

理论力学

2022 年 11 月 3 日

目录

1 推荐教材	1
2 拉格朗日力学	1
3 对称性与守恒量	4
4 哈密顿力学	8
4.1 几何构造	8
4.2 与拉格朗日力学的关系	9
4.3 哈密顿-雅可比方程	10
4.4 路径积分	10

如果说物理专业和非物理专业对物理解区的区别，大概就是从这里开始的。事实上，理论力学是现代物理的基石，它的思想对物理的影响是非常深远的。但是，它并不如一般人听说过的物理那样直观，甚至它的很多含义至今我也不敢保证完全能够理解并讲清楚。本文我还是会按照朗道的方式展开，不会如一些本科教学一样为了妥协学生的理解程度而颠倒主次。

1 推荐教材

- Landau 《Mechanics》不用多说，朗道这本小册子是当之无愧的经典。言简意赅，醍醐灌顶。我当时是用这一本入门的，时至今日我仍能回想起第一次接触理论力学时感受到的震撼。
- Goldstein 《Classical Mechanics》这本书我买来做参考，但是没怎么看过。的确，讲得比朗道全面而且细，对于朗道没怎么提的耗散和场也都有所涉及，讲法似乎也比朗道现代一点。不过，没有朗道那本大胆。

2 拉格朗日力学

理论力学主要由 2 个等价的理论构成：拉格朗日力学与哈密顿力学。

首先,有一些需要理清的共通概念。当我们描述一个系统随时间的变化时,我们往往需要用一些参量来标记系统的状态。如果我们知道了某时刻系统所处的状态,动力学方程会确定地给出系统状态随时间的演化(注意:这在量子力学中也是对的)。比如,描述粒子运动,我们要用位置坐标和速度来标记它的状态;描述电磁场,我们需要知道电势和磁矢势;描述刚体运动,我们需要知道质心的位置与速度,角方向与角速度。这些参量中,有时候一部分只是另一部分对于时间的导数。我们通常将不是导数的那一部分,称作广义坐标(q),它们的导数称作广义速度(\dot{q})。广义坐标的个数(如果有约束的话,减去约束方程的数目。物理常常讨论无约束的方程,但约束在工程中可能会经常遇到),称作自由度。比如,单粒子运动的自由度是3, N 个粒子运动的自由度是 $3N$,刚体运动的自由度是6,经典场运动的自由度是不可数无穷(空间上的每一个点都是一个自由度)。所有可能的 (q, \dot{q}) 集合,称为相空间。

拉格朗日力学最基本的原理,称为最小作用量原理(Least Action Principle)。它假定有一个作用量 Action,它是一个关于运动“轨迹”的泛函: $S[q]$ 。泛函是说,对于给定的广义坐标随时间的关系 $q(t)$,我们能通过特定的方法计算出一个实数 $S[q]$ 与这个轨迹对应。最小作用量原理则是说,真实的运动 $q(t)$ 使得泛函 $S[q]$ 达到它的极小。而关于这个泛函 $S[q]$,我们总是假定它可以写成某个关于坐标的函数对时间(或关于场的函数对时空)的积分:

$$S[q] = \int L(q, \dot{q}, t) dt \quad (1)$$

或

$$S[f] = \int \mathcal{L}(f, \partial_i f, x_i) d^4x \quad (2)$$

其中,函数 $L(q, \dot{q}, t)$ 称为拉格朗日量, $\mathcal{L}(f, \partial_i f, x_i)$ 称为拉格朗日密度。

那么拉格朗日量/密度是什么?这并没有特别的规定。一般对于拉格朗日量只有几点要求:

- 要保证作用量 S 是一个标量,即无论坐标系如何旋转,甚至洛伦茨变换都不应该改变它。这意味着在牛顿力学中拉格朗日量也是标量。而在相对论中,因为时间也不是标量,所以拉格朗日量应当适当选取来保证作用量是标量。
- 因为作用量是标量,系统总的作用量应该等于各部分之和,因此总的拉格朗日量也等于各部分的和。
- 因为我们关心的是作用量的函数形式本身,作用量可以相差一个常数(就好像势能零点可以随意取一样),因而拉格朗日量可以相差一个函数对时间的全微分。
- 拉格朗日量只与一阶导数有关,因为大量实践告诉我们知道位置和速度就已经足够描述所有运动了。除此以外,拉格朗日量是任意给定的。

