研究生规划教材

固体物理学基础教程

沈以赴 编著

化学工业出版社
教材出版中心

研 究 生 规 划 教 材

固体物理学基础教程

沈以赴 编著

化学工业出版社
教 材 出 版 中 心
·北京·

（京）新登字 039 号

图书在版编目（CIP）数据

固体物理学基础教程/沈以赴编著．—北京：化学工业出版社，2005.3
研究生规划教材
ISBN 7-5025-6717-8

Ⅰ．固…　Ⅱ．沈…　Ⅲ．固体物理学-教材　Ⅳ．048

中国版本图书馆 CIP 数据核字（2005）第 015326 号

研究生规划教材

固体物理学基础教程

沈以赴　编著

责任编辑：杨　菁

文字编辑：彭喜英

责任校对：蒋　宇

封面设计：于　兵

*

化学工业出版社
教 材 出 版 中 心　出版发行

（北京市朝阳区惠新里 3 号　邮政编码 100029）

发行电话：（010）64982530

http://www.cip.com.cn

*

新华书店北京发行所经销

北京永鑫印刷有限责任公司印刷

三河市海波装订厂装订

开本 787mm×1092mm　1/16　印张 26¾　字数 672 千字

2005 年 5 月第 1 版　　2005 年 5 月北京第 1 次印刷

ISBN 7-5025-6717-8/G·1730

定　价：52.00 元

内容提要

本书主要讲述了分析力学、量子力学及统计物理学等固体物理学的基础知识；在此基础上再讲述固体物理学的核心内容，即固体中原子的结构和运动模型，电子的运动模型等；继而进一步讨论了晶格振动等对电子运动的影响。全书共10章，包括分析力学基础、量子力学基本理论、统计物理概论、固体的结构、固体的结合、晶格振动Ⅰ——声子、晶格振动Ⅱ——热学性质、固体电子论基础、固体能带理论、固体电导理论。

本书可作为非物理类工科研究生，特别是材料学、材料加工工程研究生教材，也可作为材料物理类本科生用教材和相关科技人员参考书。

前　言

本书是为非物理类工科研究生，特别是材料学、材料加工工程学科研究生编写的教材，也可作为科技人员和其他相关专业研究生、材料物理类专业本科生的参考书。本书着重阐述固体物理学的物理基础、基本概念和基本理论。

固体物理学是物理学的一个重要分支，主要研究固体物质的微观结构、运动状态、物理性质及其相互关系。近几十年来，固体物理学的发展非常迅速，并已形成了以晶体学、晶格振动动力学、金属物理学、半导体物理学、磁学、电介质物理学、压电物理学、铁电物理学、低温物理学、高压物理学、发光学以及近期发展起来的表面物理学、超导物理学、非晶态物理学、液晶物理学、高分子物理学及低维固体物理学等分支学科，而且，新的分支尚在不断迸发。固体物理学的概念、方法和实验技术还在向相邻的学科渗透，有力地促进了材料科学和器件物理、化学物理学、生物物理学和地球物理学等广义学科的发展。

过去的大多数固体物理学教材都是为物理系学生编写的，一般要求学生有比较系统的统计物理学和量子力学的基础，这就使得普通工科研究生、大学生学习和掌握固体物理学的概念和理论感到比较困难；同时这些固体物理学教材中又或多或少地选编了固体的一些专题，如半导体、磁学及超导电性等。编者在多年的固体物理学的学习和教学过程中，深感缺乏一本适合于工科学生，特别是材料科学与工程专业学生学习固体物理基础的教材。

本书在内容处理上作了新的尝试：即首先讲授固体物理学的物理基础，主要包括分析力学、量子力学及统计物理学的基础知识；在此基础上再讲授固体物理学的核心内容，即固体中原子（离子）的结构和运动模型（固体的结构、固体的结合、晶格振动及热学性质等），固体中电子的运动模型（固体电子论、能带理论等），继而进一步讨论晶格振动等对电子运动的影响（电导理论）。而对于材料科学与工程专业学生已经掌握的晶体缺陷的性质、晶体中的扩散、相平衡与相图等则不再讲授；对于固体物理学的分支或专题，在本教程中也不介绍，学生在学完本课程后，可根据学科研究方向再选修磁性材料、非晶态材料、薄膜材料、介电材料、半导体材料、超导材料或非线性光学材料等相关课程。

在本书的编写过程中，参考了很多参考书籍和资料，这里谨向这些书籍和资料的作者们表示感谢！还要感谢南京航空航天大学孙久厚教授、肖军教授的热情支持和帮助。由于水平有限和时间仓促，书中的错误和不妥之处在所难免，恳请专家和学者批评指正。

沈以赴

2004 年 8 月 8 日

目　录

第 1 章　分析力学基础

所谓经典力学即是牛顿力学，这是一门最古老的科学之一，它的创建主要归功于伽利略（Galilei）、惠更斯（C. Huygens）及牛顿（I. Newton）等人。其中最著名的是牛顿三定律。下面介绍在牛顿所创建的原理的基础之上，经过拉格朗日（J. L. Lagrange）、哈密顿（W. R. Hamilton）等人建立的分析力学，即采用广义坐标体系从能量角度来研究经典力学。这样处理的好处，除了在许多情况下使力学问题易于解决外，尤其重要的是，它易于把经典力学与统计物理和量子力学等其他物理分支学科联系起来。分析力学这个名词来源于拉格朗日的著作（Mechanique Analytique）的标题。

