最优控制理论

L.D.伯科维茨 著

上海科学技术出版社

最优控制理论

L. D. 伯科维茨 著

贺建勋 连瑞兴 黄伙泉 曾昭磐 蔡维旋 译

上海科学技术出版社

L. D. Berkovitz

OPTIMAL CONTROL THEORY

Springer-Verlag · 1974

最优控制理论

L. D. 伯科维茨 著

贺建勋 连瑞兴 黄伙泉
曾昭磐 蔡维璇 译

上海科学技术出版社出版
(上海瑞金二路 450 号)
新华书店上海发行所发行 江苏扬中印刷厂印刷
开本 787×1092 1/32 印张 7.375 字数 194,000
1985 年 5 月第 1 版 1985 年 5 月第 1 次印刷
印数：1—10,800
统一书号：13119·1216 定价：1.65 元

原　序

本书介绍由常微分方程描述的受控过程的最优控制数学理论，它是为数学方面和应用领域里的学生和有关专业人员写的，他们对于这个课题及其跟应用的联系需要有一本广泛却又比较深刻、简明而又清晰的入门书．为了适应读者之间数学兴趣和基础知识的差别，我们把材料安排得可以使读者省略书中较高深的数学部分而不损害其理解的连贯性．对那些主要兴趣在于应用方面的读者，我们建议至少阅读第一章，第二、三、四章中引言部分推荐的各节，以及第五章全部．在每一章的引言部分都有进一步指导个别读者阅读他感兴趣的材料的说明．对于学习过相当多高等微积分教程的读者，应能理解本书中的定义和定理的表述，也应能接受其中的大部分数学推导．任何精通 Lebesque 积分和泛函分析基本概念的人，他们是能够读懂全书的．

对于希望在应用上了解更多的读者，我们建议他们阅读列在书末文献目录中的参考文献[2]、[13]、[33]、[35]和[50]．对于希望学习更多有关数学概念和某些本书没有讨论的专题的读者，可看参考文献[27]、[28]、[33]、[48]、[50]、[59]和[61]．

书中的定理，需要标明号码的方程、公式、不等式和定义等等，按下叙方式采用小数点编号：定理 III.7.2 表示第三章第 7 节中的定理 2．在第三章之外引用这个定理时，记为“定理 III.7.2”；在第三章中提到它时，简记为“定理 7.2”．需要标明号码的公式、方程、不等式等，也都采用类似的记号．

（下为誌谢部分，译略．）

L. D. 伯科维茨

美国印第安纳州，西拉法伊蒂

1974 年 8 月 5 日

译 者 的 话

本书是在1978年的译稿的基础上，由参加翻译的同志再进行认真修改，经集体讨论，然后统一定稿的．参加翻译的同志有贺建勋（序、第一章），连瑞兴（第二、五章），黄伙泉（第三章），曾昭磐（第四章），蔡维璇（第六章）；最后，由贺建勋、连瑞兴负责全书的统一定稿工作．李文清教授对原译稿进行过初校，黄淑兰同志参加过第五章原译稿的部分翻译工作．译稿曾作为厦门大学研究生和进修教师的教材，参加学习的同志对译稿提出了许多宝贵的意见．译者感谢李文清教授的热情帮助和鼓励；对大力支持本书翻译工作的厦门大学最优控制与常微分方程教研室的其他教师、进修教师和研究生亦表示深切的谢意．

本书译稿虽几经统一校阅，但限于水平，难免还会有错误和不妥之处，欢迎读者批评指正．

译 者

1981年4月

目　　录

第一章　控制问题的实例

1. 引　　言

近年来，在不同的领域内已出现了一类有着共同的数学表达形式的重要问题，这就是所谓控制问题。不管这些问题的现代起源如何，从数学的观点看，它们不过是一类已经进行了几百年研究的变分法问题的变形。

本章我们将介绍控制问题的一些实例，它们来自不同的应用领域。这样列举问题的目的，是为了说明控制问题起源的多样性，指出它们的重要性，以及诱导这些问题的数学表达形式。所举的这些实例不应该理解为我们的单子是完整的，或者认为我们在每个领域里已经选择了最有意义的问题。事实上，为了使举例不致于太复杂，我们尽量选取了十分简单的问题。

2. 生产计划问题

第一个问题是来自经济学的，它是一个生产计划方面的问题。设 T 是一个固定时间；$x(t)$ 表示在时刻 $t(0\leqslant t\leqslant T)$ 的商品存货量；$r(t)\geqslant 0$ 表示在时刻 t 时对商品的需求率，这里假定 $r(t)$ 是一个定义在 $0\leqslant t\leqslant T$ 上的时间 t 的已知连续函数；设 $u(t)$ 表示在时刻 $t(0\leqslant t\leqslant T)$ 的生产率，函数 u 由生产计划人员来选取，它就是生产计划或者叫做控制。我们将取 u 为 $0\leqslant t\leqslant T$ 上的分段连续函数。如果我们要求的所有条件被满足，则存货量 x 由微分方程

$$\frac{dx}{dt}=-r(t)+u(t),\ x(0)=x_0 \tag{2.1}$$

所确定，其中 x_0 是原来的库存水平，且 $x_0>0$。从 $x(t)$ 的实际意义来看，显然必须选取生产计划 u，使得对所有 $0\leqslant t\leqslant T$，有

$$x(t) \geqslant 0. \tag{2.2}$$

其次，因为要有一定的库存，而且生产能力受到工厂设备的限制，所以有理由要求函数 u 对所有的 $0 \leqslant t \leqslant T$ 满足约束条件：

$$0 \leqslant u(t) \leqslant A. \tag{2.3}$$

这里 $A>0$，它表示最大可能的生产率．一个满足(2.3)的生产计划 u，使得(2.1)的相应解在整个 $0 \leqslant t \leqslant T$ 上存在而且满足(2.2)时，我们就把它叫做容许计划或容许控制．

在这里自然发生这样的问题：即是否存在容许计划的问题．假如 A 足够地大，则确实存在容许计划．例如，若令

$$M = \sup[r(t): 0 \leqslant t \leqslant T],$$

且 $A>M$，则 $u(t)=A$ 就是一个容许计划．今后，我们将假定容许计划是存在的．

我们假定：每个单位时间的生产成本是生产率的函数 h．这样，在时刻 t 时的生产率是 $u(t)$，则单位时间的生产成本是 $h(u(t))$．设 $b>0$ 是单位时间贮藏单位商品的费用．于是，在时刻 t 时计算这个系统的单位时间的成本是

$$f(t, x(t), u(t)) = h(u(t)) + bx(t), \tag{2.4}$$

因此总成本由

$$C(u) = \int_0^T f(t, x(t), u(t))dt \tag{2.5}$$

给出．其中 $x(t)$ 是方程(2.1)与容许生产计划 u 相对应的解．我们采用记号 $C(u)$ 表示这个成本，因为当需求率 r 和开始的存货量 x_0 被指定以后，这个成本只依赖于函数 u 的选取．这里我们得到一个泛函的例子，它对于给定函数类中的每个函数 u 赋予一个确定的实数．

对于生产计划人员来说，问题是要选取一个容许控制 u，使得总成本 $C(u)$ 达到极小值．

3. 化 学 工 程

设 $x^1(t), \cdots, x^n(t)$ 表示反应器中几种物质在时刻 t 的浓度，

这 n 种物质在反应器中同时发生化学反应．设反应速率由微分方程组

$$\frac{dx^i}{dt}=G^i(x^1,\ \cdots,\ x^n,\ \theta(t),\ p(t)),\ x^i(0)=x_0^i,\ i=1,\ \cdots,\ n,\tag{3.1}$$

来决定，其中 $\theta(t)$ 是时刻 t 反应器内的温度，$p(t)$ 是时刻 t 反应器内的压力．我们能够在每个时刻控制温度和压力，而温度和压力满足约束条件：

$$\begin{aligned}&\theta_b\leqslant\theta(t)\leqslant\theta_a,\\&p_b\leqslant p(t)\leqslant p_a,\end{aligned}\tag{3.2}$$

其中 θ_a，θ_b，p_a 和 p_b 都是常数．它们分别表示温度和压力可以达到的最大值和最小值．

我们设反应持续进行了一段时间 T．在时刻 T 的浓度是 $x^1(T)$，$\cdots$，$x^n(T)$．与每种产品相联系的是一个经济价值或者叫做价格 c^i，$i=1$，$\cdots$，n．价格可以是负的，因为不合格的废品必须当作某种消耗来处理．于是，最终产品的价值是

$$v(p,\ \theta)=\sum_{i=1}^{n}c^ix^i(T).\tag{3.3}$$

当给定了一组初始浓度 x_0^i，假若函数 G^i 具有某些合适的性质，则最终产品的价格就可以根据函数 p 和 θ 的选取而完全确定．因此记号 $v(p,\ \theta)$是另一个泛函的例子，在这种情形，对于某个集合中的每一函数对$(p,\ \theta)$，我们得到一个确定的实数．

这里的问题是：要求在区间$[0,\ T]$上选取分段连续函数 p 和 θ，使其满足(3.2)，同时使得 $v(p,\ \theta)$取最大值．

上述问题的一个变形如下：不要求反应持续进行一个固定时间 T，而是当反应物之一，比如说 x^1 达到预先指定的浓度 x_f^1 时，即停止反应．这时终止时刻 t_f 不是预先固定的，而是方程 $x^1(t)=x_f^1$ 的最小正根．现在的问题是要使

$$v(p,\ \theta)=\sum_{i=2}^{n}c^ix^i(t_f)-k^2t_f$$

取最大值，其中项 k^2t_f 表示反应器运转的代价．

问题还有另一种变形：即当反应物中的几种反应物达到预先指定的浓度，比如说 $x^1=x_f^1, x^2=x_f^2, \cdots, x^j=x_f^j$ 时，停止反应。这时最终产品的价值是

$$\sum_{i=j+1}^{n} c^i x^i(t_f) - k^2 t_f.$$

我们注意到，在问题的最后两种变形中，当我们处理极大值问题之前，有另外的问题必须考虑，亦即使用容许函数类中的压力函数 p 和温度函数 θ 时，我们能否达到所要求的最终浓度。

4. 飞 行 力 学

在这个问题中，我们把火箭当作可变质量的点，而忽略掉它的转动惯量。假定火箭的运动发生在相对于某个固定(坐标)架的平面上，设 $y=(y^1, y^2)$ 表示火箭的位置向量，$v=(v^1, v^2)$ 表示火箭的速度向量，于是

$$\frac{dy^i}{dt}=v^i, \quad y^i(0)=y_0^i, \qquad i=1, 2, \tag{4.1}$$

其中 $y_0=(y_0^1, y_0^2)$ 表示火箭的初始位置。

设 $\beta(t)$ 表示在时刻 t 火箭燃料燃烧的速率，$m(t)$ 表示时刻 t 火箭的质量。于是

$$\frac{dm}{dt}=-\beta. \tag{4.2}$$

火箭的质量等于燃料的质量加上飞行器的质量 $a>0$，因此我们有 $m(t)\geqslant a$。

设 $\omega(t)$ 表示在时刻 t 推力向量与正 y^1-轴的夹角。燃烧速率和推力角将受约束条件

$$0\leqslant\beta_0\leqslant\beta(t)\leqslant\beta_1, \qquad \omega_0\leqslant\omega(t)\leqslant\omega_1 \tag{4.3}$$

的支配，其中 β_0, β_1, ω_0 和 ω_1 是固定的。

为了得到完整的火箭运动方程，我们分析火箭在作直线运动时的动量转移。设在时刻 t 时，火箭的质量为 m，速度为 v，则动量为 mv。设在时间区间 δt 的期间，火箭烧掉的燃料量为 $\delta\mu>0$。在时刻 $t+\delta t$ 时，设喷射燃烧生成物的速度为 v'，它们的质量显然

是 $\delta\mu$. 又在时刻 $t+\delta t$ 时设火箭的速度是 $v+\delta v$, 则它的质量显然是 $m-\delta\mu$. 设我们考虑在时刻 t 时由火箭的质量 m 和速度 v 组成的系统, 在时刻 $t+\delta t$ 时这个系统由火箭和所喷射的燃烧生成物组成. 所以这个系统在时间区间 δt 内动量的变化是

$$(\delta\mu)v'+(m-\delta\mu)(v+\delta v)-mv.$$

如果我们用 δt 去除上述表达式然后令 $\delta t\to 0$, 我们就得到这个系统动量的变化率, 它必须等于作用在这个系统上的外力之和. 因此, 假如 F 是作用在这个系统的每单位质量上的合成外力, 我们就有

$$Fm-(v'-v)\frac{d\mu}{dt}=m\frac{dv}{dt}.$$

如果假定燃烧生成物相对于火箭的速度 $(v'-v)$ 是常数 c, 同时假定我们采用 $\frac{d\mu}{dt}=\beta$, 则就得到

$$F-\frac{c\beta}{m}=\frac{dv}{dt}.$$

如果把上面的分析用于平面运动的每一个分量上, 我们就得到下面的方程:

$$\begin{cases}\dfrac{dv^1}{dt}=F^1-\dfrac{c\beta}{m}\cos\omega,\\[2mm] \dfrac{dv^2}{dt}=F^2-\dfrac{c\beta}{m}\sin\omega,\\[2mm] v^i(0)=v_0^i,\qquad i=1,\ 2.\end{cases}\tag{4.4}$$

此方程和(4.1), (4.2), (4.3)一起, 决定了平面上的火箭运动. 这里, 若假定运动发生在一个非常数的引力场, 而且假若有阻力作用在火箭上的话, 则力 F 的分量可以是 y 和 v 的函数.

涉及火箭运动的控制问题具有如下形式: 从分段连续函数类(或某个其它适当的函数类)中, 选取燃烧速率控制 β 和推力方向控制 ω 使变量 t, y, v, m 中的某些达到指定的终端值, 并从那些能使达到这些终端值的控制中去决定一个控制, 它使其余终端值的一个给定函数取极大(或极小)值. 在另一类问题中, 是使一个沿着轨线计算的积分在状态空间中取极值.

更具体地说，我们考虑"最省燃料问题"．它要求火箭从一个指定的初始点 y_0 飞到一个指定的终点 y_f 而使燃料消耗最少．这个问题是很重要的，理由如下：因为能造出且能飞升的火箭加燃料和负载的总重量根据技术的情况是受到限制的，因而要求较少的燃料消耗能使火箭携带较多的负载．从(4.2)我们有

$$m_f=m_0-\int_{t_0}^{t_f}\beta(t)dt,$$

其中 t_0 是初始时刻，t_f 是终止时刻(在此时刻火箭达到 y_f)，m_f 是最后的质量，m_0 是初始质量，所以燃料的消耗是 m_0-m_f．因此，在满足(4.1)～(4.4)的条件下，燃料消耗最少的问题是使

$$P(\beta,\ \omega)=\int_{t_0}^{t_f}\beta(t)dt \tag{4.5}$$

取最小值．这个问题等价于 m_f 取最大值的问题．在最少燃料消耗问题中，假如容许"硬着陆"的话，最终的速度向量 v_f 不必指定；假如是"软着陆"，则必须指定．而终止时刻 t_f 可以是指定的，也可以是不指定的．

另一个实例是与一个运动着的目标交会的问题．设运动目标在时刻 t 的位置向量是 $z(t)=(z^1(t),\ z^2(t))$，在时刻 t 的速度向量是 $w(t)=(w^1(t),\ w^2(t))$，其中 w^1 和 w^2 是连续函数．我们假定存在满足(4.3)的推力程序 β 和 ω，它们使得交会能够实现．在数学上这表示对应于所选取的 β 和 ω，运动方程的解 $y(t)$，$v(t)$ 能使方程

$$\begin{aligned} y(t)&=z(t),\\ v(t)&=w(t) \end{aligned} \tag{4.6}$$

有正的解．这样的控制$(\beta,\ \omega)$叫做容许的．因为对于每个容许的 β 和 ω，相应的解 y 和 v 是连续的，又因为按假设 Z 和 w 也是连续的，于是得出：对于每个容许对$(\beta,\ \omega)$，有使(4.6)成立的最小正解 $t_f(\beta,\ \omega)$，数 $t_f(\beta,\ \omega)$ 是交会的时间．这里可能出现两个问题：第一，要求从容许控制中决定一个控制，它给出最大的负载，亦即使 $m_f=m_f(\beta,\ \omega)$取最大值．第二，使交会的时间 $t_f(\beta,\ \omega)$最小．

5. 电 机 工 程

用伺服机构把一个控制面保持在某个任意位置上．设外界的各种干扰，诸如偶尔出现的一阵阵风等，它们相对于伺服机构的时间常数来说是短暂的．我们用一个直流电动机提供一个力矩把控制面引到所希望的位置．只有电动机中的电枢电压 v 是可控的．为简明起见，我们取所希望的位置为零位角，并且测量与所希望的位置发生的偏差角 θ．关于 θ 的微分方程经过适当的标准化以后可以写成

$$\frac{d^2\theta}{dt^2}+a\,\frac{d\theta}{dt}+\omega^2\theta=u,\ \theta(0)=\theta_0,\ \theta'(0)=\theta'_0. \tag{5.1}$$

这里 u 表示加在控制面上的恢复力矩，$a\,\frac{d\theta}{dt}$ 表示阻尼效应．假如不出现阻尼，则 $a=0$．因为电压源不能提供绝对值大于某个 v_0 的电压，于是恢复力矩的绝对值必须有界．因此，我们有

$$|u(t)|\leqslant A, \tag{5.2}$$

其中 A 是一个常数．

假如令 $$x^1=\theta,\ x^2=\frac{d\theta}{dt},$$

我们就可以改写方程(5.1)为下面的形式：

$$\begin{aligned}\frac{dx^1}{dt}&=x^2, & x^1(0)&=\theta_0,\\ \frac{dx^2}{dt}&=-ax^2-\omega^2x^1+u, & x^2(0)&=\theta'_0.\end{aligned} \tag{5.3}$$

问题如下：假如由静止时所希望的位置经一短暂的干扰已引起了一个偏差 $\theta=\theta_0$，$\frac{d\theta}{dt}=\theta'_0$，应如何随时改变电压以使在最短的可能时间内把控制面恢复到设定的位置 $\theta=0$，$\frac{d\theta}{dt}=0$？根据(5.3)，这个问题就是从适当的函数类中选取函数 u，比如说从分段连续函数类中选取 u，使 u 在每一时刻满足(5.2)，并且使方程(5.3)对应于这个 u 的解$(x^1,\ x^2)$在最短的时间内达到$(x^1,\ x^2)$-空间的坐标原点．

6. 捷线问题

现在我们介绍变分法中的一个问题，也就是1696年John Bernoulli提出来的所谓捷线问题．这个问题可以认为是变分法理论的出发点．Galileo看来也曾经在1630年和1638年考虑过这个问题，但是他的表达形式不是很明确的．

设在一个铅垂平面上给定P_0和P_1两点，使得P_0高于P_1．一个粒子或质点单纯由于地心引力的作用沿着连接P_0和P_1的曲线C运动；而且质点沿着曲线在点P_0具有初始速度v_0．问题是如何选取曲线C使得质点从P_0到P_1所化的时间最短．

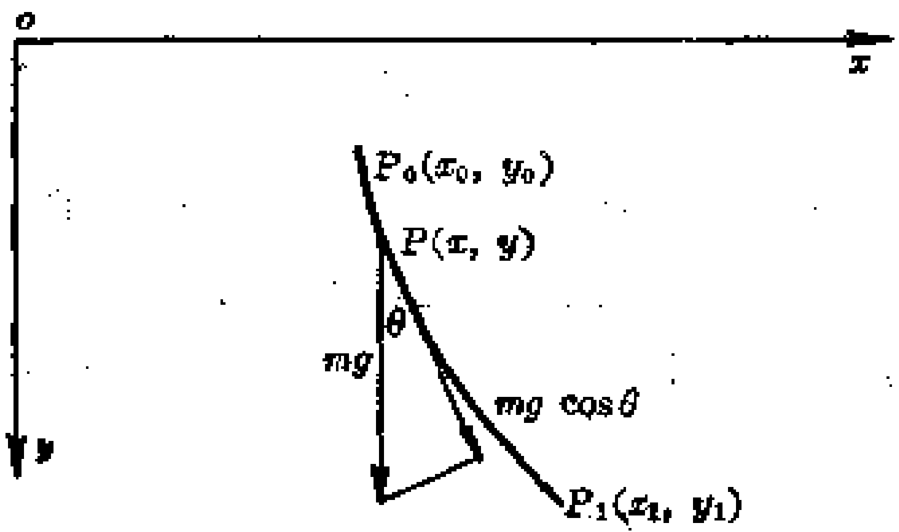

图 1

为了解析地表达这个问题，我们在平面上建立一个坐标系，如左图1所示．设点P_0具有坐标(x_0, y_0)，点P_1具有坐标(x_1, y_1)；且设曲线C的方程为$y=y(x)$；又设s_1表示P_0和P_1之间C的弧长．我们要确定从P_0到P_1经过C所需要的时间．

设P是C上具有坐标(x, y)的点，在这一点上质点因地心引力沿着曲线作用的分量是$mg\cos\theta$，其中θ是这一点的切线与向下铅垂线(正y轴)所成的夹角．从而如果设$s(t)$是质点从P_0出发在t秒钟内沿着曲线C经过的距离，我们就有

$$m\frac{d^2s}{dt^2}=mg\cos\theta.$$

假如用$2m^{-1}\frac{ds}{dt}$乘这个方程的两边，并且使用关系式$\frac{dy}{ds}=\cos\theta$，就得到

$$\frac{d}{dt}\left(\frac{ds}{dt}\right)^2=2g\frac{dy}{dt}.$$

如果令$v=\frac{ds}{dt}$，并且在初始点P_0和点P之间的曲线上积分这个

关系式，就得到

$$v^2-v_0^2=2g(y-y_0).$$

这个方程能够改写成形式

$$\frac{ds}{dt}=[2g(y-\alpha)]^{\frac{1}{2}}, \tag{6.1}$$

其中 $\alpha=y_0-(v_0^2/2g)$.

因此，若使用关系式

$$ds=[1+(y')^2]^{\frac{1}{2}}dx,$$

我们就得到沿着 C 从 P_0 到 P_1 所化的时间 τ 是

$$\tau=\int_0^\tau dt=\int_0^{s_1}\frac{ds}{[2g(y-\alpha)]^{\frac{1}{2}}}=(2g)^{-\frac{1}{2}}\int_{x_0}^{x_1}\left[\frac{1+(y')^2}{y-\alpha}\right]^{\frac{1}{2}}dx.$$

于是，除常数因子 $(2g)^{-\frac{1}{2}}$ 之外，寻找曲线 C 使质点从 P_0 到 P_1 之间转移时间最小的问题等价于下面的问题：在 $[x_0, x_1]$ 上可微且满足条件 $y(x_0)=y_0$, $y(x_1)=y_1$ 的函数类 Y 中，寻找一个函数 y，使得积分

$$\int_{x_0}^{x_1}\left(\frac{1+(y')^2}{y-\alpha}\right)^{\frac{1}{2}}dx$$

取最小值.

我们可以把这个问题表达成跟前面的问题具有相同的格式. 把独立变量 x 的符号改成 t，并令

$$y'=u,\ y(t_0)=y_0. \tag{6.2}$$

连续函数 u 叫做容许函数，假如它在 $[t_0, t_1]$ 上定义，而且假如对应于这个函数 u 方程 (6.2) 的解 $y(t)$ 满足 $y(t_1)=y_1$. 我们的问题是：在所有的容许函数 u 的类中，找出一个容许函数 u，使得

$$J(u)=\int_{t_0}^{t_1}\left(\frac{1+u^2}{y-\alpha}\right)^{\frac{1}{2}}dt$$

取最小值.

我们指出，捷线问题能够用不同的形式表示成控制问题. 由 (6.1)，质点沿着曲线 C 的速度由 $[2g(y-\alpha)]^{\frac{1}{2}}$ 给出，因此假如 θ

如图1所示，使得

$$\begin{cases}\dfrac{dx}{dt}=[2g(y-\alpha)]^{\frac{1}{2}}\sin\theta,\\ \dfrac{dy}{dt}=[2g(y-\alpha)]^{\frac{1}{2}}\cos\theta.\end{cases}$$

令 $u=\sin\theta$，则运动方程变为

$$\begin{cases}\dfrac{dx}{dt}=[2g(y-\alpha)]^{\frac{1}{2}}u, & x(t_0)=x_0,\\ \dfrac{dy}{dt}=[2g(y-\alpha)]^{\frac{1}{2}}(1-u^2)^{\frac{1}{2}}, & y(t_0)=y_0.\end{cases}\tag{6.3}$$

于是问题变为选取一个控制 u，满足 $|u|\leqslant 1$，使得在初始时刻 t_0 从 $P_0(x_0, y_0)$ 出发的点 $P(x, y)$，以最短的时间达到规定的点 $P_1(x_1, y_1)$. 假如 t_1 是达到 P_1 的时刻，则它等价于使 t_1-t_0 取最小值. 它也等价于在满足方程(6.3)和约束 $|u|\leqslant 1$ 的条件下使积分

$$\int_{t_0}^{t_1} dt$$

取最小值.

捷线问题能够用下面的方式修改：人们可以提出由 $y=y_1(x)$ 确定的曲线 Γ_1 来代替固定点 P_1，然后寻找连接 P_0 到 Γ_1 的曲线 C，使质点必须沿此曲线移动，而在最短的时间内从 P_0 到达 Γ_1. 我们也可以用曲线 Γ_0 代替 P_0，其中 Γ_0 与 Γ_1 相隔一个正的距离，要求找出一条连接 Γ_0 和 Γ_1 的曲线 C. 使质点必须沿此曲线移动，而使转移时间最短.

第二章　控制问题的表达形式

1. 引　　言

本章我们首先指出前一章诸例子的共同数学结构，这时最优控制的数学问题容许采用稍欠精确的初等表达形式.然后，在第3节才导出最优控制的数学问题的精确而又一般的表达形式. 第4节讨论问题的各种等价表达形式；第5节说明如何将某些其他控制问题提炼成第3节所给出的一般形式.作为本章的结束，我们将讨论控制问题与变分法中的问题两者之间的关系. 第4节至第6节所包括的内容在第一次阅读时可以略去，等到必要时再回过头来阅读.

2. 控制问题的初等表达形式

前一章的所有例子都具有如下形式. 系统在时刻 t 的状态用 n 维欧几里德空间里的一个点或向量

$$x(t)=(x^1(t),\ \cdots,\ x^n(t)),\ n\geqslant 1,$$

来描述. 在初始时刻 t_0，系统的状态是

$$x(t_0)=x_0=(x_0^1,\ \cdots,\ x_0^n).$$

更一般些，我们可以要求初始时刻 t_0 和初始状态 x_0 构成的点 $(t_0,\ x_0)$ 属于 $(t,\ x)$-空间里某一预先指定的集合 $\mathscr{T}_0$. 而系统依时间变化的状态遵从微分方程组

$$\frac{dx^i}{dt}=f^i(t,\ x,\ z),\ x^i(t_0)=x_0^i,\ i=1,\ \cdots,\ n, \tag{2.1}$$

其中 $z=(z^1,\ \cdots,\ z^m)$ 是实欧几里德空间 E^m 中的向量，函数 f^i 是变量 $(t,\ x,\ z)$ 的实值连续函数.

"系统按(2.1)而变化"的含意是指：当我们从值域在 m 维欧几

里德空间 E^m 中的某一指定函数类 $\mathscr{C}$（本节我们取 $\mathscr{C}$ 为分段连续函数类的子集）中选取一个函数 $u(t)$，并把 $z=u(t)$ 代入 (2.1) 的右边，我们得到常微分方程组

$$\frac{dx^i}{dt}=f^i(t,\ x,\ u(t))=F^i_u(t,\ x),\ i=1,\ \cdots,\ n, \tag{2.2}$$

这里 F^i_u 的下标 u 强调(2.2)的右边是依赖于函数 u 的选择的. 对于 $\mathscr{C}$ 中的每一个 u，假定存在点 $(t_0,\ x_0)\in\mathscr{T}_0$ 和一个定义于 $[t_0,\ t_2]$ 上取值于 E^n 中的函数 $\phi=(\phi^1,\ \cdots,\ \phi^n)$ 使得 (2.2) 满足. 亦即，要求对每一个 $t\in[t_0,\ t_2]$，满足方程组

$$\phi'^i(t)=\frac{d\phi^i}{dt}=f^i(t,\ \phi(t),\ u(t)),\ \phi^i(t_0)=x^i_0,\ i=1,\ \cdots,\ n.$$

在 u 的不连续点处，上述方程理解为单边极限成立. 函数 $\phi(t)$ 描述系统依时间 t 而变化的规律，有时就把它叫做轨线.

函数 u 的进一步要求是：在某一时刻 $t_1(t_0<t_1)$，点 $(t_1,\phi(t_1))$ 属于某一预先指定的集合 $\mathscr{T}_1$，而对于 $t\in[t_0,\ t_1)$，点 $(t,\ \phi(t))$ 不属于 $\mathscr{T}_1$. 集合 $\mathscr{T}_1$ 叫做问题的终点集合. 现在就第一章所给出的例子，其终点集合说明如下.

在生产计划问题中，$\mathscr{T}_1$ 是 $(t,\ x)$-平面上的直线 $t=T$. 在第一型化学工程问题里，$\mathscr{T}_1$ 是超平面 $t=T$；亦即 $(t,\ x)$-空间里 $x=(x^1,\ \cdots,\ x^n)$ 自由而 t 固定在 T 的点 $(T,\ x)$ 全体. 在化学工程问题的最后一种型式里，$\mathscr{T}_1$ 是 $(t,\ x)$-空间里这样一些点的集合：它的坐标 x^i 固定在 x^i_1，$i=1,\ \cdots,\ j$；而其余的坐标是自由的. 在某些问题中，要求解 $\phi(t)$ 击中可移动的目标集 $G(t)$. 亦即对某区间 $[\tau_0,\ \tau_1]$ 上的每一时刻 t，存在 x-空间里的一个集合 $G(t)$，并要求解 ϕ 在某一时刻 t 击中 $G(t)$. 解析地说，我们要求在 $[\tau_0,\ \tau_1]$ 中有一点 t_1 能使 $\phi(t_1)$ 属于 $G(t_1)$. 这类问题的一个例子是 1.4 节中的交会问题. 在变动目标集的问题里，终点集合 $\mathscr{T}_1$ 是所有点 $(t,\ x)$ 全体，其中 $t\in[\tau_0,\ \tau_1]$，而 $x\in G(t)$.

上面讨论的问题有时用稍欠精确但更形象的语言可概括成：要求控制函数 $u(t)$ 将系统从时刻 t_0 的初始状态 x_0 转移到时刻 t_1

的终止状态 x_1，其中 $(t_0, x_0)\in\mathscr{T}_0$，而 $(t_1, x_1)\in\mathscr{T}_1$. 注意，对每一个 $u(t)\in\mathscr{C}$，一般地说有多于一条的相应轨线 ϕ，这是由于初始点 (t_0, x_0) 在 $\mathscr{T}_0$ 中的不同选取或者由于没有假设系统 (2.2) 的解应满足唯一性条件所造成的.

常常进一步要求 $\mathscr{C}$ 中的函数 u 及其相应的轨线 ϕ 必须对所有 $t_0\leqslant t\leqslant t_1$ 满足一组不等式约束

$$R^i(t, \phi(t), u(t))\geqslant 0, \qquad i=1, \cdots, r, \tag{2.3}$$

其中函数 $R^1, \cdots, R^r$ 是 (t, x, z) 的已知函数. 例如，在 1.2 节讨论的生产计划问题中，约束条件可写成 $R^i\geqslant 0$, $i=1, 2, 3$，其中 $R^1(t, x, z)=x$, $R^2(t, x, z)=z$，而 $R^3(t, x, z)=A-z$. 在 1.5 节的例子中，约束条件可写成 $R^i\geqslant 0$, $i=1, 2$，其中 $R^1(t, x, z)=z+A$，而 $R^2(t, x, z)=A-z$.

第一章所有的例子都是要选择一个控制函数 u，使得某一泛函取极小或极大. 这些泛函具有如下的形式. 设 f^0 是 (t, x, z) 的实值连续函数，g_0 是定义于 $\mathscr{T}_0$ 上的实函数，而 g_1 是定义于 $\mathscr{T}_1$ 上的实函数. 对 $\mathscr{C}$ 中的每一个 u 及其相应的 (2.2) 的每一个解 ϕ，定义一个如下的代价泛函或性能指标：

$$J(\phi, u)=g_0(t_0, \phi(t_0))+g_1(t_1, \phi(t_1)) + \int_{t_0}^{t_1} f^0(s, \phi(s), u(s))ds.$$

如果要使泛函 J 取极小值，那么应找出一个 $u^*\in\mathscr{C}$ 和与之相应的 (2.2) 的解 ϕ^*，使得 $J(\phi^*, u^*)\leqslant J(\phi, u)$ 对所有的 $u\in\mathscr{C}$ 和与之相应的 ϕ 成立. 在另外的问题中是要使泛函 J 取极大值. 下一段将给出第一章所论诸例泛函 J 的形式.

在第一章的诸例子中，集合 $\mathscr{T}_0$ 常取为一个点 (t_0, x_0). 在那些例子中，除 1.3 节的例子外，所有微分方程都满足解的唯一性条件；在 1.3 节中，若假设函数 G^i 也使得解是唯一的，则在所有这些例子中 u 的选取完全决定了相应的 ϕ. 在经济学的例子中，$J(\phi, u)$ 是由 (2.5) 表示的总成本 $c(u)$，函数 f^0 由 (2.4) 给出，而 g_0 和 g_1 恒等于 0. 在 1.3 节的第一型化学工程的例子中，$J(\phi, u)$

$=v(p, \theta)$，其中 $v(p, \theta)$用(3.3)表示；函数 f^0 和 g_0 恒等于 0. 在 1.4 节的最少燃料问题中，$J(\phi, u)=P(\beta, \omega)$，其中 $P(\beta, \omega)$ 用(4.5)表示，此地 $f^0=\beta$，而 g_0 和 g_1 恒等于 0. 它的一个等价表达形式是取 $J(\phi, u)=-m_f$，这时 $f^0=0$，$g_0=0$，而 $g_1=-m_f$.

最后，我们以讨论两种推广来结束本节，它们在下一节就要给出的数学表达形式中将会出现. 首先讨论初始和终止数据. 初始点集合 $\mathscr{T}_0$ 和终止点集合 $\mathscr{T}_1$ 一起可确定 E^{2n+2} 中一个点集 $\mathscr{B}$ 如下：

$$\mathscr{B}=\{(t_0, x_0, t_1, x_1): (t_0, x_0)\in\mathscr{T}_0,\ (t_1, x_1)\in\mathscr{T}_1\}. \tag{2.4}$$

因此，要求 $(t_0, \phi(t_0))\in\mathscr{T}_0$ 和 $(t_1, \phi(t_1))\in\mathscr{T}_1$ 的一种简单推广如下：假设 $\mathscr{B}$ 是 E^{2n+2} 中已给定的集合，要求轨线 $\phi(t)$在 $[t_0, t_1]$ 上有定义，并使 $(t_0, \phi(t_0), t_1, \phi(t_1))\in\mathscr{B}$. 亦即，现在容许初始和终止数据之间可能存在相互关系. 下面我们将证明，在某种意义下，这种情形其实并不比初始和终止数据互不相关的情形更一般.

我们要讨论的第二种推广是关于控制 u 的约束的描述. 对每一(t, x)，不等式组 $R^i(t, x, z)\geqslant 0$，$i=1, \cdots, r$，确定了 m 维 z-空间的一个集合 $U(t, x)$；即

$$U(t, x)=\{z: R^i(t, x, z)\geqslant 0,\ i=1, 2, \cdots, r.\}.$$

于是，函数 u 和相应的轨线 ϕ 满足形如 (2.3) 的约束要求，可以写成

$$u(t)\in U(t, \phi(t)), \qquad t_0\leqslant t\leqslant t_1.$$

因此约束条件(2.3)是下面更一般的约束条件的特殊情况.

设 Ω 是一个函数，它把 E^{n+1} 中某适当子集的每一点 (t, x)映入 z-空间的子集，亦即

$$\Omega: (t, x)\to\Omega(t, x),$$

其中 $\Omega(t, x)$是 E^m 的子集. 那么约束条件(2.3)可以用更一般的约束

$$u(t)\in\Omega(t, \phi(t))$$

来代替.

3. 数学的表达形式

最优控制问题的精确数学表达形式将涉及到 Lebesgue 积分，这在问题的求解的研究中是本质的。读者如果希望保持比较初等的表达形式，可以用“分段连续控制”代替“可测控制”；用“分段 $C^{(1)}$ 函数”代替“绝对连续函数”，而下面方程(3.1)的解可以象本章第 2 节中的方程(2.2)的解那样来理解。

我们先建立一些符号和术语。用 t 表示实数，有时称它为时间。x 表示实欧几里德空间 $E^n(n\geqslant 1)$ 中的向量；因此 $x=(x^1, \cdots, x^n)$，向量 x 叫做状态变量。我们用上标表示向量的支量；而用下标区分不同的向量。设 z 表示 m 维欧氏空间 $E^m(m>1)$ 中的向量；因此 $z=(z^1, \cdots, z^m)$。向量 z 叫做控制变量。令 $\mathscr{R}$ 表示(t, x)-空间的区域；而 $\mathscr{U}$ 表示 z-空间的区域，这里区域的含义是指连通开集。设 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 表示 $\mathscr{R}$ 与 $\mathscr{U}$ 的笛卡尔乘积，$f^0, f^1, \cdots, f^n$ 是定义于 $\mathscr{G}$ 上的实函数。我们记

$$f=(f^1, \cdots, f^n), \hat{f}=(f^0, f^1, \cdots, f^n).$$

令 $\mathscr{B}$ 表示 E^{2n+2} 中点

$$(t_0, x_0, t_1, x_1)=(t_0, x_0^1, \cdots, x_0^n, t_1, x_1^1, \cdots, x_1^n)$$

的集合，其中 $t_1\geqslant t_0+\delta$，δ 是某一固定的正数。集合 $\mathscr{B}$ 确定所论问题的端点条件。

设 Ω 是把 $\mathscr{R}$ 里的每一点 (t, x) 映入 z-空间的区域 $\mathscr{U}$ 的子集 $\Omega(t, x)$ 的映象，则说映象 Ω 定义了一个控制约束。如果对所有的 $(t, x)\in\mathscr{R}$，都有 $\Omega(t, x)=\mathscr{U}$，我们就说不存在控制约束。

今后，我们将采用向量-矩阵的记号。这时微分方程组(2.2)可简洁地写成

$$\frac{dx}{dt}=f(t, x, u(t)),$$

这里我们按常微分方程理论的习惯，取 $\frac{dx}{dt}$ 和 $f(t, x, u(t))$ 为列向量。如果一向量是行向量或列向量是明显的，或者它是行向量或是列向量是无关紧要的，那么我们对向量和它的转置两者之间将

不加区别．向量 u 和 v 的内积写成 $\langle u, v\rangle$．我们用符号 $|x|$ 表示向量 x 的通常欧氏模，因此

$$|x|=\left(\sum_{i=1}^{n}|x^i|^2\right)^{1/2}=\langle x, x\rangle^{1/2}.$$

如果 A 和 B 是矩阵，那么它们的积写成 AB.

如果 $f=(f^1, \cdots, f^n)$ 是把某一欧几里德空间里的集合 Δ 映射到欧几里德空间 E^n 的向量函数，使得每一个实值函数 $f^1, \cdots, f^n$ 在 Δ 上连续（或 $C^{(k)}$，或可测，等等），我们就说 f 在集合 Δ 上连续（或 $C^{(k)}$，或可测，等等）．同样地，若一矩阵 A 的每一元素在某欧氏空间里的集合 Δ 上连续（或 $C^{(k)}$，或可测，等等），则说矩阵 A 在 Δ 上是连续的（或 $C^{(k)}$，或可测的，等等）.

定义 3.1 定义在区间 $[t_0, t_1]$ 上而取值于 $\mathscr{U}$ 中的可测函数 $u(t)$ 称之为在 $[t_0, t_1]$ 上的一个控制，如果存在一个定义在 $[t_0, t_1]$ 上而取值于 E^n 的绝对连续函数 $\phi(t)$，使得

（i）对所有 $t_0\leqslant t\leqslant t_1$，$(t, \phi(t))\in\mathscr{R}$；

（ii）$\phi(t)$ 是微分方程组

$$\frac{dx}{dt}=f(t, x, u(t)) \tag{3.1}$$

的解；亦即在 $[t_0, t_1]$ 上如下等式几乎处处成立：

$$\phi'(t)=f(t, \phi(t), u(t)).$$

函数 $\phi(t)$ 叫做相应于 $u(t)$ 的轨线；点 $(t_0, \phi(t_0))$ 叫做轨线的初始点；点 $(t_1, \phi(t_1))$ 叫做轨线的终止点；而点 $(t_0, \phi(t_0), t_1, \phi(t_1))$ 叫做轨线的端点.

注意，因为 ϕ 是绝对连续的，它是其自身的导函数的积分．因此（ii）意味着函数 $t\to f(t, \phi(t), u(t))$ 是在 $[t_0, t_1]$ 上 Lebesgue 可积的.

微分方程组（3.1）叫做状态方程.

关于所用的记号，我们着重指出如下这一点，字母 z 表示 $\mathscr{U}$ 中的点；而字母 u 表示取值于 $\mathscr{U}$ 中的函数.

定义 3.2 控制 u 叫做容许控制，如果存在相应于 u 的轨线

ϕ 使得

(i) $t\to f^0(t,\ \phi(t),\ u(t))\in L_1[t_0,\ t_1]$;

(ii) $u(t)\in\Omega(t,\ \phi(t))$在$[t_0,\ t_1]$上几乎处处成立;

(iii) $(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))\in\mathscr{B}$.

与定义 3.2 中的容许控制 u 相对应的轨线 ϕ 叫做容许轨线.

定义 3.3 容许控制 u 和与之相应的容许轨线 ϕ 所组成的函数对$(\phi,\ u)$叫做容许对.

注意，对一个给定的容许控制可以有多于一条的相应的容许轨线，因为我们可允许端点的不同选择. 而且，即使我们固定端点，对给定的控制也可以存在多条相应的轨线，因为我们没有要求(3.1)的解对给定的初始条件具有唯一性.

我们现在叙述最优控制问题.

问题 1 令 $\mathscr{A}$ 表示所有容许对 $(\phi,\ u)$ 的集合，并设 $\mathscr{A}$ 是非空的. 设

$$J(\phi,\ u)=g(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))+\int_{t_0}^{t_1}f^0(t,\ \phi(t),\ u(t))dt, \tag{3.2}$$

其中$(\phi,\ u)$是容许对，而 g 是定义在 $\mathscr{B}$ 上的已知实函数. 设 $\mathscr{A}_1$ 是 $\mathscr{A}$ 的非空子集，我们的问题是要在 $\mathscr{A}_1$ 中找一对 $(\phi^*,\ u^*)$，使(3.2)式在集合 $\mathscr{A}_1$ 中取极小值. 亦即，在 $\mathscr{A}_1$ 中找一元素对$(\phi^*,\ u^*)$使得

$$J(\phi^*,\ u^*)\leqslant J(\phi,\ u),\qquad \text{对所有的}(\phi,\ u)\in\mathscr{A}_1\text{成立}.$$

问题 1 的精确表达形式比较冗长. 因此，今后在考虑问题 1 时，常采用突出问题本质的如下叙述方式：在满足状态方程(3.1)，端点条件 $\mathscr{B}$，和控制约束 Ω 的条件下，使(3.2)取极小值.

上面我们是把问题 1 表达成极小化问题. 在有些应用中，可能要求泛函 J 取极大值. 但是没有必要单独考虑极大值的问题，因为 J 的极大值问题等价于 $-J$ 的极小值问题. 所以我们将只限于考虑极小值问题.

定义 3.4 问题 1 的一对解$(\phi^*,\ u^*)$叫做最优对. 轨线 ϕ^* 叫

优轨线，而控制 u^* 叫做最优控制．

(3.2)中右边第一项的函数 g 是在容许轨线的端点上取值的．因此，对每一条容许轨线它赋予一个实数，从而它是定义在容许轨线上的泛函 G_1，这里 G_1 由公式

$$G_1(\phi)=g(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))$$

定义．定义在容许轨线上的泛函的另一些例子是

$$G_2(\phi)=\max\{|\phi(t)|:t_0\leqslant t\leqslant t_1\}$$

和
$$G_3(\phi)=\max\{|\phi(t)-h(t)|:t_0\leqslant t\leqslant t_1\},$$

其中 h 是定义在区间 I 上的已知连续函数，而区间 I 包含所有容许轨线的定义区间 $[t_0, t_1]$．泛函 G_2 和 G_3 出现在这样的一些问题中，它们除了使(3.2)取极小值之外，还要求保持系统的状态始终靠近于某一预先指定的状态．

上面的讨论说明问题 1 的下述推广是合理的．

问题 2 假定除了用

$$J(\phi,\ u)=G(\phi)+\int_{t_0}^{t_1}f^0(t,\ \phi(t),\ u(t))\,dt \qquad (3.3)$$

代替 (3.2) 之外，其他一切条件如同问题 1，其中 $G(\phi)$ 是定义在容许轨线上的泛函．要求寻找一对 $(\phi^*,\ u^*)\in\mathscr{A}_1$，它使 (3.3) 在 $\mathscr{A}_1$ 中取极小值．

4．等价的表达形式

在问题 1 能形式地变换成其特殊情形的意义下，问题 1 的某些特殊情形实际上是与问题 1 等价的．这种情况在某些研究中是有用的，因为此时研究问题 1 的特殊情形之一要比研究问题 1 更方便一些．但读者务必十分小心，当变换为特殊情形时，原来问题的某些性质(诸如线性性，连续性，凸性等等性质)可能会改变．因此，在任何特定的研究中，我们必须检验对原来问题所做的合理假设，对变换后的问题是否仍然有效．

问题 1 的两种特殊情形是取 $f^0=0$ 或者 $g=0$．为了保持变分法中有关问题的术语，我们称 $f^0=0$ 的情形为 Mayer 问题；称 $g=0$

的情形为 Lagrange 问题；而第 3 节中的问题 1（即 $f^0\neq 0$, $g\neq 0$），也如变分法中那样，称为 Bolza 问题．我们通过证明问题 1 既可以表达成 Mayer 问题又可以表达成 Lagrange 问题，来说明 Mayer 形式和 Lagrange 形式与 Bolza 形式具有同样的一般性．

我们先把问题 1 表达成更高维欧氏空间中的 Mayer 问题．设 $\hat{x}=(x^0, x)=(x^0, x^1, \cdots, x^n)$，$\hat{\mathscr{R}}=E^1\times\mathscr{R}$，$\hat{\mathscr{G}}=\hat{\mathscr{R}}\times\mathscr{U}$．函数 $f^0, f^1, \cdots, f^n$ 在 $\hat{\mathscr{G}}$ 上定义且与 x^0 无关，所以它们在 $\mathscr{G}$ 上有定义．设 $\hat{\Omega}$ 是由方程 $\hat{\Omega}(t, \hat{x})=\Omega(t, x)$ 定义的 $\hat{\mathscr{R}}$ 上的映象．令

$$\hat{\mathscr{B}}=\{(t_0, \hat{x}_0, t_1, \hat{x}_1): (t_0, x_0, t_1, x_1)\in\mathscr{B}, x_0^0=0\}.$$

若 (ϕ, u) 是问题 1 的容许对，令 $\hat{\phi}=(\phi^0, \phi)$，其中 ϕ^0 是一绝对连续函数，使得

$$\phi^{0\prime}(t)=f^0(t, \phi(t), u(t)), \quad \phi^0(t_0)=0,$$

对 $[t_0, t_1]$ 中几乎所有的 t 成立．由定义 3.2 的条件 (i)，这样的函数 ϕ^0 是存在的，而且可以表示成

$$\phi^0(t)=\int_{t_0}^{t} f^0(s, \phi(s), u(s))ds.$$

则 $(\hat{\phi}, u)$ 是分别以 $\hat{\mathscr{R}}$、$\hat{\mathscr{G}}$、$\hat{\Omega}$、$\hat{\mathscr{B}}$ 代替 $\mathscr{R}$、$\mathscr{G}$、Ω、$\mathscr{B}$，而状态方程 (3.1) 用

$$\begin{cases}\dfrac{dx^0}{dt}=f^0(t, x, u(t)),\\[2mm] \dfrac{dx}{dt}=f(t, x, u(t))\end{cases} \tag{4.1}$$

代替之后的控制问题的容许对．如果令 $\hat{f}=(f^0, f)$，则方程组 (4.1) 可以写成

$$\frac{d\hat{x}}{dt}=\hat{f}(t, x, u(t)).$$

反之，对包含 $\hat{\mathscr{R}}$、$\hat{\mathscr{G}}$、$\hat{\Omega}$、$\hat{\mathscr{B}}$，并以 (4.1) 为状态方程的问题的每一容许对 $(\hat{\phi}, u)$，必存在相应于问题 1 的容许对 (ϕ, u)，其中 ϕ 是 $\hat{\phi}$ 的后 n 个支量所组成的．令

$$\hat{g}(t_0, \hat{x}_0, t_1, \hat{x}_1)=g(t_0, x_0, t_1, x_1)+x_1^0$$

和

$$\hat{J}(\hat{\phi}, u)=\hat{g}(t_0, \hat{\phi}(t_0), t_1, \hat{\phi}(t_1)).$$

则 $\hat{J}(\hat{\phi}, u)=J(\phi, u)$[注]，其中 $\hat{\phi}=(\phi^0, \phi)$. 因此，满足状态方程组 (4.1)，控制约束 $\hat{\Omega}$ 和端点条件 $\hat{\mathscr{B}}$，并使 $\hat{J}(\hat{\phi}, u)$ 取极小值的 Mayer 问题是与问题 1 等价的.

现在证明问题 1 可以表达成 Lagrange 问题. 设 $\hat{x}$、$\hat{\mathscr{R}}$、$\hat{\mathscr{G}}$、$\hat{\Omega}$ 如上一段所述. 令

$$\hat{\mathscr{B}}=\{(t_0, \hat{x}_0, t_1, \hat{x}_1):(t_0, x_0, t_1, x_1)\in\mathscr{B},\quad x_0^0=g(t_0, x_0, t_1, x_1)/(t_1-t_0)\}. \tag{4.2}$$

(记住，对 $\mathscr{B}$ 中所有的点都有 $t_1>t_0$.) 设 (ϕ, u) 是问题 1 的容许对，令 $\hat{\phi}=(\phi^0, \phi)$，其中 $\phi^0(t)\equiv g(t_0, x_0, t_1, x_1)/(t_1-t_0)$. 则 $(\hat{\phi}, u)$ 是以 $\hat{\mathscr{R}}$/ $\hat{\mathscr{G}}$、$\hat{\Omega}$ 和(4.2)式规定的 $\hat{\mathscr{B}}$ 分别代替 $\mathscr{R}$、$\mathscr{G}$、Ω 和 $\mathscr{B}$；状态方程组为

$$\begin{cases}\dfrac{dx^0}{dt}=0,\\ \dfrac{dx}{dt}=f(t, x, u(t))\end{cases} \tag{4.3}$$

的问题的容许对. 反之，对带尖顶号"∧"问题的每一个容许对 $(\hat{\phi}, u)$，必有相应于问题 1 的容许对 (ϕ, u)，其中 ϕ 由 $\hat{\phi}$ 的后 n 个支量组成. 因此，如果用 f^0+x^0 代替 f^0，并令

$$\hat{J}(\hat{\phi}, u)=\int_{t_0}^{t_1}(f^0(t, \phi(t), u(t))+\phi^0(t))dt, \tag{4.4}$$

则 $\hat{J}(\hat{\phi}, u)=J(\phi, u)$. 于是满足状态方程(4.3)，控制约束 $\hat{\Omega}$，和(4.2)式定义的端点条件 $\hat{\mathscr{B}}$，并使(4.4)取极小值的 Lagrange 问题是与问题 1 等价的.

在问题 1 里初始时刻 t_0 和终止时刻 t_1 不必是固定的. 现在证明：这时可以把问题 1 写成具有固定初始时刻和固定终止时刻的

[注] 按 $\hat{g}(t_0, \hat{x}_0, t_1, \hat{x}_1)$ 的定义和表达式 $\phi^0(t)=\int_{t_0}^{t}f^0(s, \phi(s), u(s))ds$，显然有

$$\hat{J}(\hat{\phi}, u)=\hat{g}(t_0, \hat{\phi}(t_0), t_1, \hat{\phi}(t_1))$$
$$=g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1}f^0(t, \phi(t), u(t))dt=J(\phi, u).$$

——译者注.

问题. 化变动初始时刻和终止时刻问题为固定初始时刻和终止时刻问题的办法是改变时间参数, 令

$$t=t_0+s(t_1-t_0),\ 0\leqslant s\leqslant 1.$$

并引进新的状态变量如下: 设 w 表示纯量, 考虑带状态变量 (t, x, w) 的问题, 其中 x 是 n 维向量, 而 t 是纯量. 令 s 表示时间变量, 设状态方程组为

$$\begin{cases}\dfrac{dt}{ds}=w,\ \dfrac{dw}{ds}=0;\\ \dfrac{dx}{ds}=f(t,\ x,\ \bar{u}(s))w,\end{cases}\tag{4.5}$$

其中 $\bar{u}$ 是控制, 而 f 如问题 1 所述. 令

$$\begin{aligned}\bar{\mathscr{B}}=\{(s_0,\ t_0,\ x_0,\ w_0,\ s_1,\ t_1,\ x_1,\ w_1):\ s_0=0,\ s_1=1,\\ (t_0,\ x_0,\ t_1,\ x_1)\in\mathscr{B},\ w_0=t_1-t_0\}.\end{aligned}\tag{4.6}$$

注意, 现在的初始时刻和终止时刻是固定的. 设 $\bar{\Omega}(s,\ t,\ x,\ w)=\Omega(t,\ x)$, $\bar{\phi}=(\tau,\ \xi,\ \omega)$ 是 (4.5) 相应于控制 $\bar{u}$ 的解, 其中 $(\tau,\ \xi,\ \omega)$ 与 $(t,\ x,\ w)$ 中的希腊-拉丁字母之间的对应关系乃表示 $\bar{\phi}$ 的支量和系统 (4.5) 的状态变量之间的对应关系. 令

$$\begin{aligned}\bar{J}(\bar{\phi},\ \bar{u})=g(\tau(0),\ \xi(0),\ \tau(1),\ \xi(1))\\ +\int_0^1 f^0(\tau(s),\ \xi(s),\ \bar{u}(s))\omega(s)ds.\end{aligned}\tag{4.7}$$

现在考虑满足状态方程 (4.5), 控制约束 $\bar{\Omega}$, 和端点条件 $\bar{\mathscr{B}}$, 并使 (4.7) 取极小值的固定端点时间的问题.

因为 $t_1-t_0>0$, 这表明满足条件 (4.6) 的 (4.5) 的任何解, 它的支量 $\omega(s)=t_1-t_0$ 在 $0\leqslant s\leqslant 1$ 上是正的常数. 设 $(\phi,\ u)$ 是问题 1 的容许对; 若

$$\begin{aligned}\tau(s)=t_0+s(t_1-t_0),\ \xi(s)=\phi(t_0+s(t_1-t_0)),\\ \bar{u}(s)=u(t_0+s(t_1-t_0)),\ \omega(s)=t_1-t_0,\end{aligned}$$

则容易验证 $(\bar{\phi},\ \bar{u})=(\tau,\ \xi,\ \omega,\ \bar{u})$ 是固定端点时间问题的容许对, 并有 $\bar{J}(\bar{\phi},\ \bar{u})=J(\phi,\ u)$. 反之, 若 $(\bar{\phi},\ \bar{u})$ 是固定端点时间问题的容许对, 令

$$\phi(t)=\xi((t-t_0)/(t_1-t_0)),\ \ u(t)=\bar{u}((t-t_0)/(t_1-t_0)),$$

$$t_0 \leqslant t \leqslant t_1,$$

则由于 $\tau(s)=t_0+s(t_1-t_0)$，我们可得 $t=\tau(s)$，当 $0\leqslant s\leqslant 1$ 时．且容易验证(ϕ, u)是问题 1 的容许对，而且 $J(\phi, u)=\bar{J}(\bar{\phi}, \bar{u})$[注]．所以问题 1 等价于固定端点时间问题．

因为对于固定时间问题的任何容许解总有 $\omega(s)=t_1-t_0>0$，因此对固定端点时间问题可以取 $\tilde{\mathscr{R}}=[0, 1]\times\mathscr{R}\times E^+$，其中 $E^+=\{w: w>0\}$．这种看法今后将是有用的．

如果初始和终止数据是分开给定的，则出现端点条件的特殊情形．这时 E^{n+1} 中点 (t_0, x_0) 的集合 $\mathscr{T}_0$ 和 E^{n+1} 中点 (t_1, x_1) 的集合 $\mathscr{T}_1$ 是给定的，并要求容许轨线 ϕ 满足条件

$$(t_i, \phi(t_i))\in\mathscr{T}_i, \qquad i=0, 1. \tag{4.8}$$

在此情形中，集合 $\mathscr{B}$ 由(2.4)给定．我们将证明定义 3.2 中出现的似乎更一般的条件(iii)可以化为形式(4.8)，其办法是将问题嵌入到较高维空间中．方法如下：

设 $y=(y^1, \cdots, y^n)$，y^0 是纯量．记 $\hat{y}=(y^0, y)$．问题 1 中的集合 $\mathscr{R}$ 和 $\mathscr{G}$ 用 $\tilde{\mathscr{R}}=\mathscr{R}\times E^{n+1}$ 和 $\tilde{\mathscr{G}}=\tilde{\mathscr{R}}\times\mathscr{U}$ 代替，则向量函数 $\hat{f}=(f^0, f)$ 在 $\tilde{\mathscr{G}}$ 上有定义，因为它与 $\hat{y}$ 无关．设 $\tilde{\Omega}(t, x, \hat{y})=\Omega(t, x)$．并设状态方程组为

[注] 因为 $J(\phi, u)=g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1}f^0(t, \phi(t), u(t))dt$，由关系式：

$$\phi(t)=\xi((t-t_0)/(t_1-t_0)),\ u(t)=\bar{u}((t-t_0)/(t_1-t_0)),\ \tau(s)=t_0+s(t_1-t_0),$$

显然有 $\tau(0)=t_0$，$\tau(1)=t_1$，$\phi(t_0)=\xi(0)$，$\phi(t_1)=\xi(1)$．因此

$$J(\phi, u)=g(\tau(0), \xi(0), \tau(1), \xi(1))$$
$$+\int_{t_0}^{t_1}f^0(t, \xi((t-t_0)/(t_1-t_0)), \bar{u}((t-t_0)/(t_1-t_0)))dt.$$

对上式右边第二项的积分做变换 $t=\tau(s)=t_0+s(t_1-t_0)$，则 $dt=(t_1-t_0)ds$，亦即 $dt=\omega(s)ds$，从而可得

$$J(\phi, u)=g(\tau(0), \xi(0), \tau(1), \xi(1))$$
$$+\int_0^1 f^0(\tau(s), \xi(s), \bar{u}(s))\omega(s)ds=\bar{J}(\bar{\phi}, \bar{u}).$$

——译者注．

$$\begin{cases}\dfrac{dx}{dt}=f(t,\ x,\ u(t)),\\[2mm] \dfrac{d\hat{y}}{dt}=0.\end{cases}\tag{4.9}$$

令 $\widetilde{\mathscr{T}}_0=\{(t_0,\ x_0,\ y_0^0,\ y_0):(t_0,\ x_0,\ y_0^0,\ y_0)\in\mathscr{B}\}$,

$\widetilde{\mathscr{T}}_1=\{(t_1,\ x_1,\ y_1^0,\ y_1):y_1^0=t_1,\ y_1^i=x_1^i,\ i=1,\ \cdots,\ n\}$.

定义 3.2 的条件(iii)用条件

$$(t_i,\ \tilde{\phi}(t_i))\in\widetilde{\mathscr{T}}_i,\qquad i=0,\ 1,\tag{4.10}$$

代替，其中 $\tilde{\phi}$ 是 (4.9) 的解. 则易于看出函数 u 是问题 1 的容许控制当且仅当它是系统(4.9)的满足控制约束 $\widetilde{\Omega}$ 和端点条件(4.10) 的容许控制. 此外，容许轨线 $\tilde{\phi}$ 具有形式 $\tilde{\phi}=(\phi,\ t_1,\ x_1)$. 于是，若取代价泛函为 $\tilde{J}$，其中

$$\tilde{J}(\tilde{\phi},\ u)=J(\phi,\ u),$$

则问题 1 等价于具有形如(4.8)的端点条件的问题.

5. 等周问题和参数最优化

在有些控制问题中，除了通常的约束外，还附加形如

$$\begin{cases}\displaystyle\int_{t_0}^{t_1}h^i(t,\ \phi(t),\ u(t))dt\leqslant c^i,\quad i=1,\ \cdots,\ q,\\[2mm] \displaystyle\int_{t_0}^{t_1}h^i(t,\ \phi(t),\ u(t))dt=c^i,\ i=q+1,\ \cdots,\ p,\end{cases}\tag{5.1}$$

的约束条件，其中函数 h^i 在 $\mathscr{G}$ 上定义，而常数 c^i 是预先指定的. 形如 (5.1) 的约束条件叫做等周约束. 具有等周约束的问题可以化为没有等周约束的问题，其办法如下:

引入附加状态变量 $x^{n+1},\ \cdots,\ x^{n+p}$，设 $\tilde{x}=(x,\ \bar{x})$，其中 $\bar{x}=(x^{n+1},\ \cdots,\ x^{n+p})$，则 $\tilde{x}$ 表示 E^{n+p} 中的向量. 又设状态方程组为

$$\begin{cases}\dfrac{dx^i}{dt}=f^i(t,\ x,\ u(t)),\ i=1,\ \cdots,\ n,\\[2mm] \dfrac{dx^{n+i}}{dt}=h^i(t,\ x,\ u(t)),\ i=1,\cdots,\ p.\end{cases}\tag{5.2}$$

或
$$\frac{d\tilde{x}}{dt}=\tilde{f}(t,\ x,\ u(t)),$$

其中$\tilde{f}=(f, h)$；控制约束是由方程$\tilde{\Omega}(t, \tilde{x})=\Omega(t, x)$确定的映射给出的；端点条件是满足下列条件的所有的点$(t_0, \tilde{x}_0, t_1, \tilde{x}_1)$组成的集合$\tilde{\mathscr{B}}$：

(i) $(t_0, x_0, t_1, x_1)\in\mathscr{B}$；　(ii) $x_0^i=0$, $i=n+1, \cdots, n+p$；(iii) $x_1^i\leqslant c^i$, $i=n+1, \cdots, n+q$；(iv) $x_1^i=c^i$, $i=n+q+1, \cdots, n+p$. 对于带状态变量$\tilde{x}$的系统，分别用$\tilde{\mathscr{R}}=\mathscr{R}\times E^p$和$\tilde{\mathscr{G}}=\tilde{\mathscr{R}}\times\mathscr{U}$代替$\mathscr{R}$和$\mathscr{G}$.

设(ϕ, u)是问题1的容许对，且满足等周约束(5.1)．令$\tilde{\phi}=(\phi, \bar{\phi})$，其中

$$\bar{\phi}(t)=\int_{t_0}^{t} h(s, \phi(s), u(s))ds, \qquad \bar{\phi}(0)=0.$$

则$(\tilde{\phi}, u)$是状态变量为$\tilde{x}$的系统的容许对．反之，若$(\tilde{\phi}, u)$是以$\tilde{x}$为状态变量的系统的容许对，则(ϕ, u)(其中ϕ是由$\tilde{\phi}$的前n个支量组成的)是问题1的容许对，并且满足等周约束．因此，若对$\tilde{x}$-空间里的问题取$\tilde{J}$作为代价泛函，其中$\tilde{J}(\tilde{\phi}, u)=J(\phi, u)$．我们就可以把具有等周约束(5.1)的问题按问题1的格式写成与之等价的问题．

在问题1中，决定代价泛函和微分方程组(3.1)的函数f^0, $f^1, \cdots, f^n$是看作固定的．但在某些应用中，这些函数还依赖于由我们任意安排的参数向量$w=(w^1, \cdots, w^k)$．例如，在1.4节中的火箭问题，我们可以适当地改变设计而使实际的消耗速度c在某范围$c_0\leqslant c\leqslant c_1$中变化．这时微分方程组(3.1)变成

$$\frac{dx}{dt}=f(t, x, w, u(t)), \qquad w\in W,$$

其中W是E^k中的某个指定集合．这时，对选定的控制u，相应的轨线ϕ一般还依赖于参数向量w的选择．于是，代价泛函的值$J(\phi, u, w)$也将是如此．现在问题是要在W中选取一个参数w^*，对此w^*存在容许对(ϕ^*, u^*)，使得$J(\phi^*, u^*, w^*)\leqslant J(\phi, u, w)$对一切$w\in W$和与之相应的容许对$(\phi, u)$都成立．

刚才提出的这一问题(所谓参数最优化问题)可以在$(n+k+1)$

维空间里重新表达成问题 1 的格式，其办法如下：

引入新的状态变量 $w=(w^1, \cdots, w^k)$，并考虑系统

$$\begin{cases} \dfrac{dx^i}{dt}=f^i(t, x, w, u(t)), & i=1, \cdots, n, \\ \dfrac{dw^i}{dt}=0, & i=1, \cdots, k. \end{cases} \tag{5.3}$$

令 $\tilde{x}=(x, w)$，$\tilde{\mathscr{R}}=\mathscr{R}\times E^k$，$\tilde{\mathscr{G}}=\tilde{\mathscr{R}}\times\mathscr{U}$，$\tilde{\Omega}(t, x, w)=\Omega(t, x)$；端点条件取为

$$\tilde{\mathscr{B}}=\{(t_0, x_0, w_0, t_1, x_1, w_1): \\ (t_0, x_0, t_1, x_1)\in\mathscr{B}, w_0\in W\}.$$

设 $\tilde{J}(\tilde{\phi}, u)=J(\phi, w, u)$. 容易验证，满足状态方程组(5.3)，控制约束 $\tilde{\Omega}$ 和端点条件 $\tilde{\mathscr{B}}$，并使 $\tilde{J}$ 取极小值的问题就等价于参数最优化问题.

6. 与变分法的关系

第一章 1.6 节所述的捷线问题是变分学的简单问题的一个例子，它可叙述如下：设 t 表示纯量，x 和 x' 是 E^n 中的向量，$\mathscr{G}$ 为 (t, x, x')-空间中的区域，f^0 是定义于 $\mathscr{G}$ 上的实函数，$\mathscr{B}$ 是 E^{2n+2} 中的点 (t_0, x_0, t_1, x_1) 组成的给定集合，g 是定义于 $\mathscr{B}$ 上的实函数. 容许轨线是定义于区间 $[t_0, t_1]$ 上的绝对连续函数 ϕ，使得：

(i) $(t, \phi(t), \phi'(t))\in\mathscr{G}$ 在 $[t_0, t_1]$ 上几乎处处成立；

(ii) $(t_0, \phi(t_0), t_1, \phi(t_1))\in\mathscr{B}$； (6.1)

(iii) $t\to f^0(t, \phi(t), \phi'(t))$ 在 $[t_0, t_1]$ 上是 Lebesgue 可积的.

问题是要寻找一个容许弧，它使

$$J(\phi)=g(t_0, \phi(t_0), t_1, \phi(t_1)) \\ +\int_{t_0}^{t_1} f^0(t, \phi(t), \phi'(t))dt$$

取极小值.

变分学的一般简单问题亦如捷线问题一样，可用如下方法表达成控制问题，把 x' 记作 z；亦即令 $u=\phi'$.（记住，z 表示控制变

量而 u 表示控制函数.)变分学的简单问题变成如下控制问题，在满足状态方程组

$$\frac{dx^i}{dt}=u^i(t), \qquad i=1, \cdots, n,$$

端点条件如(6.1)中的(ii)，和控制约束 Ω 的条件下，使

$$g(t_0, \phi(t_0), t_1, g(t_1))+\int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt$$

取极小值，其中

$$\Omega(t, x)=\{z: (t, x, z)\in\mathscr{G}\}.$$

变分学里的 Bolza 问题与上述的简单问题不同，它除了(6.1)之外，还要求容许弧满足微分方程组

$$F^i(t, \phi(t), \phi'(t))=0, \qquad i=1, \cdots, \mu. \tag{6.2}$$

其中 $F^1, \cdots, F^\mu$ 是定义在 $\mathscr{G}$ 上的连续函数，而 $\mu<n$.

在 Bolza 问题的理论中，其必要条件推导通常总假设函数 f^0 和 $F=(F^1, \cdots, F^\mu)$ 在 (t, x, x')-空间中的某一区域 $\mathscr{G}$ 上是属于 $C^{(1)}$ 类的，并且偏导数矩阵 $F_{x'}=(\partial F^i/\partial x'^j), i=1, \cdots, \mu, j=1, \cdots, n$, 在 $\mathscr{G}$ 上处处具有秩数 μ. 因此，在满足

$$F^i(t_2, x_2, x_2')=0, \qquad i=1, \cdots, \mu, \tag{6.3}$$

的任何点 (t_2, x_2, x_2') 的邻域中，我们总能解出 x' 的 μ 个支量，把它们表示成 t, x 和 x' 的其余 $n-\mu$ 个支量的函数，并且 x' 的这 μ 个支量是其变元的 $C^{(1)}$ 函数. 现在假定我们总能按此方式全部解出方程(6.3)，因为支量的次序可重新排列，不妨假设我们解出 x' 的前 μ 个支量(即把 x' 的前 μ 个支量表示成 t, x 和 x' 的后面 $n-\mu$ 个支量的函数)，因此可得

$$x'^i=G^i(t, x, \tilde{x}'), \qquad i=1, \cdots, \mu,$$

其中 $\tilde{x}'=(x'^{\mu+1}, \cdots, x'^n)$. 所以方程(6.2)等价于

$$\phi'^i(t)=G^i(t, \phi(t), \tilde{\phi}'(t)), \qquad i=1, \cdots, \mu,$$

其中 $\tilde{\phi}'=(d\phi^{\mu+1}/dt, \cdots, d\phi^n/dt)$. 令 $m=n-\mu$, $z=(z^1, \cdots, z^m)=(x'^{\mu+1}, \cdots, x'^n)=\tilde{x}'$. 由此可见，在上述前提下，Bolza 问题等价于如下带控制变量 $z=(z^1, \cdots, z^m)$ 的控制问题.

要使之极小化的泛函由方程

$$\bar{J}(\phi, u)=g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1}\bar{f}^0(t, \phi(t), u(t))dt$$

定义，其中

$$\bar{f}^0(t, x, z)=f^0(t, x, G^1(t, x, z), \cdots, G^\mu(t, x, z), z^1, \cdots, z^m).$$

状态方程组为

$$\begin{cases}\dfrac{dx^i}{dt}=G^i(t, x, u(t)), & i=1, \cdots, \mu;\\[2mm] \dfrac{dx^{\mu+i}}{dt}=u^i(t), & i=1, \cdots, m.\end{cases}$$

端点条件由 Bolza 问题的集合 $\mathscr{B}$ 定义. 控制约束 Ω 规定如下:

$$\Omega(t, x)=\{Z: (t, x, G^1(t, x, z), \cdots, G^\mu(t, x, z), z^1, \cdots, z^m)\in\mathscr{G}\}.$$

当然，它还要求容许对 (ϕ, u) 使得映射 $t\to\bar{f}^0(t, \phi(t), u(t))$ 是 L 可积的.