嗯,一般人看到这里应该早就晕了,因为这和从高中开始接触的牛顿力学完全不沾边,并不直观。其实可以这样理解,既然要用数学来处理现实问题,怎样将想要研究的东西抽象出来,与相应的数学结构相联系,是很重要的问题。牛顿力学实际上是非常直接的做法,即将粒子的运动状态抽象为坐标和速度之后,直接去寻找运动时这些量满足的规律。而理论力学,可以说是力学的力学,它先不管坐标和速度满足什么样的规律,而是将一种规律和函数空间中的一个函数(拉格朗日量)对应起来,使得我们可以用同一的框架去研究不同规律,研究不同的力学。所以实际上,有了最小作用量原理之后,只要给定一个拉格朗日量,就产生出一种力学。牛顿力学只不过是拉格朗日力学中的一个例子。对于接触过C++的同学,可以这样理解:

```

class 拉格朗日力学
{
public:
    double Action(double InitialTime, double FinalTime)
    {
        return Integrate(Lagrangian, InitialTime, FinalTime);
    }
    virtual double Lagrangian(vector<double> Coordinates, vector<double> Velocity,
        double Time);
};

class 牛顿力学 : public 拉格朗日力学
{
public:
    double Lagrangian(vector<double> Coordinates, vector<double> Velocity, double Time)
    {
        return Mass*DotProduction(Velocity,Velocity)/2;
    }
};

```

嗯，就是说拉格朗日力学是一个抽象类，当给定了具体的描述方式和拉格朗日量后，这个派生类才是一个力学（比如牛顿力学）。

那么我们怎么得到具体的力学的运动方程呢？用变分法。思想还是和微积分求极值很像：给轨迹一个扰动，相应拉格朗日量和作用量都有可能改变。如果对于这个扰动作用量的改变等于 0，那么意味着在这个轨迹附近作用量达到了极大或极小，这和函数出现一个峰值或谷值的情形是类似的。结论就是欧拉-拉格朗日方程：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \quad (3)$$

$$\sum_i \frac{\partial}{\partial x_i} \frac{\partial \mathcal{L}}{\partial (\partial_i f)} - \frac{\partial \mathcal{L}}{\partial f} = 0 \quad (4)$$

不过我们不知道按这个方程给出的是极小还是极大。有趣的是虽然最小作用量原理中要求极小，但实际上一般书籍中都很少去关心是不是真的极小，而把欧拉-拉格朗日方程给出的轨迹就直接当做运动方程。这里我没有仔细考虑过这个问题，只是如果想知道欧拉-拉格朗日方程给出的是极大还是极小，只需要再做一次变分，考察二阶变分的符号即可。这和判断函数的极大或极小是类似的。

作为例子，朗道给了一种“推导”牛顿力学的方法，可以感受一下物理中怎么凑一个拉格朗日量：考虑一个自由运动的粒子，因为参考系的具体选取不影响粒子的运动，所以拉格朗日量不能含坐标本身，只能包含速度。而因为空间各个方向没什么区别，所以粒子运动只能与速度的大小或 v^2 有关。又因为拉格朗日量不含坐标， $\partial L/\partial x = 0$ ，根据欧拉-拉格朗日方程， $\partial L/\partial v$ 是一个常数，由 $L = L(v^2)$ ， $\partial L/\partial v$ 后依然是 v 的函数，这意味着 v 是一个常数，即自由运动的粒子只做匀速直线运动。

那么考虑到伽利略相对性原理，换到另一个速度 ϵ 很小很小的参考系，那么粒子的速度变成 $v' =$

$\mathbf{v} - \boldsymbol{\epsilon}$ ，取一阶近似拉格朗日量变成：

$$L(\mathbf{v}') \approx L(\mathbf{v}^2) - 2 \frac{dL(\mathbf{v}^2)}{d(\mathbf{v}^2)} \mathbf{v} \cdot \boldsymbol{\epsilon} \quad (5)$$