1.1　广义坐标与拉格朗日方程

1.1.1　广义坐标

一个质点在空间的位置由它的位矢 $\boldsymbol{r}$ 所决定。位矢 $\boldsymbol{r}$ 的分量与质点的笛卡儿坐标 x、y、z 相合，如图 1.1-1 所示，质点位置除可用笛卡儿坐标 x、y、z 来表示外，也可以用其他坐标，如柱坐标 ρ、φ、z 或球坐标 r、θ、φ 等来表示。不管用哪一种坐标系，只要质点是自由的，即 x、y、z（或 ρ、φ、z 或 r、θ、φ）彼此独立，则可称该质点有 3 个自由度。例如限制某个质点在一定的平面上，则这个质点的 x、y、z 之间存在着一定的关系，这个关系就是这个质点所在的平面方程 $Ax+By+Cz+D=0$。该平面方程可称为“约束方程”。显然确定质点位置的 x、y、z 已不可能彼此无关，因为独立地选定 x、y，则 z 就确定了，即自由度为 2。如果设想某质点限制在一条直线上且这条直线为两个平面的交线，即此直线由 $A_1x+B_1y+C_1z+D_1=0$ 及 $A_2x+B_2y+C_2z+D_2=0$ 联立确定，则约束方程有两个，此时，只要独立地选定 z、x 和 y 就确定了，所以质点的自由度为 1。

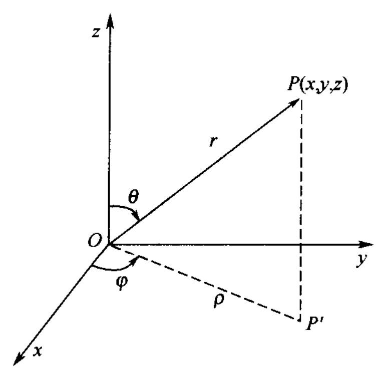

图 1.1-1　笛卡儿坐标

用“位形”这个词来简称一群质点在空间的位形。由上述可知，为了确定由 n 个质点组成的体系在空间的位形，需要给定 $3n$ 个坐标。显然 n 个质点组成的体系，如果没有任何约束，则体系的自由度为 $3n$。如果各质点的位置之间存在着某些约束，体系的位形的变化就会受到限制；也就是说，描写位形的 $3n$ 个坐标不能独立地变动。一般说来，如果由 n 个质点组成的体系存在 k 个约束，则这个体系的自由度为 $s=3n-k$，那么用 s 个独立变量就可以单值地确定该体系的位形。

广义坐标就是能够确定某体系的位形的任意一组量。若体系的自由度为 s，采用 s 个量即 q_1，$q_2\cdots q_s$，就能完全确定体系的位形。这里 q_1，$q_2\cdots q_s$ 就是该体系的广义坐标。有了广义坐标的概念，就可以把由 s 个广义坐标 q_1，$q_2\cdots q_s$ 确定的力学系用 s 维空间中的“一个

点”来表示，这个用广义坐标定义的 s 维空间就是所谓的位形空间。这样体系随时间的变化过程就可以用位形空间中的一条曲线（位轨线）来表示，位轨线上每一个点都表示系统在某一时刻的位形。位形空间是用分析方法研究力学问题所必需的一种抽象概念。引入广义坐标后，研究体系的位形变化时，就可以根据体系的特点，适当地选择广义坐标，能使变量减到最少，这就避免了由于约束所引起的许多困难。同一个体系可以选取不同的广义坐标来描写其位形。需要指出的是，广义坐标可以是笛卡儿坐标，也可以是其他的量。

广义坐标对时间的微商称为广义速度，记为 $\dot{q}_i=\dfrac{\mathrm{d}q_i}{\mathrm{d}t}$。必须指出，广义速度不一定具有长度除以时间的量纲。例如，当 q_i 代表角度 θ，$\dot{\theta}$ 代表角速度时，$\dot{q}_i$就具有角速度的量纲。

如果给定某时刻的广义坐标数值，则力学体系在该时刻的位形也就确定了，但并不能确定体系在下一时刻的位形。如果同时给定了广义坐标和广义速度，那么，根据体系的运动方程，就可以决定体系在任一时刻的广义坐标和广义速度，也就是说，体系的运动状态就完全确定了。

1.1.2 拉格朗日方程

下面研究力学体系在广义坐标中的运动方程——拉格朗日方程。

1.1.2.1 一个质点体系的拉格朗日方程

根据牛顿第二定律，质点所受的外力 $\boldsymbol{F}$ 等于质点的惯性质量 m 乘以质点的加速度，即

$$\boldsymbol{F}=m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} \tag{1.1-1}$$

式中，$\boldsymbol{r}$ 是质点的矢径。在笛卡儿坐标中，式（1.1-1）可以写成分量的形式，如

$$F_x=m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=m\ddot{x},F_y=m\frac{\mathrm{d}^2y}{\mathrm{d}t^2}=m\ddot{y},F_z=m\frac{\mathrm{d}^2z}{\mathrm{d}t^2}=m\ddot{z} \tag{1.1-2}$$

式中，F_x、F_y、F_z 分别代表力在 x、y、z 三个方向的分量。

现将质点位置用广义坐标 q_1、q_2 和 q_3 表示，并且坐标的转换满足如下关系

$$x=x(q_1,q_2,q_3),\quad y=y(q_1,q_2,q_3),\quad z=z(q_1,q_2,q_3) \tag{1.1-3}$$

为了将式（1.1-2）转换到广义坐标中去，可以用$\dfrac{\partial x}{\partial q_1}$、$\dfrac{\partial y}{\partial q_1}$、$\dfrac{\partial z}{\partial q_1}$分别乘以式（1.1-2）的三个方程，并将它们加起来，得到

$$m\left(\ddot{x}\frac{\partial x}{\partial q_1}+\ddot{y}\frac{\partial y}{\partial q_1}+\ddot{z}\frac{\partial z}{\partial q_1}\right)=F_x\frac{\partial x}{\partial q_1}+F_y\frac{\partial y}{\partial q_1}+F_z\frac{\partial z}{\partial q_1} \tag{1.1-4}$$

