欧拉和力学的变分原理*

В В Румянцев

欧拉（Leonhard Euler 1707—1783）在经典力学发展中的历史功绩，在于将数学分析方法应用于力学问题。因此，欧拉理应被认为是分析力学的奠基人之一。分析力学这门科学是以少数假定为基础，由这些假定出发，导出物体的平衡和运动的全部规律。目前，力学的变分原理是分析力学的基础，也就是数学上用变分关系的形式表示的基本的、原始的规则，力学的全部运动微分方程和定律都是作为变分原理的逻辑结果。各种变分原理，无论在形式上及变分手段上，还是在共同性上，都彼此不同，但每一个原理在其可应用范围内都是一个基础，并且好像综合了相应物质系统的整个力学。换言之，力学的某种变分原理潜在地包含了这门科学领域的全部内容，并将其所有规则统一到一个提法中。

众所周知，经典力学基于对自由物体建立的牛顿（Newton）定律和约束公理。由这些定律和公理出发可证明变分原理的正确性。同样地，任何一个变分原理都可作为公理，并由它逻辑地导出力学定律。

变分原理在形式上可分成表征任何给定时刻运动性质的微分原理和表征任何有限时间间隔运动性质的积分原理。欧拉主张以基本微分原理之一，即现今知道的达朗伯-拉格朗日（D'Alembert-Lagrange）原理作为起源，并首先在数学上给出了最小作用量积分原理的严格的富有成效的表述，其更一般的研究由拉格朗日和雅可比（Jacobi）继续和完成。

力学的变分原理促进了与其相关的强有力的数学形式体系，拉格朗日和哈密顿（Hamilton）形式体系的发展，它们与力学中的变换问题，李（Lie）群，守恒律及其他基本问题密切相联系。由于众多学者的努力，分析力学已享有所有自然科学思想发展方面优美、明确与首尾一贯的典范的声誉。

1 以前历史发展简况

很早以前就有了这样的模糊概念：自然界有可能存在具有最大或最小性质的某些普遍规律。古代哲学家在思考自然现象时就已经认为："自然界什么也不无根据地去做，并在所有自身表现方面选择最短的或最容易的路线，在这原则下他们确定了主要的终极原因，自然界趋向于这个终极原因……。亚里士多德（Aristotle）就经常提到这一信条，但是看来与其说他是从他前辈那里获得的，还不如说是他自己想出的。后来这个思想在各哲学学派中达到了如此之巩固的地位，以致由它建立了第一条哲学法规，最后，直到笛卡儿（Descartes）才试图作出他的反驳"[1]。

但是，很长时间以来谁也没有去确定"那个真正的客体，但自然界不仅在某些表现上，

* 据作者赠送给译者的抽印本译出。

而且完全在自身一切表现上总是趋向于最大限度地减小这个客体”．的确，在某些特别情形下出现“好像这个普遍原理蒙上了某种阴影．其中首先值得提到光的反射，关于这点托勒玫（Ptolemacus）早就说过，反射角总是等于入射角，并指出了这样光路最短，因此，如果反射不是这样，就会描出较长的光路．但同时要注意，对于光线的折射，不管用什么方式这种解释都不成立，在这里曲折的线不管怎样一般都不是最短路线”[1]．

科学史上第一个变分原理是几何光学中的最短时间原理，由法国学者费马（Fermat）于1662年提出．费马认为“自然界以最容易的可允许的方法起作用”[2]．他根据对以前由斯涅耳（Snell）发现的光的折射定律的研究而提出了这个原理，并证明了折射定律满足时间最短原理．但现在已证明，费马原理仅在惠更斯（Huygens）[3]的工作中得到，其中指明，光线在两点间经历的时间最短．

费马原理

$$\int_P^Q \frac{ds}{v} = \int_P^Q \mu ds = \min \tag{1}$$

乃是表示几何光学定律的最普遍的数学公式．这里 $\mu(x,y,z)$ 是光学上非均匀但各向同性介质（光在其中由点 P 到 Q 以速度 v 传播）的折射指数．

费马假设光在较疏介质中有较大速度，而在较密介质中有较小速度．笛卡儿则持相反观点．莱布尼兹（Leibniz）也试图屏弃费马的解释．莱布尼兹认为，自然界选择最容易的路线，并提出光线克服困难的概念，它由光通过任何介质时的路线长度与阻力的乘积来确定．他认为，光总是应当沿所有困难之和为最小的路线行进．莱布尼兹与笛卡儿一样，认为光线在较密介质中有较大速度．但是，莱布尼兹并没有将自己的最容易路线原理应用于任何其他情形．

力学中的作用的第一个定义由莱布尼兹提出（1669）．“运动的形式上的作用比例于质量、这些物质经历的距离与速度的乘积”[4]，即 mvs 或 $mv^2\Delta t$．

1687年牛顿发表了名著《自然哲学的数学原理》[5]，由此奠定了经典力学的基础．按拉格朗日的说法，“力学在牛顿手中变成了新科学；他的《原理》促成了这个转变的时代”[6]．我们注意到，在牛顿的《原理》一书里的许多问题中间，也有在“稀疏”介质中沿轴运动的转动物体所受的阻力为最小的形式问题．这个问题看来是力学中的第一个变分问题．最速降线问题，即最快降落曲线问题就更为著名，它于1696年由约·伯努利（Johann Bernoulli）[7]提出，而由他本人以及莱布尼兹，牛顿，雅·伯努利（Jacob Bernoulli），洛比达（L'Hospital）等人予以解决；他们都发现，速降线是悬链线．约·伯努利证明了，这条曲线也是光的传播轨道．他写道：“我因此同时解决了两个重大问题——一个是光学的，另一个是力学的……我证明了，尽管这两个问题来自完全不同的科学领域，但它们却具有同一性质”．然后写道：“自然界总是以最简单的方式起作用，如同在我们这里的情形，它借助于同一条线便为两门不同的科学效力”[8]．

实际上，这是非常特殊情况下光学-力学比拟的第一个例子．最速降线问题被认为是变分原理发生的起源，其奠基人之一就是约·伯努利的学生、年轻的欧拉．

2　欧拉在变分学和分析力学方面的某些工作

欧拉最初的大量工作花费在力学上，1736年在彼得堡发表了他的两大卷《力学，或者说

用分析描述运动的科学》．在序言中他写道："如果说分析不管在什么地方都是必要的，那么这尤其与力学相关．尽管读者会确信所提这些命题的正确性，但他并未对它们得到充分清楚、确切的理解，因此，如果稍微改变这些问题本身，那么他未必能够独立地加以解决，如果不使用分析，那么这同样的一些命题也不能用分析方法解决。恰恰当我开始熟悉牛顿的《原理》和赫曼（Hermann）的《运动论（Phoronomy）》时，我也遇到过这种情形；尽管在我看来，我好像已足够清楚地了解了许多问题的解，但是我仍不能解决稍微偏离那些问题一点点的问题．就是从那时我试图就我所能做的，从这种综合方法中将分析分离出来，对那些特别有益的命题用分析来予以研究；正由于此，我非常透彻地搞清了问题的实质．然后，我同样地研究了其他属于这门科学的、分散在许多地方的工作，我亲自给自己有计划地用同样一些方法叙述了它们，并且写成了方便的有条理的形式．在这些课题上我不仅遇到了一系列以前没有解决过的问题，现在我成功地解决了，而且还带来许多新的方法，由于这些方法，不仅力学而且看来分析本身也都极大地丰富了"[9]．

在这个工作中欧拉以合理的形式建立了点的运动方程并系统地利用它们解决了许多问题．后来，拉格朗日在评论这个工作时指出，欧拉1736年的《力学》"应该认为是对运动的研究应用了分析的第一个伟大的工作"；这个工作全部建立在公式上，这些公式描述运动时是将力分解为切向分量和法向分量[6]．

