第 1 章　拉格朗日力学

牛顿力学中的定律是关于自然界中宏观运动现象的基本规律，主要描述的是物体在力的作用下的运动规律。在牛顿力学之后，人们又发现了其他形式的力学概念体系，如拉格朗日力学和哈密顿力学。这两种力学使用了更为一般性的描述方式，但在形式上却更为抽象。在描述宏观力学现象上，二者与牛顿力学完全等价。在描述微观现象时，二者作为理论框架，更具一般性。因此，我们有必要学习这些力学体系。

按今天人们的认知，对称性在物理学中居于基础地位。今天的物理学家们从对称性出发建构基本的物理规律，人们将各种各样的相互作用看作是不同对称性的自然结果。一条数学定理——诺特定理——告诉我们，对称性也是守恒律存在的原因。由于对称性如此重要，我们需要了解对称性的一般性描述方式。

在本章的学习中，我们将先简要地回顾牛顿力学，然后从最小作用量原理出发，介绍拉格朗日力学。接着，探讨对称性与守恒律之间的关系，并引入用于描述对称性的数学语言。最后，讨论牛顿力学的时空观。

1.1　从牛顿力学到广义坐标

先回顾牛顿力学的基本内容，比较其在不同物理内容中的基本形式；然后引入广义坐标的概念；最后介绍两种常用的坐标系，即柱坐标系和球坐标系。

1.1.1　牛顿力学回顾

牛顿力学的基础是来自于日常经验的三条定律，这三条定律称为牛顿运动三定律。

（1）物体将始终处于静止或匀速直线运动状态，除非有外力改变这种状态。

（2）物体在外力的作用下将作加速运动，加速度正比于外力。比例系

数被称为质量，为物体本身固有性质。

（3）当一个物体对另一个物体施加力的作用时，它本身也受到另一个物体同等强度但方向相反的力的作用。

1. 牛顿运动定律

牛顿运动定律是关于力与运动的基础定律，是能够完备地描述宏观物体一般性运动特征的规律。由于读者大多已熟悉牛顿的运动定律，因而这里仅特别强调以下几点。

首先，在牛顿运动定律中反复出现的“物体”的概念，我们只需做直观性的理解，不宜过分深究。物理学是关于大自然的科学，其定律的得出常常依赖于直观性的概念与描述。试图以数学公理化的方式为物理学定律构建基础，严格定义所有概念，常常是极不明智的。这么说的原因主要有两点，一是研究者们研究特定物理学规律的时候，总是忽略掉大量微弱影响的因素，这使得物理规律具有天然的近似性；二是不同尺度上的物理学规律相当不同，而这种尺度上的变化没有明确的界限。因此在我们研究物理学规律时常常需要依赖一定的直观认知，比如对物体这一概念的认知。在牛顿力学定律中出现的物体概念，不宜被质点甚至粒子这样的概念替换。质点是一个纯粹的抽象概念，有一定的物理信息上的丢失。而粒子则隐含着微观性的特征，然而对于微观粒子来说，将其简单地视作“沙粒”之类的模型又是不正确的。所以，我们在描述牛顿力学定律的时候，用的是物体而不是质点或者粒子这样的概念，虽然我们解决问题时常常将物体简化为“质点”以便于讨论。质点，只是我们在处理力学问题时常用的模型中的一种。

其次，对于“力”这个概念我们也只做这种直观性的理解，因为在微观层次上“力”也并不是一个好的概念，需要被其他概念（比如场及场的相互作用）代替。同时，“力”的具体形式来自于牛顿力学定律的外部。牛顿力学定律阐明的是力和运动的关系，力的形式并不是牛顿力学处理的问题。比如，我们知道物体之间有万有引力，其大小与距离的平方成反比，性质为吸引性；带电物体之间有库仑力，其大小与距离的平方成反比，根据电荷的不同，性质为吸引或排斥。像万有引力或库仑力的具体形式，并不是我们所说的“牛顿力学”这一知识系统负责研究的内容。我们需要通过其他的理论，弄清楚这些力的形式后，直接将其放到牛顿力学体系中加以应用。值得注意的是，根据做功是否只取决于位置而与路径无关，力被分为**保守力**和**非保守力**两种。类似于万有引力和库仑力这样的力，都是保守力。保守力具有基本性，一切非保守力本质上都是保守力的某种剩余效应。

比如说，在高中的时候，我们会说一支放在桌面上的钢笔会受到桌面施加的支撑力的作用，支撑力就是非保守力。在宏观层面上的这种支撑力本质上是电磁力。我们把钢笔与桌面的连接处不断放大，最后我们会看到组成钢笔的原子与组成桌面的原子。这两组原子之间有电磁力的作用。电磁力使两者之间保持一定的距离，无法彼此穿过，因而我们在宏观上看到钢笔受桌面施加的支撑力的影响。关于这个问题我们后面的学习中还要讨论。

再次，质量是作为比例系数被定义的。要想使不同物体获得同样的加速度，则物体质量越大，需要施加的力越大；物体质量越小，需要施加的力越小。因此质量代表着改变物体运动状态的难易程度。由于在牛顿的万有引力定律中也定义了质量，因此在逻辑上，牛顿运动三定律体系与牛顿万有引力定律体系就在质量定义的问题上出现了重复定义的矛盾性。这一矛盾需要留到广义相对论的学习中解决。

最后，静止和匀速直线运动是等价的，这涉及参考系的选取。牛顿力学定律隐含定义了**惯性参考系**的概念。惯性参考系即牛顿第一定律成立的参考系。牛顿运动定律是一个一般性理论，因而我们自然会想到牛顿第一定律意味着在最一般的情况下，即一般性的空间（而不是某个特殊的房间或其他的什么空间）中，牛顿运动定律成立。这一点与我们的经验相符合。使牛顿第一定律不成立的空间总是存在着有差异的位置或特殊的方向，也就是使物体无法作匀速直线运动的位置或方向。牛顿运动定律成立即意味着没有这样的位置或特殊方向，也就是说空间均匀且各向同性。当然，最一般性的空间也不应与时间有关。我们可以把空间均匀且各向同性、时间均匀当作惯性系的定义。

2. 牛顿第二定律

从求解运动问题的角度看，牛顿运动三定律中最重要的是牛顿第二定律，即

$$\boldsymbol{F} = m\boldsymbol{a} \tag{1.1}$$

其中 $\boldsymbol{a}$ 代表加速度，即速度随时间的变化快慢。而速度又是代表物体位置变化快慢的物理量。在物理学中，变化快慢指的是某个物理量随时间的变化率，用数学的语言说即是某个物理量对时间的导数，因此我们有

$$\boldsymbol{a} = \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} \tag{1.2}$$