定义广义力

$$Q_1=F_x\frac{\partial x}{\partial q_1}+F_y\frac{\partial y}{\partial q_1}+F_z\frac{\partial z}{\partial q_1} \tag{1.1-5}$$

考虑到

$$\ddot{x}\frac{\partial x}{\partial q_1}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}\frac{\partial x}{\partial q_1}\right)-\dot{x}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x}{\partial q_1}\right),\ddot{y}\frac{\partial y}{\partial q_1}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{y}\frac{\partial y}{\partial q_1}\right)-\dot{y}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial y}{\partial q_1}\right),$$

$$\ddot{z}\frac{\partial z}{\partial q_1}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{z}\frac{\partial z}{\partial q_1}\right)-\dot{z}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial z}{\partial q_1}\right)$$

则可将式（1.1-4）改写为

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m\left(\dot{x}\frac{\partial x}{\partial q_1}+\dot{y}\frac{\partial y}{\partial q_1}+\dot{z}\frac{\partial z}{\partial q_1}\right)\right]-m\left[\dot{x}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x}{\partial q_1}\right)+\dot{y}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial y}{\partial q_1}\right)+\dot{z}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial z}{\partial q_1}\right)\right]=Q_1 \tag{1.1-6}$$

因为 x、y、z 为时间的隐函数，所以将式（1.1-3）对时间求全微商得到

$$\dot{x}=\frac{\partial x}{\partial q_1}\dot{q}_1+\frac{\partial x}{\partial q_2}\dot{q}_2+\frac{\partial x}{\partial q_3}\dot{q}_3,\dot{y}=\frac{\partial y}{\partial q_1}\dot{q}_1+\frac{\partial y}{\partial q_2}\dot{q}_2+\frac{\partial y}{\partial q_3}\dot{q}_3,$$

$$\dot{z}=\frac{\partial z}{\partial q_1}\dot{q}_1+\frac{\partial z}{\partial q_2}\dot{q}_2+\frac{\partial z}{\partial q_3}\dot{q}_3 \tag{1.1-7}$$

再分别对上式求 $\dot{x}$ 对 $\dot{q}_1$ 的偏微商得到

$$\frac{\partial \dot{x}}{\partial \dot{q}_1}=\frac{\partial x}{\partial q_1},\frac{\partial \dot{y}}{\partial \dot{q}_1}=\frac{\partial y}{\partial q_1},\frac{\partial \dot{z}}{\partial \dot{q}_1}=\frac{\partial z}{\partial q_1} \tag{1.1-8}$$

由于 x 是 q_1、q_2、q_3 的函数，故 $\frac{\partial x}{\partial q_1}$、$\frac{\partial y}{\partial q_1}$、$\frac{\partial z}{\partial q_1}$ 也分别是各个 q 的函数。于是有

$$\left.\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x}{\partial q_1}\right)&=\sum_{k=1}^{3}\frac{\partial}{\partial q_k}\left(\frac{\partial x}{\partial q_1}\right)\dot{q}_k=\frac{\partial}{\partial q_1}\sum_{k=1}^{3}\frac{\partial x}{\partial q_k}\dot{q}_k=\frac{\partial \dot{x}}{\partial q_1}\\\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial y}{\partial q_1}\right)&=\sum_{k=1}^{3}\frac{\partial}{\partial q_k}\left(\frac{\partial y}{\partial q_1}\right)\dot{q}_k=\frac{\partial}{\partial q_1}\sum_{k=1}^{3}\frac{\partial y}{\partial q_k}\dot{q}_k=\frac{\partial \dot{y}}{\partial q_1}\\\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial z}{\partial q_1}\right)&=\sum_{k=1}^{3}\frac{\partial}{\partial q_k}\left(\frac{\partial z}{\partial q_1}\right)\dot{q}_k=\frac{\partial}{\partial q_1}\sum_{k=1}^{3}\frac{\partial z}{\partial q_k}\dot{q}_k=\frac{\partial \dot{z}}{\partial q_1}\end{aligned}\right\}$$

$$\text{即}\left.\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x}{\partial q_1}\right)&=\frac{\partial \dot{x}}{\partial q_1}\\\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial y}{\partial q_1}\right)&=\frac{\partial \dot{y}}{\partial q_1}\\\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial z}{\partial q_1}\right)&=\frac{\partial \dot{z}}{\partial q_1}\end{aligned}\right\} \tag{1.1-9}$$

将式（1.1-8）和式（1.1-9）代入式（1.1-6）中，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m\left(\dot{x}\frac{\partial x}{\partial q_1}+\dot{y}\frac{\partial y}{\partial q_1}+\dot{z}\frac{\partial z}{\partial q_1}\right)\right]-m\left[\dot{x}\frac{\partial \dot{x}}{\partial q_1}+\dot{y}\frac{\partial \dot{y}}{\partial q_1}+\dot{z}\frac{\partial \dot{z}}{\partial q_1}\right]=Q_1 \tag{1.1-10}$$

如果已知质点的动能为

$$T=\frac{1}{2}mv^2=\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) \tag{1.1-11}$$

则有

$$\frac{\partial T}{\partial \dot{q}_1}=m\left(\dot{x}\frac{\partial \dot{x}}{\partial \dot{q}_1}+\dot{y}\frac{\partial \dot{y}}{\partial \dot{q}_1}+\dot{z}\frac{\partial \dot{z}}{\partial \dot{q}_1}\right) \tag{1.1-12}$$

$$\frac{\partial T}{\partial q_1}=m\left(\dot{x}\frac{\partial \dot{x}}{\partial q_1}+\dot{y}\frac{\partial \dot{y}}{\partial q_1}+\dot{z}\frac{\partial \dot{z}}{\partial q_1}\right) \tag{1.1-13}$$