由于相对性原理要求左右两个拉格朗日量没有本质区别，右边的项必须为对时间的全微分。而 \mathbf{v} 本身已经是对时间的全微分了，那么系数必须是一个常数，定义为质量 m ，那么有：

$$L = \frac{1}{2} m v^2 \quad (6)$$

如果外界有势场（通常只与坐标有关），那么势的贡献可以直接加在拉格朗日量后面，那么假设这一项为 $-V(x)$ （这里负号是为了方便），把

$$L = \frac{1}{2} m v^2 - V(x) \quad (7)$$

带入欧拉-拉格朗日方程，我们就得到了：

$$m \frac{dv}{dt} = - \frac{dV(x)}{dx} = F \quad (8)$$

这正是牛顿第二定律。由此也可以看出之前定义的质量 m 与牛顿力学中所述的质量是一致的。

实际上，相对论、麦克斯韦方程都可以放在这个框架之内。比如相对论的拉格朗日量是：

$$L = -mc^2 \quad (9)$$

只是这个积分是要对本征时 τ 而非坐标时间 t 进行。对于电磁场（重复指标表示求和）：

$$\mathcal{L} = -F^{ij} F_{ij} \quad (10)$$

其中 $F_{ij} = \partial_i A_j - \partial_j A_i$ ，其实是一个由电场强度与磁感应强度组成的矩阵。它给出的就是麦克斯韦方程的一半（另一半自然满足）。

3 对称性与守恒量

说完拉格朗日力学，一个自然而然的问题就是我们为什么需要一个研究力学的力学？想想写程序为什么需要虚基类？是为了统一接口，统一接口之后我们可以用统一的手段去处理各种不同的派生类，带来极大的方便。值得提醒的是，对于物理，普遍化的处理问题的一般方式往往都要比依赖技巧的个例重要。因而拉格朗日力学的意义之一其实也是统一接口。当有了拉格朗日量后，我们可以用同样的量子化方法过度到量子理论。而且，有一些脱离具体力学形式的结论，可以从拉格朗日力学中直接得出。最经典的一定会被讲到的结论，就是对称性与守恒量的关系。

这个关系往往被冠以诺特定理的名字被提起。一个不算严格的表述是：

- 如果有一个拉格朗日量呈现一种连续（可微）的对称性，必有一个守恒量与之对应。

这个定理背后涉及的比较复杂的数学，我无力在此深入展开。比如对可微对称性的严格定义需要李群的概念，而定理最一般的形式需要用到流形与纤维丛的概念。这些数学我并不熟悉。

对于对称这个概念，如果我们考察它的实质，实际上是一种不变性。严格来说，如果系统在某个变换操作下不变，我们称为它有一种对称性。比如，一个左右对称的图形是指将图形沿对称轴反射后，图

形仍然不变。再比如旋转对称性是指如果将图形旋转一定角度，图形仍然不变。这里反射与旋转都是对称操作。之所以会涉及群论，是因为任意的两次对称操作联合起来实施，依然是一个对称操作，这种结构与群是一致的，所以我们可以用群来描述对称操作。而如果对称操作是可微的（粗略地讲，存在“无穷小”的对称操作），则可以使用李群来描述。

那么对物理而言，最为重要的对称操作，有 3 类：

- 时间平移对称：将系统的时间整体平移，不改变系统的性质。换句话说，系统的运动性质不随时间的变化而变化。
- 空间平移对称：将系统的位置整体平移，不改变系统的性质。实质上是说空间的不同位置是没有分别的。
- 空间旋转对称：将系统旋转任意角度，不改变系统的性质。实质上是说空间的各个方向是没有分别的。

而这三个对称操作带来的守恒量，称为能量、动量与角动量。

时间平移对称意味着表征系统运动性质的拉氏量 $L(q, \dot{q}, t)$ 对于变换 $t \rightarrow t + \delta t$ 不变，即：

$$L(q, \dot{q}, t + \delta t) = L(q, \dot{q}, t) \frac{\partial L}{\partial t} = 0 \quad (11)$$

那么考虑全微分：

$$\frac{dL}{dt} = \frac{dq}{dt} \frac{\partial L}{\partial q} + \frac{d\dot{q}}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial t} = \dot{q} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{d\dot{q}}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{d}{dt} \left(\dot{q} \frac{\partial L}{\partial \dot{q}} \right) \quad (12)$$