反之，在一定条件下控制问题可写成变分学里的 Bolza 问题. 首先，我们假设对所有 (t, x), $\Omega(t, x)=E^m$, 亦即假设不存在控制约束. 引入新坐标 $y^1, \cdots, y^m$, 并令

$$\frac{dy^i}{dt}=u^i(t), \qquad i=1, \cdots, m.$$

则方程(3.1)可以写成

$$\frac{dx^i}{dt}-f^i\left(t, x, \frac{dy}{dt}\right)=0, \qquad i=1, \cdots, n.$$

如果令

$$F^i(t, x, y, x', y')=x'^i-f^i(t, x, y'), \quad i=1, \cdots, n. \tag{6.4}$$

则控制问题可以写成如下 $(m+n+1)$ 维 (t, x, y)-空间中的 Bolza 问题. 容许弧的集合是定义在 $[t_0, t_1]$ 上的绝对连续函数 $\hat{\phi}=(\phi, \eta)=(\phi^1, \cdots, \phi^n, \eta^1, \cdots, \eta^m)$ 全体，并使得:

(i) $(t, \phi(t), \eta'(t))$ 处于函数 $\hat{f}=(f^0, f)$ 的定义域中;

(ii) $(t_0, \phi(t_0), t_1, \phi(t_1))\in\mathscr{B}$, 而 $\eta(t_0)=0$;

(iii) 函数 $t\to f^0(t, \phi(t), \eta'(t))$ 是 L 可积的;

(iv) $F(t, \hat{\phi}(t), \hat{\phi}'(t))=\phi'(t)-f(t, \phi(t), \eta'(t))=0,$

$$(6.5)$$

在 $[t_0, t_1]$ 上几乎处处成立，其中 $F=(F^1, \cdots, F^n)$. 问题是要使泛函

$$g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1} f^0(t, \phi(t), \eta'(t))dt$$

在容许弧的集合中取极小值.

显然，从(6.4)可见，由控制问题所得到的Bolza问题的 $n\times(m+n)$ 偏导数矩阵 $(F_{x'}, F_{y'})$，其代表性的行是

$$(0, \cdots, 0, 1, 0, \cdots, 0, \partial f^i/\partial y'^1, \cdots, \partial f^i/\partial y'^m),$$

其中元素 1 出现在第 i 列，而前 n 列的其它元素都是 0. 因此 $n\times(m+n)$ 矩阵 $(F_{x'}, F_{y'})$ 如通常 Bolza 问题的必要条件所要求的那样，具有秩数 n.

现在我们假设出现控制约束 Ω，并设集合 $\Omega(t, x)$ 是由不等式组定义的，亦即假设存在 r 个定义于 $\mathscr{G}$ 上的 $C^{(1)}$ 类函数 $R^1, \cdots, R^r$，并定义集合

$$\Omega(t, x)=\{z: R^i(t, x, z)\geqslant 0, i=1, 2, \cdots, r\}.$$

我们对它作进一步限制，并称这些限制为约束规范:

(i) 如果 z 的支量数目 $m\leqslant r$，则对 $\mathscr{G}$ 上任意点 (t, x, z)，函数 $R^1, \cdots, R^r$ 至多有 m 个能在该点取零值.

(ii) 在点 (t, x, z) 处，设 $i_1, \cdots, i_\rho$ 表示使 $R^i(t, x, z)=0$ 的指标集；设 $R_{z,\rho}(t, x, z)$ 表示从矩阵 $R_z(t, x, z)=(\partial R^i(t, x, z)/\partial z^j)(i=1, \cdots, r, j=1, \cdots, m)$ 中取出 $i_1, \cdots, i_\rho$ 行所构成的矩阵，则 $R_{z,\rho}(t, x, z)$ 的秩为 ρ. (6.6)

为了把控制问题表达成 Bolza 问题，可按前面的方式处理. 令 $y'=z$，约束取形式 $R^i(t, \phi(t), \eta'(t))\geqslant 0$，这种限制在古典的 Bolza 问题的表达形式中是不出现的. 但是，我们可以把有此种约束的变分问题写成 Bolza 问题，这只要引入新变量 $w=(w^1, \cdots, w^r)$ 和 r 个附加状态方程

$$R^i(t, x, y') - (w'^i)^2 = 0, \qquad i = 1, \cdots, r \tag{6.7}$$

就行了．现在的 Bolza 问题是：在满足微分方程组(6.5)和(6.7)，端点条件

$$(t_0, \phi(t_0), t_1, \phi(t_1)) \in \mathscr{B}, \ \eta(t_0) = 0, \ \omega(t_0) = 0$$

(其中函数 ω 是相应于变量 w 的容许弧的支量)的条件下，使泛函

$$g(t_0, \phi(t_0), t_1, \phi(t_1)) + \int_{t_0}^{t_1} f^0(t, \phi(t), \eta'(t))dt$$

取极小值．

设

$$\begin{cases} F^i(t, x, y, w, x', y', w') = x'^i - f^i(t, x, y'), & i = 1, \cdots, n; \\ F^{n+i}(t, x, y, w, x', y', w') = R^i(t, x, y') - (w'^i)^2, & i = 1, \cdots, r. \end{cases} \tag{6.8}$$

我们要证明$(n+r)\times(n+m+r)$矩阵

$$M = \left(\frac{\partial F^q}{\partial x'^j}, \frac{\partial F^q}{\partial y'^k}, \frac{\partial F^q}{\partial w'^s}\right) = (F_{x'}, F_{y'}, F_{w'})$$

象通常 Bolza 问题的理论所要求那样，具有秩数 $n+r$.

利用公式(6.4)和(6.8)直接计算，即可看出

$$M = \begin{pmatrix} I & -f_{y'} & O_1 \\ O_2 & R_{y'} & -w \end{pmatrix},$$

其中 I 为 $n\times n$ 单位矩阵；$f_{y'}$ 是 $n\times m$ 矩阵，其代表元为 $\partial f^i/\partial y'^k$；$O_1$ 是 $n\times r$ 零矩阵；O_2 是 $r\times n$ 零矩阵；$R_{y'}$ 是 $r\times m$ 矩阵，其代表元为 $\partial R^i/\partial y'^k$；而 w 为 $r\times r$ 对角矩阵，其对角线上的元素为 $2w'^i$.

从矩阵 M 的形状可见，显然为了证明它的秩为 $n+r$，只须证明 $r\times(m+r)$ 矩阵$(R_{y'}, -w)$的秩为 r．而为了证明这点，我们假设使 $R^{i_\alpha}(t, x, y')=0$ 的指标 $i_1, \cdots, i_\rho$ 是指标 $1, \cdots, \rho$．令$(R_{y'})_\rho$ 表示由 $R_{y'}$ 的前 ρ 行构成的子矩阵，而$(R_{y'})_{r-\rho}$表示由 $R_{y'}$ 的后 $r-\rho$ 行构成的子矩阵．因此，若 $i>\rho$，则 $R^i(t, x, y')>0$；若 $i\leqslant\rho$，则 $R^i(t, x, y')=0$．于是，从

$$(w'^i)^2 = R^i(t, x, y'), \qquad i = 1, \cdots, r,$$

即可得到 $2w'^i=0$, 当 $i\leqslant\rho$; $2w'^i\neq0$, 当 $i>\rho$.

所以

$$(R_{y'},\ -w)=\begin{pmatrix}(R_{y'})_\rho, O_3, O_4\\ (R_{y'})_{r-\rho}, O_5, D\end{pmatrix},$$

其中 D 是 $(r-\rho)\times(r-\rho)$ 对角矩阵, 其对角线上的元素 $2w'^i\neq0$, $i>\rho$. 而 O_3, O_4, O_5, 是零矩阵. 由约束条件(6.6), 矩阵 $(R_{y'})_\rho$ 的秩为 ρ. 因为 D 的秩是 $r-\rho$, 可见 $(R_{y'},\ -w)$ 的秩为 r, 此即为所要求的. 结论证毕.

第三章　具有凸性假设的存在定理

1. 引　　言

本章我们要建立具有某种凸性条件的最优控制问题的基本存在定理. 全部推导的关键性定理是定理 4.1. 它保证某些轨线集合的致密性和与下半连续性有关的性质. 此外, 定理 4.1 在第四章和第六章也会用到.

定理 4.1 和建立在它之上的存在定理允许约束集依赖于时间和状态, 并要求满足由 Cesari 所提出的某种条件. 虽然这种条件在考察给定的例子中常常是不好验证的, 但是对在应用中具有意义的一大类问题, 可以证明它是满足的. 对于约束集只依赖于时间而不依赖于状态、条件比较容易验证的存在定理的问题将在第 8 节中给出. 对于这些定理的 Cesari 条件, 可用广义的 Lipschitz 条件或用极小化序列中的控制位于某个 L_p 空间 ($1\leqslant p\leqslant\infty$) 的一个固定球内的要求来代替. 对应用具有意义的问题, 这些条件通常是能满足的. 第 4 节和第 8 节的一些定理有好些重迭的部分, 但不互相包含.

本章另一个很重要的定理是定理 7.1, 它是 Filippov 引理的推广. 定理 6.2 是通常变分问题中的一个经典存在定理.

主要兴趣在应用的读者, 第一次阅读时可集中注意力于第 2 节、第 3 节到定义 3.2, 第 5 节到引理 5.2, 以及习题 5.1、6.3、6.4、6.5 和 6.6.

我们用以建立最优控制存在的数学工具, 使得我们能够保证其存在的最优控制只是可测函数. 在实际问题中, 可能希望最优控制的特性有更多的正规性. 在第五章, 我们将得到描述最优控制的一些定理. 好在有实际意义的问题中, 这些所附加的信息使我们能够得出比仅仅是可测要强而且可以实现的最优控制.

2. 最优控制的不存在和不唯一性

问题1的叙述中假定了容许对集合是非空的．给定一个状态方程组、端点条件和控制约束，并不保证容许对集合是非空的．下面用简单的例子来强调这点．

例 2.1 设 x 是一维的．状态方程是

$$\frac{dx}{dt}=u(t), \tag{2.1}$$

设 $\mathscr{B}$ 由单点 $(t_0, x_0, t_1, x_1)=(0, 0, 1, 2)$ 组成，而控制约束为

$$\Omega(t, x)=\{z: |z|\leqslant 1\}.$$

于是控制集合是定义在 $[0, 1]$ 上的实值可积函数 u 的全体．一个容许控制 $u(t)$ 是指对几乎所有 $t\in[0, 1]$，满足不等式 $|u(t)|\leqslant 1$，并使系统从时刻 $t=0$ 的状态 $x_0=0$ 转移到时刻 $t_1=1$ 的状态 $x_1=2$．显然，从 (2.1) 可见，对每一个控制 u，相应地有唯一轨线 ϕ，使得 $\phi(0)=0$，此轨线由 $\phi(t)=\int_0^t u(s)ds$ 给出．由于

$$|\phi(1)|\leqslant\int_0^1 |u(t)|dt\leqslant\int_0^1 dt=1,$$

我们要求使 $\phi(1)=2$ 的容许性变成不可能，所以 (ϕ, u) 不是容许对．

容许对的存在性问题是与可控性有关的问题，此部分内容不在本书中叙述．

如果容许控制是非空的，最优控制也未必存在．下面的例子说明最优控制不存在．

例 2.2 设 x 是一维的，状态方程是 $\frac{dx}{dt}=u(t)$，$\mathscr{B}$ 由单点 $(t_0, x_0, t_1, x_1)=(0, 1, 1, 0)$ 组成．$\Omega(t, x)=E^1$．设

$$J(\phi, u)=\int_0^1 t^2u^2(t)dt. \tag{2.2}$$

控制集合是 $L_1[0, 1]$ 中的函数的集合，每一个控制 u 有相应的唯一的满足 $\phi(0)=1$ 的轨线 ϕ，此轨线由

$$\phi(t)=1+\int_0^t u(s)ds$$

给出。

对于每一个 $0<\varepsilon<1$，规定 $u_\varepsilon(t)$ 如下：

$$u_\varepsilon(t)=\begin{cases}0, & \text{当 } \varepsilon\leqslant t\leqslant 1,\\ -\varepsilon^{-1}, & \text{当 } 0\leqslant t<\varepsilon.\end{cases}$$

设 ϕ_ε 表示对应于 u_ε 且满足 $\phi_\varepsilon(0)=1$ 的唯一轨线。因为 $\phi_\varepsilon(1)=0$ 且 tu_ε 属于 $L_2[0, 1]$，可见 $(\phi_\varepsilon, u_\varepsilon)$ 是容许对。于是容许对类 $\mathscr{A}$ 是非空的。此外，

$$J(\phi_\varepsilon, u_\varepsilon)=\int_0^\varepsilon t^2\varepsilon^{-2}dt=\varepsilon/3.$$

因为对所有容许对 (ϕ, u)，$J(\phi, u)\geqslant 0$。由此可见 $0=\inf\{J(\phi, u):(\phi, u)\in\mathscr{A}\}$。由(2.2)显然可知，当且仅当 $u(t)=0$ 在 $[0, 1]$ 上几乎处处成立时，$J(\phi, u)=0$。但 $u^*=0$ 不是容许控制，因为若 ϕ^* 是对应于 $u^*=0$ 且满足 $\phi^*(0)=1$ 的唯一轨线，则 $\phi^*(1)=1$，而不能满足所要求的 $\phi^*(1)=0$。

例 2.2(a) 设例 2.2 中的其它条件不变，只把控制约束改为

$$\Omega(t, x)=\left\{z:|z|\leqslant\frac{1}{t}\right\}, \text{ 当 } 0<t\leqslant 1;$$

$$\Omega(0, x)=E^1.$$

则例 2.2 的论述仍然正确，最优控制不存在。

例 2.2(b) 设例 2.2 的假定成立，除了控制约束改为

$$\Omega(t, x)=\left\{z:|z|\leqslant\frac{1}{t}\right\}, \text{ 当 } 0<t\leqslant 1;$$

$$\Omega(0, x)=0.$$

现在如果定义

$$u_\varepsilon(t)=\begin{cases}0, & \text{当 } \varepsilon\leqslant t\leqslant 1,\\ -\varepsilon^{-1}, & \text{当 } 0<t<\varepsilon,\\ 0, & \text{当 } t=0.\end{cases}$$

用象例 2.2 那样的论证，我们再次得到最优控制不存在的实例。

例 2.3 设状态方程为

$$\begin{cases}\dfrac{dx^1}{dt}=u^1(t),\\ \dfrac{dx^2}{dt}=u^2(t),\\ \dfrac{dx^3}{dt}=1.\end{cases}$$

端点条件为 $\mathscr{T}_0$: $(t_0,\ x_0^1,\ x_0^2,\ x_0^3)=(0,\ 0,\ 0,\ 0)$,

$\mathscr{T}_1$: $(t_1,\ x_1^1,\ x_1^2,\ x_1^3)=(1,\ 0,\ 0,\ 1)$.

设 $\Omega(t,\ x)$ 为

$$\Omega(t,\ x)\equiv\{z=(z^1,\ z^2):(z^1)^2+(z^2)^2=1\}.$$

而性能指标为

$$J(\phi,\ u)=\int_0^1[(\phi^1)^2+(\phi^2)^2]dt.$$

为了看出容许控制是非空的，我们首先指出由

$$u_k(t)=(u_k^1(t),\ u_k^2(t))=(\sin 2\pi kt,\ \cos 2\pi kt),\ k=\pm 1,\ \pm 2,\ \cdots$$

定义的任何控制 u_k 满足条件 $(u_k^1(t))^2+(u_k^2(t))^2=1$. 以

$$\phi_k^1(t)=(1-\cos 2\pi kt)/2\pi k,$$
$$\phi_k^2(t)=(\sin 2\pi kt)/2\pi k,$$
$$\phi_k^3(t)=t$$

为支量的轨线 $\phi_k(t)$ 是对应于 $u_k(t)$ 且满足端点条件的唯一轨线.

因为对所有容许的 u, $J(\phi,\ u)\geqslant 0$. 又由于

$$J(\phi_k,\ u_k)\leqslant(\pi k)^{-2},$$

可见 $0=\inf\{J(\phi,\ u):(\phi,\ u)\in\mathscr{A}\}$. 因此，若存在最优控制 u^*，则必有 $\phi^{*1}(t)=\phi^{*2}(t)=0$ 几乎处处成立，这只有 $u^{*1}=u^{*2}=0$ 几乎处处成立才有可能．这样的控制不满足 $(u^1)^2+(u^2)^2=1$. 所以本例的最优控制不存在.

例 2.4 设 x 是纯量函数，状态方程为

$$\frac{dx}{dt}=2x^2(1-t)-1+u(t),\tag{2.3}$$

$\Omega(t,\ x)=\{z:|z|\leqslant 1\}$，端点条件由 $\mathscr{T}_0=\{(0,\ x_0):0\leqslant x_0\leqslant 1\}$ 和 $\mathscr{T}_1=\{(a,\ x_1):0\leqslant x_1\leqslant(1-a)^{-2}\}$ 给出，其中 a 是大于1的固定

数. 设 $J(\phi, u)=-\phi(a)$. 因此, 若 u 是一容许控制, 而 ϕ 是相应的一容许轨线, 要求在所有的容许对 (ϕ, u) 中寻找一对 (ϕ^*, u^*), 它使 $\phi(a)$ 取最大值.

本问题的容许控制集合是使得 $|u(t)|\leqslant 1$ 几乎处处成立且在 $[0, a]$ 上可测的函数 u 的子集. 如果把 $u(t)=1$ 代入 (2.3) 的右边, 即得

$$\frac{dx}{dt}=2x^2(1-t). \tag{2.4}$$

微分方程(2.4)满足初始条件 $\phi(0)=x_0$, $(x_0\neq 0)$ 的解为

$$\phi(t)=[(1-t)^2+c]^{-1}, \tag{2.5}$$

其中 $c=(1-x_0)/x_0$. 而微分方程(2.4)满足初始条件 $\phi(0)=0$ 的解是 $\phi(t)\equiv 0$. 对应于 $u(t)=1$ 的轨线场如图 1 所示. 相应于区间 $0<x_0\leqslant 1$ 中的初始点 x_0 的常数 $c\geqslant 0$. 注意, 如果 $x_0=1$, 则 $c=0$, 而 $u(t)=1$ 不是容许控制.

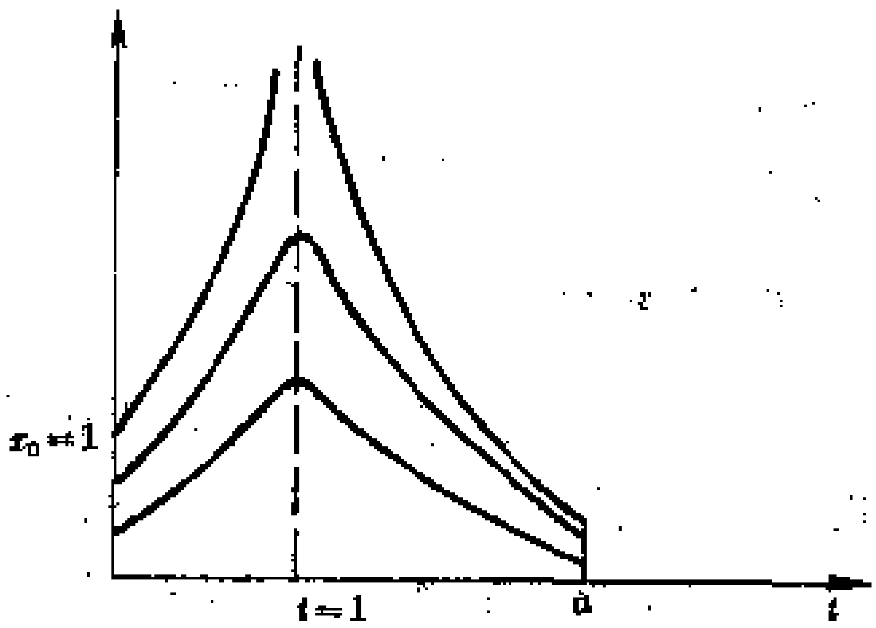

图 1

令 $\mathscr{F}$ 表示相应于控制 $u(t)=1$ 和初始条件 $x_0\in[0, 1)$ 的轨线场. 注意 $\mathscr{F}$ 并不包含在 $t_0=0$ 时从 $x_0=1$ 出发的轨线. 从(2.3)和轨线场 $\mathscr{F}$ 的性质显然可知, 若最优对 (ϕ^*, u^*) 存在, 并且 $\phi^*(0)=x_0<1$, 则必有 $u^*(t)=1$ 几乎处处成立. 而从(2.5)(参看图 1)可见 $u^*=1$ 和 $0\leqslant x_0<1$ 不可能是最优的. 因为若取新的初始状态 x_0', 其中 $x_0<x_0'<1$, 则相应于 x_0' 的方程 (2.4) 的解 ϕ 将给出 $\phi(a)>\phi^*(a)$. 另一方面, 最优轨线不能以 $x_0=1$ 为初始点, 因为如果 $x_0=1$, 则 $u(t)\equiv 1$ 不是容许控制. 此外, 一旦在一个正测度集合上取 $u(t)<1$, 其相应的轨线将进入 $\mathscr{F}$ 的内部, 那么可以修改控制以使值 $\phi(a)$ 增加. 论证的严格表述留给读者.

我们给出最优控制可以多于一个的例子来结束这一节.

例 2.5 设 x 是一维的. 状态方程是 $\frac{dx}{dt}=u(t)$. $\mathscr{B}$ 由单点 $(t_0, x_0, t_1, x_1)=(0, 0, 1, 0)$ 组成. $\Omega(t, x)=\{z: |z|\leqslant 1\}$, 且

$$J(\phi, u)=\int_0^1 (1-u^2(t))dt.$$

显然, $J(\phi, u)\geqslant 0$. 规定控制 u_1^* 如下: $u_1^*=1$, 当 $0\leqslant t<\frac{1}{2}$ 和 $u_1^*(t)=-1$ 当 $\frac{1}{2}\leqslant t\leqslant 1$. 则 u_1^* 是容许的, 且 $J(\phi_1^*, u_1^*)=0$, 其中 ϕ_1^* 是对应于 u_1^* 的唯一轨线. 从而 u_1^* 是最优的. 我们现在证明存在无穷多个最优控制. 对每个整数 $n=1, 2, 3, \cdots$, 定义控制 u_n^* 如下:

$$u_n^*(t)=(-1)^k, \text{当} \frac{k}{2^n}\leqslant t<\frac{k+1}{2^n}, k=0, 1, 2, \cdots, 2^n-1.$$

则对于每个整数 n, u_n^* 是容许的, 且 $J(\phi^*, u^*)=0$, 其中 ϕ_n^* 是对应于 u_n^* 的轨线, 所以每个 u_n^* 是最优的.

3. 凸性条件、正则性条件与弱 L_1 收敛性条件

本节介绍某种凸性条件和正则性条件, 它们在讨论最优控制和最优轨线的存在性时是必需的.

设 (t_0, x_0) 是 $\mathscr{R}$ 中一点, $\delta>0$. 设 $N_\delta(t_0, x_0)$ 表示 (t_0, x_0) 的相对闭 δ 邻域, 因此 $N_\delta(t_0, x_0)=\{(t, x):(t, x)\in\mathscr{R}, \operatorname{dist}((t, x), (t_0, x_0))\leqslant\delta\}$, 其中 $\operatorname{dist}(a, b)$ 表示 a、b 二点间的欧氏距离.

设 A 是把 $\mathscr{R}$ 的点 (t, x) 映成 E^k 的子集 $A(t, x)$ 的映射. 则 $A(N_\delta(t_0, x_0))$ 的含义如下:

$$A(N_\delta(t_0, x_0))=\cup[A(t, x):(t, x)\in N_\delta(t_0, x_0)].$$

定义 3.1 映射 A 叫做在 $\mathscr{R}$ 中的点 (t_0, x_0) 处是上半连续的, 如果

$$\bigcap_{\delta>0} \operatorname{cl} A(N_\delta(t_0, x_0))\subseteq A(t_0, x_0), \tag{3.1}$$

其中 cl 表示闭包.

注意到与(3.1)相反的包含关系总是成立的, 故若在(3.1)中用等号代替 $\subseteq$, 可得到等价的定义. 因此若 A 在 (t_0, x_0) 处是上

半连续的，则 $\Lambda(t_0, x_0)$ 必为闭集。

映射 Λ 叫做在 $\mathscr{R}$ 上是上半连续的，如果它在 $\mathscr{R}$ 的每一点是上半连续的。

上半连续映射的一个例子是由 $\Omega(t, x)=U$ 定义的映射 Ω，其中 U 是固定的闭集。虽然这个例子很平凡，但它却很重要，因为在许多应用中约束集 $\Omega(t, x)$ 不依赖于 (t, x)，而是固定的。上半连续映射的另一个例子是例 2.2(a) 的映射 Ω，它在 $(t_0, x_0)=(0, 0)$ 处是上半连续的。因为对每个 $\delta>0$，我们有 $\Omega(N_\delta(0, 0))=E^1$，由于 $\Omega(0, 0)=E^1$，(3.1) 式就立刻得到。我们请读者验证 Ω 在其它点上也是上半连续的。映射在一点上不是上半连续的例子是例 2.2(b) 的映射 Ω 在点 $(0, 0)$ 的情况。因为对每一个 $\delta>0$，总有 $\Omega(N_\delta(0, 0))=E^1$，但 $\Omega(0, 0)=0$，所以 (3.1) 式不成立。

下面是在闭集 $\mathscr{R}$ 上为上半连续映射的一个等价描述，在存在性定理的证明中将要用到它。

引理 3.1 设 $\mathscr{R}$ 是闭集，映射 Λ 在 $\mathscr{R}$ 上为上半连续的充要条件是集合 $\Delta=\{(t, x, \lambda):\lambda\in\Lambda(t, x),\ (t, x)\in\mathscr{R}\}$ 是闭的。

首先假设 (3.1) 成立。设 $\{(t_n, x_n, \lambda_n)\}$ 是 Δ 中收敛到点 (t_0, x_0, λ_0) 的叙列。因此，$\lambda_n\in\Lambda(t_n, x_n)$ 且 $(t_n, x_n)\to(t_0, x_0)$。因为 $\mathscr{R}$ 是闭的，$(t_0, x_0)\in\mathscr{R}$。此外，对每个 $\delta>0$，存在整数 $n(\delta)$，使得若 $n>n(\delta)$，则 $(t_n, x_n)\in N_\delta(t_0, x_0)$。因此 $\lambda_0\in\operatorname{cl}\Lambda(N_\delta(t_0, x_0))$。但 δ 是任意的，所以 $\lambda_0\in\bigcap_{\delta>0}\operatorname{cl}\Lambda(N_\delta(t_0, x_0))$。从而由 (3.1) 式，$\lambda_0\in\Lambda(t_0, x_0)$，故 Δ 是闭的。

反之，假如 Δ 是闭的，并且对每一个 $\delta>0$，$\lambda_0\in\operatorname{cl}\Lambda(N_\delta(t_0, x_0))$。则存在 $\mathscr{R}$ 中的点列 $\{(t_n, x_n)\}$、正数列 $\{\delta_n\}$ 和点列 $\{\lambda_n\}$，使得如下各项成立：(i) $\delta_n\to0$；(ii) $(t_n, x_n)\in N_{\delta_n}(t_0, x_0)$；(iii) $\lambda_n\in\Lambda(t_n, x_n)$；(iv) $\lambda_n\to\lambda_0$，因此 $(t_n, x_n, \lambda_n)\to(t_0, x_0, \lambda_0)$。因为 $\mathscr{R}$ 和 Δ 是闭的，$\lambda_0\in\Lambda(t_0, x_0)$，所以 (3.1) 式成立。

考虑 $\hat{f}(t, x, z)=(f^0(t, x, z), \cdots, f^n(t, x, z))$，

其中 $f^0, f^1, \cdots, f^n$ 是定义在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上的实函数。对固定的

(t, x)，当 z 取遍整个集 $\Omega(t, x)$ 时，向量 $\hat{f}(t, x, z)$ 将描出 E^{n+1} 中的一个集合，我们用 $\mathscr{Q}(t, x)$ 表示这个集合，于是

$$\mathscr{Q}(t, x)=\{\hat{y}=(y^0, y): y^0=f^0(t, x, z),\\ y=f(t, x, z),\ z\in\Omega(t, x)\}. \quad (3.2)$$

我们简记 $\mathscr{Q}(t, x)=\hat{f}(t, x, \Omega(t, x))$。我们还需要考虑和 $\mathscr{Q}(t, x)$ 有关的如下集合

$$\begin{aligned}\mathscr{Q}^+(t, x)&=\{\hat{y}=(y^0, y): y^0\geqslant f^0(t, x, z),\\ &\qquad y=f(t, x, z),\ z\in\Omega(t, x)\}\\ &=\{\hat{y}=(y^0, y): y^0\geqslant f^0(t, x, \Omega(t, x)),\\ &\qquad y=f(t, x, \Omega(t, x))\}.\end{aligned} \quad (3.3)$$

我们用例2.2和例2.3来说明这些概念的含意．在例2.2中

$$\mathscr{Q}(t, x)=\{(y^0, y): y^0=t^2z^2,\ y=z,\ z\in E^1\},$$

$$\mathscr{Q}^+(t, x)=\{(y^0, y): y^0\geqslant t^2z^2,\ y=z,\ z\in E^1\}.$$

如图2所示，对固定的 (t, x)，$t\neq0$，$\mathscr{Q}(t, x)$ 是抛物线 $y^0=t^2y^2$，它是阴影区域的边界．而 $\mathscr{Q}^+(t, x)$ 是此抛物线加上阴影的区域．显然集 $\mathscr{Q}(t, x)$ 不是凸的，而集 $\mathscr{Q}^+(t, x)$ 是凸的．如果 $t=0$，则 $\mathscr{Q}(t, x)$ 为 y 轴，而 $\mathscr{Q}^+(t, x)$ 为上半平面 $y^0\geqslant0$。

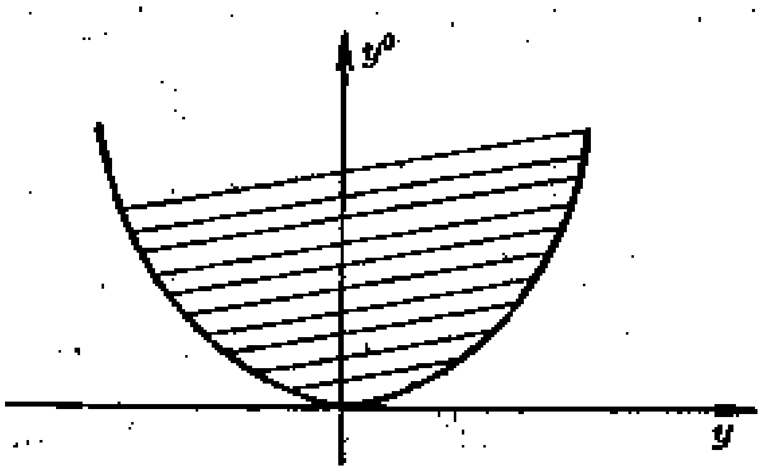

图 2

在例2.3中，

$$\mathscr{Q}(t, x)=\{(y^0, y): y^0=(x^1)^2+(x^2)^2,\ y^1=z^1,\ y^2=z^2,\\ y^3=1;\ (z^1)^2+(z^2)^2=1\},$$

$$\mathscr{Q}^+(t, x)=\{(y^0, y): y^0\geqslant(x^1)^2+(x^2)^2,\ y^1=z^1,\ y^2=z^2,\\ y^3=1;\ (z^1)^2+(z^2)^2=1\}.$$

如果固定 (t, x)，并取 $\mathscr{Q}(t, x)$ 与超平面 $y^3=1$ 的交集便得到在平面 $y^0=(x^1)^2+(x^2)^2$ 中的半径为1、中心在原点的圆周 C（见图3）。而 $\mathscr{Q}^+(t, x)$ 与超平面 $y^3=1$ 的交集是竖立在圆周 C 上的正圆柱面．因此不论是 $\mathscr{Q}(t, x)$ 或 $\mathscr{Q}^+(t, x)$，都不是凸的。

定义 3.2 映 $\mathscr{R}$ 到 E^k 的子集的映射 Λ 称为在点 $(t_0, x_0)\in\mathscr{R}$

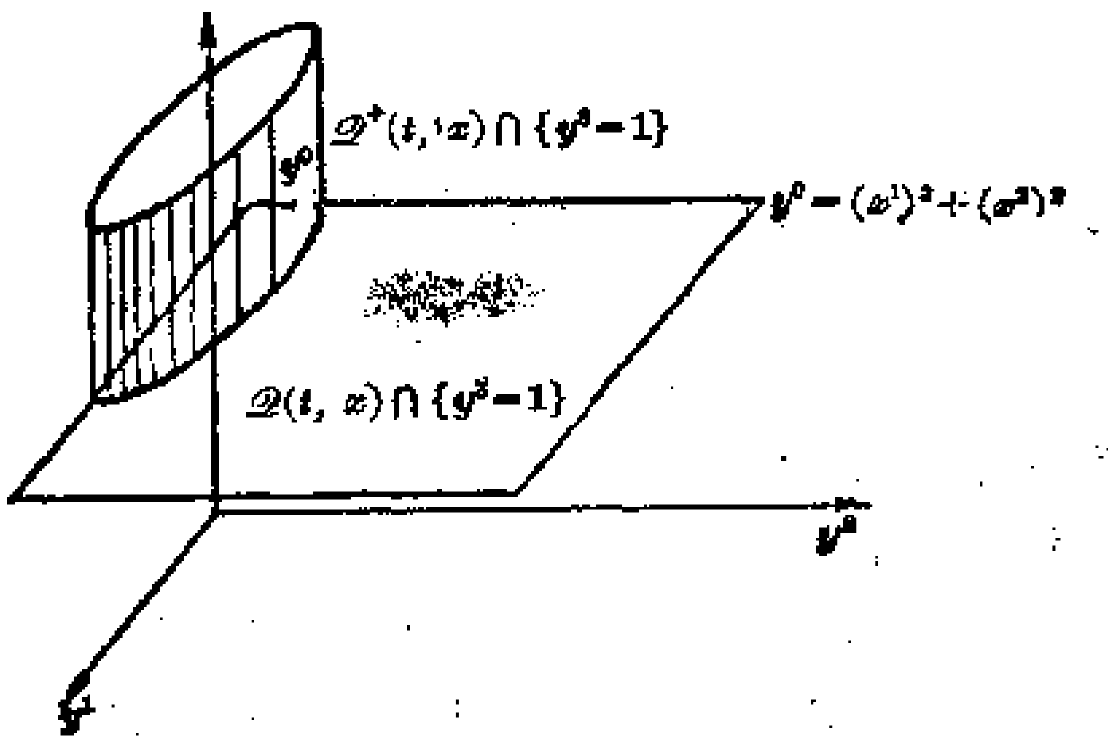

图 3

具有 Cesari 性质，如果

$$\bigcap_{\delta>0}\mathrm{cl}\ \mathrm{co}\Lambda(N_\delta(t_0, x_0))\subseteq\Lambda(t_0, x_0), \tag{3.4}$$

其中用 cl co A 表示 A 的凸壳的闭包．我们称映射 Λ 在 $\mathscr{R}$ 上具有 Cesari 性质，如果 Λ 在 $\mathscr{R}$ 上的每一点都具有 Cesari 性质.

注意到与(3.4)式相反的包含关系总是成立的，因此(3.4)式用等号代替包含关系是等价的．由此可见，如果 Λ 在(t_0, x_0)满足 Cesari 性质，则 $\Lambda(t_0, x_0)$必为闭凸集．因此在例 2.3 中所定义的映射 $\mathscr{Q}^+(t, x)$: $(t, x)\to\mathscr{Q}^+(t, x)$ 不可能满足 Cesari 性质，因为 $\mathscr{Q}^+(t, x)$不是凸的．在例 2.2 和 2.2(a) 中映射 $\mathscr{Q}^+$ 满足 Cesari 性质．另一方面，例 2.2(b)中所定义的映射 $\mathscr{Q}^+$ 不满足 Cesari 性质，虽然所有集合 $\mathscr{Q}^+(t, x)$是凸的．为了看清这一点，我们注意到对每一个 $\delta>0$，有 $\mathscr{Q}^+(N_\delta(0, 0))=\{(y^0, y): y^0>0\}\cup\{(0, 0)\}=\mathrm{co}\,\mathscr{Q}^+(N_\delta(0, 0))$，因此 $\bigcap_{\delta>0}\mathrm{cl}\ \mathrm{co}\,\mathscr{Q}^+(N_\delta(0, 0))$ 是整个闭上半平面；而 $\mathscr{Q}^+(0, 0)=\{(y^0, y): y^0\geqslant0, y=0\}$.

关于使函数 $\hat{f}$ 和 Ω 满足 Cesari 性质的充分条件，将与第 5 节和第 6 节的存在定理一道给出.

设 (t_0, x_0) 为 $\mathscr{R}$ 中的点．用 $N_{x\delta}(t_0, x_0)$ 表示 $\mathscr{R}$ 中使得 $|x-x_0|\leqslant\delta$ 的点(t_0, x)全体．令

$$A(N_{x\delta}(t_0, x_0)) = \bigcup\{A(t_0, x) : (t_0, x) \in N_{x\delta}(t_0, x_0)\}$$[注].

定义 3.3 映射 A 称为在点(t_0, x_0)满足弱 Cesari 性质，如果

$$\bigcap_{\delta>0} \operatorname{cl}\operatorname{co} A(N_{x\delta}(t_0, x_0)) \subseteq A(t_0, x_0). \tag{3.5}$$

由于(3.5)式的左边总是包含右边，因此(3.5)式可用等号代替包含关系．我们让读者验证，若 A 在点(t_0, x_0)上满足 Cesari 性质，则 A 在点(t_0, x_0)满足弱的 Cesari 性质．

任何与 x 无关的映射 $\mathscr{Q}^+$，假如 $\mathscr{Q}^+(t, x)$是闭且凸的，则它必满足弱 Cesari 性质．因此，例 2.2(b)的映射 $\mathscr{Q}^+$ 在点 (0, 0) 上满足弱 Cesari 性质，虽然它在点(0, 0)上不具有 Cesari 性质．

我们将要用到轨线相互靠近的度量．因为初始时刻和终止时刻不是固定的，故我们引入如下的度量：设 $\mathscr{X}$ 表示从 E^1 中的任何致密区间到 E^n 的连续函数族．若函数 $x\in\mathscr{X}$，定义在区间 $[a, b]$ 上，而函数 $y\in\mathscr{X}$，定义在区间$[c, d]$上，我们用如下方式把 x 的定义区间延拓到 $(-\infty, \infty)$：

$$x(t) = \begin{cases} x(a), & \text{当 } t\leqslant a, \\ x(b), & \text{当 } t\geqslant b; \end{cases}$$

并用类似的方式把 y 的定义区间延拓到$(-\infty, \infty)$．然后定义

$$\rho(x, y) = |a-c| + |b-d| + \max\{|x(t)-y(t)| : -\infty<t<+\infty\}. \tag{3.6}$$

我们可以断定 ρ 是一种度量，并且 $\mathscr{X}$ 是在此度量下的完备度量空间．

习题 3.1 证明上面刚才所作的论断．

我们列举积分学和泛函分析的入门教程中有时省略了的一些熟知的事实和定义作为本节的结束．这些事实和定义在存在性定理的证明中将要用到．对于许多主要结果的证明我们建议读者参看有关的原著．

定义 3.4 $L_1[a, b]$中函数 f 的集合 $\mathscr{F}$ 称为具有等度绝对连续的积分，如果给定 $\varepsilon>0$，存在 $\delta>0$，使得对所有 $[a, b]$ 中的具

[注] 原书误为 $A(N_{x\delta}(t_0, x_0)) = \{A(t_0, x) : (t_0, x) \in N_{x\delta}(t_0, x_0)\}$．——译者注．

有 meas$(E)<\delta$ 的 Lebesgue 可测集 E，和所有 $f\in\mathscr{F}$ 有

$$\left|\int_E f\,dt\right|<\varepsilon,$$

其中 $[a, b]=\{t: a\leqslant t\leqslant b\}$.

注意，因为 $[a, b]$ 是有限区间，并且我们是考虑 Lebesgue 测度. 由此可见如果 $\mathscr{F}$ 中的所有函数 f 具有等度绝对连续的积分，则存在常数 $K>0$ 使得对所有 $f\in\mathscr{F}$，有

$$\int_a^b |f|\,dt<K. \tag{3.7}$$

亦即集合 $\mathscr{F}$ 在 $L_1[a, b]$ 中有界.

定义 3.5 定义于 $[a, b]$ 上的绝对连续函数 f 的集合 $\mathscr{F}$ 称为等度绝对连续的. 如果给定 $\varepsilon>0$，存在 $\delta>0$，使得含于 $[a, b]$ 中的任何有限个两两不相重迭的区间 $[\alpha_i, \beta_i]$，当 $\sum_i|\beta_i-\alpha_i|<\delta$ 时，不等式 $\sum_i|f(\beta_i)-f(\alpha_i)|<\varepsilon$ 对所有 $f\in\mathscr{F}$ 成立.

我们请读者证明绝对连续函数族为等度绝对连续的，当且仅当族中所有函数的导数 f' 具有等度绝对连续的积分.

对我们来说，等度绝对连续性概念的重要性在于如下定理.

定理 3.1 设 $[a, b]$ 是一有限区间，且 $\{f_n\}$ 是 $L_1[a, b]$ 中的函数列. 则函数列 $\{f_n\}$ 弱收敛于 $L_1[a, b]$ 中的函数 f，当且仅当下列条件满足：

（i）函数列 $\{f_n\}$ 具有等度绝对连续的积分；

（ii）对 $[a, b]$ 中的每一个 t，

$$\lim_{n\to\infty}\int_a^t f_n(s)\,ds=\int_a^t f(s)\,ds.$$

我们将概述此定理的证明，至于一些详细的论述建议读者参考有关的教科书.

首先考虑条件(i)(ii)的必要性. f_n 弱收敛于 f 意味着对于每个定义在 $[a, b]$ 上的有界可测函数 g 有

$$\int_a^b gf_n\,dt\to\int_a^b gf\,dt, \tag{3.8}$$

因此，取 g 是 $[a, t]$ 上的特征函数即得到(ii)．其次，若取 g 是可测集 E 的特征函数，则 (3.8) 式的积分在任何可测集 E 上成立．那么条件 (i) 可由 Natanson[46]156 页[注]中定理 3 的推论 1 得出．

现在假定条件(i)和 (ii) 成立．我们已注意到条件 (i) 隐含着用 f_n 代替 f 时，(3.7)式成立，亦即叙列 $\{f_n\}$ 在 $L_1[a, b]$ 中有界．条件(ii) 隐含着当积分区间取在包含$[a, b]$ 中的任何区间 $[t', t'']$ 上时，条件(ii)也成立．由此和条件 (i) 可知，当积分区间取为$[a, b]$ 中任何可测集 E 时，条件(ii)成立，从而可见(3.8) 式对任何阶梯函数 g 成立． 若 g 是任意的可测函数，则 g 是阶梯函数列 $\{\sigma_k\}$ 的几乎处处收敛的极限函数．由 Egorov 定理，对每一个 $\delta>0$，存在可测集 E，满足 $\text{meas}(E)<\delta$，使得在 $[a, b]-E$ 上 $\sigma_k\to g$ 一致地成立．从上面的讨论以及在(3.8)式中用阶梯函数代替 g，$\{f_n\}$在 $L_1[a, b]$ 一致有界和 $\{f_n\}$ 的等度绝对连续性，可见 (3.8) 式对任意有界可测函数 g 是正确的．

我们将用到的另一个事实是下面的定理．

定理 3.2 设 $\mathscr{Y}$ 为 Banach 空间，$\{y_n\}$ 是 $\mathscr{Y}$ 中的元素列，它弱收敛于 $\mathscr{Y}$ 中的元素 y_0． 则对给定的整数 n 和 $\varepsilon>0$，存在有限个实数的集 $\{\alpha_i\}$，$\alpha_i\geqslant 0$，$\sum\limits_i \alpha_i=1$，使得 $\|y_0-\sum\limits_i \alpha_i y_{n+i}\|<\varepsilon$，其中 $\|\cdots\|$ 表示 $\mathscr{Y}$ 中的模．

本定理可以从 Hahan-Banach 定理的推论得知．$\mathscr{Y}$ 中每一个强闭凸子集也是弱闭的．关于后面这个命题读者可参看 Hille-Phillips[26] 定理 2.9.3，或者 Dunford-Schwartz[20] 定理 V3.13．

令 $\mu=\text{cl co}\{y_n\}$，其中闭包是按 $\mathscr{Y}$ 的模拓扑取的．则 μ 是 $\mathscr{Y}$ 中的强闭凸子集，因此它是弱闭的，由于 y_0 是弱极限点，故它在 μ 中，所以它可以在模的意义下由 $\text{co}\{y_n\}$ 中的点逼近到任意精确度．

〔注〕 中译本在 189 页．——译者注．

4. 一般存在定理

本节我们叙述和证明第二章问题 2 的一般存在定理，即定理 4.2. 另外两个一般存在定理即定理 8.3 和定理 8.6 将在第 8 节给出. 后两个定理并不包含定理 4.2, 它们也不被定理 4.2 所包含.

虽然定理 4.2 的一些假设在给定的例子中可以直接验证，但这些假设形式上并不方便. 定理 4.2 的作用在于它是一个本源性定理，从它能容易地导出某些存在定理，而这些存在定理的条件对整类问题是容易检验的. 在第 5 节和第 6 节中我们将这样做. 在第 6 节中，我们将从定理 4.2 得到通常变分法问题中的经典的存在定理. 主要兴趣在于应用的读者在继续阅读本节剩下部分之前，可先阅读第 5 节的大部分，到必要时再回过头来查阅下面的假设 4.1.

为了更好地领会以下定理的证明，我们复习一下如下的熟知结果的证明：定义在致密度量空间 $\mathscr{Y}$ 上的实值下半连续函数 f 在 $\mathscr{Y}$ 上达到它的最小值，亦即存在点 $y_0\in\mathscr{Y}$，使得 $f(y_0)\leqslant f(y)$ 对一切 $y\in\mathscr{Y}$ 成立. 如果 μ 表示 f 在 $\mathscr{Y}$ 的下确界，则 $\mu<+\infty$，且存在一个叙列 $\{y_n\}$，使得 $f(y_n)\to\mu$. 如此叙列 $\{y_n\}$ 叫做极小化叙列. 由于 $\mathscr{Y}$ 是致密的，存在一个子叙列（我们仍用 $\{y_n\}$ 表示）$\{y_n\}$ 和点 $y_0\in\mathscr{Y}$，使得 $y_n\to y_0$. 由于 f 的下半连续性，我们得到 $\liminf f(y_n)\geqslant f(y_0)$. 但由 μ 的定义，$f(y_0)\geqslant\mu$，并且 $\liminf f(y_n)=\mu$. 因此 $\mu>-\infty$，且 $f(y_0)=\mu$. 注意，我们此处只用到极小化叙列的条件致密性.

用上述论证作为指南，我们可以按下面步骤着手解决问题 2 的存在性问题. 首先给定容许对 (ϕ, u) 集合 $\mathscr{A}$ 的一个拓扑，使得在此拓扑下 $\mathscr{A}$ 的子集 $\mathscr{A}_1$ 是致密的，并证明 J 在 $\mathscr{A}_1$ 上是下半连续的. 虽然在某些特殊情形中可以这样做，但一般地说，由于状态方程中函数 f 的非线性性，使得有必要修改这一做法. 定理 4.1 以及定理 8.2 和定理 8.4 就是这样做的. 这可以按如下富有成效的方法来处理：令 $\mathscr{A}_T$ 表示所有容许轨线的集合. 我们将 $\mathscr{A}_T$ 视

为 $\mathscr{X}_p$ 中的集合，并以 $\mathscr{A}_{1T}$ 表示相应于 $\mathscr{A}_1$ 中的容许对 (ϕ, u) 的轨线集合．定理 4.1 给出集合 $\mathscr{A}_{1T}$ 在 $\mathscr{A}_T$ 中条件致密的准则．这准则导致以下事实：若 $\{\phi_k\}$ 是 $\mathscr{A}_{1T}$ 的叙列，则有一子列 $\{\phi_{kj}\}$ 和一个容许对 (ϕ, u)，使得在 $\mathscr{X}_p$ 中 $\phi_{kj}\to\phi$．注意，我们要求比 $\mathscr{A}_{1T}$ 在 $\mathscr{X}_p$ 中条件致密性更多的条件．定理 4.1 也提供适合于我们的问题的下半连续性概念的推广．

定理 4.1 是一个基础性的结果．本节的存在性定理以及第四章的广义控制的存在性定理都以它为基础．这个定理也将用以得到第四章的"可达集"的性质．

定理 4.1 以及其它几个存在性定理中将做如下的假设．

假设 4.1　(i) 存在致密集 $\mathscr{R}_0\subseteq\mathscr{R}$，使得对所有容许轨线 ϕ，当所有 $t\in[t_0, t_1]$ 时，点 $(t, \phi(t))\in\mathscr{R}_0$．(ii) 集合 $\mathscr{B}$ 是闭的．(iii) 映射 Ω 在 $\mathscr{R}_0$ 上是上半连续的．(iv) 对 $\mathscr{R}_0$ 中的每一 (t, x)，集合 $\mathscr{Q}^+(t, x)$ 是闭且凸的．(v) 存在常数 $K\geqslant 0$ 使得对所有 $(t, x, z)\in\mathscr{D}$，$f^0(t, x, z)\geqslant -K$；其中

$$\mathscr{D}=\{(t, x, z):(t, x)\in\mathscr{R}_0,\ z\in\Omega(t, x)\}. \tag{4.1}$$

(vi) 函数 f^0 在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上是下半连续的，而 f 在 $\mathscr{G}$ 上是连续的．

若对于所有 $t\in[t_0, t_1]$，$(t, \phi(t))\in\mathscr{R}'$，则称轨线 ϕ 处于 (t, x) 空间的集合 $\mathscr{R}'$ 中．于是假设 4.1-(i) 可改为所有容许轨线处于 $\mathscr{R}_0$ 中．

定理 4.1　若假设 4.1 成立，且映射 $\mathscr{Q}^+$ 在 $\mathscr{R}_0$ 的每一点上满足弱 Cesari 性质．令

$$I(\phi, u)=\int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt, \qquad (\phi, u)\in\mathscr{A}. \tag{4.2}$$

若 $\mathscr{A}_0$ 是容许对 (ϕ, u) 的集合，使得轨线 ϕ 是等度绝对连续的，并且使得 $\sup\{I(\phi, u):(\phi, u)\in\mathscr{A}_0\}<+\infty$．则在 $\mathscr{A}_0$ 中存在叙列 $\{(\phi_k, u_k)\}$，而在 $\mathscr{A}$ 中存在一个容许对 (ϕ^*, u^*)，具有以下性质：

(i) 在 $\mathscr{X}_p$ 中，ϕ_k 收敛于 ϕ^*；

(ii) $\lim_{k\to\infty} I(\phi_k, u_k) \geqslant I(\phi^*, u^*)$. (4.3)

我们着重指出，这里并没有给出关于控制 u_k 与控制 u^* 之间相互关系的论断. 此外，一般地说 (4.3) 式对所有使得 (ϕ^*, u) 为容许对的 u 是不成立的.

习题 4.1 举一个例子证明上述最后一个命题. 这种例子曾由 Cesari 于[18], p. 528 中给出过.

定理 4.1 的一个等价的定理是如下的推论 4.1, 它的侧重面稍有不同. 推论 4.1 中叙述的事实有时概括地说成在 $\mathscr{A}_0$ 中下闭包性质成立.

推论 4.1 设 $\{(\phi_k, u_k)\}$ 为 $\mathscr{A}_0$ 中的叙列，而 ϕ 为 $\mathscr{X}_p$ 中的元素，使得 $\phi_k \to \phi$ 在 $\mathscr{X}_p$ 中成立. 则存在控制 u, 使得 (ϕ, u) 是容许的，并且 $\lim\inf I(\phi_k, u_k) \geqslant I(\phi, u)$.

习题 4.2 证明推论 4.1 和定理 4.1 是等价的.

定理 4.1 的证明将在第 7 节给出. 对于定理中的证明来说，下面插入的材料是不必要的. 所以请读者可以跳过这些材料而直接阅读证明本身.

我们现在叙述并证明 II.3 的问题 2 的存在性定理.

定理 4.2 若假设 4.1 成立，并且映射 $\mathscr{Q}^+$ 在 $\mathscr{R}_0$ 上的每一点满足弱 Cesari 性质. 假定泛函 G 在 $\mathscr{A}_T$ 上是有定义、下半连续和下有界的. 把容许轨线集视为 $\mathscr{X}_p$ 的子集. 令 $\mu = \inf\{\hat{J}(\phi, u): (\phi, u) \in \mathscr{A}\}$. 若存在极小化叙列 $\{(\phi_k, u_k)\}$ 使得轨线 $\{\phi_k\}$ 是等度绝对连续的. 则 μ 是有限的，并且存在 $(\phi^*, u^*) \in \mathscr{A}$, 使得

$$\hat{J}(\phi^*, u^*) = \mu.$$

证明：因为对容许对 (ϕ, u) 而言，函数 $t \to f^0(t, \phi(t), u(t))$ 属于 $L_1[t_0, t_1]$. 由此可见 $\mu < +\infty$. 因为

$$\hat{J}(\phi, u) = G(\phi) + I(\phi, u),$$

而 G 在 $\mathscr{A}_T$ 中是下有界的，又因为 $\mu < +\infty$, 可见叙列 $\{I(\phi_k, u_k)\}$ 是上有界的. 因此，我们可以取极小化叙列（它的存在是本定理的假设）作为定理 4.1 的集合 $\mathscr{A}_0$, 从而得到如下结论：存在一个子

叙列，我们仍用$\{(\phi_k, u_k)\}$表示，以及定义于区间$[t_0^*, t_1^*]$上的$\mathscr{A}$中之元素(ϕ^*, u^*)，使得$\phi_k \to \phi^*$在$\mathscr{X}_p$中成立，并使得

$$\lim I(\phi_k, u_k) \geqslant I(\phi^*, u^*).$$

由G的下半连续性我们得到$\lim\inf G(\phi_k) \geqslant G(\phi^*)$. 因为

$$\begin{aligned}\mu = \lim \hat{J}(\phi_k, u_k) &= \lim(G(\phi_k) + I(\phi_k, u_k)) \\ &\geqslant \lim\inf G(\phi_k) + \lim I(\phi_k, u_k) \\ &\geqslant G(\phi^*) + I(\phi^*, u^*) = \hat{J}(\phi^*, u^*) = \mu.\end{aligned}$$

于是已证明μ是有限的，并且$\hat{J}(\phi^*, u^*) = \mu$.

注意，在定理 4.2 中$\mathscr{A}_1 = \mathscr{A}$. 如果$\mathscr{A}_1 \subset \mathscr{A}$，则为了得到存在定理，必须证明从定理 4.1 得到的容许对(ϕ^*, u^*)具有规定$\mathscr{A}_1$作为$\mathscr{A}$的真子集的性质.

我们可以用一种按本书来说显得过分精致的方法来表达定理 4.1 和推论 4.1，但是这种表达形式却是分布参数系统所必需的. 我们简略地介绍这种表达形式. 对于不熟悉 Sobolev 空间基本常识的读者，可以略去这些材料而无损于对定理 4.1 和推论 4.1 的理解.

我们假定假设 4.1 成立. 则存在致密区间$\mathscr{I}$，使得所有容许轨线定义在$\mathscr{I}$的子区间. 为了简化说明，我们假定所论问题具有固定的初始时刻t_0和固定的终止时刻t_1，且$\mathscr{I} = [t_0, t_1]$. 容许轨线是绝对连续函数ϕ，它的导数属于$L_1[\mathscr{I}]$. 因此每条容许轨线是 Sobolev 空间$H_1^1(\mathscr{I})$的一元素. 这里我们把构成$H_1^1(\mathscr{I})$的一个元素的等价类与它的一个表现等同起来. 因为$\mathscr{I}$是一维区间，$H_1^1(\mathscr{I})$的每一个元素有一个绝对连续的表现. 我们将把$H_1^1(\mathscr{I})$的每一元素与它的绝对连续的表现视为一致.

设$\mathscr{A}_0$是容许对(ϕ, u)的集合，使得在$\mathscr{A}_{0T}$中的轨线族ϕ是等度绝对连续的. 因此轨线族是一致有界的，从 Ascoli 定理和定理 3.1 可知轨线集合$\mathscr{A}_{0T}$是$H_1^1(\mathscr{I})$中的弱致密集. 反之，若轨线族构成$H_1^1(\mathscr{I})$中的弱致密集，则此轨线族是等度绝对连续的. 因此，在定理 4.1 的叙述中“轨线族$\{\phi\}$是等度绝对连续的”这一措辞可以用“轨线族$\{\phi\}$在$H_1^1(\mathscr{I})$中是弱致密的”来代替. 因此

定理 4.1 所给的条件尤其保证，在 $H_1^1(\mathscr{I})$ 中弱致密的轨线集 $\mathscr{A}_{0T}$ 也是在 $\mathscr{A}_T$ 中条件致密的，这里的 $\mathscr{A}_{0T}$ 和 $\mathscr{A}_T$ 被视为 $C[\mathscr{I}]$ 的子集.

5. 在致密约束的情况下的一个存在定理

本节我们将讨论当集合 $\Omega(t, x)$ 为致密时保证最优控制存在的一个定理. 此定理概括了应用中出现的各类重要问题. 它要求 $\Omega(t, x)$ 对 (t, x) 的依赖关系有比上半连续具有更多的正规性.

设 U 是 E^m 中的集合，$\varepsilon>0$. 令 $[U]_\varepsilon$ 表示 U 的闭 ε 邻域；亦即

$$[U]_\varepsilon=\{z:z\in E^m,\ \operatorname{dist}(z, U)\leqslant\varepsilon\},$$

其中 $\operatorname{dist}(z, U)=\inf\{\operatorname{dist}(z, \eta):\eta\in U\}$.

定义 5.1 将 $\mathscr{R}$ 中的点 (t, x) 映为 E^m 的子集 $\Lambda(t, x)$ 的映射 Λ 称为在 $\mathscr{R}$ 中的点 (t_0, x_0) 上关于包含上半连续的或 u. s. c. i，如果对每一个 $\varepsilon>0$，存在 $\delta>0$，使得对 $N_\delta(t_0, x_0)$ 中的每一点 (t, x)，满足

$$\Lambda(t, x)\subseteq[\Lambda(t_0, x_0)]_\varepsilon. \tag{5.1}$$

如果在 $\mathscr{R}$ 中的每一点上映射 Λ 是 u. s. c. i，则称 Λ 在 $\mathscr{R}$ 上为 u. s. c. i.

显然，如果 $\Omega(t, x)=U$，即对所有 $(t, x)\in\mathscr{R}$，$\Omega(t, x)$ 为固定集合 U，则此映射在 $\mathscr{R}$ 上为 u. s. c. i.

如果 Λ 在 (t_0, x_0) 上是 u. s. c. i 的，且 $\Lambda(t_0, x_0)$ 是闭的，则 Λ 在 (t_0, x_0) 处是上半连续的. 但是反过来未必为真. 如果 Λ 在 (t_0, x_0) 处是 u. s. c. i 的，则对每一个 $\varepsilon>0$，存在 $\delta(\varepsilon)>0$，使得若 $(t, x)\in N_{\delta(\varepsilon)}(t_0, x_0)$，则 $\Lambda(t, x)\subseteq[\Lambda(t_0, x_0)]_\varepsilon$. 因此

$$\Lambda(N_{\delta(\varepsilon)}(t_0, x_0))\subseteq[\Lambda(t_0, x_0)]_\varepsilon.$$

而且 $\operatorname{cl}\Lambda(N_{\delta(\varepsilon)}(t_0, x_0))\subseteq[\Lambda(t_0, x_0)]_\varepsilon$ 同样成立，又有

$$\bigcap_{\delta>0}\operatorname{cl}\Lambda(N_\delta(t_0, x_0))\subseteq\bigcap_{\delta(\varepsilon)>0}\operatorname{cl}\Lambda(N_{\delta(\varepsilon)}(t_0, x_0))$$

$$\subseteq\bigcap_{\varepsilon>0}[\Lambda(t_0, x_0)]_\varepsilon.$$

因为 $\Lambda(t_0, x_0)$ 是闭的，最后一个交集等于 $\Lambda(t_0, x_0)$. 因此 Λ 在 (t_0, x_0) 上是上半连续的. 映射是上半连续的但不是 u. s. c. i 的一个例子如下. 设

$$\Lambda(t, x)=\{z\in E^1: 0\leqslant z\leqslant 1\}\cup\left\{z=\frac{1}{t}\right\},\ t\neq 0;$$

$$\Lambda(0, x)=\{z\in E^1: 0\leqslant z\leqslant 1\}.$$

则 Λ 在任意点 $(0, x_0)$ 是上半连续的，但在 $(0, x_0)$ 不是 u. s. c. i 的. 注意，每一个集合 $\Lambda(t, x)$ 是致密的.

定理 5.1 假设容许对集合 $\mathscr{A}$ 非空，且如下假设成立：(i) 存在致密集 $\mathscr{R}_0\subset\mathscr{R}$，使得对所有容许轨线 ϕ，对所有 $t\in[t_0, t_1]$，$(t, \phi(t))\in\mathscr{R}_0$；(ii) 集 $\mathscr{B}$ 是闭的；(iii) 映射 Ω 在 $\mathscr{R}_0$ 上是 u. s. c. i 的；(iv) 对每一个 $(t, x)\in\mathscr{R}_0$，集合 $\Omega(t, x)$ 是致密的；(v) 对每一 $(t, x)\in\mathscr{R}_0$，集合 $\mathscr{Q}^+(t, x)$ 是凸的；(vi) 函数 f^0 在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上是下半连续的，而 f 在 $\mathscr{G}$ 上是连续的. 设 J 和 II(3.2) 中的一样，而 g 在 $\mathscr{B}$ 上为下半连续. 则存在 $(\phi^*, u^*)\in\mathscr{A}$，使得对所有 $(\phi, u)\in\mathscr{A}$，有 $J(\phi^*, u^*)\leqslant J(\phi, u)$.

一旦证明了定理 5.1 的假设能使定理 4.2 的所有假设成立，定理 5.1 便得到证明. 我们将先对定理进行讨论之后，在本节末尾给出这种证明.

附注 5.1 如果用 II(3.3) 中定义的泛函 $\hat{J}$ 代替 J，定理 5.1 仍然成立，其中 G 在 $\mathscr{A}_T$ 上是下半连续且下有界的. 正如 II. 3 节叙述问题 2 时所作的讨论中指出的，这种泛函包含了诸如 $\max|\phi(t)|$ 或 $\max|\phi(t)-h(t)|$ 的泛函.

附注 5.2 在定理 5.1 的证明中，我们将证明假设 (iii) 和 (iv) 隐含着 (4.1) 所定义的集合 $\mathscr{D}$ 是致密的，并说明致密性的真正用途. 因此 (iii) 和 (iv) 可以用 $\mathscr{D}$ 是致密的假设来代替. 但是条件 (iii) 和 (iv) 在应用中比较易于检验，因为它们直接包含问题的数据.

我们现在指出定理 5.1 可以应用于几类重要控制问题. 假定定理 5.1 关于 $\mathscr{R}_0$, Ω, g 和 $\mathscr{B}$ 的假设成立，并附加假设对每一点 $(t, x)\in\mathscr{R}_0$，集合 $\Omega(t, x)$ 是凸的. 我们将考虑当 $\Omega(t, x)$ 是凸集

时，保证 $\mathscr{Q}^+(t, x)$ 是 E^{n+1} 的凸子集的函数 $f^0, f^1, \cdots, f^n$ 的某些特殊类型.

首先，我们考虑函数 f^i 关于 x 和 z 是线性的问题，因此

$$f^i(t, x, z)=\sum_{j=1}^{n} a_j^i(t)x^j+\sum_{j=1}^{m} b_j^i(t)z^j+h_i(t),\ i=1, \cdots, n.$$

于是状态方程组为

$$\frac{dx^i}{dt}=\sum_{j=1}^{n} a_j^i(t)x^j+\sum_{j=1}^{m} b_j^i(t)u^j(t)+h_i(t),\ i=1, \cdots, n.$$

而泛函 J 给定为

$$J(\phi, u)=g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1}\left\{\sum_{j=1}^{n} a_0^j(t)\phi^j(t)+\sum_{j=1}^{m} b_0^j(t)u^j(t)+h_0(t)\right\}dt.$$

采用向量矩阵记号，状态方程变成

$$\frac{dx}{dt}=A(t)x+B(t)u(t)+h(t), \tag{5.2}$$

而代价泛函 J 为

$$J(\phi, u)=g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1}\{\langle a_0(t), \phi(t)\rangle+\langle b_0(t), u(t)\rangle+h_0(t)\}dt,$$

其中 $A(t)$、$B(t)$、$a_0(t)$、$b_0(t)$ 和 $h_0(t)$ 有明显的含义. 设实值函数 h_i, a_i^j 和 b_i^k, $i=1, \cdots, n$; $j=1, \cdots, n$; $k=1, \cdots, m$ 在某一固定区间 $[T_0, T_1]$ 上定义且连续，则集合 $\mathscr{R}$ 是 (t, x) 空间中的带形域 $T_0\leqslant t\leqslant T_1$, $-\infty<x^i<\infty$, $i=1, \cdots, n$. 集合 $\mathscr{U}$ 是整个 E^m 空间.

这里留一个练习，请读者直接证明在线性问题中，如果集合 $\Omega(t, x)$ 是凸的，则集合 $\mathscr{Q}^+(t, x)$ 也是凸的.

附注 5.3 在第六章的定理 6.2 中，我们将证明对于线性系统，为了得到最优对的存在性，在 Ω 仅仅依赖于 t 而与 x 无关的情形，可以略去集合 $\Omega(t, x)$ 为凸的要求.

接下去直到最后一段要讨论的一类问题中的一个重要问题是"具有线性对象的时间最优问题". 此问题的状态方程是 (5.2) 的

形式，它要求用最少的时间使系统从给定的初始时刻 t_0 的初始位置 x_0 引到指定的终止位置 x_1. 第一章第5节的调节器问题就是这种问题的一个例子. 如果 t_1 表示轨线到达 x_1 的时间，则我们的目的是使 t_1-t_0 最小. 从而代价泛函变成 $J(\phi, u)=t_1-t_0$. 因此我们可以把 J 看成是令 $g(t_0, x_0, t_1, x_1)=t_1-t_0$ 而 $f^0\equiv 0$ 或令 $g\equiv 0$ 而 $f^0\equiv 1$ 而得到的.

能应用定理 5.1 的另一类问题是所谓"线性对象和凸积分代价准则"的问题. 在这些问题中状态方程用 (5.2) 表示，而代价泛函为

$$J(\phi, u)=g(t_0, x_0, t_1, x_1)+\int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt,$$

其中对每一 $(t, x)\in\mathscr{R}$, f^0 在 $\Omega(t, x)$ 上是 z 的凸函数. 所谓定义在 $E^m(m\geqslant 2)$ 中的凸集 S 上的实值函数 ψ 是凸的，如果对所有 S 中的 x、y 和所有实数 $\alpha\geqslant 0$, $\beta\geqslant 0$, 使得 $\alpha+\beta=1$,

$$\psi(\alpha x+\beta y)\leqslant\alpha\psi(x)+\beta\psi(y).$$

带凸积分代价准则的一类线性问题中的一个重要问题是关于线性系统使燃料最省的问题. 在此问题中，要求线性系统从给定的初始状态 x_0 转移到指定的终止状态集合中的任意状态 x_1, 并要求在整个转移过程中燃料消耗最少. 终止时刻可以是指定的或自由的. 控制 u 需满足约束 $|u^i(t)|\leqslant 1$, $i=1, \cdots, m$, 我们以 $\beta(t)$ 表示在时刻 t 的燃料流速度，并假设它与控制向量的绝对值成正比，即

$$\beta(t)=\sum_{i=1}^{m} c^i|u^i(t)|,\quad c^i>0 \text{ 为常数}.$$

因此，系统从 x_0 转移到 x_1 的燃料总消耗量为

$$J(\phi, u)=\int_{t_0}^{t_1}\left(\sum_{i=1}^{m} c^i|u^i(t)|\right)dt,$$

J 即为要求极小化的代价泛函，其中

$$f^0(t, x, z)=\sum_{i=1}^{m} c^i|z^i|,$$

并且 f^0 关于 z 是凸的.

具有凸积分代价准则线性问题中的另一个重要问题是"二次型准则"问题，其提法如下：在固定区间$[t_0, t_1]$上指定一个绝对连续函数ξ，它通常表示系统所期望的轨线。要求选取一个容许控制u使得相应的轨线ϕ与所给定的轨线ξ在$[t_0, t_1]$上的均方误差为最小，并以能量消耗最小完成此目的．若以积分$\int_{t_0}^{t_1}|u|^2dt$作为能量消耗的度量，则导致考虑如下的代价泛函

$$J(\phi, u)=|\phi(t_1)-\xi(t_1)|^2+\int_{t_0}^{t_1}|\phi(t)-\xi(t)|^2 dt+\int_{t_0}^{t_1}|u(t)|^2 dt.$$

若令$\overline{\phi}(t)=\phi(t)-\xi(t)$，那么由于$\phi$是(5.2)式的解，$\overline{\phi}$也将是形如(5.2)的线性系统的解．因此我们可以假定泛函J具有如下形式：

$$J(\phi, u)=|\phi(t_1)|^2+\int_{t_0}^{t_1}|\phi(t)|^2 dt+\int_{t_0}^{t_1}|u(t)|^2 dt.$$

如果对轨线的坐标和u的支量赋予非负的权因子，则代价泛函变成

$$J(\phi, u)=\langle\phi(t_1), R\phi(t_1)\rangle+\int_{t_0}^{t_1}\langle\phi(t), X(t)\phi(t)\rangle dt+\int_{t_0}^{t_1}\langle u(t), U(t)u(t)\rangle dt, \tag{5.3}$$

其中X和U是具有非负对角元素的连续的对角矩阵．而R是具有非负对角元素的常数对角矩阵．

更一般地，我们可以取X和U为$[t_0, t_1]$上连续的半正定对称矩阵．以后当我们考虑非致密约束集时，矩阵U将需要是正定的．一般说，不必假设U是对角形的，但这不是本质的，下面的讨论表明了这一点．存在实正交矩阵P使得$U=P'DP$，其中D是对角矩阵，而"$'$"表示转置．在变量替换$V=Pu$之下，二次型$\langle u, Uu\rangle$变成$\langle v, Dv\rangle$，D为对角矩阵．状态方程(5.2)变成

$$\frac{dx}{dt}=A(t)x+C(t)v(t)+h(t),$$

其中$C(t)=B(t)P^{-1}(t)$. 如果X是常矩阵，则存在变量替换$y=Qx$，其中Q为正交常矩阵，使得二次型$\langle x, Xx\rangle$变为$\langle y, Yy\rangle$，Y为对角矩阵．而状态方程变成关于y和v的线性方程．

线性问题以及具有凸积分代价准则的线性问题是下面问题的特殊情形，在这种问题中，最优控制和最优轨线的存在性是定理5.1的推论．

推论 5.1 假如除(V)和(vi)之外，定理5.1的所有假设都成立；还假设$\Omega(t, x)$有进一步的性质，即对每一$(t, x)\in\mathscr{R}_0$，$\Omega(t, x)$是E^m中的凸集．假设h是从$\mathscr{R}$到E^n的连续函数，B是$\mathscr{R}$上的$n\times m$连续矩阵，f^0是$\mathscr{G}$上的实值下半连续函数，使得对每一$(t, x)\in\mathscr{R}_0$，f^0在$\Omega(t, x)$上是z的凸函数．设状态方程为

$$\frac{dx}{dt}=h(t, x)+B(t, x)u(t),$$

而代价泛函为

$$J(\phi, u)=g(t_0, \phi(t_0), t_1, \phi(t_1))+\int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt.$$

则对每一$(t, x)\in\mathscr{R}_0$，集合$\mathscr{Q}^+(t, x)$是凸的，而且泛函J在$\mathscr{A}$中达到它的最小值．

附注 5.4 在状态方程关于x是线性的一类重要问题中，假如初始点处于致密集中，那么由此初始点出发的轨线总是处于致密集中．在这些问题中定理5.1的假设(i)总是满足的．这将在下面的习题5.1和5.2中加以探讨．

为了证明推论5.1，我们只须证明对每一$(t, x)\in\mathscr{R}_0$，集合$\mathscr{Q}^+(t, x)$是凸的．从而定理5.1的所有假设都成立．((vi)的正确性可以从f^0、h和B的连续性推出．)

令$\hat{y}=(y^0, y)$[注]，在本问题中

$$\begin{aligned}\mathscr{Q}^+(t, x)=\{\hat{y}=(y^0, y): y^0\geqslant f^0(t, x, z),\\ y=h(t, x)+B(t, x)z,\ z\in\Omega(t, x)\}.\end{aligned}$$

[注] 原文误为$y=(y^0, y)$.——译者注.