将式（1.1-12）及式（1.1-13）代入式（1.1-10）得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_1}\right)-\frac{\partial T}{\partial q_1}=Q_1 \tag{1.1-14}$$

假设力学体系是保守的，也就是说存在一个势能 U，且 U 只是位置 x、y、z 的函数，那么质点所受到的力 $\boldsymbol{F}$ 是 U 的梯度的负值，即

$$\boldsymbol{F}=-\nabla U$$

$$F_x=-\frac{\partial U}{\partial x},F_y=-\frac{\partial U}{\partial y},F_z=-\frac{\partial U}{\partial z} \tag{1.1-15}$$

这样，广义力 Q_1 的表示式可写为

$$Q_1=-\left(\frac{\partial U}{\partial x}\times\frac{\partial x}{\partial q_1}+\frac{\partial U}{\partial y}\times\frac{\partial y}{\partial q_1}+\frac{\partial U}{\partial z}\times\frac{\partial z}{\partial q_1}\right)=-\frac{\partial U}{\partial q_1} \tag{1.1-16}$$

从上式可以看出，广义力 Q_1 的物理意义是变更所引起的势能的减少。将上式代入式（1.1-14）得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_1}\right)-\frac{\partial T}{\partial q_1}=-\frac{\partial U}{\partial q_1} \tag{1.1-17}$$

现在引进一个新函数 L，这个函数称为拉格朗日函数，它等于动能和势能的差，即

$$L=T-U \tag{1.1-18}$$

因为动能 T 只是速度$\dot{q}_1$、$\dot{q}_2$、$\dot{q}_3$的函数，而势能 U 只与位置q_1、q_2、q_3有关，而与$\dot{q}_1$、$\dot{q}_2$、$\dot{q}_3$无关，所以可以得到

$$\frac{\partial L}{\partial \dot{q}_1}=\frac{\partial T}{\partial \dot{q}_1},\frac{\partial L}{\partial \dot{q}_2}=\frac{\partial T}{\partial \dot{q}_2},\frac{\partial L}{\partial \dot{q}_3}=\frac{\partial T}{\partial \dot{q}_3} \tag{1.1-19}$$

代入式（1.1-17）得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_1}\right)-\frac{\partial L}{\partial q_1}=0,\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_2}\right)-\frac{\partial L}{\partial q_2}=0,\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_3}\right)-\frac{\partial L}{\partial q_3}=0 \tag{1.1-20}$$

上式即为用广义坐标表示的一个质点的运动方程，也称为拉格朗日方程。

1.1.2.2 多质点体系的拉格朗日方程

设体系有 n 个质点，自由度为 s，所以取 s 个广义坐标 q_1，$q_2\cdots q_s$。每个质点都可以按牛顿定律写出其运动方程，如第 i 个质点的运动方程为

$$F_{ix}=m_i\ddot{x}_i,F_{iy}=m_i\ddot{y}_i,F_{iz}=m_i\ddot{z}_i \tag{1.1-21}$$

坐标的转换关系为

$$x_i=x_i(q_1,q_2\cdots q_s),y_i=y_i(q_1,q_2\cdots q_s),z_i=z_i(q_1,q_2\cdots q_s) \tag{1.1-22}$$

对每个质点进行像式（1.1-4）那样的求和，得到

$$Q_1=\sum_{i=1}^{n}\left(F_{ix}\frac{\partial x_i}{\partial q_1}+F_{iy}\frac{\partial y_i}{\partial q_1}+F_{iz}\frac{\partial z_i}{\partial q_1}\right)=\sum_{i=1}^{n}m_i\left(\ddot{x}_i\frac{\partial x_i}{\partial q_1}+\ddot{y}_i\frac{\partial y_i}{\partial q_1}+\ddot{z}_i\frac{\partial z_i}{\partial q_1}\right) \tag{1.1-23}$$

体系的动能为

$$T=\sum_{i=1}^{n}\frac{1}{2}m_iv_i^2=\sum_{i=1}^{n}\frac{1}{2}m_i\left(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2\right) \tag{1.1-24}$$

体系的势能是所有质点的位置的函数，即

$$U=U(x_1,y_1,z_1\cdots x_i,y_i,z_i\cdots x_n,y_n,z_n) \tag{1.1-25}$$

相应的广义力为

$$Q_1=-\sum_{i=1}^{n}\left(\frac{\partial U}{\partial x_i}\times\frac{\partial x_i}{\partial q_1}+\frac{\partial U}{\partial y_i}\times\frac{\partial y_i}{\partial q_1}+\frac{\partial U}{\partial z_i}\times\frac{\partial z_i}{\partial q_1}\right)=-\frac{\partial U}{\partial q_1} \tag{1.1-26}$$

类似地令 $L=T-U$，则可以获得多质点体系的拉格朗日方程为

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_1}\right)-\frac{\partial L}{\partial q_1}=0,\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_2}\right)-\frac{\partial L}{\partial q_2}=0\cdots\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_s}\right)-\frac{\partial L}{\partial q_s}=0 \tag{1.1-27}$$

在有约束的情形下，拉格朗日方程的优越性显得特别清楚。对于包含 n 个质点及 k 个约束的体系，为了确定体系的运动状态，要联解 $3n$ 个牛顿方程及 k 个约束方程；但若选择适

当的广义坐标，便只需解 $3n-k$ 个拉格朗日方程。

拉格朗日方程是直接就广义坐标 q_α 写出的，这样就排除了笛卡儿坐标的特别地位，得以从普通的物理空间（在其中运动方程可能非常复杂）变换到位形空间，只要适当选取位形空间，就可使有些问题得到最大限度的简化。

用拉格朗日方程解题的一般步骤为：首先分析约束，确定自由度数；其次按自由度数选择适当的广义坐标；计算广义力或动能势能，求出拉格朗日函数；最后建立拉格朗日方程组（1.1-27）。

