

最小作用原理的一个例子：抛物体运动

考察一个在重力场中运动的粒子，为了简单起见，我们假定运动是一维的，你把它抛掷出去，它就会竖直上升而后又竖直落下。如果作一个位置-时间图，它将是一条抛物线 $x(t)$ ，其中 x 是距离地面的高度，假设它在在起始时刻 t_1 由某一高度出发，并在结束时刻 t_2 确实到达了另外某一点[图 1a]。现在，你尝试一个不同的运动：假设由这里到达那里是这样进行的，粒子在升至十分高之后又以某一特殊方式降落下来[图 1b]，但却在同一段时间内到达目的地 $\bar{x}(t)$ 。如果你算出在该路程上每一时刻的动能，减去势能，再计算出经历整条路线对时间的积分，或对其他任何我们所想出的路程积分。非常奇妙的是，你将会发现所获得的数字比起对实际运动所得的要大。也就是说，真正的路程就是那一条会使这一积分取得最小值的路程。

$$\int_{t_1}^{t_2} \left[\frac{1}{2} m \left(\frac{dx}{dt} \right)^2 - mgx \right] dt \quad (1)$$

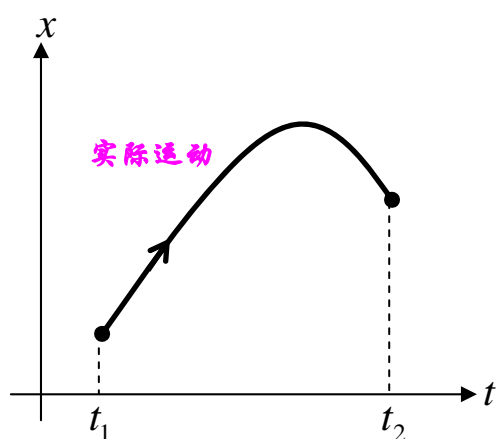


图 1a

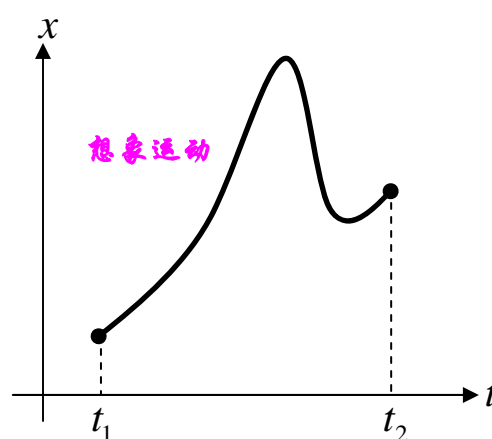


图 1b

换句话说，我们可以将动力学定律不写成 Newton 定律 $\vec{F} = m\vec{a}$ 的形式，而是写成：**粒子从一点至另一点所经行的路线其平均动能减去平均势能应尽可能地少。**

作为一个例子，考虑一个没有任何势能的自由粒子情况。我们知道粒子从一点跑至另一点的运动是匀速直线运动。该法则的意思就是；在某一定时间内从一点跑至另一点动能的积分是最少的，这是否可以告诉我们粒子就一定要以一匀速率运动呢？因为粒子是在某一定时间内由“这里”到达“那里”，因此，平均速度对于每一情况都是相同的。假如该粒子以任何其他方式运动，速度便将有时比平均值高，有时比平均值低。原来任何与平均值有偏离的东西，其均方值总是大于其平均值的平方的，所以我们见到，如果速度固定不变（当没有力作用时），则该积分便是一个极小值。正确的路线就是这样[图 2]。

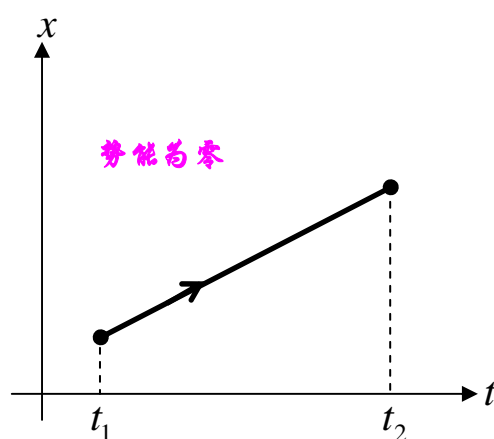


图 2

[对于任意函数，我们有

$$\frac{1}{T} \int_{t_1}^{t_2} |f(t)|^2 dt \geq \left(\frac{1}{T} \int_{t_1}^{t_2} f(t) dt \right)^2, \quad T = t_2 - t_1$$

等号当且仅当 f 为常值函数时才成立；实际上，它是另一个你所熟悉的关系

$$\frac{f_1^2 + f_2^2 + \dots + f_n^2}{n} \geq \left(\frac{f_1 + f_2 + \dots + f_n}{n} \right)^2$$

取连续极限的情况！]

回到抛物体的情况，因为粒子在重力场中具有势能，它就必须在平均上有最小的动能与势能之差值。由于在空间升得越高势能越大，如果能够尽快升到一个有高势能的地方去的话，这样该势能从动能那里扣除出去后，就可以获得一个较低的平均值。所以最好就是去选取一条能够升得高，从而可从势能处得到一个大负值的路程[图 3]。

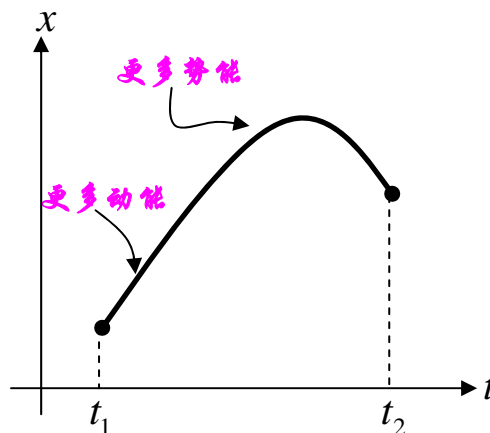


图 3

另一方面，你也不能够升得太快或跑得太远，因为这么一来你将会包罗过多的动能——你得很快达到高处使得你在可利用的剩下的那段规定时间里再落下来，到达目的地。所以你总要升得高些，但是也不应该升得太高。而事实证明，答案是粒子会在试图获得尽可能多的势能与尽可能少的额外动能之间取得某一种平衡——从而使得获得动能减去势能的差值尽可能小。