因为我们只考虑真实的物理过程，中间步骤用到了欧拉-拉格朗日方程。上式很明显可以对 t 进行积分，积分之后得到一个常数，这个常数在系统运动的过程中都不会改变，因此是一个守恒量。这个守恒量就称为能量：

$$E = \dot{q} \frac{\partial L}{\partial \dot{q}} - L \quad (13)$$

类似的，空间平移对称意味着拉氏量 $L(q, \dot{q}, t)$ 不含坐标 q ，于是由于欧拉-拉格朗日方程：

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q} = 0 \quad (14)$$

也可以进行一次积分，得到一个常量：

$$p = \frac{\partial L}{\partial \dot{q}} \quad (15)$$

这个常量称为动量。

角动量的推导稍微繁琐一点，首先能量和动量的定义是对于任何类型的系统都可以成立的，但角动量只针对三维空间中的运动。广义坐标要取为空间坐标本身。推导如果引入一个更为广泛的定理，会稍微容易一些（参见 Dubrovin 《现代几何学》 第一卷）：

- 如果一个拉氏量在某种单参数变换群下不变，那么如果表征这个变换群的向量场是 X^i ，那么动量 p_i 沿向量场的“分量”是守恒的。

我们直接用转动作为例子。对于坐标系沿 z 轴的转动，可以写成矩阵形式：

$$S_z(\theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (16)$$

这正是一个只有单个参数 θ 表征的变换群。因为转动后坐标变为：

$$r'(\theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \cos \theta + y \sin \theta \\ -x \sin \theta + y \cos \theta \\ z \end{pmatrix} \quad (17)$$

对 θ 微分：

$$\frac{dr'(\theta)}{d\theta} = \begin{pmatrix} -x \sin \theta + y \cos \theta \\ -x \cos \theta - y \sin \theta \\ 0 \end{pmatrix} = - \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} r'(\theta) \quad (18)$$

换句话说，这个变换可以看做是微分方程：

$$\frac{dr}{d\theta} = -Xr \quad (19)$$

的解，变换可以形式地写成 $r = \exp(-X\theta)$ 。而定义一个向量场 $f^i = Xr = (-y, x, 0)^T$ 足以表征这个变换。这个向量场就是我们需要的与转动相联系的向量场。那么根据定理，我们有守恒量：

$$M_z = f^i p_i = xp_y - yp_x \quad (20)$$

称为 z 方向的角动量。另外两个方向同理。综合起来，可以用叉乘将角动量写成：

$$\mathbf{M} = \mathbf{r} \times \mathbf{p} \quad (21)$$

对于能量与动量守恒，还有一点可以注意一下。以前与人讨论时曾有一个疑问，为何从牛顿第二定律 3 个方程，可以推出能量守恒、动量守恒和角动量守恒总共 7 个方程。实际上，因为牛顿是一个二阶微分方程，在求解时需要进行 2 次积分，对于 N 个方程，则需要引入 $2N$ 个积分常数。而推导能量与动量，实际上对运动方程做了一次积分，所以能量和动量其实是包含在这 $2N$ 个积分常数之中。而角动量的定义实际上线性依赖于动量，并不独立。也因此，能量、动量也称为运动方程的初积分。

以上是针对粒子系统而言的。对于场的拉格朗日密度，用类似的方式得到的与时空平移对称有关的守恒量是一个张量，称为能量-动量张量（同样，重复指标表示求和）：

$$T_{ij} = g_{jl} \partial_i f^k \frac{\partial \mathcal{L}}{\partial (\partial_l f^k)} - g_{ij} \mathcal{L} \quad (22)$$

它的结构和前面能量的定义是一致的。这个式子刻意写成了通用的形式，对于非相对论性的定义于三维欧式空间的场（比如连续介质的形变），度规 g_{ij} 就只是普通的单位矩阵 δ_{ij} ，这个张量实际上是动量张量。对于更一般的场（比如电磁场），则定义于 4 维时空之中，度规 g_{ij} 是闵可夫斯基空间的度规 $\text{diag}\{-1, 1, 1, 1\}$ 或 $\text{diag}\{1, -1, -1, -1\}$ 。这个式子算出的能量动量张量是没有对角化的。因为同样的能量动量张量之间可以相差一个张量的散度，所以一般使用时我们倾向于利用这个性质将它对角化，会有很多方便。

能量动量张量另一个重要的地方，是它与时空的弯曲相关。著名的爱因斯坦场方程表达的就是时空的曲率与能量动量张量的关系：

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \quad (23)$$

作为参考，如果从上一篇所给的几个拉式量出发，我们可以推出牛顿力学中的能量、动量：

$$L = \frac{1}{2}mv^2 - V(x) \quad (24)$$

$$p = \frac{\partial L}{\partial v} = mv \quad (25)$$

$$E = v \frac{\partial L}{\partial v} - L = vp - L = \frac{1}{2}mv^2 + V(x) \quad (26)$$