其中 $\boldsymbol{r}$ 代表物体在空间中的**位置矢量**（**位矢**）。在选取好的参考系中，根据所讨论的问题，确定好适合的坐标系，可以把位置记成坐标，这样我们

就可以做具体的数学描述了。当我们把位矢 $\boldsymbol{r}$ 当成时间 t 的函数时，它随着时间的连续变化就是物体在空间中的运动轨迹。当我们说用牛顿运动定律求解问题的时候，我们主要说的就是解出作为时间函数的 $\boldsymbol{r}(t)$ 的具体函数形式。

在讨论实际问题时，人们会引入**动量**的概念，将牛顿力学写成

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} \tag{1.3}$$

其中动量

$$\boldsymbol{p} = m\boldsymbol{v} \tag{1.4}$$

在后面的学习中，我们会看到，动量比速度和加速度更具基本性。

3. 牛顿力学对转动问题的描述

有一大类力学问题是转动问题，比如推门就是一个转动问题。人们在描述力学问题时讨论的是位置的变化，在转动问题里有一个天然的位置变量，就是角度。只要知道了角度，我们就知道了转动（比如推门）进行得怎么样。转动研究的是角度的变化问题。角度变化的快慢称为**角速度**。虽然物体运动是因为有力的存在，但在用角度描述转动时，力并不是唯一需要知道的物理量。在推门时，我们手按的位置也影响推门的效果。施加同样的力，推靠近门轴的地方就更难一些。因而，在考虑转动问题时，人们会引入**力矩**的概念。力矩被定义为

$$\boldsymbol{M} = \boldsymbol{r} \times \boldsymbol{F} \tag{1.5}$$

这里的乘法是数学上的“叉乘”。两个矢量叉乘后得到的矢量总是垂直于这两个矢量本身。转动问题可以根据牛顿第二定律写成力矩的形式，即

$$\boldsymbol{M} = \frac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t} \tag{1.6}$$

其中 $\boldsymbol{J}$ 为角动量。对于质点来说，角动量可以写成坐标与动量的叉乘：

$$\boldsymbol{J} = \boldsymbol{r} \times \boldsymbol{p} \tag{1.7}$$

在考虑单个质点绕某个定点的转动问题时

$$\boldsymbol{J} = I\boldsymbol{\Omega} \tag{1.8}$$

其中 $\boldsymbol{\Omega}$ 为角速度，即角度随时间的变化率；而 I 是质点相对于定点的**转动惯量**。若两点之间距离为 r，则

$$I = mr^2 \tag{1.9}$$

多个质点的转动问题将在本书的刚体相关内容中讨论。

4. 规律的相似性

以上内容就是我们关于牛顿力学所知道的一些基本概念与定律。有了这些基本定律与概念后，各种力学问题都可以按照先写下正确的力，再求解方程这一流程进行分析研究。我们无意讨论如何做具体的计算，因为读者通常已经在力学课程中学过相关内容。我们在这里关心的是更一般性的问题。

我们将牛顿第二定律与转动问题中的公式比较一下，会发现两者有高度的类似性。将两类问题中涉及的物理量和公式列在表 1.1中。从这个对比中我们可以很明显地看出来，一般性的运动与转动问题虽然采用了不同的物理量进行描述，但是其形式高度一致。我们只要将描述一般性问题的位矢、力、质量等概念替换成角度、力矩、转动惯量等概念，二者在规律形式上就完全一致。这使得人们猜想，有没有可能一般性地描述运动，不去区分坐标到底代表的是位矢还是角度。这种探索使人们提出了广义坐标的概念以及拉格朗日力学。

表 1.1 不同运动公式的比较

位矢	力 $\boldsymbol{F}$	动量 $\boldsymbol{p}$	质量 m	$\boldsymbol{F} = \dfrac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}$
角度	力矩 $\boldsymbol{M}$	角动量 $\boldsymbol{J}$	转动惯量 I	$\boldsymbol{M} = \dfrac{\mathrm{d}\boldsymbol{J}}{\mathrm{d}t}$

1.1.2 广义坐标

若描述一个力学系统中所有物体的位置需要 s 个独立变量，我们就说这个系统的**自由度**为 s。比如在桌面上运动的质点，其自由度就是 2。去掉桌面这个限制，在房间中运动的质点，其自由度是 3。谈论自由度时，要注意“独立”二字，若有某种约束（如限定在桌面上运动）存在，则自由度会相应变化。考虑两个质点，若它们都是独立运动的，则这样一个双质点系统的自由度就是 6。若两个质点被一根质量可忽略的细杆连接起来，则二者之间的距离将保持不变，距离保持不变这个约束条件使得这样的双质点系统的自由度变成了 5。

对于自由度为 s 的力学系统，我们可以选择 s 个坐标来描述这个系统。在以前的学习中，我们常常采用笛卡儿三维直角坐标系（简称笛卡儿坐标系或直角坐标系）来描述位置。但在有些问题里，并不需要知道所有质点的笛卡儿坐标也可以描述清楚问题，甚至还更容易。比如描述转动，我们需要的就只是角度。因而人们用广义坐标的概念来描述力学系统。

对于自由度为 s 的力学系统，能够完全描述其运动的 s 个独立变量被称为**广义坐标**。在本书中，我们一般用 $q_1, q_2, \cdots, q_s$ 来表示广义坐标。"广义"的意思在于不必拘束于笛卡儿坐标，两个坐标的间隔也不必非得是长度，也可以是角度或其他物理量。若我们描述桌面上绕着某一定点转动的笔杆，只需要选取一个角度作为广义坐标就能完整描述这个运动了。广义的意思也在于不必拘束于描述某种位置，只要能完整地描述所讨论的物理系统的自由度的物理量，都可以是广义坐标。在物理学中，非常重要及基础的一部分理论是**场论**，如电磁学或电动力学就是关于"场"的理论。在场论中，场本身就是广义坐标。场论就是用场作为广义坐标构建的理论。

根据经验，只知道坐标并不能完整了解一个力学系统的信息，通常还需要知道速度[①]。在这里我们引入广义速度的概念，广义速度就是广义坐标对时间的导数，记为 $\dot{q}_\alpha$。在后面，我们将学习用广义坐标和广义速度描述物理系统的拉格朗日力学。

为了更方便地选用适合的广义坐标来描述运动，我们有必要在讨论力学规律之前先了解一下常用的坐标系。

1.1.3 柱坐标系和球坐标系

在描述物体的运动时，人们需要选择坐标系以便定量计算。直角坐标系是最常见又最直观的坐标系，但它并不总是那么方便，特别是在描述一些具有某种对称性的物理量或规律时更是如此。比如，万有引力定律是一种平方反比力，若在直角坐标下将其写出，则是

$$F = \frac{Gm_1m_2}{x^2+y^2+z^2} \tag{1.10}$$

这里我们将某个物体的位置定义为坐标原点。由于这个力只与两点之间的距离有关,而在某一个点被定义为坐标原点的情况下,两点的距离可以写成

$$r = \sqrt{x^2+y^2+z^2} \tag{1.11}$$

显然式 (1.10) 只与一个变量 r 有关，我们却将其写成三个变量 x, y, z 而徒增烦恼。在这个定律中，三个坐标 x, y, z 互换，物理规律不变，这就是一