由上可以看出，利用拉格朗日方程来解决力学问题，首要的是写出用广义坐标表示的拉格朗日函数。其核心问题是用广义速度表示动能。在笛卡儿坐标中，体系的动能可用式（1.1-24）表示。为了将动能用广义速度表示，可将式（1.1-22）对时间取微商得到

$$\dot{x}_i=\sum_{k=1}^{s}\frac{\partial x_i}{\partial q_k}\dot{q}_k,\dot{y}_i=\sum_{k=1}^{s}\frac{\partial y_i}{\partial q_k}\dot{q}_k,\dot{z}_i=\sum_{k=1}^{s}\frac{\partial z_i}{\partial q_k}\dot{q}_k, \tag{1.1-28}$$

将上式代入式（1.1-24）中，整理后得到

$$T=\sum_{k=1}^{s}\sum_{l=1}^{s}a_{kl}\,\dot{q}_k\,\dot{q}_l \tag{1.1-29}$$

式中，a_{kl} 是 $\dot{q}$ 的各二次项的系数，它是各个广义坐标的函数。应该指出的是，体系的动能是广义速度的齐次二次函数。

实际上，在许多情况下不必采用上述步骤求动能 T。因为质点的速率是其轨迹的弧长对时间的微商，即

$$v^2=\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2=\frac{(\mathrm{d}s)^2}{(\mathrm{d}t)^2} \tag{1.1-30}$$

因此，在相应的坐标系中，若找到弧长 $\mathrm{d}s$ 的平方，将它除以 $(\mathrm{d}t)^2$，就能得到 v^2。

如在柱坐标中

$$(\mathrm{d}s)^2=(\mathrm{d}r)^2+r^2(\mathrm{d}\varphi)^2+(\mathrm{d}z)^2 \tag{1.1-31}$$

$$T=\frac{1}{2}mv^2=\frac{1}{2}m(\dot{r}^2+r^2\dot{\varphi}^2+\dot{z}^2) \tag{1.1-32}$$

在球坐标中

$$(\mathrm{d}s)^2=(\mathrm{d}r)^2+r^2(\mathrm{d}\theta)^2+r^2\sin^2\theta(\mathrm{d}\varphi)^2 \tag{1.1-33}$$

$$T=\frac{1}{2}mv^2=\frac{1}{2}m(\dot{r}^2+r^2\dot{\theta}^2+r^2\sin^2\theta\,\dot{\varphi}^2) \tag{1.1-34}$$

最后要指出的是，拉格朗日函数 L 等于力学系的动能与势能之差，是表征力学系特征的一个重要标量函数，用一个标量函数（代替矢量函数）就表征出力学系的全部特征，这是拉格朗日表述较之牛顿表述的优点之一。

1.2 广义动量与哈密顿正则方程

拉格朗日方程是处理力学问题的一个很有力的工具。但是，用能量及广义坐标表示的运动微分方程式还存在另一种形式，这就是哈密顿正则方程，它和量子力学有直接的联系。

1.2.1 广义动量

设体系的广义坐标为 q_1，$q_2\cdots q_s$，对应于每一个 q_k，可以定义一个广义动量

$$p_k=\frac{\partial L}{\partial \dot{q}_k} \tag{1.2-1}$$

式中，L 为拉格朗日函数，$\dot{q}_k$是广义速度。

那么，广义动量与一般所说的动量有什么联系呢？在特殊情况下，广义动量与一般所说的含义相同。例如，对于一个只做一维运动的自由质点，广义坐标只有一个，记为 x，其拉格朗日函数$L=\frac{1}{2}m\dot{x}^2$，其广义动量 $p_x=\frac{\partial L}{\partial \dot{x}}=m\dot{x}$，即质量乘速度。在一般情况下就不是这样，例如，受约束只能做圆周运动的质点，其自由度为 1，取方向角 θ 为广义坐标，$L=\frac{1}{2}mr^2\dot{\theta}^2$，广义动量 $p_\theta=\frac{\partial L}{\partial \dot{\theta}}=mr^2\dot{\theta}$，可见广义动量是动量矩。

引入广义动量的概念之后，就可以把一组 s 个二阶常微分方程化为一组 $2s$ 个一阶微分方程。

1.2.2 哈密顿正则方程

1.2.2.1 哈密顿函数与哈密顿正则方程

拉格朗日函数是广义坐标和广义速度的函数，它的全微分可表示为

$$\mathrm{d}L=\sum_{k=1}^{s}\frac{\partial L}{\partial q_k}\mathrm{d}q_k+\sum_{k=1}^{s}\frac{\partial L}{\partial \dot{q}_k}\mathrm{d}\dot{q}_k \tag{1.2-2}$$

根据广义动量的定义式（1.2-1）及拉格朗日方程式（1.1-27），上式可化为

$$\mathrm{d}L=\sum_{k=1}^{s}\dot{p}_k\mathrm{d}q_k+\sum_{k=1}^{s}p_k\mathrm{d}\dot{q}_k \tag{1.2-3}$$

式（1.2-3）右边第二项可写为

$$\sum_{k=1}^{s}p_k\mathrm{d}\dot{q}_k=\mathrm{d}\left(\sum_{k=1}^{s}p_k\dot{q}_k\right)-\sum_{k=1}^{s}\dot{q}_k\mathrm{d}p_k \tag{1.2-4}$$

将式（1.2-4）代入式（1.2-3）中，得到

$$\mathrm{d}\left(\sum_{k=1}^{s}p_k\dot{q}_k-L\right)=-\sum_{k=1}^{s}p_k\mathrm{d}\dot{q}_k+\sum_{k=1}^{s}\dot{q}_k\mathrm{d}p_k \tag{1.2-5}$$