假设 $\hat{y}_1$ 和 $\hat{y}_2$ 是 $\mathscr{Q}^+(t, x)$ 中的任意二点，则存在 $\Omega(t, x)$ 中的点 z_1 和 z_2 使得

$$y_1^0 \geqslant f^0(t, x, z_1),\ y_1 = h(t, x) + B(t, x)z_1,$$
$$y_2^0 \geqslant f^0(t, x, z_2),\ y_2 = h(t, x) + B(t, x)z_2.$$

设 $\alpha \geqslant 0$, $\beta \geqslant 0$, 且 $\alpha + \beta = 1$, 并令 $\hat{y}_3 = \alpha\hat{y}_1 + \beta\hat{y}_2$, 为了证明 $\mathscr{Q}^+(t, x)$ 是凸的，我们必须证明 $\hat{y}_3 \in \mathscr{Q}^+(t, x)$. 因为我们有

$$\begin{aligned} y_3 = \alpha y_1 + \beta y_2 &= h(t, x) + \alpha[B(t, x)z_1] + \beta[B(t, x)z_2] \\ &= h(t, x) + B(t, x)(\alpha z_1 + \beta z_2), \end{aligned}$$

由于 $\Omega(t, x)$ 是凸的，存在 $z_3 \in \Omega(t, x)$, 使得 $z_3 = \alpha z_1 + \beta z_2$, 从而

$$y_3 = h(t, x) + B(t, x)z_3,\ z_3 \in \Omega(t, x).$$

又由于 f^0 的凸性和 z_3 的定义，我们有

$$\begin{aligned} y_3^0 = \alpha y_1^0 + \beta y_2^0 &\geqslant \alpha f^0(t, x, z_1) + \beta f^0(t, x, z_2) \\ &\geqslant f^0(t, x, \alpha z_1 + \beta z_2) = f^0(t, x, z_3). \end{aligned}$$

所以 $\hat{y}_3 \in \mathscr{Q}^+(t, x)$.

以下我们讨论各种假设的关系.

例 2.2 说明对于假设 $\Omega(t, x)$ 是致密集是必要的. 在此例中映射 Ω 是 u. s. c. i 的，因为对所有的 (t, x), $\Omega(t, x) = E^1$, 集 $\mathscr{Q}^+(t, x)$ 是凸的. 此外如果我们考虑 $\mathscr{A}$ 的任一子类 $\mathscr{A}_0$, $\mathscr{A}_0$ 是由要求轨线处于下面形式的致密集 $\mathscr{R}_0$ 中所确定的：

$$\mathscr{R}_0 = \{(t, x): 0 \leqslant t \leqslant 1,\ a \leqslant x \leqslant b\},\qquad a \leqslant 0,\ b \geqslant 1.$$

则例 2.2 的论证表明在 $\mathscr{A}_0$ 中最小值不存在. 于是，因为 $\mathscr{B}$ 是单点组成和 $g \equiv 0$, 可见除了对所有 $(t, x) \in \mathscr{R}_0$, $\Omega(t, x)$ 为致密集的假设之外，定理 5.1 的所有假设都满足.

例 2.2(a) 表明如果集合 $\Omega(t, x)$ 的致密性在 (t, x) 空间的一个点上不满足，则定理 5.1 的结论可以不成立. 例 2.2(b) 表明即使在一点上，$\Omega(t, x)$ 不具有 u. s. c. i, 则定理 5.1 可以不成立. 还要指出，在例 2.2(a) 和 2.2(b) 中，(4.1) 式所定义的集合 $\mathscr{D}$ 是不致密的. 在例 2.2(b) 中 $\mathscr{D}$ 甚至不是闭的. 注意在形如 $(0, x)$ 的点上 Ω 不是上半连续的.

在例 2.3 中，我们再一次看到 $\mathscr{R}_0$ 的致密性仍然不能得到最

优控制的存在性．这里除了 $\mathscr{Q}^+(t, x)$ 是凸的要求之外所有假设都满足．

例 2.4 说明 $\mathscr{R}_0$ 的致密性是必要的．如果我们取 $\mathscr{R}_0$ 为 $\mathscr{F}$ 所覆盖的区域的闭包，则 $\mathscr{R}_0$ 不是致密的．因为在本例中 $f^0\equiv 0$，从(2.3)以及 Ω 的定义显然可看出对每一 (t, x)，集 $\mathscr{Q}^+(t, x)$ 是凸的．易见，对所有 (t, x)，$\Omega(t, x)$ 是致密的，且为 u. s. c. i．因为 $\mathscr{T}_0$ 和 $\mathscr{T}_1$ 是致密的，故集 $\mathscr{B}$ 是致密的．因为 $g(t_0, x_0, t_1, x_1)=x_1$，函数 g 是连续的．因此除了 $\mathscr{R}_0$ 的致密性，定理 5.1 的所有条件都满足．

下面我们讨论保证存在致密集 $\mathscr{R}_0$ 的条件，使得如假设 (i) 所要求的那样，所有容许轨线的图形都能处于 $\mathscr{R}_0$ 中．这些条件并不包含在定理的叙述中，因为它们是一种过分的限制．事实上，即使这些条件不满足，轨线也能够位于致密集中．此外，在特殊的问题中，经常可以直接证实轨线处于一致密集中．

引理 5.1 设 $\mathscr{R}$ 是包含在带形域 $a\leqslant t\leqslant b$，$-\infty<x^i<\infty(i=1, \cdots, n)$ 之中的区域．$\Delta=\{(t, x, z):(t, x)\in\mathscr{R}, z\in\Omega(t, x)\}$，函数 $f=(f^1, f^2, \cdots, f^n)$ 对一切 $(t, x, z)\in\mathscr{D}$，满足

$$|\langle x, f(t, x, z)\rangle|\leqslant K(t)(|x|^2+1), \qquad (5.4)$$

其中 $K(t)$ 是在 $[a, b]$ 上可积的函数．假设每一容许轨线至少包含属于 $\mathscr{R}$ 中的给定致密集 C 的一点 $(t_2, \phi(t_2))$．则存在致密集 $\mathscr{R}_0\subset\mathscr{R}$，使得 $\mathscr{A}$ 中的所有轨线位于 $\mathscr{R}_0$ 中．如果我们要求 $\mathscr{A}$ 中的所有轨线的初始点位于 C 中，那么我们可以略去 (5.4) 式左边的绝对值符号．

证明：对任意轨线 ϕ，令

$$\Phi(t)=|\phi(t)|^2+1=\langle\phi(t), \phi(t)\rangle+1.$$

则 $\Phi'(t)=2\langle\phi(t), f(t, \phi(t), u(t))\rangle$，而依照(5.4)

$$|\Phi'(t)|\leqslant 2K(t)(|\phi(t)|^2+1)=2K(t)\Phi(t).$$

因此

$$-2K(t)\Phi(t)\leqslant\Phi'(t)\leqslant 2K(t)\Phi(t). \qquad (5.5)$$

如果 $(t_2, \phi(t_2))$ 是轨线上属于 C 之点，则积分(5.5)式就得

$$\Phi(t)\leqslant\Phi(t_2)\exp\left(2\left|\int_{t_2}^{t}K(s)ds\right|\right)\leqslant\Phi(t_2)\exp\left(2\int_{a}^{b}K(s)ds\right)$$

对轨线上所有的点成立．因为 C 是致密的，存在常数 D 使得若 $(t, x)\in C$，则 $|x|\leqslant D$．因此

$$\Phi(t)\leqslant(D^2+1)\exp\left(2\int_{a}^{b}K(s)ds\right).$$

由于此不等式右边是一个与轨线 ϕ 无关的常数，此即表示所有轨线处于某一致密集 $\mathscr{R}_0$ 中．

如果所有初始点 (t_0, x_0) 处于致密集中，我们只须利用 (5.5) 最右边的不等式即可得到与 ϕ 无关的 $\Phi(t)$ 的界限．因此，在这种情形中我们可以略去(5.4)左边的绝对值符号．

在某些问题中，可以证明存在致密集 $\mathscr{R}_0$ 使得那些不处于 $\mathscr{R}_0$ 中的轨线给出的 J 或 $\hat{J}$ 的值比那些处于 $\mathscr{R}_0$ 中的轨线给出的更大．在这种情况下，人们可以不管那些不处于 $\mathscr{R}_0$ 中的轨线．我们简单地重新规定 $\mathscr{R}$ 为 $\mathscr{R}_0$．并重新定义容许对的集合为那些使得相应的容许轨线位于 $\mathscr{R}_0=\mathscr{R}$ 之中所有容许对．这样的一个例子将在第 6 节结合具有非致密约束的线性二次型问题在习题 6.5 中给出．

现在证明定理 5.1．我们以如下引理开始．

引理 5.2 在定理 5.1 的假设下，(4.1) 式定义的集合 $\mathscr{D}$ 是致密的，并且对每一 $(t, x)\in\mathscr{R}_0$，集合 $\mathscr{Q}^+(t, x)$ 是闭且凸的．

在上面定理 5.1 的叙述的那一段中，我们曾指出定理 5.1 的假设(iii)和(iv)隐含着 Ω 在 $\mathscr{R}_0$ 上是上半连续的，因此由引理 3.1 得 $\mathscr{D}$ 是闭的．因此为了证明 $\mathscr{D}$ 是致密的，我们必须证明它是有界的．如果 $\mathscr{D}$ 是无界的，则必存在叙列 $\{(t_n, x_n, z_n)\}$ 和 $\mathscr{R}_0$ 中的点 (t_0, x_0)，使得 $(t_n, x_n)\to(t_0, x_0)$ 和 $|z_n|\to\infty$．由于 Ω 为 u.s.c.i，故存在 n_0，使得对 $n>n_0$，$\Omega(t_n, x_n)\subseteq[\Omega(t_0, x_0)]_1$．因为 $\Omega(t_0, x_0)$ 是致密的，$[\Omega(t_0, x_0)]_1$ 亦然．由于 $z_n\in\Omega(t_n, x_n)$，可见 $|z_n|\to\infty$ 是不可能的．

由定理 5.1 的假设 (v)，每一个集 $\mathscr{Q}^+(t, x)$ 是凸的．为了证

明每一个 $\mathscr{Q}^+(t, x)$ 是闭的，假设 (t, x) 固定，而令 $\{\eta_n\}=\{y_n^0, y_n\}$ 为 $\mathscr{Q}^+(t, x)$ 的点列，它收敛于点 $\eta_0=(y_0^0, y_0)$. 按 $\mathscr{Q}^+(t, x)$ 的定义，存在点列 $z_n\in\Omega(t, x)$，使得 $y_n=f(t, x, z_n)$，且 $y_n^0\geqslant f^0(t, x, z_n)$. 因为 $\Omega(t, x)$ 是致密的，故有子叙列（我们仍记为 $\{z_n\}$）以及 $\Omega(t, x)$ 中的点 z_0，使得 $z_n\to z_0$. 由 f^0 的下半连续性和 f 的连续性可得

$$y_0^0=\lim y_n^0\geqslant\liminf f^0(t, x, z_n)\geqslant f^0(t, x, z_0),$$

$$y_0=\lim y_n=\lim f(t, x, z_n)=f(t, x, z_0).$$

因此 $(y_0^0, y_0)\in\mathscr{Q}^+(t, x)$，引理得证.

我们现在证明定理 4.2 的假设 4.1 成立. 定理 5.1 的假设 (i)、(ii) 和 (vi) 与假设 4.1 的 (i)、(ii) 和 (vi) 相同. 在引理 5.2 的证明过程中，我们已证明 Ω 在 $\mathscr{R}_0$ 上是上半连续的，所以假设 4.1 的 (iii) 成立. 因为 f^0 是下半连续的且 $\mathscr{D}$ 是致密的，所以 f^0 在 $\mathscr{D}$ 中下有界，因此假设 4.1 的 (v) 成立. 引理 5.2 的最后一个结论即为假设 4.1 的 (iv). 因此假设 4.1 都成立.

往下证明在定理 5.1 的假设下，映射 $\mathscr{Q}^+$ 在 $\mathscr{R}_0$ 中的每一点 (t, x) 上具有 Cesari 性质，自然这就隐含着弱 Cesari 性质成立. 在论证过程中我们必须用到 Carathéodory 关于凸集的一个定理. 因为这个定理在最优控制理论的某些重要场合要用到，为了便于以后参考，我们在此加以叙述. 至于该定理的证明，我们推荐读者查阅任何有关凸集理论的教科书，例如可参看 Eggleston [21].

定理 5.2 设 $\mathscr{X}$ 是 E^n 的子集，y 是 $\mathscr{X}$ 的凸壳 $\mathrm{co}\,\mathscr{X}$ 中的点. 则 y 可写成 $\mathscr{X}$ 中至多 $n+1$ 个点的凸组合. 此外，若 $\mathscr{X}$ 至多有 n 个分支，则 y 可写成 $\mathscr{X}$ 中至多 n 个点的凸组合.

设 (t, x) 固定，$y=(y^1, \cdots, y^n)$，且假设

$$\hat{y}=(y^0, y)\in\bigcap_{\delta>0}\mathrm{cl}\,\mathrm{co}\,\mathscr{Q}^+(N_\delta(t, x)).$$

则存在正实数列 $\{\delta_k\}$ 以及 E^{n+1} 中的点列 $\{\hat{y}_k\}=\{(y_k^0, y_k)\}$，使得 $\delta_k\to 0$，且

$$\hat{y}_k=(y_k^0, y_k)\in\mathrm{co}\,\mathscr{Q}^+(N_{\delta_k}(t, x)),\quad \hat{y}_k=(y_k^0, y_k)\to(y^0, y)=\hat{y}.$$

于是对每一个整数 k，存在整数 $j=j(k)$，实数组 $\alpha_{k1}, \cdots, \alpha_{kj}$，满

足

$$\alpha_{ki}\geqslant 0,\ \sum_{i=1}^{j}\alpha_{ki}=1, \tag{5.6}$$

$\mathscr{R}_0$ 中的点组 $(t_{k1}, x_{k1}), \cdots, (t_{kj}, x_{kj})$ 以及点 $\hat{y}_{k1}, \cdots, \hat{y}_{kj}$，使得

$$|(t_{ki}, x_{ki})-(t, x)|<\delta_k, \tag{5.7}$$

$$\hat{y}_{ki}\in\mathscr{Q}^+(t_{ki}, x_{ki}),\ i=1, \cdots, j;$$

$$\hat{y}_k=\sum_{i=1}^{j}\alpha_{ki}\hat{y}_{ki}. \tag{5.8}$$

从(5.7)中的第二关系式可得，存在点组

$$z_{k1}, \cdots, z_{kj},\ z_{ki}\in\Omega(t_{ki}, x_{ki}),$$

使得

$$y_{ki}^0\geqslant f^0(t_{ki}, x_{ki}, z_{ki}), \tag{5.9}$$

$$y_{ki}=f(t_{ki}, x_{ki}, z_{ki}),\ i=1, \cdots, j.$$

集 $\mathrm{co}\,\mathscr{Q}^+(N_{\delta_k}(t, x))$ 是 E^{n+1} 中的集合．因此，由定理 5.2，$\mathrm{co}\,\mathscr{Q}^+(N_{\delta_k}(t, x))$ 中每个点可表示成 $\mathscr{Q}^+(N_{\delta_k}(t, x))$ 中 $n+2$ 个点的凸组合．所以，对每一整数 k，我们假定 $j=j(k)=n+2$.

从(5.6)可见叙列 $\{(\alpha_{k1}, \cdots, \alpha_{k,n+2})\}_{k=1}^{\infty}$ 在 E^{n+2} 中是有界的．于是存在子叙列(我们仍记为 $\{(\alpha_{k1}, \cdots, \alpha_{k,n+2})\}$) 和点 $(\alpha_1, \cdots, \alpha_{n+2})$ 使得

$$(\alpha_{k1}, \cdots, \alpha_{k,n+2})\to(\alpha_1, \cdots, \alpha_{n+2}), \tag{5.10}$$

而且 $\alpha_i\geqslant 0$，$\sum_{i=1}^{n+2}\alpha_i=1$．对每一个 $i=1, \cdots, n+2$，令

$$\{(t_{ki}, x_{ki}, z_{ki})\}_{k=1}^{\infty}$$

表示相应于收敛叙列 $\{(\alpha_{k,1}, \cdots, \alpha_{k,n+2})\}$ 的子叙列．因为 $\mathscr{D}$ 是致密的，故存在子叙列和 E^m 中的点 $z_1, \cdots, z_{n+2}$，使得对所有 $i=1, \cdots, n+2$

$$(t_{ki}, x_{ki}, z_{ki})\to(t, x, z_i), \tag{5.11}$$

其中 $(t, x, z_i)\in\mathscr{D}$；亦即 $z_i\in\Omega(t, x)$.

从(5.8)和(5.9)我们有

$$y_k^0\geqslant\sum_{i=1}^{n+2}\alpha_{ki}f^0(t_{ki}, x_{ki}, z_{ki}),$$

$$y_k=\sum_{i=1}^{n+2}\alpha_{ki}f(t_{ki},\ x_{ki},\ z_{ki}).$$

从 $\hat{y}_k\to\hat{y}$, (5.10), (5.11), f 的连续性以及 f^0 的下半连续性, 我们有

$$y=\sum_{i=1}^{n+2}\alpha_i f(t,\ x,\ z_i)$$

和

$$\begin{aligned}y^0=\lim y_k^0&\geqslant\lim\left(\sum_{i=1}^{n+2}\alpha_{ki}f^0(t_{ki},\ x_{ki},\ z_{ki})\right)\\&\geqslant\sum_{i=1}^{n+2}\liminf\alpha_{ki}f^0(t_{ki},\ x_{ki},\ z_{ki})\\&\geqslant\sum_{i=1}^{n+2}\alpha_i f^0(t,\ x,\ z_i).\end{aligned}$$

因此 $\hat{y}=(y^0,\ y)\in\mathrm{co}\,\mathscr{Q}^+(t,\ x)$, 但 $\mathscr{Q}^+(t,\ x)$是凸的, 所以

$$\mathrm{co}\,\mathscr{Q}^+(t,\ x)=\mathscr{Q}^+(t,\ x).$$

因此 $\hat{y}\in\mathscr{Q}^+(t,\ x)$, 且 Cesari 性质成立.

因为所有轨线位于致密集 $\mathscr{R}_0$ 中, 且由于 $\mathscr{B}$ 是闭的, 所有容许轨线的端点集合 $\{(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))\}$ 将含于 $\mathscr{B}$ 的致密子集中. 于是我们可以假定 $\mathscr{B}$ 本身是致密的.

因为 g 在 $\mathscr{B}$ 上是下半连续的, 而 $\mathscr{B}$ 是致密的, 故 g 在 $\mathscr{B}$ 上是下有界的. 在这种情形中, 由于 $G(\phi)=g(t_0,\phi(t_0),\ t_1,\ \phi(t_1))$, 因此, 映射 G 在 $\mathscr{A}_T$ 上是下半连续的且下有界的.

最后, 我们证明 $\mathscr{A}_T$ 中的容许轨线等度绝对连续以完成定理 5.1 的证明. 这表明任何极小化序列中的轨线是等度绝对连续的, 从而定理 4.2 的所有假设都得到证实.

由于 $\mathscr{D}$ 是致密的, 而 f 是连续的, 故存在常数 $C>0$, 使得对所有 $(t,\ x,\ z)\in\mathscr{D}$, $|f(t,\ x,\ z)|\leqslant C$. 假设给定 $\varepsilon>0$, 则对任何定义于 $[t_0,\ t_1]$ 上的 $\mathscr{A}_T$ 中的 ϕ 以及任何具有 $\mathrm{meas}(E)<\varepsilon/C$ 的可测集 $E\subset[t_0,\ t_1]$, 有

$$\left|\int_E\phi'(t)dt\right|\leqslant\int_E|\phi'(t)|dt=\int_E|f(t,\ \phi(t);\ u(t))|dt\leqslant\varepsilon.$$

所以函数 $\phi'(t)$ 有等度绝对连续积分. 如定义 3.4 后面所指出的,

这表明 $\mathscr{A}_T$ 中的函数 ϕ 是等度绝对连续的.

习题 5.1 考虑状态方程如 (5.2) 的系统，其中 A、B 和 h 是区间 $[a, b]$ 上的可积函数. 假定映射 Ω 只依赖于 t，且 Ω 定义于 $[a, b]$ 上. 因此对一切 $t\in[a, b]$，$\Omega: t\to\Omega(t)$. 此外，假设 Ω 在 $[a, b]$ 上是 u. s. c. i 的，而且每一个 $\Omega(t)$ 是致密的. 证明 (5.4) 式成立，因此假如每一轨线至少有一点属于已给致密集 C，则所有的容许轨线皆位于致密集 $\mathscr{R}_0$ 中. 请不用引理 5.1 而直接证明最后这个命题. (提示：用参数变易公式.) 证明任何定义在 $[a, b]$ 上且 $u(t)\in\Omega(t)$ 几乎处处成立的可测函数 $u(t)$ 是容许控制.

习题 5.2 考虑状态方程为

$$\frac{dx}{dt}=A(t)x+B(t, u(t))$$

的系统，其中 A 在 $[a, b]$ 上可积，而 B 在 $[a, b]\times E^m$ 上连续. 其他条件如习题 5.1. 对此情形，证明习题 5.1 中所要求的结论.

习题 5.3 如果用下列假设代替 (5.4)：存在在 $\mathscr{R}$ 上有定义且为 $C^{(1)}$ 类的正函数 v 和正常数 K，使得 (i)

$$|\langle v_x(t, x), f(t, x, z)\rangle+v_t(t, x)|\leqslant Kv(t, x)$$

对所有 $(t, x, z)\in\mathscr{D}$ 成立；(ii) 对每一个 α，集合 $\{(t, x): v(t, x)\leqslant\alpha;\ (t, x)\in\mathscr{R}\}$ 是致密的. 证明引理 5.1 成立.

习题 5.4 证明定理 5.1 中的条件“$\mathscr{Q}^+(t, x)$ 是凸的”不能用如下假设所代替：集合 $\mathscr{Q}(t, x)$ 是凸的并且对每一固定的 $(t, x)\in\mathscr{R}$，f^0 是 z 的凸函数. (Cesari[17], p. 399～400.)

6. 非致密约束

本节考虑约束集 $\Omega(t, x)$ 不必是致密的问题. 定理 6.1 及其推论的主要作用是阐述保证极小化序列中的轨线的等度绝对连续的条件. 定理 6.2 是关于变分法中通常问题的经典的 Nagumo-Tonelli 存在定理. 在本节末了的习题里，我们将讨论具有“线性对象、凸积分代价准则和无界控制”这一类重要的特殊问题.

注意：若假设 4.1 成立，则存在致密区间 $\mathscr{I}=[a, b]$，使得任

何容许轨线的定义区间 $[t_0, t_1]$ 包含在 $\mathscr{I}$ 中.

定理 6.1 若假设 4.1 成立，且映射 $\mathscr{Q}^+(t, x)$ 在 $\mathscr{R}_0$ 的每一点上满足弱 Cesari 性质；设 G 在 $\mathscr{A}_T$ 上是下半连续且下有界的；$\{(\phi_k, u_k)\}$ 为极小化叙列，对每一个 $i=1, 2, \cdots, n$，假设存在定义于 $\mathscr{D}$ 上的非负函数 H^i 和常数 A^i，使得对极小化叙列中所有 (ϕ_k, u_k)，有

$$\int_{t_{0k}}^{t_{1k}} H^i(t, \phi_k(t), u_k(t))dt \leqslant A^i, \tag{6.1}$$

其中 $[t_{0k}, t_{1k}]$ 是 (ϕ_k, u_k) 的定义区间；对每一个 $i=1, \cdots, n$ 和每一个 $\varepsilon>0$，假设存在非负函数 $F^i_\varepsilon \in L_1[\mathscr{I}]$，使得对极小化叙列中所有 (ϕ_k, u_k)

$$|f^i(t, \phi_k(t), u_k(t))| \leqslant F^i_\varepsilon(t) + \varepsilon H^i(t, \phi_k(t), u_k(t)) \tag{6.2}$$

几乎处处成立. 则存在 $(\phi^*, u^*) \in \mathscr{A}$，使得 $\hat{J}(\phi^*, u^*) \leqslant \hat{J}(\phi, u)$.

定理 4.2 的所有假设除去函数族 $\{\phi_k\}$ 是等度绝对连续外，或者是定理 6.1 假设的重述，或者立刻可以由定理 6.1 的假设推出的结论. 在证明 (6.1) 和 (6.2) 或下面的 (6.3) 蕴含所需的等度绝对连续性之前，我们指出下面几点.

附注 6.1 因为 G 是下有界的，故叙列 $\{I(\phi_k, u_k)\}$ 是上有界的. 因此若取 $H^i = f^0 + K$，其中 K 如假设 4.1 的 (v)，则 (6.1) 成立.

附注 6.2 若 (6.1) 和 (6.2) 对 $H^i = f^0 + K$ 和 $A^i = A$ 成立，则可证明最优对满足 (6.1) 而相应的 $\phi^{*\prime}$ 的分量满足 (6.2). 这时 H^i 的自变量为 $(t, \phi^*(t), u^*(t))$，这一事实将在第 7 节里作为定理 4.1 的证明的一个推论而得到.

我们现在证明 (6.1) 和 (6.2) 可推出 $\{\phi_k\}$ 的等度绝对连续性. 设给定 $\eta>0$，取 (6.2) 中的 ε 等于 $\eta/2A^i$.

存在 $\delta>0$，使得如果 E 是 $\mathscr{I}$ 中的可测集且 $\text{meas}(E)<\delta$，则 $\int_E F^i_\varepsilon dt < \eta/2$. 数 δ 依赖于 F^i_ε 和 η，但由于 $\varepsilon = \eta/2A^i$，数 δ 最终仅依赖于 η. 如果我们令 $f^i(t, \phi_k(t), u_k(t)) = 0$，当 $t \notin [t_{0k}, t_{1k}]$

时．那么从(6.1)和(6.2)以及 ε 的取法，我们有

$$\int_E |f^i(t,\ \phi_k(t),\ u_k(t))|dt$$

$$\leqslant \int_E F_k^i dt+\varepsilon\int_E H^i(t,\ \phi_k(t),\ u_k(t))dt$$

$$\leqslant \eta/2+(\eta/2A^i)A^i=\eta$$

对一切 $(\phi_k,\ u_k)$ 和 $\mathscr{I}$ 中具有 $\mathrm{meas}(E)<\delta$ 的可测集 E 成立．因为 $\phi_k'(t)=f(t,\ \phi_k(t),\ u_k(t))$，则得函数 ϕ_k 的等度绝对连续性．

推论 6.1 设 Φ 是定义于 $[0,\ \infty)$ 上的正的非减函数，使得当 $\xi\to\infty$，$\Phi(\xi)\to\infty$． 令 $f_k^i=f^i(t,\ \phi_k(t),\ u_k(t))$，并假定存在常数 C^i，使得

$$\int_{t_{0k}}^{t_{1k}} |f_k^i(t)|\Phi(|f_k^i(t)|)dt\leqslant C^i, \tag{6.3}$$

则函数 ϕ_k^i 是等度绝对连续的．因此条件 (6.1) 和 (6.2) 可以用 (6.3)对任何 $i=1,\ \cdots,\ n$ 成立来代替．

我们先证明复合函数 $\Phi(|f_k^i|)$ 是可测的．设 α 是给定的．因为 Φ 是不减的．$\{\xi:\Phi(\xi)\leqslant\alpha\}$是形如$[0,\ \xi_0)$或$[0,\ \xi_0]$的集合．因为 $|f_k^i|$ 是可测的，$\{t:|f_k^i(t)|\leqslant\xi_0\}$ 或 $\{t:|f_k^i(t)|<\xi_0\}$ 是可测的，因此 $\{t:\Phi(|f_k^i(t)|)\leqslant\alpha\}$ 是可测的．

推论 6.1 可以论证从(6.3)能推出 (6.1) 和 (6.2) 成立而得到．设 $H^i(t,\ x,\ z)=|f^i(t,\ x,\ z)|\Phi(|f^i(t,\ x,\ z)|)$，$\varepsilon>0$ 是已给定的，则存在 $M_\varepsilon>0$，使得若 $\xi\geqslant M_\varepsilon$，则 $\Phi(\xi)>\dfrac{1}{\varepsilon}$． 对每一对 $(\phi_k,\ u_k)$，令

$$E_{1k}=\{t:|f^i(t,\ \phi_k(t),\ u_k(t))|\leqslant M_\varepsilon\},$$

$$E_{2k}=\{t:|f^i(t,\ \phi_k(t),\ u_k(t))|>M_\varepsilon\}.$$

对 $t\in E_{2k}$

$$|f_k^i(t)|=\frac{|f_k^i(t)|\Phi(|f_k^i(t)|)}{\Phi(|f_k^i(t)|)}\leqslant\frac{|f_k^i(t)|\Phi(|f_k^i(t)|)}{\Phi(M_\varepsilon)}$$

$$<\varepsilon H^i(t,\ \phi_k(t),\ u_k(t)).$$

因为 $[t_{0k},\ t_{1k}]=E_{1k}\cup E_{2k}\cup Z$，其中 Z 是测度为零的集合．于是

$$|f_k^i(t)|\leqslant M_\varepsilon+\varepsilon H^i(t,\ \phi_k(t),\ u_k(t))$$

几乎处处成立，这就是(6.2)．关系式(6.1)可以从(6.3)和 H^i 的定义推出．

推论 6.2　假设存在常数 $C^i>0$ 和 $p_i>1$，使得对所有极小化叙列的元素，满足

$$\int_{t_{0k}}^{t_{1k}}|\phi_k^{i\prime}(t)|^{p_i}\,dt\leqslant C^i.$$

则函数 ϕ_k^i 是等度绝对连续的．

在推论 6.1 中取 $\phi(\xi)=\xi^{p_i-1}$，可得推论 6.2．它也是 Hölder 不等式的一个直接推论．

在 II. 6 中，我们曾叙述过变分法的简单问题，并证明它为什么可以写成控制问题．现在我们进一步假设 II. 6 中的区域 $\mathscr{G}$ 具有形式 $\mathscr{R}\times E^n$，其中 $\mathscr{R}$ 是 (t, x)-空间中的区域．因此在变分问题的控制表达形式中，对所有的 (t, x)，集合 $\Omega(t, x)$ 是整个 E^n．我们现在叙述并证明变分法的一个标准存在定理．

定理 6.2　设 f^0 在 $\mathscr{G}=\mathscr{R}\times E^n$ 上是下半连续的，并且对一切 $(t, x, z)\in\mathscr{G}$，$f^0(t, x, z)\geqslant 0$．假定容许轨线集合非空且存在致密子集 $\mathscr{R}_0\subset\mathscr{R}$，使得所有的容许轨线的图形位于 $\mathscr{R}_0$ 中．设 $\mathscr{B}$ 是闭的，g 在 $\mathscr{B}$ 上是下半连续的，又假设存在定义于 $[0, \infty)$ 上的非负函数 Φ，使得当 $\xi\to\infty$ 时，$\Phi(\xi)/\xi\to\infty$；并使得对所有 $(t, x, z)\in\mathscr{R}_0\times E^n$，$f^0(t, x, z)\geqslant\Phi(|z|)$．则存在绝对连续函数 ϕ^*，它是容许弧，并使

$$J(\phi)=g(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))+\int_{t_0}^{t_1}f^0(t,\ \phi(t),\ \phi'(t))dt$$

取极小值．

可以直接验证，如果把变分问题写成控制问题，则在定理 6.2 的假设下，假设 4.1 的各款都成立．因为所有轨线都位于致密集 $\mathscr{R}_0$ 中，且 $\mathscr{B}$ 是闭的，这表明所有轨线的端点处于 $\mathscr{B}$ 的致密子集中．因此，由 g 在 $\mathscr{B}$ 上是下半连续的，可见它在端点集上是下有界的．于是由 $G(\phi)=g(t_0, \phi(t_0), t_1, \phi(t_1))$ 定义的泛函 G 在 $\mathscr{A}_T$ 上是下半连续且下有界的．由此以及不等式 $f^0(t, x, z)\geqslant 0$，

可见若 $\mu=\inf\{J(\phi):\phi\in\mathscr{A}_T\}$，则 μ 是有限的。此外，若 $\{\phi_k\}$ 为极小化叙列，则存在正常数 A，使得

$$\int_{t_{0k}}^{t_{1k}} f^0(t,\ \phi_k(t),\ u_k(t))dt\leqslant A,$$

其中 $u_k(t)=\phi_k'(t)$。注意，u_k 是几乎处处有限的。

设给定 $\varepsilon>0$，则存在正数 M_ε，使得若 $|u_k(t)|\geqslant M_\varepsilon$，有

$$\frac{f^0(t,\ \phi_k(t),\ u_k(t))}{u_k(t)}\geqslant\frac{\Phi(|u_k(t)|)}{u_k(t)}\geqslant\frac{1}{\varepsilon}.$$

于是对几乎所有 $t\in[t_{0k},\ t_{1k}]$，

$$|u_k(t)|\leqslant M_\varepsilon+\varepsilon f^0(t,\ \Phi_k(t),\ u_k(t)).$$

因为在变分法的简单问题的控制表达形式中，$f(t,\ x,\ z)=z$，可见具有 $H^i=f^0$ 的(6.1)和(6.2)对所有极小化序列的元素 $(\phi_k,\ u_k)=(\phi_k,\ \phi_k')$ 和所有 $i=1,\ \cdots,\ n$ 成立。

到此为止，我们已验证了在定理 6.2 的假设下，除了 $\mathscr{Q}^+(t,\ x)$ 满足弱 Cesari 性质的假设外，定理 6.1 的所有假设都成立。在下面的引理 6.1 中，我们将证明 $\mathscr{Q}^+(t,\ x)$ 满足 Cesari 性质，从而定理 6.1 的所有假设全部满足，这就完成了定理 6.2 的证明。

引理 6.1 在定理 6.2 的假设下，映射 $\mathscr{Q}^+(t,\ x)$ 满足 Cesari 性质。

证明：设 $(t,\ x)$ 为 $\mathscr{R}$ 中的点，$y=(y^1,\ \cdots,\ y^n)$，并假设

$$\hat{y}=(y^0,\ y)\in\bigcap_{\delta>0}\text{cl co}\,\mathscr{Q}^+(N_\delta(t,\ x)).$$

我们必须证明 $\hat{y}\in\mathscr{Q}^+(t,\ x)$。

设 $\{\delta_k\}$，$\{\hat{y}_k\}$，$\{\hat{y}_{ki}\}$，$\{z_{ki}\}$，$\{\alpha_{ki}\}$ 和 $\{(t_{ki},\ x_{ki})\}$，$i=1,\ \cdots,\ n+2$ 与定理 5.1 的证明所设相同，那里是为了证明定理 5.1 的映射 $\mathscr{Q}^+(t,x)$ 具有 Cesari 性质。因为现在 $\mathscr{D}=\mathscr{G}$ 而且不是致密的，我们必须用不同的方法证明。由于现在的 $f(t,\ x,\ z)=z$，我们有

$$\{y_{ki}\}=\{z_{ki}\},\ i=1,\ \cdots,\ n+2. \tag{6.4}$$

从(5.7)和(5.8)，我们得到

$$\begin{cases} y_k^0=\sum_{i=1}^{n+2}\alpha_{ki}y_{ki}^0, \\ y_{ki}^0\geqslant f^0(t_{ki},\ x_{ki},\ z_{ki})\geqslant 0. \end{cases} \tag{6.5}$$

因为 $\{y_k^0\}$ 收敛，故它是有界的。因为对每一个 i，$\alpha_{ki}\geqslant 0$，$y_{ki}^0\geqslant 0$，故对每个 i，叙列 $\{\alpha_{ki}, y_{ki}\}$ 也是有界的。还有，对每个 i，叙列 $\{\alpha_{ki}\}$ 是有界的。于是存在非负实数 $\alpha_1, \cdots, \alpha_{n+2}$，且 $\sum\alpha_i=1$，实数组 $\eta_1, \cdots, \eta_{n+2}$，以及指标集 $\{k\}$ 的子叙列(我们仍把它记为 $\{k\}$)，使得对每个 i，

$$\alpha_{ki}\to\alpha_i,\quad \alpha_{ki}y_{ki}^0\to\eta_i. \tag{6.6}$$

因为 $\sum\alpha_i=1$，且 $\alpha_i\geqslant 0$。可见存在正整数 $s\leqslant n+2$，若 $s<n+2$，使得当 $i=1, \cdots, s$，$\alpha_i>0$；而当 $i=s+1, \cdots, n+2$，$\alpha_i=0$。(如有必要可将 α_i 的顺序重新排列)。因此对 $i=1, \cdots, s$，

$$y_{ki}^0\to\eta_i/\alpha_i. \tag{6.7}$$

令 $y_i^0=\eta_i/\alpha_i$，按假设我们有

$$\frac{f^0(t, x, z)}{|z|}\geqslant\frac{\Phi(|z|)}{|z|}\to\infty,\ \text{当}\ |z|\to\infty. \tag{6.8}$$

从(6.5)中的不等式，以及(6.7)和(6.8)可见序列 $\{(z_{k1}, \cdots, z_{ks})\}$ 是有界的，因此存在子叙列(我们仍将它记为 $\{(z_{k1}, \cdots, z_{ks})\}$)以及点组 $z_1, \cdots, z_s$ 使得

$$z_{ki}\to z_i. \tag{6.9}$$

现在假设 $i>s$。如果序列 $\{z_{ki}\}$ 是有界的，则由于 $\alpha_{ki}\to 0$，可见 $\alpha_{ki}z_{ki}\to 0$。如果序列 $\{z_{ki}\}$ 是无界的，则存在子叙列 $\{z_{ki}\}$ 使得 $|z_{ki}|\to\infty$。那么根据(6.5)和(6.8)

$$\alpha_{ki}y_{ki}^0\geqslant\alpha_{ki}f^0(t_{ki}, x_{ki}, z_{ki})\geqslant\alpha_{ki}|z_{ki}|A_{ki},$$

其中 $A_{ki}\to\infty$，当 $k\to\infty$。但从(6.6)，$\alpha_{ki}y_{ki}^0\to\eta_i$。所以，如果 $i>s$，我们必有 $\alpha_{ki}z_{ki}\to 0$。将此结果与(6.6)和(6.9)联合起来，即得

$$\sum_{i=1}^{n+2}\alpha_{ki}z_{ki}\to\sum_{i=1}^{s}\alpha_i z_i, \tag{6.10}$$

其中 $\alpha_i\geqslant 0$，且 $\sum\limits_{i=1}^{s}\alpha_i=1$。

因为 f^0 对每一固定的 (t, x) 关于 z 是凸的，并且是下半连续的，故有

$$\begin{aligned}y^0=\lim_k y_k^0=\lim_k\sum_{i=1}^{n+2}\alpha_{ki}y_{ki}^0\\ \geqslant\lim_k\inf\sum_{i=1}^{n+2}\alpha_{ki}f^0(t_{ki},\ x_{ki},\ z_{ki})\\ \geqslant\lim_k\inf f^0\Big(t_{ki},\ x_{ki},\ \sum_{i=1}^{n+2}\alpha_{ki}z_{ki}\Big)\\ \geqslant f^0\Big(t,\ x,\ \sum_{i=1}^{s}\alpha_i z_i\Big).\end{aligned}$$

从(6.4)我们有

$$y_k=\sum_{i=1}^{n+2}\alpha_{ki}z_{ki}.$$

因为 $y_k\to y$, 从 (6.10) 可见 $y=\sum_{i=1}^{s}\alpha_i z_i$. 因此 $y^0\geqslant f^0(t,\ x,\ y)$, 而这就说明 $(y^0,\ y)\in\mathscr{Q}^+(t,\ x)$, 所以引理获得证明.

习题 6.1 设 Φ 是$[0,\infty)$上不减非负函数, 使得 $\Phi(\xi)/\xi\to\infty$, 当 $\xi\to\infty$. 设 f^0 是 $\mathscr{G}=\mathscr{R}\times E^m$ 上的实值下半连续函数, 使得对每一 $(t,\ x)\in\mathscr{R}$, f^0 是 z 的凸函数, 且 $f^0(t,\ x,\ z)\geqslant\Phi(|z|)$ 对所有 $(t,\ x,\ z)\in\mathscr{G}$ 成立. 设 $f(t,\ x,\ z)=h(t,\ x)+B(t,\ x)z$, 其中 h 和 B 对固定的 t 关于 x 是连续的. 设

$$\mathscr{Q}^+(t,\ x)=\{(\eta,\ \xi):\eta\geqslant f^0(t,\ x,\ z),\ \xi=f(t,\ x,\ z),\ z\in E^m\}.$$

证明 $\mathscr{Q}^+$ 在 $\mathscr{R}$ 中具有弱 Cesari 性质.

习题 6.2 设 $\mathscr{R}_0$ 是 $(t,\ x)$ 空间的致密子集. 除了 $K=0$ 外, Ω、$\mathscr{D}$、f^0 和 f 如假设 4.1. 函数 $f=(f^1,\ \cdots,\ f^n)$ 称为在 $\mathscr{R}_0$ 上一致地比 f^0 增长慢, 如果对每一个 $\varepsilon>0$, 存在 $K(\varepsilon)$ 使得对所有 $(t,\ x,\ z)\in\mathscr{D}$, 不等式 $|f(t,\ x,\ z)|\leqslant\varepsilon f^0(t,\ x,\ z)$ 成立, 其中 $|z|>K(\varepsilon)$. 假设恒等于 1 的函数和函数 f 二者都在 $\mathscr{R}_0$ 上一致地比 f^0 增长慢, 又假设对每一 $(t,\ x)\in\mathscr{R}_0$, $\mathscr{Q}^+(t,\ x)$ 是凸的. 证明 $\mathscr{Q}^+(t,\ x)$ 具有 Cesari 性质. (见 Cesari[18], p. 539～540.)

习题 6.3 在本习题中我们将研究具有"线性对象、凸积分代价准则以及无界控制"问题的存在定理. 本习题所述结果的另一种证明方法将在本章的习题 8.6 中讨论. 本习题中所证结果的推广将在习题 7.4 中考虑.

设 Φ 是 $[0, \infty)$ 上的不减非负函数，使得 $\Phi(\xi)/\xi\to\infty$，当 $\xi\to\infty$. 设 $\mathscr{R}$ 是 E^{n+1} 中的带形域 $[a, b]\times E^n$. 假设 f^0 是 $\mathscr{G}=\mathscr{R}\times E^m$ 上的实值下半连续函数，使得 f^0 对每一 $(t, x)\in\mathscr{R}$ 是 z 的凸函数，并且对所有 $(t, x, z)\in\mathscr{G}$，$f^0(t, x, z)\geqslant\Phi(|z|)$. 设

$$f(t, x, z)=A(t)x+B(t)z+h(t), \tag{6.11}$$

其中 A、B 和 h 在 $[a, b]$ 上是连续的. 设 $t_0=a$，而 $x_0=\xi_0$，其中 ξ_0 固定. 设 $\mathscr{X}$ 是 E^n 中的闭集，$\mathscr{B}=\{(a, \xi_0, b, x_1): x_1\in\mathscr{X}\}$. 设 $\Omega(t, x)=E^m$，容许对集合非空，G 在 $\mathscr{A}_T$ 上是下半连续且下有界的.

证明在上述假设下使 II(3.3) 所定义的泛函 $\hat{J}$ 取极小值的容许对存在.

提示：定理 6.2 的结论表明极小化叙列中的控制 u_k 具有等度绝对连续的积分，将此结果与参数变易公式以及所有容许轨线的初始点是固定的假设联合起来即可得到极小化序列中所有轨线位于致密集 $\mathscr{R}_0$ 中. 从习题 6.1 我们得知弱 Cesari 性质成立.

习题 6.4　在本习题中，我们研究一些在应用中遇到的泛函 G. 设 $\mathscr{I}$ 和 $\mathscr{X}$ 如习题 6.3 中所设，并假设 g 在 $\mathscr{X}$ 上是下半连续的. 证明由 $G(\phi)=g(\phi(b))$ 定义的泛函 G 满足习题 6.3 的假设. 对由 $G(\phi)=\max\{|\phi(t)-\xi(t)|: t\in[a, b]\}$ (其中 ξ 是定义在 $\mathscr{I}=[a, b]$ 上的已知连续函数)定义的泛函 G 证明同样的结论.

习题 6.5　本习题中我们得到"线性对象二次型积分代价准则问题"的存在定理. 此定理的另一证明方法将在习题 8.3 和 8.6 中讨论.

设 $\mathscr{R}$、$\mathscr{G}$、$\mathscr{X}$、Ω、G 和 f 如习题 6.3. 对所有 $(t, x, z)\in\mathscr{G}$，设

$$f^0(t, x, z)=\langle x, X(t)x\rangle+\langle z, R(t)z\rangle,$$

其中 X 和 R 是定义在 $[a, b]$ 上的连续函数矩阵，对所有 $t\in[a, b]$，$X(t)$ 是半正定的，而对一切 $t\in[a, b]$，$R(t)$ 是正定的.

(a) 证明当且仅当 $u(t)\in L_2[a, b]$ 时

$$\int_a^b \langle u(t),\ R(t)u(t)\rangle dt<\infty.$$

提示：对任何半正定的对称方阵 Q，我们有

$$\{\langle\xi,\ Q\xi\rangle:|\xi|=1\}=\left\{\sum_{i=1}^n \lambda_i\eta_i^2:|\eta|=1\right\},$$

其中 $\lambda_i\geqslant 0(i=1,\ \cdots,\ n)$ 是 Q 的特征值．因此若 $\lambda_1(t)$ 表示 $R(t)$ 的最大特征值而 $\lambda_n(t)$ 表示其最小特征值，则

$$|u(t)|^2\lambda_n(t)\leqslant\langle u(t),\ R(t)u(t)\rangle\leqslant|u(t)|^2\lambda_1(t)$$

对一切 $t\in[a,\ b]$ 成立．因为 λ_n 是连续的且 $\lambda_n(t)>0$，它在 $[a,\ b]$ 上是有界的且恒不为零．由此可得所需的结果．

注意，我们已证明若 $\mathscr{X}=E^n$，则容许控制集合是 $L_2[a,\ b]$．

(b) 证明在本习题的假设下，若存在容许对，则必存在使 II(3.3)中所定义的泛函 $\hat{J}$ 取极小值的容许对．

提示：可以应用习题 6.3．与习题 6.3 无关的另一证明如下．对适当的常数 M，在极小序列中我们有 $\int_a^b |u_k|^2dt\leqslant M$，由参数变易公式和 Cauchy-Schwartz 不等式，则给出极小化序列中的轨线 $\{\phi_k\}$ 的等度绝对连续性．因为所有轨线的初始点是相同的，它们都必须位于一个致密集中．习题 6.1 又给出弱 Cesari 性质．

(c) 如果对所有 t，$R(t)$ 是半正定的要求用对所有的 t，$R(t)$ 是正定的要求代替，上述结果是否正确？

7．定理 4.1 的证明

下面的定理 7.1 在定理 4.1 的证明中和最优控制理论的其它方面都是重要的．

如果 T 是测度空间，而 Z 是 Hausdorff 空间，则从 T 到 Z 的映射 Γ 叫做可测的，若在映射 Γ 之下，Z 中的一个致密集的原象是 T 中的可测集．

定理 7.1 设 T 为测度空间，Z 为 Hausdorff 空间，而 D 是拓扑空间，它是可数个致密的可度量化子集的并集．设 Γ 是 $T\to Z$ 的可测映射，Φ 是 $D\to Z$ 的连续映射，使得 $\Gamma(T)\subseteq\Phi(D)$，则存在

$T\to D$ 的可测映射 μ，使得 $\Phi*\mu=\Gamma$，其中记号"$*$"表示两个映射的复合.

定理 7.1 将在第 10 节里证明.

我们还需要下面的基本结果，它的证明留给读者作为习题.

习题 7.1 设 $\{s_n\}$ 为 E^p 中的向量序列，它收敛于极限 s. 设 $\{n_j\}$ 是整数的子序列，令

$$\sigma_{n_j}=\sum_{i=1}^{k}\alpha_{ij}s_{n_j+i},$$

其中 k 依赖于 n_j，且 $\alpha_{ij}\geqslant 0$，$\sum_i \alpha_{ij}=1$. 则 $\sigma_{n_j}\to s$.

因为所有容许轨线的图形包含在致密集 $\mathscr{R}_0$ 中，故存在致密区间 $\mathscr{I}=[a,\ b]$，使得所有容许对 $(\phi,\ u)$ 的定义区间 $[t_0,\ t_1]$ 皆含于 $\mathscr{I}$ 之中. 如果 ϕ 是定义于 $[t_0,\ t_1]$ 上的容许轨线，设 $\tilde{\phi}$ 定义如下：

$$\tilde{\phi}(t)=\begin{cases}\phi(t_0), & \text{当 } a\leqslant t\leqslant t_0,\\ \phi(t), & \text{当 } t_0\leqslant t\leqslant t_1,\\ \phi(t_1), & \text{当 } t_1\leqslant t\leqslant b.\end{cases}$$

因此 $\tilde{\phi}$ 是 ϕ 的扩张，且除了 $\tilde{\phi}$ 的定义域取为 $[a,\ b]$ 而不是 $(-\infty,\ \infty)$ 之外，所有 $\tilde{\phi}$ 通常构成度量空间 $\mathscr{X}_\rho$. 因为所有容许轨线的定义区间都包含在 $\mathscr{I}$ 之中，我们可以假定度量空间 $\mathscr{X}_\rho$ 是限于定义 $\mathscr{I}$ 上的函数全体.

$\mathscr{R}_0$ 的致密性还可推出所有轨线的端点 $(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))$ 位于 $\mathscr{B}$ 的有界子集中. 因为 $\mathscr{B}$ 是闭的，所以所有端点必位于 $\mathscr{B}$ 的致密子集中. 于是我们可以假设 $\mathscr{B}$ 本身是致密的.

我们指出如下事实，它在定理 4.1 的证明中将会用到. 根据引理 3.1，假设 4.1 的(iii)等价于集合 $\mathscr{D}$ 是闭的假设. 而且上半连续性的定义和假设 4.1 的 (iii) 隐含着对每一 $(t,\ x)\in\mathscr{R}$，集合 $\Omega(t,\ x)$ 是闭的.

在定理 4.1 的证明中，我们将选取各种序列的子序列，除非另有说明，我们将用原序列的记号表示子序列. 我们把定理 4.1 的证明分成如下几步：

第一步．存在 $\mathscr{A}_0$ 中的元素序列 $\{(\phi_k, u_k)\}$，实数 γ，点 t_0 和 $t_1 \in \mathscr{I}$，且 $t_1 > t_0$，以及点 x_0 和 $x_1 \in E^n$，使得对 $i=0, 1$，$(t_i, x_i) \in \mathscr{R}_0$，$(t_0, x_0, t_1, x_1) \in \mathscr{B}$，并且

$$\begin{aligned} &t_{ik} \to t_i, \ \phi_k(t_{ik}) \to x_i, \qquad i=0, 1, \\ &I(\phi_k, u_k) \to \gamma. \end{aligned} \tag{7.1}$$

因为 f^0 在 $\mathscr{D}$ 上是下有界的，并且所有容许对的定义区间都包含在 $\mathscr{I}$ 中，由此可见 $\{I(\phi, u):(\phi, u) \in \mathscr{A}_0\}$ 是下有界的；又按假设此数集是上有界的．因而存在实数 γ 和 $\mathscr{A}_0$ 中的元素列 $\{(\phi_k, u_k)\}$ 使得 $I(\phi_k, u_k) \to \gamma$．因为 $\mathscr{B}$ 是致密的，存在子叙列 $\{(\phi_k, u_k)\}$，使得 $\{(t_{0k}, \phi_k(t_{0k}), t_{1k}, \phi_k(t_{1k}))\}$ 收敛于 $\mathscr{B}$ 中之点 (t_0, x_0, t_1, x_1)．由于 $t_{1k} > t_{0k} + \delta$，且 $(t_{ik}, \phi_k(t_{ik})) \in \mathscr{R}_0$，由此可得 $t_1 \geqslant t_0 + \delta$，且 $(t_i, x_i) \in \mathscr{R}_0$．

第二步．存在定义于 $[t_0, t_1]$ 上的绝对连续函数 ϕ 和子序列 $\{\phi_k\}$ 使得在 $\mathscr{X}_r$ 中 $\phi_k \to \phi$，并且它们的扩张函数 $\tilde{\phi}'_k$ 在 $L_1[\mathscr{I}]$ 中弱收敛于 $\tilde{\phi}'$．此外，ϕ 满足定义 II．3.1 的条件(i)和定义 II．3.2 的(iii)．

因为所有的容许轨线的图形位于致密集 $\mathscr{R}_0$ 中，函数族 $\{\phi_k\}$ 一致有界，所以它们的扩张 $\{\tilde{\phi}_k\}$ 亦然．因为 $\{\phi_k\}$ 等度绝对连续，它们的扩张 $\{\tilde{\phi}_k\}$ 也一样是等度绝对连续的．因此从 Ascoli 定理可知，存在子序列 $\{\tilde{\phi}_k\}$ 和定义于 $\mathscr{I}$ 的函数 $\tilde{\phi}$，使得在 $\mathscr{I}$ 上 $\tilde{\phi}_k$ 一致收敛于 $\tilde{\phi}$，并且 $\tilde{\phi}$ 是绝对连续的，所以 $\tilde{\phi}'$ 存在且属于 L_1．因此

$$\tilde{\phi}(t) = \tilde{\phi}(a) + \int_a^t \tilde{\phi}'(s)\,ds, \quad a \leqslant t \leqslant b. \tag{7.2}$$

令

$$\phi(t) = x_0 + \int_{t_0}^t \tilde{\phi}'(s)\,ds, \quad t_0 \leqslant t \leqslant t_1. \tag{7.3}$$

我们断定 $\tilde{\phi}$ 是这个 ϕ 的扩张．为了弄清这一事实，首先假设 $t_0 > a$ 而 $t_2 < t_0$，因为 $t_{0k} \to t_0$，可见对充分大的 k，$t_{0k} > t_2$，于是 $\tilde{\phi}_k(t_2) = \phi_k(t_{0k})$．因为 $\tilde{\phi}_k(t_2) \to \tilde{\phi}(t_2)$，而 $\phi_k(t_{0k}) \to x_0$，我们得到对一切 $a \leqslant t_2 < t_0$，$\tilde{\phi}(t_2) = x_0$，由 $\tilde{\phi}$ 的连续性得知 $\tilde{\phi}(t_0) = x_0$．如果 $t_0 = a$，则对一切 k，$\tilde{\phi}_k(t_0) = \tilde{\phi}_k(a) = \phi_k(t_{0k})$．现在若令 $k \to \infty$，即得

$\tilde{\phi}(t_0)=x_0$. 因此对所有 $t\in[a, t_0]$, $\tilde{\phi}(t)=\tilde{\phi}(t_0)=x_0$. 用类似方法可以证明对于 $t_1\leqslant t\leqslant b$, $\tilde{\phi}(t)=\tilde{\phi}(t_1)=x_1$. 因此在 $\mathscr{I}=[t_0, t_1]$ 上 $\tilde{\phi}'(t)=0$. 从(7.3)我们有 $\phi(t_0)=x_0$, 所以对 $a\leqslant t\leqslant t_0$, $\tilde{\phi}(t)=\phi(t_0)$. 因为 $\tilde{\phi}(a)=x_0$, 从 (7.2) 以及在 $[t_0, t_1]$ 外面 $\tilde{\phi}'$ 等于零, 我们得知对于 $t\in[t_0, t_1]$,

$$\tilde{\phi}(t)=x_0+\int_a^t \tilde{\phi}'(s)ds=x_0+\int_{t_0}^t \tilde{\phi}'(s)ds=\phi(t).$$

对于 $t\geqslant t_1$ 我们有

$$\tilde{\phi}(t)=x_0+\int_{t_0}^{t_1} \tilde{\phi}'(s)ds=\phi(t_1).$$

于是 $\tilde{\phi}$ 是 ϕ 的扩张. 因此,对 $t\in[a, b]$, 我们有 $\phi'(t)=\tilde{\phi}'(t)$, 且 $\phi(t)=\tilde{\phi}(t)$.

因为 $t_{ik}\to t_i$, , $i=0, 1$, 我们已证明了在 $\mathscr{X}_\rho$ 中 $\phi_k\to\phi$.

从关系式

$$\tilde{\phi}_k(t)=\phi_k(t_{0k})+\int_a^t \tilde{\phi}_k'(s)ds,\ a\leqslant t\leqslant b,$$

由(7.2) 中以 x_0 代替 $\tilde{\phi}(a)$, 并由 (7.1) 式,以及 $\tilde{\phi}_k\to\tilde{\phi}$ 的收敛性, 可知对一切的 $t\in[a, b]$,

$$\int_a^t \tilde{\phi}_k'(s)ds\to\int_a^t \tilde{\phi}'(s)ds.$$

因为函数 $\tilde{\phi}_k$ 是等度绝对连续的, 它的导函数 $\tilde{\phi}_k'$ 具有等度绝对连续的积分. 因此由定理 3.1, $\tilde{\phi}_k'$ 在 $L_1[\mathscr{I}]$ 中弱收敛于 $\tilde{\phi}'$.

因为 $\tilde{\phi}$ 是满足定义 II. 3.1 的(i) 和定义 II. 3.2 的 (iii) 的函数列的一致极限,故这些条件对 $\tilde{\phi}$ 从而也对 ϕ 同样成立.

第三步. 存在于 $[t_0, t_1]$ 上可积的实值函数 λ, 使得 $(\lambda(t), \phi'(t))\in\mathscr{Q}^+(t, \phi(t))$ 在 $[t_0, t_1]$ 上几乎处处满足,并使

$$\int_{t_0}^t \lambda(s)ds\leqslant\gamma. \tag{7.4}$$

因为 $\tilde{\phi}_k'$ 在 L_1 中弱收敛于 $\tilde{\phi}'$, 从定理 3.2 可得如下命题. 对每一个整数 j, 存在整数 n_j, 整数集 $i=1, \cdots, k$, 其中 $k=k(j)$ 依赖于 j, 以及满足

$$\alpha_{ij}\geqslant 0,\ i=1,\cdots,k;\qquad \sum_{i=1}^{k}\alpha_{ij}=1 \tag{7.5}$$

的实数组 $\alpha_{1j},\cdots,\alpha_{kj}$，使得 $n_{j+1}>n_j+k(j)$，且当 $j\to\infty$ 时

$$\int_a^b\left|\tilde{\phi}'-\sum_{i=1}^{k}\alpha_{ij}\phi'_{n_j+i}\right|dt\to 0. \tag{7.6}$$

令

$$\psi_j=\sum_{i=1}^{k}\alpha_{ij}\tilde{\phi}'_{n_j+i}.$$

回顾对每一个整数 q，如果 $t\notin[t_{0q},t_{1q}]$，则 $\tilde{\phi}'_q(t)=0$，而且 u_q 和 ϕ_q 只在 $[t_{0q},t_{1q}]$ 上定义。如果对 $t\notin[t_{0q},t_{1q}]$，我们就规定

$$f(t,\phi_q(t),u_q(t))=0,$$

于是我们能将 ψ_j 写成：

$$\psi_j(t)=\sum_{i=1}^{k}\alpha_{ij}f(t,\phi_{n_j+i}(t),u_{n_j+i}(t)). \tag{7.7}$$

按 ψ_j 的定义，(7.6) 式说明在 $L_1[\mathscr{I}]$ 中 $\psi_j\to\tilde{\phi}'$。于是存在子序列 $\{\psi_j\}$，使得

$$\psi_j(t)\to\tilde{\phi}'(t) \tag{7.8}$$

在 $\mathscr{I}$ 中几乎处处成立。我们现在假设(7.7)式是这种子序列。相应于序列(7.7)我们规定序列 $\{\lambda_j\}$ 如下：

$$\lambda_j(t)=\sum_{i=1}^{k}\alpha_{ij}f^0(t,\phi_{n_j+i}(t),u_{n_j+i}(t)), \tag{7.9}$$

其中若 $t\notin[t_{0q},t_{1q}]$，我们令 $f^0(t,\phi_q(t),u_q(t))=0$，并且对每一个 j，数 α_{ij}，下标 n_j+i 以及函数 ϕ_{n_j+i} 和 u_{n_j+i} 如(7.7)中所示。注意如果 $t\notin[t_0,t_1]$，存在 $j_0=j_0(t)$，使得若 $j>j_0$，则 $\psi_j(t)=0$ 和 $\lambda_j(t)=0$。

定义

$$\lambda(t)=\liminf\lambda_j(t). \tag{7.10}$$

因为 $f^0\geqslant -K$，所以 $\lambda_j\geqslant -K$。于是 $\lambda\geqslant -K$。此外，如果 $t\notin[t_0,t_1]$，则 $\lambda(t)=0$。因此，若令 $f_q^0(t)\equiv f^0(t,\phi_q(t),u_q(t))$，并利用 Fatou 定理，我们得到

$$\int_{t_0}^{t_1}\lambda\,dt=\int_a^b\lambda\,dt\leqslant\liminf_{j\to\infty}\left[\sum_{i=1}^{k}\alpha_{ij}\int_a^b f_{n_j+i}^0\,dt\right]$$
$$=\liminf_{j\to\infty}\left[\sum_{i=1}^{k}\alpha_{ij}\int_{t_0,n_j+i}^{t_1,n_j+i} f_{n_j+i}^0\,dt\right]$$
$$=\liminf_{j\to\infty}\left[\sum_{i=1}^{k}\alpha_{ij}I(\phi_{n_j+i},\ u_{n_j+i})\right].$$

从(7.1)我们有 $I(\phi_{n_j+i},\ u_{n_j+i})\to\gamma$, 当 $j\to\infty$. 然后从(7.5)和习题(7.1)可得

$$\lim_{j\to\infty}\sum_{i=1}^{k}\alpha_{ij}I(\phi_{n_j+i},\ u_{n_j+i})=\gamma.$$

因此(7.4)式得证. 因为 λ 是下有界的, 由此可见 $\lambda(t)\in L_1[\mathscr{I}]$, 并且在 $[t_0,\ t_1]$ 上几乎处处有限.

我们现在证明 $(\lambda(t),\ \phi'(t))\in\mathscr{Q}^+(t,\ \phi(t))$ 在 $[t_0,\ t_1]$ 上几乎处处满足. 设 T_1 表示 $[t_0,\ t_1]$ 中的点集, 在其上 $\lambda(t)$ 是有限的, 且 $\psi_j(t)\to\phi'(t)$. 对每一个整数 k, 定义集合 E_k 如下:

$$E_k=\{t: t\in[t_{0k},\ t_{1k}],\ u_k(t)\notin\Omega(t,\ \phi_k(t))\}.$$

则按定义 II. 3.2 的(ii), $\mathrm{meas}(E_k)=0$. 令 E 表示所有 E_k 的并集, 并令 T_2 表示 $[t_0,\ t_1]$ 中不属于 E 的点的集合, $T'=T_1\cap T_2$, 显然 $\mathrm{meas}\,T'=t_1-t_0$.

设 t 是 T' 中固定元素, $t\neq t_i$, $i=0,1$, 则存在子序列 $\{\lambda_j(t)\}$, 一般地说它依赖于 t, 使得 $\lambda_j(t)\to\lambda(t)$. 对于相应的子叙列 $\{\psi_j(t)\}$, 从(7.8)以及 t 是 $(t_0,\ t_1)$ 的内点的事实, 我们有 $\psi_j(t)\to\phi'(t)$. 因为 t 是 $(t_0,\ t_1)$ 内点, 且 $t_{ik}\to t_i$, $i=0,\ 1$, 可见存在 j_0, 使得若 $j>j_0$, 则 $t\in(t_{0,n_j+i},\ t_{1,n_j+i})$. 对每一个 $\delta>0$, 存在整数 $k_0=k_0(\delta)$, 使得如果 $k>k_0(\delta)$, 则 $|\phi_k(t)-\phi(t)|<\delta$. 于是, 对 $k>k_0$,

$$(t,\ \phi_k(t))\in N_{\delta}(t,\ \phi(t)).$$

因此对充分大的 j

$$\hat{f}(t,\ \phi_{n_j+i}(t),\ u_{n_j+i}(t))\in\mathscr{Q}^+(N_{\delta}(t,\ \phi(t))),$$

其中 $\hat{f}=(f^0,\ f)$. 因此, 由(7.7)、(7.9)和(7.5),

$$(\lambda_j(t),\ \psi_j(t))\in\mathrm{co}\,\mathscr{Q}^+(N_{\delta}(t,\ \phi(t))).$$

因为 $\lambda_j(t)\to\lambda(t)$, $\psi_j(t)\to\phi'(t)$, 我们有

$$(\lambda(t),\ \phi'(t))\in \mathrm{cl\ co}\,\mathscr{Q}^+(N_{\varepsilon\delta}(t,\ \phi(t))).$$

因为 δ 是任意的, 对每一个 $\delta>0$,

$$(\lambda(t),\ \phi'(t))\in \mathrm{cl\ co}\,\mathscr{Q}^+(N_{x\delta}(t,\ \phi(t))),$$

从而 $(\lambda(t),\ \phi'(t))$ 必在这些集合的交集之中, 由于 $\mathscr{Q}^+(t,\ \phi(t))$ 满足弱 Cesari 性质, 我们得知 $(\lambda(t),\ \phi'(t))\in\mathscr{Q}^+(t,\ \phi(t))$. 因为 t 是 T' 中不同于 t_0 或 t_1 的任意点, 我们得到所需的结果.

第四步. 存在定义于 $[t_0,\ t_1]$ 上的可测函数 u, 使得对几乎所有的 t: (i) $\phi'(t)=f(t,\ \phi(t),\ u(t))$; (ii) $u(t)\in\Omega(t,\ \phi(t))$; (iii) $\lambda(t)\geqslant f^0(t,\ \phi(t),\ u(t))$.

满足第四步结论的函数 u 的存在性是 $(\lambda(t),\ \phi'(t))\in\mathscr{Q}^+(t,\ \phi(t))$ 的另一种说法. 问题是要去证明存在可测函数 u 具有此性质. 我们将用定理 7.1 完成这个证明.

设 $T=\{t:(\lambda(t),\ \phi'(t))\in\mathscr{Q}^+(t,\ \phi(t))\}$, 令

$$Z=E^1\times E^n\times E^n\times E^1,$$

$$D=\{(t,\ x,\ z,\ \eta):(t,\ x,\ z)\in\mathscr{D},\ \eta\geqslant f^0(t,\ x,\ z)\}.$$

集合 T 是 Lebesgue 可测的, 因此它是具有 Lebesgue 测度的测度空间. 空间 Z 是 Hausdorff 空间. 在第一步前面的讨论中, 我们已指出集 $\mathscr{D}$ 是闭的. 因为 f^0 是下半连续的, 可见集 D 也是闭的. 于是 D 可以写成可数个致密子集 D_i 的并集, 其中 D_i 是 D 与中心在 E^{2n+2} 的原点而半径为 i 的闭球的交集.

令 Γ 表示由 $t\to(t,\ \phi(t),\ \phi'(t),\ \lambda(t))$ 所定义的 $T\to Z$ 的映象. 因为每一个函数 ϕ、ϕ'、λ 是可测的, 故 Γ 是可测映射. 设 Φ 表示由 $(t,\ x,\ z,\ \eta)\to(t,\ x,\ f(t,\ x,\ z),\ \eta)$ 所定义的从 D 到 Z 的映象. 因为 f 是连续的, 故 Φ 也是连续的. 又从第三步我们得知 $\Gamma(T)\subseteq\Phi(D)$. 因此定理 7.1 的所有假设都满足. 于是, 存在从 T 到 D 的可测映象 μ, 即

$$\mu:\ t\to(\tau(t),\ x(t),\ u(t),\ \eta(t)),$$

使得对一切 $t\in T$

$$\Phi(\mu(t))=\Phi(\tau(t),\ x(t),\ f(\tau(t),\ x(t),\ u(t)),\ \eta(t))^{[注]}$$
$$=\Gamma(t)=(t,\ \phi(t),\ \phi'(t),\ \lambda(t)),$$

由此得到第四步的结论.

第五步. 证明的完成.

设$(\phi^*, u^*)=(\phi, u)$, 其中ϕ是第二步得到的函数, 而u是第四步得到的函数. 于第二步中我们已证明存在序列$\{(\phi_k, u_k)\}$, 使得在$\mathscr{X}_p$中$\phi_k\to\phi^*$, 并且满足定义 II. 3.1 的(i)和定义 II. 3.2 的(iii). 第四步中的命题(i)断定(ϕ^*, u^*)满足定义 II. 3.1 的(ii). 因此为了证明(ϕ^*, u^*)是容许对, 只须再证明$t\to f^0(t, \phi(t), u(t))$在$[t_0, t_1]$上是 Lebesgue 可积的. 因为$\phi$和$u$是可测的且$f^0$是下半连续的, 故$t\to f^0(t, \phi(t), u(t))$是可测的. 因为$\lambda$可积且$f^0$是下有界的, 于是由第四步的命题(iii)可知$t\to f^0(t, \phi(t), u(t))$是可积的.

从第四步的(iii)和(7.4)式, 我们得到

$$I(\phi^*, u^*)=\int_{t_0}^{t_1} f^0(t, \phi^*(t), u^*(t))dt\leqslant\gamma.$$

但是由(7.1)式, 我们有$I(\phi_k, u_k)\to\gamma$. 将最后这两个命题结合起来就得到(4.3)式, 因此定理 4.1 的证明完毕.

在附注 6.2 中, 我们断定如果当$H^i=f^0+K$时, (6.1)和(6.2)对一切极小化序列中的(ϕ_k, u_k)成立, 则(6.1)和(6.2)对(ϕ^*, u^*)成立. 我们现在证明此论断.

设$\{(\phi_k, u_k)\}$为极小化序列, 那么如果

$$\mu=\inf\{J(\phi, u):(\phi, u)\in\mathscr{A}\},$$

我们有$I(\phi_k, u_k)\to\mu$. 于是, 若取$\mathscr{A}_0=\{(\phi_k, u_k)\}$, 则由第一步我们有$\mu=\gamma$. 从(6.1)和$H^i=f^0+K$的假设, 可见$I(\phi_k, u_k)+K(t_{1k}-t_{0k})\leqslant A$. 因为$I(\phi^*, u^*)=\mu$, 以及对适当的子序列$t_{ik}\to t_i$, $i=0, 1$, 我们得到$I(\phi^*, u^*)+K(t_1-t_0)\leqslant A$. 而这就是说

$$\int_{t_0}^{t_1}[f^0(t, \phi^*(t), u^*(t))+K]dt\leqslant A,$$

[注] 原书误为$(\tau(t), x(t), f(\tau(t), x(t), u(t)), \eta(t))$.——译者注.

故(ϕ^*, u^*)满足具有 $H^i=f^0+K$ 的(6.1).

如果当 $H^i=f^0+K$ 时 (6.2) 对极小化序列 $\{(\phi_k, u_k)\}$ 成立，则从(7.7), (7.5)和关系式 $H^i=f^0+K$ 我们有

$$
\begin{aligned}
|\psi_j^i(t)| &\leqslant \sum_{q=1}^{k} \alpha_{qj} |f^i(t, \phi_{n_j+q}(t), u_{n_j+q}(t))| \\
&\leqslant F_\varepsilon^i(t)+\varepsilon \sum_{q=1}^{k} \alpha_{qj}[f^0(t, \phi_{n_j+q}(t), u_{n_j+q}(t))+K].
\end{aligned}
$$

现在从(7.9)可得

$$
|\psi_j^i(t)| \leqslant F_\varepsilon^i(t)+\varepsilon[\lambda_j(t)+K].
$$

设 t 是 T_1 中的点，此处 T_1 是第三步里定义的在 $[t_0, t_1]$ 中的点集，在其上 $\lambda(t)$ 有限且 $\psi_j(t)\to\phi'(t)$. 于是存在与 t 有关的子叙列 $\{\lambda_j(t)\}$，使得 $\lambda_j(t)\to\lambda(t)$. 因此对 $t\in T_1$

$$
|\phi^{i\prime}(t)| \leqslant F_\varepsilon^i(t)+\varepsilon[\lambda(t)+K]. \tag{7.11}
$$

因为 $\operatorname{meas} T_1=t_1-t_0$, 不等式(7.11)在 $[t_0, t_1]$ 上几乎处处成立.