下面我将要证明粒子确实就是这样运动的，我们把动能减去势能并对时间的积分叫做作用量 S ：

$$\text{作用量} = S = \int_{t_1}^{t_2} (\text{动能} - \text{势能}) dt \quad (2)$$

记住势能和动能两者都是时间的函数。每一不同的可能路线你将获得关于这一作用量的不同值。我们的数学问题是去寻找使这个数值最小的那一条曲线。

这是一个崭新的课题，它与寻常微积分学中的极大和极小问题是很不一样的。你不能在算出了作用量之后取微分来能找出那个极小值。微积分中我们研究的对象称为函数 $f(x)$ ，或者说是这样一个法则 f ，给定自变量 x 的一个数值，这个法则就告诉我们另一个数值 $f(x)$ 。而需要去找出其中该函数是最小或最大的那个变量 x 的数值。例如，我们讲的势能函数，对于每一点就有一对应的势能，而我们必须去找出势能最低的那一点。但现在对于空间每一条路线我们是有一个数值的 $S[x(t)]$ ，而我们得去找出那一条会使该数值极小的空间路线 $x(t)$ 。实际上，我们现在要研究的乃是以路径作为“自变量”的函数，或者说它是函数的函数，因此我们将它称为泛函（推广的函数）——这是完全不同的一个数学分支，它并不是寻常的微积分，而是属于所谓变分学的范畴。

有许多问题属于这一种数学。例如，圆周往往被定义为所有与一固定点的距离为一定值的点的轨迹，但对于圆周还有另一个定义的方法：圆周乃是具有某给定长度而又能容纳最大面积的那一条曲线。对于某一周长来说，任何其他曲线所容纳的面积比圆周都要小。因此，若我们提出这样一个问题：试找出对某一给定周长能容纳最大面积的那条曲线，我们就有一个变分学问题——与我们所熟悉的有所不同的一种微积分。再举物理学中的两个例子。当光从空气进入玻璃时，光所行进的路线可以由 Snell 定律

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (3)$$

来确定。但是，你也可以这样确定路线，如果你计算光沿着某条曲线从一点到另一点所用的时间，那么正确的运动是沿着使得这个时间取最小值的那条路线。这就是一个变分的问题。再比如，不必通过给出关于场的一个微分方程，而是通过申明某一积分是一极大或极小值，就能够描述静电学。在电荷密度处处为已知的情况，问题就是要找出空间每一处的电势 ϕ ，你知道答案应该是：

$$\nabla^2 \phi = -\rho / \epsilon_0 \quad (4)$$

但陈述这同一事情的另一种办法却是：计算这么一个积分 U^* ：

$$U^* = \int \left[\frac{\epsilon_0}{2} (\nabla \phi)^2 - \rho \phi \right] dV \quad (5)$$

这是一个要对全部空间取的体积分，对于该正确的势分布 $\phi(x, y, z)$ ， U^* 是一极小值。

尽管这是一个不同于微积分的崭新的数学问题，但是，为了找出正确路线，看一下微积分中极值的含义是有益处的。比如，我有一个势能函数，那么你是怎么知道它在哪一点取极值的呢？当然，

你已经知道如果有一个极值点，那么势能对坐标的微分在这一点应该等于零，为什么呢？假设我们已经知道取极小值的点，现在当我使自变量在某种方式上稍为有点偏离时，函数就有一个与该偏离成正比的变化。这变化大概会使函数变得大些，否则我们就不会得到一个极小值了。但此时若该变化与偏离成正比，

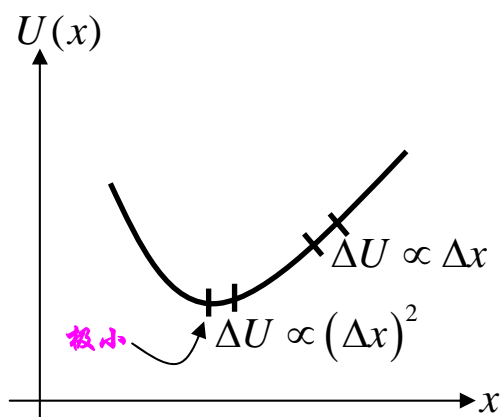


图 4

则倒转偏离的符号将会使函数变得较小。我们会使函数沿一个方向增加而沿另一个方向减少。唯一能够有一个真正极小值的办法是在第一次近似上不作任何改变，而改变乃是与对极小值点的偏离的平方成正比的。在该曲线的任何其他部分若移开原位置一个小距离，则函数之值也将按第一级变化。[图 4]。

现在，假设 $x(t)$ 代表该真实路线——即我们要试图寻找的那条曲线。取一

一条尝试路线 $\bar{x}(t)$ ，与该真实路线有一微小差别，这差别我们称之为 $\eta(t)$ ，即 $\bar{x}(t) = x(t) + \eta(t)$ ，作用量就有一个与该偏离成正比的变化（当然，应该是偏离 $\eta(t)$ 乘某个东西对时间的积分，不过这丝毫不会影响我们下面的讨论）。同样的，为了使得由 $x(t)$ 给出的作用量是一个极小值，这变化应该会使作用量变得大些。与前面同样的理由，若该变化与偏离成正比，则倒转偏离的符号将会使作用量变得较小。唯一能够有一个真正极小值的办法是在第一次近似上不作任何改变，而改变乃是与对极小值点的偏离的平方成正比的。也就是说，极小值有这么一个特点：若在第一级上离开该正确路线，则作用量与极小值的偏差只是属于第二级的。对于任何想象的路线，若移开一个小距离，则作用量之值也将按第一级变化。但在正确路线上，一个小的偏离在第一次近似上将不会产生任何差异[图 5]。

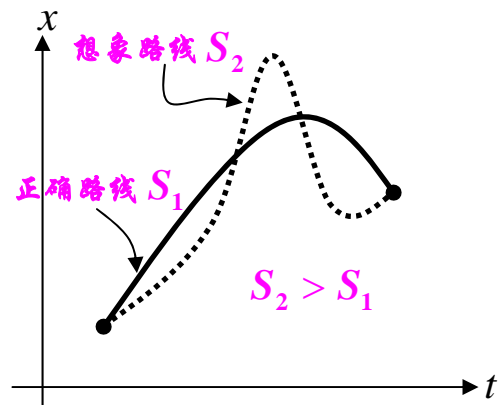


图 5

这就是我们将要用来算出那其实路线的办法。如果已有一条真实路线，那么在一条与之只有小小差别的曲线上的作用量在第一次近似上将不会造成任何差别。若确实有一个极小值的话，则任何差别都将落在第二次近似上。

现在我们的想法是：若对路线 $\bar{x}(t)$ 计算该作用量 \bar{S} ，则这个与我们对路线 $x(t)$ 所算出的作用量 S 之差，即 \bar{S} 与 S 之差，在小 η 的第一次近似上应等于零。这差别可以是第二级的，但在第一级上这差别应该为零。而且，这对于任何一个 η 都应该是正确的，只要它是足够小的。不过，还不仅如此。当然还有一个条件，那就是你所考虑的各路线都有彼此相同的起点和