定义

$$H=\sum_{k=1}^{s}p_k\dot{q}_k-L \tag{1.2-6}$$

于是式（1.2-5）变为

$$\mathrm{d}H=-\sum_{k=1}^{s}\dot{p}_k\mathrm{d}q_k+\sum_{k=1}^{s}\dot{q}_k\mathrm{d}p_k \tag{1.2-7}$$

由上式可以看出，H 只是各个 q_k 及 p_k 的函数，称为哈密顿函数，即

$$H=H(q,p) \tag{1.2-8}$$

式中，q 及 p 代表各个 q_k 及 p_k。

对 $H(q,p)$取微分可得

$$\mathrm{d}H=\sum_{k=1}^{s}\frac{\partial H}{\partial q_k}\mathrm{d}q_k+\sum_{k=1}^{s}\frac{\partial H}{\partial p_k}\mathrm{d}p_k \tag{1.2-9}$$

将式（1.2-9）与式（1.2-7）对比得到

$$\frac{\partial H}{\partial q_k}=-\dot{p}_k,\frac{\partial H}{\partial p_k}=\dot{q}_k(k=1,2\cdots s) \tag{1.2-10}$$

上式称为哈密顿正则方程或哈密顿运动方程，简称哈密顿方程。式（1.2-10）中显示有 s 个自由度的系统可由 s 个广义坐标和 s 个广义动量的一阶微分方程来描述其运动状态。对上述包括 $2s$ 个形式完全相同的一阶微分方程（正则方程）求解，会出现 $2s$ 个积分常数，这些积分常数由起始条件决定，若已知在某时刻的各个广义坐标及广义动量的数值，根据这 $2s$ 个已知值就可决定 $2s$ 个积分常数，于是体系在各个时刻的状态，也就是体系的运动状态就完全确定了。

由此可见，一个自由度为 s 的力学体系可以用包含 $2s$ 个一阶微分方程的哈密顿方程来描写其运动状态，也可以用包含 s 个二阶微分方程的拉格朗日方程来描写。这两种方法是完全等效的，它们的积分解都包含 $2s$ 个积分常数，要由 $2s$ 个起始条件来决定。但是，哈密顿方程把牛顿力学中最突出的特征——决定论表达得最为明显，因为决定论（也可以说是严格的因果关系）正是全部经典物理学的特征，这一点在学过量子力学之后，可能会得到更深的体会。

1.2.2.2 哈密顿函数的物理意义

在作一般的讨论之前，先看一个简单的情形，即体系是只有一个质点的保守系，取笛卡儿坐标为广义坐标。这时有

$$L=T-U=\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)-U(x,y,z) \tag{1.2-11}$$

根据哈密顿函数的定义式

$$H=(p_x\dot{x}+p_y\dot{y}+p_z\dot{z})-L \tag{1.2-12}$$

广义动量 p_x,p_y,p_z 分别为 $m\dot{x},m\dot{y},m\dot{z}$，于是

$$H=(m\dot{x}^2+m\dot{y}^2+m\dot{z}^2)-L=2T-L=T+U \tag{1.2-13}$$

显然在这种情形下，哈密顿函数就是体系的动能与位能之和，即体系的能量。

对于一般情形，即体系是具有 s 个广义坐标 $q_1,q_2\cdots q_s$ 的保守系。已经证明，此时体系的哈密顿函数 H 只是各个 q_k 及 p_k 的函数，H 对时间的变化率为

$$\dot{H}=\sum_{k=1}^{s}\frac{\partial H}{\partial q_k}\dot{q}_k+\sum_{k=1}^{s}\frac{\partial H}{\partial p_k}\dot{p}_k \tag{1.2-14}$$

将哈密顿方程式（1.2-10）代入上式，得到

$$\dot{H}=\sum_{k=1}^{s}\frac{\partial H}{\partial q_k}\times\frac{\partial H}{\partial p_k}-\sum_{k=1}^{s}\frac{\partial H}{\partial p_k}\times\frac{\partial H}{\partial q_k}=0 \tag{1.2-15}$$

这就是说，在保守系中，哈密顿函数不随时间变化。体系的能量不随时间变化，那么哈密顿函数是否就是体系的能量呢？回答是肯定的。下面给出简单证明。

因为保守系的势能只是坐标的函数，即势能 U 与广义速度没有明显的关系，于是

$$p_k=\frac{\partial L}{\partial \dot{q}_k}=\frac{\partial T}{\partial \dot{q}_k} \tag{1.2-16}$$

所以有

$$H=\sum_{k=1}^{s}\frac{\partial T}{\partial \dot{q}_k}\dot{q}_k-L \tag{1.2-17}$$

根据式（1.1-29）可知，T 是广义速度的齐次二次函数。将所有 $\dot{q}$ 都乘以因子 λ，则必有

$$\lambda^2 T=\sum_{k=1}^{s}\sum_{l=1}^{s}a_{kl}(\lambda\dot{q}_k)(\lambda\dot{q}_l) \tag{1.2-18}$$

所以有
$$\lambda^2 T(\dot{q}_1,\dot{q}_2\cdots\dot{q}_s)=T(\lambda\dot{q}_1,\lambda\dot{q}_2\cdots\lambda\dot{q}_s) \tag{1.2-19}$$

将上式两边对 λ 求导数，得到

$$2\lambda T(\dot{q})=\dot{q}_1\frac{\partial T(\lambda\dot{q})}{\partial(\lambda\dot{q}_1)}+\dot{q}_2\frac{\partial T(\lambda\dot{q})}{\partial(\lambda\dot{q}_2)}+\cdots+\dot{q}_s\frac{\partial T(\lambda\dot{q})}{\partial(\lambda\dot{q}_s)} \tag{1.2-20}$$

