理论力学

孙天阳

中国科学技术大学数学科学学院

tysun@mail.ustc.edu.cn

2025 年 11 月 3 日

目录

目录 1

1 拉格朗日力学 **2**

1 最小作用量原理 2

2 变分法和拉格朗日方程 4

Chapter 1

拉格朗日力学

1 最小作用量原理

首先我们要回答一个问题, 如何描述物体运动的完整状态? 或者说, 预测该物体在任何其他时刻（未来或过去）的运动所需要的最小完备信息是什么? 牛顿力学启发我们, 只知道位置是不够的, 但知道速度就足够了. 因为描述物体运动的牛顿第二定律是一个二阶方程, 只需要且必须由两个初始条件决定. 所以我们用 $(q, \dot{q})$ 来描述物体的状态. 但需要知道, 这里的速度 $\dot{q}$ 并不是一个来自于真实运动轨迹的速度, 而是说, 我用位置 $q = 0$ 且速度 $\dot{q} = 1$ 或位置 $q = 0$ 且速度 $\dot{q} = -1$ 作为初始来描述状态. 所以说, 状态空间的正确概念, 其实不只是 q 的全体, 流形 M, 而是 $(q, \dot{q})$ 的全体, TM, 也就是 M 的切丛. 我认为这是拉格朗日力学的理论的起点.

定义 1.1. *设 M 是一个流形, M 上的一个不含时拉格朗日量是指一个光滑函数 $L\colon TM \to \mathbb{R}$, M 上的一个含时拉格朗日量是指一个光滑函数 $L\colon \mathbb{R}_t \times TM \to \mathbb{R}$.*

单单从这个定义来看, 即使考虑上我们接下来马上要定义的作用量, 考虑上最小作用量原理, 考虑上从最小作用量原理推导出的欧拉-拉格朗日方程, 这一切不过是无聊的数学游戏, 我们可以随便选择一个流形 M, 随便选择一个光滑函数 L, 都可以展开研究. 但拉格朗日力学成功在于, 当我们假设 $L = T - V$, 我们就能推导出牛顿力学, 当我们假设 L 是麦克斯韦-洛伦兹拉格朗日量, 我们就能推导出麦克斯韦方程组和洛伦兹力, 所以或许是其这种普适性, 使得其定义上必须抽象.

定义 1.2. *设 $\gamma\colon [t_1, t_2] \to M$ 是一条光滑曲线, 则 $\dot{\gamma}$ 是从 $[t_1, t_2]$ 到 TM 的光滑曲线. 定义作用量 S 为 $S[\gamma] = \displaystyle\int_{t_1}^{t_2} L(t, \dot{\gamma}(t))\,\mathrm{d}t = \int_{t_1}^{t_2} L(t, q^i(t), \dot{q}^i(t))\,\mathrm{d}t$, 这是一个泛函.*

注记. *设 M 是一个 n 维流形, 则 TM 是一个 $2n$ 维流形. $\gamma(t)$ 是 M 中的一个点, 需要用 n 个坐标来表达, $\dot{\gamma}(t)$ 是 TM 中的一个点, 需要用 $2n$ 个坐标来表达, 前 n 个坐标描述它是哪个点处的切空间中的切向量, 后 n 个坐标描述它具体是这个切空间中的哪个切向量. 如果 $\gamma(t)$ 在局部坐标下的表达是 $(q^1(t), \cdots, q^n(t))$, 那么 $\dot{\gamma}(t)$ 在局部坐标下的表达就是 $(q^1(t), \cdots, q^n(t), \dot{q}^1(t), \cdots, \dot{q}^n(t))$.*

公理 1.3 (最小作用量原理/哈密顿原理)**.** *给定拉格朗日系统 (M, L), 在所有固定端点的曲线 $\gamma\colon [t_1, t_2] \to M, \gamma(t_1) = p_1, \gamma(t_2) = p_2$ 中, 系统总是沿着使作用量泛函 S 取驻值的曲线运动.*

注记. *S 是一个泛函, 那它的定义域是什么呢? 我觉得你可以说是对任意的 $t_1 < t_2$ 然后 $\gamma\colon [t_1, t_2] \to M$ 这种光滑曲线, 也可以说是对固定的 $t_1 < t_2$ 然后 $\gamma\colon [t_1, t_2] \to M$ 这种光滑曲线. 总之它这个*

定义域的数学结构还是蛮复杂的, 至少没有线性结构, 因为 γ 的值域 M 是流形, 而流形上没有加法和数乘. 严格来说可能是实 Fréchet 流形. 尽管 S 的定义域比较大, 但最小作用量原理中谈论 γ 是 S 的极值时, 并不是在整个定义域的意义下考虑的, 或者说没有考虑所有的切方向, 而是在加了 $\gamma(t_1) = p_1, \gamma(t_2) = p_2$ 这个端点固定的约束的意义下考虑的, 从切向量的角度考虑就是容许的切方向只有端点为 0 的切方向.

