{mixed boundary}) \end{cases} \quad (4.21)$$

其作用量为

$$A[\phi] = \int_V (p|\nabla\phi|^2 + f\phi^2 - 2\rho\phi) dV + \int_S p(\alpha\phi^2 - 2g\phi) dS \quad (4.22)$$

• 证明：

$$\begin{aligned} \delta A[\phi] &= 2 \int_V (p\nabla\phi \cdot \nabla\delta\phi + f\phi\delta\phi - \rho\delta\phi) dV + 2 \int_S p(\alpha\phi\delta\phi - g\delta\phi) dS \\ &= 2 \int_V \delta\phi [-\nabla(p\nabla\phi) + f\phi - \rho] dV \\ &\quad + 2 \int_{\partial V} p\delta\phi \nabla\phi \cdot \hat{n} dS + 2 \int_S p\delta\phi(\alpha\phi - g) dS \\ &= 2 \int_V \delta\phi [-\nabla(p\nabla\phi) + f\phi - \rho] dV \\ &\quad + 2 \int_{\partial V-S} p\delta\phi \frac{\partial\phi}{\partial\hat{n}} dS + 2 \int_S p\delta\phi \left[\frac{\partial\phi}{\partial\hat{n}} + \alpha\phi - g\right] dS \end{aligned} \quad (4.23)$$

由  $\delta A[\phi] = 0$  即可得到 Sturm-Liouville 方程及边界条件。

• 有限元素法特点：

Sturm-Liouville 边值问题可以等价地化为泛函极值（变分）问题。可以看到，对泛函求极值会自动保证满足边界条件。因此，有限元素法对边界条件与偏微分方程同等地处理。

## 4.2.2 二维场的有限元素法

离散化过程首先要对区域进行网格划分。 $\dim = 2$  时, 采用的是三角形网格划分, 并且满足

- (1) 相邻两个三角形必定共两点一边;
- (2) 边界三角形有且仅有两点在边界上

说明: 这样的三角形网格划分可以保证插值线性函数的连续性。

## 1. 元素内线性插值函数

$$\phi(x, y) = g_1 + g_2x + g_3y, \quad (x, y) \in \text{triangle} \quad (4.24)$$

where

$$\begin{cases} g_1 = \frac{1}{2\Delta} (a_i\phi_i + a_j\phi_j + a_m\phi_m) \\ g_2 = \frac{1}{2\Delta} (b_i\phi_i + b_j\phi_j + b_m\phi_m) \\ g_3 = \frac{1}{2\Delta} (c_i\phi_i + c_j\phi_j + c_m\phi_m) \end{cases}$$

and

$$\Delta = \frac{1}{2} \begin{vmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_m & y_m \end{vmatrix}, \quad \begin{cases} a_i = x_jy_m - x_my_j \\ b_i = y_j - y_m \\ c_i = x_m - x_j \end{cases}$$

其中, 三角形元素的三个顶点为  $(x_i, y_i)$ ,  $(x_j, y_j)$ ,  $(x_m, y_m)$ 。引入三角形型函数

$$N_l(x, y) = \frac{1}{2\Delta} (a_l + b_lx + c_ly), \quad (l = i, j, m) \quad (4.25)$$

则三角形元素内线性插值函数可以表达为

$$\phi(x, y) = \sum_{l=i,j,m} \phi_l N_l(x, y) \quad (4.26)$$

2. 作用量泛函  $\rightarrow$  作用量函数

$$A[\phi] = \int_{\mathcal{D}} \left\{ \frac{\varepsilon}{2} \left[ \left( \frac{\partial \phi}{\partial x} \right)^2 + \left( \frac{\partial \phi}{\partial y} \right)^2 \right] - \rho \phi \right\} dx dy \quad (4.27)$$

离散化后，有

$$A(\phi) = A_1(\phi) - A_2(\phi) = \sum_e \left[ A_{1e}(\phi) - A_{2e}(\phi) \right] \quad (4.28)$$

定义与元素  $e$  关联的向量及矩阵

$$\Phi^{(e)} = \begin{pmatrix} \phi_i \\ \phi_j \\ \phi_m \end{pmatrix}, \quad N^{(e)} = \begin{pmatrix} N_i \\ N_j \\ N_m \end{pmatrix} \quad (4.29)$$

•  $A_{1e}(\phi)$  推导:

$$\left. \begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{1}{2\Delta} \sum_{l=i,j,m} b_l \phi_l \\ \frac{\partial \phi}{\partial y} &= \frac{1}{2\Delta} \sum_{l=i,j,m} c_l \phi_l \end{aligned} \right\} \Rightarrow \begin{pmatrix} \frac{\partial \phi}{\partial x} \\ \frac{\partial \phi}{\partial y} \end{pmatrix} = B^{(e)} \Phi^{(e)}$$

where

$$\begin{aligned} B^{(e)} &= \frac{1}{2\Delta} \begin{pmatrix} b_i & b_j & b_m \\ c_i & c_j & c_m \end{pmatrix} \\ \Downarrow \\ \therefore A_{1e}(\phi) &= \frac{1}{2} \Phi^{(e)T} \left[ \int_e \varepsilon B^{(e)T} B^{(e)} dx dy \right] \Phi^{(e)} \quad (4.30) \\ &\quad \downarrow \\ &\quad \text{defined as } K^{(e)} \end{aligned}$$

•  $A_{2e}(\phi)$  推导:

$$\begin{aligned} A_{2e}(\phi) &= \Phi^{(e)T} \left[ \int_e \rho(x, y) N^{(e)}(x, y) dx dy \right] \quad (4.31) \\ &\quad \downarrow \\ &\quad \text{defined as } P^{(e)} \end{aligned}$$

• 作用量函数:

定义网格节点列向量  $\Phi$ ，则元素  $e$  的顶点列向量  $\Phi^{(e)}$  与之关系为  $\Phi^{(e)} = R^{(e)} \Phi$ 。

定义

$$K = \sum_e R^{(e)T} K^{(e)} R^{(e)}, \quad P = \sum_e R^{(e)T} P^{(e)} \quad (4.32)$$

则作用量函数可以写成如下的二次型

$$A(\phi) = \frac{1}{2} \Phi^T K \Phi - \Phi^T P \quad (4.33)$$

### 3. 极值方程组

$$\frac{\partial A}{\partial \Phi^T} = 0 \quad \longrightarrow \quad K \Phi = P \quad (4.34)$$