为解一系列振动理论问题，欧拉（1740）将雅·伯努利-赫曼原理推广应用于复摆振动问题，即"在其作用下的推动力应当是构成摆的重物，为使有可能一起运动，这些推动力等价于在重力作用下所得到的那些力；因此，如果它们的方向相反，则前者与后者应处于平衡"[6]．这个原理是拉格朗日形式的达朗伯普遍原理的前奏．

在此期间欧拉还研究了等周问题，即在给定长度的所有简单闭曲线中确定一条线使之保持最大面积的问题，并发表了一系列文章，这些文章奠定了数学新领域——变分学的基础．1744年欧拉发表了变分学的著名论文[10]，其中特别给出了定积分

$$J=\int_a^b F(x,y,y')dx,\qquad y(a)=\alpha,\ y(b)=\beta \tag{2}$$

取极值的必要条件．

为了导出上述条件，欧拉将区间$[a,b]$分成带横坐标$x_0=a,x_1,x_2,\cdots,x_n,x_{n+1}=b$的相等小区间，用带纵坐标$y_s=f(x_s)$的折线代替未知曲线$y=f(x)$，用有限差之比$f'(x_s)=(y_{s+1}-y_s)/(x_{s+1}-x_s)$代替导数$f'(x_s)$，用有限和

$$S'=\sum_{s=0}^{n}F(x_s,y_{s+1},y'_s)(x_{s+1}-x_s)$$

代替积分（2）．

为了确定函数S'的极值，取其对y_{s+1}的导数并让它等于零，得到方程

$$\left[\frac{\partial F}{\partial y}-\frac{\Delta}{\Delta x}\left(\frac{\partial F}{\partial y'}\right)\right]\Big|_{x=x_s}=0\qquad (s=0,1,\cdots,n,n+1)$$

且$\Delta x=x_{s+1}-x_s$．在$\Delta x\to 0$的极限下，得到欧拉方程

$$\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'}=0\qquad (a\leqslant x\leqslant b) \tag{3}$$

由欧拉给出的对方程（3）的推导并不是十分严格的，因为包含毫无根据的极限过程。后来拉格朗日给出了方程（3）的严格推导。

3 最小作用量原理

最小作用量原理首先由莫佩蒂（Maupertuis）于1744年提出．他研究光的直射和反射运动，得出了结论：这些运动"依赖于形而上学定律，即自然界总是利用最简单方式产生其作用"，也就是"光所遵循的路线是使作用量为最小的路线"[11]．后来莫佩蒂证明了，光的所有折射和反射现象都与此原理相一致。

稍后，也在1744年，欧拉发表了作为自己变分学论文[10]的第二个附录的著作《用最大和最小方法确定抛体在无阻力介质中的运动》[12]。

欧拉这个著作开头写道："因为一切自然现象都应遵守某个最大或最小定律，所以，对于在某些力作用下抛体所描出的曲线，毫无疑义地具有某种最大或最小的性质……看来合乎真理的是……抛体发生的所有运动的总合应该是最小．尽管可以证明，这个结论理由不充分，但如果我证明这个结论与先验的已知真理相合，那么它就获得了那样的力量，使得一切疑问……完全消失"．继而他写道："我确信，物体描出的曲线应是在包含同样一些界限之间所有其他线中使$\int M ds\sqrt{v}$取最小，或者因M是常数，$\int ds\sqrt{v}$取最小……取走过微元距离ds的时间为dt，因$ds=dt\sqrt{v}$，我们就有

$$\int ds\sqrt{v}=\int v dt$$

因此，对于抛体描出的曲线，物体在各个时刻所具有的活力之和为最小"[12]．我们注意到，欧拉在这里用$\sqrt{v}$表示速度。

根据这个原理欧拉开始研究物体在均匀重力情形下的运动．积分$\int ds\sqrt{a+gx}$或（因$dy=pdx$而$ds=dx\sqrt{1+p^2}$）积分

$$\int dx\sqrt{(a+gx)(1+p^2)}$$

取最小，欧拉求得抛物线方程

$$y=(2/g)\sqrt{c(a-c+gx)}$$

进而，他研究了一系列其他情形的运动，借助于最小作用量原理解决了问题并确定此原理是正确的．在他后来的工作中[13]，欧拉研究了液体平衡问题和其他一系列问题。

1746年莫佩蒂发表了回忆录，在其开头提到所引用的欧拉著作[12]时写道："著名几何学家证明，在物体在有心力作用下描出的轨道上，曲线元增大的速度总是最小……．现在我试图从同样的根源中找出最高一类更重要的真理"[14]．在此著作中莫佩蒂首先作了批判性的评价："由自然界的奇妙中找到上帝存在的证明研究"，由此得出结论："必须在自然界的一般规律中找到上帝存在的证明．规律以最高理智的属性为基础，按照规律，运动可以维持、分布和消灭"．而后又写道："我……发现了一切规律都以其为基础的普适性原理．此原理可同样推广到刚体和弹性体，所有物质实体（本体）的运动和静止也都依赖此原理．这就是最小作用量原理，多么明哲、多么当之无愧的最高实质的原理．看来，自然界的这个原理应是无时无刻无所不在的；自然界不仅在其自身的变化中遵循它，而且在其自身的永恒存

在中竭力保持它”[14]．进而莫佩蒂表述了这个原理：“当自然界中发生某种变化时，这种变化所必要的作用量是最小可能的．作用量就是物体质量与其速度以及与其走过距离的乘积”[14]．最后，莫佩蒂将此原理应用于刚体和弹性体的撞击．

莫佩蒂将目的论引入力学的企图招致一系列学者激烈的反击，他的著作引起了超越力学范畴的讨论，数学家、力学家、哲学家、政论家都加入了[莫佩蒂，考尼希（König），达西（d'Arcy），库蒂弗龙（Courtivron），欧拉，达朗伯，伏尔泰（Voltaire），弗里德里希二世（Friedrich Ⅱ）等]．讨论中，优先权问题，运动度量的自然哲学问题和物理学问题，以及世界观基本问题，相互交错．讨论的核心是关于物质世界各种现象的因果制约性问题或关于按创世者智慧的目的论的既定方针问题．欧拉在讨论中支持莫佩蒂．例如，在《关于最小作用量原理的学位论文》（1753）[1]中，欧拉批评了考尼希和其他学者对莫佩蒂强调的最小作用量原理的攻击．同时，欧拉写道：他自己“理解这个重要性质如同所说是非先验的而是根据经验的”．“只是在大量试验之后，我才得出在这类运动中取最小值的那个公式．因此，我没有对它花费比我研究其他情形更大的力气”[1]．

说到莫佩蒂原理，欧拉强调指出，“所有动力学和水动力学仅用一个最大和最小方法就可以非常容易地揭示出来”[1]．这一工作的结尾写道：“自然界在所有自身表现上趋于某个最小的，并且这个最小的……无疑是作用量的概念”[1]．伏尔泰（1753）也参加了讨论，写出了辛辣嘲笑莫佩蒂及其目的论的抨击文章[15]，欧拉作为莫佩蒂的支持者也受到了“批评”．按普鲁士国王弗里德里希二世的命令，伏尔泰的这篇抨击文章被刽子手焚烧了．

达朗伯在刊载在《百科全书》上的宇宙起源论的文章（1754）中写道，把这讨论作一总结，好像莫佩蒂的著作“在1752年引起了很热烈的争论……莫佩蒂是证明在折射时作用量是最小的第一人……欧拉证明了，这个原理对于受指向定点的引力或推力作用的物体所描出的曲线是成立的．这个漂亮的命题将莫佩蒂原理也推广到描述光的微粒从一个介质运动到通过另一介质时的小曲线，因此从这一观点看原理一般说是正确的和没有限制的……欧拉还指明了在另外许多情形中这个原理可以很容易和很优美地得到应用”[16]．

考尼希在对莫佩蒂发现这个原理的优先权提出异议时，于1751年发表了雅可比写给莱布尼兹的信的摘录，它局部地提出了最小作用量原理．但是，普鲁士科学院宣称这个摘录是伪造的．