对于狭义相对论：

$$S = - \int mc^2 d\tau = - \int mc^2 \sqrt{1 - \frac{v^2}{c^2}} dt = \frac{\partial L}{\partial v} = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (27)$$

$$E = vp - L = \frac{mv^2}{\sqrt{1 - \frac{v^2}{c^2}}} + mc^2 \sqrt{1 - \frac{v^2}{c^2}} = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (28)$$

（这是用坐标时间定义的普通速度表示的，如果用本征时定义 4-速度 $u^i = dx^i/d\tau$ ，会更为简洁。注意到 $\tau = t\sqrt{1 - v^2/c^2}$ ， $u^i = v^i dt/d\tau$ $p^i = mu^i$ ， $E = mc^2 u^0$ 。）对于电磁场，能量动量张量为：

$$T^{ik} = \frac{1}{4\pi c} (-F^{il} F_l^k + \frac{1}{4} g^{ik} F_{lm} F^{lm}) \quad (29)$$

等等。

一个更有趣的例子是平方反比力场中的粒子运动（开普勒问题）：

$$L = \frac{1}{2}mv^2 + \frac{k}{r} \quad (30)$$

它除了通常的能量守恒和对于原点的角动量守恒外，还隐藏着一个额外的守恒量：拉普拉斯-隆格-楞次矢量：

$$\mathbf{A} = \mathbf{p} \times \mathbf{M} - mk \frac{\mathbf{r}}{r} \quad (31)$$

因为 $\mathbf{p} \times \mathbf{M} = \mathbf{p} \times (\mathbf{r} \times \mathbf{p}) = m^2 \mathbf{v} \times (\mathbf{r} \times \mathbf{v}) = m^2 v^2 \mathbf{r} - m^2 (\mathbf{v} \cdot \mathbf{r}) \mathbf{v}$ ，这个矢量始终位于位置矢量 \mathbf{r} 与速度 \mathbf{v} 所张成的平面上。另外由这个矢量可以看出粒子的运动必须是圆锥曲线。然而，造成这个守恒量的对称操作并不显然，它源于与开普勒问题等价的一个 4 维空间中的 3 维球面上的运动问题。详细的细节可以参考 wiki: Laplace-Runge-Lenz vector。

4 哈密顿力学

4.1 几何构造

哈密顿力学是和拉格朗日力学等价的力学体系。然而，一般的书中介绍哈密顿力学时，总是讲它是如何从拉格朗日力学中“推导”过来的。但既然等价，这种推导应该是不必要的，应该有其他更基本的构建哈密顿力学的方式。这种方式大概涉及到相空间的几何性质，这我可能无法写得很清楚。

经验中确定一个自由度为 N 的系统的运动状态需要知道 $2N$ 个参数，其中 N 个取为广义坐标 q 。那么另外 N 个是什么，拉格朗日力学中取为广义坐标对时间的导数——广义速度，而哈密顿力学则取为 N 个我们也不知道是什么东西的参数，叫做广义动量 p 。回忆一下相空间的概念，相空间是系统所有运动状态的集合，那么对于自由度为 N 的系统，相空间的维度是 $2N$ 。既然我们选用广义坐标 q 和广义动量 p 作为系统状态的标记，不妨记 $x = (q, p) = (q_1, q_2, \dots, q_N, p_1, p_2, \dots, p_N)$ 为相空间上的点。我们假设相空间上的度规是这样的简单形式（参见 Dubrovin 《现代几何学》第一卷）

$$g = \begin{pmatrix} 0_N & I_N \\ -I_N & 0_N \end{pmatrix} \quad (32)$$

那么哈密顿力学相当于假设系统是一个梯度系统，其运动方程为微分方程：

$$\frac{dx}{dt} = \nabla H(q, p, t) \quad (33)$$

方程中的“势场” $H(q, p, t)$ 称为哈密顿量。而回忆梯度的定义：

$$(\nabla H)_i = \sum_j g_{ij} \frac{\partial H}{\partial x_j} \quad (34)$$

这个方程给出的就是正则方程（英文 Canonical 其实是正统的意思……物理学家在这种地方总能给人一种中二的感觉）：

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad (35)$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (36)$$