从(7.4), 等式 $\mu=\gamma$ 以及第四步的(iii)我们有

$$
I(\phi, u)=\int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt \leqslant \int_{t_0}^{t_1} \lambda(t)dt \leqslant \mu.
$$

因为$(\phi^*, u^*)=(\phi, u)$, 且 $I(\phi^*, u^*)=\mu$, 我们有

$$
\int_{t_0}^{t_1}[\lambda(t)-f^0(t, \phi(t), u(t))]dt=0.
$$

现在从第四步的 (iii) 可得 $\lambda(t)=f^0(t, \phi^*(t), u^*(t))$ 几乎处处成立. 如果将此关系式代入 (7.11) 式中的 $\lambda(t)$, 并将 (7.11) 中的 $\phi^{i\prime}(t)$改写成 $\phi^{*i\prime}(t)$, 我们就得到(6.2)对(ϕ^*, u^*)成立.

习题 7.2 证明在定理 4.1 中可以用如下较弱的假设代替假设 4.1 的(v)：存在实值非负函数 $\psi\in L_1[\mathscr{I}]$ 和 n 维向量 b 使得对所有 $(t, x, z)\in\mathscr{D}$, $f^0(t, x, z)\geqslant-\psi(t)+\langle b, f(t, x, z)\rangle$.

习题 7.3 证明在定理 4.1 中可以用如下较弱的假设代替假设 4.1 的(vi)：设 $\mathscr{I}$ 是 E^1 中的致密区间, $\mathscr{R}=\mathscr{I}\times E^n$, $\mathscr{U}=E^m$. 对每一个 $t\in\mathscr{I}$, 设 $\hat{f}=(f^0, f)$是 E^{n+m} 上(x, z) 的连续函数，而对每一 $(x, z)\in E^{n+m}$, f 是 $\mathscr{I}$ 中 t 的可测函数.

提示：由 Scorza-Dragoni [55] 和 Vainberg ([57]§ 定理 18.2,

p. 148) 似乎是独立地发现的定理，我们可得如下命题．对每一个 $\varepsilon>0$，存在开集 $G\subset\mathscr{I}$ 使得 $\operatorname{meas}G<\varepsilon$，并使 $\hat{f}=(f^0, f)$ 在 $(\mathscr{I}-G)\times E^n\times E^m$ 上是连续的．由定理 4.1 的证明中的第四步，首先得知 $T-(G\cap T)$ 中可测函数 u 的存在性．又由于 ε 是任意的，于是我们得知在 $\mathscr{I}$ 上存在具有所需性质的可测函数 u.

习题 7.4 用习题 7.3 的结果将习题 6.3 的结果推广到 A 和 h 在 $\mathscr{I}$ 上是可积的而 B 在 $\mathscr{I}$ 上是有界可测的情形．

8．没有 Cesari 性质的存在定理

本节我们将叙述并证明两个存在定理．在这两个定理中没有假设弱 Cesari 性质成立．但 $\mathscr{Q}^+(t, x)$ 仍然假设是凸的．本节的两个存在定理都假设约束映射 Ω 仅与 t 有关．其中一个定理假设函数 $\hat{f}$ 满足广义 Lipschitz 条件．另一个定理假设极小化序列中的控制都位于某 L_p 空间 $(1\leqslant p\leqslant\infty)$ 的一个球内．

在第一个定理中，除了假设 4.1 之外，我们再做如下假设，并把它列为假设 8.1．注意，假设 8.1 的 (i) 取代了假设 4.1 的 (vi)．

假设 8.1 (i) 函数 $\hat{f}=(f^0, f)$ 是连续的．(ii) 集 $\Omega(t, x)$ 与 x 无关；亦即对给定的 t，$\Omega(t, x)=\Omega(t, x')=\Omega(t)$ 对一切使得 (t, x) 和 (t, x') 属于 $\mathscr{R}$ 的 x 和 x' 成立．(iii) 存在定义在 $[0, \infty)$ 上的不减函数 ω 使得 $\lim\limits_{\delta\to0}\omega(\delta)=0$，以及定义于 $E^1\times\mathscr{U}$ 上的非负函数 L 使得

$$|\hat{f}(t, x, z)-\hat{f}(t, x', z)|\leqslant L(t, z)\omega(|x-x'|) \tag{8.1}$$

对所有 $\mathscr{D}$ 中的 (t, x, z) 和 (t, x', z) 都成立．

注意，若 $\mathscr{D}$ 是致密的，$\hat{f}$ 在 $\mathscr{D}$ 上一致连续，则 (8.1) 对 $L=1$ 且 ω 取连续模成立．如果 $\hat{f}$ 对 x 满足 Lipschitz 条件，则 (8.1) 对 $\omega(\delta)=\delta$ 而 L 等于 Lipschitz 常数成立．

下面的定理可代替定理 4.1．我们将在叙述定理 8.3 之后给予证明．

定理 8.1 若假设 4.1 和 8.1 成立．令 $\mathscr{A}_0$ 为容许对 (ϕ, u)

的集合，使得轨线 ϕ 是等度绝对连续的，且 $\sup\{I(\phi, u):(\phi, u)\in\mathscr{A}_0\}$ 是有限的，并使得对于一切属 $\mathscr{A}_0$ 中的对 (ϕ, u) 中的容许控制 u，有

$$\int_{t_0}^{t_1} L(t, u(t))dt \leqslant A, \tag{8.2}$$

其中 $[t_0, t_1]$ 是 (ϕ, u) 的定义区间，而 A 是与 u 无关的常数．则存在 $\mathscr{A}_0$ 中的序列 $\{(\phi_k, u_k)\}$ 和 $\mathscr{A}$ 中的容许对 (ϕ^*, u^*) 具有如下性质：(i) 在 $\mathscr{X}_p$ 中，$\phi_k \to \phi^*$；(ii)

$$\varliminf_{k\to\infty} I(\phi_k, u_k) \geqslant I(\phi^*, u^*). \tag{8.3}$$

我们现在给出弱 Cesari 性质不成立但定理 8.1 的假设成立的例子．

例 8.1 设 $x=(x^1, x^2)$，z 为实数，$\Omega(t, x)=E^1$，$f^0\equiv 0$，而 $f=(z, x^1, z)$．假设端点条件如下：$t_0=0$，$t_1=1$，$x_0^1=0$，$0\leqslant x_0^2\leqslant 1$，$x_1^1$ 和 x_1^2 自由．令 $\mathscr{R}=\{(t, x):0\leqslant t\leqslant 1,\ |x^i|<M,\ i=1, 2\}$，其中 M 是适当大的正常数．

对每一点 $(t, x)\in\mathscr{R}$，

$$\mathscr{Q}^+(t, x)=\{(\eta, \xi)=(\eta, \xi^1, \xi^2):\eta>0,\ \xi^1=z,\ \xi^2=x^1z\}. \tag{8.4}$$

集合 $\mathscr{Q}^+(t, x)$ 显然是闭且凸的．此外，容易验证假设 4.1 中的其他条件以及假设 8.1 的(i)和(ii)都成立．而且 $\hat{f}=(f^0, f)$ 满足假设 8.1 的(iii)，只须取 $L(t, z)=|z|$ 和 $\omega(|x-x'|)=|x-x'|$．

设 $\mathscr{A}$ 是容许对 (ϕ, u) 的集合，集 $\mathscr{A}$ 非空．设 $\mathscr{A}_0$ 为 $\mathscr{A}$ 的任意子集，使得属于 $\mathscr{A}_0$ 的容许对 (ϕ, u) 的容许轨线 ϕ 是等度绝对连续的．从关系式 $\phi^{1\prime}(t)=u(t)$ 几乎处处成立，可见函数 u，从而函数 $|u|$ 具有等度绝对连续的积分．于是，因为积分区间是 $[0, 1]$，故存在常数 A 使得 $\int_0^1 |u|dt\leqslant A$ 对一切 $(\phi, u)\in\mathscr{A}_0$ 的 u 成立．因此，定理 8.1 的假设都满足．

另一方面，对所有 $\delta>0$，$\mathrm{cl\ co}\,\mathscr{Q}^+(N_{x\delta}(t, x))=\{(\eta, \xi):\eta\geqslant 0,\ \xi\in R^2\}$．由此和(8.4)式，我们看出弱 Cesari 性质不成立．

下面的定理 8.2 是关于问题 2 的与定理 4.2 相对应的存在性

定理，正如定理 4.2 可以从定理 4.1 得出那样，此定理也可由定理 8.1 得到.

定理 8.2 设假设 4.1 和 8.1 成立. 设 G 在 $\mathscr{A}_T$ 上是下半连续且下有界的，且存在极小化序列 $\{(\phi_k, u_k)\}$，使得 ϕ_k 是等度绝对连续的，并使(8.2)式成立. 则存在$(\phi^*, u^*)\in\mathscr{A}$ 使得 $\hat{J}(\phi^*, u^*)\leqslant\hat{J}(\phi, u)$ 对一切 $(\phi, u)\in\mathscr{A}$ 成立.

习题 8.1 证明定理 8.2.

从定理 8.2 可以得到定理 5.1 的稍为减弱的形式(按其假设稍为增强的意义下). 这将在下面习题里讨论.

习题 8.2 假如定理 5.1 的假设成立，再附加假设 Ω 与 x 无关，f^0 是连续的而不只是下半连续的. 证明定理 5.1 的结论可以从定理 8.2 得到.

非致密约束情形的存在定理在下面的定理中讨论，它的证明类似于定理 6.1，除了现在要用定理 8.2 代替原来证明中的定理 4.2 以外.

定理 8.3 设除了 $\mathscr{Q}^+(t, x)$ 满足弱 Cesari 性质之外，定理 6.1 的假设成立. 又假设 8.1 成立，且极小化序列中的控制 u_k 满足

$$\int_{t_{0k}}^{t_{1k}} L(t, u_k(t))\,dt\leqslant A,$$

其中 A 是常数. 则存在最优对 $(\phi^*, u^*)\in\mathscr{A}$.

我们现在证明定理 8.1. 证明过程直至包括 (7.9) 中 λ_j 的定义为止的部分，都与定理 4.1 的证明完全一样. 用于证明第三步的论证的剩下部分是不相同的. 读者务必十分留心各种子序列选取的顺序.

对应于 ψ_j 和 λ_j，定义函数列 σ_j 和 θ_j 如下

$$\begin{aligned}\sigma_j(t)&=\sum_{i=1}^{k}\alpha_{ij}f(t,\ \phi(t),\ u_{n_j+i}(t)),\\ \theta_j(t)&=\sum_{i=1}^{k}\alpha_{ij}f^0(t,\ \phi(t),\ u_{n_j+i}(t));\end{aligned}\tag{8.5}$$

其中如果 $t\notin[t_{0q}, t_{1q}]$，我们令 $\hat{f}(t, \phi(t), u_q(t))=0$，并且对每一整数 j，数 α_{ij}、下标 n_j+i 以及函数 u_{n_j+i} 与(7.7)中的一样. 函数

σ_j 和 θ_j 是可测的。

设 $M_k=\max\{|\tilde{\phi}_k(t)-\tilde{\phi}(t)|:t\in\mathscr{I}\}$，因为 $\tilde{\phi}_k$ 在 $\mathscr{I}$ 上一致收敛于 $\tilde{\phi}$，$M_k\to 0$ 当 $k\to\infty$。令 $f_q^*(t)=f(t,\ \phi(t),\ u_q(t))$，而 $f_q=f(t,\ \phi_q,\ u_q)$。注意，当 $t\notin[t_{0q},\ t_{1q}]$ 时，$f_q^*(t)=f_q(t)=0$。因为 σ_j 和 θ_j 是可测的，利用(8.5)、(8.1)和(8.2)，我们得到

$$\begin{aligned}\int_a^b|\sigma_j-\psi_j|\,dt&\leqslant\sum_{i=1}^k\alpha_{ij}\int_a^b|f^*_{n_j+i}-f_{n_j+i}|\,dt\\&=\sum_{i=1}^k\alpha_{ij}\int_{t_{0,n_j+i}}^{t_{1,n_j+i}}|f^*_{n_j+i}-f_{n_j+i}|\,dt\\&\leqslant\sum_{i=1}^k\alpha_{ij}\omega(M_{n_j+i})\int_{t_{0,n_j+i}}^{t_{1,n_j+i}}L(t,\ u_{n_j+i}(t))\,dt\\&\leqslant A\sum_{i=1}^k\alpha_{ij}\omega(M_{n_j+i}).\end{aligned}$$

因此 $\sigma_j-\psi_j\in L_1[\mathscr{I}]$。因为 $M_k\to 0$，$\omega(\delta)\to 0$，当 $\delta\to 0$ 时，我们得到在 $L_1[\mathscr{I}]$ 中 $\sigma_j-\psi_j\to 0$。用类似方法可证明 $\theta_j-\lambda_j\in L_1[\mathscr{I}]$，且在 $L_1[\mathscr{I}]$ 中 $\theta_j-\lambda_j\to 0$。于是存在子叙列使得

$$\sigma_j(t)-\psi_j(t)\to 0,\ \theta_j(t)-\lambda_j(t)\to 0 \tag{8.6}$$

几乎处处成立。以下我们取 (8.5)，(7.7) 和 (7.9) 中的函数为这些子序列中的函数。

现在我们规定 $\lambda(t)$ 如(7.10)，并且如 (7.10) 下面一段所论证的，$\lambda(t)\in L_1[\mathscr{I}]$，且(7.4)成立。

象定理 4.1 的证明中那样，令 T' 表示 $[t_0,\ t_1]$ 中的点集，在其上 $\lambda(t)$ 是有限的，$\psi_j(t)\to\phi'(t)$，且对一切 k，$u_k(t)\in\Omega(t)$。回顾 Ω 仅依赖于 t，此集合的测度为 t_1-t_0。设 T 为 T' 中使得(8.6)成立的点集，则 $\operatorname{meas}T=t_1-t_0$。

设 t 为 T 中任意固定的点。因为 $\psi_j(t)\to\phi'(t)$，由(8.6)可见 $\sigma_j(t)\to\phi'(t)$。从 λ 的定义可见存在子序列 $\{\lambda_j(t)\}$，它一般与 t 有关，使得 $\lambda_j(t)\to\lambda(t)$。从(8.6)，我们得 $\theta_j\to\lambda(t)$。对于相应的子序列我们仍然有 $\sigma_j(t)\to\phi'(t)$。因为 $T\subset T'$ 且 Ω 与 x 无关，可见对一切 j 和 i，$u_{n_j+i}(t)\in\Omega(t)$。因此

$$\hat{f}(t,\ \phi(t),\ u_{n_j+i}(t))\in\mathscr{Q}^+(t,\ \phi(t)).$$

因为 $\mathscr{Q}^+(t, x)$是凸的，故点$(\theta_j(t), \sigma_j(t))$属于 $\mathscr{Q}^+(t, \phi(t))$。因为 $\mathscr{Q}^+(t, \phi(t))$是闭的，且 $(\theta_j(t), \sigma_j(t))\to(\lambda(t), \phi'(t))$，我们得到$(\lambda(t), \phi'(t))\in\mathscr{Q}^+(t, \phi(t))$. 由于 t 是 T 中的任意点，我们有$(\lambda(t), \phi'(t))\in\mathscr{Q}^+(t, \phi(t))$在$[t_0, t_1]$上几乎处处成立.

现在，剩下的证明就与定理 4.1 的证明完全一样了.

习题 8.3 对"线性对象二次型积分代价准则"问题应用定理 8.2 以得习题 6.5 中的存在定理.

我们现在讨论第二个存在定理. 下面的定理可取代定理 4.1.

定理 8.4 设假设 4.1 中除了条件 (vi) 外都成立. 又设 $\hat{f}=(f^0, f)$ 是连续的；且集合 $\Omega(t, x)$ 与 x 无关，亦即对给定的 t，$\Omega(t, x)=\Omega(t, x')=\Omega(t)$对所有使得$(t, x)$ 和 (t, x') 属于 $\mathscr{R}$ 的 x 和 x' 成立. 设 $\mathscr{A}_0$ 是容许对 (ϕ, u) 的集合，使得 $\mathscr{A}_0$ 中的所有轨线 ϕ 是等度绝对连续的. $\sup\{I(\phi, u):(\phi, u)\in\mathscr{A}_0\}$ 是有限的，并且使得属于 $\mathscr{A}_0$ 的所有容许控制 u 满足

$$\|u\|_p\leqslant M, \tag{8.7}$$

其中 $1\leqslant p\leqslant\infty$，$M>0$ 是固定的，而 $\|u\|_p$ 表示 $L_p[t_0, t_1]$ 中 u 的模. 则存在 $\mathscr{A}_0$ 中的序列 $\{(\phi_k, u_k)\}$ 和容许对 $(\phi^*, u^*)\in\mathscr{A}$，具有如下性质：(i) 在 $\mathscr{X}_0$ 中，$\phi_k\to\phi^*$；(ii) $\lim I(\phi_k, u_k)\geqslant I(\phi^*, u^*)$.

本定理的证明将在定理 8.6 的叙述之后给出.

附注 8.1 在例 8.1 中定理 8.4 的假设是满足的.

下面的定理 8.5 是相应于定理 4.2 和 8.2 的关于问题 2 的存在定理. 并且按照从定理 4.1 和 8.1 分别得到定理 4.2 和 8.2 一样的方法，定理 8.5 可以从定理 8.4 得出.

定理 8.5 设假设 4.1 除了条件 (vi) 以外各条均成立. 又设 $\hat{f}$ 和 Ω 如定理 8.4 所述. 设 G 在 $\mathscr{A}_T$ 上是下半连续的且下有界的. 若存在极小化序列 $\{(\phi_k, u_k)\}$ 使得所有 ϕ_k 是等度绝对连续的且使(8.7)成立. 则存在$(\phi^*, u^*)\in\mathscr{A}$，使得

$$\hat{J}(\phi^*, u^*)\leqslant\hat{J}(\phi, u)$$

对一切$(\phi, u)\in\mathscr{A}$ 成立.

相应于定理 6.1 和 8.3 的定理是如下的定理 8.6. 其证明类似于定理 6.1, 除了现在用定理 8.5 代替定理 4.2 之外.

定理 8.6 设除了 $\mathscr{Q}^+(t, x)$ 满足弱 Cesari 性质和假设 4.1 的条件 (vi) 之外, 定理 6.1 的假设成立. 又设 $\hat{f}$ 和 Ω 如定理 8.4 所述. 设极小化序列中的控制 u_k 满足 (8.7). 则存在最优对 $(\phi^*, u^*)\in\mathscr{A}$.

我们现在证明定理 8.4. 下面引理是证明的关键.

引理 8.1 设 h: $(t, \xi)\to h(t, \xi)$ 是从 $[\alpha, \beta]\times E^r$ 到 E^1 的连续映射. $\{v_k\}$ 和 $\{w_k\}$ 是 $L_p[\alpha, \beta]$ 中的叙列, $1\leqslant p\leqslant\infty$, 使得 $\|v_k\|_p\leqslant M$, $\|w_k\|_p\leqslant M$, 对某一 $M>0$ 成立, 并使得在 $[\alpha, \beta]$ 上依测度有 $(v_k-w_k)\to 0$. 则在 $[\alpha, \beta]$ 上依测度

$$h(t, v_k(t))-h(t, w_k(t))\to 0.$$

我们推迟引理 8.1 的证明而继续概要地叙述定理 8.5 的证明. 其证明从第一步到第二步如定理 4.1 的证明那样进行. 而第三步修改如下.

设 p 如(8.7)中所示. 因为 $\tilde{\phi}_k\to\tilde{\phi}$ 在 $[a, b]$ 上一致成立. 且所有轨线位于固定的致密集中, 可见存在常数 M' 使得 $\|\tilde{\phi}_k\|_p\leqslant M'$, $\|\tilde{\phi}\|_p\leqslant M'$, 其中 $\|\cdot\|_p$ 表示 $L_p[a, b]$ 上的模. 设 $\tilde{u}_k$ 由 $t\notin[t_{0k}, t_{1k}]$ 时令 $\tilde{u}_k(t)=0$ 所得到的 u_k 的从 $[t_{0k}, t_{1k}]$ 到 $[a, b]$ 的扩展. 因为由(8.7), 所有 u_k 位于 $L_p[t_{0k}, t_{1k}]$ 中半径为 M 的球内. 我们得到存在常数 $A>0$, 使得对一切 k, 函数 $v_k=(\tilde{\phi}_k, \tilde{u}_k)$ 和 $w_k=(\tilde{\phi}, u_k)$ 位于 $L_p[a, b]$ 中半径为 A 的球中. 还注意到, 在 $[a, b]$ 的所有点上 $v_k(t)-w_k(t)\to 0$.

令

$$\hat{\Delta}_k(t)=\hat{f}(t, \phi(t), u_k(t))-\hat{f}(t, \phi_k(t), u_k(t)), \qquad (8.8)$$

其中, 如果 $t\notin[t_{0k}, t_{1k}]$, 令 $\hat{\Delta}_k=0$. 那么按引理 8.1 的结果(取 $\xi=(x, z)$)以及 $t_{ik}\to t_i$, $i=0, 1$, 可知在 $[a, b]$ 上依测度 $\hat{\Delta}_k\to 0$. 因为依测度 $\hat{\Delta}_k\to 0$, 故存在子叙列 $\{(\phi_k, u_k)\}$ 使得

$$\hat{\Delta}_k(t)\to 0 \qquad (8.9)$$

在 $[a, b]$ 中几乎处处成立.

其次，定义函数 ψ_j，λ_j 和 λ 如定理 4.1 的证明第四步中那样，且已证过(7.4)满足．然后，如(8.5)那样，定义序列 σ_j 和 θ_j．如果用 $\Delta_k^0(t)$ 表示向量 $\hat{\Delta}_k(t)$ 的第一个支量，而剩下的 n 个支量组成的向量用 $\Delta_k(t)$表示．利用(8.5)，(8.6)，(7.7)和(7.9)我们得到

$$\sigma_j(t)-\psi_j(t)=\sum_{i=1}^{k}\alpha_{ij}\Delta_{n_j+i}(t),$$

$$\theta_j(t)-\lambda_j(t)=\sum_{i=1}^{k}\alpha_{ij}\Delta_{n_j+i}^0(t).$$

那么，从(8.9)和习题 7.1 可知(8.6)成立。

剩下的证明是定理 8.1 的证明的最后四段的逐字重复。

现在我们证明引理 8.1．必须证明对任意给定 $\eta>0$ 和 $\varepsilon>0$，存在一个整数 N，使得若 $n>N$ 时，则

$$\operatorname{meas}\{t:|f(t,\ v_n(t))-f(t,\ w_n(t))|\geqslant\eta\}<\varepsilon. \tag{8.10}$$

设

$$A=M(2/\varepsilon)^{1/p}, \tag{8.11}$$

其中，我们把 $1/\infty$ 当零解释．令 G_A 表示 E^r 中使得 $|\xi|\leqslant A$ 的点 ξ 的集合．因为 f 在 $[\alpha,\ \beta]\times G_A$ 上是一致连续的，可见存在 $\delta>0$ 使得若 $|\xi-\xi'|<\delta$，且 ξ 和 ξ' 属于 G_A，则

$$|f(t,\ \xi)-f(t,\ \xi')|<\eta \tag{8.12}$$

对所有 $t\in[\alpha,\ \beta]$ 成立．

设 I_n 表示在 $[\alpha,\ \beta]$ 中的点 t 的集合，在其上或者使 $|v_n(t)|>A$ 或者 $|w_n(t)|>A$．令

$$G_n=\{t:|v_n(t)-w_n(t)|\geqslant\delta\},$$

从(8.12)我们得知，当 $t\notin I_n\cup G_n$ 时，有

$$|f(t,\ v_n(t)-f(t,\ w_n(t)|<\eta.$$

因此要建立引理 8.1，必须证明对充分大的 n，$\operatorname{meas}\{I_n\cup G_n\}<\varepsilon$．

对 $p<\infty$，我们有

$$M\geqslant\left(\int_\alpha^\beta|v_n|^p dt\right)^{1/p}>\left(\int_{I_n}A^p dt\right)^{1/p}=A(\operatorname{meas}I_n)^{1/p}.$$

由此式和(8.11)式可得 $\operatorname{meas}I_n<\varepsilon/2$．当 $p=\infty$ 时，从(8.11)式我们有 $\operatorname{meas}I_n=0$．因为 v_n-w_n 依测度收敛于零，存在整数 N

使得当 $n>N$ 时，$\text{meas}\, G_n<\varepsilon/2$. 因此 $\text{meas}(I_n\cup G_n)<\varepsilon$. 引理 8.1 证毕.

习题 8.4 证明在定理 8.4 中可以用习题 7.4 所做的较弱的假设来代替 $\hat{f}$ 连续的假设.

习题 8.5 用定理 8.6 推出 Nagumo–Tonelli 定理，即定理 6.2.

习题 8.6 对"线性对象凸积分代价准则和无界控制问题"，用定理 8.4~8.6 得出习题 6.3 所给的存在定理. 对习题 6.5 和 6.6 中所论的"线性对象二次型积分代价准则问题"做同样的证明.

习题 8.7 在某些应用中，对线性对象二次型积分准则问题中出现形如

$$\int_a^b |u|^2 dt<M \tag{8.13}$$

的等周约束，其中 M 是正常数. 此种等周约束起源于能量利用的限制. 请推出习题 6.5 中所提出的问题的存在性定理，这时对它再附加约束条件(8.13)式.

9. 极小化序列中控制的性状

在各种存在性定理的叙述和证明中没有论述或引出有关属于极小化序列中控制的特点的结论. 本节中我们给出一个例子，它表明对于极小化序列中的轨线列在 $\mathscr{X}_\rho$ 中是可能收敛于最优轨线的，而没有相应的控制子序列按任何通常意义下收敛于最优控制.

例 9.1 设状态方程组为

$$\begin{aligned}
\frac{dx^1}{dt}&=p^1(t)v^1(t)+p^2(t)v^3(t),\\
\frac{dx^2}{dt}&=p^1(t)v^2(t)+p^2(t)v^4(t),\\
\frac{dx^3}{dt}&=1.
\end{aligned} \tag{9.1}$$

设 $f^0(t,x,z)=(x^1)^2+(x^2)^2$, $\mathscr{T}_0=\{(t_0,x_0):x_0=0,\ t_0=0\}$, $\mathscr{T}_1$

$=\{(t_1, x_1): t_1=x_1^3=1,\ x_1^1=x_1^2=0\}$. 控制函数是 $u=(v^1, v^2, v^3, v^4, p^1, p^2)$. 控制约束是

$$
\begin{gathered}
(v^1(t))^2+(v^2(t))^2=1,\ (v^3(t))^2+(v^4(t))^2=1, \\
p^1(t)\geqslant 0,\ p^2(t)\geqslant 0,\ p^1(t)+p^2(t)=1.
\end{gathered}
\tag{9.2}
$$

问题是使 $$J(\phi, u)=\int_0^1[(\phi^1(t))^2+(\phi^2(t))^2]dt$$

取极小值.

容易验证定理5.1的一切假设都满足. 特别, 为了看出集合 $\mathscr{Q}^+(t, x)$是凸的, 注意对固定的 t, 形如$(y^1(t), y^2(t))$的向量集合可以写成

$$(y^1(t), y^2(t))=p^1(t)(v^1(t), v^2(t))+p^2(t)(v^3(t), v^4(t)),$$

其中 $y^1(t)$用(9.1)第一方程右边的式子表示, $y^2(t)$用(9.1)第二方程右边的式子表示, 且满足条件(9.2)式. 因此集合$(y^1(t), y^2(t))$是单位圆的凸壳, 它是单位圆盘. 由此可见 $\mathscr{Q}^+(t, x)$ 的凸性是显然的.

设 $\phi^*(t)=(0, 0, t)$, $0\leqslant t\leqslant 1$. θ 为 $[0, 2\pi]$ 中固定的实数, 而 u_θ^* 为如下规定的控制:

$$
\begin{gathered}
(v^1(t),\ v^2(t))=(\cos\theta,\ \sin\theta), \\
(v^3(t),\ v^4(t))=(-\cos\theta,\ -\sin\theta), \\
p^1(t)=p^2(t)=1/2.
\end{gathered}
$$

则(ϕ^*, u_θ^*)是容许对且 $J(\phi^*, u_\theta^*)=0$. 因为对所有容许对(ϕ, u), 有 $J(\phi, u)\geqslant 0$, 可见(ϕ^*, u_θ^*)是最优的. 注意, 相应于 $[0, 2\pi]$ 中的每一个 θ 值, 存在一个最优对, 故存在无穷多个最优对(ϕ^*, u_θ^*). 此外, 所有最优对有同样的轨线. 实际上, 从 J 的形式和端点条件, 显然 ϕ^* 是唯一可能的最优轨线. 因此最优轨线是唯一的, 而最优控制却不是唯一的.

现在我们来构造一个极小化叙列. 令

$$
\begin{gathered}
v_k^1(t)=v_k^3(t)=\cos 2\pi kt, \\
v_k^2(t)=v_k^4(t)=\sin 2\pi kt, \\
p_k^1(t)=p_k^2(t)=1/2,\ k=1,\ 2,\ 3,\ \cdots,
\end{gathered}
$$

则 (9.1) 变成

$$\frac{dx^1}{dt}=\cos 2\pi kt,$$

$$\frac{dx^2}{dt}=\sin 2\pi kt,$$

$$\frac{dx^3}{dt}=1.$$

如例 2.3 中我们所看到的，相应的容许对序列 $\{(\phi_k, u_k)\}$ 是极小化序列，且在[0, 1]上 ϕ_k 一致地收敛于 ϕ^*.

设 f 为 $L_1[0, 1]$ 中的任一函数，则由 Riemann-Lebesgue 引理，有

$$\int_0^1 f(t)\sin 2k\pi t\,dt\to 0,\quad \int_0^1 f(t)\cos 2\pi kt\,dt\to 0. \qquad (9.3)$$

设对 $[0, 2\pi)$ 中某一 θ, $v_k=(v_k^1, \cdots, v_k^4)$ 的子叙列或按 $L_p(1\leqslant p\leqslant\infty)$, 或依测度、或几乎处处收敛于函数

$$\tilde{u}_\theta^*=(\cos\theta,\ \sin\theta,\ -\cos\theta,\ -\sin\theta).$$

注意，$\tilde{u}_\theta^*$ 是由 u_θ^* 的前四个支量而得到的．因为函数列 v_k 是一致有界的，在所有的情形中，我们有

$$\int_0^1 v_k\,dt\to\int_0^1 \tilde{u}_\theta^* dt=(\cos\theta,\ \sin\theta,\ -\cos\theta,\ -\sin\theta).$$

因为对所有 θ, $\tilde{u}_\theta^*\neq 0$, 在(9.3)中取 $f=1$, 我们得到矛盾．

若在(9.3)式中取 f 为 [0, 1] 上的有界可测范围内的函数，则序列 $\{v_k\}$ 依 $L_1[0, 1]$ 的弱拓扑收敛于零．因此，没有 $\{v_k\}$ 的子叙列能按此拓扑收敛于 $\tilde{u}_\theta^*$. 同样，若取 f 为 $L_q[0, 1]$ 中的所有函数 $(1<q<\infty)$, 我们得知没有 $\{v_k\}$ 的子叙列能按 $L_q[0, 1]$ 中的弱拓扑收敛于 $\tilde{u}_\theta^*$. 最后我们注意到若让 f 取遍 $L_1[0, 1]$ 中所有函数，我们得知没有 $\{v_k\}$ 的子叙列能按由 $L_1[0, 1]$ 诱导的 $L_\infty[0, 1]$ 上的弱拓扑收敛于 $\tilde{u}_\theta^*$.

10．定理 7.1 的证明

我们首先假设 D 是开区间 $(0, \infty)$ 的闭子集 L 的情形下来证明定理 7.1.

对于每一个正整数 k，我们可以将 Φ 的值域 $\Phi(L)$ 分成至多可数个集合 $B_j^k(j=1, 2, \cdots)$ 的并集如下．令

$$A_j^k=\Phi(L\cap[0, j2^{-k}]),\qquad j=1, 2, \cdots,$$

而

$$B_j^k=A_j^k-A_{j-1}^k.$$

有些 B_j^k 集可以是空的．每个非空集 B_j^k 是由 $\Phi(L)$ 中那些属于 $L\cap[(j-1)2^{-k}, j2^{-k}]$ 的点的象点，但不是 $L\cap[0, (j-1)2^{-k}]$ 中之点的象点所组成的．因此非空的集合 B_j^k 是不相交的，并且它们的并集就是整个 $\Phi(L)$．注意，集合 A_j^k 是致密的，而且

$$B_j^k=B_{2j}^{k+1}\cup B_{2j-1}^{k+1}. \tag{10.1}$$

对每一个正整数 k，我们定义从 T 到 L 的函数 μ_k 如下．因为 $\Gamma(T)\subseteq\Phi(L)$，可见对 T 中的每一个 t，存在唯一整数 j，使得 $\Gamma(t)\in B_j^k$．设

$$\mu_k(t)=\inf\Phi^{-1}(B_j^k),$$

其中 B_j^k 是包含 $\Gamma(t)$ 的 $\Phi(L)$ 中的集合．函数 μ_k 是可测的，因为按假设它至多是可数个值的集合，$\inf\Phi^{-1}(B_j^k)$，$j=1, 2, \cdots$，和每一个集 $\mu_k^{-1}(\inf\Phi^{-1}(B_j^k))$ 按下面的论证也是可测的．对每一个 j，我们有 $\mu_k^{-1}(\inf\Phi^{-1}(B_j^k))=\Gamma^{-1}(B_j^k)$，但是

$$\Gamma^{-1}(B_j^k)=\Gamma^{-1}(A_j^k-A_{j-1}^k)=\Gamma^{-1}(A_j^k)-\Gamma^{-1}(A_{j-1}^k).$$

集合 A_j^k 和 A_{j-1}^k 是致密的．因此，由于 Γ 是可测的，故集合 $\Gamma^{-1}(A_j^k)$ 和 $\Gamma^{-1}(A_{j-1}^k)$ 是可测的．从而它们的差集也是可测的．

我们现在证明从 T 到 L 的可测映射序列 $\{\mu_k\}$ 是递增叙列．如果 t 是使得 $\Gamma(t)\in B_j^k$，则由(10.1)可见，对于 $i=2j$ 或 $i=2j-1$，$\Gamma(t)\in B_i^{k+1}$．因此 $B_i^{k+1}\subseteq B_j^k$．从而

$$\inf\Phi^{-1}(B_i^{k+1})\geqslant\inf\Phi^{-1}(B_j^k).$$

于是 $\mu_{k+1}(t)\geqslant\mu_k(t)$．

因为 t 使得 $\Gamma(t)\in B_j^k$，由此又可得对于所有 $q\geqslant k$，有

$$\mu_q\leqslant j2^k. \tag{10.2}$$

特别，如果 $\Gamma(t)\in B_j^0$，我们对于所有 k，有 $\mu_k(t)\leqslant j$．因此叙列 $\{\mu_k(t)\}$ 是有上界的，并且函数 $\mu_k(t)$ 收敛于可测函数 μ．因为 L 是闭的，且数 $\mu_k(t)$ 在 L 中，故 μ 的值域包含在 L 中．

我们现在证明对所有 T 中的 t, $\Phi(\mu(t))=\Gamma(t)$. 假如此命题不真, 则由于 Z 是 Hausdorff 空间, 存在 T 中的 t 和 Z 中的开集 $\mathscr{O}$ 使得 $\Phi(\mu(t))\in\mathscr{O}$ 而 $\Gamma(t)\notin\mathscr{O}$. 因为 Φ 是连续的, 故存在 $\mu(t)$ 在 L 中的邻域包含在 $\Phi^{-1}(\mathscr{O})$ 中. 因此存在整数 k 和 j_0 使得

$$\begin{aligned}&\mu(t)\in L\cap((j_0-1)2^{-k},\ j_02^{-k}),\\&L\cap[(j_0-1)2^{-k},\ j_02^{-k}]\subseteq\Phi^{-1}(\mathscr{O}).\end{aligned}\tag{10.3}$$

因为存在整数 j 使得 $\Gamma(t)\in B_j^k$. 如果 $j<j_0$, 则从 (10.2) 我们得到对一切 $q\geqslant k$, $\mu_q(t)\leqslant j2^{-k}\leqslant(j_0-1)2^{-k}$. 因此

$$\mu(t)\leqslant(j_0-1)2^{-k}.$$

但按照(10.3)的第一式我们有 $\mu(t)>(j_0-1)2^{-k}$, 从而 $j\geqslant j_0$. 若 $>j_0$, 则 $\Gamma(t)$是$((j-1)2^{-k},\ j2^{-k}]$ 中的某一点 λ 在 Φ 之下的象 j点而不是 $[0,\ (j-1)2^{-k}]$ 中任何点的象点. 于是存在 $q\geqslant k$, 使得 $\mu_q(t)>j_02^{-k}\geqslant\mu(t)$. 但由于 $\mu_q(t)$是收敛于 $\mu(t)$的递增序列, 这是不可能的. 于是 $j=j_0$, 因此 $\Gamma(t)$是 $L\cap((j_0-1)2^{-k},\ j_02^{-k}]$ 中某一点在 Φ 之下的象点. 因此, 由(10.3)的第二式我们有

$$\Gamma(t)\in\Phi(L\cap[(j_0-1)2^{-k},\ j_02^{-k}])\subseteq\mathscr{O},$$

它与 $\Gamma(t)\notin\mathscr{O}$ 的假设相矛盾. 因此, 对 D 是 $(0,\ \infty)$ 的闭子集 L 的情况定理得证.

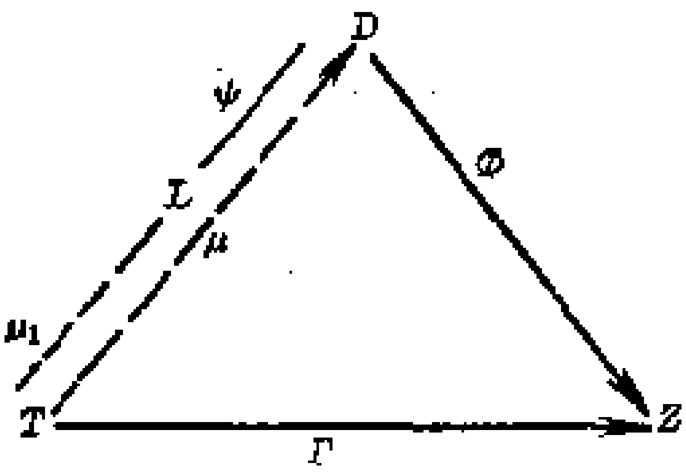

图 4

其次, 我们对 D 是$(0,\ \infty)$中闭集 L 的连续象的情形证明定理. 这时 $D=\psi(L)$, 其中 ψ 是把 L 映上 D 的连续映射. 我们所要论证的事实如图 4 所示.

映射 $\Phi*\psi$ 是从 L 到 Z 的连续映射, 其中$*$表示复合. 因为 $\Gamma(T)\subseteq\Phi(D)$, 且

$$\Phi(D)=\Phi(\psi(L)),$$

由此可见 $\Gamma(T)\subseteq(\Phi*\psi)(L)$, 因此, 由于我们已对 $D=L$ 的情况证明了定理, 故存在从 T 到 L 的可测函数 μ_1, 使得 $(\Phi*\psi)*\mu_1=\Gamma$. 令 $\mu=\psi*\mu_1$. 则 μ 是从 T 到 D 的映射, 使得 $\Phi*\mu=\Gamma$. 我们断定 μ 是可测的. 设 C 是 D 中的致密集, 则 $\psi^{-1}(C)$在 L 中是闭

的并且是可数多个集合 $L\cap[0,k]\cap\psi^{-1}(C)$，$k=1,2,\cdots$ 的并集．于是 $\mu_1^{-1}(\psi^{-1}(C))$ 是可测的，从而 $\mu^{-1}(C)$ 也是可测的．

现在设 $D=\bigcup_{i=1}^{\infty}C_i$，其中 C_i 是致密度量空间．按 Urysohn 定理，每一个致密度量空间是 Cantor 集的连续象．关于这一事实的证明读者可参阅（[29]定理 3.28, p. 127）．于是对每一个整数 k，存在包含在 $[2k-1, 2k]$ 中的 Cantor 集 L_k 的平移，以及从 L_k 到 C_k 上的连续象 ψ_k．注意，每个 L_k 是闭的．令 $L=\bigcup L_k$，而如果 $t\in L_k(k=1,2,\cdots)$，ψ 是由公式 $\psi(t)=\psi_k(t)$ 定义的从 L 到 D 的映射．则 L 是 $(0,\infty)$ 的闭子集，而函数 ψ 在 L 上是连续的，并且 $\psi(L)=D$．因此前一种情况的假设都满足，从而存在可测函数 μ 使得 $\Phi_*\mu=\Gamma$．

11．关于状态是线性的系统中控制的存在性

我们考虑状态方程为

$$\frac{dx}{dt}=A(t)x+h(t,u(t)) \tag{11.1}$$

的系统，在对系统作适当假设下将利用定理 7.1 证明控制的存在．

引理 11.1　设 $\mathscr{I}$ 是 E^1 中的固定闭区间，$\Omega: t\to\Omega(t)$ 是从 $\mathscr{I}$ 到 E^m 的闭子集的映射，它在 $\mathscr{I}$ 上是上半连续的．则存在可测函数 $u:\mathscr{I}\to E^m$ 使得 $u(t)\in\Omega(t)$ 几乎处处满足．

证明：令

$$\Delta=\{(t,z): t\in\mathscr{I},\ z\in\Omega(t)\}. \tag{11.2}$$

因为 Ω 是上半连续的，从引理 3.1 可见 Δ 是闭的．因此集合可以写成可数个致密集 $\Delta_i=\Delta\cap B_i$ 的并集，其中 B_i 是 E^{m+1} 中半径为 i 的闭球．

设 $Z=\mathscr{I}\times E^1$，而 $T=\mathscr{I}$，则 Z 是 Hausdorff 空间，而 T 是具有 Lebesgue 测度的测度空间．设 Φ 是用公式

$$\Phi(t,z)=(t,|z|)$$

定义的从 Δ 到 Z 的映射，则 Φ 显然是连续的．对每一个 $t\in\mathscr{I}$，令

$$d(t)=\inf\{|z|:z\in\Omega(t)\}.$$

设 Γ 是用公式 $\Gamma(t)=(t,\ d(t))$ 定义的从 T 到 Z 的映射. 因为对每一个 t 集合 $\Omega(t)$ 是闭的，故至少存在一点 $z\in\Omega(t)$，使得 $|z|=d(t)$. 因此 $\Gamma(T)\subseteq\Phi(\Delta)$[注].

其次，我们通过证明对任意实数 α，集合 $E_\alpha=\{t:d(t)\leqslant\alpha\}$ 是闭的来证明映射 $d(t)$ 在 $\mathscr{I}$ 上是下半连续的. 假如对某一 α，集 E_α 不是闭的，则将存在 E_α 中的点列 $\{t_n\}$ 收敛于 $t_0\in\mathscr{I}$，并使得 $d(t_0)=\beta$，其中 $\beta>\alpha$. 我们证明这将导致矛盾.

因为 Ω 是上半连续的，

$$\Omega(t_0)\supseteq\bigcap_{\delta>0}\mathrm{cl}\Big(\bigcup_{|t-t_0|<\delta}\Omega(t)\Big).\tag{11.3}$$

设 $\varepsilon=(\beta-\alpha)/2$. 因为 $d(t_n)\leqslant\alpha$，由此可见在每一个集合 $\Omega(t_n)$ 中存在 z_n 使得 $|z_n|<\alpha+\varepsilon$. 选取子序列 $\{z_n\}$ 使之收敛于 E^m 中某一点 z_0. 显然

$$|z_0|\leqslant\alpha+\varepsilon<\beta=d(t_0).\tag{11.4}$$

相应的序列 $\{t_n\}$ 仍然收敛于 t_0. 因此对每一个 $\delta>0$，存在 $n_0(\delta)$，使得当 $n>n_0(\delta)$，我们有 $|t_n-t|<\delta$. 从而对所有 $n>n_0(\delta)$，

$$z_n\in\bigcup_{|t-t_0|<\delta}\Omega(t).\tag{11.5}$$

因此，对每一个 $\delta>0$，z_0 属于 (11.5) 右边的闭包，所以 z_0 属于 (11.3) 的右边. 但按 (11.4)，$|z_0|<d(t_0)$，所以 z_0 不能属于 $\Omega(t_0)$. 然而，这与(11.3)矛盾.

因为 $d(t)$ 在 $\mathscr{I}$ 上是下半连续的，映射 Γ 当然是可测的. 因此定理 7.1 的假设对 $D=\Delta$ 的情形是满足的. 于是存在可测映射 $\mu:t\to(z(t),\ u(t))$，使得 $u(t)\in\Omega(t)$ 且

$$(\Phi*\mu)(t)=(\tau(t),\ |u(t)|)=\Gamma(t)=(t,\ d(t)).$$

于是 $\tau(t)=t$ 且 $u(t)\in\Omega(t)$. 映射 $u:t\to u(t)$ 就是所要求的映射. 注意

$$|u(t)|=\inf\{|z|:z\in\Omega(t)\}.$$

引理 11.2 设 $\mathscr{I}$ 是 E^1 中的闭区间，A 是 $\mathscr{I}$ 上的 $n\times n$ 连续

[注] 原书误为 $\Phi(D)$.——译者注.

矩阵. 设 $\mathscr{U}$ 是 E^m 中的区域, h 是从 $\mathscr{I}\times\mathscr{U}$ 到 E^n 的连续映射. 设 $\Omega: t\to\Omega(t)$ 是从 $\mathscr{I}$ 到 E^m 的子集的映射, 它在 $\mathscr{I}$ 上是上半连续的. 设存在定义于 $\mathscr{I}$ 上的非负可积函数 μ 使得

$$|h(t, z)|\leqslant\mu(t) \tag{11.6}$$

对所有 $(t, z)\in\Delta$ 成立, 其中 Δ 如 (11.2) 所示. 则存在控制 u, 在区间 $\mathscr{I}$ 上满足 $u(t)\in\Omega(t)$.

证明: 按照引理 11.1, 存在定义于 $\mathscr{I}$ 上的可测函数 u, 使得 $u(t)\in\Omega(t)$. 因为 h 是连续的, 故映射 $t\to h(t, u(t))$ 是可测的. 因此, 由(11.6)可见 $h(t, u(t))$是可积的. 因为所得的系统(11.1)关于 x 是线性的, 因此相应于上述的 $u(t)$, 它的解在整个区间 $\mathscr{I}$ 上存在. 这表明 $u(t)$是容许控制, 且 $u(t)\in\Omega(t)$. 证毕.

附注 11.1 若假定映象是 u. s. c. i 的, 而且约束集 $\Omega(t)$ 是致密的, 则由引理 5.2 可知集 Δ 是致密的, 且(11.6)式成立.

第四章　没有凸性的存在定理

1. 引　　言

上一章的存在定理要求问题的原始资料具有某种正规性，并且要求集合$\mathscr{Q}^+(t, x)$是凸的．在非致密约束的情形，还要求轨线族是等度绝对连续的，而保证这一点的合理的条件已叙述过了．加在问题上的所有条件，也许除了集合$\mathscr{Q}^+(t, x)$为凸这一要求外，可以证明是恰当的．集合$\mathscr{Q}^+(t, x)$为凸的要求，实质上是把我们限制于那样的系统，其状态方程关于控制是线性的，代价函数f°关于z是凸的，而且约束集$\Omega(t, x)$也是凸的．在本章中，我们考察集合$\mathscr{Q}^+(t, x)$不必为凸的系统．

第二节中，我们不要求集合$\mathscr{Q}^+(t, x)$为凸，而实质上是用对控制的变化率所加的限制来代替它．这是控制具有惯性的系统的数学理想化．所得的结果是，对这样的控制，集合$\mathscr{Q}^+(t, x)$为凸的假定对于存在性是不必要的．

第三节中，我们用满足凸性假设的"松弛问题"代替原始问题．原始问题的所有轨线都是松弛轨线．通过引入松弛问题来解决存在性问题的情况概述如下．虽然集合$\mathscr{Q}^+(t, x)$没有凸性妨碍我们断言原始问题有解，但我们能断定松弛问题有解．于是就希望松弛问题的解的进一步刻划会得到这样的结论：所说的松弛解实际上是原始问题的解．当然，事实并非必定如此，松弛问题可以有解而原始问题却没有解．可是，在第四节中我们要证明，在对系统加上一些合理的假设之后，原始问题的轨线在松弛问题的轨线集合内稠密．这样，松弛轨线可以用寻常轨线一致地逼近．另一方面，松弛代价泛函的下确界能严格地小于原始代价泛函的下确界．第四节的材料在第六章第二节至第七节证明最大值原理时还要用到．

某些作者是在可达集性质的基础上进行最优控制问题的研究的．我们把这个程序颠倒过来，在第五节应用定理 III 4.1 来推得可达集的基本性质．

在第六节中，我们研究 f^0 和 f 关于状态变量为线性，且约束集与状态变量无关的问题．我们证明，对这样的系统，集合 $\mathscr{Q}^+(t, x)$ 为凸的要求可以去掉．这里的基本事实是：松弛可达集和寻常可达集相等且二者都是致密的和凸的．这一事实也被用于得到线性系统的 bang-bang 原理．除了需要参考定义 5.2 和有关的讨论外，第六节可以单独阅读．

主要兴趣在于应用的读者可以首先只阅读第二节，定理 6.3、定理 6.4 的叙述以及其后的讨论，然而在阅读定理 6.4 时，要回过头去参考定义 5.2、3.2 以及有关的讨论．

2. 惯性控制器

在第三章中，控制 u 是可测函数．这样，控制 u 的变化速率就没有限制．在物理上，这意味着假定控制没有惯性．虽然这对电气控制是合适的模型，但对力学系统和经济系统就不那么合适了．所以我们考察有惯性的控制．具有分段连续的导数 u'，使 $|u'(t)| \leqslant k$（k 为常数）的连续函数 u，看来是惯性控制器的合适模型．在 u' 的不连续点，不等式 $|u'(t)| \leqslant k$ 成立的意思是对左右极限 $u'(t-0)$ 和 $u'(t+0)$ 都满足．这样我们就有了控制的变化速率的界限．但事实证明，刚才描述过的函数 u 的类型对于证明惯性控制器问题的存在定理来说，是限制太强了．为了证明这种存在定理，只须把惯性控制器的模型取为使得 $|u'(t)| \leqslant k$ 几乎处处满足的绝对连续函数 u，其中 k 是与 u 无关的常数．

现在叙述惯性控制器的最小化问题．

问题 3 在下列条件下：

$$dx/dt = f(t, x, u(t)),$$

$$u(t) \in \Omega(t, x),$$

$$(t_0, \phi(t_0), t_1, \phi(t_1)) \in \mathscr{B},$$

u 是 $[t_0, t_1]$ 上的绝对连续函数，

$|u'(t)| \leqslant K$ 在 $[t_0, t_1]$ 上几乎处处成立，

使泛函

$$J(\phi, u) = G(\phi) + \int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt$$

取极小值，其中 K 为预先给定的常数，G 是定义在 $\mathscr{X}_p$ 上的泛函。

问题 3 的容许对 (ϕ, u) 是满足定义 II 3.1 与 II 3.2 的条件，且使 $|u'(t)| \leqslant K$ 几乎处处成立的绝对连续函数 ϕ 和绝对连续函数 u。

定理 2.1 设问题 3 的容许对集合 $\mathscr{A}$ 非空，且下列假定成立：(i) 存在致密集 $\mathscr{R}_0 \subset \mathscr{R}$，使得对所有容许轨线 ϕ 和所有 $[t_0, t_1]$ 中的 t，点 $(t, \phi(t)) \in \mathscr{R}_0$. (ii) 集合 $\mathscr{B}$ 为闭. (iii) 映射 Ω 在 $\mathscr{R}_0$ 上是 u. s. c. i. 的. (iv) 对 $\mathscr{R}_0$ 中每一点 (t, x)，集合 $\Omega(t, x)$ 是致密的. (v) 函数 f^0 在 $\mathscr{G} = \mathscr{R} \times \mathscr{U}$ 上是下半连续的，而函数 f 在 $\mathscr{G}$ 上连续. 设 G 在 $\mathscr{A}_T$ 上是下半连续且有下界的. 则在 $\mathscr{A}$ 中存在问题 3 的最优对 (ϕ^*, u^*)。

注 2.1 注意，除了现在不要求集合 $\mathscr{Q}^+(t, x)$ 为凸外，本定理的假设与定理 III 5.1 的假设相同。

我们把问题 3 改写成形如问题 2 的等价问题，然后应用定理 III 5.1 来证明定理 2.1。

我们取 z 作为状态变量，并取导数 u' 作为控制，则系统方程变为

$$\begin{aligned} \frac{dx}{dt} &= f(t, x, z), \\ \frac{dz}{dt} &= v(t). \end{aligned} \tag{2.1}$$

(2.1)的解记作(ϕ, ψ)；这样，$\phi'(t) = f(t, \phi(t), \psi(t))$，$\psi'(t) = v(t)$ 几乎处处成立. 注意到由于 f、ϕ 与 ψ 是连续的，

$$\phi'(t) = f(t, \phi(t), \psi(t))$$

处处成立. 设 $\mathscr{C}^m = \{w \in E^m: |w| \leqslant K\}$. 令

$$\tilde{\mathscr{R}} = \{(t, x, z): (t, x) \in \mathscr{R}, z \in \Omega(t, x)\},$$

$$\widetilde{\mathscr{B}}=\{(t_0, x_0, z_0, t_1, x_1, z_1):(t_0, x_0, t_1, x_1)\in\mathscr{B},$$
$$z_i\in\Omega(t_i, x_i),\ i=0, 1\},$$
$$\widetilde{\Omega}(t, x, z)=\{w\in E^m: |w|\leqslant K\}=\mathscr{C}^m.$$

考虑如下定义的问题3'. 在满足方程(2.1), $(t, \phi(t), \psi(t))\in\widetilde{\mathscr{R}}$, $v(t)\in\mathscr{C}^m$ 几乎处处成立, 且 $(t_0, \phi(t_0), \psi(t_0), t_1, \phi(t_1), \psi(t_1))\in\widetilde{\mathscr{B}}$ 的条件下, 使泛函

$$\mathscr{J}(\phi, \psi, v)=G(\phi)+\int_{t_0}^{t_1}f^0(t, \phi(t), \psi(t))dt$$

取极小值. 问题3'的每一个容许对(ϕ, ψ, v), 对应着问题3的唯一容许对$(\phi, u)=(\phi, \psi)$, 使$J(\phi, u)=\mathscr{J}(\phi, \psi, v)$. 反之, 问题3的每一个容许对$(\phi, u)$, 对应着问题3'的具有$\psi(t)=u(t)$的容许对$(\phi, \psi, v)$, 且使$J(\phi, u)=\mathscr{J}(\phi, \psi, v)$. 因此, 问题3'等价于问题3. 在下一段我们将证明问题3'满足定理III 5.1的假设. 因此问题3'有解(ϕ^*, ψ^*, v^*), 从而问题3有解(ϕ^*, u^*), 其中$u^*(t)=\psi^*(t)$.

现在验证问题3'满足定理III 5.1的假设. 首先注意到, 因为u连续且Ω是u. s. c. i. 的, 所以$u(t)\in\Omega(t, \phi(t))$几乎处处满足蕴含$u(t)\in\Omega(t, \phi(t))$处处满足. 因此, 由于$\psi(t)=u(t)$, 故对所有$t$值, $\psi(t)\in\Omega(t, \phi(t))$. 因为问题3的所有轨线的图象都位于致密集$\mathscr{R}_0\subset\mathscr{R}$之内, 由此得出, 如果令$\widetilde{\mathscr{R}}_0=\mathscr{D}$, 这里$\mathscr{D}$由III(4.1)定义, 那么问题3'的一切轨线$(\phi, \psi)$的图象位于$\widetilde{\mathscr{R}}_0$之内. 易见, $\widetilde{\mathscr{R}}_0\subset\widetilde{\mathscr{R}}$. 此外, 因为$\Omega$是u. s. c. i. 的, 且每一个$\Omega(t, x)$致密, 故由引理III 5.2得知, 集合$\widetilde{\mathscr{R}}_0$是致密的. 因而问题3'的所有轨线都位于致密集$\widetilde{\mathscr{R}}_0\subset\widetilde{\mathscr{R}}$之内.

为了证明$\widetilde{\mathscr{B}}$为闭, 设$\{(t_{0n}, x_{0n}, z_{0n}, t_{1n}, x_{1n}, z_{1n})\}$是$\widetilde{\mathscr{B}}$中收敛于点$(t_0, x_0, z_0, t_1, x_1, z_1)$的点列. 由于$\mathscr{B}$为闭, 故$(t_0, x_0, t_1, x_1)\in\mathscr{B}$. 设给定$\varepsilon>0$, 由于$\Omega$是u. s. c. i., 故知对充分大的$n$,

$$z_{in}\in\Omega(t_{in}, x_{in})\subseteq[\Omega(t_i, x_i)]_\varepsilon,\qquad i=0, 1.$$

因此, 对任意的$\varepsilon>0$和$i=0, 1$, $z_i\in[\Omega(t_i, x_i)]_\varepsilon$. 因为每一集合$\Omega(t_i, x_i)$ $(i=0, 1)$是闭的, 故知$z_i\in\Omega(t_i, x_i)$, $i=0, 1$. 因而$\widetilde{\mathscr{B}}$

为闭．由于对 $\tilde{\mathscr{R}}_0$ 中每一点 (t, x, z) 有 $\tilde{\Omega}(t, x, z)=\mathscr{C}^m$，故得 $\tilde{\Omega}$ 是 u. s. c. i.，且每一集合 $\tilde{\Omega}(t, x, z)$ 是致密的．问题 3′ 的集合 $\tilde{\mathscr{Q}}^+(t, x, z)$ 由下式确定：

$$\tilde{\mathscr{Q}}^+(t, x, z)=\{(y^0, y, \eta): y^0 \geqslant f^0(t, x, z),$$
$$y=f(t, x, z),\ \eta=w,\ w\in\mathscr{C}^m\},$$

显然它是凸的．最后，函数 f^0，f 和 G 满足所需的连续性假设．

习题 2.1 不用定理 III 5.1 和定理 III 4.1，直接证明定理 2.1．

提示：若 $\{(\phi_k, u_k)\}$ 是最小序列，则函数 u_k 有等度绝对连续的积分．

3. 松弛问题

本节我们叙述对应于问题 2 的松弛问题，并证明松弛问题的存在定理．问题 2 将看作原始问题．如在引言中指出过的，松弛问题与原始问题有关，即使原始问题不满足第三章的凸性假设，它也满足凸性假设．因此，在适当的条件之下，我们可以对不能保证原始问题有解的情形，证明松弛问题有解．在某些情况下，我们能应用松弛问题解存在的知识，来证明原始问题有解．例如，当解是由第五章的必要条件刻划时，可以证明松弛解实际上是寻常解．应用松弛问题来得到原始问题的解存在的另一个例子，可以在下面的定理 6.2 中找到．当然，原始问题没有解时，松弛问题可以有解．

设 q 为正整数，S 为任一集合．记号 $[S]^q$ 表示 S 与其自身的 q 重笛卡儿乘积．以 $\tilde{z}$ 表示 $[E^m]^{n+2}$ 中的向量．这样，$\tilde{z}=(z_1, \cdots, z_{n+2})$，这里 $z_i=(z_i^1, \cdots, z_i^m)\in E^m$，$i=1, \cdots, n+2$．设

$$\Gamma=\{\pi: \pi=(\pi^1, \cdots, \pi^{n+2}),\ \pi^i\geqslant 0,\ \sum\pi^i=1\},$$
$$\tilde{\Omega}(t, x)=[\Omega(t, x)]^{n+2}\times\Gamma,$$
$$\tilde{f}^0(t, x, \tilde{z}, \pi)=\sum_{i=1}^{n+2}\pi^i f^0(t, x, z_i), \tag{3.1}$$
$$\tilde{f}(t, x, \tilde{z}, \pi)=\sum_{i=1}^{n+2}\pi^i f(t, x, z_i).$$

其中 z_i 为 $\mathscr{U}$ 中的向量．设 $\mathscr{B}$ 和 $\mathscr{R}$ 与第二章的问题 2 中的相同．如果 $u_1, \cdots, u_{n+2}$ 是定义在公共区间 $[t_0, t_1]$ 上的一组可测函数，而且每一个 u_i 在 E^m 中取值，则定义从 $[t_0, t_1]$ 到 $[E^m]^{n+2}$ 的可测映射 $\tilde{u}$ 为：$\tilde{u}=(u_1, \cdots, u_{n+2})$．

定义 3.1 定义在区间 $[t_0, t_1]$ 上的形如

$$v=(\tilde{u}, p)=(u_1, \cdots, u_{n+2}, p^1, \cdots, p^{n+2})$$

的可测函数 v 称为松弛控制，如果下述条件成立：对每一个 $i=1, \cdots, n+2$，函数 u_i 的值域包含在 $\mathscr{U}$ 中；对每一个 $i=1, \cdots, n+2$，函数 p^i 取实值，而且在 $[t_0, t_1]$ 上几乎处处有 $p^i(t)\geqslant 0$ 和 $\sum_{i=1}^{n+2} p^i(t)=1$．存在一个定义在 $[t_0, t_1]$ 上的绝对连续函数 $\psi=(\psi^1, \cdots, \psi^n)$，使得

（i）对所有 $t\in[t_0, t_1]$，$(t, \psi(t))\in\mathscr{R}$；

（ii）ψ 是微分方程组

$$\frac{dx}{dt}=\tilde{f}(t, x, \tilde{u}(t), p(t))=\sum_{i=1}^{n+2}p^i(t)f(t, x, u_i(t)) \qquad (3.2)$$

的解，即 $\psi'(t)=\tilde{f}(t, \psi(t), \tilde{u}(t), p(t))$ 在 $[t_0, t_1]$ 上几乎处处成立．函数 ψ 称为对应于 v 的松弛轨线．微分方程组 (3.2) 称为松弛状态方程组．

定义 3.2 松弛控制 v 称为容许松弛控制，如果存在对应于 v 的松弛轨线 ψ，使得：

（i）映射 $t\to\tilde{f}^0(t, \psi(t), \tilde{u}(t), p(t))=\sum_{i=1}^{n+2}p^i(t)f^0(t, \psi(t), u_i(t))$ 属于 $L_1[t_0, t_1]$；

（ii）$v(t)=(\tilde{u}(t), p(t))\in\tilde{\Omega}(t, \psi(t))$ 在 $[t_0, t_1]$ 上几乎处处满足；

（iii）$(t_0, \psi(t_0), t_1, \psi(t_1))\in\mathscr{B}$．

这时，轨线 ψ 称为容许松弛轨线，而对 (ψ, v) 称为容许松弛对．

现叙述对应于问题 2 的松弛问题．

问题 2R 以 $\mathscr{A}_R$ 表示所有容许松弛对组成的集合．设 $\mathscr{A}_R$ 非空．令

$$\tilde{J}(\psi,\ v)=G(\psi)+\int_{t_0}^{t_1}\tilde{f}^0(t,\ \psi(t),\ \tilde{u}(t),\ p(t))dt,$$

其中$(\psi,v)\in\mathscr{A}_R$，G是定义在$\mathscr{X}_\rho$上的泛函．设$\mathscr{A}_{1R}$是$\mathscr{A}_R$的非空子集．在$\mathscr{A}_{1R}$中求一个最优容许松弛对$(\psi^*,\ v^*)$，即在$\mathscr{A}_{1R}$中求出一个元素$(\psi^*,\ v^*)$，使得对$\mathscr{A}_{1R}$中的所有$(\psi,\ v)$，有

$$\tilde{J}(\psi^*,\ v^*)\leqslant\tilde{J}(\psi,\ v).$$

例 III 9.1 的控制问题是对应于例 III 2.3 的松弛问题．

设$(\phi,\ u)$是问题 2 的容许对．令$v=(u_1,\cdots,u_{n+2},1,0,\cdots,0)$，这里$u_i=u$，$i=1,\ \cdots,\ n+2$．那么，如果取$\psi=\phi$，对$(\psi,\ v)$就是一个松弛对．因而如果$\mathscr{A}$非空，则$\mathscr{A}_R$也非空．不严格地说，每一个容许对也是一个容许松弛对．下面的注现在是显而易见的了．

注 3.1 如果$\tilde{\mu}=\inf\{\tilde{J}(\psi,\ v):(\psi,v)\in\mathscr{A}_R\}$，$\mu=\inf\{J(\phi,\ u):(\phi,\ u)\in\mathscr{A}\}$，则$\tilde{\mu}\leqslant\mu$．

在习题 3.1 中，我们将考察一个原始问题与松弛问题二者都有解，而$\tilde{\mu}$严格小于μ的例子．在本章习题 4.1，我们将给出保证$\tilde{\mu}=\mu$的准则．

现在叙述本节的主要结果．我们将只考虑致密约束集$\Omega(t,x)$，而把非致密约束集的相应结果的叙述与证明留给读者．

定理 3.1 设存在致密集$\mathscr{R}_0\subset\mathscr{R}$，使所有容许松弛轨线都位于$\mathscr{R}_0$中．设原始问题关于本定理的$\mathscr{R}_0$满足定理 III 5.1 的假设 (ii)、(iii)、(iv)．设f^0与f满足定理 III 5.1 的假设 (vi)，而G和注 III 5.1 中的相同．则松弛问题在$\mathscr{A}_R$中有解．

验证松弛问题满足定理 III 5.1 的除(v)之外的所有假设是简单的事情．为了证明 (v) 成立，必须证明对$\mathscr{R}_0$中每一$(t,\ x)$，集合

$$\tilde{\mathscr{Q}}^+(t,\ x)=\{(y^0,\ y):y^0\geqslant\tilde{f}^0(t,\ x,\ \tilde{z},\ \pi),$$
$$y=\tilde{f}(t,\ x,\ \tilde{z},\ \pi),\ (\tilde{z},\ \pi)\in\tilde{\Omega}(t,\ x)\}$$

是凸的．设

$$\tilde{\mathscr{Q}}(t,\ x)=\{(y^0,\ y):y^0=\tilde{f}^0(t,\ x,\ \tilde{z},\ \pi),$$

$$y=\tilde{f}(t,\ x,\ \tilde{z},\ \pi),\ \tilde{z}\in[\Omega(t,\ x)]^{n+2},\ \pi\in\Gamma\};$$

而
$$\mathscr{Q}(t,\ x)=\{(y^0,\ y):y^0=f^0(t,\ x,\ z),$$
$$y=f(t,\ x,\ z),\ z\in\Omega(t,\ x)\}.$$

则 $\tilde{\mathscr{Q}}(t,\ x)\subset\mathrm{co}\,\mathscr{Q}(t,\ x)$. 根据 Caratheodory 定理(定理 III 5.2), $\mathrm{co}\,\mathscr{Q}(t,\ x)$ 中的每一点都可以写成 $\mathscr{Q}(t,\ x)$ 中至多 $(n+2)$ 个点的凸组合. 由 $\tilde{\mathscr{Q}}(t,\ x)$ 的定义和(3.1)可以看出 $\mathrm{co}\,\mathscr{Q}(t,\ x)\subseteq\tilde{\mathscr{Q}}(t,\ x)$. 因此, $\tilde{\mathscr{Q}}(t,\ x)=\mathrm{co}\,\mathscr{Q}(t,\ x)$, 从而 $\tilde{\mathscr{Q}}(t,\ x)$ 是凸的. 而如果 $\tilde{\mathscr{Q}}(t,\ x)$ 是凸的, 则 $\tilde{\mathscr{Q}}^+(t,\ x)$ 也是凸的. 所以对松弛问题, 定理 III 5.1 的(v)成立. 这样, 对松弛问题, 定理 III 5.1 的所有假设都成立, 因而松弛问题有解.

习题 3.1　考虑具有状态方程组

$$dx^1/dt=(x^2)^2-(u(t))^2,$$
$$dx^2/dt=u(t),$$
$$dx^3/dt=(x^2)^4$$

的问题, 其中 u 是满足约束 $|u(t)|\leqslant 1$ 的实值函数. 设 $\mathscr{B}=\{(t_0,\ x_0;\ t_1,\ x_1):t_0=0;\ x_0=0,\ t_1=1,\ x_1^3=0\}$, $G(\phi)=\phi^1(1)$, $f^0\equiv 0$. 证明定理 III 5.1 的凸性假设不满足, 但仍有最优对存在. 对松弛问题找一个最优对, 并证明 $\tilde{\mu}<\mu$. 求一寻常轨线序列 ϕ_k 和相应的控制 u_k, 使得除端点条件外, 所有的约束都满足, 且使 $G(\phi_k)\to\tilde{\mu}$ 和 $\phi_k\to\psi^*$ 在 $[0,\ 1]$ 上一致地成立, 这里 ψ^* 是最优容许松弛轨线. 注意, 对原始问题的容许对 $(\phi_k,\ u_k)$ 不可能做到这一点.

习题 3.2　假设除 f^0 取为连续外, 定理 3.1 的假设成立; 则不失一般性可假定 $f^0\equiv 0$. 在这些假定之下, 下述松弛轨线的定义等价于正文中给出的定义. 松弛轨线是使

$$\psi'(t)\in\mathrm{co}\,f(t,\ \psi(t),\ \Omega(t,\ \psi(t)))$$

几乎处处成立的绝对连续函数 ψ, 其中

$$f(t,\ x,\ \Omega(t,\ x))=\{y:\ y=f(t,\ x,\ z),\ z\in\Omega(t,\ x)\}.$$

4. 颤振引理; 松弛控制的逼近

定理 4.1 在第六章推导最大值原理时将要用到. 在本节也要

用它来证明定理 4.3 的如下事实：在适当的前提下，一个控制系统的寻常轨线在该系统的松弛轨线集合内稠密．在定理 4.1 的这两项应用中，我们将看到下述的 f_i 关于 t 仅为可测的假设，是由于应用的需要而赋予我们的．定理 4.1 有时称为“颤振引理”，这样称呼的理由将在下面的注 4.4 中讨论．

定理 4.1 设 $\mathscr{I}$ 是有界闭区间，$\mathscr{X}$ 是 E^n 中的致密集．假定 $f_1, \cdots, f_q$ 是定义在 $\mathscr{I}\times\mathscr{X}$ 上而在 E^n 中取值的函数，它们具有如下性质：

（i）对 $\mathscr{X}$ 中的每一点 x，每一个 f_i 是 $\mathscr{I}$ 上的可测函数；

（ii）对 $\mathscr{I}$ 中的每一个 t，每一个 f_i 在 $\mathscr{X}$ 上连续；

（iii）存在定义在 $\mathscr{I}$ 上的可积函数 μ，使得对 $\mathscr{I}\times\mathscr{X}$ 中的所有点 (t, x) 与 (t, x')，以及 $i=1, \cdots, q$，有

$$\begin{aligned}&|f_i(t, x)|\leqslant\mu(t),\\&|f_i(t, x)-f_i(t, x')|\leqslant\mu(t)|x-x'|.\end{aligned}\tag{4.1}$$

设 $p^i(i=1, \cdots, q)$ 是定义在 $\mathscr{I}$ 上的非负的实值可测函数，几乎处处满足

$$\sum_{i=1}^{q} p^i(t)=1.\tag{4.2}$$

则对每一 $\varepsilon>0$，存在把 $\mathscr{I}$ 分割成一组有限的互不交叠的区间 $E_j(j=1, \cdots, k)$ 的一种剖分，以及对每个 E_j 赋以函数 $f_1, \cdots, f_q$ 之一的值的一种办法，使下述事实成立：如果以 f_{E_j} 表示分配给 E_j 的函数，而 f 是在每个 E_j 内部 E_j^0 与函数 f_{E_j} 相等的函数，即

$$f(t, x)=f_{E_j}(t, x)，若 t\in E_j^0, j=1, \cdots, k,$$

则对 $\mathscr{I}$ 中任意 t'、t'' 和 $\mathscr{X}$ 中的所有 x，有

$$\left|\int_{t'}^{t''}\left(\sum_{i=1}^{q}p^i(t)f_i(t, x)-f(t, x)\right)dt\right|<\varepsilon.\tag{4.3}$$

注 4.1 设 $E_j=[\tau_j, \tau_{j+1}]$，$j=1, \cdots, k$．若令 $f(\tau_j, x)=f_{E_j}(\tau_j, x)$，$j=1, \cdots, k$，$f(\tau_{k+1}, x)=f_{E_k}(\tau_{k+1}, x)$，则 (4.3) 仍成立，而且下述命题为真：函数 f 满足 (4.1)．如果函数 f_i 对 t 的某些值在 $\mathscr{X}$ 上属于类 $C^{(r)}$，则 f 对同样的 t 值属于类 $C^{(r)}$．

对 $n\times m$ 矩阵 M，取 $|M|$ 作为线性变换的模，这里 M 确定

了 E^n 和 E^m 的标准基底间的关系. 如果对 $\mathscr{I}$ 中的每个 t, f_i 在 $\mathscr{X}$ 中属于类 $C^{(1)}$, 而且对于每个 i 和 $\mathscr{X}$ 中的 x,

$$|\partial f_i(t, x)/\partial x| \leqslant \mu(t),$$

其中 $\partial f_i/\partial x$ 表示 f_i 的一阶偏导数的 Jacobi 矩阵, 则 $|\partial f(t, x)/\partial x| \leqslant \mu(t)$.