达朗伯曾写道：“一般地说关于作用量的所有定理……都不过或多或少地是一般的数学定理，而非哲学原理……因此，莫佩蒂原理与其他原理一样只是数学原理……关于作用量的这场争论……有点像宗教上关于苦痛的一些争论，而参加进去的人也经受了苦痛，因此没有什么意义”[16]．

因此，总括起来，应当说，欧拉是给出了受指向固定中心引力作用的自由物体的最小作用量原理的严格表述的第一人，即他是建立数学上表述特殊情形下最小作用量原理的奠基人．可是，对于一般情形，则只是拉格朗日给出了完全的表述，在建立和发展最小作用量原理方面起决定性作用的也是他．

在1760—1761年，拉格朗日发表了著作[17]，其中提出了确定积分的最大和最小的新方法，并应用这个方法去解各种动力学问题．实际上，从拉格朗日开始了变分学的新阶段：他在研究有效算法时不仅给出了解决以前提出那些问题的简单形式，而且应用它求解了力学的

一系列复杂问题。特别是，拉格朗日首先提出了函数变分的概念并严格地导出了积分取驻值的充要条件，即现今大家熟悉的欧拉-拉格朗日方程（3）。下面就来简要推导出这个方程。

设 $y=f(x)$ 为泛函（2）的连续可微极值曲线，与它一起研究一些通过同样一些有限点的如下形式的邻近曲线：

$$\overline{f(x)}=f(x)+\varepsilon\varphi(x)$$

其中 $\varphi(x)$ 为某任意函数，且 $\varphi(a)=\varphi(b)=0, \delta y=\overline{f(x)}-f(x)=\varepsilon\varphi(x)$ 为函数 $f(x)$ 的变分。于是

$$\delta F(x,y,y')=\varepsilon\left(\frac{\partial F}{\partial y}\varphi+\frac{\partial F}{\partial y'}\varphi'\right)$$

分部积分后得到

$$\delta J=\int_a^b \delta F dx=\int_a^b\left(\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'}\right)\varphi dx=0$$

由此考虑到函数 $\varphi(x)$ 的任意性而得出欧拉-拉格朗日方程（3）。

拉格朗日在其后续的著作，即[17]的续篇中，援引欧拉的著作[12]，推广了最小作用量原理并表述了其“普遍原理”如下：“设有任意多个物体 $M, M', M'', \cdots$，它们以任何方式相互作用，此外，它们在与某些距离的函数成正比的有心力作用下运动；设 s, s', s''，…表示物体在时刻t所通过的空间，而 u, u', u''，…为此时刻末的速度，表达式

$$M\int u ds+M'\int u' ds'+M''\int u'' ds''+\cdots$$

总是最大或最小”[18]。

在证明此原理的正确性时，拉格朗日假设对于所有可比较的运动满足能量守恒律

$$T-U=h=\text{const} \tag{4}$$

并且 $\delta h=0$。这里T和U分别为系统的动能（活力）和所施力的力函数。而后拉格朗日给出下列各种运动的一系列问题的解：物体在指向固定中心的引力作用下的运动；相互联系物体的运动；非柔软线（以及可压缩的和弹性的）的运动；非弹性和弹性流体的运动规律。

因此，拉格朗日利用能量守恒律将欧拉对自由质点表述的原理推广到彼此联系、以任何方式相互作用的任意质点系情形。雅可比以此为基础着重指出，拉格朗日最小作用量原理是分析力学之母[19]。

我们注意到，在谈到自己的最小作用量原理时，拉格朗日写道，把它看成“不是形而上学的原理，而是力学定律中简单的和普遍的结论……这个原理在与活力原理相结合并按变分学规则发展时，将给出为解每个问题所必要的全部方程；由此就使涉及物体运动的一般解题方法变得同样简单了”[6]。

1788年拉格朗日发表了自己的名著《分析力学》，后来哈密顿称之为“一种数学的诗”[20]。在这本基础性著作中，拉格朗日“抱定目的是将力学的理论以及与其相关问题的解法归结为一般公式，其简单的发展给出为解每个问题所必要的全部方程……所叙述的……方法既不需要建立，也不需要几何或力学的推理；只需按计划好的同一形式的方法进行代数运算。所有爱好分析的人们都会高兴地看到，力学已成为分析的一个新的部分”[6]。这里拉格朗日重新叙述了自己的最小作用量原理，给出了它的下述提法：“任何物体系在相互引

力的作用下或者在指向固定中心并与距离函数成比例的力的作用下运动时，各个物体所描出的曲线以及这些物体的速度应是这样的：在每条曲线的起点和终点为给定，因此对应于这些点的坐标的变分为零的条件下，速度乘以曲线元的积分再乘以各个质量之积的和是最大或最小”[6]，即

$$\delta \sum_{\nu} m_{\nu} \int v_{\nu} ds_{\nu} = 0$$

或者

$$\delta \int_{t_0}^{t} 2T dt = 0, \quad \delta q_i = 0 : t = t_0, t \tag{5}$$

并且对于所有的比较曲线通过同样一个常数 h 而满足定律（4）. 接着，拉格朗日指出，“所得出的定理不仅包含着物体运动的重要性质，而且也可以用来确定这个运动”[6]. 他证明，最小作用量原理引向由达朗伯-拉格朗日原理所表示的拉格朗日动力学普遍公式

$$\sum (\mathbf{F}_{\nu} - m_{\nu} \mathbf{W}_{\nu}) \cdot \delta \mathbf{r}_{\nu} = 0 \tag{6}$$

其中 $\delta \mathbf{r}_{\nu}$ 为在给定时刻约束所允许的虚位移. 应当指出，达朗伯于1743年在其动力学专题论文[21]中叙述的原理，他利用它解决了许多问题，特别是岁差问题，拉格朗日这样表述：“如果几个物体具有由于它们之间有相互作用而强制改变的运动，那么……这些运动就可以看成是由这些物体实际得到的运动和由其他消失的运动所组成的运动；由此得知，后者应是这样的：如果这些物体处于不存在它们的作用的情形，那么它们将彼此平衡”[6].

拉格朗日补充道，这个原理给出“直接的、普遍的方法，借助这个方法可将……所有力学问题表为方程形式……但确定那些应消失的力以及这些力的平衡规律的困难，常常使这个原理的应用带来不便并使人厌倦”[6]. 在这个意义上讲，基于“力和由这些力引起的取相反方向的运动之间”的平衡的另一方法是可取的. 正如拉格朗日指出的，这个方法意味着“复原到赫曼和欧拉的方法，它可用于解许多力学问题”[6]，这个方法表达为形式（6）. 实际上，这个原理乃是达朗伯原理与可能位移原理

$$\sum_{\nu} \mathbf{F}_{\nu} \cdot \delta \mathbf{r}_{\nu} = 0 \tag{7}$$

的结合，后者在解静力学问题时有很大通用性，由约·伯努利（1717）首先提出并为瓦里尼翁（Varignon）所应用. 这个原理给出莫佩蒂“静止定律”出现的理由，后来（1751）由欧拉发展并推广[22]. 拉格朗日看到了这个原理“还具有珍贵的仅为它固有的优越于其他原理的好处，它可将包括按物体平衡问题建立的所有问题表为一般形式”[6]，并给出静力学的普遍公式（7）.