① 这是经验，而不是数学定理提出的要求。

种对称性。选择合适的坐标系统，可以帮助我们将公式写成只与一个变量 r 有关的形式。因此人们还需要直角坐标系以外的其他坐标系，如柱坐标系和球坐标系。

1. 柱坐标系

在描述某些具有轴对称性的问题时，人们常常选用**柱坐标系**来描述问题。柱坐标系可以看作是平面上的极坐标系和一个直角坐标系的轴的组合。我们在直角坐标系中选出一个维度，比如 z 轴所代表的维度，则直角坐标系可用一个二维空间（xy 平面所代表的空间）与 z 轴来描述。对于 xy 平面，我们可以用极坐标系来描述。极坐标系用坐标 (ρ,φ) 来代表平面上的点，其中 ρ 代表某点到原点的距离，φ 代表该点与原点的连线与选定的基准轴的夹角。因而，三维空间中的任何一个点可以表示成 (ρ,φ,z)，这种表示空间中位置的坐标系就是柱坐标系。显然，柱坐标系中的坐标 (ρ,φ,z) 与直角坐标系中的坐标 (x,y,z) 有如下关系：

$$\begin{cases} x=\rho\cos\varphi \\ y=\rho\sin\varphi \\ z=z \end{cases} \tag{1.12}$$

柱坐标系适合用来讨论相对于 z 轴有轴对称性的问题。

2. 球坐标系

球坐标系是极常用的一种坐标系，如图 1.1 所示，因为它可以将长度部分与角度部分分开。这样的特点有助于它方便地处理：①被积函数只依赖于径向部分（长度）的空间积分；②具有球对称性的微分方程问题（同样也是强调仅有径向依赖这一特点）。

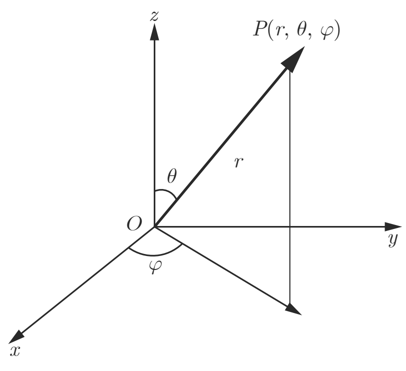

图 1.1 球坐标系

球坐标系用一个坐标 r 来表示三维空间中一个点（记为 P 点）距离原点（记为 O 点）的距离，再用两个角度来表示原点到 P 点的射线在空间中的相对位置。为方便起见，我们可以借助直角坐标系来说明球坐标系，并同时给出两个坐标系的变换关系。在球坐标系中需要定义一个特殊的方向，比如我们可以将这个方向定义为 z 轴的正向。定义之后我们将这个轴称为**极轴**，极轴与 OP 之间的夹角被称为**极角**，用 θ 表示。利用极角，可以将射线 OP 投影到 z 轴以及与 z 轴垂直的 xy 平面上，两个投影的长度分别为 $r\cos\theta$ 和 $r\sin\theta$。将 OP 在 xy 平面上的投影与 x 轴的夹角称为**方位角**，记为 φ。显然，OP 在 xy 平面上的投影（长度为 $r\sin\theta$）在 x 轴上的投影长 $r\sin\theta\cos\varphi$，在 y 轴上的投影长 $r\sin\theta\sin\varphi$。我们用坐标 (r,θ,φ) 就能完全表示 (x,y,z)，即

$$\left.\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{aligned}\right\} \tag{1.13}$$

直角坐标 (x,y,z) 可代表空间中任意一点，因而与之等价的 (r,θ,φ) 也可以代表空间中任意一点。用一个长度 r 与两个角度 θ,φ 构成的坐标系被称为球坐标系。人们一般将 r 的变化方向称为径向。很明显，极角 θ 变化的方向垂直于径向，也垂直于方位角 φ 变化的方向；方位角 φ 的变化方向也垂直于径向。同直角坐标系一样，三个坐标变化的方向彼此垂直，这种坐标变化的方向彼此垂直的坐标系被称为**正交坐标系**。

3. 不同坐标系中的无限小变化与体积微元

在处理物理问题时，我们常常使用无限小的量，如无限小的长度，无限小的体积等。用无限小的好处是可以建立简单的关系。在无限小的情况下，**不均匀可被看作均匀，弯曲可被看作平直**，因而便于建立物理量之间的关系。将无限小的量求和就得到了有限情况下的物理量。无限小的求和就是积分。

在直角坐标系中，三个无限小长度很简单，分别是 $\mathrm{d}x,\mathrm{d}y$ 和 $\mathrm{d}z$。由于直角坐标系中的三个方向互相垂直，因而直角坐标系中的无限小体积可以表示成 $\mathrm{d}x\mathrm{d}y\mathrm{d}z$。

柱坐标系是平面极坐标系和一个笛卡儿坐标的组合。笛卡儿坐标很简单，对应的无限小长度就是 $\mathrm{d}z$，将其乘以极坐标系的面积微元就得到了柱坐标系的体积微元。在平面极坐标系中，当长度 ρ 与角度 φ 分别变化 $\mathrm{d}\rho$ 与 $\mathrm{d}\varphi$ 时，在径向（ρ 的方向）上的长度变化就是 $\mathrm{d}\rho$，在横向（φ 的变

化方向）上的长度变化是 $\rho\mathrm{d}\varphi$（弧长公式），如图 1.2 所示。径向与横向彼此垂直，因而面积（矩形）微元为 $\rho\mathrm{d}\rho\mathrm{d}\varphi$。将面积微元乘以 $\mathrm{d}z$ 就得到了柱坐标系中的体积微元

$$\mathrm{d}V_{柱} = \rho\mathrm{d}\rho\mathrm{d}z\mathrm{d}\varphi \tag{1.14}$$

其中 ρ 的取值范围为零到无穷大，φ 的取值范围为 $0 \sim 2\pi$。

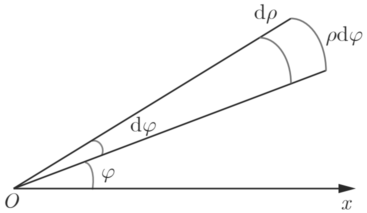

图 1.2　极坐标系的面积微元

球坐标系的体积微元也可以类似求出。在球坐标系中有三个方向，即 r 的变化方向，极角 θ 的变化方向和方位角 φ 的变化方向，三者互相垂直。这三个变量的无穷小变化 $\mathrm{d}r$，$\mathrm{d}\theta$ 与 $\mathrm{d}\varphi$ 组成了这三个彼此垂直的方向上的无穷小长度变化。在径向 r 的方向上长度变化就是 $\mathrm{d}r$。其他两个方向上的无穷小长度变化都可以像极坐标系一样用弧长公式求出。因而极角 θ 的变化带来的无穷小长度是 $r\mathrm{d}\theta$；而方位角 φ 的变化对应的“半径”是 $r\sin\theta$，因而带来的无穷小长度变化就是 $r\sin\theta\mathrm{d}\varphi$，如图 1.3 所示。将这三个无穷小长度相乘就得到了球坐标系中的体元

$$\mathrm{d}V_{球} = r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \tag{1.15}$$