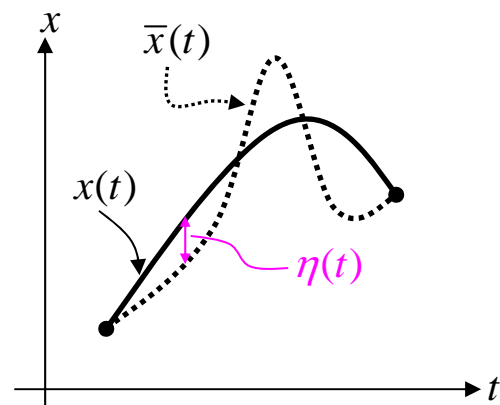


图 6

终点——每条路线都在时刻 t_1 从某点出发而在时刻 t_2 到达某另一点，这些地点和时间都保持固定不变。因此我们的偏离 η 在每一端都应等于零，即 $\eta(t_1) = \eta(t_2) = 0$ 。有了这个条件，我们的数学问题才告确定。

于是我们的打算是，把 $\bar{x}(t) = x(t) + \eta(t)$ 代入该作用量公式中

$$\bar{S} = \int \left[\frac{m}{2} \dot{\bar{x}}^2 - U(\bar{x}) \right] dt \quad (6)$$

式中 $U(\bar{x})$ 代表势能。而 $\dot{\bar{x}}$ 这个微商当然就是 $x(t)$ 的微商再加上 $\eta(t)$ 的微商，所以对于作用量会得到这么一个表达式：

$$\bar{S} = \int \left[\frac{m}{2} (\dot{x} + \dot{\eta})^2 - U(x + \eta) \right] dt \quad (7)$$

现在我们必须写得更详尽些。对于该平方项我们得到

$$\dot{x}^2 + 2\dot{x}\dot{\eta} + \dot{\eta}^2 \quad (8)$$

我们还需要一个在 $x + \eta$ 上的势能 U 。我认为 η 是小的，因而可以将 $U(\bar{x})$ 写成 Taylor 级数。它近似地等于 $U(x)$ ，在次一级近似上（按微商的通常性质）则该修正应该是 η 乘以 U 对 x 的变率，如此等等：

$$U(x + \eta) = U(x) + \eta U'(x) + \frac{\eta^2}{2} U''(x) + \dots \quad (9)$$

为了简化书写，我已将 U 对 x 的微商写成 U' 。可是对高于第一次幂之项我们并不在意，因而将所有含有 η^2 和更高次幂的项都取出来并放进一个称为“二阶和更高阶项”的项中（记为 $O(\eta^2)$ ）。从积分中的这些高阶项我只得到属于第二次或者更高次幂的，而我们便不须对之操心了。将所有这一切都凑合起来，得到：

$$\bar{S} = \int_{t_1}^{t_2} \left[\frac{m}{2} \dot{x}^2 - U(x) + m\dot{x}\dot{\eta} - \eta U'(x) + O(\eta^2) \right] dt \quad (10)$$

现在，如果你观察得仔细些，便会见到我在这里安排好的头两项相当于用那真实路线 x 所计算出来的作用量 S ，而我要集中注意力的东西乃是 S 的变化—— \bar{S}

与对该正确路线所应得的 S 之间的差别。当我们丢开那些“二阶和更高阶项”时，就将这一差值写成 δS ，并称之为 S 的变分。【这类似于微积分中，如果丢开 $f(x + \varepsilon) - f(x)$ 中 ε 的那些“二阶和更高阶项”，或者说当 $\varepsilon \rightarrow 0$ 时，那么我们将此差值称为函数 $f(x)$ 的一阶微分，并记为 $df(x)$ ，或简单地写为 df 。】因此，我对于 δS 便有

$$\delta S = \int_{t_1}^{t_2} [m\dot{x}\dot{\eta} - \eta U'(x)] dt \quad (11)$$

现在的问题是：这里有某一积分。虽然我还不知道 x 是什么，但我确实知道不管 η 是什么，这一积分总应该等于零。你试想想，这可能发生的唯一途径就是那个乘上 η 的东西应当是零。可是那含有 $\dot{\eta}$ 的第一项又该怎么办呢？如果 η 终归可以是一任意的东西，那它的微商也该是任意的了，因而你可以断定 $\dot{\eta}$ 的系数也应当等于零。但这可就不对了。之所以不对是因为在 η 与它的微商之间存在某一种关系；它们并非完全互相独立。

这里我们就会遇到在变分学中解决一切问题总要用到的同一普遍原则。即首先对你所要变化的东西作一个偏离（象我们上面通过如上 η 而做到的那样），目的是寻找第一级的项；然后又总要把积分安排成含有“某种东西乘以该偏离 η ”，而又不含有其他微商（没有 $\dot{\eta}$ ）的那种形式。因此，我们必须重新安排以便得到“某一件东西”乘以 η 。过了一会你将会看出这样做的巨大价值。

怎样才能将该 $\dot{\eta}$ 项重新安排使其含有 η 呢？回答是通过分部积分就可以做到。事实证明，变分学的全部巧妙就在于包括写下 S 的变分，然后利用分部积分就可以将 η 的微商清除掉。在微商会出现的每一问题中总是采取这一办法的。

如果你有任一个函数 f 乘以 $\dot{\eta}$ 并对 t 积分，你可以写下 $f\dot{\eta}$ 的微商：

$$\frac{d}{dt}(f\dot{\eta}) = \dot{\eta}f + f\ddot{\eta} \quad (12)$$

你所要积分的是对末一项而积的，因此

$$f\dot{\eta} = \frac{d}{dt}(\eta f) - \dot{f}\eta \quad (13)$$

在我们关于 δS 的公式中，该函数 f 就是 $m\dot{x}$ ；因此，我得到下列关于 δS 的公式：

$$\delta S = m\dot{x}\eta(t)\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt}(m\dot{x})\eta(t)dt - \int_{t_1}^{t_2} U'(x)\eta(t)dt \quad (14)$$