拉格朗日由他的动力学普遍公式（6）出发，将力学系统的运动方程表为两种形式，现今它们分别称为第一类和第二类拉格朗日方程. 第二类拉格朗日方程

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \qquad (i = 1, \cdots, n) \tag{8}$$

的杰出作用在于它们相对于变量 q_i 的点变换的不变性（协变性）. 这里 q_i 为系统独立的广义坐标，点的笛卡儿坐标可用它们表为显函数 $x_i = x_i(t, q_1, \cdots, q_n)$. 方程（8）在适当选取

q_i下在任何选取的运动参考系中仍然是对的．拉格朗日方程是“不变性原理”的第一个例子，而这个原理是数学上主导思想之一，它在现代物理学中也起着头等作用．

完成了最小作用量原理研究工作的是雅可比．在他的《动力学讲义》[23]中，他发表意见：拉格朗日形式的最小作用量原理不够明了，因为包含时间作为独立变量，这时系统在有限点经历的时间依赖于该点运动的轨道，根据活力定律，因此作用积分的上限不是常数．与此相关，雅可比（1837）[24]提出借助活力积分（4）消去（5）的被积函数中的时间，并建立了几何形式的最小作用量原理

$$\delta\int_P^Q\sqrt{2(U+h)}\sqrt{\sum_\nu m_\nu ds_\nu^2}=0$$

应当强调，借助积分（4）消去（5）的被积函数中的时间，等价于将（4）当作（5）的补充条件，因此雅可比原理等价于拉格朗日原理．

在广义坐标q_i下系统的动能

$$T=\frac{1}{2}\sum_\nu m_\nu\left(\frac{ds_\nu}{dt}\right)^2$$

可以表为

$$T=\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}(q)\dot q_i\dot q_j\qquad\left(\dot q_i\equiv\frac{dq_i}{dt}\right)$$

在按公式

$$ds^2=\sum_{i,j=1}^{n}a_{ij}dq_idq_j\tag{9}$$

定义的位形空间中的度量下，雅可比原理表为

$$\delta\int_P^Q\sqrt{2(U+h)}\,ds=0\tag{10}$$

并且在系统位置的始点P和终点Q处$\delta q_i=0,\delta h=0$．

雅可比原理将完整保守系统的运动研究引向在带度量（9）的黎曼（Riemann）空间中寻求变分问题（10）的那些极值的几何问题，它们就是系统的真实轨道．因此，雅可比原理揭示出完整保守系统的运动和空间黎曼几何之间实际存在的密切联系．如果当$U=0$时，运动在没有给定力时发生，那么系统以常速度沿位形空间的测地线运动．这个事实是伽利略（Galileo）惯性定律的推广．当$U\neq0$时，系统运动的研究也归结为带度量

$$d\sigma^2=2(U+h)ds^2$$

的黎曼空间测地线的确定问题，因为此时雅可比原理（10）有形式

$$\delta\int_P^Q d\sigma=0$$

需注意，对单个质点情形，当线元ds是三维欧氏空间元时，雅可比原理乃是光学中费马原理（1）的力学比拟．比较（1）与（10），我们看到，如果$\mu=\sqrt{2(U+h)}$，则点的轨道与光线的轨道重合．但是，这种比拟仅是对于力学中运动点和光学中光线二者的轨道本身的，而过程随时间的进行在两种问题中则完全不同[25]．

必须注意，雅可比关于最小作用量原理中的最小问题的结论“不是在系统中两个任意位

置间成立，而只是在彼此充分接近的终了位置和初始位置间成立”[23]。如果在系统的轨道上取某一点A，并研究另一条通过A并与前一轨道成一小角度而与其交于点B的轨道，那么当轨道间角度无限减小时B的极限位置称为点A的动理学焦点。雅可比证明了，如果积分区间的终了点Q处于初始点P的动理学焦点之前的轨道上，那么作用量是最小。

4 哈密顿原理与光学-力学比拟

1834—1835年哈密顿[20,26]建立了比欧拉-拉格朗日-雅可比原理更为普遍的最小作用量原理。研究受定常几何约束限制的力学系统，在具有力函数U(U可显含t)的那些力作用下，哈密顿证明了形如

$$\delta\int_{t_0}^{t_1} L dt=0,\quad \delta q_i=0:t=t_0,t_1;\quad L=T+U \qquad (11)$$

的最小作用量原理是正确的。稍后不久(1848)，奥斯特罗格拉德斯基(Остроградский)[27]论证了对于非定常约束情形以及有非势力$\mathbf{F}_\nu$作用情形的原理(11)，代替(11)时原理取形式

$$\int_{t_0}^{t_1}\left(\delta T+\sum_\nu \mathbf{F}_\nu\cdot\delta \mathbf{r}_\nu\right)dt=0 \qquad (12)$$

因此，对于拉格朗日函数$L(t,q,\dot q)$和加在系统上的约束明显依赖于时间并且活力积分不存在的情形，原理(11)也是正确的。

哈密顿原理(11)的极值方程具有在表征系统状态的广义坐标q_i和广义速度$\dot q_i$下系统运动的拉格朗日方程(8)的形式。哈密顿提出取q_i和广义动量

$$p_i=\partial L/\partial \dot q_i \qquad (i=1,\cdots,n) \qquad (13)$$

作为基本变量。因(13)的右端对$\dot q_i$的雅可比行列式异于零，所以由方程(13)可解出$\dot q_i=\varphi_i(t,q,p)$。研究时引入函数

$$H(t,q,p)=\sum_{i=1}^n p_i\dot q_i-L \qquad (14)$$

此函数现今称为哈密顿函数，方程(8)就可以变换为正则方程的形式

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i},\quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i} \qquad (i=1,\cdots,n) \qquad (15)$$

它由哈密顿给出[20]。注意，考虑到(14)，原理(11)也可以表为

$$\delta\int_{t_0}^{t_1}\left(\sum_{i=1}^n p_i\dot q_i-H\right)dt=0,\qquad \delta q_i=0:t=t_0,t_1 \qquad (16)$$

上式称为哈密顿原理的第二形式。问题(16)的极值方程是哈密顿方程(15)。尽管方程(15)完全等价于方程(8)，但与(8)相比有很大优越性，因为(15)的右端仅用一个不含q,p对t的导数的函数$H(t,q,p)$来表示，而q,p对t的导数仅在方程(15)的左端出现。哈密顿根据自己对动力学的大量研究提出了这些方程。

哈密顿从光学开始了(1824)自己的科学研究[28]，而后将他得到的基本结果推广到动力学，建立了光学-力学比拟。

根据光的波动理论，对光的传播的数学描述是用惠更斯(1690)[3]提出的光线方法和波前方法。

据惠更斯波原理，光波在时刻 t 的每个波前 Σ 是新的次级波的激发源，在时刻 t' ($t'>t$) 绕过这些次级波的是波前 Σ'. 由惠更斯原理出发，哈密顿在研究中引入光学上各向同性介质的特征函数

$$V(x,y,z,x',y',z')=t'-t \tag{17}$$

它确定波由点 x,y,z 向点 x',y',z' 传播的时间间隔. 用 l,m,n 表示曲面 Σ 在点 x,y,z 的法线方向余弦，用 l',m',n' 表示波前 Σ' 在点 x',y',z' 的相应方向余弦. 因为 Σ 和 Σ' 为绕过次级波的，因此方程

$$\frac{1}{l}\frac{\partial V}{\partial x}=\frac{1}{m}\frac{\partial V}{\partial y}=\frac{1}{n}\frac{\partial V}{\partial z} \tag{18}$$

以及方程

$$\frac{1}{l'}\frac{\partial V}{\partial x'}=\frac{1}{m'}\frac{\partial V}{\partial y'}=\frac{1}{n'}\frac{\partial V}{\partial z'} \tag{19}$$

都成立. 方程 (17) — (19) 连同方程

$$l'^2+m'^2+n'^2=1 \tag{20}$$

一起组成了确定量 x',y',z',l',m',n' 作为 x,y,z,l,m,n 的函数的 6 个方程的方程组. 这些方程仅借助一个函数 $V(x,y,z,x',y',z')$ 就完全确定了光线的行为. 因此，所有光学问题都归结为确定介质的特征函数[20].

从数学观点上看，函数 (17) 确定把任何一个波前 Σ 转移到某个波前 Σ' 的切变换或接触变换，而 Σ' 由 Σ 在时间 $t'-t$ 内介质中激发的传播来得到. 后来哈密顿证明了，无穷小变换的解析表达，即波前由一个位置向无限接近于该位置的另一个位置的运动由以下方程来确定:

$$\frac{dx}{dt}=\frac{\partial H}{\partial \xi},\ \cdots,\ \frac{d\xi}{dt}=-\frac{\partial H}{\partial x},\ \cdots \tag{21}$$

关系 (17) — (20) 是方程组 (21) 的积分. 但方程 (21) 有动力学方程 (15) 的形式，而这就建立了哈密顿的光学-力学比拟.