同样地，r 的取值范围为零到无穷大，极角 θ 的变化范围为 $0 \sim \pi$，方位角 φ 的变化范围为 $0 \sim 2\pi$，这样就能穷尽整个空间且无重复。

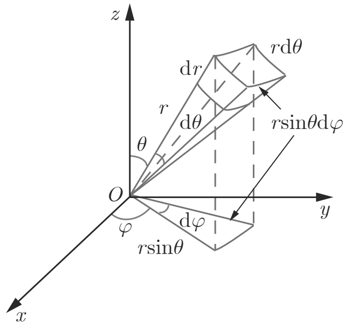

图 1.3　球坐标系的体积微元

4. 哈密顿算符

在讨论物理学问题时，我们经常会讨论到某种物理量随空间位置的变化[①]，这时我们就会用到**哈密顿算符**[②]。哈密顿算符一般用符号 ∇ 表示，它是一个矢量型算符。根据定义，在直角坐标系中，哈密顿算符被记为

$$\nabla=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\quad \text{或}\quad \frac{\partial}{\partial x}\boldsymbol{e_x}+\frac{\partial}{\partial y}\boldsymbol{e_y}+\frac{\partial}{\partial z}\boldsymbol{e_z} \tag{1.16}$$

这里的 $\boldsymbol{e_x},\boldsymbol{e_y},\boldsymbol{e_z}$ 指的是三个直角坐标方向的单位矢量，又称基矢。从这个定义式中，我们可以很明显地看出哈密顿算符的意义，即物理量在不同方向上随坐标改变的变化率。因此我们可以根据前面对每个方向上无穷小长度的分析直接给出柱坐标系和球坐标系中的哈密顿算符。在柱坐标系中

$$\nabla=\boldsymbol{e_\rho}\frac{\partial}{\partial\rho}+\boldsymbol{e_\varphi}\frac{\partial}{\rho\partial\varphi}+\boldsymbol{e_z}\frac{\partial}{\partial z} \tag{1.17}$$

在球坐标系中

$$\nabla=\boldsymbol{e_r}\frac{\partial}{\partial r}+\boldsymbol{e_\theta}\frac{1}{r}\frac{\partial}{\partial\theta}+\boldsymbol{e_\varphi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi} \tag{1.18}$$

平面直角坐标系是一种固定坐标系，即其基矢是固定的。但并非所有的坐标系都有固定基矢的特征。我们很明显地可以看到，球坐标系中三个坐标的变化方向就是始终变动的，但是其彼此之间的垂直关系一直保持。在进行求导运算时，要注意柱坐标系与球坐标系的基矢本身也随坐标的变化而改变，因而也应被求导。

5. 梯度，散度与旋度

若某个物理量只有大小没有方向，这样的物理量被称为**标量**。当一个标量型的物理量在空间的不同位置处大小不同时，我们可以用这个物理量的**梯度**来表示其在空间不同方向上的变化情况。比如，一座山上不同位置处的温度不同，我们将温度记为 $T(x,y,z)$，则

$$\nabla T(x,y,z)=\frac{\partial T}{\partial x}\boldsymbol{e_x}+\frac{\partial T}{\partial y}\boldsymbol{e_y}+\frac{\partial T}{\partial z}\boldsymbol{e_z} \tag{1.19}$$

---

① 遍布于空间并在空间不同位置处有不同大小的物理对象的基本形态一般被称作“场”，相关的理论被称为场论。场是非常基本的物理学概念，在物理学中有广泛应用。

② 算符，也有人将其称为**算子**，指的是某种数学操作，比如求导操作。符号 ∇ 读作 nabla。

反映了温度在各个方向上随坐标变化的情况。梯度体现了一个物理量在空间中上梯次变化的特征。

有些物理量除了大小还有方向，比如电场强度 $\boldsymbol{E}$，我们将这种物理量称为矢量。对于矢量，我们可以用**散度**和**旋度**描述其在空间中的分布特征。

散度在数学上指的是哈密顿算符与矢量型物理量的点乘，如 $\nabla\cdot\boldsymbol{E}$。散度在物理上体现的是物理量的“源”的特征。数学上有一个高斯定理：某个矢量（如 $\boldsymbol{E}$）的散度的体积分等于该矢量在该空间表面上的面积分，即[1]

$$\iiint \nabla\cdot\boldsymbol{E}\mathrm{d}V = \oiint \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{S} \tag{1.20}$$

因而，若一个矢量的散度不为零，如

$$\nabla\cdot\boldsymbol{E} = \rho \tag{1.21}$$

则

$$\oiint \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{S} = \iiint \rho\mathrm{d}V \tag{1.22}$$

矢量 $\boldsymbol{E}$ 在表面上的大小取决于 ρ 的分布情况。散度代表的是“源” ρ 导致的物理量 $\boldsymbol{E}$ 向不同方向发散开来的情况。若“无源”，即 ρ 等于零，则

$$\oiint \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{S} = 0 \tag{1.23}$$

即在所封闭的表面上没有 $\boldsymbol{E}$ 流出。散度体现的是“源”的分布。

旋度指的是哈密顿算符与矢量的叉乘。根据二维的高斯定理，即斯托克斯定理，一个矢量的旋度，如 $\nabla\times\boldsymbol{E}$，在某个区域上的面积分等于该矢量在围绕该区域的边界上的线积分，即

$$\iint (\nabla\times\boldsymbol{E})\cdot\mathrm{d}\boldsymbol{S} = \oint \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{l} \tag{1.24}$$

因此，若矢量 $\boldsymbol{E}$ 的旋度 $\nabla\times\boldsymbol{E}$ 不为零，则线积分不为零，这意味着该矢量呈现出一种“绕圈”或“螺旋”的特征。旋度体现的是一个物理量“螺旋”的强度。