首项应该是在 t_1 和 t_2 两限上算出来的。然后我还必须有那个从分部积分剩下的积分。末项则是照抄下来的，没有什么改变。

现在碰到一件总会发生的事情——被积出的那一部分不见了。（事实上，如果该被积出的部分不消失，那你就应当重申该原理，加上一些条件而保证其如此）我们已经说过，在路线两端 η 必须是零，因为该原理要求只有在该变化曲线开始并终结于那些选定点上时作用量才是一极小值。这条件就是 $\eta(t_1) = \eta(t_2) = 0$ 。所以该项积分结果为零。我们将其他各项都凑集起来并得到：

$$\delta S = \int_{t_1}^{t_2} [-m\ddot{x} - U'(x)]\eta(t)dt \quad (15)$$

S 的变分现在就成为我们所希望得到的形式了——在该方括号内的各项全都乘上了 $\eta(t)$ 并从 t_1 积至 t_2 。

我们得到了某种东西乘以 $\eta(t)$ 的积分总等于零，而不论 η 是什么，结果始终为零。这意味着这个东西就是零。为什么呢？你可以想象这样对于任意做够小的 $\eta(t)$ 都等于零的积分

$$\int_{t_1}^{t_2} F(t)\eta(t)dt = 0 \quad (16)$$

这意味着当你将 $\eta(t)$ 取为 $\alpha F(t)$ 时（其中 α 是一个足够小的数），上面的关系只有在 $F(t)$ 始终恒等于零的情况下才是成立的。

因此，只有对于满足这复杂微分方程

$$-m\ddot{x} - U'(x) = 0 \quad (17)$$

或者说

$$-U'(x) = m\ddot{x} \quad (18)$$

的那条路线，该作用量积分才将是一极小值。这个方程你是见过的，它不过就是 $F = m\ddot{x}$ 罢了。因为左边势能微商的负值，那就是力。

所以，至少对于一个保守系统来说，我们已证明了最小作用量原理会给出该正确答案；也就是说，具有最小作用量的那条路线就是满足了 Newton 定律的那条路线，就是粒子真实的运动路线。

上面的论述实际上并未证明过它是一个极小点——也许是一个极大点。事实上，它的确不必是一个极小点。所谓“最小”我们确实指的是，当你改变路线时， S 值的第一级变化为零。并不要求是一个真正的“极小”。在物理学中当你遇到，某个“最小”原理的时候，通常所指的都是这个含义。

如果足够细心的话，你还会发现上面的推导在一个小的细节上看起来不那么正确，至少在数学上看来似乎不是特别严格。那就是，在我丢掉的“二阶和更高阶项”中包含着正比于 η^2 这样的项，而你知道， η 是小量并不一定意味着 $\dot{\eta}$ 也是一个量。因此丢掉的正比于 $\dot{\eta}^2$ 的这一项并不见得就比我保留下来的那些量来得小，确实如此。但是，如果我们取这样的想象路径，它不仅与真实路径的差别 η 很小，而且在任一时刻它们速度的差别 $\dot{\eta}$ 也是很微小的，那么，上面的推导就没有任何问题了。在物理上我们总是这样做的：如果在推导过程中遇到某些任意函数，人们通常总是假定它们有足够好的“行为”。在前面的推导中，我们这里所说的足够好的“行为”首先就是指 η 是可微的，其次，还必须要求 $\dot{\eta}$ 也是很微小的。实际上，如果你愿意，你可以假设 η 有任意阶的导数，而且其各阶导数都是很小的。重要的不是这些数学的细节，而是这样的假定能保证我们得到正确的物理结论！

现在，我要对最小作用原理作一些讨论。这个原理宣称从一处到另一处的某一积分是一极小值——这会告诉我们有关全过程的某种东西——它是关于整个

路线的全面描述。它与一个宣称当你沿路线行进时，有一个力在使它加速的定律相比在其特性上是有很大差别的。现在，我想要解释的第一件事是：在有了这一类最小作用原理时，为什么还会有微分定律。原因是这样的：试考虑在时间和空间中的那条实际路线。沿这一真实路线， S 是

一极小值。让我们假定已有了该真实路线，而它在空间和时间上既通过某一点 a ，又通过另一附近之点 b [见图 7]。现在，如果从 t_1 至 t_2 整个积分是一极小值，那么沿 a 至 b 的那一小段该积分也就应该是一极小值。因为不然的话，你仅仅通过把这一部分路程拨来拨去就可使整个积分值稍为降低一点。

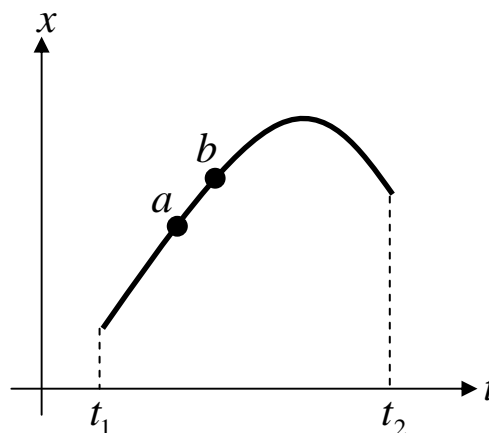


图 7

所以在这条路线中的每一小段上也应该都是一个极小值。并且不管该小段如何短，这都必然正确。因此整个路线会给出一个极小值的原理，也可说成是路线上的每一无限小线段也各具有能给出一个最小作用量的那种曲线。现在若我们取路线上足够短的一段——在十分靠近的 a 与 b 两点之间——那么，因为你在那整整一小段路线上几乎总是呆在同一个地点，在遥远处势能如何逐点变化就是无足轻重的了，你必须加以讨论的唯一东西就是势能中的第一阶变化。答案只能取决于势能的微商而不是在各处的势能。所以关于整条路线的总性质的陈述就变成对一小段路线会发生的事情的陈述——一种微分式描述。而这一种微分式描述就仅仅涉及到势能的微商，也就是在该点上粒子受到的力。这就是在总体定律与微分定律之间的关系的一种定性解释。

我想要说明的另一件事是：从微分的观点，粒子的运动是容易理解的。由于粒子受到力的作用，所以它的速度发生了变化，又由于粒子有速度，所以位置也改变了。换言之，每一时刻当粒子获得一加速度时仅仅知道在该时刻应该做什么。可是如果你讲粒子会做出决定以选取将能给出最小作用量的那条路线，这就完全是另一回事了。粒子怎么能知道周围其他的路线作用量来得要更大呢？或者说粒子会对不同路线的作用量进行比较吗？答案是肯定的！这必须用量子的观点来解释。整个量子力学（对于非相对论情况，并略去电子自旋）是这样工作的：一个

粒子从时刻 t_1 所处的点 **1** 出发，将在 t_2 到达点 **2** 的几率，等于一个几率幅的平方。每一个可能路线都会对几率幅有贡献，而总幅则是所有这些贡献之和。对于每个