式中，$T(\dot{q})$表示 $T(\dot{q}_1,\ \dot{q}_2\cdots\dot{q}_s)$，$T\ (\lambda\dot{q})$ 表示 $T\ (\lambda\dot{q}_1,\ \lambda\dot{q}_2\cdots\lambda\dot{q}_s)$。令 $\lambda=1$，便得到

$$\sum_{k=1}^{s}\frac{\partial T}{\partial\dot{q}_k}\dot{q}_k=2T \tag{1.2-21}$$

由此得到
$$H=T+U \tag{1.2-22}$$

上式说明哈密顿函数等于体系的总能量。但是需要强调的是，哈密顿函数是用广义坐标 q 和广义动量 p 来表示的函数，而不是用 q 及$\dot{q}$ 来表示的。

前面讨论的全部是针对保守体系的，也就是说，假设存在一个势能 U，U 只是位置的函数，质点所受到的力是 U 的梯度的负值。下面简要介绍非保守体系的情形。

设存在一个势能 U，质点所受到的力为 $-U$ 的梯度，但 U 不仅是位置的函数，而且还是时间的函数。这时，广义力 Q_l 的表示式（1.1-26）、拉格朗日方程式（1.1-27）和哈密顿正则方程式（1.2-10）仍然成立。拉格朗日函数的定义 $L=T-U$ 仍然不变。哈密顿函数 H 仍然是体系的动能与势能之和 $T+U$。不同的是，这时 L 及 H 都是时间的函数了，即

$$L=L(q,\dot{q},t) \tag{1.2-23}$$

$$H=H(q,p,t) \tag{1.2-24}$$

H 对时间的微商$\dot{H}$不再为零了。

1.3 哈密顿原理

1.3.1 哈密顿原理及其变分形式

1.3.1.1 哈密顿原理

根据前面的讨论可知，一个力学体系的运动可以用一组微分方程来描述，这组微分方程可以是牛顿方程、拉格朗日方程或者是哈密顿正则方程，它们彼此是等价的。将微分方程积分，根据体系在某时刻的状态决定积分常数，便可确定体系在任何时刻的状态。反过来说，如果力学体系的初态和终态已确定，则体系将按照确定的路径由初态运动至终态，由力学体系的状态所决定的各种特性（如体系的动能、势能、拉格朗日函数等）对时间的自始态到终态的积分也是确定的。于是提出了一个问题，这些积分遵循什么规律？哈密顿原理回答了这个问题。

哈密顿原理可表述为：体系由始态到终态其拉格朗日函数沿实际运动路径对时间的积分取极值（一般为极小值）。极值的含义是：与其他任何假想的路径比较起来，沿真实路径的积分值为极小或极大。形象地说，力学体系会自动地选择一条路径，使拉格朗日函数从

t_0（始态）到 t_e（终态）的积分 $\int_{t_0}^{t_e} L\mathrm{d}t$ 取极值。

这里不妨以一维运动为例做一简单说明。如图 1.3-1 所示，图中以 A 表示体系在 t_0 时刻的状态，B 表示在 t_e 时刻的状态。体系实际走的路径为 ACB，而 ADB 及 AEB 表示两条任意的“变更”的路径。设体系的拉格朗日函数为 $L=L(x,\dot{x})$，式中，x 及 $\dot{x}$ 都是 t 的函数。路径 ACB 对应于 $x=x_C(t)$，ADB 对应于 $x=x_D(t)$，AEB 对应于 $x=x_E(t)$。按照哈密顿原理，$\int_{t_0}^{t_e} L[x_C(t),\ \dot{x}_C(t)]\mathrm{d}t$ 比 $\int_{t_0}^{t_e} L[x_D(t),\ \dot{x}_D(t)]\mathrm{d}t$ 及 $\int_{t_0}^{t_e} L[x_E(t),\dot{x}_E(t)]\mathrm{d}t$ 都小或都大。

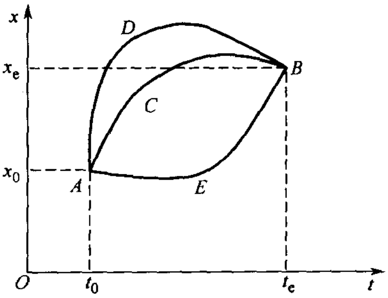

图 1.3-1　哈密顿原理示意图

1.3.1.2　哈密顿原理的变分形式

通常把 $\int_{t_0}^{t_e} L\mathrm{d}t$ 称为哈密顿作用量 J。路径“变更”意味着 $x(t)$ 的函数形式发生变更，J 的值也随之变更。J 是一个依赖于 $x(t)$ 而改变的变量，而 $x(t)$ 可以是 $x_C(t)$、$x_D(t)$ 或是 $x_E(t)$。依赖于函数的变更而变化的变量称为泛函数。在这里遇到的仅是泛函数取极值的问题，这是变分学的问题。

$x(t)$ 的微小变更称为 $x(t)$ 的变分，记为 δx。变分 δx 与微分 $\mathrm{d}x$ 的含义是不同的。在这里 x 是 t 的函数，$\mathrm{d}x$ 代表 t 变化所引起的变化，而 δx 代表函数形式本身的变更。在上述的例子中，$x(t)$ 代表通过 A 及 B 的任一个函数，并且 $x(t_0)=x_0, x(t_e)=x_e$。若 $x(t)$ 变更为 $x(t)+\alpha\eta(t)$，其中 $\eta(t)$ 是 t 的任意函数但满足条件 $\eta(t_0)=\eta(t_e)$，因而 $x(t)+\alpha\eta(t)$ 也是通过 A 及 B 的函数；当参数 α 的绝对值足够小时，$\alpha\eta(t)$ 代表函数的微小变更，所以 $\delta x=\alpha\eta(t)$。