---

[1] 这里为了使初学者看得明白，而特别对三维积分用了三个积分号，二维积分用了两个积分号。这样做毫无必要。以后我们一般只写一个积分号。

6. 拉普拉斯算符

在很多物理问题中,我们都会用到“梯度的散度”这个算符,即 $\nabla\cdot\nabla=\nabla^2$,有时被记作 $\triangle$。这个算符被称为拉普拉斯算符[①]。

直角坐标系是一个基矢固定的坐标系,因而两个算符的点乘可直接得到。直角坐标系中

$$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2} \tag{1.25}$$

柱坐标系和球坐标系中的情况复杂一些,因为其基矢方向会变化。

我们先看二维球坐标系即极坐标系的情况。极坐标系中有两个基矢,即 $\boldsymbol{e_\rho}$ 与 $\boldsymbol{e_\varphi}$。显然,当 ρ 变化时,两个基矢的方向都不会发生变化,用数学的语言写出来就是

$$\frac{\partial \boldsymbol{e_\rho}}{\partial\rho}=0,\quad \frac{\partial \boldsymbol{e_\varphi}}{\partial\rho}=0 \tag{1.26}$$

这里使用了偏导数,因为我们需要讨论基矢分别随着不同变量改变的变化。

当 φ 大小变化时,两个基矢都会变化。当 φ 变为 $\varphi+\Delta\varphi$ 时,变化后的径向基矢可用平行四边形法则和弧长公式得到,应为 $\boldsymbol{e_\rho}+\Delta\varphi\boldsymbol{e_\varphi}$。也就是说,基矢 $\boldsymbol{e_\rho}$ 改变了 $\Delta\varphi\boldsymbol{e_\varphi}$ 这么多。类似地,我们也可以用平行四边形法则和弧长公式得出另一个基矢 $\boldsymbol{e_\varphi}$ 的改变量 $\Delta\boldsymbol{e_\varphi}$ 为 $-\Delta\varphi\boldsymbol{e_\rho}$。如图 1.4 所示,由 $\Delta\varphi$ 引起的两个基矢的变化为

$$\Delta\boldsymbol{e_\rho}=\Delta\varphi\boldsymbol{e_\varphi},\quad \Delta\boldsymbol{e_\varphi}=-\Delta\varphi\boldsymbol{e_\rho} \tag{1.27}$$

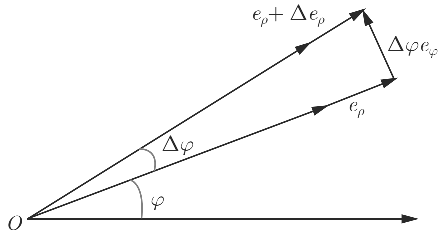

图 1.4 极坐标系中的单位矢量 $\boldsymbol{e_\rho}$ 会随着 φ 的变化而变化。当 φ 改变了 $\Delta\varphi$ 时,$\boldsymbol{e_\rho}$ 变到了 $\boldsymbol{e_\rho}+\Delta\boldsymbol{e_\rho}$。显然这个变化的方向为 $\boldsymbol{e_\varphi}$,变化的大小为单位矢量的长度(即 1)乘以变化的角度 $\Delta\varphi$。因此,$\Delta\boldsymbol{e_\rho}=\Delta\varphi\boldsymbol{e_\varphi}$。

---

① 在物理学中,有些符号具有通用性,比如 ∇;而有些符号用得比较少,不应将其视作通用符号,比如拉普拉斯算符 $\triangle$。另外,人们常常用 Δ 代表变化。比如我们在这里将使用 $\Delta\varphi$ 代表 φ 这个变量的变化量。本书中我们一般默认变化量 $\Delta\varphi$ 是个无限小变化量。

在无限小变化下，可写成偏导数形式：

$$\frac{\partial \boldsymbol{e_\rho}}{\partial \varphi} = \boldsymbol{e_\varphi}, \quad \frac{\partial \boldsymbol{e_\varphi}}{\partial \varphi} = -\boldsymbol{e_\rho} \tag{1.28}$$

利用这些基矢随坐标变化的公式,就可以直接计算哈密顿算符的点乘，得到极坐标系中的拉普拉斯算符

$$\begin{aligned}\nabla^2_{\text{极}} &= \left(\boldsymbol{e_\rho}\frac{\partial}{\partial \rho} + \boldsymbol{e_\varphi}\frac{1}{\rho}\frac{\partial}{\partial \varphi}\right)\cdot\left(\boldsymbol{e_\rho}\frac{\partial}{\partial \rho} + \boldsymbol{e_\varphi}\frac{1}{\rho}\frac{\partial}{\partial \varphi}\right)\\ &= \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial^2 \varphi}\\ &= \frac{1}{\rho}\frac{\partial}{\partial \rho}\rho\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial^2 \varphi}\end{aligned} \tag{1.29}$$

这里用到了求导的链式法则和基矢的正交归一性，即

$$\boldsymbol{e_\rho}\cdot\boldsymbol{e_\rho} = \boldsymbol{e_\varphi}\cdot\boldsymbol{e_\varphi} = 1, \quad \boldsymbol{e_\rho}\cdot\boldsymbol{e_\varphi} = 0 \tag{1.30}$$

对于柱坐标系来说，由于 z 轴垂直于 ρ 与 φ 的平面，因此彼此没有影响，完全独立。这使得我们可以立即从极坐标的公式得到柱坐标系中的拉普拉斯算符：

$$\nabla^2_{\text{柱}} = \frac{1}{\rho}\frac{\partial}{\partial \rho}\rho\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial^2 \varphi} + \frac{\partial^2}{\partial z^2} \tag{1.31}$$

同样的分析方式还可应用于球坐标系。首先，径向部分的变化显然不会引起任何方向的变化，因而

$$\frac{\partial \boldsymbol{e_r}}{\partial r} = 0, \quad \frac{\partial \boldsymbol{e_\theta}}{\partial r} = 0, \quad \frac{\partial \boldsymbol{e_\varphi}}{\partial r} = 0 \tag{1.32}$$

至于 θ 的变化引起的 $\boldsymbol{e_r}$ 与 $\boldsymbol{e_\theta}$ 的变化，同极坐标的情况完全一样。又由于 θ 变化的平面始终垂直于 φ 所在的平面，因而不会影响 $\boldsymbol{e_\varphi}$。于是我们有

$$\frac{\partial \boldsymbol{e_r}}{\partial \theta} = \boldsymbol{e_\theta}, \quad \frac{\partial \boldsymbol{e_\theta}}{\partial \theta} = -\boldsymbol{e_r}, \quad \frac{\partial \boldsymbol{e_\varphi}}{\partial \theta} = 0 \tag{1.33}$$