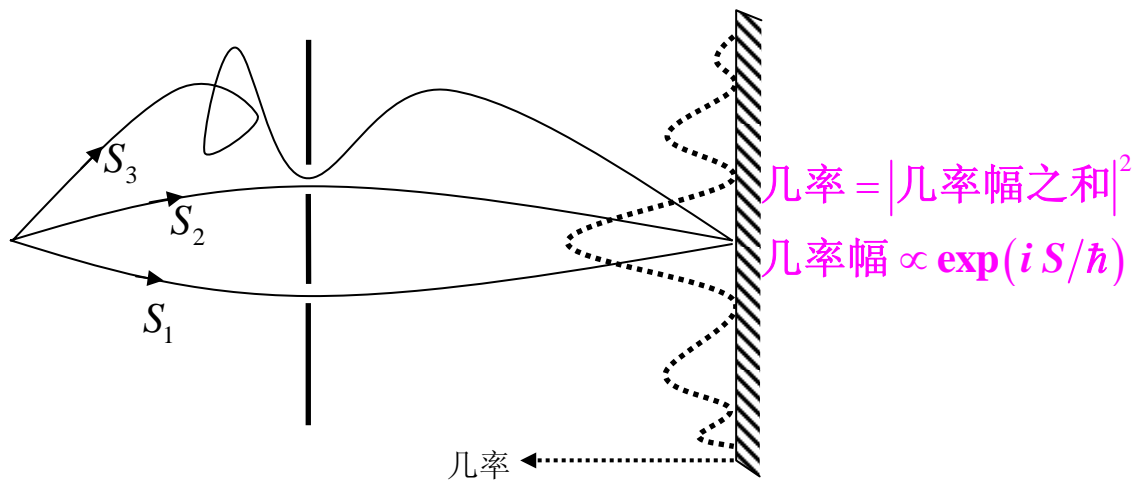


图 8

我们能想象出来的轨道 $x(t)$ ，就得算出一个幅来，然后再把它们都相加起来。上述的作用量积分就告诉我们关于一单独路线所应有之幅如何，它正比于某一常数乘 $\exp(iS/\hbar)$ ，其中 S 就是对于该路线的作用量。这就是说，如果我们用一复数来表示幅的周相，那相角就是 S/\hbar 。作用量 S 具有能量乘时间的那种量纲，而普朗克常数 \hbar 也具有这同一量纲。它是判定量子力学在什么时候才显得重要的一个常数。

这里就是它工作的原理：假设对所有路线， S 比起 \hbar 来都很大（譬如说点 **1** 和点 **2** 相距很远，或者说当所考察的空间尺度很大时，另一种情况，如果质量很大时）。当然此时每一路线都会贡献某一定之幅。但是，因为 S 非常巨大，即使一个小小的 S 变化也就意味着一个完全不同之相——因为 \hbar 竟是那么小。所以在取该总和时，互相靠近的路线一般都会将其效应互相抵消——除了一个区域以外，而这个区域就是当一条路线与其邻近路线在第一次近似上全都会给出同一个周相时（更准确地说，在 \hbar 范围内的同一个作用量）。只有那些路线才算重要。因此，在普朗克常数 \hbar 趋于零的极限情况下，正确的量子力学可以总结成简单的一句话：“忘记所有这些几率幅吧。该粒子会在一条特殊路线上行进，那就是在

其中 S 不会在第一次近似上发生变化的路线”。这就是最小作用原理与量子力学之间的关系。我刚才讲粒子会选择路径或者说会对不同的路径进行比较，现在我们看到，这种选择、比较是通过位相的相消来实现的。

接下来，我将说明在前面推导过程中的所遇到的一些量的正式的概念。首先，那个经过对时间积分就可以得到作用量 S 的函数称为 Lagrange 函数，在抛物体的例子中，它只是粒子的速度和位置的函数，而对于其他更加一般的情形，它还可能明显地依赖于时间，因此，Lagrange 函数通常用

$$L = L(x, \dot{x}, t) = T - U \quad (19)$$

表示。其次，尝试路线 $\bar{x}(t)$ 与真实路线 $x(t)$ 的微小差别 $\eta(t)$ 称之为 $x(t)$ 的变分，通常记为 $\delta x(t)$ ，或简单地写为 δx [尽管 $x(t)$ 是一个函数，而这里我们将其变化记为 δx 而不是 dx ，其原因在于这种变化并非由自变量 t 的改变所引起，它是函数 $x(t)$ 本身作为整体的一个变化]，至于 $\dot{\eta}(t)$ 则记为 $\delta \dot{x}$ ，即

$$\delta \dot{x} \equiv \frac{d(\delta x)}{dt} \quad (20)$$

这是 $\delta \dot{x}$ 的确切含义。很多书上也把上式解释为

$$\delta \frac{dx}{dt} = \frac{d}{dt} \delta x \quad (21)$$

并把它作为变分运算的一个法则，也就是说：对于 $x(t)$ 来说，其对时间的微商与变分这两种运算是可以交换的。

利用这些记号，最小作用原理也就可以写成

$$\delta S = \delta \int_{t_1}^{t_2} L(x, \dot{x}, t) dt = 0 \quad (22)$$

这个 δS 是指当两个相邻路径的作用量差值的一阶项。当然我们要求

$$\delta x(t_1) = \delta x(t_2) = 0 \quad (23)$$

由于

$$\begin{aligned}\delta S &= \delta \int_{t_1}^{t_2} L(x, \dot{x}, t) dt = \int_{t_1}^{t_2} [L(x + \delta x, \dot{x} + \delta \dot{x}, t) - L(x, \dot{x}, t)] dt \\ &= \int_{t_1}^{t_2} \delta L(x, \dot{x}, t) dt\end{aligned}\quad (24)$$

而括号中的这一项正是 Lagrange 函数 L 的变分，即

$$\delta L = L(x + \delta x, \dot{x} + \delta \dot{x}, t) - L(x, \dot{x}, t) \quad (25)$$

你也可以将方程(24)看作变分的另一个法则：变分与积分运算时可以交换的，当然，你知道它不过是作用变分的定义而已。与前面同样的理由，函数 L 的这种变化起源于路径 $x(t)$ 本身的改变，而不是来自自变量 t 的改变，因此我们不把它写作 dL ，在变分运算中， t 仅仅起着参量的作用。另外一点，在微积分中，当你看到记号 dx 时，那就意味着这是一个无穷小量，也就是说，当你求此时函数 $f(x)$ 的变化时是要取极限的，或者说，你仅仅需要保留 dx 的那些一阶项。类似地，在变分的计算中， δ 这个符号总是意味着所作的运算应该仅仅保留 δx 的一阶小量，或者说 δx 是一个无穷小的函数。因此，当我们把方程(25)右边第一项按照 δx 和 $\delta \dot{x}$ 写成 Taylor 级数并且只保留到一阶项时，便得到了

$$\delta L = \frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} \quad (27)$$