定理证明的第一步是建立下述引理.

引理 4.1 设 $\mathscr{I}$ 和 $\mathscr{X}$ 如定理所述, f 是从 $\mathscr{I}\times\mathscr{X}$ 到 E^n 的函数, 它具有与定理中的 $f_1, \cdots, f_q$ 相同的性质. 则对任一 $\varepsilon>0$, 存在一个从 $\mathscr{I}\times\mathscr{X}$ 到 E^n 的依赖于 ε 的连续函数 g, 使得对 $\mathscr{X}$ 中的所有 x, 有

$$\int_{\mathscr{I}}|f(t, x)-g(t, x)|dt<\varepsilon. \tag{4.4}$$

证明: 从(4.1)可推得: 对 $\mathscr{X}$ 中的 x 和 x',

$$\int_{\mathscr{I}}|f(t, x)-f(t, x')|dt \leqslant |x-x'|\int_{\mathscr{I}}\mu(t)dt.$$

因此, 对任意的 $\varepsilon>0$, 只要 $|x-x'|<\frac{\varepsilon}{2}\left[\int_{\mathscr{I}}\mu(t)dt\right]^{-1}$[注], 总有

$$\int_{\mathscr{I}}|f(t, x)-f(t, x')|dt<\varepsilon/2. \tag{4.5}$$

因为 $\mathscr{X}$ 致密, 故存在 $\mathscr{X}$ 的有限开覆盖 $\mathscr{O}_1, \cdots, \mathscr{O}_k$, 使得如果 x 和 x' 同在一个 $\mathscr{O}_i$ 中, 则(4.5)成立.

设 $\{x_1, \cdots, x_k\}$ 为使 $x_i\in\mathscr{O}_i$ 的有限点集. 对每个 $i=1, \cdots, k$, 存在定义在 $\mathscr{I}$ 上的连续函数 h_i, 使得

$$\int_{\mathscr{I}}|f(t, x_i)-h_i(t)|dt<\varepsilon/2. \tag{4.6}$$

设 $\gamma_1, \cdots, \gamma_k$ 是对应于 $\mathscr{X}$ 的有限开覆盖 $\mathscr{O}_1, \cdots, \mathscr{O}_k$ 的单位划分, 即: $\gamma_1, \cdots, \gamma_k$ 是 $\mathscr{X}$ 上的实值连续函数, 使得

(i) $\gamma_i(x)\geqslant 0$, 对一切 $x\in\mathscr{X}$;

(ii) $\gamma_i(x)=0$, 若 $x\notin\mathscr{O}_i$; (4.7)

(iii) $\sum_{i=1}^{k}\gamma_i(x)=1$.

[注] 原文误为 $\varepsilon/2\left[\int_{\mathscr{I}}\mu(t)dt\right]$. ——译者注.

关于对应于局部致密的Hausdorff空间的致密子集的有限开覆盖，存在一个单位划分的证明，见Rudin([54], p. 40).

定义 $$g(t,x)=\sum_{i=1}^{k}\gamma_i(x)h_i(t).$$

则g在$\mathscr{I}\times\mathscr{X}$上连续．下面证明$g$满足(4.4)，因而是所要求的函数．

$$\begin{aligned}&\int_{\mathscr{I}}|g(t,x)-f(t,x)|\,dt\\ &\leqslant\int_{\mathscr{I}}\left|\sum_{i=1}^{k}\gamma_i(x)h_i(t)-\sum_{i=1}^{k}\gamma_i(x)f(t,x_i)\right|dt\\ &\quad+\int_{\mathscr{I}}\left|\sum_{i=1}^{k}\gamma_i(x)f(t,x_i)-\sum_{i=1}^{k}\gamma_i(x)f(t,x)\right|dt\\ &\leqslant\sum_{i=1}^{k}\gamma_i(x)\int_{\mathscr{I}}|h_i(t)-f(t,x_i)|\,dt\\ &\quad+\sum_{i=1}^{k}\gamma_i(x)\int_{\mathscr{I}}|f(t,x_i)-f(t,x)|\,dt.\end{aligned}$$

由(4.6)，右边的第一个和号中的每个积分小于$\varepsilon/2$，由此及(4.7)-(iii)就可推得右边的第一个和小于$\varepsilon/2$．现估计右边第二个和号下的第i个被加项．若$x\notin\mathscr{O}_i$，则由(4.7)-(iii)，$\gamma_i(x)=0$，故该被加项为零；若$x\in\mathscr{O}_i$，则由(4.5)，积分小于$\varepsilon/2$，因此由(4.7)-(i)，该被加项小于$\varepsilon\gamma_i(x)/2$．所以，第二个和中每一被加项小于$\varepsilon\gamma_i(x)/2$．至此从(4.7)-(iii)可知第二个和小于$\varepsilon/2$．因此g满足(4.4)，引理得证．

设$\bar{\varepsilon}>0$给定．令

$$\varepsilon=\bar{\varepsilon}/2(2+q+|\mathscr{I}|).\tag{4.8}$$

其中$|\mathscr{I}|$表示$\mathscr{I}$的测度．以后若A为可测集，就用$|A|$表示A的测度．由引理4.1得知，对每个$i=1,\cdots,q$，有一定义在$\mathscr{I}\times\mathscr{X}$上、在$E^n$中取值的连续函数$g_i$，使得

$$\int_{\mathscr{I}}|f_i(t,x)-g_i(t,x)|\,dt<\varepsilon.\tag{4.9}$$

由于每个g_i在$\mathscr{I}\times\mathscr{X}$上连续，且$\mathscr{I}$和$\mathscr{X}$致密，故每个$g_i$在$\mathscr{I}\times\mathscr{X}$上一致连续．因此存在$\delta>0$，使得若$|t-t'|<\delta$，则

$$|g_i(t, x)-g_i(t', x)|<\varepsilon. \tag{4.10}$$

此外，还可假定 δ 是这样的数，使得若 E 是 $\mathscr{I}$ 的可测子集，而 $|E|<\delta$，则

$$\int_E \mu(t)\,dt<\varepsilon. \tag{4.11}$$

设 $\{I_k\}$ 是把 $\mathscr{I}$ 分成有限个不相交叠区间的剖分，使得对每个区间 I_k，$|I_k|<\delta$. 此外，设 $I_k=[t_k, t_{k+1}]$，且 $\cdots<t_{k-1}<t_k<t_{k+1}<t_{k+2}<\cdots$. 对每个 I_k，可构造出把 I_k 划分为不相交叠子区间 $E_{k1}, \cdots, E_{kq}$ 的剖分，使

$$|E_{ki}|=\int_{I_k} p^i(t)\,dt. \tag{4.12}$$

这是可能的，因为

$$\sum_{i=1}^{q}|E_{ki}|=\sum_{i=1}^{q}\int_{I_k} p^i(t)\,dt=\int_{I_k}\left(\sum_{i=1}^{q} p^i(t)\right)dt=|I_k|.$$

此式的最后一个等式由(4.2)推得.

定义

$$f(t, x)=f_i(t, x), \text{ 当 } t\in E_{ki}^0, \tag{4.13}$$

其中 E_{ki}^0 表示 E_{ki} 的内部. 这样，除区间 E_{ki} 的端点外，f 在 $\mathscr{I}$ 的一切点上有定义. 在这些端点上，f 可象注 4.1 那样定义，或用其它任意一种方法定义. 令

$$\lambda(t, x)=\sum_{i=1}^{q} p^i(t) f_i(t, x)-f(t, x). \tag{4.14}$$

当 k 跑遍区间 I_k 的下标集合，i 跑遍集合 $\{1, \cdots, q\}$ 时，区间组 $\{E_{ki}\}$ 构成把 $\mathscr{I}$ 分成有限个不相交叠子区间的一个剖分. 这个剖分（我们重新记作 $\{E_j\}$）的存在性正是定理所要证明的. 如果区间 E_j 原来是区间 E_{ki}，则指定给 E_j 的函数 f_{E_j} 就是 f_i. 如果把定义 λ 的式(4.14)与(4.3)加以比较，并注意到(4.13)，即可看出，要证明定理，必须证明对 $\mathscr{I}$ 中任意的 t' 和 t''，以及 $\mathscr{X}$ 中的所有 x，有

$$\left|\int_{t'}^{t''}\lambda(t, x)\,dt\right|<\bar{\varepsilon}. \tag{4.15}$$

不失一般性可设 $t'<t''$. 点 t' 将属于剖分 $\{I_k\}$ 的某个区间

I_α，而点 t'' 将属于某个区间 I_β. 若 $I_\alpha \neq I_\beta$，以 s_1 表示 I_α 的右端点 $t_{\alpha+1}$，以 s_2 表示 I_β 的左端点 t_β. 那么如果 J 表示下标集 $\{\alpha+1, \alpha+2, \cdots, \beta\}$，就有

$$[s_1, s_2] = [t_{\alpha+1}, t_\beta] = \bigcup_{j\in J} I_j.$$

见图 1. 因此有

$$\left|\int_{t'}^{t''}\lambda\,dt\right| \leqslant \left|\int_{t'}^{s_1}\lambda\,dt\right| + \left|\int_{s_1}^{s_2}\lambda\,dt\right| + \left|\int_{s_2}^{t''}\lambda\,dt\right| = A+B+C.$$

从(4.14)、注 4.1、(4.2)、(4.11) 以及 t' 与 s_1 在同一区间之内且 $|I_\alpha|<\delta$ 这一事实，可得出

$$A \leqslant \int_{t'}^{s_1}\left(\sum_{i=1}^{q}|p^i f_i| + |f|\right)dt = \int_{t'}^{s_1}\sum_{i=1}^{q}p^i|f_i|\,dt + \int_{t'}^{s_1}|f|\,dt$$

$$\leqslant \int_{t'}^{s_1}\left(\sum_{i=1}^{q}p^i\right)\mu\,dt + \int_{t'}^{s_1}\mu\,dt = 2\int_{t'}^{s_1}\mu\,dt < 2\varepsilon.$$

注意到若 t' 和 t'' 在同一区间 I_α 之内，则上述估计与(4.8)结合起来便得出(4.15).

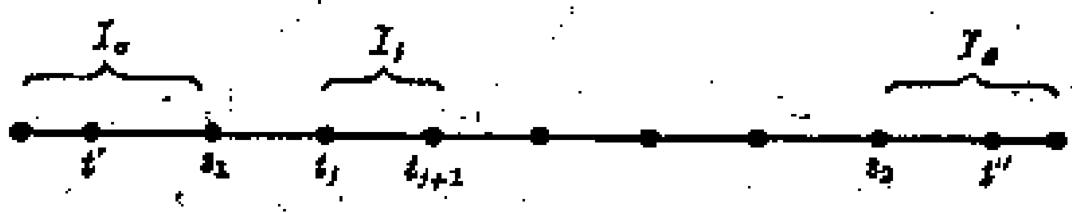

图 1

用与上述类似的论证可证得 $C<2\varepsilon$.

现估计 B. 因 $I_k = [t_k, t_{k+1}]$，故

$$B = \left|\int_{s_1}^{s_2}\lambda\,dt\right| \leqslant \sum_{j\in J}\left|\int_{t_j}^{t_{j+1}}\lambda\,dt\right|.$$

对 $t\in E_{ji}$，令 $g(t, x) = g_i(t, x)$，这里 $i=1, \cdots, q$, $j\in J$；则可估计上式右边的每个被加项如下：

$$\left|\int_{t_j}^{t_{j+1}}\lambda\,dt\right| \leqslant \left|\int_{t_j}^{t_{j+1}}\left(\sum_{i=1}^{q}p^i(f_i-g_i)\right)dt\right|$$

$$+\left|\int_{t_j}^{t_{j+1}}\left(\sum_{i=1}^{q}p^i g_i - g\right)dt\right| + \left|\int_{t_j}^{t_{j+1}}(g-f)\,dt\right|$$

$$= A_j + B_j + C_j.$$

因此

$$B\leqslant\sum_{j\in J}(A_j+B_j+C_j). \tag{4.16}$$

从 p^i 的非负性和(4.2)得出

$$A_j\leqslant\int_{t_j}^{t_{j+1}}\left(\sum_{i=1}^{q}p^i|f_i-g_i|\right)dt\leqslant\sum_{i=1}^{q}\int_{t_j}^{t_{j+1}}|f_i-g_i|dt.$$

由 f 和 g 的定义得出

$$C_j\leqslant\int_{t_j}^{t_{j+1}}|g-f|dt\leqslant\sum_{i=1}^{q}\int_{t_j}^{t_{j+1}}|f_i-g_i|dt.$$

所以

$$\begin{aligned}\sum_{j\in J}(A_j+C_j)&\leqslant 2\sum_{i=1}^{q}\int_{s_1}^{s_2}|f_i-g_i|dt\\&\leqslant 2\sum_{i=1}^{q}\int_{\mathscr{I}}|f_i-g_i|dt<2q\varepsilon,\end{aligned} \tag{4.17}$$

其中最后一个不等式从(4.9)推得.

现考虑 B_j.

$$B_j=\left|\int_{t_j}^{t_{j+1}}\left(\sum_{i=1}^{q}p^ig_i-g\right)dt\right|=\left|\sum_{i=1}^{q}\int_{t_j}^{t_{j+1}}p^ig_i\,dt-\sum_{i=1}^{q}\int_{E_{ji}}g_i\,dt\right|.$$

在每个集合 E_{ji} 中取一点 t_{ji}. 因为 $E_{ji}\subset I_j$, $|I_j|<\delta$, 故从(4.10)推得：对 I_j 中一切 t, $\mathscr{X}$ 中一切 x 和所有 $i=1, \cdots, q$,

$$g_i(t, x)=g_i(t_{ji}, x)+\eta_i(t, x),$$

其中 $|\eta_i(t, x)|<\varepsilon$. 所以，应用(4.12)可得

$$\begin{aligned}B_j&=\left|\sum_{i=1}^{q}\left(\int_{t_j}^{t_{j+1}}(p^i(t)g_i(t_{ji}, x)+p^i(t)\eta_i(t, x))\,dt\right.\right.\\&\qquad\left.\left.-\int_{E_{ji}}(g_i(t_{ji}, x)+\eta_i(t, x))dt\right)\right|\\&=\left|\sum_{i=1}^{q}\left(g_i(t_{ji}, x)|E_{ji}|-g_i(t_{ji}, x)|E_{ji}|\right.\right.\\&\qquad\left.\left.+\int_{t_j}^{t_{j+1}}p^i(t)\eta_i(t, x)\,dt-\int_{E_{ji}}\eta_i(t, x)\,dt\right)\right|\\&<\sum_{i=1}^{q}\left(\varepsilon\int_{t_j}^{t_{j+1}}p^i\,dt+\varepsilon|E_{ji}|\right)=2\varepsilon|I_j|.\end{aligned} \tag{4.18}$$

因此
$$\sum_{j\in J}B_j<2\varepsilon|s_2-s_1|\leqslant 2\varepsilon|\mathscr{I}|.$$

把此不等式与(4.17)、(4.16)联合起来就得出 $B<2\varepsilon(q+|\mathscr{I}|)$. 若把这个估计与关于 A 和 C 的估计结合起来，并应用(4.8)，便得

$$\left|\int_{t'}^{t''}\lambda\,dt\right|<2\varepsilon(2+q+|\mathscr{I}|)=\bar{\varepsilon},$$

这就是所要证明的(4.15). 定理证毕.

在下一个定理中我们要证明,如果把 $\mathscr{X}$ 中的向量 x 换成等度绝对连续函数族中的函数 ψ, 则(4.3)仍然正确.

定理 4.2 设 $f_1, \cdots, f_q$ 与 $p^1, \cdots, p^q$ 与定理 4.1 中的相同, Ψ 是定义在 $\mathscr{I}$ 上而在 $\mathscr{X}$ 中取值的等度绝对连续函数族. 则对每个 $\varepsilon>0$, 存在一种把 $\mathscr{I}$ 分为有限个不相交叠区间的剖分,和对每个区间 E_j 赋以函数 $f_1, \cdots, f_q$ 之一的值的一种分配, 使下述事实成立: 以 f_{E_j} 表示分配给 E_j 的函数, 如果 f 为在 E_j 的内部 E_j^0 与 f_{E_j} 相等的函数, 即

$$f(t, x)=f_{E_j}(t, x),\ t\in E_j^0,$$

则对 $\mathscr{I}$ 的每一 t' 和 t'' 以及 Ψ 中的每一个函数 ψ, 有

$$\left|\int_{t'}^{t''}\left(\sum_{i=1}^{q}p^i(t)f_i(t, \psi(t))-f(t, \psi(t))\right)dt\right|<\varepsilon. \tag{4.19}$$

证明: 设给定 $\varepsilon>0$. 因为 Ψ 中的函数是等度绝对连续的, 且 $\mathscr{I}$ 是致密的, 故存在把 $\mathscr{I}$ 分成有限个不相交叠子区间

$$\{I_j\}=\{[t_j, t_{j+1}]\}\quad (j=1, \cdots, k)$$

的一种剖分, 使得 $\cdots<t_{j-1}<t_j<t_{j+1}<t_{j+2}<\cdots$, 而且对 Ψ 中的所有函数 ψ 与 $I_j(j=1, \cdots, k)$ 中的一切 t, 有

$$|\psi(t)-\psi(t_j)|<\varepsilon\left(2(q+1)\int_{\mathscr{I}}\mu\,dt\right)^{-1}=\varepsilon'. \tag{4.20}$$

现对 $f_1, \cdots, f_q$ 和 $p^1, \cdots, p^q$ 应用定理 4.1, 把其中的 ε 换成 $\varepsilon/2k$, 则存在定理 4.1 描述过的函数 f, 使得对 $\mathscr{X}$ 中的一切 x 和 $\mathscr{I}$ 中的所有 t'、t'', 有

$$\left|\int_{t'}^{t''}\lambda(t, x)dt\right|<\varepsilon/2k, \tag{4.21}$$

其中 λ 由(4.14)确定. 我们通过证明对 Ψ 中的一切 ψ 和 $\mathscr{I}$ 中的 t' 与 t'', 有

$$\left|\int_{t'}^{t''}\lambda(t, \psi(t))dt\right|<\varepsilon,$$

而完成本定理的证明. 定义

$$\hat{\lambda}(t)=\lambda(t,\ \psi(t_j)),\ t_j\leqslant t\leqslant t_{j+1},\ j=1,\ \cdots,\ k,$$

并令 $\hat{\lambda}(t_{k+1})=\lambda(t_{k+1},\ \psi(t_k))$. 则

$$\begin{aligned}&\left|\int_{t'}^{t''}\lambda(t,\ \psi(t))dt\right|\\&\leqslant\left|\int_{t'}^{t''}(\lambda(t,\ \psi(t))-\hat{\lambda}(t))dt\right|+\left|\int_{t'}^{t''}\hat{\lambda}(t)dt\right|\\&=A+B.\end{aligned}\tag{4.22}$$

设 $t'<t''$, $t'\in I_\alpha=[t_\alpha,\ t_{\alpha+1}]$, $t''\in I_\beta=[t_\beta,\ t_{\beta+1}]$. 以 J 表示下标集 $\{\alpha,\ \alpha+1,\ \alpha+2,\ \cdots,\ \beta\}$, 则

$$\begin{aligned}A&\leqslant\int_{t'}^{t''}|\lambda(t,\ \psi(t))-\hat{\lambda}(t)|dt\leqslant\int_{t_\alpha}^{t_{\beta+1}}|\lambda(t,\ \psi(t))-\hat{\lambda}(t)|dt\\&=\sum_{j\in J}\int_{t_j}^{t_{j+1}}|\lambda(t,\ \psi(t))-\lambda(t,\psi(t_j))|dt\\&=\sum_{j\in J}\int_{t_j}^{t_{j+1}}\Big|\sum_{i=1}^{q}p^i(t)\{f_i(t,\ \psi(t))-f_i(t,\ \psi(t_j))\}\\&\qquad\qquad+f(t,\ \psi(t_j))-f(t,\ \psi(t))\Big|dt\\&\leqslant\sum_{j\in J}\int_{t_j}^{t_{j+1}}\Big(\sum_{i=1}^{q}p^i(t)\mu(t)|\psi(t)-\psi(t_j)|+\mu(t)|\psi(t)-\psi(t_j)|\Big)dt,\end{aligned}$$

这里最后一个不等式从 (4.1) 和注 4.1 推出. 由关系式 $0\leqslant p^i(t)\leqslant1$ 与(4.20)可看出最后这个和又小于

$$\sum_{j\in J}(q+1)\varepsilon'\int_{t_j}^{t_{j+1}}\mu\,dt\leqslant\varepsilon'(q+1)\int_{\mathscr{I}}\mu\,dt=\varepsilon/2.$$

这就证明了 $A<\varepsilon/2$.

为了估计 B, 写出下列不等式:

$$\begin{aligned}B\leqslant&\left|\int_{t'}^{t_{\alpha+1}}\lambda(t,\ \psi(t))dt\right|+\sum_{j=\alpha+1}^{\beta-1}\left|\int_{t_j}^{t_{j+1}}\lambda(t,\ \psi(t_j))dt\right|\\&+\left|\int_{t_\beta}^{t''}\lambda(t,\ \psi(t_\beta))dt\right|.\end{aligned}$$

由于(4.21), 右端的每个被加项都小于 $\varepsilon/2k$. 因为最多只有 k 个被加项(区间 I_j 的数目), 故得 $B<\varepsilon/2$. 把这个估计式与 A 的估计式联合起来,并代入(4.22),便得到所希望的结果.

下一个定理的证明要用到一个不等式, 它就是文献中有名的

Gronwall 不等式，这个不等式在微分方程的研究中很有用处。

引理 4.2 设 ρ 与 μ 是在 $[0, \infty)$ 上连续的非负实值函数，使得对一切 $[0, \infty)$ 上的 t_0、t，不等式

$$\rho(t) \leqslant \alpha + \int_{t_0}^{t} \mu(s)\rho(s)ds, \quad \alpha \geqslant 0 \tag{4.23}$$

成立，则

$$\rho(t) \leqslant \alpha \exp\left(\int_{t_0}^{t} \mu(s)ds\right). \tag{4.24}$$

证明：设 $\alpha > 0$. 则(4.23)的右边严格为正，故得

$$\rho(t)\mu(t)\left[\alpha + \int_{t_0}^{t} \mu(s)\rho(s)ds\right]^{-1} \leqslant \mu(t).$$

从 t_0 到 t 积分此不等式两边，并应用(4.23)，可得

$$\log \rho(t) \leqslant \log\left[\alpha + \int_{t_0}^{t} \mu\rho\, ds\right] \leqslant \log \alpha + \int_{t_0}^{t} \mu\, ds.$$

由此即得(4.24).

若 $\alpha = 0$，则(4.23)对一切 $\alpha_1 > 0$ 成立. 因此(4.24)对一切 $\alpha_1 > 0$ 成立. 令 $\alpha_1 \to 0$，便得 $\rho(t) \equiv 0$. 因此(4.24)显然为真.

注 4.2 引理的证明表明，若 $\alpha > 0$ 且(4.23)中的严格不等式成立，则(4.24)中也是严格不等式成立.

定理 4.3 设 $\mathscr{I}$ 是 E^1 中的致密区间，$\mathscr{X}$ 是 E^n 中的致密区间. 令 $\mathscr{R} = \mathscr{I} \times \mathscr{X}$，$\mathscr{G} = \mathscr{R} \times \mathscr{U}$，这里 $\mathscr{U}$ 为 E^m 中的区域. 设 f 为从 $\mathscr{G}$ 到 E^n 的连续映射，Ω 是从 $\mathscr{R}$ 到 E^m 的子集的映象，它与 x 无关，即对 $\mathscr{X}$ 中一切 x 和 x'，$\Omega(t, x') = \Omega(t, x) = \Omega(t)$. 设 $\mathscr{D}$ 与III(4.1)相同. 假定存在定义在 $\mathscr{I}$ 上的可积函数 μ，使得对一切 $\mathscr{D}$ 中的(t, x, z)，有

$$|f(t, x, z)| \leqslant \mu(t),$$

且对 $\mathscr{D}$ 中所有的(t, x, z)与(t, x', z)，有

$$|f(t, x, z) - f(t, x', z)| \leqslant \mu(t)|x - x'|. \tag{4.25}$$

设 $\mathscr{I}_1 = [t_0, t_1]$ 是包含在 $\mathscr{I}$ 内部的致密区间，$\mathscr{X}_1$ 是在 $\mathscr{X}$ 内部的致密区间. 设 $\mathscr{R}_1 = \mathscr{I}_1 \times \mathscr{X}_1$. 设

$$v = (\tilde{u}, p) = (u_1, \cdots, u_{n+2}, p^1, \cdots, p^{n+2})$$

是与控制系统

$$\frac{dx}{dt}=f(t,\ x,\ u(t))$$

相对应的松弛系统

$$\frac{dx}{dt}=\sum_{i=1}^{n+2}p^i(t)f(t,\ x,\ u_i(t))$$

在 $\mathscr{I}_1$ 上的一个松弛控制. 设这两个系统的初始点都是 $(t_0,\ x_0)$, ψ 是与 v 对应的在 $\mathscr{I}_1$ 上的松弛轨线, 而且对 $[t_0,\ t_1]$ 中的一切 t, $\psi(t)\in\mathscr{X}_1$. 则存在一个 $\varepsilon_0>0$, 使得对每一满足 $0<\varepsilon<\varepsilon_0$ 的 ε, 存在一个定义在 $\mathscr{I}_1$ 上的控制 u_ε, 它具有如下性质: 对 $\mathscr{I}_1$ 中几乎所有的 t, $u_\varepsilon\in\Omega(t)$, 对应于 u_ε 的轨线 ϕ_ε 位于 $\mathscr{I}_1\times\mathscr{X}_1$ 之内, 且对 $\mathscr{I}_1$ 中一切 t, $|\phi_\varepsilon(t)-\psi(t)|<\varepsilon$.

注 4.3　定理 4.3 表明, 在适当的假设之下, 系统的寻常轨线在松弛轨线的集合内按 $[t_0,\ t_1]$ 上的一致拓扑是稠密的. 因此对 $[t_0,\ t_1]$ 上任一松弛轨线 ψ, 存在控制序列 $\{u_k\}$ 与相应的轨线序列 $\{\phi_k\}$, 使得 $u_k(t)\in\Omega(t)$ 几乎处处满足, 而且在 $[t_0,\ t_1]$ 上一致地有 $\phi_k\to\psi$. 我们提醒读者, 对于一个具体的控制问题, 若 ψ 是容许的松弛轨线, 但对 $(\phi_k,\ u_k)$ 关于原始问题却不一定是容许的, 因为或者映射 $t\to f^0(t,\ \phi_k(t),\ u_k(t))$ 可能不可积, 或者 ϕ_k 的端点可能不满足端点条件. 请读者回顾控制 (定义 II 3.1) 与容许控制 (定义 II 3.2) 之间的差别.

注意, 这里对约束集 $\Omega(t)$ 的性质没有作任何假定.

证明: 令 ε_0 表示 $\partial\mathscr{X}$ 与 $\partial\mathscr{X}_1$ 之间的距离, 这里对任一集合 A, 记号 ∂A 表示 A 的边界. 那么 $\varepsilon_0>0$. 令

$$k=\int_{t_0}^{t_1}\mu\,dt,\tag{4.26}$$

并设 ε 为满足 $0<\varepsilon<\varepsilon_0$ 的任意数. 对 $\mathscr{I}_1\times\mathscr{X}$ 中的 $(t,\ x)$ 和 $i=1,\ \cdots,\ n+2$, 令

$$f_i(t,\ x)=f(t,\ x,\ u_i(t)).\tag{4.27}$$

不难验证, 由于本定理的假设, 函数 f_i 满足定理 4.1 和 4.2 的假设. 特别要指出, 因为 f 在 $\mathscr{D}$ 上连续, 且每个 u_i 可测, 故函数 f_i 对

$\mathscr{X}_1$ 中每一固定的 x 在 $\mathscr{I}_1$ 上可测.

令 $\varepsilon'=\varepsilon e^{-k}$. 下面把定理 4.2 应用于刚刚定义的函数 $f_1, \cdots, f_{n+2}$, 松弛控制中的函数 $p^1, \cdots, p^{n+2}$ 以及由单个元素即松弛控制 ψ 组成的函数族 Ψ, 且取定理 4.2 中的 ε 值等于 ε'. 我们得知存在函数 $\hat{f}$, 使得对 $x\in\mathscr{X}_1$ 当 $t\in\mathscr{I}_1$, 有

$$\hat{f}(t, x)=f_{E_j}(t, x), \qquad t\in E_j^0; \tag{4.28}$$

而对 $\mathscr{I}_1$ 中任意的 t' 与 t'', 有不等式

$$\left|\int_{t'}^{t''}\left(\sum_{i=1}^{n+2} p^i(t)f_i(t, \psi(t))-\hat{f}(t, \psi(t))\right)dt\right|<\varepsilon'. \tag{4.29}$$

从 f_i 的定义和(4.28)可推得

$$\hat{f}(t, x)=f_{E_j}(t, x)=f(t, x, u_{E_j}(t)), \quad t\in E_j^0. \tag{4.30}$$

定义
$$u_s(t)=u_{E_j}(t), \text{ 若 } t\in E_j^0.$$

则由于 u_{E_j} 是 $u_1, \cdots, u_{n+2}$ 中的一个, 且每个 u_i 在 $\mathscr{I}_1$ 上几乎处处满足 $u_i(t)\in\Omega(t)$, 于是得知 $u_s(t)\in\Omega(t)$ 在 $\mathscr{I}_1$ 上几乎处处满足. 由 u_s 的定义和(4.30)即可得到

$$\hat{f}(t, x)=f(t, x, u_s(t)).$$

考虑带初始点(t_0, x_0)的系统

$$\frac{dx}{dt}=f(t, x, u_s(t))=\hat{f}(t, x). \tag{4.31}$$

由于 f 满足(4.25), 故知假如把 u_s 扩充成在 $\mathscr{I}$ 上有定义且可测, 则对 $\mathscr{I}\times\mathscr{X}$ 内部的每一点(t_2, x_2), 有(4.31)的唯一解通过它. 特别, 存在(4.31)的以(t_0, x_0)为初始点的唯一解 ϕ_s. 这个解在以 t_0 作为其内点的某开区间上有定义. 以 $\mathscr{I}_{\max}=(a, b)$ 表示 ϕ_s 有定义的最大区间. 如果 $[a, b]\subset\mathscr{I}_1$, 则 $\limsup\limits_{t\to b}\phi_s(t)$ 必为 $\mathscr{X}$ 的边界点, 因为否则可以把解 ϕ_s 延展到一个包含 $\mathscr{I}_{\max}$ 在其内部的区间上. 这与 $\mathscr{I}_{\max}$ 的最大性矛盾. 我们将证明对 $\mathscr{I}_{\max}$ 中的一切 t, 不等式 $|\phi_s(t)-\psi(t)|<\varepsilon$ 成立. 因为对 $[t_0, t_1]$ 中一切 t, $\psi(t)\in\mathscr{X}_1$, 且因 $\varepsilon<\varepsilon_0=\operatorname{dist}(\partial\mathscr{X}_1, \partial\mathscr{X})$, 由此可得 $[a, b]=\mathscr{I}_1$ 且 ϕ_s 在整个 $\mathscr{I}_1$ 上有定义. 此外, 在整个 $\mathscr{I}_1$ 上将有

$$|\phi_s(t)-\psi(t)|<\varepsilon.$$

因为ψ在整个$\mathscr{I}_1$上有定义，且$\psi(t_0)=\phi_\varepsilon(t_0)=x_0$，故对$[t_0, b]$中的一切$t$有

$$\begin{aligned}|\psi(t)-\phi_\varepsilon(t)| &= \left|\int_{t_0}^{t}(\psi'(s)-\phi_\varepsilon'(s))ds\right| \\ &= \left|\int_{t_0}^{t}\left(\sum_{i=1}^{n+2}p^i(s)f_i(s, \psi(s))-\hat{f}(s, \phi_\varepsilon(s))\right)ds\right| \\ &\leqslant \left|\int_{t_0}^{t}\left(\sum_{i=1}^{n+2}p^i(s)f_i(s, \psi(s))-\hat{f}(s, \psi(s))\right)ds\right| \\ &\quad + \left|\int_{t_0}^{t}(\hat{f}(s, \psi(s))-\hat{f}(s, \phi_\varepsilon(s)))ds\right| \\ &< \varepsilon'+\int_{t_0}^{t}|\hat{f}(s, \psi(s))-\hat{f}(s, \phi_\varepsilon(s))|ds,\end{aligned}$$

其中最后的不等式从(4.29)推得．现在从(4.30)与(4.25)可推得

$$\int_{t_0}^{t}|\hat{f}(s, \psi(s))-\hat{f}(s, \phi_\varepsilon(s))|ds\leqslant\int_{t_0}^{t}\mu(s)|\psi(s)-\phi_\varepsilon(s)|ds.$$

由引理 4.2、注 4.2、式(4.26)和ε'的定义可断言：

$$|\psi(t)-\phi_\varepsilon(t)|<\varepsilon'\exp\left(\int_{t_0}^{t}\mu\,ds\right)\leqslant\varepsilon'e^K=\varepsilon,$$

定理证毕．

注 4.4 从定理 4.3 的证明可以看出为什么必须假定定理 4.1 的函数f_i关于t可测且关于x连续，而不假定关于(t, x)连续．因为只假定控制u是可测的，无论我们假定f的性质如何正则，也只能保证由(4.27)定义的函数f_i关于t可测．

现在能给出把定理 4.1 称为"颤振引理"的理由了．在大多数应用中，函数$f_1, \cdots, f_q$象在定理 4.3 中那样获得，亦即对状态方程为$dx/dt=f(t, x, u(t))$的系统，我们取q个控制$u_1, \cdots, u_q$，并用方程(4.27)定义函数$f_1, \cdots, f_q$．定理 4.1 的函数f是用与本定理获得函数$\hat{f}$的同样方式得到的，也就是把基本区间$\mathscr{I}$划分为大量的小区间，并在每个子区间上取控制$u_1, \cdots, u_q$中的一个来构造控制u_ε．在物理系统中，控制u_ε相当于在各个控制$u_1, \cdots, u_q$之间的来回快速转换．在工程术语中，这种系统就叫做"颤振"．所以控制u_ε有时就叫做"颤振控制"．

从定理 4.3 的证明，我们知道了比刚才提到的更多的事实，这就是松弛轨线能用寻常轨线加以逼近，而且我们要多近就可以多近．我们还知道，利用由据以定义问题的松弛控制的那些控制构造出来的颤振控制，能够实现这种逼近．

注 4.5 如果取 $\mathscr{X}=\mathscr{X}_1=E^n$，本定理仍然有效．

习题 4.1 考虑问题 2，其中的 f^0、f 与 Ω 满足定理 4.3 的假设，而 $\mathscr{X}_1=\mathscr{X}=E^n$．设端点集 $\mathscr{B}$ 给定如下：t_0、x_0、t_1 固定，x_1 为 $\mathscr{X}=E^n$ 中的任一元素；G 为 $C[t_0, t_1]$ 上的连续泛函．证明如果问题 2 有解 (ϕ^*, u^*)，则 (ϕ^*, u^*) 也是对应的松弛问题（问题 2R）的解．

5. 可 达 集

稍微不精确地说，控制系统

$$\frac{dx}{dt}=f(t, x, u(t)) \tag{5.1}$$

以 (t_0, x_0) 为初始点的在时刻 t_1 的可达集，是系统的轨线在时刻 t_1 所能到达的点的集合．我们将从定理 III 4.1 和定理 III 8.1 推得可达集的性质．所以，在研究(5.1)的轨线时，假设系统(5.1)是关于 $f^0\equiv0$ 的控制问题的状态方程组是有益的．今后，在本节内将作这一假设．另外，引入度量空间的致密子集所构成的空间上的某种度量也是有益的．

定义 5.1 设 $\mathscr{S}$ 是度量空间，A 和 B 是 $\mathscr{S}$ 的致密子集．令

$$h(A, B)=\frac{1}{2}\{\max_{a\in A} d(a, B)+\max_{b\in B} d(A, b)\},$$

其中 d 为 $\mathscr{S}$ 上的度量．

能够证明，h 在由 $\mathscr{S}$ 的致密子集构成的空间 $\mathscr{T}$ 上定义了一种度量．度量 h 称为 Hausdorff 度量．

习题 5.1 证明 h 定义了 $\mathscr{T}$ 上的一种度量的论断．

现在给出可达集的精确定义．以 $\mathscr{A}^+$ 表示所有对 (ϕ, u) 的集合，这里 u 为控制，ϕ 为对应于 u 的轨线，它满足容许对定义中除

$(t_0, \phi(t_0), t_1, \phi(t_1)) \in \mathscr{B}$ 之外的所有要求. 注意到因为 $f^0 \equiv 0$, 映象 $t \to f^0(t, \phi(t), u(t))$ 在 $L_1[t_0, t_1]$ 中的要求总是满足的.

定义 5.2 系统(5.1)以 (t_0, x_0) 为初始点的在时刻 $t \geqslant t_0$ 的可达集定义为所有这样的点 x 的集合, 使得对属于 $\mathscr{A}^+$ 中的对 (ϕ, u) 且满足 $\phi(t_0) = x_0$ 的某一轨线 ϕ, 关系式 $\phi(t) = x$ 成立. 可达集记作 $\mathscr{K}(t, t_0, x_0)$[注].

设 (t_0, x_0) 与 $t_1 > t_0$ 均固定. 以 $\mathscr{A}^+(t_0, x_0, t_1)$ 表示 $\mathscr{A}^+$ 中定义在 $[t_0, t_1]$ 上且满足 $\phi(t_0) = x_0$ 的对 (ϕ, u) 的集合. 以下假定 $\mathscr{A}^+(t_0, x_0, t_1)$ 非空, 而且在各种叙述中不再重复这个假定. 对按状态为线性的系统, 保证 $\mathscr{A}^+(t_0, x_0, t_1)$ 非空的条件已在第三章第二节给出了. 对区间 $t_0 < t < t_1$ 中的每个 t, 可以用与定义 $\mathscr{A}^+(t_0, x_0, t_1)$ 相同的方法定义对 (ϕ, u) 的集合 $\mathscr{A}^+(t_0, x_0, t)$. 若 $t < t'$, 则 $\mathscr{A}^+(t_0, x_0, t) \supset \mathscr{A}^+(t_0, x_0, t')$. 因为当 $t_0 < t \leqslant t_1$ 时 $\mathscr{A}^+(t_0, x_0, t)$ 非空, 故对于 $t_0 \leqslant t \leqslant t_1$, 可达集 $\mathscr{K}(t, t_0, x_0)$ 非空.

如果下面我们只集中注意力于区间 $[t_0, t_1]$, 而不考虑定义在这个区间之外的轨线和轨线段, 则 $\mathscr{A}^+$ 的下述定义与上面那个定义是一致的:

$$\mathscr{A}^+ = \bigcup_{t_0 \leqslant t \leqslant t_1} \mathscr{A}^+(t_0, x_0, t).$$

今以 $\mathscr{A}_T^+$ 表示属于 $\mathscr{A}^+$ 中的对 (ϕ, u) 的轨线 ϕ 的集合; 而属于 $\mathscr{A}^+(t_0, x_0, t)$ 的对 (ϕ, u) 的轨线 ϕ 的集合则记作 $\mathscr{A}_T^+(t_0, x_0, t)$.

下面两个定理是讨论可达集结构的一般定理. 定理 5.1 的假设对应于定理 III 4.1 的假设. 正如用定理 III 4.1 研究存在问题那样, 我们将使用下面的定理 5.1 和 5.2 来得到关于某些问题的可达集的知识, 而这些问题的假设所包含的条件比下述定理 5.1 与 5.2 的条件容易验证.

定理 5.1 假设(5.1)中的函数 f 是连续的, 约束映射 Ω 在 $\mathscr{R}$ 上是上半连续的. 对 $\mathscr{R}$ 中的每一点 (t, x), 设集合

$$\mathscr{Q}(t, x) = \{y : y = f(t, x, z),\ z \in \Omega(t, x)\} \quad . \tag{5.2}$$

[注] 即 $\mathscr{K}(t, t_0, x_0) = \{x : x = \phi(t), \phi(t_0) = x_0, (\phi, u) \in \mathscr{A}^+\}$. ——译者注.

是闭和凸的，而且映射 $\mathscr{Q}^+$ 在 $\mathscr{R}$ 中的每一点上满足弱 Cesari 性质．设轨线集合 $\mathscr{A}_T^+$ 等度绝对连续．则对 $[t_0, t_1]$ 中的每个 t，集合 $\mathscr{K}(t, t_0, x_0)$ 是致密的．此外，如果对 $[t_0, t_1]$ 中的每个 t，$\mathscr{A}_T^+(t_0, x_0, t)$ 中的每条轨线是 $\mathscr{A}_T^+(t_0, x_0, t_1)$ 中的轨线的限制[注]，则映射 $t\to\mathscr{K}(t, t_0, x_0)$ 是从 $[t_0, t_1]$ 映入 E^n 中赋予了 Hausdorff 度量的致密子集所构成的空间的连续映射．

首先证明集合 $\mathscr{K}(t, t_0, x_0)$ 是致密的．如果 $t=t_0$，则 $\mathscr{K}(t, t_0, x_0)=x_0$，故它是致密的．现设 $t>t_0$．因为对 $\mathscr{A}_T^+$ 中所有 ϕ，$\phi(t_0)=x_0$，而且因为 $\mathscr{A}_T^+$ 中的函数等度绝对连续，故知 $\mathscr{A}_T^+$ 中所有轨线 ϕ 位于致密集 $\mathscr{R}_0\subseteq\mathscr{R}$ 中．因此集合 $\mathscr{K}(t, t_0, x_0)$ 有界．为了证明它们是致密的，只须证明它们为闭．

设 x 为 $\mathscr{K}(t, t_0, x_0)$ 的极限点．则存在序列 $\{x_k\}$，使得 $x_k\in\mathscr{K}(t, t_0, x_0)$ 且 $x_k\to x$．因而存在 $\mathscr{A}_T^+$ 中的序列 $\{(\phi_k, u_k)\}$，使得 $\phi_k(t)=x_k$．假定除了(5.1)是关于 $f^0\equiv0$ 的问题的状态方程外，还加上这个问题的端点条件是由集合 $\mathscr{B}=\{(t_0', x_0', t_1', x_1'): t_0'=t_0, t_1'=t, x_0'=x_0$，而 x_1' 为 E^n 中任一向量$\}$确定的．则 $\mathscr{A}_T^+(t_0, x_0, t)$ 中的对 (ϕ, u) 是这个问题的容许对．特别，上述序列中的对 (ϕ_k, u_k) 是容许的．容易验证，在本定理的假设之下，把序列 $\{(\phi_k, u_k)\}$ 取作定理 III 4.1 的集合 $\mathscr{A}_0$ 时，定理 III 4.1 就可以应用到这里来．从而得到存在 $\mathscr{A}_T^+(t_0, x_0, t)$ 中的对 (ϕ^*, u^*)，使得 $\phi_k\to\phi^*$ 在 $[t_0, t_1]$ 上一致地成立．因此，$\phi^*(t)=x$，所以 $\mathscr{K}(t, t_0, x_0)$ 为闭．

映射 $t\to\mathscr{K}(t, t_0, x_0)$ 的连续性证明如下．设给定 $\varepsilon>0$．因为 $\mathscr{A}_T^+$ 中的函数等度绝对连续，故存在 $\delta>0$，使得对 $[t_0, t_1]$ 中满足 $|t-t'|<\delta$ 的 t 和 t'，不等式 $\left|\int_t^{t'}\phi'(s)ds\right|<\varepsilon$ 成立．设 x 为 $\mathscr{K}(t, t_0, x_0)$ 中的点．则存在既在 $\mathscr{A}^+(t_0, x_0, t)$ 中，也在 $\mathscr{A}^+(t_0, x_0, t_1)$ 中的对 (ϕ, u)，使得 $\phi(t)=x$．设 t' 为 $[t_0, t_1]$ 中使 $|t'-t|<\delta$ 的点．则

［注］ $\phi\in\mathscr{A}_T^+(t_0, x_0, t)$ 是 $\tilde{\phi}\in\mathscr{A}_T^+(t_0, x_0, t_1)$ 的限制，这是指：$[t_0, t]\subseteq[t_0, t_1]$，使得当 $s\in[t_0, t]$ 时，有 $\phi(s)=\tilde{\phi}(s)$．此时 $\tilde{\phi}$ 又称为 ϕ 的扩张．——译者注

$$|\phi(t')-\phi(t)|=\left|\int_t^{t'}\phi'(s)\,ds\right|<\varepsilon.$$

因此 $d(x, \mathscr{K}(t', t_0, x_0))<\varepsilon$，这里 d 表示欧氏距离。由于 x 是 $\mathscr{K}(t, t_0, x_0)$ 的任一点，故有 $\max\{d(x, \mathscr{K}(t', t_0, x_0)); x\in\mathscr{K}(t, t_0, x_0)\}<\varepsilon$. 类似地，$\max\{d(x, \mathscr{K}(t, t_0, x_0)); x\in\mathscr{K}(t', t_0, x_0)\}<\varepsilon$. 因此对 $[t_0, t_1]$ 中使 $|t-t'|<\delta$ 的一切 t 和 t'，都有

$$h(\mathscr{K}(t, t_0, x_0), \mathscr{K}(t', t_0, x_0))<\varepsilon.$$

推论 5.1 对 $[t_0, t_1]$ 中每个 t，集合 $\mathscr{A}^+_T(t_0, x_0, t)$ 是 $[t_0, t_1]$ 上的连续函数空间 $C[t_0, t_1]$ 按一致拓扑的致密子集.

证明：因为 $\mathscr{A}^+_T(t_0, x_0, t)$ 中的轨线族 $\{\phi\}$ 等度绝对连续，故它一致有界且等度连续. 因此据 Ascoli 定理推得，如果 $\{\phi_k\}$ 为 $\mathscr{A}^+_T(t_0, x_0, t)$ 中的轨线序列，则存在一个子序列(重新记作 $\{\phi_k\}$)和 $C[t_0, t]$ 中的连续函数 ϕ，使得在 $C[t_0, t]$ 中 $\phi_k\to\phi$. 象证明 $\mathscr{K}(t, t_0, x_0)$ 是闭的那样，应用定理 III 4.1 可以证明 ϕ 在 $\mathscr{A}^+_T(t_0, x_0, t)$ 之内.

定理 5.2 设除了映射 $\mathscr{Q}^+$ 满足弱 Cesari 性质的假定之外，定理 5.1 的假设成立. 再假设 III 8.1 的 (ii)(iii) 成立，而且对 $\mathscr{A}^+$ 中的一切 (ϕ, u)，III(8.2)式成立. 则定理 5.1 的结论成立.

定理 5.2 的证明与定理 5.1 类似，只是证明定理 5.1 时，定理 III 4.1 所起的作用现在由定理 III 8.1 来承担. 也可以提出一个类似于推论 5.1 的推论，其细节留给读者.

还可以提出基于定理 III 8.1 与基于定理 III 8.4 的与定理 5.1 类似的一些定理，这些也留给读者.

现在考察集合 $\Omega(t, x)$ 为致密时可达集的一些性质. 下面的定理 5.3 及其推论的假设蕴含定理 5.1 的假设成立. 在一些具体例子中，验证定理 5.3 的假设的正确性通常比验证定理 5.1 的假设的正确性容易. 这里，定理 5.3 与定理 5.1 的关系，和定理 III 5.1 与定理 III 4.1 的关系一样.

定理 5.3 设(5.1)中的函数 f 连续；存在致密集 $\mathscr{R}_0\subset\mathscr{R}$，使 $\mathscr{A}^+_T$ 中的一切轨线 ϕ 都位于 $\mathscr{R}_0$ 之内；映射 Ω 在 $\mathscr{R}_0$ 上为 u. s. c. i.,

且对 $\mathscr{R}_0$ 中每一点 (t, x)，集合 $\Omega(t, x)$ 致密．设对 $\mathscr{R}_0$ 中每一点 (t, x)，由(5.2)定义的集合 $\mathscr{Q}(t, x)$是凸的．则对 $[t_0, t_1]$ 中的每个 t，集合 $\mathscr{K}(t, t_0, x_0)$ 在 E^n 中致密，且集合 $\mathscr{A}_T^+(t_0, x_0, t)$ 在 $C[t_0, t]$ 中致密．此外，如果对 $[t_0, t_1]$ 中的每个 t，$\mathscr{A}_T^+(t_0, x_0, t)$ 中的轨线是 $\mathscr{A}_T^+(t_0, x_0, t_1)$中的轨线的限制，则映射 $t \to \mathscr{K}(t, t_0, x_0)$ 是从 $[t_0, t_1]$ 映入 E^n 的具有 Hausdorff 度量的致密子集所构成的空间的连续映射．

此定理的证明类似于定理 III 5.1 的证明，就是证明本定理的假设包含了定理 5.1 的假设．其细节留给读者．

在第三章第五节中我们指出过，在应用中具有重要性的许多特殊类型的问题，都满足定理 III 5.1 的假设．因为定理 5.3 中关于 f、$\mathscr{R}_0$、Ω 和 $\mathscr{Q}$ 的假设，与定理 III 5.1 中涉及状态方程的那些假设相同，故知定理 5.3 适用于第三章第五节的特殊问题．特别，有定理 5.3 的如下推论，它类似于推论 III 5.1 且包含许多在应用中有意义的问题．

推论 5.3.1 设系统(5.1)有如下形式：

$$\frac{dx}{dt} = h(t, x) + B(t, x)u(t), \tag{5.3}$$

其中 h 和 B 在 $\mathscr{R}$ 上连续．假设存在致密集 $\mathscr{R}_0 \subset \mathscr{R}$，使 $\mathscr{A}_T^+$ 中所有轨线 ϕ 都位于 $\mathscr{R}_0$ 之中，映射 Ω 在 $\mathscr{R}_0$ 上为 u. s. c. i.，而且对 $\mathscr{R}_0$ 中每一点 (t, x)，集合 $\Omega(t, x)$ 是致密的且为凸的．则定理 5.3 的结论成立．

对线性系统有如下命题．

推论 5.3.2 设系统(5.1)有形式

$$\frac{dx}{dt} = A(t)x + B(t)u(t), \tag{5.4}$$

其中 A 和 B 在 $[t_0, t_1]$ 上连续．假设映射 Ω 仅与 t 有关，且在 $[t_0, t_1]$ 上为 u. s. c. i.．设每个集合 $\Omega(t)$ 是致密且凸的．则可达集是致密的和凸的，而且映射 $t \to \mathscr{K}(t, t_0, x_0)$ 是连续的．对 $[t_0, t_1]$ 中的每个 t，集合 $\mathscr{A}_T^+(t_0, x_0, t)$ 在 $C[t_0, t_1]$ 中是凸的和致密的．

证明：由于 Ω 是 u. s. c. i. 的且每个集合 $\Omega(t)$ 致密，故由引理 III 5.2 得出，定义在 $[t_0, t_1]$ 上且几乎处处满足 $u(t)\in\Omega(t)$ 的任一可测函数 u 是本性有界的．因为系统关于 x 线性，故知定义在 $[t_0, t_1]$ 上且满足 $u(t)\in\Omega(t)$ 的任一可测函数 u 使方程 (5.4) 产生满足 $\phi(t_0)=x_0$ 的唯一解 ϕ．而且，这个解在整个区间 $[t_0, t_1]$ 上有定义．由引理 III 11.1，至少存在一个定义在 $[t_0, t_1]$ 上的可测函数 u，使 $u(t)\in\Omega(t)$ 几乎处处满足．所以，定义在 $[t_0, t_1]$ 的子区间 $[t_0, t_2]$ 上且满足约束 $u(t)\in\Omega(t)$ 的任一可测函数 u 能够扩展为在 $[t_0, t_1]$ 上有定义，且在 $[t_0, t_1]$ 上满足约束条件的可测函数．因此对 $[t_0, t_1]$ 中每个 t，$\mathscr{A}_T^+(t_0, x_0, t)$ 中的轨线是 $\mathscr{A}_T^+(t_0, x_0, t_1)$ 中轨线的限制．由习题 III 5.1 推得，在本推论的假设下，$\mathscr{A}_T^+$ 中的所有轨线 ϕ 都位于一致密集之中，除了集合 $\mathscr{K}(t, t_0, x_0)$ 与 $\mathscr{A}_T^+(t_0, x_0, t)$ 为凸的论断之外，推论 5.3.2 可从推论 5.3.1 推出．集合 $\mathscr{K}(t_0, x_0, t)$ 与 $\mathscr{A}_T^+(t_0, x_0, t)$ 的凸性是集合 $\Omega(t)$ 的凸性以及(5.4)的解 ϕ 的如下参数变易公式的结果：

$$\phi(t)=\Phi(t)\left\{x_0+\int_{t_0}^{t}\Phi^{-1}(s)B(s)u(s)ds\right\},$$

其中 Φ 是齐次系统 $dx/dt=A(t)x$ 的使 $\Phi(t_0)=I$ 的基本解矩阵．

集合 $\mathscr{Q}^+(t, x)$ 的凸性是定理 5.3 所要求的．如果这些集合非凸，我们就考察与(5.1)相对应的松弛轨线．下面的二个定理涉及松弛轨线的可达集的结构以及它们与原始问题的可达集的关系．

与(5.1)对应的松弛系统是

$$\frac{dx}{dt}=\tilde{f}(t, x, v(t))=\sum_{i=1}^{n+1}p^i(t)f(t, x, u_i(t)). \tag{5.5}$$

因为我们把(5.1)看作是具有 $f^0\equiv0$ 的控制问题的状态方程，故系统(5.5) 给出相应的松弛系统的状态方程．对松弛系统，定义 $\mathscr{A}_R^+$ 为所有松弛对(ψ, v)的集合，此处 v 为松弛控制，ψ 为对应于 v 的松弛轨线，它们满足松弛容许对定义中除 $(t_0, \psi(t_0), t_1, \psi(t_1))\in\mathscr{B}$ 之外的一切要求．松弛问题的集合 $\mathscr{A}_R^+$ 对应于原始问题的

集合 $\mathscr{A}^+$. 用类似的方法可定义关于松弛问题的集合 $\mathscr{K}_R(t, t_0, x_0)$、$\mathscr{A}_R^+(t_0, x_0, t)$、$\mathscr{A}_{RT}^+$ 和 $\mathscr{A}_{RT}^+(t_0, x_0, t)$, 这些集合分别对应于原始系统的集合 $\mathscr{K}(t, t_0, x_0)$、$\mathscr{A}^+(t_0, x_0, t)$、$\mathscr{A}_T^+$ 和 $\mathscr{A}_T^+(t_0, x_0, t)$. 集合 $\mathscr{K}_R(t, t_0, x_0)$ 称为松弛可达集.

定理 5.4 设(5.1)中的函数 f 连续; 存在致密集 $\mathscr{R}_0\subset\mathscr{R}$, 使得 $\mathscr{A}_{RT}^+$ 中的所有松弛轨线 ψ 都位于 $\mathscr{R}_0$ 之中; 映射 Ω 在 $\mathscr{R}_0$ 上为 u. s. c. i., 且对 $\mathscr{R}_0$ 中每一点 (t, x), 集合 $\Omega(t, x)$ 是致密的. 则对 $[t_0, t_1]$ 中每个 t, 集合 $\mathscr{K}_R(t, t_0, x_0)$ 在 E^n 中致密, 且集合 $\mathscr{A}_{RT}^+(t_0, x_0, t)$ 在 $C[t_0, t]$ 中致密. 此外, 如果对 $[t_0, t_1]$ 中每个 t, $\mathscr{A}_{RT}^+(t_0, x_0, t)$ 中的轨线是 $\mathscr{A}_{RT}^+(t_0, x_0, t_1)$ 中的轨线的限制, 则映射 $t\to\mathscr{K}_R(t, t_0, x_0)$ 是从 $[t_0, t_1]$ 映入 E^n 中具有 Hausdorff 度量的致密子集所构成的空间的连续映射.

定理 5.4 是定理 5.3 以及注意到集合

$$\tilde{\mathscr{Q}}(t, x)=\{y: y=\tilde{f}(t, x, \tilde{z}, \pi), \tilde{z}\in[\Omega(t, x)]^{n+1}, \pi\in\Gamma\}$$

为凸的直接结果. 关于 $\tilde{f}$, $\tilde{z}$, π 与 Γ 的定义见(3.1).

注 5.1 对于与(5.3)对应的松弛系统, 可提出对应于推论 5.3.1 的推论, 其假设是 h 与 B 连续, 所有松弛轨线位于一致密集之中, Ω 为 u. s. c. i., 以及每个集合 $\Omega(t, x)$是致密的. 注意, 这里集合 $\Omega(t, x)$ 没有假定是凸的. 我们可断言定理 5.4 的结论成立.

注 5.2 对于与线性系统(5.4)相应的松弛系统, 可以提出相应于推论 5.3.2 的推论. 除了集合 $\Omega(t)$没有假定为凸之外, 关于 A、B 和 Ω 的假定与推论 5.3.2 一样. 我们断言, 松弛可达集为致密的和凸的, 而且映射 $t\to\mathscr{K}_R(t, t_0, x_0)$连续; 集合 $\mathscr{A}_{RT}^+(t_0, x_0, t)$在 $C[t_0, t]$中是致密的和凸的.

习题 5.2 证明集合 $\mathscr{K}_R(t, t_0, x_0)$ 和 $\mathscr{A}_{RT}^+(t_0, x_0, t_1)$ 为凸的断言.

提示: 应用定理 III 5.2、定理 III 7.1 和参数变易公式.

定理 5.5 设 $\mathscr{I}$ 为 E^1 中的区间, $\mathscr{X}$ 为 E^n 中的区间, 令 $\mathscr{R}=\mathscr{I}\times\mathscr{X}$. 假设(5.1)中的函数 f 在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上连续, 这里 $\mathscr{U}$

是 E^n 中的区域. 设 Ω 与 x 无关且在 $\mathscr{I}$ 上为 u. s. c. i.. 对 $\mathscr{I}$ 中的每个 t, 设 $\Omega(t)$ 为致密. 设 $\mathscr{D}$ 和 III(4.1) 中相同, 且存在定义在 $\mathscr{I}$ 上的可积函数 μ, 使得对 $\mathscr{D}$ 中的 (t, x, z) 和 (t, x', z), 有

$$|f(t, x, z)-f(t, x', z)| \leqslant \mu(t)|x-x'|.$$

设 $\mathscr{I}_1=[t_0, t_1]$ 是 $\mathscr{I}$ 内的致密区间, x_0 是 $\mathscr{X}$ 中的一个固定点. 设存在致密区间 $\mathscr{X}_1 \subset \mathscr{X}$, 使得与 (5.1) 相对应的松弛系统的以 (t_0, x_0) 为初始点且在 $\mathscr{I}_1$ 的子区间上定义的所有轨线都位于 $\mathscr{I}_1 \times \mathscr{X}_1$ 之内. 则对 $\mathscr{I}_1$ 中所有 t, 集合 $\mathscr{K}_R(t, t_0, x_0)$ 和 $\mathscr{A}^+_{RT}(t_0, x_0, t)$ 致密, 而且

$$\mathscr{K}_R(t, t_0, x_0)=\mathrm{cl}\, \mathscr{K}(t, t_0, x_0),$$
$$\mathscr{A}^+_{RT}(t_0, x_0, t)=\mathrm{cl}\, \mathscr{A}^+_T(t_0, x_0, t),$$

其中 $\mathscr{A}^+_T(t_0, x_0, t)$ 的闭包是在 $C[t_0, t]$ 空间中取的.

由于(5.1)的每一条轨线也是松弛轨线, 故得

$$\mathscr{K}(t, t_0, x_0) \subseteq \mathscr{K}_R(t, t_0, x_0), \quad \mathscr{A}^+_T(t_0, x_0, t) \subseteq \mathscr{A}^+_{RT}(t_0, x_0, t). \tag{5.6}$$

容易看出, 可把定理 5.4 应用于现在的系统, 从而得知, 对 $[t_0, t_1]$ 中的每个 t, 集合 $\mathscr{K}_R(t, t_0, x_0)$ 在 E^n 中致密, 而且集合 $\mathscr{A}^+_{RT}(t_0, x_0, t)$ 在 $C[t_0, t]$ 中致密. 设 $\mathscr{X}_2$ 是使 $\mathscr{X}_1 \subset \mathscr{X}_2 \subset \mathscr{X}$ 的任一致密集. 从引理 III 5.2 以及 Ω 的 u. s. c. i. 性质与在 $[t_0, t_1]$ 上的致密性可推得, 集合 $\mathscr{D}_1=\{(t, x, z):(t, x) \in \mathscr{I}_1 \times \mathscr{X}_2, z \in \Omega(t)\}$ 是致密的. 因此 f 在 $\mathscr{D}_1$ 上有界, 而且在对函数 μ 作可能的重新定义之后, 定理 4.3 的假设成立. 于是从定理 4.3 得出, $\mathscr{A}^+_R(t_0, x_0, t) \subseteq \mathrm{cl}\, \mathscr{A}^+_T(t_0, x_0, t)$, $\mathscr{K}_R(t, t_0, x_0) \subseteq \mathrm{cl}\, \mathscr{K}(t, t_0, x_0)$. 由于集合 $\mathscr{A}^+_{RT}(t_0, x_0, t)$ 和 $\mathscr{K}_R(t, t_0, x_0)$ 在适当的空间中为闭, 且 (5.6) 成立, 故得定理的结论.

习题 5.3 证明: 如果把 Ω 为 u. s. c. i. 和 $\Omega(t)$ 致密的假设换成下述假设: 映射 $t \to \Omega(t)$ 上半连续; 对 $\mathscr{I} \times \mathscr{X}_1$ 中的每一点 (t, x), 集合

$$\mathscr{Q}(t, x)=\{y: y=f(t, x, z), z \in \Omega(t)\}$$

为闭集; 对 $\mathscr{D}$ 中的所有点 (t, x, z), 有

$$|f(t, x, z)| \leqslant \mu(t).$$

则定理 5.5 仍然正确.

提示：应用定理 5.2.

6. 关于状态变量为线性的系统

就应用中出现的问题来说，存在定理 III 5.1 和本章的定理 5.3 的假设可以认为是合理的. 可是，关于集合 $\mathscr{Q}^+(t, x)$ 的凸性假定，实质上是把我们限制在关于控制为线性且约束集 $\Omega(t, x)$ 为凸的系统. 在本节我们要证明，对关于 x 为线性的系统，在假定映射 Ω 仅与 t 有关的前提下，集合 $\mathscr{Q}^+(t, x)$ 为凸的要求可以去掉. 对线性系统，这意味着不必要求集合 $\Omega(t)$ 为凸. 我们的分析还将得到线性系统的"bang-bang 原理".

我们先证明一条引理，它可以看作是将在定理 6.4 中讨论的"bang-bang 原理"的特殊情形.

引理 6.1 设 E 为直线上的可测子集，且 E 的测度有限. 设 y 是定义在 E 上取值于 E^k 中的函数，而且使得 $y \in L_1(E)$. 设 w 为定义在 E 上的实值可测函数，使得 $0 \leqslant w \leqslant 1$. 则存在可测子集 $F \subset E$，使得

$$\int_E y(t) w(t) dt = \int_F y(t) dt.$$

在引理 6.1 的证明和本节的其它地方，需要用到 Krein-Milman 定理，我们把它叙述成下面的引理 6.2. 这个定理的证明请读者查阅 Dunford-Schwartz 的书([20]，定理 V 8.4，440 页). 引理 6.1 的证明还要用到关于 Banach 空间的弱*拓扑的某些基本事实. 关于这些问题，请读者参考 Dunford-Schwartz 的书([20]，第五章，420～423 页). 这里用到的某些知识的简要易读的论述还可在 Hermes 与 LaSalle 的书([27]，1～22 页)中找到.

引理 6.2 (Krein-Milman). 设 $\mathscr{C}$ 是局部凸拓扑向量空间中的致密凸集，则 $\mathscr{C}$ 是它的极端点[注]集合的闭凸壳.

[注] 原文为 extreme point，常译为"极点". 为避免与传递函数的极点混淆，故译为"极端点". ——译者注.

引理 6.1 的证明：定义映射 T 如下．对 $L_\infty(E)$ 中的每个实函数 ρ，令 $T_\rho=\int_E y(t)\rho(t)dt$．这样定义的映射 T 是从具有弱* 拓扑的 $L_\infty(E)$ 到具有欧氏拓扑的 E^k 的连续映射．令

$$a=\int_E y(t)W(t)dt.$$

则 $T^{-1}(a)$ 是 $L_\infty(E)$ 的凸弱*闭集．以 Σ 表示 $T^{-1}(a)$ 与 $L_\infty(E)$ 中的单位球的交集．由于 $W\in\Sigma$，故集合 Σ 非空．弱*拓扑是 Hausdorff 拓扑．由于 $L_\infty(E)$ 中的单位球是弱*致密的，而 $T^{-1}(a)$ 是弱*闭的，故集合 Σ 为弱*致密，而且它还是凸的．所以，由 Krein–Milman 定理，Σ 有极端点集．我们要证明 Σ 的极端点集是 E 的可测子集 F 的特征函数 χ_F．因为这样一来就有

$$a=\int_E y\chi_F\,dt=\int_F y\,dt,$$

因而引理便得到证明了．

证明的办法是对 y 的值域的维数 k 施行归纳法．我们将给出一般的归纳步骤．证明的第一步，即 $k=1$，实质上与一般步骤相同，故留给读者．

假定引理对 $k-1$ 为真．设 θ 为 Σ 的一个极端点，而 θ 不是某个集合 F 的特征函数．则存在 $\varepsilon>0$ 和测度为正的可测集 $E_1\subset E$，使得对 E_1 中几乎所有的 t，有 $\varepsilon\leqslant\theta(t)\leqslant1-\varepsilon$．设 E_2 和 E_3 是 E_1 的两个子集，E_2、E_3 有正的测度，而且 $E_3=E_1-E_2$．把归纳假定用于 E_2 和 E_3，我们便知存在可测集 F_2 和 F_3，使得 $F_2\subset E_2$，$F_3\subset E_3$，而且

$$\frac{1}{2}\int_{E_j}y^i(t)dt=\int_{F_j}y^i(t)dt,\qquad i=1,\cdots,k-1;\ j=2,3.$$

令 $h_2=2\chi_{F_2}-\chi_{E_2}$，$h_3=2\chi_{F_3}-\chi_{E_3}$，则 h_2、h_3 在 E_1 上不恒等于零，其绝对值不超过 1，且对 $i=1,\cdots,k-1$ 满足

$$\int_{E_2}y^ih_2\,dt=\int_{E_3}y^ih_3\,dt=0.$$

令 $h(t)=\alpha h_2(t)+\beta h_3(t)$，这里 α 与 β 取得使 $|\alpha|<\varepsilon$，$|\beta|<\varepsilon$，

$\alpha^2+\beta^2>0$，而且

$$\int_{E_1} y^k h\, dt=\alpha\int_{E_1} y^k h_2\, dt+\beta\int_{E_1} y^k h_3\, dt=0.$$

这总是可以做得到的．这样定义的函数还满足

$$\int_{E_1} y^i h\, dt=0, \qquad i=1, \cdots, k-1,$$

和 $|h(t)|<\varepsilon$．因此，$0<\theta\pm h<1$，而

$$\int_{E_1} y(\theta\pm h)\, dt=\int_{E_1} y\theta\, dt=\alpha.$$

于是 $\theta+h$ 与 $\theta-h$ 在 Σ 之内．但这样一来，θ 就不可能是 Σ 的极端点，因为它是以 $\theta+h$ 和 $\theta-h$ 为端点的线段的中点．所以，θ 必为某个集合 F 的特征函数．

定理 6.1 假设 $\mathscr{I}$ 是 E^1 中的闭区间，A 是在 $\mathscr{I}$ 上连续的 $n\times n$ 矩阵函数，$\mathscr{U}$ 是 E^m 中的区域，h 是从 $\mathscr{I}\times\mathscr{U}$ 到 E^n 的连续映射．设 Ω 与 x 无关且在 $\mathscr{I}$ 上是上半连续的．对每个 $\mathscr{I}$ 中的 t，设集合 $h(t, \Omega(t))$ 是闭的．令

$$\Delta=\{(t, z): t\in\mathscr{I}, z\in\Omega(t)\},$$

并设存在一个定义在 $\mathscr{I}$ 上的非负可积函数 μ，使得对 Δ 中的一切 (t, z)，有

$$|h(t, z)|\leqslant\mu(t). \tag{6.1}$$

设状态方程为

$$\frac{dx}{dt}=A(t)x+h(t, u(t)), \tag{6.2}$$

初始点为 (t_0, x_0)．则映射 $t\to\mathscr{K}(t, t_0, x_0)$ 是连续的，集合 $\mathscr{K}_R(t, t_0, x_0)$ 与 $\mathscr{K}(t, t_0, x_0)$ 是非空、致密且凸的，并且

$$\mathscr{K}_R(t, t_0, x_0)=\mathscr{K}(t, t_0, x_0).$$

证明：根据引理 III 11.2，至少存在定义在 $\mathscr{I}$ 上的一个控制和相应的轨线.[9] 所以，对所有的 $t_1>t_0$ 和一切 $[t_0, t_1]$ 中的 t，集合 $\mathscr{K}(t, t_0, x_0)$ 与 $\mathscr{K}_R(t, t_0, x_0)$ 非空．而且，每一个定义在区间 $[t_0, t]$ 上的控制能够扩张成定义在整个区间 $[t_0, t_1]$ 上的控制．

由 (6.2) 的解的参数变易公式和 (6.1) 可得出，$\mathscr{A}_1^+$ 中的轨线族等度绝对连续．在定理 5.1 中用于证明映射 $t\to\mathscr{K}(t, t_0, x_0)$ 在

那种情况下连续的论证方法，能够用来证明映射 $t\to\mathscr{K}(t, t_0, x_0)$ 在现在的情形下是连续的.

下面证明集合 $\mathscr{K}_R(t, t_0, x_0)$ 是凸的和致密的. 对应于(6.2)的松弛系统是

$$\frac{dx}{dt}=A(t)x+\sum_{i=1}^{n+1}p^i(t)h(t, u_i(t)), \tag{6.3}$$

这里对 $\mathscr{I}$ 中几乎所有的 t 有

$$\sum_{i=1}^{n+1}p^i(t)=1,\ 0\leqslant p^i(t)\leqslant 1,\ i=1, \cdots, n+1, \tag{6.4}$$

而且 $u_i(t)\in\Omega(t)$. 因此正如 (6.2) 的满足 $\phi(t_0)=x_0$ 的解那样，(6.3)的满足 $\psi(t_0)=x_0$ 的解对 $\mathscr{I}$ 中所有的 t 有定义.

以 Θ 表示 $L_1[t_0, t]$ 中几乎处处满足

$$\theta(s)\in\operatorname{co}h(s, \Omega(s)) \tag{6.5}$$

的函数 θ 的集合. 我们证明集合 Θ 也能表示为 $L_1[t_0, t]$ 中使

$$\theta(s)=\sum_{i=1}^{n+1}p^i(s)h(s, u_i(s)) \tag{6.6}$$

几乎处处成立的一切函数 θ 的集合,其中 $p^1, \cdots, p^{n+1}$ 为满足关系式 (6.4) 且在 $[t_0, t]$ 上可测的实值函数，而 $u_1, \cdots, u_{n+1}$ 是满足 $u_i(s)\in\Omega(s)$ 的可测函数. 形如(6.6)的任一函数满足(6.5)，且根据(6.4)和(6.1)，它是可积的. 反之，由 Caratheodory 定理（定理 III 5.2)，无须关于函数 p^i 与 $u_i(i=1, \cdots, n+1)$ 的可测性的任何断言，满足(6.5)的任一函数 θ 都可以写成(6.6)的形式. 可是，由定理 III 7.1，我们能够断定 p^i 与 u_i 可以取为可测的. 因此，(6.5)与(6.6)给出 Θ 的等价描述.