根据这些结果，哈密顿[20] 证明了动力学方程可表为只依赖于一个未知作用函数的积分形式

$$S(t,q_i,q_i^0)=\int_{t_0}^{t}L\,dt \tag{22}$$

这里 q_i^0 表示坐标 q_i 在 $t=t_0$ 时的值，而积分是沿方程组 (15) 的某一轨道进行. 哈密顿称函数 (22) 为主函数; 它在动力学中的作用与函数 (17) 在光学中的作用一样. 实际上，在初始给定的那些小变化下，由一个运动向另一运动过渡时，S 的改变按 (22) 将等于

$$\delta S=\sum_i p_i\delta q_i-\sum_i p_i^0\delta q$$

另一方面

$$\delta S=\sum_i\frac{\partial S}{\partial q_i}\delta q_i+\sum_i\frac{\partial S}{\partial q_i^0}\delta q_i^0$$

在任意值 $\delta q_i,\delta q_i^0$ 下比较这两个表达式，我们得到

$$p_i=\frac{\partial S}{\partial q_i},\quad -p_i^0=\frac{\partial S}{\partial q_i^0}\qquad (i=1,\cdots,n)\tag{23}$$

因此，为寻求系统的运动定律 $q_i=q_i(t,q_s^0,p_s^0)$ 以及相应的值p_i，即方程组（15）的通解，只要知道主函数 S 就够了。

为此必须事先知道系统的轨道而揭示按公式（22）确定主函数的困难，哈密顿引入一阶偏微分方程

$$\frac{\partial S}{\partial t}+H\left(t,q,\frac{\partial S}{\partial q}\right)=0\tag{24}$$

主函数S对（24）来说是全积分。雅可比[23]补充了哈密顿的研究，证明，如果已知此方程的任何全积分 $S(t,q_i,\alpha_i)$，$\left\|\frac{\partial^2 S}{\partial q_i\partial\alpha_i}\right\|\neq 0$，那么关系

$$\frac{\partial S}{\partial q_i}=p_i,\quad \frac{\partial S}{\partial \alpha_i}=\beta_i\qquad (i=1,\cdots,n)\tag{25}$$

（其中 α_i,β_i 为任意常数）就是哈密顿正则方程（15）的第一积分完全组。关系（25）的第二组使我们可以确定函数 $q_i=q_i(t,\alpha_j,\beta_j)$，其第一组则可以确定函数 $p_i=p_i(t,\alpha_j,\beta_j)$，即方程（15）的通解。因此，正则方程（15）的积分问题乃是寻求哈密顿方程（24）的全积分的等价问题。尽管后一问题一般说不太简单，但用这种方法仍可解许多动力学问题。

研究并解方程（15）的非常有效的方法同样是作为相空间坐标的变量 q_i,p_i 的正则变换或接触变换方法。问题在于，代替直接对方程（15）积分，来寻求某一新变量组Q_i,P_i，使得任何方程组（15）重又变换为哈密顿方程组，但具有更简单的哈密顿函数。

实际上，研究由变量 q_i,p_i 变到新变量Q_i,P_i的变换，满足关系

$$\sum_i p_i\delta q_i-\sum_i P_i\delta Q_i=\delta W(t,q_i,Q_i)\tag{26}$$

哈密顿原理的第二形式（16）在新变量下取形式

$$\delta\int_{t_0}^{t_1}\left[\sum_i P_i dQ_i-\left(H+\frac{\partial W}{\partial t}\right)dt\right]+\delta\int_{t_0}^{t_1}dW=0,\quad \delta Q_i=0:t=t_0,t_1$$

由此得到，对新变量的微分方程也有正则形式（15），并有新的哈密顿函数

$$\widetilde{H}=H+\frac{\partial W}{\partial t}\tag{27}$$

满足关系（26）的变换称为正则变换；函数 $W(t,q_i,Q_i)$ 称为正则变换的母函数。由关系（26）得到方程

$$p_i=\frac{\partial W}{\partial q_i},\quad -P_i=\frac{\partial W}{\partial Q_i}\qquad (i=1,\cdots,n)\tag{28}$$

它们确定带有给定母函数 $W(t,q_i,Q_i)$的正则变换，并且导数$\partial W/\partial q_i(i=1,\cdots,n)$可看作变量 Q_i 的函数，它们是独立的，由此得到

$$\left\|\frac{\partial^2 W}{\partial q_i\partial Q_i}\right\|\neq 0$$

比较关系（26）和哈密顿主函数的变分

$$\delta S=\sum_i p_i\delta q_i-\sum_i p_i^0\delta q_i^0$$

可以断定变量 q_i, p_i 在初始时刻 t_0 和在时刻 t 的值彼此用正则变换相联系，并且哈密顿主函数起母函数 W 的作用．由此得知，由方程（15）描述的系统的运动可看作变量 q_i, p_i 的正则变换链．如果研究无穷小正则变换 $P_i=p_i+\Delta p_i$，$Q_i=q_i+\Delta q_i$，那么不难得到形如

$$\frac{\Delta q_i}{\Delta t}=\frac{\partial H}{\partial p_i},\quad \frac{\Delta p_i}{\Delta t}=-\frac{\partial H}{\partial q_i}\qquad (i=1,\cdots,n)$$

的方程，当 $\Delta t\to 0$ 时，它们过渡到方程（15）．因此，哈密顿正则方程使我们可以断定，系统的运动乃是变量 q_i, p_i 的无穷小正则变换的连续序列．

对于满足关系

$$\frac{\partial W}{\partial t}+H=0 \qquad (29)$$

的母函数 $W(t,q_i,Q_i)$，函数 $\tilde{H}=0$ 且正则方程取如下形式：

$$\frac{dQ_i}{dt}=0,\quad \frac{dP_i}{dt}=0\qquad (i=1,\cdots,n)$$

即 Q_i, P_i 是常量．变换公式（28）在此情形将关系（29）化为哈密顿偏微分方程（24），并使母函数 $W(t, q_i, Q_i)$ 具有这方程的全积分的意义，同时公式（28）本身就得到了雅可比定理的内容（25）。

正则变换的理论基础由雅可比（1837）[34] 奠定，他证明了在任何动力学问题中，在正则变换下运动方程的哈密顿形式保持不变．

基于这些结果，力学系统的运动的全过程可看成是正则变换的连续展开，看成是相空间到自己的连续变化的映射．

因此，根据哈密顿的光学-力学比拟，在19世纪中叶出现分析力学的新分支，这个分支兼有变分学方法和变换群理论．在这一方面已取得许多重大成果，但对这些成果的综述评论已超出本文范围．

按契塔耶夫（Четаев）[31] 的说法，哈密顿的光学-力学比拟确定了分析力学进步的世纪．

本节最后我们简要叙述一下由契塔耶夫[30-32] 发展的光学-力学比拟的未竟工作．正如前面说到的，哈密顿的光学-力学比拟是保守系统动力学与惠更斯的光波动说之间的比拟．但光学理论并未停止在惠更斯理论上，光学中替代它的是菲涅耳(Fresnel)，柯西(Cauchy)，麦克斯韦（Maxwell）的理论．

柯西将哈密顿的光学-力学比拟的进一步发展作为目标，得到了在弹性介质振动领域的这一比拟．柯西从分析动力学出发，对继续发展惠更斯之后的光线理论的比拟问题得到了自己的发现．大概克莱茵（Klein）是第一位注意到这个情况的人，并提出了光学-力学比拟的进一步发展问题．

在菲涅耳之后光被当作某个振动过程来考虑．因此，光学-力学比拟的发展，按契塔耶夫[31] 的说法，必须在振动运动的领域中，在保守系统稳定运动的性质中去寻找．进而，他把两种现象间比拟的新看法表述为一个现象的变换群与另一个现象的变换群相重合．