2 变分法和拉格朗日方程

回忆在数学分析中, 函数 f 在某点 x_0 处的微分 $\mathrm{d}f(x_0)$ 其实是关于增量 $\mathrm{d}x$ 的线性函数（如果要用微分形式的语言, 那 $\mathrm{d}f$ 就是微分形式, 而所谓增量应该是切空间中的元素, 但这里朴素理解就好）, 并且这个线性函数是线性主部, 也就是说 $f(x_0+\mathrm{d}x)-f(x_0)-\mathrm{d}f(x_0)(\mathrm{d}x)$ 是一个关于 $\mathrm{d}x$ 的二阶小量. 读者如果能清晰且熟练地把握住这一点, 那便有希望不被变分搞乱. 让我们来进行逐字对应: 泛函 S 在某点 $\gamma(t)$ 处的变分 $\delta S[\gamma(t)]$ 其实是关于增量 $\delta\gamma(t)$ 的线性泛函.

另一套重要的语言是, 如果我们想求函数 f 在某点 x_0 处沿方向 v 的方向导数, 只需要找一条经过 x_0 并在 x_0 处的切向量为 v 的曲线 $l(\varepsilon)$, 即要满足 $l(0)=x_0$ 且 $l'(0)=v$, 然后将 f 与 l 复合起来并关于 ε 在 $\varepsilon=0$ 处求导. 这在泛函与变分这里是一样的, 如果我们想计算泛函 S 在 $\gamma(t)$ 处沿切向量 $\delta\gamma(t)$ 的方向导数, 只需要找一族关于 ε 的曲线 $\gamma_\varepsilon(t)$ 或者 $\gamma(t,\varepsilon)$, 使得 $\gamma_0(t)=\gamma(t)$, 并且关于 ε 在 $\varepsilon=0$ 处求导时得到的切向量是 $\delta\gamma(t)$, 将这族曲线与 S 复合然后关于 ε 在 $\varepsilon=0$ 处求导.

有了以上铺垫, 我们可以开始计算上一节中定义的作用量泛函 S 的变分. 首先我们考虑 S 的定义域是所有定义在 $[t_1,t_2]$ 上到 M 的光滑曲线的情形. 考虑一族曲线 $\gamma(t,\varepsilon)$, 则 $S[\gamma_\varepsilon]$ 现在是一个关于 ε 的函数, 可以关于 ε 求导,

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} S[\gamma_\varepsilon]=\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0}\int_{t_1}^{t_2} L(t,q^i(t,\varepsilon),\dot{q}^i(t,\varepsilon))\,\mathrm{d}t=\int_{t_1}^{t_2}\frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial\varepsilon}+\frac{\partial L}{\partial\dot{q}^i}\frac{\partial\dot{q}^i}{\partial\varepsilon}\mathrm{d}t$$

$$=\int_{t_1}^{t_2}\frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial\varepsilon}\bigg|_{\varepsilon=0}+\frac{\partial L}{\partial\dot{q}^i}\frac{\partial^2 q^i}{\partial\varepsilon\partial t}\bigg|_{\varepsilon=0}\mathrm{d}t=\int_{t_1}^{t_2}\frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial\varepsilon}\bigg|_{\varepsilon=0}+\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{q}^i}\frac{\partial q^i}{\partial\varepsilon}\right)\bigg|_{\varepsilon=0}-\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{q}^i}\right)\frac{\partial q^i}{\partial\varepsilon}\bigg|_{\varepsilon=0}\mathrm{d}t$$

$$=\int_{t_1}^{t_2}\left(\frac{\partial L}{\partial q^i}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{q}^i}\right)\frac{\partial q^i}{\partial\varepsilon}\bigg|_{\varepsilon=0}\mathrm{d}t+\left(\frac{\partial L}{\partial\dot{q}^i}\frac{\partial q^i}{\partial\varepsilon}\right)\bigg|_{\varepsilon=0}\bigg|_{t_1}^{t_2}$$

上述推导, 是基于 $\gamma_\varepsilon(t)$ 都是定义在 $[t_1,t_2]$ 上的曲线, 也就是固定积分边界的变分. 如果额外假设 $\gamma_\varepsilon(t_1)=p_1,\gamma_\varepsilon(t_2)=p_2$, 即固定端点的变分, 那么上面的边界项会消失.

朱界杰老师的讲义里还提到可动积分边界的变分, 这必须小心处理. 首先必须定义清楚此时 S 的定义域是什么, 我认为合理的定义是三元组 $(a,b,\gamma(t))$, 其中 $a<b$ 而 γ 是 $[a,b]\to M$ 的光滑曲线. 我们现在还是在三元组的空间中取一族对象 $(a(\varepsilon),b(\varepsilon),\gamma(t,\varepsilon))$. 注意, 我们并没有给出三元组上的拓扑, 也没有给出三元组上的光滑结构, 所以 $(a(\varepsilon),b(\varepsilon),\gamma(t,\varepsilon))$ 对 ε 的光滑依赖性是没有被定义的, 但我自认为已经在严谨与直观之间做了一个很好的取舍. 我们仍是对 ε 求导

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0}\int_{a(\varepsilon)}^{b(\varepsilon)} L(t,q^i(t,\varepsilon),\dot{q}^i(t,\varepsilon))\,\mathrm{d}t=L(b,\gamma(b))b'(0)-L(a,\gamma(a))a'(0)+\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0}\int_a^b L(t,q^i(t,\varepsilon),\dot{q}^i(t,\varepsilon))\,\mathrm{d}t$$