这里我们发现了变分 δ 的一个最重要的运算法则，那就是：变分 δ 作用于函数上时，其规则几乎与微分 d 的规则相同，即同样的满足链式法则，差别仅仅在于前者后者不作用在时间 t 上，或者说 δt 是等于零的。具体一点，对于函数 $L(x, \dot{x}, t)$ （当然你也可以取任意别的函数），你知道它的微分等于

$$dL = \frac{\partial L}{\partial x} dx + \frac{\partial L}{\partial \dot{x}} d\dot{x} + \frac{\partial L}{\partial t} dt \quad (28)$$

如果把所有微分符号“ d ”换成变分符号“ δ ”，并令 δt 等于零，那么我们就得到了函数 L 的变分。

稍微要加以解释的是，尽管 x 和 \dot{x} 是满足某种关系的，但是，在用 δ 或者 d 作用在函数 $L(x, \dot{x}, t)$ 时，我们是将它们当作独立变量处理的。这样做有道理吗？举个例子最容易明白。某个函数，比如 $f(x) = x^3$ 的微分我们知道是 $df = 3x^2 dx$ ，但是，你完全可以把这个函数写为 $f(x, x^2) = x \cdot x^2$ ，并在求微分时把 x 和 x^2 都看作独立变量：

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial x^2} dx^2 = x^2 dx + x dx^2 = 3x^2 dx \quad (29)$$

结果当然是一样的。

继续我们前面的讨论。现在利用变分符号以及 Lagrange 函数我可以将作用量的变分写为

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} \right) dt \quad (30)$$

接下来我们需要把括号中的项变为 δx 乘以某个东西，你还记得 $\delta \dot{x}$ 的含义就是

$\frac{d}{dt}(\delta x)$ ，因此，对第二项做分部积分我们就得到

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) \delta x dt + \left. \frac{\partial L}{\partial \dot{x}} \delta x \right|_{t_1}^{t_2} = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) \delta x dt \quad (31)$$

)

同样的，完全积出的项（所谓端点项）等于零。由于上面的积分对于任何的 δx 都等于零，因此，我们就得到了，满足微分方程

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = 0 \quad (32)$$

的路线就是那条使得作用量取最小值的路线，也就是那条真实的运动路线。

当然，以上所作的一切不过是重复我们以前的工作，不过在这个过程中仍然得到了一些非常重要的东西。首先，我们得到了一个直接用 Lagrange 函数表示的方程(32)，这个方程就称为 Lagrange 方程，你很快就会领略到将动力学方程写

成这种形式所带来的巨大优势。其次，同样特别重要的，我们得到了变分“ δ ”运算法则，总结一下，

1. 当作用于路径 $x(t)$ 上时， δ 与 d/dt 可交换，即

$$\delta \dot{x} = \delta \frac{d}{dt} x = \frac{d}{dt} \delta x$$

实际上，你也很容易证明，对任意的函数这个关系都是成立的，即

$$\frac{d}{dt} \delta f = \delta \frac{d}{dt} f$$

2. δ 与积分运算可以交换。
3. δ 作用于函数上时满足链式法则，即

$$\delta f(x, \dot{x}, t) = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial \dot{x}} \delta \dot{x} + \frac{\partial f}{\partial t} \delta t$$

4. δ 对时间没有作用，即 $\delta t = 0$ ，因此

$$\delta f(x, \dot{x}, t) = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial \dot{x}} \delta \dot{x}$$

在微积分中，我们有了微分“ d ”的基本性质或法则之后，对于任意函数的微分就可以很容易地得到，而不必再求助于微分的定义通过极限来计算了。了解了这一点，我想你也就很容易明白上面这些法则的重要性了。

现在我们看一下，怎样将 Lagrange 方程变为前面所得到的 Newton 方程。对于势能为 $U(x)$ 的作一维运动的粒子，其 Lagrange 函数为

$$L = T - U = \frac{1}{2} m \dot{x}^2 - U(x) \quad (33)$$

因此

$$\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x} \quad (34)$$

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x} \quad \Rightarrow \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \frac{d}{dt}(m\dot{x}) = m\ddot{x}$$

你看到 Lagrange 方程实际上就是

$$m\ddot{x} = -\frac{\partial U}{\partial x} = F \quad (35)$$

正是 Newton 方程。

在得到 $\frac{d}{dt} \frac{\partial L}{\partial \dot{x}}$ 这个过程中，我们是否可以先计算 $\frac{dL}{dt}$ ，然后再求它对于的偏

导数呢？试着做一下，拿我们前面抛物体的运动为例，此时 $L = \frac{1}{2}m\dot{x}^2 - mgx$ ，

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \frac{d}{dt}(m\dot{x}) = m\ddot{x} \quad (36)$$

而

$$\frac{\partial}{\partial \dot{x}} \frac{dL}{dt} = \frac{\partial}{\partial \dot{x}}(m\dot{x}\ddot{x} - mg\dot{x}) = m\ddot{x} - mg \quad (37)$$

因此，一般来说，

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \neq \frac{\partial}{\partial \dot{x}} \frac{dL}{dt} \quad (38)$$

这一点是要特别提醒的！

最后，我想就上述问题谈论一些普遍化可能性。第一，可以推广至三维，即不只是 x ，我宁愿有 x_1 、 x_2 和 x_3 各作为 t 之函数；此时作用量更为复杂些。对于三维运动，你必须用到完整的动能式—— $m/2$ 倍于整个速度的平方。这就是说，

$$T = \frac{1}{2}m\dot{x}_i\dot{x}_i \quad (39)$$

并且，势能也是 x_1 、 x_2 和 x_3 的函数。而路线究竟如何呢？路线是在空间中某一

条普遍曲线，不容易画出来的，但意思却都是一样。不过以前的 δx 又是怎么回事？噢，这时候它应该有三个对应的分量。你可以在 x_1 、 x_2 或 x_3 方向上——也可以同时所有三个方向上移动路线。所以 $\delta \vec{r} = (\delta x_1, \delta x_2, \delta x_3)$ 该是一个矢量。然而这样做实际上并未曾把事情弄得过于复杂。由于只有第一阶的变分必须为零，我们便可以通过三个连续移动而进行计算。可仅仅在 x_1 方向上移动 δx_1 ，而申明它的系数应等于零。这样就得到一个方程，然后又在 x_2 方向上移动而得到另一个方程。又在 x_3 方向上得到第三个。当然或者按照你所喜欢的任一种次序进行。无论如何你已得到了三个方程。而牛顿定律实际上就是在三维中的三个方程——对每一分量就有一个。我想你实际上能够明白，这一定会行得通，但这个三维问题仍留给你自己去作证明。顺便提一下，你本来也可以来用任一种你所喜欢的坐标系，诸如极坐标或其他坐标，从而看出如果在半径、角度、或其他坐标上各有一个移动将会发生的事情，便会立即得出适用于该坐标系的牛顿定律。