契塔耶夫[32]证明了对于保守系统的稳定运动，庞加莱（Poincaré）变分方程允许线性变换的幺模群，正如他说明的，这个群具有洛伦茨（Lorentz）完全群的概念。因为后者正是柯西和麦克斯韦光理论的基础，所以这个结果就是光学-力学比拟的发展。

先前（1958）契塔耶夫发表了著作[30]，其中确定了柯西的光的数学理论与保守系统的运动稳定性之间的比拟。

设系统的广义坐标和动量为 $q_i, p_i (i=1,\cdots,n)$，力函数为 $U(q_1,\cdots,q_n)$ 而动能为

$$T=\frac{1}{2}\sum_{i,j}^{n} g_{ij}p_ip_j$$

1945 年契塔耶夫[31]证明了如下定理：如果未扰运动是稳定的，那么庞加莱变分方程将化为带常系数的方程组，并且所有特征数等于零。

利用这个定理，契塔耶夫研究了当哈密顿方程（24）的全积分 $S=-ht+V(q,\alpha)$ 中的常数 α_i 为固定值时，由常数 β_i 的变化所确定的那些系统的扰动运动，并求得椭圆型偏微分方程

$$\sum_{i,j}\frac{\partial}{\partial q_i}\left(g_{ij}\frac{\partial V}{\partial q_j}\right)=0$$

它表示所有特征数为零的条件。如果研究某个二次可微函数 $\varphi(-ht+V)$，它依赖于哈密顿方程的全积分 $-ht+V$，那么所指条件就有波动方程形式

$$\frac{2(U+h)}{h^2}\frac{\partial^2\varphi}{\partial t^2}=\sum_{i,j}\frac{\partial}{\partial q_i}\left(g_{ij}\frac{\partial\varphi}{\partial q_j}\right) \tag{30}$$

这个波动方程建立了柯西光理论与保守系统稳定运动之间的比拟。

5　非完整系统

赫兹（Hertz）（1894）[33]引进了术语“完整系统和非完整系统”。前者只用有限个约束和可归结为有限个约束来限制，后者则用微分不可积约束来限制。前面所叙述的结果是属于完整系统的。

力学的积分原理对非完整系统的可应用性问题已有很长历史并有大量文献，我们仅指出其中的某些工作。众所周知，这些原理起初是对完整系统建立的；当试图将其推广到非完整系统时发生了严重困难，看来还是赫兹[33]首先注意到了这一点。他得到了关于哈密顿原理对非完整系统不可应用的结论，并且指出了并非位形空间的所有两点都能用非完整系统的轨道来联结。赫尔德（Hölder 1896）[34]提出了新的积分原理如下：

$$\int_{t_0}^{t_1}\left(2T\frac{d\delta t}{dt}+\delta T+\sum_{\nu}\mathbf{F}_{\nu}\cdot\delta\mathbf{r}_{\nu}\right)dt=0 \tag{31}$$

其中记号 δ 表示变分，在这种变分下虚位移 $\delta\mathbf{r}_{\nu}$ 趋于时刻 t 的原始轨道的每一点，结果发生了新的轨道，这轨道的点与原始轨道的点相对应，但不是时刻 t 而是时刻 $t+\delta t$ 的点，这里 δt 为无穷小函数，并且在 $t=t_0, t_1$ 时 $\delta\mathbf{r}_{\nu}=0$。在变分的进一步专门选取下，设 $\delta t=0$，赫尔德由（31）得到了哈密顿-奥斯特罗格拉德斯基原理（12），而设

$$\delta T=\sum_{\nu}\mathbf{F}_{\nu}\cdot\delta\mathbf{r}_{\nu}$$

时，则得到了推广形式的拉格朗日最小作用量原理

$$\delta\int_{t_0}^{t_1} T dt = 0 \tag{32}$$

至于如何确定，应当遍历系统变更过的位置的连续序列．当存在能量积分（4）时，原理（32）等价于原理（5）．要注意，变更过的运动不满足不可积约束方程．由此特别地得到，在有势力下，对非完整系统的哈密顿原理没有像对完整系统那样使拉格朗日函数积分取驻值的形式（11），而有取拉格朗日函数变分的对时间的积分为零的形式

$$\int_{t_0}^{t_1} \delta L dt = 0, \quad \delta q_i = 0 : t = t_0, t_1 \tag{33}$$

1901年同时发表了沃龙涅茨（Воронец）[35]和苏斯洛夫（Суслов）[36]的工作，其中对线性约束情形提出了两个新的积分原理形式，它们不同于赫尔德原理的形式，并且前一位作者对他提出的原理既没有给出论证也没有取一个名称，后一位作者则称他的原理为达朗伯原理的变形并强调指出它"根本不是哈密顿原理"．

沃龙涅茨提出的原理有形式

$$\int_{t_0}^{t_1}\left[\delta(\Theta+U)+\sum_{l=1}^{r}\frac{\partial T}{\partial \dot{q}_{k+l}}(\delta\dot{q}_{k+l}-\delta\varphi_l)\right]dt=0 \tag{34}$$

并假定对所有速度满足关系

$$\frac{d}{dt}\delta q_i = \delta \dot{q}_i \qquad (i=1,\cdots,n) \tag{35}$$

Θ 为借助约束方程用独立速度表示的动能 T．

但苏斯洛夫假设条件（35）不是对所有速度满足，而只对独立速度 $\dot{q}_s$ $(s=1,\cdots,k=n-r)$ 成立，这时对不独立速度有条件

$$\frac{d}{dt}\delta q_{k+l}-\delta\dot{q}_{k+l}=\sum_{s=1}^{k}A_s^{k+l}\delta q_s \qquad (l=1,\cdots,r) \tag{36}$$

其中量 A_s^{k+l} 由约束方程确定．苏斯洛夫得到的原理有形式

$$\int_{t_0}^{t_1}\left(\delta L+\sum_{l=1}^{r}\frac{\partial T}{\partial \dot{q}_{k+l}}\sum_{s=1}^{k}A_s^{k+l}\delta q_s\right)dt=0 \tag{37}$$

1911年查普雷金（Чаплыгин）[37]证明了，引入由方程 $Ndt=d\tau$ 确定的新独立变量 τ 替代 t，其中 $N(q,q_1)$ 即所谓约化乘子，依赖于两个自由参数 q, q_1 的非完整系统的运动方程可写成哈密顿正则方程组．当由某两个相容方程确定的约化乘子 N 存在时，在积分

$$\tau=\int_0^t N dt$$

不变条件下，变形的哈密顿原理

$$\delta\int_0^t LN dt = 0 \tag{38}$$

成立．

克尔纳（Kerner）[38]将受线性微分约束的系统的运动方程与满足约束方程的曲线类中哈密顿作用量驻定的拉格朗日变分问题的欧拉方程作了比较，并证明了这些方程当且仅当约束完全可积时才是等价的．根据这个结果，他得出结论：看作驻定作用量变分原理的哈密顿

原理仅对完整系统才是对的。

卡蓬（Capon）[39]注意到，赫尔德的形式体系不同于变分学的形式体系（比较曲线不满足约束方程，这时变分受约束条件的限制），因此他宣告了赫尔德的结果没有根据。同时，杰弗里斯（Jeffreys）[40]，帕尔斯（Pars）[41]反对这篇文章，发表支持赫尔德结果的意见。杰弗里斯由物理前提出发，指出非完整系统的哈密顿原理类似于带非有势力的完整系统的原理，不是驻定作用量原理，但赫尔德形式（33）是对的。帕尔斯以精密的分析考查了这个问题，并证明，对于完整系统，无论驻定作用量原理形式（11），还是赫尔德形式（33）都是对的，而对于非完整系统则只是后一形式才是对的。

诺沃谢洛夫（Новоселов）[42]指出对于具有非线性契塔耶夫约束[43]系统积分原理的正确性，并且证明，对于小积分域，这些积分原理给出作用量在真实轨道上有极小，但比较曲线不满足约束。纽马克（Неймарк）和福法耶夫（Фуфаев）指出，稳定作用量原理的表达形式"……依赖于对交换关系的观点"[44]，而对于非完整系统，特别地，赫尔德形式是对的。