最后我们讨论 φ 的变化引起的基矢变动。φ 在固定平面上变动，我们可以把它当作是极坐标的角度。为了分析变化情况，引入这个平面上的“径向”矢量，即引入三维径向矢量在 φ 平面上的投影的单位矢量 $\boldsymbol{e_\perp}$。$\boldsymbol{e_r}$、

$\boldsymbol{e_\theta}$ 与 $\boldsymbol{e_\perp}$ 构成的平面就是 θ 变化的平面。因此，我们可以直接用 $\boldsymbol{e_r}$ 与 $\boldsymbol{e_\theta}$ 表示出 $\boldsymbol{e_\perp}$，如图 1.5 所示。

$$\boldsymbol{e_\perp} = \sin\theta \boldsymbol{e_r} + \cos\theta \boldsymbol{e_\theta} \tag{1.34}$$

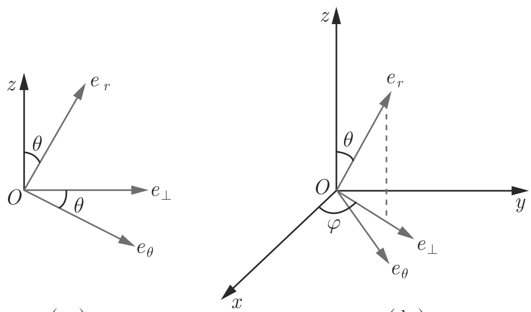

图 1.5　图 (a) 为图 (b) 中的一个剖面，即 θ 变化的面。$\boldsymbol{e_\perp}$ 为矢量在 φ 平面上投影的基矢，可用 $\boldsymbol{e_r}$ 与 $\boldsymbol{e_\theta}$ 在 $\boldsymbol{e_\perp}$ 上的投影表示出 $\boldsymbol{e_\perp}$。同样地，$\boldsymbol{e_\perp}$ 在 $\boldsymbol{e_r}$ 与 $\boldsymbol{e_\varphi}$ 方向上的分量分别为 $\sin\theta\boldsymbol{e_\perp}$ 与 $\cos\theta\boldsymbol{e_\perp}$。

有了 $\boldsymbol{e_\perp}$，我们就可以按照极坐标的方式，直接写出 φ 的变化对 $\boldsymbol{e_\varphi}$ 的影响：

$$\frac{\partial \boldsymbol{e_\varphi}}{\partial \varphi} = -\boldsymbol{e_\perp} = -\sin\theta \boldsymbol{e_r} - \cos\theta \boldsymbol{e_\theta} \tag{1.35}$$

同样地，我们有

$$\frac{\partial \boldsymbol{e_\perp}}{\partial \varphi} = \boldsymbol{e_\varphi} \tag{1.36}$$

而这个变化在 $\boldsymbol{e_r}$ 与 $\boldsymbol{e_\theta}$ 方向上的分量就是 φ 的变动对这两个方向的影响。直接写成偏导的形式，有

$$\frac{\partial \boldsymbol{e_r}}{\partial \varphi} = \sin\theta \frac{\partial \boldsymbol{e_\perp}}{\partial \varphi} = \sin\theta \boldsymbol{e_\varphi}, \quad \frac{\partial \boldsymbol{e_\theta}}{\partial \varphi} = \cos\theta \frac{\partial \boldsymbol{e_\perp}}{\partial \varphi} = \cos\theta \boldsymbol{e_\varphi} \tag{1.37}$$

既然已经有了所有的单位矢量偏导运算结果，就可以直接计算出球坐标系下的拉普拉斯算符 $\nabla^2_{球}$:

$$\begin{aligned}
&\left(\boldsymbol{e_r}\frac{\partial}{\partial r} + \boldsymbol{e_\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \boldsymbol{e_\varphi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \varphi}\right)\cdot\left(\boldsymbol{e_r}\frac{\partial}{\partial r} + \boldsymbol{e_\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \boldsymbol{e_\varphi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \varphi}\right)\\
=&\frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial \varphi^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{\cos\theta}{r^2\sin\theta}\frac{\partial}{\partial \theta}\\
=&\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial \varphi^2}
\end{aligned} \tag{1.38}$$

为便于记忆，我们把它写成了最后一行中较为紧凑的形式。

本节讨论的这些有关不同坐标系中算符 ∇ 和算符 ∇^2 的运算，以及后面可能用到的算符 ∇ 的梯度、散度与旋度运算，都可以更为一般性地用**正交曲线坐标系**理论中的**拉梅系数**等公式表示出来。但对于初学者来说，一开始先建立一个直观的几何概念更有助于后续的学习。有兴趣的读者可自行进一步学习。

1.2 最小作用量原理与欧拉-拉格朗日方程

牛顿力学形式的公式能够在人类日常生活的尺度上以极高的精确度描述大自然，因而很快被人们所接受。到了 18 世纪的时候，人们发现可以用其他形式的理论体系描述宏观运动。拉格朗日力学就是这样一种力学体系，其出发点为最小作用量原理。

1.2.1 最小作用量原理

人们发现，宏观运动所遵循的物理规律可以通过最小作用量原理得出。

最小作用量原理：一个物理系统可由主函数（被称为拉格朗日量）$L(q_1, q_2, \cdots, q_s; \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_s; t)$ 描述，其中 q_α 与 $\dot{q}_\alpha$ 分别为广义坐标与广义速度，s 为系统自由度；主函数 L 在时间 $t_{初}$ 与 $t_{末}$ 之间的积分 S 被称为作用量，其中 $t_{初}$ 与 $t_{末}$ 分别为物理过程的初始时刻和末了时刻；真实的物理过程总是使作用量 S 为最小值。

最小作用量原理使我们可以在拥有拉格朗日量（简称拉氏量）L 的情况下获得描述物理过程的运动方程。至于一个物理系统的拉氏量应该是什么样子的，人们只能根据经验以及一些相当一般性的限制写出来[①]。

根据最小作用量原理的表述，作用量

$$S = \int_{t_{初}}^{t_{末}} L(q_1, q_2, \cdots, q_s; \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_s; t)\mathrm{d}t \tag{1.39}$$

对于这样定义的作用量，若 $q_\alpha(t)$ 取不同的函数，则对应的作用量就不同。真实的物理世界具有明显的规律性，因而虽然 $q_\alpha(t)$ 可以有无数种形式，但真实的物理过程总是对应某个特定形式的 $q_\alpha(t)$。最小作用量原理告诉我

① 如何写出一个物理系统的拉格朗日量是所有初学者最大的困扰。目前没有什么原理或程序帮我们做这件事。

们，能使 S 取最小值的 $q_\alpha(t)$ 才是真实的物理。我们想知道的是满足何种限制的 $q_\alpha(t)$ 能使作用量 S 取最小值。

若某一组 $q_\alpha(t)$ 使 S 取得了最小值，那么对 $q_\alpha(t)$ 的任何变动都会使作用量 S 变大。根据最小作用量原理的这个特点，我们可以得到描述运动的基本方程，这个方程被称为欧拉-拉格朗日方程。为了得到运动方程，我们需要研究当广义坐标 $q_\alpha(t)$ 的函数形式变化时作用量的变化情况，因而需要用到被称为变分法的数学工具。