举一个例子。考虑各向同性平面（二维）谐振子的运动。这时动能

$$T = \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2) \quad (40)$$

而势能

$$U = \frac{1}{2}k(x_1^2 + x_2^2) \quad (41)$$

因此 Lagrange 函数

$$L = T - U = \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2) - \frac{1}{2}k(x_1^2 + x_2^2) \quad (42)$$

每个坐标（这里我们先用笛卡尔坐标）对应一个 Lagrange 方程，所以运动方程为

$$0 = \frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = -kx_i - \frac{d}{dt}(m\dot{x}_i) \quad (43)$$

即

$$m\ddot{x}_i + kx_i = 0, \quad (i = 1, 2) \quad (44)$$

这两个方程对应两个有相同频率(这正是各向同性的含义)的 1 维谐振子的运动。

但是，你完全可以采用别的坐标系，比如极坐标系，Lagrange 函数用极坐标表示就是

$$L = T - U = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{1}{2}kr^2 \quad (45)$$

而 r 和 θ 的运动仍然遵循 Lagrange 方程，即

$$\frac{\partial L}{\partial r} - \frac{d}{dt} \frac{\partial L}{\partial \dot{r}} = 0, \quad \frac{\partial L}{\partial \theta} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} = 0 \quad (46)$$

把 Lagrange 函数代入就得到了 r 满足的方程

$$mr\dot{\theta}^2 - kr - m\ddot{r} = 0 \Rightarrow m(\ddot{r} - r\dot{\theta}^2) - kr = 0 \quad (47)$$

它正是我们熟悉的 Newton 方程的径向部分，而 θ 满足的方程则是切向的 Newton 方程

$$\frac{d}{dt}(mr^2\dot{\theta}) = 0 \Rightarrow m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = 0 \quad (48)$$

这是一个中心力问题，粒子受到的力的切向分量为零。上面第一个式子括号中的量实际上正是粒子相对于力心的角动量，因此，这个方程的含义是讲：角动量是一个守恒量！

同样，这一方法也可推广至任何数目的粒子。例如，如果你有 n 个粒子，在它们之间作用着力，因而就有一个相互势能，那么你只要将这 n 粒子的动能相加并取它们间的互作用作为其势能。对此你想要变化什么东西呢？势必变化 n 个粒子的路线（第 a 个粒子位置的变化为 $\delta\vec{r}_a$ ）。于是，对于在三维中运动的 n 个粒子的体系，就总共有 $3n$ 个方程。你可以在 x_1 方向、 x_2 方向和 x_3 方向上分别变更每一个粒子的位置，因而就有 $3n$ 个方程。而这正是理该如此的。其中每一个粒子在受力作用时的加速度都由三个方程确定。

再进一步，这种方案是否还可以推广到连续体系呢？答案是肯定的，不过有关这方面的内容我将在这门课的最后再来讨论。实际上，我们还可以推广到相对论情形，这种推广并不难做到，因此，在这一章的后面我们就会来处理它。

实际上，我们还可以推广到相对论情形，例如具有势能 $U(\vec{r})$ 的粒子的运动使得下面这个 Lagrange 函数

$$L = -m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} - U(\vec{r}) = -m_0 c^2 \sqrt{1 - \frac{\dot{x}_i \dot{x}_i}{c^2}} - U(x_1, x_2, x_3)$$

的积分取最小值。这个作用量积分的第一部分是粒子的静质量 m_0 乘以 c^2 再乘以对于速度函数 $\sqrt{1 - v^2/c^2}$ 的积分。我将把这一作用量公式事实上确能给出那些正确的相对论性运动方程这一问题的证明留给你们中作为练习。实际上，你将得到运动方程 $d\vec{p}/dt = -\nabla U$ 的各分量，其中你会记起 $\vec{p} = m\vec{v}/\sqrt{1 - v^2/c^2}$ 。

我想要强调在一般情况下，比如在相对论公式中，那作用量的被积函数不再具有动能减去势能的形式。那只有在非相对论性的近似上才正确的。例如， $m_0 c^2 \sqrt{1 - v^2/c^2}$ 这一项就不是我们所称之为动能的了。对于在任一特定情况下作用量应该是什么的问题必须通过某种尝试的办法来确定。这与首先确定什么是运动定律的问题恰好相同。你只要对所已知的一些方程盘来盘去，看你能否把它们纳入于最小作用量的形式之中。

我刚才说过，我们得到了牛顿定律。这并非十分正确，因为牛顿定律还包括象摩擦一类的非保守力。牛顿说 ma 等于任何 F ，可是最小作用原理却只适用于保守力系——那里所有的力都可以从一个势能函数获得。然而，你知道，在微观的一级——即在物理学最深入的那一级——并没有非保守力。象摩擦力那样的非保守力之所以出现乃是由于我们忽略了微观上的复杂性——实在有太多的粒子有待分析。而所有基本的相互作用都是保守的，都可以置于最小作用原理的那种形式上面。另一个在经典力学中很重要但这里我们还未加讨论的问题是体系运动受限的情况，也就是所谓的约束问题。譬如限制在球面上的粒子的运动。不过，假设你能在原子的水平上观测一下这个问题，那就会发现它们并不是紧贴在一起的，粒子在运动过程中并不是严格地限制在球面上，而是不停地上下起伏，也就是说，在原子的尺度上，并没有也不可能有什么约束存在，究其根源是因为 Heisenberg 测不准原理。实际上，当你想要在原子的尺度上描述一个宏观物体的时候，就会发现你甚至连“物体”究竟指的是什么都无法说个明白，因为每一时

刻都有一些原理脱离或者进入“物体”！

但是，即便对于这两种情况，我们仍然可以将它纳入到变分原理的框架中。在后面的内容中，我将把注意力放在所有这些前面提到的推广上（暂不考虑相对论效应）。

附录：关于“最小”的一点说明

对于粒子在重力场中的运动（为简单起见，我仍将假定粒子只有垂直方向的运动），方程(7)所给出的任一假象路线 $\bar{x}(t)$ 的作用量为

$$\begin{aligned}
 \bar{S} &= \int_{t_1}^{t_2} \left[\frac{1}{2} m (\dot{x} + \dot{\eta})^2 - mg(x + \eta) \right] dt \\
 &= \int_{t_1}^{t_2} \left[\left(\frac{1}{2} m \dot{x}^2 - mgx \right) + (m\dot{x}\dot{\eta} - mg\eta) + \frac{1}{2} m \dot{\eta}^2 \right] dt \\
 &= S + \int_{t_1}^{t_2} (m\dot{x}\dot{\eta} - mg\eta) dt + \int_{t_1}^{t_2} \frac{1}{2} m \dot{\eta}^2 dt \quad (7) \\
 &= S + \cancel{m\dot{x}\eta} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} (\cancel{m\ddot{x}} + mg) \eta dt + \int_{t_1}^{t_2} \frac{1}{2} m \dot{\eta}^2 dt \\
 &= S + \int_{t_1}^{t_2} \frac{1}{2} m \dot{\eta}^2 dt
 \end{aligned}$$