不难看出[45]，原理（33），（34）和（37）是等价的并可彼此变换。研究以拉格朗日函数 $L=T+U$ 和一般说对速度 $\dot{q}_i$ 是非线性的理想不可积独立约束

$$f_l(t,q,\dot{q})=0 \qquad (l=1,\cdots,r) \tag{39}$$

来表征的非完整系统。约束（39）也可表为形式

$$\dot{q}_{k+l}=\varphi_l(t,q_i,\dot{q}_1,\cdots,\dot{q}_k) \qquad (l=1,\cdots,r) \tag{40}$$

虚位移 δq_i 满足契塔耶夫条件[43]

$$\sum_{i=1}^{n}\frac{\partial f_l}{\partial \dot{q}_i}\delta q_i=0 \qquad (l=1,\cdots,r)$$

如果认为对于所有速度都满足交换性条件（35），那么

$$\delta f_l=\sum_{i=1}^{n}\left(\frac{\partial f_l}{\partial q_i}-\frac{d}{dt}\frac{\partial f_l}{\partial \dot{q}_i}\right)\delta q_i \qquad (l=1,\cdots,r) \tag{41}$$

在约束可积情形，$\delta f_l=0$；在约束（39）不可积情形，表达式（41）不等于零，但如果约束（39）是非线性的，它们就有可能按系统的运动方程

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}-\frac{\partial L}{\partial q_i}=\sum_l \mu_l\frac{\partial f_l}{\partial \dot{q}_i} \qquad (i=1,\cdots,n) \tag{42}$$

而变为零[42]。不等式 $\delta f_l\neq 0$ 意味着变分过的路线不满足约束方程。

在沃龙涅茨原理（34）中量 $\Theta(t,q_i,\dot{q}_1,\cdots,\dot{q}_k)$ 表示借助于方程（40）消去不独立速度的动能 $T(t,q_i,\dot{q}_i)$。考虑到

$$\delta T=\delta\Theta+\sum_l\frac{\partial T}{\partial \dot{q}_{k+l}}(\delta\dot{q}_{k+l}-\delta\varphi_l) \tag{43}$$

并在表达式 $\delta L=\delta T+\delta U$ 中量 δT 用等式（43）的右边替代，我们就可将（33）变换为等式（34），而因此它就是哈密顿原理的沃龙涅茨形式。

另一方面，如果考虑到在苏斯洛夫采用的方法中变分 $\delta\dot{q}_{k+l}=\delta\varphi_l$，则方程（43）取形式 $\delta T=\delta\Theta$，由此苏斯洛夫原理化为形式

$$\int_{t_0}^{t_1}\Big[\delta(\Theta+U)+\sum_{l=1}^{r}\frac{\partial T}{\partial \dot{q}_{k+l}}\sum_{s=1}^{k}A_s^{k+l}\delta q_s\Big]dt=0$$

这就是等式（34）．于是证明了赫尔德形式（33），沃龙涅茨形式（34）和苏斯洛夫形式（37）的积分原理的等价性，它们都是非完整系统的哈密顿原理的各种不同形式．

对照满足约束方程（39）的曲线类中作用积分（11）取极值的拉格朗日问题

$$\delta\int_{t_0}^{t_1}\Big(L(t,q_i,\dot{q}_i)+\sum_l \kappa_l f_l(t,q_i,\dot{q}_i)\Big)dt=0 \tag{44}$$

其中 $\kappa_l(t)$ 为拉格朗日不定乘子．问题（44）的欧拉-拉格朗日方程（3）

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}-\frac{\partial L}{\partial q_i}=\sum_l \kappa_l\Big(\frac{\partial f_l}{\partial q_i}-\frac{d}{dt}\frac{\partial f_l}{\partial \dot{q}_i}\Big)-\sum_l \dot{\kappa}_l\frac{\partial f_l}{\partial \dot{q}_i}\qquad (i=1,\cdots,n) \tag{45}$$

乃是对q_i为二阶、对κ_l为一阶的微分方程．方程组（45），（39）的通解依赖于$2n$个任意常数，可是运动方程（42），（39）的通解依赖于$2n-r$个任意常数，由此显然，方程（45）和（42）对于非完整系统是不等价的．

不难证明[45]，使方程（42），（39）的某个解 $q_i(t)$ 处于方程（45），（39）的解中的充要条件是

$$\sum_{l,i}\kappa_l\Big(\frac{\partial f_l}{\partial q_i}-\frac{d}{dt}\frac{\partial f_l}{\partial \dot{q}_i}\Big)\delta q_i=0 \tag{46}$$

由此得知，当且仅当非完整系统的运动满足条件（46）时，哈密顿原理才具有驻定作用量变分原理（11）的特征．这个结论对于拉格朗日形式和雅可比形式的最小作用量原理也是对的[46]．

在按公式（36）取变分时，在条件[47]

$$\sum_l \frac{\partial T}{\partial \dot{q}_{k+l}}A_s^{k+l}=0\qquad (s=1,\cdots,k) \tag{47}$$

下，原理（37）具有驻定作用量原理的特征．必须强调，条件（46），（47）对于非完整系统仅在极个别情形中才满足．

对于真实的运动，哈密顿作用量取驻值的问题与积分等价于方程（42）的正则运动方程

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i},\quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i}+\sum_l \mu_j\frac{\partial f_l}{\partial \dot{q}_i}\qquad (i=1,\cdots,n) \tag{48}$$

的哈密顿-雅可比方法对于非完整系统的推广问题密切相关．借助于变量变换

$$\pi_i=p_i+\sum_l \kappa_l\frac{\partial f_l}{\partial \dot{q}_i}\qquad (i=1,\cdots,n)$$

哈密顿函数可表为[46]

$$H(t,q_i,p_i)=H_1(t,q_i,\pi_i)-\sum_{l,i}\kappa_l\frac{\partial f_l}{\partial \dot{q}_i}\dot{q}_i$$

对于广义哈密顿方程

$$\frac{\partial S}{\partial t}+H_1\Big(t,q,\frac{\partial S}{\partial q}\Big)=0\qquad (i=1,\cdots,n) \tag{49}$$

特征方程有形式

$$\frac{dq_i}{dt}=\frac{\partial H_1}{\partial \pi_i},\quad \frac{d\pi_i}{dt}=-\frac{\partial H_1}{\partial q_i} \qquad (i=1,\cdots,n) \tag{50}$$

如果 $S(t,q,\alpha)$ 是方程（49）的全积分，那么按雅可比定理，关系式

$$\frac{\partial S}{\partial q_i}=\pi_i,\quad \frac{\partial S}{\partial \alpha_i}=\beta_i \qquad (i=1,\cdots,n) \tag{51}$$

就是方程组（50）的积分。

可以证明[46,48]，仅当满足条件（46）时，方程（50）的解也是运动方程（48），（39）的解。

因此，与拉格朗日乘子法相结合的广义哈密顿-雅可比方法，当且仅当哈密顿原理具有驻定作用量的变分原理的特征时，才可应用于非完整系统的积分。

6　谢道夫变分方程

哈密顿原理如同其他积分原理，不仅可应用于有限自由度系统，这前面已说过，而且可应用于带分布参数系统和连续介质。在可逆物理过程中，这些原理是某些泛函取驻值的变分原理，而在不可逆过程的情形则是变分方程。

我们简略地叙述谢道夫（Седов）变分方程[49]，它是基于变分方法对不可逆过程的推广。谢道夫方程有形式

$$\delta\int_{t_0}^{t_1}\int_V \Lambda dv dt+\delta W^*+\delta W=0 \tag{52}$$

其中 V 为连续介质的任意体积，$[t_0,t_1]$ 为任意时间间隔，$\Lambda=\rho[(1/2)v_iv^i-U]$ 为拉格朗日函数，δW^*和δW是如下形式的泛函：

$$\delta W^*=\int_{t_0}^{t_1}\left[\int_V \rho T dS dv-\delta G'+\delta A_{\text{об}}^{(e)}\right]dt$$

$$\delta W=-\left[\int_V \rho v_i\delta x^i dv\right]\Bigg|_{t_0}^{t_1}+\int_{t_0}^{t_1}\delta A_{\text{пов}}^{(e)}dt$$