从上述讨论可以得出，松弛轨线 ψ 也可以定义为满足下式的绝对连续函数：

$$\psi'(s)=A(s)\psi(s)+\theta(s),$$

其中 θ 为 Θ 中的某个函数. 由参数变易公式得：

$$\psi(t)=\Psi(t)\left[x_0+\int_{t_0}^{t}\Psi^{-1}(s)\theta(s)ds\right],$$

其中 Ψ 是齐次方程组 $dx/dt=A(t)x$ 满足 $\Psi(t_0)=I$ 的基本解矩

阵，I 为 $n\times n$ 单位矩阵．因此集合 $\mathscr{K}_R(t, t_0, x_0)$ 可表示为 E^n 中的使

$$x=\Psi(t)\left[x_0+\int_{t_0}^{t}\Psi^{-1}(s)\theta(s)ds\right] \tag{6.7}$$

的点 x 的集合，其中 θ 为 Θ 中的某个函数．因此，要证明 $\mathscr{K}_R(t, t_0, x_0)$ 为凸和致密，只需证明 E^n 中形如

$$\left\{x: x=\int_{t_0}^{t}\Psi^{-1}(s)\theta(s)ds,\ \theta\in\Theta\right\} \tag{6.8}$$

的点集是凸的和致密的．

由(6.5)可直接得出集合 Θ 是 $L_1[t_0, t]$ 的凸子集．$\mathscr{K}_R(t, t_0, x_0)$ 的凸性现在可以从(6.7)推出．

定义从 $L_1[t_0, t]$ 到 E^n 的映象 T 如下：

$$T_\rho=\int_{t_0}^{t}\Psi^{-1}(s)\rho(s)ds,\ \rho\in L_1[t_0, t].$$

映象 T 是从 $L_1[t_0, t]$ 到 E^n 的(按)强拓扑的连续映象．所以它也是从 $L_1[t_0, t]$ 到 E^n 的(按)弱拓扑的连续映象．因此 $L_1[t_0, t]$ 中的弱致密集的象在 E^n 中致密．所以，要证明集合(6.8)，从而 $\mathscr{K}_R(t, t_0, x_0)$ 是致密的，只需证明 Θ 是弱致密的．现在就来证明这一事实．

我们已指出集合 Θ 是 $L_1[t_0, t]$ 的凸子集．正如下面的论证所表明的，集合 Θ 在 $L_1[t_0, t]$ 中还是闭的．设 $\{\theta_k\}$ 是 Θ 的在 $L_1[t_0, t]$ 中收敛于元素 θ_0 的序列．则存在一个子序列(重新记作 $\{\theta_k\}$)，使得除了一个测度为零的子集 E_0 中的点外，在 $[t_0, t]$ 的所有点上有 $\theta_k\to\theta_0$．以 E_{-1} 记使 $\mu(s)=\infty$ 的点 s 的集合，这里 μ 和(6.1)中的一样．对每个整数 k，以 E_k 表示使(6.6)不成立的点的集合．则 $E=\bigcup\limits_{k=-1}^{\infty}E_k$ 的测度为零．设 s 为 $[t_0, t_1]$ 上的不在 E 中的点，则存在一个整数的子列，它仍记作 $\{k\}$，使得对所有 $i=1, \cdots, n+1$，序列

$$\{p_k^i(s)\}_{k=1}^{\infty} \text{和} \{h(s, u_{ik}(s))\}_{k=1}^{\infty}$$

收敛．设 $p^i(s)=\lim\limits_{k\to\infty}p_k^i(s)$，$i=1, \cdots, n+1$．则 $0\leqslant p^i(s)\leqslant 1$，且

$\sum p^i(s)=1$. 由于 $h(s, \Omega(s))$ 为闭，故知对各个 $i=1, \cdots, n+1$，$\lim\limits_{k\to\infty} h(s, u_{ik}(s))$ 在 $h(s, \Omega(s))$ 之中. 因此 θ_0 在点 s 满足 (6.5). 因为 s 是 $[t_0, t]-E$ 中的任意点，故得(6.5)几乎处处成立，且 Θ 在 $L_1[t_0, t]$ 中是闭的.

由于 Θ 是 $L_1[t_0, t]$ 中的凸闭集，故它也是弱闭的. 从 (6.6)、(6.4)和(6.1)可以得出，对 Θ 中的所有 θ，$|\theta(s)|\leqslant\mu(s)$. 因此 Θ 中的函数具有等度绝对连续的积分. 所以 Θ 是弱序列条件致密的 ([20] 推论 IV. 8.11, 294 页). 因此根据 Eberlein-Smulian 定理 ([20] 定理 V. 6.1)，Θ 的闭包是弱致密的. 由于 Θ 弱闭，故它也是弱致密的. 因此 $\mathscr{K}_R(t, t_0, x_0)$ 是致密的.

由于系统 (6.2) 的轨线也是松弛轨线，故有 $\mathscr{K}(t, t_0, x_0)\subseteq\mathscr{K}_R(t, t_0, x_0)$. 所以要完成定理的证明，只需证明 $\mathscr{K}_R(t, t_0, x_0)\subseteq\mathscr{K}(t, t_0, x_0)$.

设 x 是 $\mathscr{K}_R(t, t_0, x_0)$ 中的元素. 则 x 由(6.7)给出. 设

$$a=\Psi^{-1}(t)x-x_0.$$

点 a 在由(6.8)定义的集合之内. 集合 $T^{-1}(a)$ 在 $L_1[t_0, t]$ 中是弱闭且凸的. 以 Σ 记 $T^{-1}(a)$ 和 Θ 的交集. 则 Σ 是弱闭且凸的. 根据 Krein-Milman 定理(引理 6.2)，Σ 有极端点 θ_0. 因为 $\theta_0\in\Theta$，故知 θ_0 可用(6.6)右边表示. 现证明不存在 $[t_0, t]$ 的测度为正的子集 E，能使得对数集 $\{1, \cdots, n+1\}$ 中的某个 i 和某个 $\varepsilon>0$，在 E 上有 $0<\varepsilon\leqslant p^i(s)\leqslant 1-\varepsilon$. 这个断言蕴含 θ_0 可写成

$$\theta_0(s)=h(s, u(s)), \tag{6.9}$$

其中 u 是 $[t_0, t]$ 上的可测函数，它满足 $u(s)\in\Omega(s)$. 一旦建立了式(6.9)，定理便证明完毕，因为这样一来就有

$$a=\Psi^{-1}(t)x-x_0=\int_{t_0}^{t}\Psi^{-1}(s)h(s, u(s))\,ds,$$

这就是说，$x\in\mathscr{K}(t, t_0, x_0)$.

假设 θ_0 可表示为形式 (6.6)，且存在 $\varepsilon>0$ 和可测集 $E\subseteq[t_0, t]$，使得 $\text{meas}(E)>0$，而且在 E 上有 $\varepsilon\leqslant p^1(s)\leqslant 1-\varepsilon$[注]. 则

【注】 为明确起见，不妨设 $i=1$. ——译者注.

至少有另一个上标 i，在 E 上也有 $\varepsilon\leqslant p^i(s)\leqslant 1-\varepsilon$. 为确定起见，假定 $i=2$ 是一个这样的上标. 由于 a 属于由 (6.8) 定义的集合，从(6.6)得

$$a=\int_{t_0}^{t}\left\{\Psi^{-1}(s)\sum_{i=1}^{n+1}p^i(s)h(s,\ u_i(s))\right\}ds$$
$$=\int_{t_0}^{t}\left\{\sum_{i=1}^{n+1}p^i(s)\left(\Psi^{-1}(s)h(s,\ u_i(s))\right)\right\}ds.$$

下面我们取 $w=\frac{1}{2}$，并对此 w 以及由式

$$y(s)=\begin{pmatrix}\Psi^{-1}(s)h(s,\ u_1(s))\\ \Psi^{-1}(s)h(s,\ u_2(s))\end{pmatrix}$$

确定的在 E^{2n} 中取值的 $L_1(E)$ 中的向量函数 y 应用引理 6.1，我们得出：存在集合 $F\subseteq E$，使得

$$\frac{1}{2}\int_E\Psi^{-1}(s)h(s,\ u_i(s))ds$$
$$=\int_F\Psi^{-1}(s)h(s,\ u_i(s))ds,\qquad i=1,\ 2.$$

以 χ_F 表示集合 F 的特征函数，以 χ_E 表示集合 E 的特征函数. 设函数 γ 由下式定义：

$$\gamma(s)=2\chi_F(s)-\chi_E(s).$$

在集合 E 上函数 γ 的绝对值等于 1. 令

$$\pi_1^1=p^1+\varepsilon\gamma,\ \pi_1^2=p^2-\varepsilon\gamma,\ \pi_1^i=p^i,\qquad i=3,\ \cdots,\ n+1,$$
$$\pi_2^1=p^1-\varepsilon\gamma,\ \pi_2^2=p^2+\varepsilon\gamma,\ \pi_2^i=p^i,\qquad i=3,\ \cdots,\ n+1.$$

则 $0\leqslant\pi_1^i\leqslant 1$，$0\leqslant\pi_2^i\leqslant 1$，$i=1,\ \cdots,\ n+1$. 而且 $\sum\pi_1^i=1$，$\sum\pi_2^i=1$.

令
$$\theta_1(s)=\sum_{i=1}^{n+1}\pi_1^i(s)h(s,\ u_i(s)),$$
$$\theta_2(s)=\sum_{i=1}^{n+1}\pi_2^i(s)h(s,\ u_i(s)).$$

则 θ_1 与 θ_2 在 Θ 中，而且

$$a=\int_{t_0}^{t}\Psi^{-1}(s)\theta_1(s)ds=\int_{t_0}^{t}\Psi^{-1}(s)\theta_2(s)ds.$$

因此 θ_1 与 θ_2 都在 Σ 中. 但是 $\theta_0=(\theta_1+\theta_2)/2$，这与 θ_0 是极端点的假定矛盾，定理得证.

当约束集 $\Omega(t)$ 致密时，有下面的作为定理 6.1 的推论的结果.

定理 6.2 设 $\mathscr{I}$ 是 E^1 中的致密区间，A 是在 $\mathscr{I}$ 上连续的 $n\times n$ 矩阵函数. 设 $\mathscr{U}$ 是 E^m 中的区域，h 是从 $\mathscr{I}\times\mathscr{U}$ 到 E^n 的连续映射. 假设 Ω 与 x 无关且在 $\mathscr{I}$ 上是 u. s. c. i. 的，对每个 $\mathscr{I}$ 中的 t，设集合 Ω 致密. 则对带初始点 (t_0, x_0) 和满足约束条件 $u(t)\in\Omega(t)$ 的系统(6.2)，下述事实成立. 对 $\mathscr{I}$ 中的一切 t, $t>t_0$，可达集 $\mathscr{K}(t, t_0, x_0)$ 是非空、致密和凸的，而且

$$\mathscr{K}(t, t_0, x_0)=\mathscr{K}_R(t, t_0, x_0),$$

映象 $t\to\mathscr{K}(t, t_0, x_0)$ 在任一区间 $[t_0, t_1]\subset\mathscr{I}$ 上连续.

由于 $\mathscr{I}$ 致密且 Ω 在 $\mathscr{I}$ 上是 u. s. c. i. 的，故从引理 III 5.2 推得，定理 6.1 的叙述中定义的集合 Δ 是致密的. 因为 h 连续，h 在 Δ 上有界. 因此(6.1)成立，因而定理 6.1 的假设成立.

对关于状态变量为线性的系统，有下面的不要求集合 $\mathscr{Q}^+(t, x)$ 为凸的存在定理.

定理 6.3 设 $\mathscr{I}=[t_0, t_1]$ 是 E^1 中的固定致密区间，$\mathscr{U}$ 为 E^m 中的区域，h 是从 $\mathscr{I}\times\mathscr{U}$ 到 E^n 的连续映射，A 是在 $\mathscr{I}$ 上连续的 $n\times n$ 矩阵函数. 设映射 a_0: $\mathscr{I}\to E^n$ 与 h_0: $\mathscr{I}\times\mathscr{U}\to E^1$ 是连续的；Ω 与 x 无关，且在 $\mathscr{I}$ 上为 u. s. c. i.. 设对 $\mathscr{I}$ 中每一个 t，集合 $\Omega(t)$ 致密. 又设 g 为实值函数，在 E^n 中连续，而 x_0 固定. 则满足端点条件 $\mathscr{B}=\{(t_0', x_0', t_1', x_1'): t_0'=t_0, x_0'=x_0, t_1'=t_1\}$，控制约束 Ω 和系统方程

$$\frac{dx}{dt}=A(t)x+h(t, u(t)),$$

且使泛函

$$J(\phi, u)=g(\phi(t_1))+\int_{t_0}^{t_1}(\langle a_0(t), \phi(t)\rangle+h_0(t, u(t)))dt$$

取极小值的问题有解.

注 6.1 从引理 III 11.2 和注 11.1 得到，这时容许对的集合非空.

注 6.2 对线性系统有 $h(t, z)=B(t)z$，定理 6.2 使我们可略去至今为止所用的约束 $\Omega(t)$ 为凸的要求.

证明：通过引入附加坐标 x^0 和状态方程 $dx^0/dt=\langle a_0(t), x\rangle+h_0(t, u(t))$，不失一般性可假定 a_0、h_0 恒等于零(见第二章第四节)．由定理 6.2，集合 $\mathscr{K}(t_1, t_0, x_0)$ 致密．因为 g 在可达集 $\mathscr{K}(t_1, t_0, x_0)$ 上连续，它在 $\mathscr{K}(t_1, t_0, x_0)$ 的某个点，比如说 x_1^* 上达到它的极小值．按定义，$\mathscr{K}(t_1, t_0, x_0)$ 恰好由等于 $\phi(t_1)$ 的那些点组成，这里 ϕ 为某容许轨线．设 ϕ^* 对应于 x_1^*．则对一切容许对 (ϕ, u)，$J(\phi^*, u^*)=g(\phi^*(t_1))\leqslant g(\phi(t_1))=J(\phi, u)$，定理证毕.

定理 6.1 的另一推论是所谓 "bang-bang 原理"；它包含在下面的定理 6.4 中．使用这一术语的理由和这个原理在应用中的意义，将在给出定理 6.4 的证明之后讨论.

如果 $\mathscr{C}$ 是 E^m 中的致密凸集，则以 $\mathscr{C}_e$ 表示 $\mathscr{C}$ 的极端点的集合．由 Krein-Milman 定理，$\mathscr{C}_e$ 非空且 $\mathscr{C}=\mathrm{cl\,co}(\mathscr{C}_e)$.

定理 6.4 设 $\mathscr{I}=[t_0, t_1]$ 是 E^1 中的致密区间，A 是 $\mathscr{I}$ 上的 $n\times n$ 连续矩阵，B 是 $\mathscr{I}$ 上的 $n\times m$ 连续矩阵．设 $\mathscr{C}$ 是 E^m 中的致密凸集，其极端点集合 $\mathscr{C}_e$ 是闭的．以 $\mathscr{K}(t_1, t_0, x_0)$ 表示系统

$$\frac{dx}{dt}=A(t)x+B(t)u(t) \tag{6.10}$$

的具初始点 (t_0, x_0) 和控制约束 $u(t)\in\mathscr{C}$、在 t_1 的可达集，而 $\mathscr{K}_e(t_1, t_0, x_0)$ 表示系统 (6.10) 的带初始点 (t_0, x_0) 和控制约束 $u(t)\in\mathscr{C}_e$ 的可达集，则 $\mathscr{K}_e(t_1, t_0, x_0)$ 非空，且

$$\mathscr{K}(t_1, t_0, x_0)=\mathscr{K}_e(t_1, t_0, x_0).$$

证明：因为由 $u(t)=z_0$ (这里 z_0 为 $\mathscr{C}_e$ 中的任意点) 确定的函数对于带初始点 (t_0, x_0) 及控制约束 $u(t)\in\mathscr{C}_e$ 的系统 (6.10) 是容许控制，故知 $\mathscr{K}_e(t_1, t_0, x_0)$ 非空．由于 $\mathscr{C}$ 为致密凸集，故 $\mathscr{C}=\mathrm{cl\,co}(\mathscr{C}_e)$．因为 $\mathscr{C}$ 致密，故 $\mathscr{C}_e$ 有界．又按假设 $\mathscr{C}_e$ 是闭的．故集合 $\mathscr{C}_e$ 是致密的．由定理 III 5.2，$\mathrm{co}(\mathscr{C}_e)$ 中的任一点都可写成至多 $(n+1)$ 个 $\mathscr{C}_e$ 中的点的凸组合．因而 $\mathrm{co}(\mathscr{C}_e)$ 是 $\Gamma\times(\mathscr{C}_e)^{n+1}$ 的连续象，这里 Γ 与 (3.1) 中一样．因此 $\mathrm{co}(\mathscr{C}_e)$ 是致密的，所以

$\mathscr{C}=\mathrm{clco}(\mathscr{C}_e)=\mathrm{co}(\mathscr{C}_e)$. 因而使 $u(t)\in\mathscr{C}$ 的任一控制都可以写成

$$u(t)=\sum_{i=1}^{m+1}p^i(t)u_i(t),$$

其中 $0\leqslant p^i(t)\leqslant 1$, $\sum p^i(t)=1$, $u_i(t)\in\mathscr{C}_e$. 由定理 III 7.1, 函数 p^i 和 u_i 可取为可测的. 因此集合 $\mathscr{K}(t_1, t_0, x_0)$ 包含在与 $\mathscr{K}_e(t_1, t_0, x_0)$ 对应的松弛可达集 $\mathscr{K}_{eR}(t_1, t_0, x_0)$ 之内.

反之, 带控制约束 $u(t)\in\mathscr{C}_e$ 的系统 (6.10) 的每一松弛控制, 都是带控制约束 $u(t)\in\mathscr{C}$ 的系统 (6.10) 的控制. 因此据上面讨论得 $\mathscr{K}(t_1, t_0, x_0)=\mathscr{K}_{eR}(t_1, t_0, x_0)$. 容易验证, 带初始点 (t_0, x_0) 和控制约束 $u(t)\in\mathscr{C}_e$ 的系统(6.10)满足定理 6.1 的假设. 因此, $\mathscr{K}_{eR}(t_1, t_0, x_0)=\mathscr{K}_e(t_1, t_0, x_0)$. 于是, 本定理得证.

在许多应用中, 约束集 $\mathscr{C}$ 是 E^m 中的致密凸多面体, 或者甚至是立方体. 极端点集合 $\mathscr{C}_e$ 是多面体的顶点的集合, 所以是闭的. 在这种情况下, 定理 6.4 表明, 如果具有值 $u(t)\in\mathscr{C}$ 的控制 u 把系统从时刻 t_0 的点 x_0 转移到时刻 t_1 的点 x_1, 则存在一个具有在 $\mathscr{C}_e$ 中取值 $u_e(t)$ 的控制 u_e, 也可以实现这种转移. 因此, 在设计控制系统时, 设计者只需要考虑与 $\mathscr{C}$ 的顶点相对应的有限个控制位置. 用术语 "bang-bang" 描述在顶点取值的控制, 起源于 $\mathscr{C}$ 是一维区间的情形. 在这种情形, 具有 $u_e(t)\in\mathscr{C}_e$ 的控制 u_e 是取值 $+1$ 和 -1 的控制. 这样的控制代表控制装置的极端位置, 所以在工程的行话里常称之为 "bang-bang" 控制. 在控制理论的文献里, 这一术语已被引入诸如定理 6.4 的一些定理中.

第五章　最大值原理及其某些应用

1. 引　　言

本章我们叙述最大值原理，并用它去刻画几类重要问题的最优控制. 最大值原理的证明将在下一章给出.

在第2节里，我们用动态规划理论来推导最大值原理. 虽然这些论证数学上是正确的，但它的假定却使最有意义的问题被排除在外. 这一节的目的是把定理叙述得合乎情理一些，并再给定理作一些剖析和说明. 从逻辑推理的观点来看，除了一个概念之外，第2节可以略去. 这个概念就是最优综合或最优反馈控制，它是在第2节中引入，而在第9节中还要用到.

在第3节里，我们给出Lagrange型控制问题的最大值原理的精确叙述. 对其它控制问题(诸如第二章里所论及的各种形式)，其最大值原理的表达形式都将放在习题里去讨论. 对于许多重要的特殊情形，经常可以给出最优对的更精确的描述，其中的一部分也放在习题里讨论. 这一节的习题是对一般理论的重要补充.

在第4节里，我们使用最大值原理和我们的存在性定理之一来确定一个具体例子的最优对，其目的是为了说明如何运用最大值原理，以及在大量的问题中预计会遇到一些什么样的困难.

本章剩下的几节则专门致力于最大值原理对于一些特殊类型问题的应用. 第5节证明如何从最大值原理推出古典变分法中的一阶必要条件，而在该节习题中，我们研究变分法里的经典Bolza问题与最优控制问题的关系. 第6节研究关于状态变量是线性的控制问题. 第7节专门讨论线性问题；第8节进一步专门讨论线性时间最优问题. 关于这些问题的标准结果，往往可以作为最大

值原理的相当简单的推论而得到，读者由此可明显地看出最大值原理的威力.

在第9节里，我们讨论所谓"线性对象二次代价准则"问题.这里，我们再一次从最大值原理得到最优对的标准特征，同时还证明了必要条件是充分的，并得到最优控制的标准综合.

2. 最大值原理的动态规划推导

本节我们要在很多限制的前提下推出最大值原理. 这些前提，在论证过程中按其需要将给以清楚的说明. 读者务必注意，本节所做的假设在一些重要的问题中经常是不满足的. 虽然有些论证在较少的限制前提下也可以成立，但我们这里不打算这样做.本节的目的是为了导出下面第3节定理3.1所给出的最大值原理的精确描述，并对最大值原理为什么似乎从表面上看来是正确的这一点，给予某些剖析.

设 $\mathscr{R}_1$ 是 (t, x)-空间的区域，而 $\mathscr{R}$ 是 $\mathscr{R}_1$ 的子域，使得 $\mathscr{R}$ 的闭包含于 $\mathscr{R}_1$ 中. 对 $\mathscr{R}$ 中的每一点 (τ, ξ)，考虑如下问题：

在满足状态方程组

$$\frac{dx}{dt}=f(t, x, u(t)) \tag{2.1}$$

和控制约束 $u(t)\in\Omega(t)$ 以及端点条件

$$(t_0, \phi(t_0))=(\tau, \xi), \qquad (t_1, \phi(t_1))\in\mathscr{T}$$

的条件下，求泛函

$$J(\phi, u)=g(t_1, \phi(t_1))+\int_{t_0}^{t_1} f^0(t, \phi(t), u(t))dt \tag{2.2}$$

的极小值.

我们假设终端集合 $\mathscr{T}$ 是 q 维的 $C^{(1)}$ 流形，其中 $0\leqslant q\leqslant n$，并且 $\mathscr{T}$ 是 $\mathscr{R}$ 的边界的一部分. 见图1. 为了简单起见，我们又假设 $\mathscr{T}$ 可以用单一坐标补缀来表示. 亦即假设 $\mathscr{T}$ 是由所有的形如 (t_1, x_1) 的点所组成的，其中

$$t_1=T(\sigma), \quad x_1=x(\sigma), \tag{2.3}$$

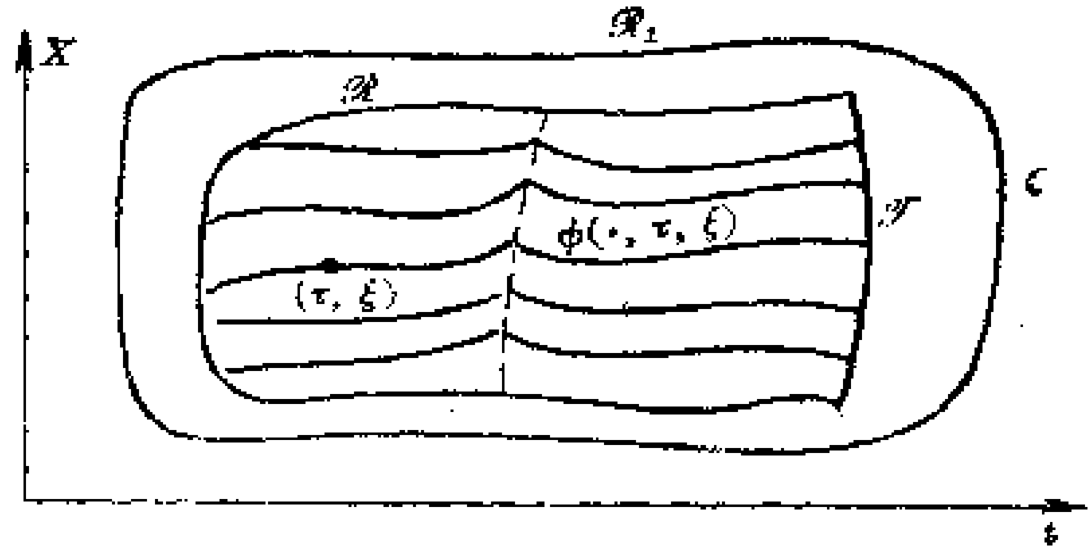

图 1

这里 $T(\sigma)$ 和 $x(\sigma)$ 是定义在 E^q 空间里的某一开的平行六面体 Σ 中的 $C^{(1)}$ 函数. 且假设映射(2.3)的雅可比矩阵

$$\frac{\partial(T,x)}{\partial\sigma}=\begin{pmatrix}\dfrac{\partial T}{\partial\sigma^1}\cdots\dfrac{\partial T}{\partial\sigma^q}\\ \dfrac{\partial x^1}{\partial\sigma^1}\cdots\dfrac{\partial x^1}{\partial\sigma^q}\\ \vdots \qquad\quad \vdots\\ \dfrac{\partial x^n}{\partial\sigma^1}\cdots\dfrac{\partial x^n}{\partial\sigma^q}\end{pmatrix}$$

在 Σ 的所有点上秩数为 q. 我们假定(2.2)中的函数 g 在 $\mathscr{T}$ 的邻域中有定义且属于 $C^{(1)}$, 而 f^0 和 f 是 $\mathscr{G}_1=\mathscr{R}_1\times\mathscr{U}$ 上的 $C^{(1)}$ 函数. 注意, 此处假定约束映射 Ω 与 x 无关而只依赖于 t.

假定对 $\mathscr{R}$ 里的每一点 (τ,ξ), 上面考虑的问题有唯一解. 我们用 $\phi(\cdot,\tau,\xi)$ 表示以 (τ,ξ) 为初始点的唯一最优轨线, 相应的唯一最优控制用 $u(\cdot,\tau,\xi)$ 表示. 假设函数 $u(\cdot,\tau,\xi)$ 是分段连续的, 在不连续点 t_d 上 $u(\cdot,\tau,\xi)$ 的值是它的右极限; 因此 $u(t_d,\tau,\xi)=u(t_d+0,\tau,\xi)$. 在轨线上的点 (t,x) 处满足关系 $x=\phi(t,\tau,\xi)$. 特别有

$$\xi=\phi(\tau,\tau,\xi).$$

在时刻 t 最优控制的值是 $u(t)=u(t,\tau,\xi)$.

对于 $\mathscr{R}$ 里的每一点 (τ,ξ), 令 $W(\tau,\xi)$ 表示由唯一最优对 $(\phi(\cdot,\tau,\xi),u(\cdot,\tau,\xi))$ 给出的泛函(2.2)的值. 因此, 若用

$\mathscr{A}(\tau,\ \xi)$表示以$(\tau,\ \xi)$为初始点的问题的容许对$(\phi,\ u)$的集合，则

$$W(\tau,\ \xi)=\min\{J(\phi,\ u):(\phi,\ u)\in\mathscr{A}(\tau,\ \xi)\}. \tag{2.4}$$

$W(\tau,\ \xi)$叫做所论问题的值函数.

设$\tau_1>\tau$，而$(\tau_1,\ \xi_1)$是最优轨线$\phi(\cdot,\ \tau,\ \xi)$上的点，则$\xi_1=\phi(\tau_1,\ \tau,\ \xi)$. 我们可以断定，从点$(\tau_1,\ \xi_1)$出发的最优对是由$(\phi(\cdot,\ \tau,\ \xi),\ u(\cdot,\ \tau,\ \xi))$给出的. 亦即对$t\geqslant\tau_1$，有

$$\begin{aligned}\phi(t,\ \tau,\ \xi)&=\phi(t,\ \tau_1,\ \xi_1),\\ u(t,\ \tau,\ \xi)&=u(t,\ \tau_1,\ \xi_1).\end{aligned} \tag{2.5}$$

换言之，最优轨线具有这样的性质，从它上面任一点出发的轨线仍然是最优的. 为了弄清楚这一点，我们写出

$$\begin{aligned}W(\tau,\ \xi)=\int_\tau^{\tau_1}f^{0*}(t,\ \tau,\ \xi)dt+\int_{\tau_1}^{t_1}f^{0*}(t,\ \tau,\ \xi)dt&\\ +g(t_1,\ \phi(t_1,\ \tau,\ \xi)),&\end{aligned} \tag{2.6}$$

其中

$$f^{0*}(t,\ \tau,\ \xi)=f^0(t,\ \phi(t,\ \tau,\ \xi),\ u(t,\ \tau,\ \xi)). \tag{2.7}$$

如果对$t>\tau_1$，$(\phi(\cdot,\ \tau,\ \xi),\ u(\cdot,\tau,\ \xi))$不是最优的，则由$(\tau_1,\ \xi_1)$代替$(\tau,\ \xi)$的(2.4)式，以及最优对的唯一性假设，我们将有$W(\tau_1,\ \xi_1)$严格小于(2.6)式右边后两项之和. 因此，定义控制

$$u(t)=\begin{cases}u(t,\ \tau,\ \xi), & \text{当 } \tau\leqslant t<\tau_1,\\ u(t,\ \tau_1,\ \xi_1), & \text{当 } \tau_1\leqslant t\leqslant t_1.\end{cases}$$

对如此定义的控制u和相应的轨线ϕ会使$J(\phi,\ u)<W(\tau,\ \xi)$，这与(2.4)式相矛盾. 顾及$u(\cdot,\ \tau_1,\ \xi_1)$是以$(\tau_1,\ \xi_1)$为初始点的问题的最优控制，于是(2.5)式成立.

在$\mathscr{R}$上我们定义如下的函数U:

$$U(\tau,\ \xi)=u(\tau,\ \tau,\ \xi).$$

如果在(2.5)的第二个方程中令$t=\tau_1$，并根据函数U的定义，我们得出对一切$\tau_1>\tau$

$$u(\tau_1,\ \tau,\ \xi)=U(\tau_1,\ \xi_1), \tag{2.8}$$

其中$\xi_1=\phi(\tau_1,\ \tau,\ \xi)$. 因此，$U$在$\mathscr{R}$里的每一点$(\tau,\ \xi)$上的取值$U(\tau,\ \xi)$，就是与通过这一点的唯一的最优轨线相对应的唯一最优

控制函数的取值．函数 U 叫做最优控制的综合或最优综合函数，也叫做最优反馈控制．

现在假设函数 W 在 $\mathscr{R}$ 上是 $C^{(1)}$ 类的，我们要导出 W 必须满足的偏微分方程．考虑 $\mathscr{R}$ 里的点 (τ, ξ) 和区间 $[\tau, \tau+\Delta t]$，此处 $\Delta t>0$．令 v 是定义于 $[\tau, \tau+\Delta t]$ 上的连续控制，满足 $v(t)\in\Omega(t)$．我们假定 $\Delta t>0$ 充分小，使得状态方程(2.1)中的 $u(t)$ 用 $v(t)$ 代替时，它在 $[\tau, \tau+\Delta t]$ 上有解 $\psi(t)$，而且有 $\psi(\tau)=\xi$．令 $\Delta x=\psi(\tau+\Delta t)-\psi(\tau)$，因此控制 $v(t)$ 在时间间隔 $[\tau, \tau+\Delta t]$ 内把系统从 ξ 转移到 $\xi+\Delta x$．对于 $t\geqslant\tau+\Delta t$，我们用以 $(\tau+\Delta t, \xi+\Delta x)$ 为起点的最优控制 $u(\cdot, \tau+\Delta t, \xi+\Delta x)$，相应的轨线为 $\phi(\cdot, \tau+\Delta t, \xi+\Delta x)$．令 $\tilde{u}$ 表示在 $[\tau, \tau+\Delta t]$ 上用控制 $v(t)$，而随后用 $u(\cdot, \tau+\Delta t, \xi+\Delta x)$ 所得到的控制，$\tilde{\phi}$ 表示相应于 $\tilde{u}$ 的轨线，则 $(\tilde{\phi}, \tilde{u})\in\mathscr{A}(\tau, \xi)$，而且

$$
\begin{aligned}
W(\tau, \xi)\leqslant J(\tilde{\phi}, \tilde{u})=\int_{\tau}^{\tau+\Delta t} f^0(s, \psi(s), v(s))ds
&+\int_{\tau+\Delta t}^{t_1} f^{0*}(s, \tau+\Delta t, \xi+\Delta x)ds \\
&+g(t_1, \phi(t_1, \tau+\Delta t, \xi+\Delta x)),
\end{aligned}
$$

其中 f^{0*} 是(2.7)式定义的函数．因此，上式右边后两项之和等于 $W(\tau+\Delta t, \xi+\Delta x)$．于是

$$
W(\tau+\Delta t, \xi+\Delta x)-W(\tau, \xi)\geqslant-\int_{\tau}^{\tau+\Delta t} f^0(s, \psi(s), v(s))ds.
$$

因为 W 是 $\mathscr{R}$ 上的 $C^{(1)}$ 函数，我们能对上述不等式左边应用 Taylor 定理，从而得到

$$
\begin{aligned}
&W_\tau(\tau, \xi)\Delta t+\langle W_\xi(\tau, \xi), \Delta x\rangle+o(|(\Delta t, \Delta x)|) \\
&\geqslant-\int_{\tau}^{\tau+\Delta t} f^0(s, \psi(s), v(s))ds,
\end{aligned}
\tag{2.9}
$$

其中 (W_τ, W_ξ) 表示 W 的偏导数向量；而 $o(|(\Delta t, \Delta x)|)/|(\Delta t, \Delta x)|\to 0$，当 $|(\Delta t, \Delta x)|\to 0$．从关系式

$$
\frac{\Delta x}{\Delta t}=\frac{1}{\Delta t}\int_{\tau}^{\tau+\Delta t} f(s, \psi(s), v(s))ds,
$$

以及 f，ψ 和 v 的连续性，可得

$$\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t}=f(\tau,\ \psi(\tau),\ v(\tau))=f(\tau,\ \xi,\ v(\tau)).$$

因此，如果在(2.9)式的两边同除以 Δt，然后让 $\Delta t\to 0$，即得

$$W_\tau(\tau,\ \xi)+\langle W_\xi(\tau,\ \xi),\ f(\tau,\ \xi,\ v(\tau))\rangle\geqslant -f^0(\tau,\ \xi,\ v(\tau)). \tag{2.10}$$

在上面的分析中，如果在 $[\tau,\ \tau+\Delta t]$ 上采用控制 $v(s)=u(s,\ \tau,\ \xi)$，则在论证的每一步中等号成立．因此利用(2.8)式，我们可得关系

$$W_\tau(\tau,\ \xi)=-f^0(\tau,\ \xi,\ U(\tau,\ \xi))-\langle W_\xi(\tau,\ \xi),\ f(\tau,\ \xi,\ U(\tau,\ \xi))\rangle. \tag{2.11}$$

现在进一步假设约束映射 Ω 是充分光滑的，使得对每一向量 $z\in\Omega(t)$，存在定义于某区间 $[\tau,\ \tau+\Delta t]$ $(\Delta t>0)$ 上的连续函数 v，具有 $v(\tau)=z$，而在 $[\tau,\ \tau+\Delta t]$ 上 $v(s)\in\Omega(s)$．特别，如果 Ω 是常值映射，即对一切 t，$\Omega(t)=\mathscr{C}$．那么，对一切 $s\in[\tau,\ \tau+\Delta t]$，我们可以取 $v(s)=z$．在刚才对 Ω 所做的假设下，我们联合(2.10)和(2.11)两式可以得出关系

$$W_\tau(\tau,\ \xi)=\max_{z\in\Omega(\tau)}\{-f^0(\tau,\ \xi,\ z)-\langle W_\xi(\tau,\ \xi),\ f(\tau,\ \xi,\ z)\rangle\}, \tag{2.12}$$

它在 $z=U(\tau,\ \xi)$ 处达到最大值．方程(2.12)有时叫做 Bellman 方程．方程(2.11)叫做 Hamilton-Jacobi 方程．

方程(2.11)和(2.12)可以写得更紧凑些．首先，用公式

$$H(t,\ x,\ z,\ p^0,\ p)=p^0f^0(t,\ x,\ z)+\langle p,\ f(t,\ x,\ z)\rangle \tag{2.13}$$

在 $E^1\times E^n\times E^m\times E^1\times E^n$ 上定义一实值函数 H．现在，如果不用 $(\tau,\ \xi)$ 而改用 $(t,\ x)$ 表示 $\mathscr{R}$ 里的任意点，则(2.11)可以用 H 的记号写成

$$W_t(t,\ x)=H(t,\ x,\ U(t,\ x),\ -1,\ -W_x(t,x)). \tag{2.14}$$

方程(2.12)可写成

$$W_t(t,\ x)=\max_{z\in\Omega(t)}\{H(t,\ x,\ z,\ -1,\ -W_x(t,\ x))\}. \tag{2.15}$$

现在假设函数 W 是属于 $C^{(2)}$ 类的. 在此附加的假设之下,我们将导出 Pontryagin 最大值原理. 设 (τ,ξ) 仍为 $\mathscr{R}$ 里的固定点,考虑由公式

$$F(x)=W_t(\tau,x)+f^0(\tau,x,U(\tau,\xi))\\ +\langle W_x(\tau,x),f(\tau,x,U(\tau,\xi))\rangle \tag{2.16}$$

所定义的 $\mathscr{R}$ 上的函数 $F(x)$,其中 W_t 是 W 关于时间 t 的偏导数,而 W_x 表示 W 关于状态变量的偏导数向量. 由(2.11)式可见 $F(\xi)=0$. 另一方面,由于 $U(\tau,\xi)\in\Omega(\tau)$, 在(2.12)式中用$(\tau,x)$代替$(\tau,\xi)$我们可得如下不等式

$$W_t(\tau,x)\geqslant -f^0(\tau,x,U(\tau,\xi))\\ -\langle W_x(\tau,x),f(\tau,x,U(\tau,\xi))\rangle.$$

这说明 $F(x)\geqslant 0$. 因此, $F(x)$在 $x=\xi$ 处有最小值. 由于 W 是 $C^{(2)}$ 类函数, F 是 $C^{(1)}$ 函数. 于是,因为 ξ 是 F 的定义域之内点,并且 F 在 $x=\xi$ 达到最小值,故有 $F_x(\xi)=0$. 如果用(2.16)式计算 F 关于状态变量 x 的偏导数,然后在 $x=\xi$ 处令该偏导数等于0,我们就得到

$$\frac{\partial^2 W}{\partial t\partial x^i}+\frac{\partial f^0}{\partial x^i}+\sum_{j=1}^{n}\frac{\partial^2 W}{\partial x^i\partial x^j}f^j+\sum_{j=1}^{n}\frac{\partial W}{\partial x^j}\frac{\partial f^j}{\partial x^i}=0,\quad i=1,\cdots,n, \tag{2.17}$$

其中 W 的偏导数是在(τ,ξ)处取值的, 而 f^i 及其偏导数是在$(\tau,\xi,U(\tau,\xi))$处取值的. 因为(τ,ξ)是 $\mathscr{R}$ 中任意指定的点,这表明方程(2.17)对变元(t,x)和$(t,x,U(t,x))$成立,这里(t,x)为 $\mathscr{R}$ 中任一点.

在继续进行我们的分析之前,先引入一些有用的术语.

定义 2.1 如果 $h:(t,x,z)\to h(t,x,z)$是从 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 到 $E^k(k\geqslant 1)$的函数, 那么"函数 h 沿轨线 $\phi(\cdot,\tau,\xi)$取值"的含意是指复合函数 $t\to h(t,\phi(t,\tau,\xi),u(t,\tau,\xi))$. 同样,若 w 是定义在 $\mathscr{R}$ 上的函数, 那么"函数沿轨线 $\phi(\cdot,\tau,\xi)$取值"的含意是指复合函数 $t\to w(t,\phi(t,\tau,\xi))$.

现在假设$(\tau,\ \xi)$是$\mathscr{R}$中的一个固定点，并考虑偏导数向量$W_x=(W_{x^1},\ \cdots,\ W_{x^n})$沿着从$(\tau,\ \xi)$出发的最优轨线的性质．我们定义一个从$[\tau,\ t_1]$到$E^n$的函数$\lambda(\cdot,\ \tau,\ \xi): t\to\lambda(t,\ \tau,\ \xi)$如下，

$$\lambda(t,\ \tau,\ \xi)=-W_x(t,\ \phi(t,\ \tau,\ \xi)). \tag{2.18}$$

因为W是$C^{(2)}$函数，所以λ对t可微．由于$\phi'(t,\ \tau,\ \xi)=f(t,\ \phi(t,\ \tau,\ \xi),\ u(t,\ \tau,\ \xi))$，我们得到

$$\frac{d\lambda^i}{dt}=-\frac{\partial^2 W}{\partial t\partial x^i}-\sum_{j=1}^{n}\frac{\partial^2 W}{\partial x^i\partial x^j}f^j,\quad i=1,\ \cdots,\ n, \tag{2.19}$$

其中W的偏导数以及f的支量是沿轨线$\phi(\cdot,\ \tau,\ \xi)$取值的．如果将(2.19)代入(2.17)并利用(2.18)，我们有

$$\frac{d\lambda^i}{dt}=-\left\{-\frac{\partial f^0}{\partial x^i}+\sum_{j=1}^{n}\lambda^j\frac{\partial f^j}{\partial x^i}\right\},\quad i=1,\ \cdots,\ n.$$

用向量-矩阵符号表示即为

$$\frac{d\lambda}{dt}=-\left\{-\frac{\partial f^0}{\partial x}+\frac{\partial f}{\partial x}\lambda\right\}, \tag{2.20}$$

其中$\dfrac{d\lambda}{dt}$，$\dfrac{\partial f^0}{\partial x}$和$\lambda$是列向量；而$\dfrac{\partial f}{\partial x}$是偏导数矩阵，它的第$i$行和第$j$列的元素为$\partial f^j/\partial x^i$．(2.20)式中的偏导数是沿轨线$\phi(\cdot,\ \tau,\ \xi)$取值的．总而言之，我们已证明了对于每一最优轨线$\phi(\cdot,\ \tau,\ \xi)$存在一个函数$\lambda(\cdot,\ \tau,\ \xi)$使得(2.20)式成立．我们指出，由于$\partial f^0/\partial x$和$\partial f/\partial x$是沿$\phi(\cdot,\ \tau,\ \xi)$取值的，它们是定义于$[\tau,\ t_1]$上的$t$的函数．因此方程组(2.20)可看成是函数$\lambda(\cdot,\ \tau,\ \xi)$必须满足的变系数线性微分方程组．该方程组的初始条件将在下面讨论．

用(2.13)式定义的函数H表示，方程(2.20)变为

$$\lambda'(t,\ \tau,\ \xi)$$
$$=-H_x(t,\ \phi(t,\ \tau,\ \xi),\ u(t,\ \tau,\ \xi),\ -1,\ \lambda(t,\ \tau,\ \xi)), \tag{2.21}$$

其中“$'$”表示对时间t的导数．按(2.13)中H的定义又有

$$\phi'(t,\ \tau,\ \xi)$$
$$=H_\lambda(t,\ \phi(t,\ \tau,\ \xi),\ u(t,\ \tau,\ \xi),\ -1,\ \lambda(t,\ \tau,\ \xi)). \tag{2.22}$$

由(2.21)和(2.22)可见函数$\phi(\cdot,\ \tau,\ \xi)$和$\lambda(\cdot,\tau,\ \xi)$满足微分方程组

$$\begin{cases}\dfrac{dx}{dt}=H_p(t,\ x,\ u(t,\ \tau,\ \xi),\ -1,\ p),\\ \dfrac{dp}{dt}=-H_x(t,\ x,\ u(t,\ \tau,\ \xi),\ -1,\ p).\end{cases}\tag{2.23}$$

我们可将(2.14)和(2.15)结合起来又得

$$\begin{aligned}&H(t,\ x,\ U(t,\ x),\ -1,\ -W_x(t,\ x))\\&=\max_{z\in\Omega(t)}H(t,\ x,\ z,\ -1,\ -W_x(t,\ x)).\end{aligned}\tag{2.24}$$

因为(2.24)式对所有 $\mathscr{R}$ 中的$(t,\ x)$成立,故它沿着最优轨线 $\phi(\cdot,\ \tau,\ \xi)$也成立. 而在轨线 $\phi(\cdot,\ \tau,\ \xi)$上的点$(t,\ x)$处,关系

$$x=\phi(t,\ \tau,\ \xi)$$

成立. 由此关系和(2.8)式可见对这种点,关系式

$$u(t,\ \tau,\ \xi)=U(t,\ \phi(t,\ \tau,\ \xi))\tag{2.25}$$

成立. 关系式(2.18)沿最优轨线也成立. 因此,对所有 $\tau\leqslant t\leqslant t_1$,其中 t_1 是 $\phi(\cdot,\tau,\ \xi)$到达 $\mathscr{T}$ 的时刻,有关系式

$$\begin{aligned}&H(t,\ \phi(t,\ \tau,\ \xi),\ u(t,\ \tau,\ \xi),\ -1,\ \lambda(t,\ \tau,\ \xi))\\&=\max_{z\in\Omega(t)}H(t,\ \phi(t,\ \tau,\ \xi),\ z,\ -1,\ \lambda(t,\ \tau,\ \xi)).\end{aligned}\tag{2.26}$$

方程组(2.23)已具有边界条件 $\phi(\tau,\ \tau,\ \xi)=\xi$, 而 λ 的边界条件将与刻画最优轨线 $\phi(\cdot,\ \tau,\ \xi)$和最优控制 $u(\cdot,\ \tau,\ \xi)$的关系式(2.26)一起用下面的方法确定之. 与 $\phi(\cdot,\ \tau,\ \xi)$和 $u(\cdot,\ \tau,\ \xi)$相联系, 存在函数 $\lambda(\cdot,\ \tau,\ \xi)$, 使得 $\lambda(\cdot,\ \tau,\ \xi)$和$\phi(\cdot,\ \tau,\ \xi)$是方程组(2.23)带适当边界条件的解,并使(2.26)在$[\tau,\ t_1]$上成立. 在现在的假设下, 方程组(2.23)及其适当的边界条件和关系式(2.26)一起构成所谓 Pontryagin 的最大值原理. 它们是最优对必须满足的一组必要条件. 更精确和更一般的表达形式将在下面定理 3.1 中给出.

方程(2.23)和(2.26)中含有最优控制 $u(\cdot,\ \tau,\ \xi)$, 我们改写这些方程, 使之含有最优综合 U. 如果把(2.25)代入(2.21)和(2.22),我们看出 $\phi(\cdot,\ \tau,\ \xi)$和 $\lambda(\cdot,\ \tau,\ \xi)$满足方程组

$$\left.\begin{aligned}\frac{dx}{dt}&=H_p(t,\ x,\ U(t,\ x),\ -1,\ p),\\ \frac{dp}{dt}&=-H_x(t,\ x,\ U(t,\ x),\ -1,\ p).\end{aligned}\right\}\tag{2.27}$$

从(2.26)式我们看到$\phi(\cdot,\ \tau,\ \xi)$和$\lambda(\cdot,\ \tau,\ \xi)$还满足

$$H(t,\ x,\ U(t,\ x),\ -1,\ p)=\max_{z\in\Omega(t)}H(t,\ x,\ z,\ -1,\ p).\tag{2.28}$$

其次，我们推导"横截条件"．这些条件是值函数W及其偏导数必须满足的边界条件．横截条件也是函数$\lambda(\cdot,\ \tau,\ \xi)$必须满足的边界条件．

设$(\tau_1,\ \xi_1)$是以$(\tau,\ \xi)$为起点的最优轨线的终止点．则有Σ中的点σ_1，使得$\tau_1=T(\sigma_1)$，$\xi_1=x(\sigma_1)$．设Γ^i是$\mathscr{T}$上的通过点$(\tau_1,\ \xi_1)$的曲线，它的参数方程为

$$t_1=T(\sigma_1^1,\ \cdots,\ \sigma_1^{i-1},\ \sigma^i,\ \sigma_1^{i+1},\ \cdots,\ \sigma_1^q),$$
$$x_1=x(\sigma_1^1,\ \cdots,\ \sigma_1^{i-1},\ \sigma^i,\ \sigma_1^{i+1},\ \cdots,\ \sigma_1^q),$$

其中σ^i取值于某开区间$(a^i,\ b^i)$上．曲线Γ^i可以这样得到：向量σ的所有支量除第i个支量外保持固定，而让第i个支量在整个区间$(a^i,\ b^i)$上变化．曲线Γ^i有时叫做在$\mathscr{T}$上的第i个坐标曲线．

现在假设$\mathscr{T}$是n维的，$\mathscr{T}$的每一点是唯一最优轨线的终止点．又假设W能在$\mathscr{T}$的邻域内开拓成为$C^{(1)}$函数．那么，从(2.2)和W的定义可见对$(t_1,\ x_1)\in\mathscr{T}$，有

$$W(t_1,\ x_1)=g(t_1,\ x_1).\tag{2.29}$$

所以(2.29)式沿着每一曲线$\Gamma^i(i=1,\ \cdots,\ n.\ )$成立．因此，我们可以对(2.29)沿着$\Gamma^i$关于$\sigma^i$求导数而得到

$$W_t\frac{\partial T}{\partial\sigma^i}+\left\langle W_x,\ \frac{\partial x}{\partial\sigma^i}\right\rangle=g_t\frac{\partial T}{\partial\sigma^i}+\left\langle g_x,\ \frac{\partial x}{\partial\sigma^i}\right\rangle$$

沿着Γ^i成立．我们将此方程改写为

$$\left\langle(W_t-g_t,\ W_x-g_x),\ \left(\frac{\partial T}{\partial\sigma^i},\ \frac{\partial x}{\partial\sigma^i}\right)\right\rangle=0,\qquad i=1,\ \cdots,\ n.\tag{2.30}$$

特别，(2.30)式在 $\sigma^i=\sigma_1^i$ 成立．因此，我们可以取 W_t, W_x, g_t 和 g_x 的自变量为 (τ_1, ξ_1)，而 $\dfrac{\partial T}{\partial\sigma^i}$ 和 $\dfrac{\partial x}{\partial\sigma^i}$ 的自变量为 σ_1．利用(2.14)式我们可以把方程(2.30)改写成

$$\left\langle (H-g_t,\ W_x-g_x),\ \left(\frac{\partial T}{\partial\sigma^i}, \frac{\partial x}{\partial\sigma^i}\right)\right\rangle=0, \qquad i=1, \cdots, n, \tag{2.31}$$

其中 H 是在 $(\tau_1, \xi_1, u(\tau_1, \tau, \xi), -1, W_x(\tau, \xi))$ 处取值的．

如果利用(2.18)式，方程(2.31)可写成

$$\left\langle (H-g_t,\ -\lambda-g_x),\ \left(\frac{\partial T}{\partial\sigma^i}, \frac{\partial x}{\partial\sigma^i}\right)\right\rangle=0, \qquad i=1, \cdots, n, \tag{2.32}$$

其中 $\lambda=\lambda(\tau_1, \tau, \xi)$．当我们把方程(2.32)写出来时，它变成

$$\left(H-\frac{\partial g}{\partial t}\right)\frac{\partial T}{\partial\sigma^i}-\sum_{j=1}^{n}\left(\lambda^j+\frac{\partial g}{\partial x^j}\right)\frac{\partial x^j}{\partial\sigma^i}=0, \qquad i=1, \cdots, n.$$

因为 $H=-f^0+\langle\lambda, f\rangle$，我们可以把这组方程改写如下：

$$-\left(f^0+\frac{\partial g}{\partial t}\right)\frac{\partial T}{\partial\sigma^i}-\sum_{j=1}^{n}\frac{\partial g}{\partial x^j}\frac{\partial x^j}{\partial\sigma^i}=\sum_{j=1}^{n}\left(\frac{\partial x^j}{\partial\sigma^i}-\frac{\partial T}{\partial\sigma^i}f^j\right)\lambda^j, \tag{2.33}$$

其中函数 f^0, f^j, $\partial g/\partial t$ 和 $\partial g/\partial x^j$ 是在轨线 $\phi(\cdot, \tau, \xi)$ 的端点 (τ_1, ξ_1) 上取值的，T 和 x 的偏导数是在 Σ 内相应于 (τ_1, ξ_1) 的点 σ_1 处取值的．如果 $\mathscr{T}$ 是 q 维流形 $(0\leqslant q\leqslant n)$，则(2.33)由 q 个方程而不是 n 个方程组成．从我们这里的论证是作不出此断言的，但在下一章将证明其成立．

因为 Γ^i 的单位切向量是 $(\partial T/\partial\sigma^i,\ \partial x/\partial\sigma^i)$ 方向上的单位向量，方程(2.30)说明在 (τ_1, ξ_1) 处，向量 $(W_t-g_t,\ W_x-g_x)$ 或者是零向量或者正交于 Γ^i．我们假设正交性成立，从(2.31)和(2.32)我们看到这一命题等价于说 $(H-g_t,\ W_x-g_x)$ 或者 $(H-g_t,\ -\lambda-g_x)$ 在 (τ_1, ξ_1) 处正交于 Γ^i．因为这一事实对每一坐标曲线 $\Gamma^i(i=1, \cdots, n)$ 都正确，又由于在 (τ_1, ξ_1) 处 Γ^i 的切向量生成 $\mathscr{T}$ 在 (τ_1, ξ_1) 处的切平面，所以下述命题正确：向量 $(W_t-g_t,\ W_x$

$-g_x$)或等价的向量($H-g_t$, W_x-g_x)或等价的向量($H-g_t$, $-\lambda-g_x$)在最优轨线的端点(τ_1, ξ_1)的取值与 $\mathscr{T}$ 在该点上正交. 这是横截条件的几何表述，它的解析表达式由(2.30)或(2.31)或(2.32)组成.

方程(2.29)和(2.30)是偏微分方程(2.14)的边界条件.

方程(2.32)，或等价地,(2.33)确定了 $\lambda(\cdot, \tau, \xi)$ 在 $t=\tau_1$ 的值. 因此它们提供了系统(2.23)和(2.27)的至今还缺少的边界条件. 注意，方程(2.32)和(2.33)关于 λ 是线性的. 因此 λ 满足带线性边界条件的线性微分方程组. 我们指出，带边界条件 $\phi(\tau, \tau, \xi)=\xi$ 和(2.32)的系统(2.23)构成一个两点边值问题，其中 ϕ 的值在初始时刻给定，而 λ 的值在终止时刻给定.

附注 2.1 如果解出满足边界条件(2.29)和(2.30)的偏微分方程(2.14)，就可发现此问题的特征方程即为方程(2.27). 我们把这事实的验证留给读者.

3. 最大值原理的表述

本节我们要给出第二章的问题 1 的最大值原理的精确表述，其证明将在第六章中给出. 我们假设 $g\equiv 0$. 如第二章第 4 节所指出的，$g\not\equiv 0$ 的问题可以化为带 $g\equiv 0$ 的问题. 当 $g\not\equiv 0$ 以及其他问题的最大值原理的表述将放在本节末的习题中去讨论.

下面的假设将是讨论问题的出发点.

假设 3.1 (i)区域 $\mathscr{R}$ 具有形式 $\mathscr{I}_0\times\mathscr{X}_0$，其中 $\mathscr{I}_0$ 是 E^1 中的开区间，而 $\mathscr{X}_0$ 是 E^n 中的开区间. 区域 $\mathscr{U}$ 是 E^m 中的开区间. (ii)函数 $\hat{f}=(f^0, f)=(f^0, f^1, \cdots, f^n)$ 对每一点 $(t, z)\in\mathscr{I}_0\times\mathscr{U}$，在 $\mathscr{X}_0$ 上是 $C^{(1)}$ 类的，而对每一 $x\in\mathscr{X}_0$，它在 $\mathscr{I}_0\times\mathscr{U}$ 上是 Borel 可测的. (iii)对每一致密区间 $\mathscr{X}\subset\mathscr{X}_0$ 和容许控制 u，存在一个定义在 $\mathscr{I}_u=[t_0, t_1]$ 上的函数 $\mu=\mu(\cdot, \mathscr{X}, u)$(此地 $\mathscr{I}_u$ 表示 u 的定义区间)，使得 $\mu\in L_1[\mathscr{I}_u]$ 并且对几乎所有 $t\in\mathscr{I}_u$ 和所有 $x\in\mathscr{X}$，有

$$|\hat{f}(t, x, u(t))|\leqslant\mu(t),$$
$$|\hat{f}_x(t, x, u(t))|\leqslant\mu(t).$$

假设 3.2 映射 Ω 与 x 无关，亦即对所有 $\mathscr{X}_0$ 里的 x 和 x'，$\Omega(t, x)=\Omega(t, x')$. 因此 $\Omega: t\to\Omega(t)$.

假设 3.3 定义问题的端点条件的集合 $\mathscr{B}$ 是 q 维的 $C^{(1)}$ 流形，其中 $0\leqslant q\leqslant 2n+1$.

因为最大值原理的表述只涉及到在最优轨线端点的邻域内 $\mathscr{B}$ 的性质. 不失一般性，可设 $\mathscr{B}$ 能用单一坐标补缀来表示，亦即我们假设 $\mathscr{B}$ 是 E^q 中一个开的平行六面体 Σ 在映射

$$t_0=T_0(\sigma),\ x_0=x_0(\sigma),$$

$$t_1=T_1(\sigma),\ x_1=x_1(\sigma)$$

之下的象，其中函数 T_i 和 $x_i(i=0,1)$ 在 Σ 上是 $C^{(1)}$ 类的，并且 Jacobian 矩阵

$$(T_{0\sigma}, x_{0\sigma}, T_{1\sigma}, x_{1\sigma})$$

在 Σ 上处处有秩为 q.

如第 2 节那样，令 H 是用下面公式表示的、在 $E^1\times E^n\times E^m\times E^1\times E^n$ 上定义的实值函数：

$$H(t, x, z, p^0, p)=p^0f^0(t, x, z)+\langle p, f(t, x, z)\rangle$$

$$=\sum_{i=0}^{n}p^if^i(t, x, z). \tag{3.1}$$

若令 $\hat{p}=(p^0, p)$，(3.1)可写成等价形式

$$H(t, x, z, \hat{p})=\langle\hat{p}, \hat{f}(t, x, z)\rangle.$$

因为本章剩下的部分我们以最优对为主要论题，因此，对最优对将不用特别记号，诸如打上星号或上标. 除非我们觉得会与其他结果相混淆时才给以区别.

定理 3.1(积分型最大值原理). 设 $g\equiv0$，并且假设 3.1 至 3.3 成立. 设 (ϕ, u) 为定义在区间 $[t_0, t_1]$ 上的最优对，则存在一常数 $\lambda^0\leqslant0$ 和定义于 $[t_0, t_1]$ 上的绝对连续向量函数 $\lambda=(\lambda^1, \cdots, \lambda^n)$，使得如下事实成立：

(i) 向量 $\hat{\lambda}(t)=(\lambda^0, \lambda(t))$ 在 $[t_0, t_1]$ 上恒不为零；

(ii) 对 $[t_0, t_1]$ 中几乎所有的 t，有

$$\begin{cases}\phi'(t)=H_p(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t)),\\ \lambda'(t)=-H_x(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t));\end{cases}\tag{3.2}$$

(iii) 对定义于 $[t_0,\ t_1]$ 上的任何容许控制 $v(t)$,

$$\int_{t_0}^{t_1} H(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t))dt$$

$$\geqslant\int_{t_0}^{t_1} H(t,\ \phi(t),\ v(t),\ \hat{\lambda}(t))dt;\tag{3.3}$$

(iv) 如果映射 $t\to\hat{f}(t,\ \phi(t),\ u(t))$ 在 $t=t_i(i=0,\ 1)$ 连续，则 $(2n+2)$ 维向量

$$(H(\pi(t_0)),\ -\lambda(t_0),\ -H(\pi(t_1)),\ \lambda(t_1))\tag{3.4}$$

在点 $(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))$ 处正交于 $\mathscr{B}$，其中

$$\pi(t_i)=(t_i,\ \phi(t_i),\ u(t_i),\ \hat{\lambda}(t_i)),\qquad i=0,\ 1.$$

定理 3.1 将在第六章中证明.

推论 3.1 (点式最大值原理)　如果对所有 t, $\Omega(t)=\mathscr{C}$, 其中 $\mathscr{C}$ 是固定集合，而 $\hat{f}$ 在 $\mathscr{R}\times\mathscr{U}$ 上连续，则

$$H(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t))\geqslant H(t,\ \phi(t),\ z,\ \hat{\lambda}(t))\tag{3.5}$$

对几乎所有 $t\in[t_0,\ t_1]$ 和所有 $z\in\mathscr{C}$ 成立.

推论 3.1 也在第六章中证明. 注意，如果 (3.5) 式成立，那么定理中的 (3.3) 当然成立.

推论 3.2　若对所有 t, $\Omega(t)=\mathscr{C}$, 其中 $\mathscr{C}$ 是一个固定集合. 设 $\hat{f}$ 在 $\mathscr{R}\times\mathscr{U}$ 上连续而在 $\mathscr{R}$ 上连续可微，u 在 $[t_0,\ t_1]$ 上有界. 则存在定义于 $[t_0,\ t_1]$ 上的绝对连续函数 $h(t)$, 使得在 $[t_0,\ t_1]$ 上几乎处处满足

$$h(t)=H(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t))$$

和

$$h'(t)=H_t(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t)).$$

此外，如果所论问题是定常的；亦即 $\hat{f}$ 与 t 无关且 $\mathscr{B}$ 与 t_0 和 t_1 无关. 则

$H(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t))=C$ 在 $[t_0,\ t_1]$ 上几乎处处成立，

其中 C 是常数.

如果 u 是分段连续的，而 $\mathscr{C}$ 是闭的，则 $t\to H(t,\ \phi(t),\ u(t),$

$\hat{\lambda}(t)$)是绝对连续的，并且它的导数是映象 $t \to H_t(t, \phi(t), u(t), \hat{\lambda}(t))$.

推论 3.2 也会在第六章证明.

定义 3.1 满足定理 3.1 结论的容许对 (ϕ, u) 叫做极值对. 函数 ϕ 叫做极值轨线，而控制 u 叫做极值控制.

方程组(3.2)说明函数 $\hat{\lambda}(t)$ 和 $\phi(t)$ 是方程组

$$\begin{cases} \dfrac{dx}{dt} = H_p(t, x, u(t), p^0, p), \\ \dfrac{dp}{dt} = -H_x(t, x, u(t), p^0, p) \end{cases} \tag{3.6}$$

的解.

如果 (ϕ, u) 是极值对，那么我们称向量 $(t, \phi(t), u(t), \hat{\lambda}(t))$ 为 t 处的极值元，其中 $\hat{\lambda}(t)$ 如定理 3.1 中所述. 我们把极值元表示成

$$\pi(t) = (t, \phi(t), u(t), \hat{\lambda}(t)). \tag{3.7}$$

因此，方程(3.2)可以写成

$$\phi'(t) = H_p(\pi(t)), \quad \lambda'(t) = -H_x(\pi(t)).$$

我们曾经用极值元写出横截条件.

因为 $H_p = f$, (3.2)中第一个方程可写成 $\phi'(t) = f(t, \phi(t), u(t))$ 几乎处处成立，此即表示 ϕ 是相应于控制 u 的轨线. (3.2)中第二个方程可以写成

$$\lambda'(t) = -[\hat{f}_x(t, \phi(t), u(t))]\hat{\lambda}(t), \tag{3.8}$$

其中 $\hat{f}_x$ 是 $n \times (n+1)$ 矩阵，它的第 i 行和第 j 列元素为 $\partial f^j/\partial x^i$ $(i=1, \cdots, n, j=0, 1, \cdots, n.)$，这些元素都是沿轨线 ϕ 取值的. 因此 λ 满足一个线性微分方程组，当把它用分量形式写出时，即为

$$\frac{d\lambda^i}{dt} = -\lambda^0 \frac{\partial f^0}{\partial x^i}(t, \phi(t), u(t)) - \sum_{j=1}^{n} \frac{\partial f^j}{\partial x^i}(t, \phi(t), u(t))\lambda^{j[注]}, \qquad i=1, \cdots, n.$$

[注] 原书误为 $\frac{dp^i}{dt} = -\lambda^0 \frac{\partial f^0}{\partial x^i} - \sum_{j=1}^{n} \frac{\partial f^j}{\partial x^i} p^j$. ——译者注.

PDF 文件使用 "pdfFactory" 试用版本创建 www.fineprint.com.cn

(3.4)式在轨线ϕ的端点正交于$\mathscr{B}$的命题通常叫做"横截条件". 横截条件的解析表达式放在下面习题3.4中讨论.

最大值原理是最优性的必要条件, 它不是充分条件. 如下面的例子所示.

例 3.1 设

$$f^0(t, x, z)=az^2-4bxz^3+2btz^4, \qquad a>0,\ b>0.$$

$\Omega(t, x)=E^1$, 状态方程为$dx/dt=u(t)$, 而$\mathscr{B}$由单点$(t_0, x_0, t_1, x_1)=(0, 0, 1, 0)$组成. 问题是要使

$$J(\phi, u)=\int_0^1 f^0(t, \phi(t), u(t))dt$$

取极小值. 而函数H是$H(t, x, z, \hat{p})=p^0(az^2-4bxz^3+2btz^4)+pz$.

方程(3.6)变为

$$\begin{cases}\dfrac{dx}{dt}=H_p=u(t),\\[2mm] \dfrac{dp}{dt}=-H_x=4p^0b(u(t))^3,\end{cases}\tag{3.9}$$

因为$\mathscr{B}$是零维的, 横截条件不能给出关于λ^0, $\lambda(0)$或$\lambda(1)$的信息.

设ϕ, u和λ在$[0, 1]$上恒为0, 并取$\lambda^0=-1$. 则(3.9)满足, 而且

$$H(t, \phi(t), z, \hat{\lambda}(t))=-z^2(a+2btz^2).$$

因为$a>0$, $b>0$, 而$t\geqslant 0$, 上式右边对$[0, 1]$中的一切t是非正的. 另一方面, $H(t, \phi(t), u(t), \hat{\lambda}(t))=0$, 所以(3.5)成立. 因此$(\phi, u)$是极值对. 注意, 在本例中, 每当$z\neq 0$时, (3.5)中严格不等号成立.

为了看一下$(\phi, u)=(0, 0)$虽为极值对, 但不是最优的, 我们首先注意到$J(\phi, u)=J(0, 0)=0$. 今取k为满足$0<k<1$的固定数, 而令h为小的正数, 定义控制u_h如下:

$$u_h(t)=\begin{cases}k/h, & \text{当 } 0\leqslant t<h,\\ -k/(1-h), & \text{当 } h\leqslant t<1.\end{cases}$$

设ϕ_h为相应于u_h的轨线, 则经直接计算(其计算留给读者)可以

证明 $\lim\limits_{h\to 0} J(\phi_h, u_h) = -\infty$[注]. 因此$(\phi, u)$不可能是最优的。

习题 3.1 证明 $J(\phi_h, u_h) \to -\infty$, 当 $h\to 0$.

在下面的习题中和本书的剩下部分，我们将用$e(\phi)$表轨线ϕ的端点，因此

$$e(\phi) = (t_0, \phi(t_0), t_1, \phi(t_1)). \tag{3.10}$$

习题 3.2 设$\mathscr{T}_0$是一点，而$\mathscr{T}_1$是n维的$C^{(1)}$流形. 试证若极值轨线不与$\mathscr{T}_1$相切，则$\lambda^0\neq 0$，并且$\lambda(t_1)$是唯一的.

习题 3.3 证明如果$\lambda^0\neq 0$，则可假设$\lambda^0=-1$.

习题 3.4 证明"$(2n+2)$维向量(3.4)在端点$e(\phi)$正交于$\mathscr{B}$"等价于向量$(\lambda^0, \lambda(t_0), \lambda(t_1))$满足如下方程组:

$$\lambda^0\left(f^0(P_0)\frac{\partial T_0}{\partial\sigma^i} - f^0(P_1)\frac{\partial T_1}{\partial\sigma^i}\right) + \sum_{j=1}^{n}\lambda^j(t_0)\left(f^j(P_0)\frac{\partial T_0}{\partial\sigma^i} - \frac{\partial x_0^j}{\partial\sigma^i}\right)$$

$$-\sum_{j=1}^{n}\lambda^j(t_1)\left(f^j(P_1)\frac{\partial T_1}{\partial\sigma^i} - \frac{\partial x_1^j}{\partial\sigma^i}\right) = 0,$$

$i=1, \cdots, q$, 其中

$$P_i = (t_i, \phi(t_i), u(t_i)), \qquad i=0, 1,$$

而T_i和$x_i (i=0, 1)$的偏导数是在映入端点$e(\phi)$的σ处取值的.

习题 3.5 如果设$g\neq 0$，并假设g在E^{2n+2}中的$\mathscr{B}$之邻域中是$C^{(1)}$函数，则最大值原理的表述如定理3.1，除了把横截条件改写成向量

$$(H(\pi(t_0)) - \lambda^0 g_{t_0}, \ -\lambda(t_0) - \lambda^0 g_{x_0}, \ -H(\pi(t_1))$$

$$-\lambda^0 g_{t_1}, \ \lambda(t_1) - \lambda^0 g_{x_1})$$

在端点$e(\phi)$上正交于$\mathscr{B}$，其中g的偏导数是在$e(\phi)$上取值的.