也就是说，不管 η 具有什么样的形式，也不管它是否是一个“小”的偏离，想象路线的作用量总是比真实路线的要大。因此，在这个具体的例子中，真实路线的作用量确实是最小的。这里关键处在于势能是坐标的线性函数。而由于在一个点的充分小的领域中，势能函数是随着坐标线性变化的，由此我们可以得到一个结论：当起点和终点相距足够近时，确切地说，当时间间隔非常小时（你可以考虑一下粒子作周期性运动的情形），真实路线的作用量都是一个局部的极小值。当时间间隔很长时，真实路线的作用量通常是一个鞍点（扭转点），也就是说，与真实路线的作用量比较，某些假想路线的作用量要大，某些要小，而另一些则会给出同样的数值。实际上，当时间间隔足够长时，真实路线的作用量除了在某些

势场（譬如重力场）中是一个真正的极小值外，在其他场合都会是一个鞍点，它不可能是一个极大值，换句话说，不可能任何一条想象路线给出的作用量都比真实路线的要小。这一结论有赖于势能 $U(x)$ 二阶导数，我想具体的讨论就留给你作为思考吧！

下面我们就来看一个真实路线的作用量取鞍点的例子。考察一个一维的谐振子，如果令 $k = m\omega^2$ ，那么它的周期就可以表示为 $\tau = 2\pi/\omega$ ，假设初始时粒子静止并处在距离平衡位置为 x_0 的地方，在一个周期之后它又回到了这个位置（你也可以考虑粒子初始时刻具有某一个速度的一般情形），真实的路线为

$$x(t) = x_0 \cos \omega t$$

现在，我们想象如下一些运动，它们都从点 x_0 出发又在时间 τ 后回到同一点 [图9]：

$$x_1(t) = x(t) + \alpha \sin \frac{1}{2} \omega t$$

$$x_2(t) = x(t) + \alpha \sin \omega t$$

$$x_3(t) = x(t) + \alpha \sin \frac{3}{2} \omega t$$

如果分别用 S 以及 S_i ($i = 1, 2, 3$) 表示真实路线和上面的三类想象路线的作用量（将每一时刻的动能减去势能，然后从 $t = 0$ 到 $t = \tau$ 在一个周期内积分），通过直接计算你就可以很容易得到

$$S_1 - S = -\frac{3}{8} \pi m \omega \alpha^2$$

$$S_2 - S = 0$$

$$S_3 - S = \frac{3}{2} \pi m \omega \alpha^2$$

也就是说，不管 α 是多大的非零常数，第一类路线的作用量总是比小，第三类路线的作用量总是比大，而第二类路线的作用量则恰好等于真实路线的作用量。当然，当 α 很小时，你可以看到想象路线与真实路线的作用量之差在 α 的第一

次近似上是等于零的。

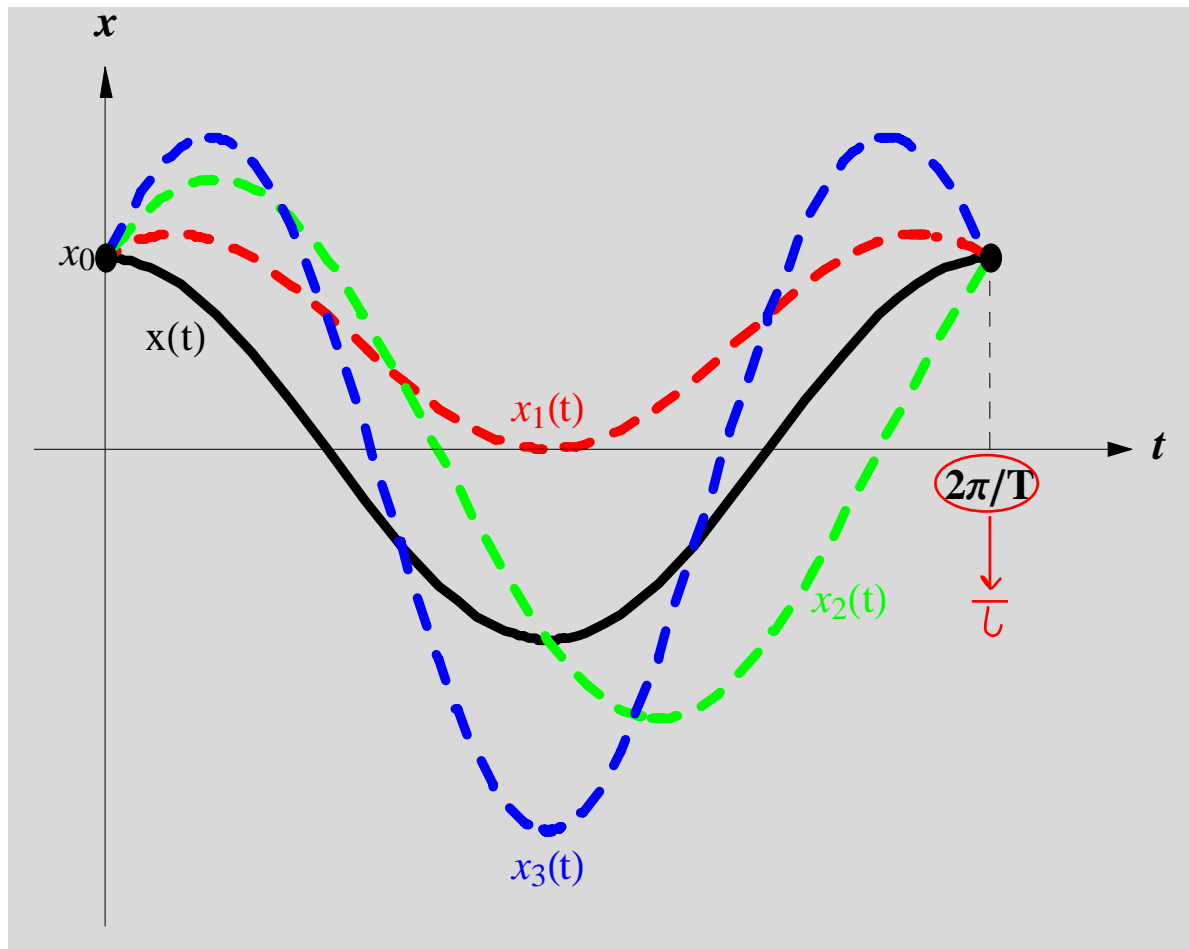


图9