并且 ρ 为质量密度，U 和 S 为单位质量的内能密度和熵密度，x^i 为坐标，$v^i=dx^i/dt$ 为速度，T 为温度，$\delta G'$ 为非补偿热，$\delta A_{\text{об}}^{(e)}$和 $\delta A_{\text{пов}}^{(e)}$ 分别为外体力和面力的功。

在方程（52）中，拉格朗日函数 Λ 和泛函 δW^* 是给定量，泛函 δW 由方程（52）确定，它的计算相应于建立状态方程。

变分方程（52）实际上是考虑到热力学第一、第二定律所描述的连续介质的动力学普遍方程对时间的积分。

变分方程（52）由谢道夫（1965）给出，与建立带复杂性质连续介质新模型的问题相关。谢道夫将这个变分方程作为连续介质力学的原始公设。在变分方法范围内建立新的模型，包括确定选取待定函数并给出拉格朗日函数Λ和泛函 δW^*[50]。

7　力学的微分原理

还要简述一下力学的微分原理。达朗伯-拉格朗日原理（6）是变分方程，没有极值特性。高斯（Gauss 1829）[51]赋予达朗伯-拉格朗日原理以极值原理的形式。引入仅在加速度上不同于真实运动的想象速度，高斯证明了对于真实运动来说，拘束有极小：

$$\sum_{\nu} \frac{1}{2} m_{\nu} \left(\ddot{\mathbf{r}}_{\nu} - \frac{\mathbf{F}_{\nu}}{m_{\nu}} \right)^2 = \min$$

赫兹（1894）[33]在建立无力力学时提出了直路（或最小曲率）原理

$$\frac{1}{2} \sum_{\nu} m_{\nu} \ddot{\mathbf{r}}_{\nu}^2 = \min$$

它实际上是特殊情形 $\mathbf{F}_{\nu}=0$时的高斯原理。

契塔耶夫（1941）[52]给出了高斯原理的变形，证明对于真实运动来说，如果力学系统是完全自由的，那么由在作用力场中高斯的正向想象运动和为建立真实运动所需力场中的倒退（反向）运动两者组成的基本循环上的功A是极大：

$$A = \max A_{\mu}, \quad A_{\mu} = \sum_{\nu} (\mathbf{F}_{\nu} - m_{\nu} \ddot{\mathbf{r}}_{\nu}) \cdot \left(\dot{\mathbf{r}}_{\nu} + \mathbf{w}_{\nu}^{\mu} \frac{dt}{2} \right) dt$$

其中 $\mathbf{w}_{\nu}^{\mu}$为想象运动中的加速度。

注意到这个原理是赫曼和欧拉思想的直接变形，拉格朗日在叙述达朗伯原理时发展了这个思想。运用热力学卡诺（Carnot）原理的办法，契塔耶夫原理可扩充通常研究的力学系统的特性。

这个原理在物理系统方面的推广有形式

$$\delta \left\{ -\frac{dt^2}{2} \sum_{\nu} \frac{1}{2m_{\nu}} \left[(\mathbf{F}_{\nu} - m_{\nu} \mathbf{w}_{\nu})^2 + \Delta Q - \Delta U \right] \right\} = 0$$

其中 U 为内能，ΔQ 为热流并利用热力学第一、第二定律。这就是物理系统的契塔耶夫原理[53,54]。

达朗伯-拉格朗日原理，高斯原理和契塔耶夫原理也可推广到连续介质[55]。

8　庞特里雅金极大值原理

最后我们简述最优控制[56]的最简单问题，这个问题是在问题

$$J = \int_{t_0}^{t_1} f^0(t, x_1, \cdots, x_n; u_1, \cdots, u_r) dt = \min$$

中，在条件

$$\frac{dx^i}{dt} = f^i(t, x_1, \cdots, x_n; u_1, \cdots, u_r), \quad t_0 \leqslant t \leqslant t_1$$

$$x(t_0) = x_0, \quad x(t_1) = x_1, \quad u(t) \in U$$

下，来寻求必要条件。

设可允控制类由带第一类间断点、在某闭域 U 上取值的所有分段连续有界函数 $u(t)$ 组成。这个问题就是庞特里雅金（Понтрягин）提法下变分学的拉格朗日问题。类似问题的本质特征是事先确定在闭域U上取值的可允控制类。

如果在研究中引入庞特里雅金函数

$$\Pi(\phi(t), x(t), t, u(t)) = \sum_{\alpha=0}^{n} \phi_{\alpha} f^{\alpha}(t, x, u)$$

那么，相变量x^i和起拉格朗日乘子作用的辅助变量 ϕ_i的方程可写成正则形式

$$\frac{dx_i}{dt}=\frac{\partial \Pi}{\partial \phi_i},\quad \frac{d\phi_i}{dt}=-\frac{\partial \Pi}{\partial x^i}\qquad (i=0,1,\cdots,n)$$

下述定理成立[56]：设 $u(t)$，$t_0\leqslant t\leqslant t_1$ 是相点由位置 x_0 到位置 x_1 过渡的可允控制，而 $x(t)$ 是相应的轨道，于是 $x(t_0)=x_0$，$x(t_1)=x_1$（时刻 t_0，t_1 固定）．为使 $u(t)$ 给出带固定时间的最优问题的解，必须存在与 $u(t)$ 和 $x(t)$ 对应的、非零连续矢量函数 $\phi(t)=(\phi_0(t),\phi_1(t),\cdots,\phi_n(t))$，使得

①对所有 t，$t_0\leqslant t\leqslant t_1$，变量 $u\in U$ 的函数 $\Pi(\phi(t),x(t),t,u)$ 在点 $u=u(t)$ 达到极大

$$\Pi(\phi(t),x(t),t,u(t))\underset{u\in U}{=}M(\phi(t),x(t),t)$$

②函数 $\phi_0(t)=\text{const}$ 为非正．

显然，庞特里雅金极大值原理属于微分变分原理范畴．当控制的可允值集合为开集时，极大值原理等于著名的维尔斯特拉（Weierstrass）条件．但在控制落到域 U 的边界上时，维尔斯特拉条件一般说不满足，而极大值原理此时也是对的．

可以证明[57]，最优原理也可表为积分变分原理形式

$$\delta\int_{t_0}^{t_1}\left(\sum_{i=0}^{n}\phi_i\dot{x}^i-\Pi\right)dt\geqslant 0,\quad \delta x_i=0:\ t=t_0,t_1$$

它类似于哈密顿原理．

在一般情形下，由这个积分原理可导出运动方程，边界条件和横截性方程，以及庞特里雅金极大值原理．

*　　*　　*

最后我们再一次强调，力学的变分原理将运动方程和场方程表为简单的不变形式，包括综合了运动的连续的和离散的观点，并且是力学中广义的因果性原理．它们具有巨大的启发作用，特别是在连续介质力学，相对论，量子物理中，并且可以得到现代核物理以及基本粒子理论的信赖．力学的变分原理已在亥姆霍兹（Helmholtz），玻耳兹曼（Boltzmann），吉布斯（Gibbs），庞加莱，爱因斯坦（Einstein），玻尔（Bohr），索末菲（Sommerfeld），德布罗意（de Broglie），薛定谔（Schrödinger），海森伯（Heisenberg），狄喇克（Dirac），波恩（Born）以及其他许多学者的著作中得到了应用、推广和发展．力学变分原理的极大普遍性，在许多非力学领域推广的可能性，与守恒律和李群的相关性，都使这些原理在解许多物理的基础问题中具有中心的位置．但这些非常有意义的问题已超出本文的范围．

参考文献（57篇，略）

100081 北京理工大学**梅凤翔**译自：Развитие идей Леонарда Эйлера и современная наука. Сборник статей, Москва «Наука» (1988)：180—207.（**董务民**校）