那么对于任何一个函数 $f(p, q, t)$ ，它随时间的变化其实是：

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial q_i} \frac{dq_i}{dt} + \sum_i \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \sum_i \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} = \frac{\partial f}{\partial t} + \langle \nabla F, \nabla H \rangle \quad (37)$$

最后一项的 $\langle \cdot, \cdot \rangle$ 是内积符号，在给定度规下，内积写作：

$$\langle \nabla F, \nabla H \rangle = \sum_{i,j} g_{ij} \frac{\partial F}{\partial x_i} \frac{\partial H}{\partial x_j} \quad (38)$$

那么定义泊松括号：

$$\{A, B\} = \langle \nabla A, \nabla B \rangle \quad (39)$$

任何函数随时间的变化都可以写成：

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\} \quad (40)$$

如果函数 f 不显含时间，那么有：

$$\frac{df}{dt} = \{f, H\} \quad (41)$$

特别的，对于哈密顿量本身，直接计算容易验证

$$\{H, H\} = 0 \quad (42)$$

所以

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \quad (43)$$

只要哈密顿量不显含时间，那么哈密顿量不随时间变化。

4.2 与拉格朗日力学的关系

哈密顿力学形式的构建到这里应该就完整了，但是这样我们看不到任何与具体问题的联系。这种联系我们还是需要从与拉格朗日力学的等价关系中寻找。首先，从上面看来，显然哈密顿量和拉格朗日量类似，也是足以表征系统运动性质的函数。因此当哈密顿量不显含时间的时候，也表明时间不影响系统运动的规律，系统具有时间平移的对称性。在拉格朗日力学中，这个对称性带来能量守恒量：

$$E = \dot{q} \frac{\partial L}{\partial \dot{q}} - L \quad (44)$$

而在哈密顿力学中，这意味着哈密顿量不随时间变化，也就是说，哈密顿量是此时的守恒量。那么，最简单的情况下，等价性意味着哈密顿量正是能量：

$$H = E = \dot{q} \frac{\partial L}{\partial \dot{q}} - L \quad (45)$$

如果再注意到拉格朗日力学中，动量 $p = \frac{\partial L}{\partial \dot{q}}$ ，那么等式的右边正好像是一个勒让德变换的形式：

$$H(q, p, t) = \dot{q}p - L(q, \dot{q}, t) \quad (46)$$

(严格来说，勒让德变换应该是 $H(q, p, t) = \sup_{\dot{q}}(p\dot{q} - L(q, \dot{q}, t))$ 。要求 $\frac{\partial^2 L}{\partial \dot{q}^2} > 0$ 一般都能满足，而对于给定的 p ，这最大值就取在 $p = \frac{\partial L}{\partial \dot{q}}$ 的位置，所以这还可以算是个勒让德变换。)

而勒让德变换，我们知道，正是将一组自变量中的函数与另一组自变量中的函数相联系的方式。而 (q, p) 坐标和 (q, \dot{q}) 坐标描述的是同一个空间， $p = \frac{\partial L}{\partial \dot{q}}$ 正好提供了一个将两组坐标相联系的变换。在这个变换下，哈密顿量和拉格朗日量通过勒让德变换相联系，这样我们就算是明确了两套力学之间的关系。

如果从这个勒让德变换出发，我们也可以从欧拉-拉格朗日方程推导出正则方程。这个计算还是很简单。这里做一些注记：

- 同样是运动方程，欧拉-拉格朗日方程是 N 个二阶常微分方程组，而正则方程是 $2N$ 个一阶常微分方程组。解这些方程组时都会引入 $2N$ 个常数，它们确定了系统的初始状态。这些地方是一样的。

- 同拉格朗日量类似，哈密顿量虽然等于能量，但它的绝对数值并没有意义，有意义的仍然是它的函数形式。因此，哈密顿量也可以相差一个常数而没有任何实际区别。
- 不是所有的系统的哈密顿量与拉格朗日量都有解析的表达式。比如，对于化学反应系统，它的哈密顿量有解析表达式： $H(q, p) = \sum_j a_j(q)(e^{p \cdot v_j} - 1)$ ，但它的拉格朗日量只是数值可算，没有解析的表达式。这个哈密顿量中， j 是化学反应的编号，向量 q 是系统中各种分子的个数， v_j 是反应 j 的状态改变矢量， $a_j(q)$ 是当系统的状态为 q 时，第 j 个反应发生的概率。