习题 3.6(等周约束) 假如我们给容许对(ϕ, u)加上如下附

[注] 此时 $\phi_h(t) = \begin{cases} kt/h, & \text{当 } 0\leqslant t<h, \\ k(1-t)/1-h, & \text{当 } h\leqslant t<1. \end{cases}$

$$J(\phi_h, u_h) = ak^2\left(\frac{1}{h} - \frac{1}{1-h}\right) + bk^4\left(\frac{h-4}{h^3} + \frac{3-2h}{(1-h)^3}\right),$$

由此不难看出

$$\lim_{h\to 0} J(\phi_h, u_h) = -\infty.$$

——译者注。

加约束:

$$\int_{t_0}^{t_1} h^i(t,\ \phi(t),\ u(t))dt = c^i,\qquad i=1,\ \cdots,\ r.$$

其中 $h=(h^1,\ \cdots,\ h^r)$ 具有与定理 3.1 中的 f 一样的性质，而 $c=(c^1,\ \cdots, c^r)$ 是给定的向量．令

$$\widetilde{H}(t,\ x,\ z,\ \hat{p},\ \rho)=H(t,\ x,\ z,\ \hat{p})+\langle\rho,\ h(t,\ x,\ z)\rangle.$$

试证：如果 $g\equiv0$，而 $(\phi,\ u)$ 是刚才提出的等周问题的最优对，其定义区间为 $[t_0,\ t_1]$．则下面最大值原理成立：

存在常数 $\lambda^0\leqslant0$ 和绝对连续向量函数 $\lambda(t)=(\lambda^1(t),\ \cdots,\ \lambda^n(t))$，以及常向量 $r=(r^1,\ \cdots,\ r^r)$．使得如下事实成立：

(i) 向量 $(\hat{\lambda}(t),\ r)=(\lambda^0,\ \lambda(t),\ r)$ 在 $[t_0,\ t_1]$ 上恒不为 0．

(ii) 对几乎所有 $[t_0,\ t_1]$ 中的 t，有

$$\begin{cases}\phi'(t)=\widetilde{H}_p(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t),\ r),\\ \lambda'(t)=-\widetilde{H}_x(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t),\ r).\end{cases}$$

(iii) 对任何定义于 $[t_0,\ t_1]$ 上的容许控制 $v(t)$，有

$$\int_{t_0}^{t_1}\widetilde{H}(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t),\ r)dt \geqslant\int_{t_0}^{t_1}\widetilde{H}(t,\ \phi(t),\ v(t),\ \hat{\lambda}(t),\ r)dt.$$

(iv) 如果映射 $t\to(\hat{f}(t,\ \phi(t),\ u(t)),\ h(t,\ \phi(t),\ u(t)))$ 在 $t=t_i(i=0,\ 1)$ 上连续，则 $(2n+2)$ 维向量

$$(\widetilde{H}(\tilde{\pi}(t_0)),\ -\hat{\lambda}(t_0),\ -\widetilde{H}(\tilde{\pi}(t_1)),\ \hat{\lambda}(t_1))$$

在端点 $e(\phi)$ 上正交于 $\mathscr{B}$，其中

$$\tilde{\pi}(t_i)=(t_i,\ \phi(t_i),\ u(t_i),\ \hat{\lambda}(t_i),\ r),\qquad i=0,\ 1.$$

习题 3.7（参数最优化）　考虑在 II.5 中描述的控制与参数最优化问题．设 $\hat{f}=(f^0,\ f)$ 是定义于 $\mathscr{I}_0\times\mathscr{X}_0\times\mathscr{W}_0\times\mathscr{U}$ 上的函数，其中 $\mathscr{I}_0$，$\mathscr{X}_0$ 和 $\mathscr{U}$ 如假设 3.1，而 $\mathscr{W}_0$ 是 E^k 中的开区间．对每一 $(t,\ z)\in\mathscr{I}_0\times\mathscr{U}$，假设 $\hat{f}$ 在 $\mathscr{X}_0\times\mathscr{W}_0$ 上是 $C^{(1)}$ 类的，而对每一 $(x,\ w)\in\mathscr{X}_0\times\mathscr{W}_0$，$\hat{f}$ 在 $\mathscr{I}_0\times\mathscr{U}$ 上是 Borel 可测的．设 μ 如假设 3.1，而且

$$|\hat{f}(t,\ x,\ w,\ u(t))|\leqslant\mu(t),\qquad|\hat{f}_x(t,\ x,\ w,\ u(t))|\leqslant\mu(t),$$

$$|\hat{f}_w(t, x, w, u(t))| \leqslant \mu(t).$$

设 W 是 $\mathscr{W}_0$ 里的开集，并限制参数值 w 属于 W，(ϕ_0, u_0, w_0) 为所论问题的解，其中 (ϕ_0, u_0) 定义在区间 $[t_0, t_1]$ 上．令

$$\widetilde{H}(t, x, w, z, \hat{p}) = \langle \hat{p}, \hat{f}(t, x, w, z)\rangle.$$

试证：存在常数 $\lambda^0 \leqslant 0$ 和绝对连续向量 $\lambda(t) = (\lambda^1(t), \cdots, \lambda^n(t))$，使得下列事实成立：

(i) 向量 $\hat{\lambda}(t) = (\lambda^0, \lambda(t))$ 在 $[t_0, t_1]$ 上恒不为 0；

(ii) 对几乎所有 $[t_0, t_1]$ 中的 t，有

$$\begin{cases}\phi_0'(t) = \widetilde{H}_p(t, \phi_0(t), w_0, u_0(t), \hat{\lambda}(t)),\\ \lambda'(t) = -\widetilde{H}_x(t, \phi_0(t), w_0, u_0(t), \hat{\lambda}(t));\end{cases}$$

(iii) 对任何定义于区间 $[t_0, t_1]$ 上的容许控制 $v(t)$，有

$$\int_{t_0}^{t_1} \widetilde{H}(t, \phi_0(t), w_0, u_0(t), \hat{\lambda}(t))dt$$

$$\geqslant \int_{t_0}^{t_1} \widetilde{H}(t, \phi_0(t), w_0, v(t), \hat{\lambda}(t))dt;$$

(iv) 如果映射 $t \to \hat{f}(t, \phi_0(t), w_0, u_0(t))$ 在 $t = t_i (i = 0, 1)$ 上连续，则 $(2n+2)$ 维向量

$$(\widetilde{H}(\tilde{\pi}(t_0)), -\lambda(t_0), -\widetilde{H}(\tilde{\pi}(t_1)), \lambda(t_1))$$

在 $e(\phi_0)$ 上正交于 $\mathscr{B}$，其中

$$\tilde{\pi}(t_i) = (t_i, \phi_0(t_i), w_0, u_0(t_i), \hat{\lambda}(t_i)). \qquad i = 0, 1.$$

此外，

$$\int_{t_0}^{t_1} \langle \hat{\lambda}(t), \frac{\partial \hat{f}}{\partial w^j}(t, \phi_0(t), w_0, u_0(t))\rangle dt = 0, \qquad j = 1, \cdots, k.$$

习题 3.8 设 (ϕ, u) 是极值对，证明若 u 是逐段连续的，则由 $H^*(t) = H(\pi(t))$ 定义的映射 H^* 是绝对连续的，并有

$$\frac{dH^*}{dt} = H_t,$$

其中方程两边都是在 $\pi(t)$ 处取值的．

4. 一个例子

本节我们将举例说明如何用最大值原理和存在性定理去寻找

最优控制. 本例对指出在处理更复杂的系统中所遇到的困难也是有益的. 读者应注意, 这里最大值原理中的所有信息都被用来解决问题.

我们考虑如下资源分配问题. 设 $x_P(t)$ 表示在时刻 t 商品(比如钢)的生产速度, 而 $x_I(t)$ 表示该商品的投资速度. 我们假设全部投资都用于增加生产能力. 对于产钢而言, 投资可以想象为用来扩建新钢厂、添置采矿设备、以及运输设备. 所有这些费用的总和就是投资. 设 $x_O(t)$ 表示在时刻 t 商品的销售速度. 在产钢的情况中, 销售可想象为用户的商品(如汽车)的生产. 我们假设

$$x_P(t)=x_I(t)+x_O(t),\ x_I(t)\geqslant 0,\ x_O(t)\geqslant 0, \tag{4.1}$$

或

$$x_P(t)=C+\int_0^t x_I(s)\,ds, \tag{4.2}$$

其中 C 为初始生产速度. 方程(4.1)意味着假定在给定的周期内所生产的商品必须或者分配给投资或者分配给销售. 方程(4.2)则意味着假定商品全部分配给投资用以增加生产能力. 设 $T>0$ 是已给的正数. 那么在 $[0, T]$ 时间内, 总销售量为

$$P=\int_0^T x_O(s)\,ds.$$

在 $[0, T]$ 内的每一瞬时 t, 生产计划者必须选择 $x_I(t)$ 和 $x_O(t)$ 使 P 取极大值.

现在我们把这个问题表达成控制问题. 设 $x(t)$ 表示在时刻 t 的商品生产速度, 因为不能没有商品, $x(t)\geqslant 0$. 设 $u(t)$ 表示在时刻 t 所生产的商品分配给投资的部分, 因此 $0\leqslant u(t)\leqslant 1$, 而 $1-u(t)$ 表示分配给销售的部分, 于是

$$x_I(t)=u(t)x(t),$$
$$x_O(t)=(1-u(t))x(t).$$

因此前面提出的问题等价于如下控制问题: 在满足方程

$$\frac{dx}{dt}=u(t)x,\ x_0=C, \tag{4.3}$$

和约束

$$0 \leqslant u(t) \leqslant 1, \ x \geqslant 0, \tag{4.4}$$

(其中 $C>0$)的条件下，使泛函

$$J(\phi, u) = -\int_0^T (1-u(s))\phi(s)ds \tag{4.5}$$

取极小值．这里 T 是固定的，而终止状态 x_1 是非负的任意值．

在控制问题的标准表达形式中，这相当于说，对所有的 t，$\Omega(t)=[0, 1]$，而端点条件为

$$\mathscr{B}=\{(t_0, x_0, t_1, x_1): t_0=0, \ x_0=C, \ t_1=T, \ x^1 \geqslant 0\}. \tag{4.6}$$

因此 $\mathscr{B}$ 和 Ω 满足定理 III.5.1 的假设．而且 $f^0(t, x, z)=-(1-z)x$，$f(t, x, z)=zx$，故 f^0 和 f 也满足定理 III.5.1 的假设．从状态方程和初始条件(4.3)可推知，对任何满足(4.4)的控制 u，相应的轨线 ϕ 满足

$$C \leqslant \phi(t) \leqslant Ce^t. \tag{4.7}$$

因此约束 $x \geqslant 0$ 总是满足的，所以往后的讨论中可以略去这一限制．因为 $t_1=T$，(4.7)表明所有轨线处于致密集合中．在本例中，集合 $\mathscr{Q}^+$ 由

$$\begin{aligned}\mathscr{Q}^+(t, x) &= \{(y^0, y): y^0 \geqslant (z-1)x, \ y=zx, \ 0 \leqslant z \leqslant 1\} \\ &= \{(y^0, y): y^0 \geqslant y-x, \ 0 \leqslant y \leqslant x\}\end{aligned}$$

给出．对每一点(t, x)，集合 $\mathscr{Q}^+(t, x)$是闭和凸的，因此定理 III.5.1 的所有假设都满足，故最优对存在．

现在设(ϕ, u)表示最优对，我们要用定理 3.1 和推论 3.1 来确定(ϕ, u)，并在此过程中同时证明(ϕ, u)是唯一的．

在本例中，(3.1)式所定义的 H 为

$$H(t, x, z, \hat{p}) = -p^0(1-z)x+pzx.$$

因此 $\qquad H_p=zx, \quad -H_x=p^0(1-z)-pz.$

定理 3.1 中的 $\hat{\lambda}(t)$ 以及满足方程(3.2)的最优对(ϕ, u)，在本例中变成

$$\begin{cases}\phi'(t)=u(t)\phi(t),\\ \lambda'(t)=\lambda^0(1-u(t))-\lambda(t)u(t).\end{cases}\tag{4.8}$$

此外，由于 $\hat{f}$ 是连续的，而 $\Omega(t)=[0,\ 1]$ 对所有 t 成立，推论 3.1 可以套用，并且(3.5)式成立．而在本例中(3.5)式变为

$$(\lambda^0+\lambda(t))\phi(t)u(t)\geqslant(\lambda^0+\lambda(t))\phi(t)z\tag{4.9}$$

对一切 $z\in[0,\ 1]$ 和几乎所有 $t\in[0,\ T]$ 成立．

由于 $\mathscr{B}$ 是用(4.6)给定的，在 $\mathscr{B}$ 中任一点上，与 $\mathscr{B}$ 相切的向量必为(0, 0, 0, 1)的倍数．因此，在 ϕ 的端点 $e(\phi)$ 处正交于 $\mathscr{B}$ 的横截条件(3.4)式在本例中化为条件

$$\lambda(T)=0.\tag{4.10}$$

因为 $\hat{\lambda}(t)=(\lambda^0,\ \lambda(t))$ 恒不为 0，故得 $\lambda^0\neq0$．由于 $\lambda^0\leqslant0$，所以 $\lambda^0<0$．于是可在(4.8)的第二个方程和不等式(4.9)的两边同除以 $(-\lambda^0)>0$．如果我们令 $\tilde{\lambda}=\lambda(t)/-\lambda^0$，然后把 $\tilde{\lambda}$ 重新记为 λ，我们就得到，ϕ、u 和 λ 应满足

$$\begin{cases}\phi'(t)=u(t)\phi(t),\\ \lambda'(t)=-(1-u(t))-\lambda(t)u(t),\end{cases}\tag{4.11}$$

和不等式

$$(-1+\lambda(t))\phi(t)u(t)\geqslant(-1+\lambda(t))\phi(t)z\tag{4.12}$$

对一切 $0\leqslant z\leqslant1$ 和几乎所有 $t\in[0,\ T]$ 成立．将(4.8)、(4.9)与(4.11)、(4.12)相比较，即可看出在(4.8)和(4.9)中可以取 $\lambda^0=-1$．一般地，我们总可以证明，如果 $\lambda^0\neq0$，则可以取 $\lambda^0=-1$ (见习题 3.3)．

从(4.11)可见，ϕ 和 λ 是方程组

$$\begin{cases}\dfrac{dx}{dt}=u(t)x,\\[2mm] \dfrac{dp}{dt}=-(1-u(t))-u(t)p\end{cases}\tag{4.13}$$

的解．其边界条件由(4.10)和 $\phi(0)=C$ 给出．所以我们能知道 ϕ 的初值以及 λ 的终值；但不能同时知道 ϕ 和 λ 在同一点上的值．我们将看到，这就造成问题的困难．

因为 $C>0$，由(4.7)可见对一切 t，$\phi(t)>0$．因此尽管我们

不知道 ϕ 在 $t=T$ 的值 $\phi(T)$，我们仍然可知 $\phi(T)>0$.

从(4.12)式我们看到，对几乎所有 t 和 $z=u(t)$，能使表达式

$$(-1+\lambda(t))\phi(t)z \tag{4.14}$$

在 $0\leqslant z\leqslant 1$ 上取到最大值. 因此(4.14)式中 z 的系数的符号是最重要的，如果该系数 >0，则 $z=1$ 使(4.14)式取最大值；如果该系数 <0，则 $z=0$ 使它取最大值. 但已知对几乎所有 $t\in[0, T]$，$\phi(t)>0$，所以起决定作用的因子是 $(-1+\lambda(t))$.

设 $\phi(T)=\xi$. 由于 $\xi>0$ 和 $\lambda(T)=0$，故在 $t=T$ 处(4.14)式中 z 的系数是负的. 此外，由 $\phi(t)$ 和 $\lambda(t)$ 的连续性，必有形如 $(T-\delta, T]$ 的最大区间，使得 z 的系数在此区间上是负的. 因此，对一切 $t\in(T-\delta, T]$，$z=0$ 能使(4.14)达到最大值，所以对一切 $T-\delta<t\leqslant T$，$u(t)=0$.

在初始点 $t=0$ 处不能做如上的分析，因为 $\lambda(0)$，从而 z 的系数的符号不知道. 这就迫使我们试图从终端流形 $t=T$ 上任一点 $\xi>0$ 出发反向进行上一段所做的分析.

我们已经指出，存在一个最大区间 $(T-\delta, T]$，于其上 $u(t)=0$，它是以 T 为右端点的使 $-1+\lambda(t)<0$ 的最大区间. 现在来确定 δ. 由(4.13)的第一个方程中令 $u(t)=0$，以及按假设 $\phi(T)=\xi$，可见在 $(T-\delta, T]$ 上 $\phi(t)=\xi$. 又由(4.13)第二个方程中令 $u(t)=0$，以及从(4.10)式我们可以看出，在 $(T-\delta, T]$ 上 $\lambda(t)=T-t$. 因此若 $T>1$，则当 $T-1<t\leqslant T$ 时，$-1+\lambda(t)<0$，而 $-1+\lambda(T-1)=0$. 所以 $\delta=1$. 综上所述，我们已确立如下事实，如果 $T>1$，则在 $(T-1, T]$ 上有

$$u(t)=0,\ \phi(t)=\xi,\ \lambda(t)=T-t.$$

如果规定 $u(T-1)=0$，则上面三个等式在 $[T-1, T]$ 上成立.

如果 $T\leqslant 1$，则 $u(t)=0$，$\phi(t)=\xi$(其中 $\xi=C$)是最优对. 显然，从最优对的构成可见它是唯一的. 因此如果“规划范围”(T 的值)太短，最优策略是把产品全部付予销售.

我们现在转向 $T>1$ 的情形，并确定 $T-1$ 左边的 (ϕ, u). 我们建议读者在区间 $[T-1, T]$ 上画出函数 ϕ, u 和 λ 的图形，并且

当这些函数在 $T-1$ 的左边确定之后把它们画完整. 我们把(4.13)中的第二个方程改写成

$$\frac{dp}{dt}=-1+(1-p)u(t), \tag{4.15}$$

在区间$[0, T-1]$上考虑这个关于 λ 的微分方程. 因为 λ 在$[0, T]$上连续, 从下面直到最后一段的讨论中可以取 $\lambda(t)$ 的终止条件为 $\lambda(T-1)=1$.

因为 $\lambda(T-1)=1$ 和 $0\leqslant u(t)\leqslant 1$, 从(4.15)式和 $\lambda(t)$ 的连续性可见在$(T-1-\delta_1, T-1)$上有 $\lambda'(t)<0$. 因此在此区间上, 当我们按时间反向进行分析时 $\lambda(t)$ 是递增函数. 所以对某区间$(T-1-\delta_2, T-1)$中的 t, $\lambda(t)>1$. 又因为对所有的 t, $\phi(t)>0$, 于是从(4.14)式可见 $z=1$ 使其在$(T-1-\delta_2, T-1)$上达到最大值. 从而在该区间上 $u(t)=1$, 并且方程(4.13)变成

$$\frac{dx}{dt}=x,\ \phi(T-1)=\xi;$$

$$\frac{dp}{dt}=-p,\ \lambda(T-1)=1.$$

因此在区间$(T-1-\delta_2, T-1)$上, 有

$$\lambda(t)=\exp(T-1-t),\ \phi(t)=\xi\exp(t-T+1). \tag{4.16}$$

然而, 由于对所有 $0\leqslant t<T-1$, $\lambda(t)>1$. 所以$(T-1-\delta_2, T-1)=(0, T-1)$. 因此在$(0, T-1)$上 $\lambda(t)$ 和 $\phi(t)$ 可用(4.16)表示, 而且 $u(t)=1$.

在 $t=0$ 处我们有 $\phi(0)=\xi\exp(-T+1)$. 我们要求 $\phi(0)=C$, 因此 $\xi=C\exp(T-1)$, 从而在$[0, T-1]$上

$$\phi(t)=Ce^t.$$

我们上面确定的(ϕ, u)是一对极值对, 从用以确定(ϕ, u)的方法可见(ϕ, u)是唯一的. 因为我们已知道最优对是存在的而且它必定为极值对, 故知(ϕ, u)确实是最优对, 而且它是唯一的. 我们指出, 虽然存在性定理保证了可测最优控制的存在, 而应用最大值原理所得到的控制却是分段连续的.

在上述例子中所用的方法是一种常用的方法, 它说明困难发

生在 ϕ 的值给定在初始点，而 λ 的值给定在终止点．在少数问题中，可以从任意终止点按时间反向着手进行分析，并调整积分常数以便得到给定的初始状态．但对大多数问题或者必须采用计算机处理的问题，这样做并不那么容易．这里我们就不讨论这些问题了．

5．与变分法的关系

本节研究最大值原理与变分法中一阶必要条件之间的关系．我们详细地证明怎样从最大值原理得到关于变分法中简单问题的经典的一阶必要条件．在习题 5.1 中我们请读者从最大值原理推出 Bolza 问题的一阶必要条件．在习题 5.2 中，我们请读者从习题 5.1 叙述的结果导出某类问题的最大值原理．

在 II.6 中我们曾证明了变分法中的简单问题可以表达成如下控制问题．在满足约束条件：

$$\frac{dx}{dt}=u(t),\ e(\phi)\in\mathscr{B},\ \Omega(t,\ x)=\mathscr{U}$$

的条件下，使

$$J(\phi)=g(e(\phi))+\int_{t_0}^{t_1}f^0(t,\ \phi(t),\ u(t))dt$$

取极小值．其中 $\mathscr{B}$ 是 E^{2n+2} 中给定的集合，$e(\phi)$ 表示 ϕ 的端点 $(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))$，而 $\mathscr{U}$ 是 E^n 中的开集．我们假设 f^0 在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上是 $C^{(1)}$ 类的，$\mathscr{B}$ 是 E^{2n+2} 中的 q 维 $C^{(1)}$ 流形 $(0\leqslant q\leqslant 2n+1)$，而 $g\equiv 0$．我们将证明定理 3.1 可化成极小曲线必须满足的通常一阶变分的必要条件．

在现在的情形，函数 H 由公式

$$H(t,\ x,\ z,\ \hat{p})=p^0f^0(t,\ x,\ z)+\langle p,\ z\rangle$$

给出．设 ϕ 是变分问题的解，其定义区间为 $[t_0,\ t_1]$，则 $(\phi,\ u)=(\phi,\ \phi')$ 是相应控制问题的解．因此 $(\phi,\ u)$ 满足定理 3.1 的条件．此外，因为约束集合是固定的，推论 3.1 成立．所以存在纯量 $\lambda^0\leqslant 0$ 和定义于 $[t_0,\ t_1]$ 上的绝对连续向量函数 $\lambda(t)=(\lambda^1(t),\ \cdots,\ \lambda^n(t))$，使得对一切 $[t_0,\ t_1]$ 中的 t，$(\lambda^0,\ \lambda(t))\neq 0$，并且对几乎所

有 $t\in[t_0, t_1]$，有

$$\begin{cases}\phi'(t)=H_p(\pi(t))=u(t), & (5.1)\\ \lambda'(t)=-H_x(\pi(t))=-\lambda^0 f_x^0(t, \phi(t), u(t)), & (5.2)\end{cases}$$

而

$$H(t, \phi(t), z, \hat{\lambda}(t))=\lambda^0 f^0(t, \phi(t), z)+\langle\lambda(t), z\rangle, \tag{5.3}$$

它在 $\mathscr{U}$ 上的 $z=u(t)$ 处取到最大值．此外，向量

$$(H(\pi(t_0)), -\lambda(t_0), -H(\pi(t_1)), \lambda(t_1)) \tag{5.4}$$

在点 $e(\phi)$ 上正交于 $\mathscr{B}$．（记住，$g\equiv 0$．）

我们断定 $\lambda^0\neq 0$．因为若 $\lambda^0=0$，那么从(5.3)可见对几乎所有 $t\in[t_0, t_1]$，

$$\langle\lambda(t), u(t)\rangle\geqslant\langle\lambda(t), z\rangle$$

对一切 $z\in\mathscr{U}$ 成立．此即表示，对几乎所有 $t\in[t_0, t_1]$，线性泛函 $z\to\langle\lambda(t), z\rangle$ 在开集 $\mathscr{U}$ 的某一点 $u(t)\in\mathscr{U}$ 上达到最大值．这只有当 $\lambda(t)=0$ 才可能发生．但那样一来就有 $(\lambda^0, \lambda(t))=0$．这是不可能的．

因为 $\lambda^0\neq 0$，今后在(5.1)～(5.4)中可以取 $\lambda^0=-1$．（见习题 3.3．）从(5.2)式可得

$$\lambda(t)=\int_{t_0}^{t_1} f_x^0(s, \phi(s), u(s))ds+C, \tag{5.5}$$

其中 C 是某一常向量．从(5.3)可得知映射 $z\to H(t, \phi(t), z, -1, \lambda(t))$ 在 $z=u(t)$ 处取最大值．因为 $\mathscr{U}$ 是开的，$u(t)$ 是 $\mathscr{U}$ 的内点，又映射 $z\to H(t, \phi(t), z, -1, \lambda(t))$ 是可微的，它的导数在 $z=u(t)$ 处等于 0．所以

$$H_z(t, \phi(t), u(t), -1, \lambda(t))=0.$$

从而有

$$\lambda(t)=f_z^0(t, \phi(t), u(t)) \tag{5.6}$$

对几乎所有 $t\in[t_0, t_1]$ 成立．

从(5.1)、(5.5)和(5.6)我们得到

$$f_z^0(t, \phi(t), \phi')=\int_{t_0}^{t_1} f_x^0(s, \phi(s), u(s))ds+C \tag{5.7}$$

在 $[t_0, t_1]$ 上几乎处处成立．方程(5.7)有时叫做积分型的 Euler

方程或 du-Bois Reymond 方程. 在初等理论中假设 ϕ' 是分段连续的，方程(5.7)则处处成立. 而用公式

$$f_z^{0*}(t)=f_z^0(t,\ \phi(t),\ \phi'(t)) \tag{5.8}$$

定义的函数 f_z^{0*} 甚至在 ϕ 的角点上也是连续的. 所谓 ϕ 的角点是指 ϕ' 在该点具有跳跃的不连续性. 这一结果是熟知的 Weiestrass-Erdmann 角点条件.

从(5.7)我们可得

$$\frac{d}{dt}(f_z^{0*})=f_x^{0*},\text{在}[t_0,\ t_1]\text{上几乎处处成立}, \tag{5.9}$$

其中 f_z^{0*} 由(5.8)式定义，而 f_x^{0*} 表示映射 $t\to f_x^0(t,\ \phi(t),\ \phi'(t))$. 方程(5.9)即为 Euler 方程. 如果假设 ϕ' 是分段连续的，则(5.9)在 ϕ 的角点之间成立.

以下我们讨论横截条件. 为了简单起见，我们假设 u 在 $t=t_i$ $(i=0,\ 1)$ 上连续，并且(5.1)在这些点上成立. 如果在(5.4)中取 $\lambda^0=-1$，同时利用(5.1)和(5.6)，则(5.4)变成

$$\begin{aligned}(-f^{0*}(t_0)+\langle f_z^{0*}(t_0),\ \phi'(t_0)\rangle,\ -f_z^{0*}(t_0),\ f^{0*}(t_1)\\ -\langle f_z^{0*}(t_1),\ \phi'(t_1)\rangle,\ f_z^{0*}(t_1)),\end{aligned} \tag{5.10}$$

其中 $f^{0*}(t)=f^0(t,\ \phi(t),\ \phi'(t))$，而 f_z^{0*} 由(5.8)给出. 现在的横截条件就是(5.10)在 $e(\phi)$ 处正交于 $\mathscr{B}$.

如果取 $\lambda^0=-1$ 并利用(5.1)和(5.6)，则 $z=u(t)$ 使(5.3)式于 $\mathscr{U}$ 上取最大值的命题变成如下命题：对几乎所有 $t\in[t_0,\ t_1]$ 和所有 $z\in\mathscr{U}$，有

$$\begin{aligned}&-f^0(t,\ \phi(t),\ \phi'(t))+\langle f_z^0(t,\ \phi(t),\ \phi'(t)),\ \phi'(t)\rangle\\ &\geqslant -f^0(t,\ \phi(t),\ z)+\langle f_z^0(t,\ \phi(t),\ z),\ z\rangle.\end{aligned} \tag{5.11}$$

若引入如下函数 $\mathscr{E}$：

$$\mathscr{E}(t,\ x,\ z,\ y)\underset{\text{定}}{=}f^0(t,\ x,\ z)-f^0(t,\ x,\ y)-\langle f_z^0(t,\ x,\ y),\ z-y\rangle,$$

则(5.11)等价于命题

$$\mathscr{E}(t,\ \phi(t),\ z,\ \phi'(t))\geqslant 0 \tag{5.12}$$

对几乎所有 t 和所有 $z\in\mathscr{U}$ 成立. 不等式(5.12)是著名的 Weierstrass 条件. 注意，(5.12)左边是映射 $t\to f^0(t,\ \phi(t),\ z)$ 关

于点 $z=\phi'(t)$ 的 Toylor 展开式的一次项所组成的.

若假设 ϕ' 是分段连续的，则(5.12)在 ϕ 的角点之间必然成立. 如果 $t=\tau$ 是一角点，则令 $t\to\tau+0$ 和 $t\to\tau-0$，我们得知对于令 $t\to\tau\pm0$ 所得到的单边极限(5.12)仍然成立.

现在假设对固定的 $(t, x)\in\mathscr{R}$，函数 $f^0(t, x, z)$ 在 $\mathscr{U}$ 上是 $C^{(2)}$ 类的. 因此对每个确定的 t，映射 $z\to H(t, \phi(t), z, \hat\lambda(t))$ 在 $\mathscr{U}$ 上是 $C^{(2)}$ 类的. 对几乎所有的 t，函数 H 在 $z=u(t)$ 达到最大值. 因为 $\mathscr{U}$ 是开集，$z=u(t)$ 必为 $\mathscr{U}$ 的内点. 因此由矩阵 $H_{zz}(t, \phi(t), u(t), \hat\lambda(t))$ 确定的二次型对几乎所有 t 是半负定的. 但根据(5.3)以及可以取 $\lambda^0=-1$ 的事实，可得

$$H_{zz}(t, \phi(t), u(t), \hat\lambda(t))=-f^0_{zz}(t, \phi(t), \phi'(t)).$$

因此由 $f^0_{zz}(t, \phi(t), \phi'(t))$ 所确定的二次型对几乎所有的 t 是半正定的. 从而对所有 $\eta=(\eta^1, \cdots, \eta^n)\neq0$ 和几乎所有 t，有

$$\langle\eta, f^0_{zz}(t, \phi(t), \phi'(t))\eta\rangle\geqslant0. \tag{5.13}$$

这是周知的 Legendres 条件.

我们继续假设对 $\mathscr{R}$ 中固定的 (t, x)，函数 f^0 在 $\mathscr{U}$ 上是 $C^{(2)}$ 类的. 设 ϕ' 分段连续. 我们说 ϕ 是非奇异的，如果对 $[t_0, t_1]$ 中非角点的 t 矩阵 $f^0_{zz}(t, \phi(t), \phi'(t))$ 是非奇异的. 我们将证明如果 ϕ 是非奇异的，则在角点之间 ϕ 是 $C^{(2)}$ 函数. 这个结果是著名的 Hilbert 可微性定理.

设 τ 是 $[t_0, t_1]$ 中的不是角点的点，考虑如下关于 n 维向量 w 的方程

$$f^0_z(t, \phi(t), w)-\int_{t_0}^{t}f^0_x(s, \phi(s), \phi'(s))ds-C=0, \tag{5.14}$$

其中 C 如(5.7)所示. 从(5.7)我们看到在 $t=\tau$ 处 $w=\phi'(\tau)$ 是(5.14)的解. 此外，按假设矩阵 $f^0_{zz}(\tau, \phi(\tau), \phi'(\tau))$ 是非奇异的. 此矩阵是方程组(5.14)的 Jacobian 矩阵，因此由隐函数定理得知方程组(5.14)在不含有 ϕ 的任何角点的区间 $(\tau-\delta, \tau+\delta)$ 内有唯一的 $C^{(1)}$ 类解 w. 另一方面，从(5.7)我们又可看出 ϕ' 是(5.14)在

$(\tau-\delta, \tau+\delta)$中的解．因此 $\phi'=w$．从而在$(\tau-\delta, \tau+\delta)$上 ϕ' 是 $C^{(1)}$ 类的而 ϕ 是 $C^{(2)}$ 类的．因为 τ 是角点之间的任一点，这就证明了在角点之间 ϕ 是 $C^{(2)}$ 类的．

现在，若进一步假设 f^0 在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上是 $C^{(2)}$ 类的，而 ϕ 在角点之间是 $C^{(2)}$ 类的．我们可以用链式法则[注]从(5.9)得到如下角点之间的方程

$$f^{0*}_{x'x'}\phi''+f^{0*}_{x'x}\phi'+f^{0*}_{x't}-f^{0*}_{x}=0,$$

其中星号表示函数是在$(t, \phi(t), \phi'(t))$取值的，函数 ϕ' 和 ϕ'' 是在 t 上取值的．

习题 5.1 考虑如 II.6 所述的 Bolza 问题．现在假设微分方程 II(6.2)左边的函数 $F=(F^1, \cdots, F^\mu)$是由

$$F^i(t, x, x')=x'^i-G^i(t, x, \tilde{x}'), \qquad i=1, \cdots, \mu \quad (5.15)$$

给定的，其中 $\tilde{x}'=(x'^{\mu+1}, \cdots, x'^n)$．假设函数 f^0 和 F 在 $\mathscr{G}=\mathscr{R}\times\mathscr{U}$ 上是 $C^{(1)}$ 类的．注意，由(5.15)式，这相当于假设函数 $G=(G^1, \cdots, G^\mu)$在$(t, x, \tilde{x}')$-空间的适当区域上是 $C^{(1)}$ 类的．集合 $\mathscr{B}$ 假设是 q 维的 $C^{(1)}$ 流形$(0\leqslant q\leqslant 2n+1)$，而函数 g 在 $\mathscr{B}$ 的邻域中是 $C^{(1)}$ 的，令$\hat{\rho}=(\rho^0, \rho)=(\rho^0, \rho^1, \cdots, \rho^\mu)$和

$$\begin{aligned}L(t, x, x', \hat{\rho})&=-\rho^0 f^0(t, x, x')+\sum_{i=1}^{\mu}\rho^i F^i(t, x, x')\\&=-\rho^0 f^0(t, x, x')+\langle\rho, F(t, x, x')\rangle.\end{aligned}$$

在这些假设下试证：如果 ϕ 是 Bolza 问题的极小化函数，则下列诸条件成立：

(a)（Lagrange 乘子法则）．存在常数 $\psi^0\leqslant 0$ 和定义在$[t_0, t_1]$上的可测函数 $\psi=(\psi^1, \cdots, \psi^\mu)$，使得对所有 $t\in[t_0, t_1]$，$(\psi^0, \psi(t))\neq 0$，且

$$\begin{aligned}&L_{x'}(t, \phi(t), \phi'(t), \hat{\psi}(t))\\&=\int_{t_0}^{t_1}L_x(s, \phi(s), \phi'(s), \hat{\psi}(s))ds+C,\end{aligned}$$

其中 C 是一个适当的常向量而 $\hat{\psi}(t)=(\psi^0, \psi(t))$．此外，如果

$$P(t)=(t, \phi(t), \phi'(t), \hat{\psi}(t)),\quad A=L-\langle x', L_{x'}\rangle,$$

[注] 所谓链式法则，也就是复合函数的微分法则．——译者注

则$(2n+2)$维向量

$$(-A(p(t_0))-\psi^0 g_{t_0}(e(\phi)),\ -L_{x'}(p(t_0))-\psi^0 g_{x_0}(e(\phi)),$$
$$A(p(t_1))-\psi^0 g_{t_1}(e(\phi)),\ L_{x'}(p(t_1))-\psi^0 g_{x_1}(e(\phi)))$$

在$e(\phi)$上正交于$\mathscr{B}$.

最后这个命题通常叫做横截条件.

(b) (Weierstrass 条件)如果

$$\mathscr{E}(t, x, x', z, \hat{p})=L(t, x, z, \hat{p})-L(t, x, x', \hat{p})$$
$$-\langle F_{x'}(t, x, x'), z-x'\rangle^{[注]},$$

则对所有z和几乎所有t,有

$$\mathscr{E}(t, \phi(t), \phi'(t), z, \hat{\psi}(t))\geqslant 0,$$

其中ϕ和$\hat{\psi}$如(a)所述.

(c) (Clebsch 条件)不等式

$$\langle \eta, L_{x'x'}(t, \phi(t), \phi'(t), \hat{\psi}(t))\eta\rangle\geqslant 0$$

对几乎所有t和所有使

$$L_{x'}(t, \phi(t), \phi'(t), \hat{\psi}(t))\eta=0$$

的非零向量$\eta=(\eta^1, \cdots, \eta^n)$成立,其中$\phi$和$\hat{\psi}$如(a)所述.

习题 5.2 考虑具有如 II.6 的控制约束的控制问题,并假设约束条件 II.(6.6) 成立. 假定端点集合$\mathscr{B}$是q维$C^{(1)}$流形$(0\leqslant q\leqslant 2n+1)$,函数$g$、$f^0$、$f$和$R$在它们的定义域上是$C^{(1)}$类的. 假如对变分法中的 Bolza 问题习题 5.1 的结论成立,证明如下定理:

设(ϕ, u)是定义于$[t_0, t_1]$上的最优对,则存在常数$\lambda^0\leqslant 0$,定义于$[t_0, t_1]$上的绝对连续函数$\lambda(t)=(\lambda^1(t), \cdots, \lambda^n(t))$和定义于$[t_0, t_1]$上的可测函数$\nu(t)=(\nu^1(t), \cdots, \nu^r(t))$,使得如下诸项成立.

(i) 向量$\hat{\lambda}(t)=(\lambda^0, \lambda(t))$恒不为 0,而$\nu(t)\leqslant 0$在$[t_0, t_1]$上几乎处处成立;

[注] 在此不等式以及下面的不等式中,字母z在原书中是用X'表示的,这里为避免与x'混淆,我们改用z。——译者注。

(ii) 对几乎所有 $t\in[t_0, t_1]$, 有

$$\phi'(t)=H_p(t, \phi(t), u(t), \hat{\lambda}(t)),$$
$$\lambda'(t)=-H_x(t, \phi(t), u(t), \hat{\lambda}(t))-R_x(t, \phi(t), u(t))\nu(t),$$
$$H_z(t, \phi(t), u(t), \hat{\lambda}(t))+R_z(t, \phi(t), u(t))\nu(t)=0,$$
$$\nu^i(t)R^i(t, \phi(t), u(t))=0, \qquad i=1, \cdots, r;$$

(iii) 对几乎所有 $t\in[t_0, t_1]$,

$$H(t, \phi(t), u(t), \hat{\lambda}(t))\geqslant H(t, \phi(t), z, \hat{\lambda}(t))$$

对所有满足 $R(t, \phi(t), z)\geqslant 0$ 的 z 成立;

(iv) 如习题 3.5 中给定的横截条件成立;

(v) 对 $[t_0, t_1]$ 中的每一个 t 值，令 $\hat{R}$ 表示由 R 在 $(t, \phi(t), u(t))$ 处取零值的那些分量所构成的向量. 则对几乎所有 t,

$$\langle e, (H(\pi(t))+\langle\nu(t), R(t, \phi(t), u(t))\rangle)_{zz}e\rangle\geqslant 0$$

对所有使 $\hat{R}_z(t, \phi(t), u(t))e=0$ 的向量 $e=(e^1, \cdots, e^m)$ 成立.

6. 关于状态变量为线性的系统

本节应用最大值原理来解决在满足状态方程

$$\frac{dx}{dt}=A(t)x+h(t, u(t)), \tag{6.1}$$

控制约束 Ω 和端点约束 $\mathscr{B}$ 的条件下，使

$$J(\phi, u)=g(e(\phi))+\int_{t_0}^{t_1}(\langle a_0(s), \phi(s)\rangle+h_0(s, u(s)))ds$$

取极小值的问题. 本节将做如下假设.

假设 6.1 (i) 约束映射 Ω 是常值映射; 亦即对所有 t, $\Omega(t)=\mathscr{C}$, 其中 $\mathscr{C}$ 是 E^m 中的固定集合; (ii) 集合 $\mathscr{B}$ 是 q 维的 $C^{(2)}$ 流形, 其中 $0\leqslant q\leqslant 2n+1$; (iii) 向量函数 a_0 和矩阵函数 A 是有界可测的; (iv) 函数 h_0: $(t, z)\to h_0(t, z)$ 和 h: $(t, z)\to h(t, z)$ 关于 t 可测而关于 z 是连续的; (v) 函数 g 在 E^{2n+2} 中 $\mathscr{B}$ 的邻域里是 $C^{(1)}$ 类的.

假定上述问题的最优对 (ϕ, u) 存在, 其定义区间为 $[t_0, t_1]$. 关于 (ϕ, u) 我们做如下假设.

假设 6.2 (i)在 $t=t_0$ 和 $t=t_1$ 处映射

$$t\rightarrow A(t)\phi(t)+h(t,\ u(t)),$$

$$t\rightarrow\langle a_0(t),\ \phi(t)\rangle+h_0(t,\ u(t))$$

是连续的. 自然,这使得我们能够应用横截条件.

(ii) 在端点 $e(\phi)$ 上向量

$$(g_{t_0},\ g_{x_0},\ g_{t_1},\ g_{x_1}) \tag{6.2}$$

既不为零也不正交于 $\mathscr{B}$.

定理 IV.6.3 是关于状态变量是线性的系统的存在性定理. 在此定理中矩阵 A 和函数 h, h_0 和 a_0 假定是连续的. 如果 A、a_0、h_0 和 h 满足假设 6.1 的条件,容易验证定理 IV.6.3 是成立的.

正如定理 IV.6.3 的证明中第一句话所指出的,不失一般性,我们可以假设 $f^0\equiv 0$. 于是泛函 J 由

$$J(\phi,\ u)=g(e(\phi)) \tag{6.3}$$

给定. 往后我们都假设 J 是用(6.3)表示的.

现在我们研究本问题的最大值原理的形式. 函数 H 变成

$$H(t,\ x,\ z,\ \hat{p})=\langle p,\ A(t)x\rangle+\langle p,\ h(t,\ z)\rangle. \tag{6.4}$$

因此,若以 a^{ij} 表示 A 的第 i 行第 j 列元素,即可得

$$H_{x^j}(t,\ x,\ z,\ \hat{p})=\sum_{i=1}^{n}p^i a^{ij}(t).$$

所以(3.6)中的第二个方程变成

$$\frac{dp}{dt}=-A(t)^*p, \tag{6.5}$$

其中星号表示"转置".

设 λ 是方程(6.5)之满足初始条件

$$\lambda(t_0)=\eta \tag{6.6}$$

的解. Ψ 表示(6.5)满足 $\Psi(t_0)=I$ 的基本解矩阵,则

$$\Psi'(t)=-A^*(t)\Psi(t),\ \Psi(t_0)=I, \tag{6.7}$$

并且有

$$\lambda(t)=\Psi(t)\eta. \tag{6.8}$$

系统(6.5)是系统

$$\frac{dx}{dt}=A(t)x \tag{6.9}$$

的共轭系统. 设 Φ 表示(6.9)的满足条件 $\Phi(t_0)=I$ 的基本解矩阵，则由乘积的微分法则和(6.9)式可得

$$(\Phi^*\Psi)'=(\Phi^*)'\Psi+\Phi^*\Psi'=(\Phi^*A^*)\Psi+\Phi^*(-A^*\Psi)=0.$$

于是对所有的 t, $\Phi^*\Psi=C$, 其中 C 是常数矩阵. 但是 $\Phi^*(t_0)=\Psi(t_0)=I$, 所以 $C=I$, 从而可得

$$\Psi(t)=\Phi^{-1*}(t) \tag{6.10}$$

对所有的 t 成立. 因此方程(6.8)又可写成

$$\lambda(t)=\Phi^{-1*}(t)\eta.$$

对于 g 不恒等于 0 的问题的横截条件已在习题 3.5 中给出，如果令

$$f^*(t_i)=A(t_i)\phi(t_i)+h(t_i,\ u(t_i)),\quad i=0,\ 1,$$

则习题 3.5 中给出的横截条件在现在的情形取如下形式，向量

$$(\langle\lambda(t_0),\ f^*(t_0)\rangle-\lambda^0 g_{t_0},\ -\lambda(t_0)-\lambda^0 g_{x_0},$$
$$-\langle\lambda(t_1),\ f^*(t_1)\rangle-\lambda^0 g_{t_1},\ \lambda(t_1)-\lambda^0 g_{x_1})$$

在点 $e(\phi)$ 处正交于 $\mathscr{B}$, 其中 g 的所有偏导数是在 $e(\phi)$ 处取值的. 按照(6.6)和(6.8), 此向量也可以写成

$$(\langle\eta,\ f^*(t_0)\rangle-\lambda^0 g_{t_0},\ -\eta-\lambda^0 g_{x_0},\ -\langle\Psi(t_1)\eta,\ f^*(t_1)\rangle-\lambda^0 g_{t_1},$$
$$\Psi(t_1)\eta-\lambda^0 g_{x_1}). \tag{6.11}$$

现在证明 $\eta\neq 0$. 因为若 $\eta=0$, 那么由(6.8)可见在 $[t_0,\ t_1]$ 上 $\lambda(t)=0$. 因此按最大值原理 $\lambda^0\neq 0$. 而且若 $\eta=0$, (6.11)变成

$$\lambda^0(g_{t_0},\ g_{x_0},\ g_{t_1},\ g_{x_1}).$$

由于 $\lambda^0\neq 0$, 则横截条件蕴含着向量(6.2)或者是零，或者在点 $e(\phi)$ 正交于 $\mathscr{B}$. 然而，这在开始时已被排除在外.

习题 6.1 试用向量(6.11)与 $\mathscr{B}$ 在 $e(\phi)$ 处的 q 个线性无关切向量中的每一个向量作内积，并令其等于 0, 以得到将横截条件表示成关于未知量 $(\lambda^0,\ \eta^1,\ \cdots,\ \eta^n)$ 的 q 个线性方程的方程组的解析表达式.

在现在的讨论中，对所有的 t, $\Omega(t)=\mathscr{C}$, 所以推论 3.1 可套

用．从(6.4)式我们看出在目前的讨论中，(3.5)等价于下面的不等式

$$\langle\lambda(t),\ h(t,\ u(t))\rangle\geqslant\langle\lambda(t),\ h(t,\ z)\rangle \tag{6.12}$$

对$[t_0,\ t_1]$中几乎所有的t和$\mathscr{C}$中所有的z成立．从(6.8)和(6.10)我们可得

$$\begin{aligned}\langle\lambda(t),\ h(t,\ z)\rangle&=\langle\Psi\eta,\ h(t,\ z)\rangle=\langle\eta,\ \Psi^*(t)h(t,\ z)\rangle\\&=\langle\eta,\ \Phi^{-1}(t)h(t,\ z)\rangle.\end{aligned}$$

因此，(6.12)等价于不等式

$$\langle\eta,\ \Phi^{-1}(t)h(t,\ u(t))\rangle\geqslant\langle\eta,\ \Phi^{-1}(t)h(t,\ z)\rangle. \tag{6.13}$$

我们用下面的定理来概括前面的讨论，它给出关于状态变量是线性的系统的最大值原理．

定理 6.1 设$(\phi,\ u)$满足状态方程(6.1)，控制约束Ω和端点集合$\mathscr{B}$，并使泛函(6.3)取极小值；又设假设 6.1 和 6.2 成立，令Φ表示满足$\Phi(t_0)=I$的(6.9)的基本解矩阵．则存在E^n中的非零向量η和纯量$\lambda^0\leqslant0$，使得向量(6.11)在$e(\phi)$处正交于$\mathscr{B}$，并使之对$[t_0,\ t_1]$中几乎所有的t，

$$\max\{\langle\eta,\ \Phi^{-1}(t)h(t,\ z)\rangle : z\in\mathscr{C}\}$$

在$z=u(t)$上达到．

附注 6.1 依照(6.10)，上面待取最大值的量也可以写成$\langle\eta,\ \Psi^*(t)h(t,\ z)\rangle$.

附注 6.2 注意，在原则上对于状态变量是线性的系统，为了确定极值轨线，我们仅需函数λ的初始值η．当然，知道λ的终止值也行，因为按照(6.8)，$\eta=\Psi^{-1}(t_1)\lambda(t_1)$．为了确定$\eta$，我们可利用横截条件，这包含要知道$e(\phi)$和$u(t_i)$ $(i=0,\ 1)$之值．从而，一旦$u(t)$已知，由参数变易公式即可给出极值轨线．

7. 线性系统

线性系统是指系统(6.1)中的函数h由

$$h(t,\ z)=B(t)z+d(t) \tag{7.1}$$

给出，其中 B 是 $n\times m$ 矩阵，d 是 n 维向量．因此线性系统的状态方程为

$$\frac{dx}{dt}=A(t)x+B(t)z+d(t), \tag{7.2}$$

为了使假设 6.1 的(iii)和(iv)成立，B 和 d 必须是有界可测函数．

线性系统的最大值原理是定理 6.1 和(7.1)式的直接推论．

定理 7.1 设 (ϕ, u) 满足状态方程(7.2)、控制约束 Ω 和端点约束 $\mathscr{B}$，并使泛函(6.3)取极小值；若假设 6.1 和 6.2 成立，令 Φ 为(6.9)之满足 $\Phi(t_0)=I$ 的基本解矩阵．则存在 E^n 中的非零向量 η 和纯量 $\lambda^0\leqslant 0$，使得向量(6.11)在 $e(\phi)$ 处正交于 $\mathscr{B}$，并使得对几乎所有 $t\in[t_0, t_1]$，

$$\max\{\langle\eta\Phi^{-1}(t)B(t), z\rangle : z\in\mathscr{C}\} \tag{7.3}$$

在 $z=u(t)$ 上达到．

附注 7.1 根据(6.10)，可以用

$$\max\{\langle\eta\Psi^*(t)B(t), z\rangle : z\in\mathscr{C}\}$$

代替(7.3)．在这两种情况中，我们都得到当 $z=u(t)$ 时，使得 z 的一个线性形式在集合 $\mathscr{C}$ 上取到最大值．为了强调这一点，我们令

$$L(t, \eta, z)=\langle\eta\Phi^{-1}(t)B(t), z\rangle=\langle\eta\Psi^*(t)B(t), z\rangle. \tag{7.4}$$

设 ψ_j 表示基本解矩阵 Ψ 的第 j 列向量函数，$\eta=(\eta^1,\cdots,\eta^n)$，$B^j(t)$ 表示 $B(t)$ 的第 j 列，而 b^{ik} 表示 B 的第 i 行第 k 列元素．则(7.4)中 z^j 的系数为

$$\sum_{i,k=1}^{n}\eta^i\psi_i^k(t)b^{kj}(t)=\eta\Psi^*(t)B^j(t).$$

在许多问题中，集合 $\mathscr{C}$ 是下述立方体：

$$\mathscr{C}=\{z: |z^i|<1,\ i=1,\ \cdots,\ m\}. \tag{7.5}$$

在这种情形，最优控制常常可以表述得很简单．

推论 7.1 设 $\mathscr{C}$ 是由(7.5)给定的，对每个 $j=1, \cdots, m$，设

$$E_j(\eta)=\{t: \eta\Psi^*(t)B^j(t)=0\}^{[注]}$$

具有零测度．则对几乎所有 $t\in[t_0, t_1]$，

［注］原书误为 $\eta\Psi^*(t)B^j(t)\neq 0$．——译者注．

$$u^j(t)=\operatorname{sign}\eta\Psi^*(t)B^j(t). \qquad (7.6)$$

因为对几乎所有 $t\in[t_0, t_1]$，$u(t)$ 使映射 $z\to L(t, \eta, z)$ 在 $\mathscr{C}$ 上取最大值．又由于 $\eta\Psi^*(t)B^j(t)$ 是 z^j 的系数，故推论为真．

定义 7.1 线性系统(7.2)称为在区间 $[t_0, t_1]$ 上关于 $\mathscr{C}$ 是正规的，如果对每一个非零 n 维向量 μ 和几乎所有 $t\in[t_0, t_1]$，最大值

$$\max\{L(t, \mu, z): z\in\mathscr{C}\}$$

在 $\mathscr{C}$ 中的唯一点 $z^*(t)$ 上达到．

注意，一个系统是否在给定区间上是正规的，是由矩阵 A 和 B 以及约束集合 $\mathscr{C}$ 所确定的．本节末了我们将建立包含 A、B 和 $\mathscr{C}$ 的条件的比较容易验证的正规性准则．

如果 $\mathscr{C}$ 是由(7.5)给定的，则系统是正规的充要条件是对 E^n 中的每一个 μ 和每一个 $j=1, \cdots, m$，集合 $E_j(\mu)$ 具有零测度．因此推论 7.1 说明如果一个系统关于 $\mathscr{C}$ 是正规的且 $\mathscr{C}$ 是由(7.5)给定的，则最优控制由(7.6)给出．

我们现在研究当 $\mathscr{C}$ 是致密凸集时，最优控制的结构．如果 $\mathscr{C}$ 是致密的和凸的，由 Krein-Milman 定理(引理 IV.6.2)，集合 $\mathscr{C}$ 的极端点集合 $\mathscr{C}_e$ 是非空的．下面定理 7.1 的推论成立．

推论 7.2 设 $\mathscr{C}$ 是致密的凸集，系统(7.2)是正规的，若 $u(t)$ 是最优控制，则对几乎所有 $t\in[t_0, t_1]$，$u(t)\in\mathscr{C}_e$．

证明：如果结论不真，则对一个具有正测度的集合 E 中的 t，$u(t)\notin\mathscr{C}_e$．因此对 $t\in E$，存在 $\mathscr{C}$ 中的点 $z_1(t)$ 和 $z_2(t)$ 以及满足 $\alpha(t)+\beta(t)=1$ 的实数 $\alpha(t)>0$，$\beta(t)>0$，使得 $u(t)=\alpha(t)z_1(t)+\beta(t)z_2(t)$．因为系统是正规的，对 E 中几乎所有的 t，线性函数 $L(t, \eta, \cdot)$ 在 $\mathscr{C}$ 中的唯一点 $z^*(t)$ 上达到它的最大值．又由最大值原理该最大值应在 $u(t)$ 上达到，所以 $z^*(t)=u(t)$．于是

$$L(t, \eta, u(t))>L(t, \eta, z_1(t)), \quad L(t, \eta, u(t))>L(t, \eta, z_2(t))$$

在 E 中几乎处处成立．从而

$$\begin{aligned}L(t, \eta, u(t))&=\alpha(t)L(t, \eta, u(t))+\beta(t)L(t, \eta, u(t))\\&>\alpha(t)L(t, \eta, z_1(t))+\beta(t)L(t, \eta, z_2(t))\end{aligned}$$

$$=L(t, \eta, \alpha(t)z_1(t)+\beta(t)z_2(t))$$
$$=L(t, \eta, u(t)).$$

这就产生矛盾.

定义 7.2 设 $\mathscr{C}$ 是以 $e_1, \cdots, e_k$ 为顶点的致密多面体 $\mathscr{P}$, 控制 u 称为区间 $[t_0, t_1]$ 上的 bang-bang 控制, 如果对几乎所有 $t\in[t_0, t_1]$, $u(t)$ 等于这些顶点之一.

如果 $\mathscr{C}=\mathscr{P}$, 推论 7.2 可以重新叙述如下.

推论 7.3 设系统是正规的, 约束集合是致密的多面体 $\mathscr{P}$, 则任何最优控制必是 bang-bang 控制.

附注 7.2 bang-bang 控制原理(定理 IV.6.4)告诉我们, 如果没有假设系统是正规的, 若 u 是一个最优控制, 则存在另一个 bang-bang 的最优控制. 另一方面, 推论 7.3 告诉我们, 如果系统是正规的, 则任何最优控制必为 bang-bang 控制. 同样地, 对于任意的致密凸的约束集合 $\mathscr{C}$, bang-bang 原理告诉我们, 若 u 是最优控制, 则存在最优控制 u^*, 使得 $u^*\in\mathscr{C}_e$. 如果系统是正规的, 则推论 7.2 说, 假若 u 是最优控制, 则 u 本身必具有性质 $u(t)\in\mathscr{C}_e$.

上述结果并不保证正规系统最优控制的唯一性. 下面的定理给出最优控制是唯一的适当条件.

定理 7.2 设 $\mathscr{C}$ 是致密的凸集, 系统是正规的, $\mathscr{B}$ 是 E^{2n+2} 中一个线性流形的相对开的凸子集, g 由

$$g(t_0, x_0, t_1, x_1)=g_1(x_0, x_1)+g_2(t_0, t_1) \tag{7.7}$$

给定, 其中 g_1 是凸的. 设 u_1 和 u_2 是两个定义在同一区间 $[t_0, t_1]$ 上的最优控制. 则在 $[t_0, t_1]$ 上几乎处处有 $u_1(t)=u_2(t)$.

证明: 设 ϕ_1 是对应于 u_1 的轨线, ϕ_2 是对应于 u_2 的轨线. 定义 $u_3(t)=(u_1(t)+u_2(t))/2$, 因为 $\mathscr{C}$ 是凸的, $u_3\in\mathscr{C}$. 设 ϕ_3 是对应于 u_3 的轨线, 它满足初始条件 $\phi_3(t_0)=(\phi_1(t_0)+\phi_2(t_0))/2$. 则

$$\phi_3(t)=\Phi(t)\left\{\phi_3(t_0)+\frac{1}{2}\int_{t_0}^{t}\Phi^{-1}(s)\{[B(s)u_1(s)+B(s)u_2(s)]+2d(s)\}ds\right\}=(\phi_1(t)+\phi_2(t))/2,$$

且 $$e(\phi_3)=(e(\phi_1)+e(\phi_2))/2.$$
因为 $\mathscr{B}$ 是线性流形的凸子集，可见 $e(\phi_3)\in\mathscr{B}$. 因此 (ϕ_3, u_3) 是容许对.

设 $\mu=\inf\{J(\phi, u):(\phi, u)$ 是容许对$\}$，则
$$\mu=J(\phi_1, u_1)=J(\phi_2, u_2).$$
从 μ 的定义，由(6.3)[注]、(7.7)以及 g_1 的凸性，且 ϕ_1 和 ϕ_2 具有相同起始和终止时间的假设，我们可得
$$\mu\leqslant J(\phi_3, u_3)=g(e(\phi_3))=g[(e(\phi_1)+e(\phi_2))/2]$$
$$\leqslant\frac{1}{2}g(e(\phi_1))+\frac{1}{2}g(e(\phi_2))=\mu.$$
因此 $J(\phi_3, u_3)=\mu$，从而 (ϕ_3, u_3) 是最优对. 由推论 7.2, $u_3(t)\in\mathscr{C}$，几乎处处成立. 这与 u_3 的定义相矛盾，除非 $u_1(t)=u_2(t)$ 几乎处处成立.

附注 7.3 对于具有固定的 t_0 和 t_1 的问题，g 自然有形式(7.7)，其中 $g_2\equiv 0$. 如果我们假设 g 是 (t_0, x_0, t_1, x_1) 的凸函数，则 g 具有形式(7.7)的假设即可放弃.

定义 7.3 线性系统(7.2)称为在区间 $[t_0, t_1]$ 上关于约束集合 $\mathscr{C}$ 是强正规的，如果对 E^n 中的每个非零向量 μ，除了 $[t_0, t_1]$ 中的有限个点的集合外，对此区间上所有其它的点，$\max\{L(t,\mu,z): z\in\mathscr{C}\}$ 在 $\mathscr{C}$ 中的唯一点 $z^*(t)$ 上达到.

定义 7.4 控制 u 称为在区间 $[t_0, t_1]$ 上分段常值的，如果存在有限个互不相交的开子区间 (τ_j, τ_{j+1})，使得闭子区间 $[\tau_j, \tau_{j+1}]$ 的并集为 $[t_0, t_1]$，且 u 在每一个开子区间 (τ_j, τ_{j+1}) 上取常数.

下面的定理给出具有实际意义的强正规系统之最优控制的特征. 关于强正规性的简单判别法将在定理 7.4 及其推论中给出.

定理 7.3 设 (ϕ, u) 是最优对，假设 6.1 和 6.2 成立，矩阵 B 是连续的，约束集 $\mathscr{C}$ 是致密的多面体 $\mathscr{P}$，系统(7.2)在 (ϕ, u) 的定义区间 $[t_0, t_1]$ 上是强正规的. 则 u 在 $[t_0, t_1]$ 上是分段常值的，其值取在 $\mathscr{P}$ 的顶点上.

[注] 原书误为(6.6).——译者注.

证明：如果我们去掉 t_0, t_1 和 $L(t, \eta, z)$ 不能在唯一点 $z^*(t)$ 达到最大值的那有限个点的集合，我们得到互不相交的有限个开子区间 (τ_j, τ_{j+1})，使得闭区间 $[\tau_j, \tau_{j+1}]$ 的并集为 $[t_0, t_1]$．设 J 表示区间 (τ_j, τ_{j+1}) 中的一个，从推论 7.2 的证明中可见对 J 中的每一个 t，$u(t)$ 等于 $\mathscr{P}$ 的顶点 $e_i(i=1, \cdots, k)$ 中的一个．设 M_i 表示 J 中满足 $u(t)=e_i$ 的 t 的集合，则 M_i, $i=1, \cdots, k$，不全部是空的，集 M_i 是两两不相交的，且 $J=\cup M_i$．我们现在证明如果 M_i 是不空的，则它是开集．因为若 $\tau \in M_i$，则

$$L(\tau, \eta, e_i) > L(\tau, \eta, e_j), \quad 对所有\ j \neq i. \tag{7.8}$$

因为对固定的 η, e_i，映射 $t \to L(t, \eta, e_i)$ 是连续的，于是 (7.8) 在 τ 的邻域中成立，从而该邻域中所有的点均在 M_i 中，所以 M_i 是开的．因为 J 是连通的，又因为 $J=\cup M_j$，其中 M_j 是两两不相交的开集．可见对 $j \neq i$，M_j 必为空集，于是在 J 中 $u(t)=e_i$，定理证毕．

定理 7.3 的结论比推论 7.3 的结论强得多，这里我们断言最优控制是分段常值的，其值取在 $\mathscr{P}$ 的顶点 $e_1, \cdots, e_k$ 上，而在推论 7.3 中我们仅能断定最优控制是可测的，其值取在 $\mathscr{P}$ 的诸顶点上．当然，这里的前提也更严格了．

下面我们介绍强正规性的判别准则作为本节的结束．

定理 7.4 设状态方程为 (7.2)，A 在致密区间 $\mathscr{I}$ 上是 $C^{(n-2)}$ 类的，而 B 在 $\mathscr{I}$ 上是 $C^{(n-1)}$ 类的，控制的约束集合是致密的多面体 $\mathscr{P}$．设

$$B_1(t) = B(t) \tag{7.9}$$

$$B_j(t) = -A(t)B_{j-1} + B'_{j-1}(t), \quad j=2, \cdots, n.$$

如果对 E^m 中每一个平行于 $\mathscr{P}$ 的棱的向量 w，向量组

$$B_1(t)w,\ B_2(t)w,\ \cdots,\ B_n(t)w \tag{7.10}$$

对所有 $t \in \mathscr{I}$ 是线性无关的，则系统 (7.2) 在 $\mathscr{I}$ 上关于 $\mathscr{P}$ 是强正规的．

证明：假如结论不真，则存在 E^n 中的非零向量 η 和 $\mathscr{I}$ 中的无穷点集 E 使得对 E 中的 t，$L(t, \eta, z)$ 在 $\mathscr{P}$ 上的最大值不是在 $\mathscr{P}$

中的唯一点 $z^*(t)$ 上达到. 因为对固定的 (t, η) 映射 $z\to L(t, \eta, z)$ 是线性的, 而且 $\mathscr{P}$ 是致密的多面体, 故 $L(t, \eta, z)$ 在 $\mathscr{P}$ 上的最大值应在 $\mathscr{P}$ 的某一面上达到. 又因为 $\mathscr{P}$ 只有有限个面, 所以存在一个无穷点集 $E_1\subset E$ 和 $\mathscr{P}$ 的一个面 $\mathscr{P}_F$, 使得对 $t\in E_1$, $L(t, \eta, z)$ 在整个 $\mathscr{P}$ 上的最大值于 $\mathscr{P}_F$ 上取到. 因此, 若 e_1 和 e_2 是 $\mathscr{P}_F$ 中的两个不同顶点, 则对所有 $t\in E_1$, $L(t, \eta, e_1)=L(t, \eta, e_2)$ 成立, 那么如果取 $w=e_1-e_2$, 就得

$$L(t, \eta, w)=\langle \eta\Psi^*(t), Bw\rangle=0$$

对所有 $t\in E_1$ 成立. 从(7.9)中的第一个方程我们又可得

$$L(t, \eta, w)=\langle \eta\Psi^*(t), B_1(t)w\rangle=0 \tag{7.11}$$

对所有的 $t\in E_1$ 成立. 因为 E_1 是无穷点集且 $\mathscr{I}$ 是致密的, 于是在 $\mathscr{I}$ 中必有 E_1 的极限点 τ, 从(7.11)和 B_1、Ψ^* 的连续性可得

$$L(\tau, \eta, w)=\langle \eta\Psi^*(\tau), B_1(\tau)w\rangle=0. \tag{7.12}$$

由假设矩阵 A 是 $C^{(n-2)}$ 类的, 因此, 系统(6.9)的共轭系统的基本解矩阵是 $C^{(n-1)}$ 类的, 因为 $B_1=B$ 而 B 假设是 $C^{(n-1)}$ 类的. 故从(7.11)的第一个等号可见, 映射 $t\to L(t, \eta, w)$ 在 $\mathscr{I}$ 上是 $C^{(n-1)}$ 类, 且有

$$L'(t, \eta, w)=\langle \eta\Psi^{*\prime}(t), B_1(t)w\rangle+\langle \eta\Psi^*(t), B_1'(t)w\rangle,$$

从(6.7)我们得知

$$\Psi^{*\prime}(t)=-\Psi^*(t)A(t),$$

如果把此关系式代入上面的方程, 即得

$$L'(t, \eta, w)=\langle \eta\Psi^*(t), (-A(t)B_1(t)+B_1'(t))w\rangle.$$

根据(7.9)的第二个方程, 我们得到

$$L'(t, \eta, w)=\langle \eta\Psi^*(t), B_2w\rangle. \tag{7.13}$$

在函数的任意两个零点之间函数的导函数必有一个零点, 因此, 对以 τ 为极限点的无穷点集 E_2 中所有 t 值, 有 $L'(t, \eta, w)=0$, 从(7.13)和 $\Psi^*(t)$、$B_2(t)$ 的连续性可见

$$\langle \eta\Psi^*(\tau), B_2(\tau)w\rangle=0.$$

我们可以按此方式归纳而得

$$\langle \eta\Psi^*(\tau), B_j(\tau)w\rangle=0, \quad j=1, \cdots, n.$$

因为 n 个向量 $B_1w, B_2w, \cdots, B_nw$ 假设是线性无关的，故必有 $\eta\Psi^*(\tau)=0$，但这是不可能的，因为 $\eta\neq0$ 且 $\psi^*(\tau)$是非奇异的．这个矛盾证明定理为真．

推论 7.4.1 设 A 和 B 是常数矩阵，如果对 E^m 中的每一个平行于 $\mathscr{P}$ 的棱的向量 w，向量 $Bw, ABw, A^2Bw, \cdots, A^{n-1}Bw$ 是线性无关的，则系统(7.2)在 $\mathscr{I}$ 上关于 $\mathscr{P}$ 是强正规的．

本推论可以从如下的观察而得知：如果 A 和 B 是常数矩阵，则

$$B_j=(-A)^{j-1}B, \qquad j=1, 2, \cdots, n.$$

如果集合 $\mathscr{P}$ 是其棱平行于坐标轴的平行六面体，则只需把我们所论的向量 w 看成是 E^m 中的标准基底向量 $w_1, \cdots, w_m$，其中 w_i 是第 i 个支量为 1 而其他支量为 0 的 m 维向量．设 b^j 表示矩阵 B 的第 j 列．则 $b^j=Bw_j$，而且推论 7.4.1 给出如下的推论：

推论 7.4.2 设 A 和 B 是常数矩阵，且 $\mathscr{P}$ 是平行六面体，它的棱平行于坐标轴．设 b^j 表示 B 的第 j 列，对每一个 $j=1, \cdots, m$，设

$$b^j, Ab^j, A^2b^j, \cdots, A^{n-1}b^j$$

线性无关，则系统(7.2)在 $\mathscr{I}$ 上关于 $\mathscr{P}$ 是强正规的．

8．线性时间最优问题

线性时间最优问题，就是要求按线性系统的发展规律，把一个给定的点 x_0 用最短时间转移到另一个给定的点 x_1．更精确地说，所谓“线性时间最优问题”，乃是要求在满足状态方程(7.2)，控制约束 Ω 和端点条件 $\mathscr{B}$ 的条件下，使

$$J(\phi, u)=t_1-t_0\text{[注]}$$

取极小值．现在的函数 g 取 $g(t_1)=t_1-t_0$，其中

$$\mathscr{B}=\{(t_0, x_0, t_1, x_1): t_0=t_0', x_0=x_0', x_1=x_1'\},$$

而 t_0', x_0', x_1' 是事先给定的．

如果 $\Omega(t)=\mathscr{C}$，其中 $\mathscr{C}$ 是固定的致密凸集，并假设系统(7.2)

[注] 原书误为 $J(\phi, u)=t_1$．——译者注．

关于 $\mathscr{C}$ 是正规的，则由推论 7.2，任何最优控制 u，对几乎所有 t 都具有形式 $u(t)\in\mathscr{C}_e$. 如果 $\mathscr{C}$ 是致密的多面体 $\mathscr{P}$，则任何最优控制 u 是 bang-bang 控制. 我们还可断言它是唯一的. 如果 u_1 和 u_2 是两个不同的最优控制，则因为问题是使 t_1-t_0 最小化，可见 u_1 和 u_2 都定义在同一个区间 $[t_0, t_1^*]$ 上，其中 t_1^* 是所需的最小时刻. 因为函数 g 现在由 $g(t_1)=t_1-t_0$ 给出，它是取 $g_1\equiv 0$ 的(7.7)形式. 注意到 g_1 是凸的，并且对所有的 t_1，有

$$\nabla g(t_1)=(0,\ 0_n,\ 1,\ 0_n), \tag{8.1}$$

其中 0_n 是 n 维零向量. 在此情形中，(8.1)右边的向量也是 $\mathscr{B}$ 的单位切向量. 因此 $\nabla g(t_1)\neq 0$ 且 $\nabla g(t_1)$ 不正交于 $\mathscr{B}$. 最后，$\mathscr{B}$ 是 n 维的线性流形. 因此定理 7.2 的所有假设都满足，故 $u_1(t)=u_2(t)$. 现把我们的结果概括成下面的定理.

定理 8.1 在线性时间最优问题中，如果控制约束集是固定的致密凸集 $\mathscr{C}$，系统关于 $\mathscr{C}$ 是正规的，则最优控制 u 是唯一的，并且对几乎所有的 t，有 $u(t)\in\mathscr{C}_e$.

大量线性时间最优问题具有唯一极值控制的性质，这个性质也就保证了由存在性定理断定存在的极值控制必为最优的. 实际上，对这类系统的问题，无须引用存在性定理，可以使用证明极值控制唯一性的方法，直接证明极值控制是唯一的和最优的.

定理 8.2 设系统方程是由(7.2)中取 $d\equiv 0$ 给出的；$\mathscr{C}$ 是致密的凸集，它包含 E^m 的原点为其内点；设系统关于 $\mathscr{C}$ 是正规的；又设 (ϕ_1, u_1) 是以 $x_1=0$ 为终止状态的时间最优问题的极值对，ϕ_1 达到原点的终止时间为 t_1-t_0；设 (ϕ_2, u_2) 是容许对，它用时间 t_2-t_0 将 x_0 转移到原点，则 $t_2\geqslant t_1$，等号成立的充要条件是对几乎所有 t，$u_1(t)=u_2(t)$.

证明：假设存在容许对 (ϕ_2, u_2)，对它来说 $t_2<t_1$，由参数变易公式可得

$$0=\Phi(t_1)\left\{x_0+\int_{t_0}^{t_1}\Phi^{-1}(s)B(s)u_1(s)ds\right\},$$

$$0=\Phi(t_2)\left\{x_0+\int_{t_0}^{t_2}\Phi^{-1}(s)B(s)u_2(s)ds\right\},$$

其中 $\Phi(t)$是系统(6.9)满足 $\Phi(t_0)=I$ 的基本解矩阵，如果我们用 $\Phi^{-1}(t_1)$左乘第一个方程，用 $\Phi^{-1}(t_2)$左乘第二个方程即得

$$-x_0=\int_{t_0}^{t_1}\Phi^{-1}(s)B(s)u_1(s)ds=\int_{t_0}^{t_2}\Phi^{-1}(s)B(s)u_2(s)ds. \tag{8.2}$$

因为 u_1 是极值控制，存在 E^n 中的非零向量 η 使得对$[t_0, t_1]$中几乎所有的 t，$u_1(t)$使 $L(t, \eta, z)$在 $\mathscr{C}$ 上达到最大值。如果按(8.2)计算$\langle\eta, -x_0\rangle$，可得

$$\int_{t_0}^{t_2}+\int_{t_2}^{t_1}\langle\eta\Phi^{-1}(s)B(s), u_1(s)\rangle ds$$

$$=\int_{t_0}^{t_2}\langle\eta\Phi^{-1}(s)B(s), u_2(s)\rangle ds.$$

因此

$$\int_{t_0}^{t_2}\{L(s, \eta, u_1(s))-L(s, \eta, u_2(s))\}ds$$

$$=-\int_{t_2}^{t_1}L(s, \eta, u_1(s))ds. \tag{8.3}$$

因为 u_1 是极值控制且系统关于 $\mathscr{C}$ 是正规的，$u_1(t)\in\mathscr{C}_e$ 几乎处处成立，又由于 0 是 $\mathscr{C}$ 的内点，$u_1(t)\neq 0$ 几乎处处成立，且

$$L(t, \eta(t), u_1(t))>L(t, \eta, 0)=0.$$

因此(8.3)的右边$\leqslant 0$. 当且仅当 $t_1=t_2$ 时等号成立。另一方面，因为系统是正规的，于是

$$L(t, \eta, u_1(t))\geqslant L(t, \eta, u_2(t))$$

对 t 几乎处处成立。等号成立当且仅当 $u_1(t)=u_2(t)$几乎处处成立。于是(8.3)左边的积分 $\geqslant 0$，等号成立当且仅当 $u_1(t)=u_2(t)$几乎处处成立。因此(8.3)[注]的每一边都等于0，这就推出 $t_2=t_1$，而且 $u_1(t)=u_2(t)$几乎处处成立。定理证毕。

9. 线性二次准则问题

本节所研究的这类问题，其状态方程为

[注] 原书误为(8.2)。——译者注。

$$\frac{dx}{dt}=A(t)x+B(t)z+d(t) \tag{9.1}$$

函数 f^0 由

$$f^0(t, x, z)=\langle x, Q(t)x\rangle^{[注]}+\langle z, R(t)z\rangle \tag{9.2}$$

给定，这种问题的存在定理已在第三章中给出．具有致密控制约束的问题在推论 III.5.1 中考虑过，具有无界控制的问题在习题 III.6.5 中也考虑过。

除非另有说明，下面的假设在本节都有效．

假设 9.1 (i) (9.1)和(9.2)中的矩阵 A, B, Q, R 和(9.1)中的向量 d 都是区间 $[a, b]$ 上的连续函数．

(ii) 对 $[a, b]$ 中的每一个 t, 矩阵 $Q(t)$ 是对称半正定的，矩阵 $R(t)$ 是对称正定的．

(iii) 对 $[a, b]$ 中的每一个 t, $\Omega(t)=\mathscr{O}$, 其中 $\mathscr{O}$ 是 E^n 中的固定开集．

(iv) 集 $\mathscr{B}$ 具有如下形式：

(t_0, x_0) 固定；$(t_1, x_1)\in\mathscr{T}_1$, 其中 $\mathscr{T}_1$ 是 n 维 $C^{(1)}$ 流形；t_0, $t_1\in[a, b]$.