$x(t)$ 变更将引起 $\dot{x}(t)$、$L(x,\dot{x})$ 及 $\int_{t_0}^{t_e} L\mathrm{d}t$ 变更，分别以 $\delta\dot{x}$、δL、$\delta\int_{t_0}^{t_e} L\mathrm{d}t$ 表示。定义

$$\delta\dot{x}=\frac{\mathrm{d}}{\mathrm{d}t}(x+\delta x)-\frac{\mathrm{d}x}{\mathrm{d}t} \tag{1.3-1}$$

$$\delta L(x,\dot{x})=L(x+\delta x,\dot{x}+\delta\dot{x})-L(x,\dot{x}) \tag{1.3-2}$$

$$\delta\int_{t_0}^{t_e} L(x,\dot{x})\mathrm{d}t=\int_{t_0}^{t_e} L(x+\delta x,\dot{x}+\delta\dot{x})\mathrm{d}x-\int_{t_0}^{t_e} L(x,\dot{x})\mathrm{d}x \tag{1.3-3}$$

根据这些变分符号的含义，可得下面的变分运算规则

$$\delta\dot{x}=\frac{\mathrm{d}}{\mathrm{d}t}\delta x \tag{1.3-4}$$

$$\delta\int_{t_0}^{t_e} L\mathrm{d}t=\int_{t_0}^{t_e}\delta L\mathrm{d}t \tag{1.3-5}$$

亦即变分运算可与对 t 求导数及积分运算交换次序。还要指出的是，当 $x(t)$ 变更时，t 并不变化，所以

$$\delta t=0 \tag{1.3-6}$$

变分与微分也还有相似之处。当求由于 x 的微小变更所引起的 $f(x)$ 的变更时，其方法与求 $f(x)$ 的微分一样，只要将微分改为变分，即

$$\delta f(x)=\frac{\mathrm{d}f}{\mathrm{d}x}\delta x \tag{1.3-7}$$

现在来看哈密顿原理所说的哈密顿作用量 $J=\int_{t_0}^{t_e} L\mathrm{d}t$ 在实际路径上取极值用变分学的语言怎样表达。路径的变更相当于函数 $x(t)$ 的变更。如果把 J 看做 x 的函数，按照式(1.3-7)，$\delta J=\frac{\mathrm{d}J}{\mathrm{d}x}\delta x$，在极值处，$\frac{\mathrm{d}J}{\mathrm{d}x}=0$，于是 $\delta J=0$。因此，哈密顿原理可以用变分学的语言表达为：力学体系的哈密顿作用量的变分

$$\delta J=\delta\int_{t_0}^{t_e} L\mathrm{d}t=0 \tag{1.3-8}$$

式中，L 是体系的拉格朗日函数。

以上讨论的是一维问题，只有一个函数变更。当有 s 个广义坐标时，相应有 s 个函数变更，上述的讨论仍然适用，只要将拉格朗日函数及相应的变分作如下改写，即

$$L=L(q_1,q_2\cdots q_s,\dot{q}_1,\dot{q}_2\cdots\dot{q}_s,t) \tag{1.3-9}$$

$$\begin{aligned}\delta L=&L(q_1+\delta q_1,q_2+\delta q_2\cdots q_s+\delta q_s,\dot{q}_1+\delta\dot{q}_1,\dot{q}_2+\delta\dot{q}_2\cdots\dot{q}_s+\delta\dot{q}_s,t)-\\&L(q_1,q_2\cdots q_s,\dot{q}_1,\dot{q}_2\cdots\dot{q}_s,t)\end{aligned} \tag{1.3-10}$$

$$\delta\dot{q}_k=\frac{\mathrm{d}}{\mathrm{d}t}\delta q_k \tag{1.3-11}$$

$$\delta f(q_1,q_2\cdots q_s)=\sum_{k=1}^{s}\frac{\partial f}{\partial q_k}\delta q_k \tag{1.3-12}$$

哈密顿原理仍然表达为

$$\delta J=\delta\int_{t_0}^{t_e} L\mathrm{d}t=\int_{t_0}^{t_e}\delta L\mathrm{d}t=0 \tag{1.3-13}$$

1.3.2 哈密顿原理的推演

历史上这个原理是由哈密顿当做公理提出来的。以这个原理为出发点可推演出运动的微分运动方程式而不必应用牛顿定律；另一方面，也可以牛顿定律为出发点推演得到哈密顿原理。

设体系的拉格朗日函数为

$$L=L(q_1,q_2\cdots q_s,\dot{q}_1,\dot{q}_2\cdots\dot{q}_s,t) \tag{1.3-14}$$

式中，$q_1,q_2\cdots q_s$ 代表广义坐标（自由度为 s），$\dot{q}_1,\dot{q}_2\cdots\dot{q}_s$ 代表相应的广义速度，作为较普遍形式的 L 还显含时间 t。体系经过给定的始态（t_0）及终态（t_e）的哈密顿作用量为

$$J=\int_{t_0}^{t_e} L\mathrm{d}t \tag{1.3-15}$$

当路径变更时，J 的变分为

$$\delta J=\delta\int_{t_0}^{t_e} L\mathrm{d}t=\int_{t_0}^{t_e}\delta L\mathrm{d}t \tag{1.3-16}$$

s 个 q_k 的变更由变分 $\delta q_k(k=1,2\cdots s)$ 表示，相应的 $\dot{q}_k$ 变更为 $\delta\dot{q}_k$。q_k 的变更不引起 t 的变化，所以

$$\delta J=\int_{t_0}^{t_e}\sum_{k=1}^{s}\left(\frac{\partial L}{\partial q_k}\delta q_k+\frac{\partial L}{\partial\dot{q}_k}\delta\dot{q}_k\right)\mathrm{d}t \tag{1.3-17}$$

注意到 $\delta\dot{q}_k=\frac{\mathrm{d}}{\mathrm{d}t}\delta q_k$，并使用分部积分方法，可将积分号下括弧内的第二项化为