4.3 哈密顿-雅可比方程

当然，哈密顿力学远远不只有这么点点内容。如果给定系统的初态，将已经极小化的作用量当做末态的函数，考虑哈密顿量与之的关系，有（嗯，这里记号和概念会有点乱，注意一下积分符号里面的 q, p 是过程的函数，而积分符号外面的 q, p 是末态的坐标。）：

$$S(q, t) = \int_{t_0}^t L dt = \int_{t_0}^t (p\dot{q} - H(q, p, t)) dt \quad (47)$$

所以：

$$\frac{dS}{dt} = L \quad (48)$$

对路径做变分，考虑欧拉-拉格朗日方程：

$$\delta S = \int_{t_0}^t \delta L dt = \left[\frac{\partial L}{\partial q} \delta q \right]_{t_0}^t = \frac{\partial L}{\partial q} \delta q(t) = p \delta q \quad (49)$$

得到：

$$\frac{\partial S}{\partial q} = p \quad (50)$$

那么由上面几式，

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial q} \dot{q} = \frac{\partial S}{\partial t} + p\dot{q} = L \quad (51)$$

$$\frac{\partial S}{\partial t} = L - p\dot{q} = -H \quad (52)$$

这样我们可以进一步用作用量 S 来代替动量 p ，得到的也是一个运动方程，叫做哈密顿-雅可比方程：

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0 \quad (53)$$

关于相空间中的几何还有众多 topic，比如表明相空间中演化中区域体积不变的刘维尔定理，保持正则方程不变的正则变换等等。因为我也没有完全整理好它们的 significance，这里就不详述了。

4.4 路径积分

在将经典力学过渡到量子力学时，相空间的坐标不再是实数，而要变成算符，系统的状态改由希尔伯特空间中的矢量——态来描述。由于算符的对易子所属的李代数与泊松括号的李代数同构，可以将经典力学中的方程翻译成量子力学。将经典的哈密顿量中的函数 q, p 改成算符 \hat{q}, \hat{p} ，得到的就是量子力学的哈密顿量；将泊松括号改成对易子，得到的就是量子力学的运动方程——海森堡方程：

$$\frac{d\hat{F}}{dt} = \frac{\partial \hat{F}}{\partial t} + \frac{1}{i\hbar} [\hat{F}, \hat{H}] \quad (54)$$

更深层次的关系，则要通过路径积分来理解：一个系统从状态 $|a\rangle$ 演化到 $|b\rangle$ 的概率，写成对中间状态 $|a'\rangle$ 的积分，等于：

$$\langle a|b\rangle = \int \langle a|a'\rangle \langle a'|b\rangle da' \quad (55)$$

如果我们加入很多中间状态：

$$\langle a|b\rangle = \int \langle a|a_1\rangle \langle a_1|a_2\rangle \langle a_2|a_3\rangle \cdots \langle a_n|b\rangle da_1 da_2 \cdots da_n \quad (56)$$

并将 $n \rightarrow \infty$ ，得到的相当于是一个对于路径的积分。注意这里有测度的问题，因为一般的无穷重积分是没有定义的。严格定义路径积分需要引入维纳测度，细节我就不清楚了。而对于任意接近的两个无限接近的状态，我们假设概率为：

$$\langle a_i|a_{i+1}\rangle = \exp\left(\frac{iL}{\hbar} dt\right) \quad (57)$$

那么总的概率可以写成：

$$\langle a|b\rangle = \int \prod_{\text{Path}} \exp\left(\frac{iL}{\hbar} dt\right) Dx = \int \exp\left(\frac{i}{\hbar} \int L(x(t), \dot{x}(t), t) dt\right) Dx = \int \exp\left(\frac{i}{\hbar} S[x]\right) Dx \quad (58)$$

这里 Dx 表示积分是对于所有可能的路径积分。那么以这个积分为出发点，结合经典力学，也可以完全构建出量子力学。从这里也可以粗糙地看出量子力学如何过渡到经典力学：如果作用量 S 比起 \hbar 大太多，对上面的积分有贡献的基本只有经典的路径，因为经典路径对应于 S 极小的路径。而偏离经典路径越远， S 越大，附近路径的 $\exp\left(\frac{i}{\hbar} S[x]\right)$ （注意其实是三角函数）震荡的频率也越快，求和后贡献基本相消，于是实际上有贡献的只有经典路径及其周围很窄的范围，因而此时粒子运动只走确定的路径。