1.2.2 变分法

变分法是一种处理极值问题的一般性方法。变分和微分的区别在于，变分是函数形式的变化，微分是数的变化。我们使用微分时，分析的是一个变量的数值变化对函数值的影响。函数 $f(x)$ 的微分 $\mathrm{d}f$ 等于 $f'(x)\mathrm{d}x$，$\mathrm{d}x$ 是变量的数值变化，$\mathrm{d}f$ 是函数值的数值变化。变分则不同，变分研究的是函数形式的变化。比如一个函数 $F(x(t))$，F 是 x 的函数，而 x 又是 t 的函数。当 $x(t)$ 的函数形式变化时，F 的变化就是变分法要研究的内容。

这里需要注意的是变分法研究的内容不同于中学时学过的复合函数，复合函数研究的仍是数的变化对函数值的影响，只不过函数上多套了一层。再次强调，变分考虑的问题是“函数形式”的变化。所谓的函数形式的变化，例如指的是 $x(t)$ 从 t^2 变化到 $\sin t$，或者从 $(2t-\cos t)$ 变化到 $(t^4-\log t)$ 这种变化。显然，函数的形式有无数种，所以一般来说变分处理的是一个“无穷可能”的问题。无穷可能当然无法简单直接地处理，但人们仍可根据一些特定的限制，利用变分得到有用的结果。

在使用变分时，我们先了解它的基本运算关系。人们一般用 δ 表示变分，如 $\delta x(t)$ 就是 $x(t)$ 的变分，也就是 $x(t)$ 函数形式的变化；正如人们一般用符号 d 表示微分，如 $\mathrm{d}x$ 表示的就是 x 的微分。根据最一般性的要求，变分自然地满足与微分类似的一些计算公式，如：

$$\delta(A+B)=\delta A+\delta B \tag{1.40}$$

$$\delta(AB)=(\delta A)B+A\delta B \tag{1.41}$$

$$\delta\left(\frac{A}{B}\right)=\frac{(\delta A)B-A\delta B}{B^2} \tag{1.42}$$

一个多变量的函数，其变分等于其在各个方向上变分的和，即

$$\delta f=\sum_i \frac{\partial f}{\partial x_i}\delta x_i \tag{1.43}$$

利用这些运算公式，我们就能从最小作用量原理出发，通过变分法得到运动方程。

1.2.3 欧拉-拉格朗日方程

根据最小作用量原理，真实的物理过程对应的是使拉氏量取最小值的路径。路径指的是 $q_\alpha(t)$（$\alpha = 1, 2, \cdots, s$，s 为系统自由度）的具体函数形式。所有 $q_\alpha(t)$ 的函数形式都确定以后，我们就有了一条路径，或者说轨道。此时，我们知道物理系统中所有的物体在任意时刻的位置。最小值点是一个稳定点，物理上的稳定点处对应的物理量变化率为零。一个物理量（某个函数）的变化等于变量的变化乘以变化率，因而在稳定点处的任意偏离造成的物理量的变化都应该是零①。假定 $q_\alpha(t)$ 就是物理路径，那么对它的任何变动 $\delta q_\alpha(t)$ 都会使作用量的变化 δS 等于零。因而，我们也可以将最小作用量原理表示成一个方程：

$$\delta S = 0 \tag{1.44}$$

这个方程具有极强的一般性，我们甚至可以说，所有的基本物理规律都符合这一方程，而不同物理的区别在于拉氏量的不同。

从上面最小作用量原理方程中，我们可以得出一个运动方程。最小作用量原理要求物理规律满足作用量变分为零。对作用量作变分

$$\delta S = \delta \int_{t_{初}}^{t_{末}} L \mathrm{d}t = \int_{t_{初}}^{t_{末}} \sum_{\alpha=1}^{s} \left[\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right] \mathrm{d}t \tag{1.45}$$

这里分别有广义坐标和广义速度的变分，不太容易处理。若能将二者合并到一起，似乎就容易了。只需要利用求导的链式法则，我们就可以做到这点。变分和微分都是普通的线性运算，可交换顺序，即

$$\delta \dot{q}_\alpha = \delta \frac{\mathrm{d}}{\mathrm{d}t} q_\alpha = \frac{\mathrm{d}(\delta q_\alpha)}{\mathrm{d}t} \tag{1.46}$$

这样我们就可以将上面表达式中的第二项利用求导的链式法则改写为

$$\frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha} \frac{\mathrm{d}(\delta q_\alpha)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \right) - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \delta q_\alpha \tag{1.47}$$

将改写后的表达式代回 δS 中，则有

① 稳定点就是一阶导数等于零的极小值点。极大值点对应的一阶导数也为零，但是极大值点可以不是一个稳定点，很小的偏离就可能导致对极大值点的彻底偏离。进一步地，人们通常还根据某个极值是不是最小值点将其分为稳定点和亚稳定点。物理系统总是有向最稳定状态变化的倾向。

$$\begin{aligned}\delta S &= \int_{t_{初}}^{t_{末}} \sum_{\alpha=1}^{s} \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \right) + \left(\frac{\partial L}{\partial q_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \right) \delta q_\alpha \right] \mathrm{d}t \\ &= \sum_{\alpha=1}^{s} \frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \Big|_{t_{初}}^{t_{末}} + \int_{t_{初}}^{t_{末}} \sum_{\alpha=1}^{s} \left[\left(\frac{\partial L}{\partial q_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \right) \delta q_\alpha \right] \mathrm{d}t \end{aligned} \tag{1.48}$$

其中第一项由于已经被写成了全导数的形式，因而被直接积分出来了。我们所讨论的是在给定初末点位置的情况下的运动规律是什么，因而初末点是确定的，即为一个具体的常数。常数不能变动，或者说变动为零，即

$$\delta q_\alpha \Big|_{t_{初}} = \delta q_\alpha \Big|_{t_{末}} = 0 \tag{1.49}$$

因而

$$\sum_{\alpha=1}^{s} \frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \Big|_{t_{初}}^{t_{末}} = 0 \tag{1.50}$$

若要作用量做变分等于零，则有

$$\delta S = \int_{t_{初}}^{t_{末}} \sum_{\alpha=1}^{s} \left[\left(\frac{\partial L}{\partial q_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \right) \delta q_\alpha \right] \mathrm{d}t = 0 \tag{1.51}$$

其中的 δq_α 代表任意的变动。所以，要想整个表达式在任意的变动下都为零，则必须有

$$\frac{\partial L}{\partial q_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) = 0 \tag{1.52}$$

这样就得到了一组（s 个）基本方程，这组方程被称为**欧拉-拉格朗日方程**。基于这一方程的力学体系被称为**拉格朗日力学**。在拉格朗日力学框架下研究问题，就是要求解欧拉-拉格朗日方程，得到广义坐标的具体形式。

1.2.4 拉格朗日量

拉格朗日力学的核心是拉格朗日量（简称拉氏量）。在讨论上面的内容时，我们并没有提到拉氏量怎样写出。拉氏量只能根据经验和猜测写出，这些经验和猜测来自一些一般性的认识和要求。下面，我们通过具体例子来看如何写出一个物理系统的拉氏量。

1. 自由质点的拉氏量

我们先来讨论一个自由的质点的拉氏量应该是什么样子的。自由指的是不受任何外部影响，因而自由质点的运动应该仅考虑一般性的时间与空间性质。

在没有外部影响的情况下，我们相信空间具有**均匀性**且**各向同性**。均匀性指的是在空间中不同位置的物理规律应该都是一样的；各向同性指的是物理规律应该没有方向性，即不存在某一个特殊的方向，在那个方向上物理规律的形式有某种特殊性。除空间的基本性质外，我们还相信，时间也具有均匀性。时间均匀性即物理规律在各个时刻都是一样的，一百年前的物理规律与今天的物理规律没有任何差别。关于时空的这三个要求具有一般性，这些要求帮我们定义了**惯性参考系**，即满足空间均匀性、空间各向同性以及时间均匀性的参考系。

在上述要求下，自由质点的拉氏量的形式受到了强烈的限制。首先，自由质点拉氏量不能是位置与时间的函数，不然在不同位置或不同时刻拉氏量就不一样，因而不具有空间和时间的均匀性。这使得自由质点的拉氏量只能是其速度的函数。但是，速度具有方向性，这将破坏空间各向同性的要求。结合这几点，自由质点的拉氏量只能是速度大小的函数。

若我们把“不同惯性系下的物理规律应该是一样的”当作一个基本要求，则由于物理规律由作用量决定，而作用量是拉氏量的积分，因而两个不同惯性系下的拉氏量要么相同，要么相差一个时间的全导数项。这是因为全导数项可直接积分出来，对作用量的贡献是一个常数，而常数的变分为零，因而不影响运动方程，或者说运动规律。在这种要求下，我们看看自由质点的拉氏量应受到哪些限制。

我们可以用无穷小的方法分析自由质点拉氏量的形式。若有两个惯性系以一个无穷小速度 $\boldsymbol{\varepsilon}$ 相对于彼此运动，则在两个惯性系中，某自由质点的拉氏量分别为 $L(v)$ 和 $L(v')$，其中

$$v' = |\boldsymbol{v}'| = |\boldsymbol{v} + \boldsymbol{\varepsilon}| \tag{1.53}$$

由于绝对值运算不太方便做小量展开，因而我们把它写成等价的形式，即平方开根号：

$$v' = \sqrt{\boldsymbol{v}'^2} = \sqrt{(\boldsymbol{v} + \boldsymbol{\varepsilon})^2} = \sqrt{\boldsymbol{v}^2 + 2\boldsymbol{v} \cdot \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^2} \tag{1.54}$$

将 v' 按小量 ε 展开，得到

$$v' = v + \frac{\boldsymbol{v} \cdot \boldsymbol{\varepsilon}}{v} + \mathcal{O}(\varepsilon^2) \tag{1.55}$$

由于我们要求两个拉氏量相等或相差一个时间的全导数项，因此，通过将 $L(v')$ 展开的方式比较其与 $L(v)$ 的差别。将

$$L(v') = L\left(v + \frac{\boldsymbol{v} \cdot \boldsymbol{\varepsilon}}{v} + \mathcal{O}(\varepsilon^2)\right) \tag{1.56}$$

在 v 处展开，得到

$$L(v') = L(v) + \frac{\mathrm{d}L(v)}{\mathrm{d}v}\frac{\boldsymbol{v}\cdot\boldsymbol{\varepsilon}}{v} + \mathcal{O}(\varepsilon^2) \tag{1.57}$$

这里的 $\mathcal{O}(\varepsilon^2)$ 项既包括了 v' 展开中的 ε^2 及以上阶的项，也包括了 $L(v')$ 的二阶导数及以上阶的项（也是 ε^2 及以上阶）。在忽略高阶小量的情况下，$L(v')$ 与 $L(v)$ 的差别就是

$$\frac{\mathrm{d}L(v)}{\mathrm{d}v}\frac{\boldsymbol{v}\cdot\boldsymbol{\varepsilon}}{v} \tag{1.58}$$

速度 $\boldsymbol{v}$ 本身就是位矢 $\boldsymbol{r}$ 对时间的全导数项 $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$，而 $\boldsymbol{\varepsilon}$ 为任意的无限小常数。因而，若该项整体为时间的全导数项，则 $\boldsymbol{v}\cdot\boldsymbol{\varepsilon}$ 的系数（v 的函数）只能为常数，即

$$\frac{1}{v}\frac{\mathrm{d}L(v)}{\mathrm{d}v} = 2\frac{\mathrm{d}L(v)}{\mathrm{d}v^2} = \text{常数} \tag{1.59}$$

这告诉我们，L 是 v^2 的线性函数。将上式中的常数记为 m，则 L 可被写作

$$L = \frac{1}{2}mv^2 \tag{1.60}$$

待定常数 m 应该取决于所讨论的自由质点本身的性质。

由于我们已经熟知牛顿力学，一眼就看出这就是牛顿力学中的自由质点动能，m 就是自由质点的质量。但是在上面的推导中，我们并未讨论 m 的意义。在拉格朗日力学中，m 只是一个应该取决于质点本身性质的待定常数，其物理意义将通过与具体的实验比对定出。从这里我们可以看出拉格朗日力学的特点，它从抽象的概念出发，得到特定的物理内容，这些物理内容的具体含义将在与实验对比时变得明晰。事实上，所有的物理学理论或所有的科学理论也多如此，即从一个理论出发，得到某些内容，通过实验，明白得到的这些内容的含义。一些物理学理论，如牛顿力学，其用以构建理论本身的概念或名词具有日常生活气息，比如“力”，所以我们很容易理解它。然而我们并不应该预先假定我们已知道了某些物理量，如 m 的意义。物理量的意义是在与实验比对过程中确定下来的，物理理论本身并不一定总需要用实验上含义清楚的物理量来构建。

最后，由于我们总是要选择坐标系来讨论问题，根据不同坐标系中长度微元的形式，可以分别写出不同坐标系中的速度公式，从中可得到自由