(v) 函数 $g: (t_1, x_1)\to g(t_1, x_1)$ 在 E^{n+1} 上是 $C^{(1)}$ 类的．

所要研究的问题是在满足状态方程(9.1)，控制约束 Ω 和端点条件 $\mathscr{B}$ 的条件下，求

$$J(\phi, u)=g(t_1, \phi(t_1))+\frac{1}{2}\int_{t_0}^{t_1}\{\langle\phi(s), Q(s)\phi(s)\rangle + \langle u(s), R(s)u(s)\rangle\}ds \tag{9.3}$$

的极小值．其中问题的依据是满足假设 9.1.

我们现在借助于最大值原理来刻画最优对．函数 H 由

$$H(t, x, z, \hat{p})=\frac{p^0}{2}\{\langle x, Q(t)x\rangle+\langle z, R(t)z\rangle\} + \langle p, A(t)x\rangle+\langle p, B(t)z\rangle+\langle p, d(t)\rangle \tag{9.4}$$

给定，因此

$$H_x(t, x, z, \hat{p})=p^0Q(t)x+A^*(t)p.$$

[注] 原书字母 Q 是用 X 表示的，为了避免混淆我们改用 Q. ——译者注．

其中星号表示转置.

现在考虑最优对(ϕ, u). 我们做如下假设.

假设 9.2 轨线ϕ在点$(t_1, \phi(t_1))$上不与$\mathscr{T}_1$相切.

从(9.1)和(9.2)右边的形式以及假设9.1-(ii)可见, 为了能应用横截条件于本问题, 必须假设映射$t\to u(t)$在t_0和t_1处是连续的. 我们以后将看到, 用不包含横截条件的论证方法, 最优控制$u(t)$必须是连续的. 此刻, 我们假设已经证明了这一点.

按照假设9.1的(iv), 集合$\mathscr{B}$是由形如(t_0, x_0, t_1, x_1)的点所组成的n维流形, 其中(t_0, x_0)固定, 而(t_1, x_1)在指定的n维$C^{(1)}$流形$\mathscr{T}_1$中. 因此习题3.5中所给的横截性条件现在取如下形式: 向量

$$(-H(\pi(t_1))-\lambda^0 g_{t_1}, \lambda(t_1)-\lambda^0 g_{x_1})$$

在$(t_1, \phi(t_1))$处与$\mathscr{T}_1$正交. 其中g的偏导数是在$(t_1, \phi(t_1))$上取值的.

如果我们假设轨线ϕ在它的终止点$(t_1, \phi(t_1))$与$\mathscr{T}_1$不相切, 则$\lambda^0\neq 0$. 因为若$\lambda^0=0$, 横截条件表明

$$(-\langle\lambda(t_1), f(t_1)\rangle, \lambda(t_1)) \tag{9.5}$$

在$(t_1, \phi(t_1))$处与$\mathscr{T}_1$正交. 其中$f(t_1)$表示(9.1)的右边在$(t_1, \phi(t_1), u(t_1))$处的值. 曾用以建立习题3.2的结果的方法在这里是适用的, 并表明如果轨线ϕ在$(t_1, \phi(t_1))$不与$\mathscr{T}_1$相切, (9.5)所示的向量在该点不能与$\mathscr{T}_1$正交. 因为$\lambda^0\neq 0$, 由习题3.3可见, 我们可以取$\lambda^0=-1$. 于是横截条件可以叙述成: 向量

$$(g_{t_1}+f_1^0-\langle\lambda(t_1), f_1\rangle, g_{x_1}+\lambda(t_1)) \tag{9.6}$$

在$(t_1, \phi(t_1))$处与$\mathscr{T}_1$正交. 其中g的偏导数在$(t_1, \phi(t_1))$处取值, f_1^0表示(9.2)右边在$(t_1, \phi(t_1), u(t_1))$上的值, f_1表示(9.1)右边在$(t_1, \phi(t_1), u(t_1))$上的值. 注意, 利用习题3.2的结论就表明$\lambda(t_1)$是唯一的.

方程(3.2)现在变成

$$\begin{cases}\dfrac{d\phi}{dt}=A(t)\phi(t)+B(t)u(t)+d(t),\\ \dfrac{d\lambda}{dt}=Q(t)\phi(t)-A^*(t)\lambda(t).\end{cases}\tag{9.7}$$

因为控制约束集是固定的，(3.5)成立．从(9.4)我们看到，就现在的论题而言，(3.5)等价于不等式

$$-\frac{1}{2}\langle u(t),\ R(t)u(t)\rangle+\langle\lambda(t),\ B(t)u(t)\rangle$$
$$\geqslant-\frac{1}{2}\langle z,\ R(t)z\rangle+\langle\lambda(t),\ B(t)z\rangle$$

对所有 $z\in\mathscr{O}$ 和几乎所有 $t\in[t_0,\ t_1]$ 成立．因此，对几乎每一个 $t\in[t_0,\ t_1]$，映射

$$z\to-\frac{1}{2}\langle z,\ R(t)z\rangle+\langle\lambda(t),\ B(t)z\rangle\tag{9.8}$$

在 $\mathscr{O}$ 上的最大值必在 $z=u(t)$ 处达到．但是 $\mathscr{O}$ 是开集，所以映象(9.8)的导数在 $z=u(t)$ 处为 0，因此

$$-R(t)u(t)+B^*(t)\lambda(t)=0.$$

因为对所有 t，$R(t)$ 是非奇异的，由此我们得到

$$u(t)=R^{-1}(t)B^*(t)\lambda(t)，几乎处处成立.\tag{9.9}$$

注意，因为 B, R 和 λ 是连续的，最优控制也是连续的．

如果把(9.9)代入(9.7)中的第一个方程，从最大值原理我们得到如下的定理．

定理 9.1 设$(\phi,\ u)$是以$[t_0,\ t_1]$为其定义区间的最优对，假设 9.2 成立，则存在定义于$[t_0,\ t_1]$上的绝对连续函数 $\lambda=(\lambda^1,\ \cdots,\ \lambda^n)$，使得$(\phi,\ \lambda)$是线性系统

$$\begin{cases}\dfrac{dx}{dt}=A(t)x+B(t)R^{-1}(t)B^*(t)p+d(t),\\ \dfrac{dp}{dt}=Q(t)x-A^*(t)p\end{cases}\tag{9.10}$$

的解；并使向量(9.6)在$(t_1,\ \phi(t_1))$处与 $\mathscr{T}_1$ 正交；最优控制由(9.9)给出．

我们现在专门考虑取 $\mathscr{T}_1$ 为超平面 $t_1=T$ 的问题；亦即

$$\mathscr{T}_1=\{(t_1, x_1): t_1=T, x_1 \text{ 自由}\}, \tag{9.11}$$

并且取 g 为

$$g(x_1)=\frac{1}{2}\langle x_1, Gx_1\rangle, \tag{9.12}$$

其中 G 是半正定对称矩阵.

附注 9.1 如果 (9.11) 成立, 则 $\mathscr{T}_1$ 的每个切向量的第一个支量等于 0, 另一方面, 轨线 ϕ 的切向量的第一个分量总不为 0. 此外, 从(9.1)和最优控制 u 的连续性可见轨线 ϕ 在所有点上都有切向量, 因此若(9.11)成立, 则假设 9.2 自然满足.

推论 9.1 如果(9.11)和(9.12)成立, 则 ϕ 和 λ 满足带边界条件

$$\phi(t_0)=x_0, \lambda(t_1)=-G\phi(t_1) \tag{9.13}$$

的系统(9.10). 这里第一个条件是初始条件的重述, 它已经有了. 从(9.6)与 $\mathscr{T}_1$ 在轨线终止点 $(t_1, \phi(t_1))$ 处的正交性, 以及(9.12)可得出第二个条件.

满足定理 9.1 条件的容许对 (ϕ, u) 叫做极值对. 如果(9.11)和(9.12)成立, 则极值对满足(9.13).

在下一个定理中, 我们证明如果(9.11)和(9.12)成立, 则极值对是唯一的且是最优的. 这将不必引用前面所建立的任何存在定理.

定理 9.2 若 (9.11) 和 (9.12) 成立, 设 (ϕ, u) 是极值对, (ϕ_1, u_1) 是任何其他容许对, 则 $J(\phi_1, u_1)\geqslant J(\phi, u)$. 等号成立当且仅当 $u=u_1$, 这时还有 $\phi=\phi_1$.

证明: 首先注意到, 因为系统(9.1)是线性的, 并且 (t_0, x_0) 是固定的, 如果 $u=u_1$, 则 $\phi=\phi_1$. 令

$$\phi_f=\phi(T), \phi_{1f}=\phi_1(T).$$

因为对所有的 t, $Q(t)$ 是半正定的, $R(t)$ 是正定的. 又因为 G 是半正定的, 我们得到

$$0\leqslant\langle(\phi_{1f}-\phi_f), G(\phi_{1f}-\phi_f)\rangle+\int_{t_0}^{T}\{\langle(\phi_1-\phi), Q(\phi_1-\phi)\rangle$$
$$+\langle(u_1-u), R(u_1-u)\rangle\}dt,$$

当且仅当 $u_1=u$ 时等号成立. 因此

$$0\leqslant 2J(\phi_1,\ u_1)+2J(\phi,\ u)-2\langle\phi_{1f},\ G\phi_f\rangle$$
$$-2\int_{t_0}^{T}\{\langle\phi_1,\ Q\phi\rangle+\langle u_1,\ Ru\rangle\}dt,$$

我们把它改写成

$$J(\phi_1,\ u_1)+J(\phi,\ u)$$
$$\geqslant\langle\phi_{1f},\ G\phi_f\rangle+\int_{t_0}^{T}\{\langle\phi_1,\ Q\phi\rangle+\langle u_1,\ Ru\rangle\}dt. \tag{9.14}$$

因为 $(\phi,\ u)$ 是极值对, 存在绝对连续向量函数 $\lambda(t)$, 使得 λ 和 ϕ 是(9.10)带边界条件(9.13)的解, 并使(9.9)式成立. 我们现在对(9.14)右边积分号里的 $Q\phi$ 用(9.10)的第二个方程代替, 而(9.14)右边积分号里的 u 用(9.9)式代替, 即可得到

$$J(\phi_1, u_1)+J(\phi,\ u)$$
$$\geqslant\langle\phi_{1f},\ G\phi_f\rangle+\int_{t_0}^{T}\{\langle\phi_1,\ \lambda'+A^*\lambda\rangle+\langle u_1,\ B^*\lambda\rangle\}dt. \tag{9.15}$$

(9.15)右边的积分可以写成

$$\int_{t_0}^{T}\{\langle\phi_1,\ \lambda'\rangle+\langle A\phi_1+Bu_1,\ \lambda\rangle\}dt.$$

因为 $(\phi_1,\ u_1)$ 是容许对, 从(9.1)我们有

$$A\phi_1+Bu_1=\phi_1'-d.$$

将此关系式代入上面的积分, 便给出

$$\int_{t_0}^{T}\{\langle\phi_1,\ \lambda'\rangle+\langle\phi_1',\ \lambda\rangle-\langle d,\ \lambda\rangle\}dt.$$

因此, 我们可以把(9.15)改写成

$$J(\phi_1,\ u_1)+J(\phi,\ u)\geqslant\langle\phi_{1f},\ G\phi_{1f}\rangle+\langle\phi_1,\ \lambda(t_1)\rangle$$
$$-\langle x_0,\ \lambda(t_0)\rangle-\int_{t_0}^{T}\langle d,\ \lambda\rangle dt.$$

如果现在利用(9.13), 即可得到

$$J(\phi_1, u_1)+J(\phi,\ u)\geqslant-\langle x_0,\ \lambda(t_0)\rangle-\int_{t_0}^{T}\langle d,\ \lambda\rangle dt.$$

回顾等号成立当且仅当 $u_1=u$, 而这时 $\phi_1=\phi$. 因此若取 $u_1=u$, 上面的不等式变成

$$2J(\phi, u) = -\int_{t_0}^{T} \langle d, \lambda \rangle dt - \langle x_0, \lambda(t_0) \rangle. \tag{9.16}$$

把(9.16)代入前面的不等式中，即得

$$J(\phi_1, u_1) \geqslant J(\phi, u).$$

当且仅当 $u_1 = u$ 等号成立.

习题 9.1 考虑 t_0, t_1 固定，$\mathscr{T}_1$ 为 n 维线性流形，$g: x_1 \to g(x_1)$ 是凸函数的线性二次问题. 假设控制约束集 $\mathscr{O}$ 是凸的，不借助于定理 9.1 的最大值原理，证明如果(ϕ, u)是最优对，则(ϕ, u)是唯一的.

本节提出的线性二次准则问题，其中 $\mathscr{T}_1$ 如(9.11)，g 如(9.12)，能使我们得到一个很精致而又相对简单的最优控制的综合函数，这种综合函数的确定将在本节剩下的部分讨论.

对具有固定的初始点(τ, ξ) $(a \leqslant \tau < T, \xi \in R^n)$的问题，从习题 III.6.5 可见存在最优对$(\phi(\cdot, \tau, \xi), u(\cdot, \tau, \xi))$. 由定理 9.1，此最优对是极值的，并且对 $\tau \leqslant t \leqslant T$,

$$u(t, \tau, \xi) = R^{-1}(t) B^*(t) \lambda(t, \tau, \xi).$$

然后由定理 9.2 可得带固定初始点(τ, ξ)问题的最优对是唯一的. 因此，如第二节所述，我们得到一个最优轨线场 $\mathscr{F}$. 我们得到最优控制的综合函数或最优反馈控制 U 如下：

$$U(\tau, \xi) = u(\tau, \tau, \xi) = R^{-1}(\tau) B^*(\tau) \lambda(\tau, \tau, \xi). \tag{9.17}$$

上式对所有 $a \leqslant \tau < T$ 和所有 $\xi \in R^n$ 成立. 因为我们可以选取任何这样的点(τ, ξ)作为所论问题的初始点.

(9.17)的反馈律是不能令人满意的，因为它需要知道伴随变量，亦即乘子向量 λ 在初始点的值. 如果第二节的表达式是正确的，则由(2.18)我们有 $W_x(\tau, \xi) = -\lambda(\tau, \tau, \xi)$，从而(9.17)可写成

$$U(\tau, \xi) = -R^{-1}(\tau) B^*(\tau) W_x(\tau, \xi). \tag{9.18}$$

这就导致我们就现在的问题来研究值函数 W. 象第二节一样，我们将继续形式地假设所有函数具有所需要的阶数的导数存在，并且是连续的. 用此方法我们将首先得到作为反馈控制结构的一些

洞察和猜测．然后，用其他方法严格地证明这些推测是正确的。

今后我们假设在(9.1)中 $d=0$.

函数 W 满足 Hamilton-Jacobi 方程(2.14)，就现在的情形它变成

$$W_t=-\frac{1}{2}\langle x,\ Qx\rangle-\frac{1}{2}\langle U,\ RU\rangle-\langle W_x,\ Ax\rangle-\langle W_x,\ BU\rangle.$$

在此等式及下面各等式中我们将略去函数所包含的自变量．利用(9.18)式，我们可以改写此方程如下：

$$W_t=-\frac{1}{2}\langle x,\ Qx\rangle-\frac{1}{2}\langle R^{-1}B^*W_x,\ B^*W_x\rangle-\langle W_x,\ Ax\rangle+\langle W_x,\ BR^{-1}B^*W_x\rangle,$$

因此

$$W_t=-\frac{1}{2}\langle x,\ Qx\rangle-\frac{1}{2}\langle W_x,\ BR^{-1}B^*W_x\rangle-\langle W_x,\ Ax\rangle. \tag{9.19}$$

从方程(9.19)的形式，启发人们猜测 Hamilton-Jacobi 方程(2.14)存在一个形如

$$W(t,\ x)=\frac{1}{2}\langle x,\ P(t)x\rangle \tag{9.20}$$

的解，其中对每一个 t, $P(t)$是对称矩阵．于是

$$W_x=P(t)x,\quad W_t=\frac{1}{2}\langle x,\ P'(t)x\rangle. \tag{9.21}$$

并对适当选择的 $P(t)$，我们可以使(9.19)式左边的二次型等于右边的二次型.

若假定有形如(9.20)的解，把(9.21)代入(9.19)，并注意到 $P^*=P$，我们得到

$$\frac{1}{2}\langle x,\ P'x\rangle=-\frac{1}{2}\langle x,\ Qx\rangle+\frac{1}{2}\langle x,\ PBR^{-1}B^*Px\rangle-\langle x,\ PAx\rangle. \tag{9.22}$$

因为对任何矩阵 M 可以表示成

$$M=(M+M^*)/2+(M-M^*)/2.$$

因此，对所有的 x

$$\langle x,\ Mx\rangle=\frac{1}{2}\langle x,\ (M+M^*)x\rangle.$$

如果将此结论应用于(9.22)中的矩阵 PA，即得

$$\frac{1}{2}\langle x,\ P'x\rangle=-\frac{1}{2}\langle x,\ Qx\rangle+\frac{1}{2}\langle x,\ PBR^{-1}B^*Px\rangle$$
$$-\frac{1}{2}\langle x,\ (PA+A^*P)x\rangle.$$

因此，如果 Hamilton-Jacobi 方程存在形如(9.20)的解，矩阵 $P(t)$ 必须满足如下微分方程

$$P'=-Q+PBR^{-1}B^*P-(PA+A^*P). \tag{9.23}$$

此外，因为

$$W(T,\ x_1)=g(x_1)=\frac{1}{2}\langle x_1,\ Gx_1\rangle,$$

从(9.20)可见(9.23)的解必须满足条件

$$P(T)=G. \tag{9.24}$$

方程(9.23)有时叫做矩阵 Riccati 方程. 如果(9.23)满足(9.24)的解存在，则从(9.21)的第一个关系式和(9.18)，我们所期望的最优综合函数或最优反馈控制律将由

$$U(t,\ x)=-R^{-1}(t)B^*(t)P(t)x \tag{9.25}$$

给出. 注意，控制律关于 x 是线性的，而它的确定仅需要解 Riccati 微分方程(9.23).

我们现在证明，上面最后几段分析的情况确实是对的.

定理 9.3 若 $\mathscr{T}_1$ 如(9.11)，g 如(9.12)，设(9.1)中的 $d=0$，约束集合 $\mathscr{O}$ 包含原点. 则满足状态方程(9.1)，控制约束 Ω 和端点条件 $\mathscr{B}$，而使泛函(9.3)取极小值的问题有最优综合函数，这里问题的依据是满足假设 9.1. 此综合函数是由(9.25)给出的，并且对所有 $a\leqslant t\leqslant T$ 和所有 R^n 中的 x 成立. 矩阵 $P(t)$ 对每一个 t 是对称的，且矩阵函数 $P(t)$ 是满足条件(9.24)的 Riccati 矩阵微分方程(9.23)的解，它在整个区间 $[a,\ T]$ 上有定义.

证明：从常微分方程解的标准存在和唯一性定理可见，方程(9.23)在某区间 $(T-\delta,\ T]$ 上存在满足(9.24)的唯一解. 注意，

如果 P 是(9.23)之满足(9.24)的解，则 P^* 亦然，于是由解的唯一性即知 $P=P^*$，所以 P 是对称的。

设 τ 是 $(T-\delta,\ T)$ 上的任意点，ξ 是 E^n 中的任意点．我们将使用前一段所得到的解 $P(t)$ 来构造以 $(\tau,\ \xi)$ 为初始点的问题的极值对．由定理 9.2，此极值对将是唯一的；并将实现对带初始点 $(\tau,\ \xi)$ 的问题的极小化．

考虑满足初始条件 $x(\tau)=\xi$ 的线性系统

$$\frac{dx}{dt}=A(t)x-B(t)R^{-1}(t)B^*(t)P(t)x. \tag{9.26}$$

我们用 $\phi(\cdot,\ \tau,\ \xi)$ 表示这个系统的解．此解在 P 的定义区间上有定义且唯一．

令

$$\lambda(t,\ \tau,\ \xi)=-P(t)\phi(t,\ \tau,\ \xi),\ \tau\leqslant t\leqslant T. \tag{9.27}$$

注意，

$$\lambda(T,\ \tau,\ \xi)=-P(T)\phi(T,\ \tau,\ \xi)=-G\phi(T,\ \tau,\ \xi),$$

其中后一个等式从(9.24)而得．因此 λ 满足(9.13)的第二式．

如果我们微分(9.27)式，然后利用(9.23)和(9.26)，即得

$$\begin{aligned}\frac{d\lambda}{dt}&=-\frac{dP}{dt}\phi-P\frac{d\phi}{dt}\\&=(Q-PBR^{-1}B^*P+PA+A^*P)\phi-P(A-BR^{-1}B^*P)\phi\\&=Q\phi+A^*P\phi.\end{aligned}$$

再利用(9.27)就得出

$$\frac{d\lambda}{dt}=Q\phi-A^*\lambda.$$

于是按照定理 9.1 和推论 9.1，$\phi(\cdot,\ \tau,\ \xi)$ 和 $\lambda(\cdot,\ \tau,\ \xi)$ 可确定一个极值对 $(\phi(\cdot,\ \tau,\ \xi),\ u(\cdot,\ \tau,\ \xi))$．其中

$$u(t,\ \tau,\ \xi)=R^{-1}(t)B^*(t)\lambda(t,\ \tau,\ \xi).$$

从而，由定理 9.2 可见此极值对是所论问题的唯一的最优对．从(9.27)式所定义的 λ，以及上面的方程可得

$$u(t,\ \tau,\ \xi)=-R^{-1}(t)B^*(t)P(t)\phi(t,\ \tau,\ \xi).$$

因此，由 $\phi(\tau,\ \tau,\ \xi)=\xi$，就得

$$u(\tau, \tau, \xi) = -R^{-1}(\tau)B^*(\tau)P(\tau)\xi.$$

因为(τ, ξ)是$(T-\delta, T)\times E^n$中的任意点，并且由(τ, ξ)确定的最优对是唯一的，令

$$U(\tau, \xi) = u(\tau, \tau, \xi),$$

我们就得到了最优控制的综合．所以在$(T-\delta, T)\times E^n$中的最优综合函数可以写成

$$U(t, x) = -R^{-1}(t)B^*(t)P(t)x.$$

其中，我们用一般点(t, x)代替(τ, ξ)．然而这正好是关系式(9.25)．

最后，我们证明(9.23)满足初始条件(9.24)的解$P(t)$是定义在全区间$[a, T]$上的，从而得到(9.25)式对所有$a\leqslant t\leqslant T$和E^n中的所有x都成立．

假设$\delta>0$，它使得$(T-\delta, T]$是方程(9.23)的解P的以T为右端点的最大定义区间．则从常微分方程理论中的标准存在性定理[注]和方程(9.23)的形式可以得知，从点$T-\delta$的右边当$t\to T-\delta$时$P(t)$必为无界．下面我们将证明若$T-\delta\geqslant a$，则当$t\to T-\delta$时$P(t)$却是有界的，由此自然得知$P(t)$在整个$[a, T]$上有定义，从而(9.25)对所有$a\leqslant t\leqslant T$和R^n中的所有x成立．

从线性二次问题的存在性定理(习题III.6.5)和定理9.2可知对所有$[a, T]\times E^n$中的(τ, ξ)，函数$W(\tau, \xi)=J(\phi(\cdot, \tau, \xi), u(\cdot, \tau, \xi))$有定义，其中$(\phi(\cdot, \tau, \xi), u(\cdot, \tau, \xi))$是带初始点$(\tau, \xi)$的线性二次问题的唯一最优对．如此定义的函数$W$叫做值函数或$J$的值．令$\tilde{\phi}(\cdot, \tau, \xi)$表示所论问题相应于控制$\tilde{u}$的轨线，其中在$[\tau, T]$上$\tilde{u}(t)=0$，则

$$0\leqslant W(\tau, \xi)\leqslant J(\tilde{\phi}(\cdot, \tau, \xi), \tilde{u}), \tag{9.28}$$

其中最左边的不等式由假设9.1的条件和(9.12)式而来．从(9.1)中令$d=0$，我们得知

$$\tilde{\phi}(t, \tau, \xi) = \Phi(t, \tau)\xi.$$

[注] 此段论证中所援用的实际上是常微分方程论中通常的所谓延拓定理的结论．——译者注．

其中$\Phi(\cdot,\tau)$是系统$\frac{dx}{dt}=A(t)x$满足$\Phi(\tau,\tau)=I$的基本解矩阵，所以，从(9.3)和(9.12)得到

$$J(\tilde{\phi}(\cdot,\tau,\xi),\tilde{u})=\langle\xi\Phi^*(T,\tau),G\Phi(T,\tau)\xi\rangle+\int_\tau^T\langle\xi\Phi^*(s,\tau),Q(s)\Phi(s,\tau)\xi\rangle ds.$$

由此可见对给定在E^n中的致密集$\mathscr{X}$，存在一个依赖于$\mathscr{X}$的常数M，使得对所有$a\leqslant\tau\leqslant T$和$\mathscr{X}$中所有的$\xi$

$$J(\tilde{\phi}(\cdot,\tau,\xi),\tilde{u})\leqslant M.$$

由此不等式和(9.28)表明：对给定在E^n中的致密集$\mathscr{X}$，存在一个依赖于$\mathscr{X}$的常数M，使得对所有$a\leqslant\tau\leqslant T$和所有$\xi\in\mathscr{X}$，有

$$0\leqslant W(\tau,\xi)\leqslant M. \tag{9.29}$$

在前面证明定理9.2的过程中，我们曾导出关于$J(\phi(\cdot,\tau,\xi),u(\cdot,\tau,\xi))$的一个表达式(9.16)．因为现在$d=0$，故由(9.16)可得

$$W(\tau,\xi)=-\frac{1}{2}\langle\xi,\lambda(\tau,\tau,\xi)\rangle. \tag{9.30}$$

其中λ是具有初始点(τ,ξ)的问题的伴随函数或乘子函数．这里没有假设λ是由(9.27)给出的．

现在考虑点(τ,ξ)，其中τ满足不等式$T-\delta<\tau\leqslant T$．对这样的点(9.27)成立．于是

$$\lambda(\tau,\tau,\xi)=-P(\tau)\phi(\tau,\tau,\xi)=-P(\tau)\xi.$$

把此式代入(9.30)即得

$$W(\tau,\xi)=\frac{1}{2}\langle\xi,P(\tau)\xi\rangle. \tag{9.31}$$

它对$T-\delta<\tau\leqslant T$和所有ξ成立．由此关系式和(9.29)，我们得知对所有$T-\delta\leqslant\tau\leqslant T$和致密集$\mathscr{X}$中所有的$\xi$，必满足

$$0\leqslant\frac{1}{2}\langle\xi,P(\tau)\xi\rangle\leqslant M.$$

因此$P(\tau)$在$T-\delta\leqslant\tau\leqslant T$上必是有界的，定理证毕．

在证明定理9.3的过程中，我们也证明了如下推论．

推论 9.3 对所有$a\leqslant\tau\leqslant T$和$E^n$中的$\xi$，值函数$W$是由(9.31)给定的．

第六章　最大值原理的证明

1. 引　　言

本章致力于最大值原理，即定理V.3.1的证明．我们实际上将证明一个比最大值原理更一般的定理，而最大值原理就成为这个定理的一个特殊情形．最优轨线的基本性质促使我们引入所谓$\mathscr{F}$-$\mathscr{N}$极值曲线的概念．接着叙述$\mathscr{F}$-$\mathscr{N}$极值曲线的一个必要条件(本章定理3.1)，并说明如何由它导出定理V.3.1.

为了证明下面的定理3.1，我们按如下步骤进行：首先在控制所对应的轨线的端点上建立扰动的作用．然后考虑扰动的某些凸集，并证明一阶扰动轨线的端点生成欧氏空间中一个锥体．接着证明，作为最优性假设的结果，这个锥体被某一半直线所分离．这个分离性的解析推论就构成本章的定理3.1，由此可得出定理V.3.1.

本章最后一节，我们将证明推论V.3.1和V.3.2，这些推论给出最大值原理的加强形式，它们对一些重要情形是可以应用进去的．

2. $\mathscr{F}$-$\mathscr{N}$极值曲线

在定理V.3.1中假设$g\equiv 0$．证明这个定理采用Mayer表达形式是更方便的，这两种表达形式的等价性已在II.4中证明过了．

令$\hat{f}=(f^0, f)=(f^0, f^1, \cdots, f^n)$；$\hat{x}=(x^0, x)=(x^0, x^1, \cdots, x^n)$以及令$\hat{\phi}=(\phi^0, \phi)=(\phi^0, \phi^1, \cdots, \phi^n)$，现在考虑如下问题：

问题2.1　设

$$J(\hat{\phi}, u)=\phi^0(t_1). \tag{2.1}$$

在满足方程组

$$\frac{d\hat{x}}{dt}=\hat{f}(t,\ x,\ u(t)),\tag{2.2}$$

控制约束 $u(t)\in\Omega(t)$ 和端点条件

$$\{(t_0,\ \phi(t_0),\ t_1,\ \phi(t_1))\in\mathscr{B},\ \phi^0(t_0)=0,\ \phi^0(t_1)\text{自由}\}\tag{2.3}$$

的情况下，使 J 取极小值.

注意(2.2)的右端是 $n+1$ 维向量且 $\hat{f}$ 只与 $x=(x^1,\ \cdots,\ x^n)$ 有关而与 x^0 无关. 还要注意 $(\hat{\phi},\ u)$ 关于问题 2.1 是容许对当且仅当 $(\phi,\ u)$ 关于原来的 Lagrange 问题是容许对.

设记号和 V.3 中的一样. 对每个控制 u 存在一个定义在区间 $[t_0,\ t_1]\subset\mathscr{I}_0$ 上的轨线 ϕ，使得 $(\phi,\ u)$ 是容许的，我们得到定义在 $[t_0,\ t_1]\times\mathscr{X}_0$ 上的函数 $\hat{F}_u$ 如下:

$$\hat{F}_u(t,\ x)=\hat{f}(t,\ x,\ u(t)).\tag{2.4}$$

由假设 V.3.1-(ii)得到，对 $\mathscr{X}_0$ 内每个 x，函数 $t\to\hat{F}_u(t,\ x)$ 在 $[t_0,\ t_1]$ 上可测；而对 $[t_0,\ t_1]$ 中每个 t，函数 $x\to\hat{F}_u(t,\ x)$ 在 $\mathscr{X}_0$ 上是 $C^{(1)}$ 类的. 由假设 V.3.1-(iii)我们得到对每个致密集 $\mathscr{X}\subset\mathscr{X}_0$ 和每个控制 u，在 $L_1[t_0,\ t_1]$ 中存在非负函数 $\mu=\mu(\cdot,\ \mathscr{X},\ u)$，使得对 $[t_0,\ t_1]$ 中几乎所有的 t，满足

$$|\hat{F}_u(t,\ x)|\leqslant\mu(t),\quad\left|\frac{\partial\hat{F}_u}{\partial x}(t,\ x)\right|\leqslant\mu(t).\tag{2.5}$$

如果对 $t\notin[t_0,\ t_1]$，令 $\hat{F}_u(t,\ x)=0$，则 $\hat{F}_u$ 在 $\mathscr{I}_0\times\mathscr{X}_0$ 上有定义，且上一段的叙述用 $\mathscr{I}_0$ 代替 $[t_0,\ t_1]$ 时仍成立. 我们下面就假定是这种情形.

令 $\mathscr{F}'$ 表示上段所得函数 $\hat{F}_u$ 的集合. 集 $\mathscr{F}'$ 是让 u 跑遍所有容许控制时所得到的"(2.2)的右边"的集合. 于是，问题 2.1 可改述如下:

问题 2.2 在满足

$$\frac{d\hat{x}}{dt}=\hat{F}_u(t,\ x),\quad\hat{F}_u\in\mathscr{F}'\tag{2.6}$$

和端点条件(2.3)的条件下，使(2.1)取极小值.

由常微分方程理论的标准存在唯一性定理及(2.5)的第二个不等式推得，一旦 $\hat{F}_u$ 取定且 $(t_0, \phi(t_0))$ 指定后，则轨线 ϕ 就唯一确定。因此问题 2.2 中有由我们选取的函数 $\hat{F}_u$ 和点 $(t_0, \phi(t_0))$。记住 $\phi^0(t_0)=0$ 是已经指定了的。函数 $\hat{F}_u$ 和点 $(t_0, \phi(t_0))$ 要选取得使(2.3)成立且 $\phi^0(t_1)$ 达到极小值。

设 (ϕ^*, u^*) 是问题 2.1 的一个最优对，则相应于 u^* 我们得到函数 $\hat{F}^*=\hat{F}_{u^*}$，它关于问题 2.2 是最优的。

现在我们介绍最优对 (ϕ^*, u^*) 的性质，它在定理 V.3.1 的证明中将要用到。

令

$$\mathcal{N}'=\{(t_0, \hat{x}_0, t_1, \hat{x}_1):(t_0, x_0, t_1, x_1)\in\mathscr{B},\\ x_0^0=0,\ x_1^0\leqslant\phi^{*0}(t_1)\},\tag{2.7}$$

$$\mathcal{M}'=\{(t_0, \hat{x}_0, t_1, \hat{x}_1):(t_0, x_0, t_1, x_1)\in\mathscr{B},\\ x_0^0=0,\ x_1^0=\phi^{*0}(t_1)\},\tag{2.8}$$

$$\mathscr{E}'=\{e(\hat{\phi}):\hat{\phi}\in\mathscr{A}_T\},$$

这里 $\mathscr{A}_T$ 表示问题 2.1 的容许轨线集合。

如果我们设 $\mathcal{N}_1'$ 由这样的点 $(t_0, \hat{x}_0, t_1, \hat{x}_1)$ 组成：$(t_0, x_0, t_1, x_1)\in\mathscr{B}$，$x_0^0=0$ 且对某个 $\varepsilon>0$，$\phi^{*0}(t_1)-\varepsilon<x_1^0<\phi^{*0}(t_1)+\varepsilon$，则由于 $\mathscr{B}$ 是 q 维 $C^{(1)}$ 流形，下面的结论是正确的。集合 $\mathcal{N}_1'$ 是 $q+1$ 维 $C^{(1)}$ 流形。集合 $\mathcal{M}'$ 是包含在 $\mathcal{N}_1'$ 中的 q 维 $C^{(1)}$ 流形，且把 $\mathcal{N}_1'$ 分为两个集合，其中一个是 $\mathcal{N}'$，另一个是 $\mathcal{N}_1'-\mathcal{N}'$。注意 $\mathcal{M}'$ 包含在 $\mathcal{N}'$ 的边界中，而 $\mathcal{N}'$ 可以认为是位于 $\mathcal{M}'$ 的一边的 $\mathcal{N}_1'$ 的那一部分。

如果存在一个流形 $\mathcal{N}_1$ 使得 $\mathcal{N}_1$、$\mathcal{N}$ 和 $\mathcal{M}$ 的关系与 $\mathcal{N}_1'$、$\mathcal{N}'$ 和 $\mathcal{M}'$ 的关系一样，我们称 $\mathcal{N}$ 是具有上部边界 $\mathcal{M}$ 的一个流形。我们把正式定义的表达形式留给读者。

$(\hat{\phi}^*, u^*)$ 是问题 2.1 的最优对的命题，意味着在

$$e(\hat{\phi}^*)=(t_0, \hat{\phi}^*(t_0), t_1, \hat{\phi}^*(t_1))$$

的邻域内没有端点在此邻域内的其它容许轨线，使其零上标分量比 $\phi^{*0}(t_1)$ 小。实际上，这个性质表示局部极小值。因此，若 $(\hat{\phi}^*, u^*)$

是最优对，则 $e(\hat{\phi})$ 有下列性质：

(i) $e(\hat{\phi}^*)\in\mathscr{M}'$；

(ii) 存在 $e(\hat{\phi}^*)$ 的一个邻域 $\mathscr{O}'$ 使得 $(\mathscr{O}'\cap\mathscr{E}'\cap\mathscr{N}')\subset\mathscr{M}'$.

图 2.1 说明当初始点 $\hat{x}=(x^0, x^1)$ 固定，而终端集合是一维流形 $\mathscr{T}$ 的情形.

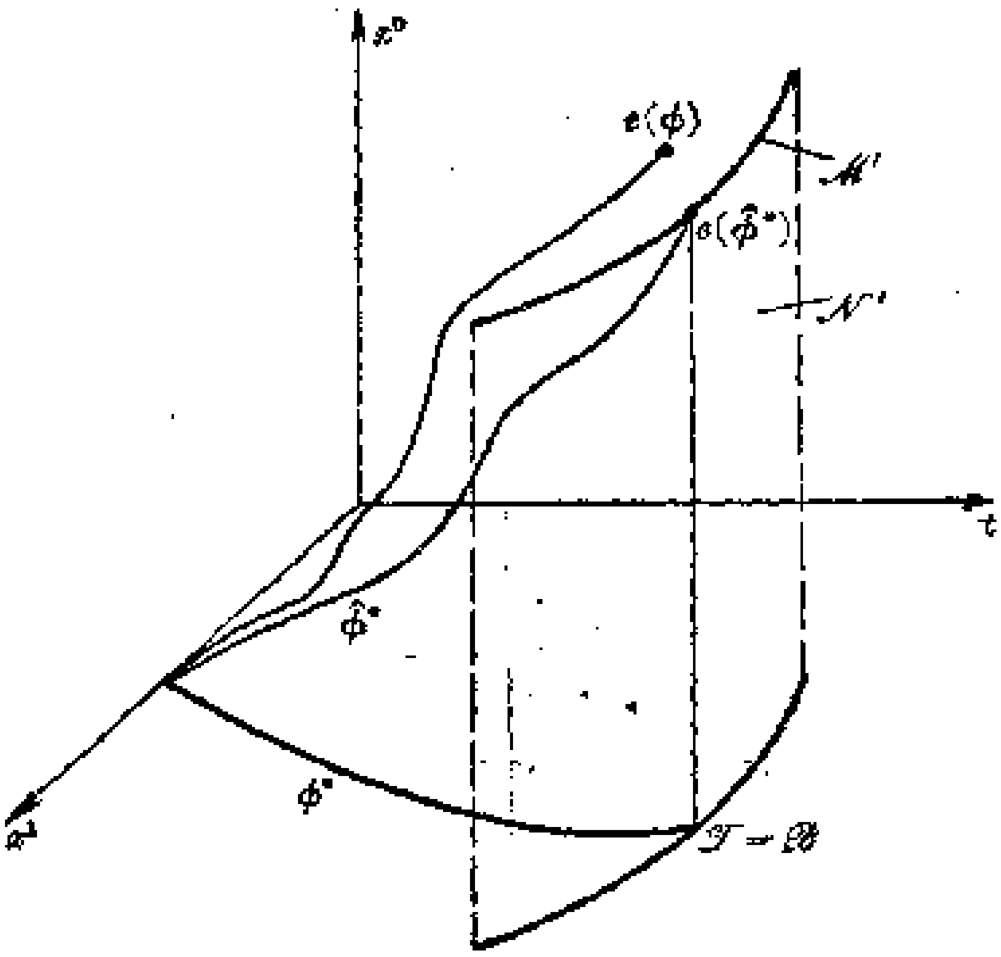

图　1

上面的讨论导出如下的考虑. 设 $\mathscr{Y}_0$ 是 E^{n+1} 中的一个开区间，而 E^{n+1} 中的点记为 $y=(y^0, y^1, \cdots, y^n)$. 又设 $\mathscr{F}$ 是定义在 $\mathscr{I}_0\times\mathscr{Y}_0$ 上取值于 E^{n+1} 中的函数族. 我们用 F 表示 $\mathscr{F}$ 中的元素并作如下假设：

假设 2.1 (i) $\mathscr{F}$ 中的每个 F 对 $\mathscr{Y}_0$ 中每个固定的 y 在 $\mathscr{I}_0$ 上可测，而对 $\mathscr{I}_0$ 中的每个固定的 t 在 $\mathscr{Y}_0$ 上是 $C^{(1)}$ 类的.

(ii) 对 $\mathscr{F}$ 中每个 F 和致密子集 $\mathscr{Y}\in\mathscr{Y}_0$，在 $L_1[\mathscr{I}_0]$ 中存在一个非负函数 $\mu=\mu(\cdot, \mathscr{Y}, F)$，使得

$$|F(t, y)|\leqslant\mu(t),\quad |F_y(t, y)|\leqslant\mu(t)\tag{2.9}$$

在 $\mathscr{I}_0$ 中几乎处处成立.

注 2.1 $\mathscr{F}'$ 中的函数 $\hat{F}$ 是定义在 $\mathscr{I}_0\times\mathscr{X}_0$ 上的. 因此，如果我们令 $\mathscr{Y}_0=E'\times\mathscr{X}_0$ 并取

$$y=(y^0,\ x),\ y^0\in E',\ x\in\mathscr{X}_0, \tag{2.10}$$

则由(2.5)式及其前面的叙述中所指出的函数 $\hat{F}_u$ 的连续性和可测性可知族 $\mathscr{F}'$ 满足假设 2.1.

考虑微分方程组的集合

$$y'=F(t,\ y),\ F\in\mathscr{F},\ y(t_0)=y_0. \tag{2.11}$$

设 ψ 是(2.11)定义在区间 $[t_0,\ t_1]\subset\mathscr{I}_0$ 上的解；$\mathscr{S}$ 表示(2.11)的解及其定义区间的集合，于是

$$\mathscr{S}=\{[\psi,\ t_0,\ t_1]:\psi'(t)=F(t,\ \psi(t)),\ F\in\mathscr{F},\ t\in[t_0,\ t_1]\}.$$

令
$$e(\psi)=(t_0,\ \psi(t_0),\ t_1,\ \psi(t_1));$$
$$\mathscr{E}=\{e(\psi):\psi\in\mathscr{S}\}.$$

设 $\mathscr{N}$ 是一个其边界为子集 $\mathscr{M}$ 的可微流形，这里 $\mathscr{M}$ 也是一个可微流形．集合 $\mathscr{N}$ 和 $\mathscr{M}$ 均在 $E^{2(n+1)+2}$ 中．

定义 2.1 对 $\mathscr{F}$ 中某个函数 $\bar{F}$，(2.11) 的解 $\bar{\psi}$ 称为 $\mathscr{F}$-$\mathscr{N}$ 极值曲线，如果：

(i) $e(\bar{\psi})\in\mathscr{M}$；

(ii) 存在 $e(\bar{\psi})$ 的一个邻域 $\mathscr{O}$ 使得 $(\mathscr{O}\cap\mathscr{E}\cap\mathscr{N})\subset\mathscr{M}$.

注 2.2 如果 $(\hat{\phi}^*,\ u^*)$ 是问题 2.1 的最优对，则 $\hat{\phi}^*$ 对问题 2.2 是最优的，且 $\hat{\phi}^*$ 是 $\mathscr{F}'$-$\mathscr{N}'$ 极值曲线．我们将导出 $\mathscr{F}$-$\mathscr{N}$ 极值曲线所要满足的必要条件，这些条件就将给出最优对 $(\hat{\phi}^*,\ u^*)$ 所要求的必要条件．

今对 $\mathscr{F}$ 作进一步假定．设

$$P^r=\left\{\alpha:\alpha\in E^r,\ \alpha^i\geqslant 0,\ \sum_{i=1}^r\alpha^i=1\right\}. \tag{2.12}$$

令 $\quad \mathrm{co}\mathscr{F}=\{h:h=\sum_{i=1}^r\alpha^iF_i,\ F_i\in\mathscr{F},\ \alpha\in P^r$, 对某个整数 $r\}$.

集合 $\mathrm{co}\mathscr{F}$ 称为 $\mathscr{F}$ 的凸壳．

定义 2.2 函数 F 的集合 $\mathscr{F}$ 称为拟凸的，如果 $\mathscr{F}$ 满足假设 2.1，且有下列性质：

设已给 $\varepsilon>0$，致密集 $\mathscr{Y}\subset\mathscr{Y}_0$，$\mathscr{F}$ 中的有限个元素组 F_1, …, F_r 及 P^r 中的 α，则存在定义在 $\mathscr{I}_0\times\mathscr{Y}$ 上，依赖于 α, ε, $\mathscr{Y}$ 和

$F_1, \cdots, F_r$ 的函数 $g(\cdot, \cdot, \alpha, \varepsilon)$，具有下列性质：

(i) 函数 g 对 $\mathscr{Y}$ 中固定的 y 关于 t 是可测的，对 $\mathscr{I}_0$ 中固定的 t 在 $\mathscr{Y}$ 上是 $C^{(1)}$ 类的；

(ii) 由

$$F_\alpha(t, y)=\sum_{i=1}^{r}\alpha^i F_i(t, y)-g(t, y, \alpha, \varepsilon)$$

定义的函数 F_α 在 $\mathscr{F}$ 中；

(iii) 在 $L_1[\mathscr{I}_0]$ 中存在依赖于 $\mathscr{Y}$ 和 $F_1, \cdots, F_r$，但不依赖于 α 或 ε 的函数 μ，使得

$$|g(t, y, \alpha, \varepsilon)|\leqslant\mu(t), \quad |g_y(t, y, \alpha, \varepsilon)|\leqslant\mu(t) \qquad (2.13)$$

对 $\mathscr{Y}$ 中所有的 y、所有的 $\alpha\in P^r$ 及 $\mathscr{I}_0$ 中几乎所有的 t 成立；

(iv) 如果 $\mathscr{Z}$ 是定义在 $\mathscr{I}_0$ 上取值于 $\mathscr{Y}$ 的等度连续函数族，则对 $\mathscr{Z}$ 中的每个 z 及 $\mathscr{I}_0$ 中的每个 t 和 t'

$$\left|\int_t^{t'} g(s, z(s), \alpha, \varepsilon)ds\right|\leqslant\varepsilon^2; \qquad (2.14)$$

(v) 如果 $\{\alpha_n\}$ 是 P^r 中的点列，使得 $\alpha_n\to\alpha$，则对 $\mathscr{Y}$ 中的每个 y，

$$g(\cdot, y, \alpha_n, \varepsilon)\longrightarrow g(\cdot, y, \alpha, \varepsilon)$$

在 $\mathscr{I}_0$ 上依测度成立.

注意，由于 P^r 是闭的，故在(v)中所述的点 α 在 P^r 中.

引理 2.1 族 $\mathscr{F}'$ 是拟凸的.

由于 $\mathscr{F}'$ 的元素是由(2.4)定义的函数 $\hat{F}_u$，因此，$\mathscr{F}'$ 中的有限个元素组可以写成由

$$\hat{f}_i(t, x)=\hat{f}(t, x, u_i(t)), \quad i=1, 2, \cdots, r$$

所定义的函数 $\hat{f}_1, \cdots, \hat{f}_r$ 的集合，其中 $\hat{f}$ 是(2.2)的右端函数，而 $u_1, \cdots, u_r$ 是问题 2.1 的控制.

在注 2.1 中我们指出过，$\mathscr{F}'$ 中的函数 $\hat{F}_u$ 满足假设 2.1. 存在函数 g 满足定义 2.2 的(i)～(iv)是定理 IV.4.2 的推论. 在那个定理中把 n 换成 $n+1$，函数 f_i 取成现在的 $\hat{f}_i$ 以及取

$$P^i(t)=\alpha, \quad q=r,$$

则所要求的函数 g 由

$$g(t,\ y,\ \alpha,\ \varepsilon)=\lambda(t,\ x)$$

给出，其中 λ 由 IV.(4.14)确定，而 y 与(2.10)中的一样.

为了证明(v)成立，我们指出：由于

$$g(t,\ y,\ \alpha_n,\ \varepsilon)=\sum_{i=1}^{r}\alpha_n^i\hat{f}_i(t,\ x)-\hat{f}_{\alpha_n}(t,\ x)$$

而
$$g(t,\ y,\ \alpha,\ \varepsilon)=\sum_{i=1}^{r}\alpha^i\hat{f}_i(t,\ x)-f_\alpha(t,\ x),$$

故只须证明

$$\hat{f}_{\alpha_n}(\cdot,\ x)\longrightarrow f_\alpha(\cdot,\ x) \tag{2.15}$$

在 $\mathscr{I}_0$ 上依测度成立就够了.

为看出这一点，我们首先回顾定理 IV.4.1 中函数 $\hat{f}_\alpha$ 和 $\hat{f}_{\alpha_n}$ 的定义. 先把区间 $\mathscr{I}_0$ 写成适当的有限个子区间的并集：

$$\mathscr{I}_0=\bigcup_{j=1}^{k}I_j,$$

其中 $I_j=[t_j,\ t_{j+1}]$，$\cdots<t_{j-1}<t_j<t_{j+1}<\cdots$，而 k 依赖于 ε. 然后为了定义 $\hat{f}_\alpha$，把每个 I_j 分割成有限个不交迭的子区间，从左到右记为 E_{j1}，$\cdots$，E_{jr}，而 E_{ji} 的长度是 $\alpha^i|I_j|$. 对于 E_{ji} 内部的 t，令 $\hat{f}_\alpha(t,\ x)=\hat{f}_i(t,\ x)$，$i=1$，$\cdots$，$r$. 在端点上 $\hat{f}_\alpha$ 可用任意方式定义. 类似地，为定义 $\hat{f}_{\alpha_n}$，把每个 I_j(对所有 α_n 和 α，I_j 是相同的)分割成有限个不交迭的子区间，从左到右记为 E_{j1}^n，E_{j2}^n，$\cdots$，E_{jr}^n. E_{ji}^n 的长度是 $\alpha_n^i|I_j|$. 对于 E_{ji}^n 内部的 t，令 $\hat{f}_{\alpha_n}(t,\ x)=f_i(t,\ x)$，$i=1$，2，$\cdots$，$r$；在端点上 $\hat{f}_{\alpha_n}$ 以任意方式定义.

由于只存在有限个区间 I_j，只要证明在每个 I_j 上(2.15)依测度成立就够了. 为简化记号，我们考虑 I_1 并设 I_1 的长度为 1，于是$E_{1i}=[\alpha^{i-1},\alpha^i]$和 $E_{1i}^n=[\alpha_n^{i-1},\alpha_n^i]$，其中 $\alpha_n^0=\alpha^0$ 是 I_1 的左端点. 由于 $\alpha_n\to\alpha$，故知对每个 $\eta>0$ 存在正整数 n_0，使得对 $n>n_0$ 及所有的 $j=1$，$\cdots$，r，点 α_n^{j-1} 或者属于$[\alpha^{j-1},\ \alpha^j]$，或者属于$[\alpha^{j-2},\ \alpha^{j-1}]$；点 α_n^j 或者属于$[\alpha^{j-1},\ \alpha^j]$，或者属于$[\alpha^j,\ \alpha^{j+1}]$；而且

$$\sum_{i=1}^{r}|\alpha^i-\alpha_n^i|<\eta.$$

这里具有负数或零上标的 α 表示 I_1 的左端点. 现在由 $\hat{f}_{\alpha_n}$ 和 $\hat{f}_\alpha$

的定义得到除测度小于 η 的集合中的 t 之外，$f_{\alpha_n}(t, x)=f_\alpha(t, x)$. 于是(2.15)在 I_1 上依测度收敛.

3. $\mathscr{F}$-$\mathscr{N}$ 极值曲线的必要条件

这一节的定理 3.1，给出 $\mathscr{F}$-$\mathscr{N}$ 极值曲线必须满足的必要条件. 最大值原理即定理 V.3.1，是作为这个定理的推论而得到的. 定理 3.1 的证明将在本章 4～7 节中给出.

对 $\mathscr{F}$ 中的每个函数 $F=(F^0, F^1, \cdots, F^n)$，我们定义函数 H_F 如下：

$$H_F(t, y, l)=\sum_{i=0}^{n} l^i F^i(t, y)=\langle l, F(t, y)\rangle. \tag{3.1}$$

注意，l 是 $(n+1)$ 维向量，每个函数 H_F 是实值的，关于 l 是线性的，关于 y 是 $C^{(1)}$ 类，而关于 t 是可测的.

定理 3.1 设 $\mathscr{F}$ 是拟凸集. 令 $\bar\psi$ 为定义在 $[t_0, t_1]$ 上的 $\mathscr{F}$-$\mathscr{N}$ 极值曲线，而且

$$\bar\psi'(t)=\bar F(t, \bar\psi(t)) \tag{3.2}$$

几乎处处成立；令

$$\bar H(t, y, l)=H_{\bar F}(t, y, l).$$

则存在定义在 $[t_0, t_1]$ 上的绝对连续函数 $\eta=(\eta^0, \eta^1, \cdots, \eta^n)$，使得

(i) $\eta(t)$ 在 $[t_0, t_1]$ 上恒不为零；

(ii) 对 $[t_0, t_1]$ 中几乎所有的 t，成立

$$\begin{cases}\bar\psi'(t)=\bar H_l(t, \bar\psi(t), \eta(t))=\bar F(t, \bar\psi(t)), & (3.3)\\ \eta'(t)=-\bar H_y(t, \bar\psi(t), \eta(t)) & \\ \qquad=-\langle\eta(t), \bar F_y(t, \bar\psi(t))\rangle; & (3.4)\end{cases}$$

(iii) 对 $\mathscr{F}$ 中所有的 F，不等式

$$\int_{t_0}^{t_1}\bar H(s, \bar\psi(s), \eta(s))ds\geqslant\int_{t_0}^{t_1}H_F(s, \bar\psi(s), \eta(s))ds \tag{3.5}$$

成立. 而且，如果映射 $t\to\bar F(t, \bar\psi(t))$ 在端点 t_0 和 t_1 是连续的，则 $2(n+1)+2$ 维向量

$$(\bar H(t_0, \bar\psi(t_0), \eta(t_0)), -\eta(t_0), \\ -\bar H(t_1, \bar\psi(t_1), \eta(t_1)), \eta(t_1)) \tag{3.6}$$

在 $e(\bar{\psi})$ 上与 $\mathscr{M}$ 正交.

现由这个定理来推导最大值原理(即定理 V.3.1).

在注 2.2 中我们已指出问题 2.1 的解 $(\hat{\phi}^*, u^*)$ 有这样的性质,即 $\hat{\phi}^*$ 是 $\mathscr{F}'$-$\mathscr{N}'$ 极值曲线. 因此,用 $\mathscr{F}'$ 代替 $\mathscr{F}$, $\hat{\phi}^*$ 代替 $\bar{\psi}$, 而 y 如(2.10)中所示,对定理 3.1 还是适用的. $\mathscr{F}'$ 中的函数 $\hat{F}$ 由(2.4)给出,而 F 代之以问题 2.2 的最优函数 $\hat{F}^*=\hat{F}^*_u$, 函数 H_F 为 $H_{\hat{F}^*}$ 所代替. 从(3.1)和(2.4)得知 $H_{\hat{F}^*}$ 由公式

$$H_{\hat{F}^*}(t, y, l)=\sum_{i=0}^{n} l^i\hat{F}^i_u(t, x)=\sum_{i=0}^{n} l^i f^i(t, x, u(t)) \tag{3.7}$$

给出, 其中 $\hat{f}=(f^0, f)=(f^0, f^1, \cdots, f^n)$ 是(2.2)的右端函数,而 u 是容许控制. 类似地,可得到

$$H_{\hat{F}^*}(t, y, l)=\sum_{i=0}^{n} l^i f^i(t, x, u^*(t)). \tag{3.8}$$

从(3.8)可看出 $H_{\hat{F}^*}$ 与 y^0 无关. 因此, 向量方程(3.4)的第一个分量是

$$\frac{d\eta^0}{dt}=0.$$

于是 η^0 是常数,把它记为 λ^0. 令

$$\lambda=(\eta^1, \cdots, \eta^n) \quad \text{和} \quad \hat{\lambda}=(\lambda^0, \lambda)=(\lambda^0, \lambda^1, \cdots, \lambda^n).$$

如果注意到 $\mathscr{F}'$ 中的 $\hat{F}_u$ 与 y_0 无关,则(3.4)后几个分量给出

$$\lambda'(t)=-\lambda^0 f^0_x(t, \phi^*(t), u^*(t)) -\langle\lambda(t), f_x(t, \phi^*(t), u^*(t))\rangle.$$

(3.3)的后 n 个分量给出

$$\phi^{*\prime}(t)=f(t, \phi^*(t), u^*(t)).$$

上面的后两个方程正好是定理 V.3.1 的方程 V.(3.2), 其中最优对已记为 (ϕ^*, u^*), 它代替了 (ϕ, u).

由(3.5)、(3.7)和(3.8)我们得到定理 V.3.1 的(3.3).

根据(3.6)在 $e(\bar{\psi})$ 上与 $\mathscr{M}=\mathscr{M}'$ 的正交性以及(2.8)可得 V(3.4)在 $e(\phi^*)$ 上与 $\mathscr{B}$ 正交.

设 ν 是 $E^{2(n+1)+2}$ 中形如

$$(\tau-t_0, \hat{\xi}_0-\hat{\phi}^*(t_0), \tau-t_1, \hat{\xi}_1-\hat{\phi}^*(t_1))$$

的任意向量，其中

$$\xi_1^0-\phi^{*0}(t_1)<0, \tag{3.9}$$

而且在过 $e(\hat{\phi}^*(t_1))$ 的 $\mathscr{N}$ 的切平面内．在证明定理 3.1 的过程中，将说明(3.6)与任何这样的 ν 的内积总是 $\geqslant 0$．一个这样的 ν 是除 $\xi_1^0-\phi^{*0}(t_1)$ 以外，全部分量为零的向量，这个特殊的 ν 和(3.6)的内积是

$$\eta^0(t_1)(\xi_1^0-\phi^{*0}(t_1))\geqslant 0.$$

由此式与(3.9)就可推得 $\eta^0(t_1)\leqslant 0$，但是 $\eta^0(t_1)=\lambda^0$，所以 $\lambda^0\leqslant 0$，从而定理 3.1 的所有论断都成立．

4．极值轨线的扰动

这一节，我们首先叙述将要用到的微分方程组

$$\frac{dy}{dt}=G(t,\ y,\ \beta),\ \ y(\tau)=\eta \tag{4.1}$$

的解的一些结果，其中 $t\in\mathscr{I}_0$，$y\in\mathscr{Y}_0$，而 β 在欧氏空间某开区域 B 内．我们将以 $\theta=\theta(\cdot,\ \tau,\ \eta,\ \beta)$ 表示(4.1)的解．由于初始条件 $y(\tau)=\eta$，所以有 $\eta=\theta(\tau,\ \tau,\ \eta,\ \beta)$．

虽然下面引理 4.1 和引理 4.2 的内容是熟知的，但在任何一本单独的标准参考书中，都不可能全部按我们所给的方式叙述．为使读者方便起见，我们将指定一本标准参考书当作唯一的材料来源，并指明如何得到引理中那些在所给的参考书中找不到的论断．

引理 4.1 对 $\mathscr{Y}_0\times B$ 中的每个 $(y,\ \beta)$，设 $G(\cdot,\ y,\ \beta)$ 在 $\mathscr{I}_0$ 上可测，且对 $\mathscr{I}_0$ 中的每个 t 和 $\beta\in B$，$G(t,\ \cdot,\ \beta)$ 在 $\mathscr{Y}_0$ 上是 $C^{(1)}$ 类的．设存在致密区间 $\mathscr{I}\subset\mathscr{I}_0$ 和存在包含在 $\mathscr{Y}_0$ 中的致密区间 $\mathscr{Y}$ 以及 $L_1[\mathscr{I}]$ 中的函数 μ，使得对所有 $\mathscr{I}\times\mathscr{Y}\times B$ 中的 (t,y,β)，满足

$$|G(t,\ y,\ \beta)|\leqslant\mu(t),\quad |G_y(t,\ y,\ \beta)|\leqslant\mu(t), \tag{4.2}$$

则对 $\mathscr{I}\times\mathscr{Y}\times B$ 内部的每个点 $(\tau,\ \eta,\ \beta)$，存在(4.1)的定义在最大区间 $(\omega_-(\tau,\ \eta,\ \beta),\ \omega_+(\tau,\ \eta,\ \beta))$ 上的唯一解 $\theta(\cdot,\ \tau,\ \eta,\ \beta)$．此外，单边极限

$$\lim_{\tau\to\omega_-}\theta(t,\ \tau,\ \eta,\ \beta)=\theta_-$$
$$\lim_{\tau\to\omega_+}\theta(t,\ \tau,\ \eta,\ \beta)=\theta_+ \tag{4.3}$$

存在，且点$(\omega_-,\ \theta_-)$、$(\omega_+,\ \theta_+)$为$\mathscr{I}\times\mathscr{Y}$的边界点.

对B中每个固定的β，(4.1)的解在某区间$[\tau-\varepsilon(\eta,\ \beta),\ \tau+\varepsilon(\eta,\ \beta)]$上(其中$\varepsilon(\eta,\ \beta)>0$)的存在性由[38]的定理68.4得到. 现证$\theta$在其有定义的任何区间上是唯一的. 设$\psi$是(4.1)的另一个解，并把$\theta(t,\ \tau,\ \eta,\ \beta)$简记为$\theta(t)$，则由关系式$\psi(\tau)=\theta(\tau)=\eta$、中值定理和(4.2)可以得知，对在$\theta$和$\psi$共同有定义的区间中的任何$t$，有

$$|\theta(t)-\psi(t)|\leqslant\int_\tau^t|G(s,\ \theta(s),\ \beta)-G(s,\ \psi(s),\ \beta)|ds$$
$$\leqslant\int_\tau^t\mu(s)|\theta(s)-\psi(s)|ds.$$

由 Gronwall 引理(引理 IV.4.2)即得欲证的唯一性.

证明了存在包含τ在其内部的闭区间，使得在此闭区间上存在(4.1)的唯一解以后，由延拓定理便可得出存在一个包含在$\mathscr{I}$中的最大开区间$(\omega_-(\tau,\ \eta,\ \beta),\ \omega_+(\tau,\ \eta,\ \beta))$，使得在其上(4.1)存在唯一解. 今证单边极限$\theta_-$和$\theta_+$存在. 设$t_1$、$t_2$是任意满足$\tau<t_1<t_2<\omega_+(\tau,\ \eta,\ \beta)$的两点，则由(4.2)有

$$|\theta(t_2)-\theta(t_1)|\leqslant\int_{t_1}^{t_2}|G(s,\ \theta(s),\ \beta)|ds\leqslant\int_{t_1}^{t_2}\mu(s)ds.$$

如果令$t_1\to\omega_+$，$t_2\to\omega_+$，则从μ的可积性和 Cauchy 准则，得知θ_+存在. 类似地论证可得θ_-存在.

如果点$(\omega_+(\tau,\ \eta,\ \beta),\ \theta_+)$是$\mathscr{I}\times\mathscr{Y}$的内点，则由上述的存在唯一性结论，我们得知问题

$$\frac{dy}{dt}=G(t,\ y,\ \beta),\quad y(\omega_+)=\theta_+$$

存在唯一的解，它在含ω_+为其内部的某区间上有定义；而在ω_+的左边这个解应与θ一致. 这样一来，就与$(\omega_-(\tau,\ \eta,\ \beta),\ \omega_+(\tau,\ \eta,\ \beta))$是最大存在区间的假设矛盾，因此，$(\omega_+(\tau,\ \eta,\ \beta),\ \theta_+)$是

$\mathscr{I}\times\mathscr{Y}$ 的边界点. 类似地论证, 可得$(\omega_-(\tau, \eta, \beta), \theta_-)$为 $\mathscr{I}\times\mathscr{Y}$ 的边界点.

引理 4.2 设G如引理 4.1 所述, 只是对 $\mathscr{I}_0$ 中每个 t, 我们设 $G(t, \cdot, \cdot)$在 $\mathscr{Y}_0\times B$ 上是 $C^{(1)}$ 类. 设$(\tau, \bar{\eta}, \bar{\beta})$为 $\mathscr{I}\times\mathscr{Y}\times B$ 内部的一点, t'、t'' 为 $\mathscr{I}$ 中两点, 使得 $\omega_-(\tau, \bar{\eta}, \bar{\beta})<t'<t''<\omega_+(\tau, \bar{\eta}, \bar{\beta})$, 则下述论断是正确的:

(i) 存在正的 ε_1 和 δ 使得对 $\mathscr{Y}\times B$ 中满足

$$|\eta-\bar{\eta}|<\delta,\ |\beta-\bar{\beta}|<\delta \tag{4.4}$$

的每一对(η, β), 存在(4.1)的定义在区间

$$t'-\varepsilon_1\leqslant t\leqslant t''+\varepsilon_1 \tag{4.5}$$

上的唯一解 $\theta(\cdot, \tau, \eta, \beta)$;

(ii) 函数 $\theta(t, \tau, \cdot, \cdot,)$关于 η 和 β 是可微的, 而且偏导数 θ_η 和 θ_β 在由(4.4)和(4.5)所确定的集合上是连续的;

(iii) 对每一对满足(4.4)的(η, β), 矩阵函数 $\theta_\eta(\cdot, \tau, \eta, \beta)$ 是方程组

$$\nu'=G_y(t, \theta(t, \tau, \eta, \beta), \beta)\nu$$

在区间(4.5)上的基本解矩阵, 且 $\theta_\eta(\cdot, \tau, \eta, \beta)$满足初始条件

$$\theta_\eta(\tau, \tau, \eta, \beta)=I,$$

其中 I 是$(n+1)\times(n+1)$单位矩阵.

引理 4.2 的论断(i)和(ii)是从[38]的定理 69.4 的(c_1)和(c_2)两部分得来的; 论断 (iii) 是在[38]定理 69.4 的证明过程中建立的; 参看[38]的方程(P)及第 362 页上的有关原文.

推论 4.1 设

$$B=\left\{\beta:\beta=(\varepsilon\alpha,\ \varepsilon)=\varepsilon(\alpha^1, \cdots, \alpha^r, 1),\ \alpha\in P^r,\ -\frac{1}{2}<\varepsilon<\frac{1}{2}\right\}, \tag{4.6}$$

其中 P^r 如(2.12)所示; 又设

$$\theta_\varepsilon=\theta_\beta\frac{\partial\beta}{\partial\varepsilon},\ G_\varepsilon=G_\beta\frac{\partial\beta}{\partial\varepsilon}, \tag{4.7}$$

则对满足(4.4)的每对(η, β), 函数 $\theta_\varepsilon(\cdot, \tau, \eta, \beta)$ 是问题

$$\begin{aligned}\nu' = G_y(t,\ \theta(t,\ \tau,\ \eta,\ \beta),\ \beta)\nu \\ +G_\varepsilon(t,\ \theta(t,\ \tau,\ \eta,\ \beta),\ \beta), \\ \nu(\tau)=0\end{aligned} \tag{4.8}$$

在区间(4.5)上的解.

证明：考虑系统

$$\begin{aligned}&\frac{dy}{dt}=G(t,\ y,\ (\varepsilon\alpha,\ \varepsilon)),\ y(\tau)=\eta;\\ &\frac{d\varepsilon}{dt}=0, \qquad\qquad \varepsilon(\tau)=\varepsilon;\end{aligned} \tag{4.9}$$

其中 α 是 P^r 中的一个固定元素. 由于我们考虑的是一个固定的 α, 系统(4.9)不包含参数. 如果把 β 看做固定的，经过明显的修改，引理 4.2 仍成立. 因此，在引理 4.2 的叙述中删去参数 β, 就可应用于系统(4.9). 由引理 4.2 可得，对所有满足(4.4)的 η 和满足 $|\varepsilon|\leqslant\varepsilon_2$ 的 $\varepsilon(\varepsilon_2>0)$, $\theta(\cdot,\ \tau,\ \eta,\ \beta)$ 和 $\varepsilon(\cdot,\ \tau,\ \eta,\ \beta)=\varepsilon$ (其中 $\beta=(\alpha\varepsilon,\ \varepsilon)$)是(4.9)在区间 $t'-\varepsilon_1\leqslant t\leqslant t''+\varepsilon_1$ 上的唯一解. 由引理 4.2 的(iii)得到对于 $t'-\varepsilon_1\leqslant t\leqslant t''+\varepsilon_1$,

$$\begin{pmatrix}\theta_\eta & \theta_\varepsilon\\ \varepsilon_\eta & \varepsilon_\varepsilon\end{pmatrix}'=\begin{pmatrix}G_y & G_\varepsilon\\ 0 & 0\end{pmatrix}\begin{pmatrix}\theta_\eta & \theta_\varepsilon\\ \varepsilon_\eta & \varepsilon_\varepsilon\end{pmatrix},$$

其中在矩阵方程中的撇“′”表示关于 t 的导数，函数 θ_η、θ_ε、ε_η、ε_ε 及其导数在$(t,\ \tau,\ \eta,\ \beta)$处取值，而函数 G_y 和 G_ε 在$(t,\ \theta(t,\ \tau,\ \eta,\ \beta),\ \beta)$处取值. 这里 θ_η 和 G_y 是$(n+1)\times(n+1)$矩阵，θ_ε 和 G_ε 是$(n+1)$维列向量，ε_η 是$(n+1)$维行向量，ε_ε 是纯量. 由关系式 $\varepsilon(t,\ \tau,\ \eta,\ \beta)=\varepsilon$, 可得 $\varepsilon_\varepsilon=1$. 由此和上述矩阵方程推得，$\theta_\varepsilon(\cdot,\ \tau,\ \eta,\ \beta)$满足微分方程(4.8). 从引理 4.2(iii)推出的关系式

$$\begin{pmatrix}\theta_\eta(\tau,\ \tau,\ \eta,\ \beta) & \theta_\varepsilon(\tau,\ \tau,\ \eta,\ \beta)\\ \varepsilon_\eta(\tau,\ \tau,\ \eta,\ \beta) & \varepsilon_\varepsilon(\tau,\ \tau,\ \eta,\ \beta)\end{pmatrix}=\begin{pmatrix}I & 0\\ 0 & 1\end{pmatrix},$$

可见，$\theta_\varepsilon(\tau,\ \tau,\ \eta,\ \beta)=0$. 于是，(4.8)中的所有论断都成立.

注 4.1 设 B 如(4.6)所定义，δt_0 和 δt_1 为纯量，而 δ_w 为$(n+1)$维向量，使得$(\delta t_0,\ \delta t_1,\ \delta_w)$在 E^{n+3} 中一个固定致密集 K 内. 又设 τ 和 η 固定，且 $\bar\beta=0$, 则 $\bar\beta$ 相应于 $\varepsilon=0$. 设$[t_0,\ t_1]$是包

含在开区间$(\omega_-(\tau, \bar{\eta}, 0), \omega_+(\tau, \bar{\eta}, 0))$内部的一个闭区间. 记住$(\omega_-(\tau, \bar{\eta}, 0), \omega_+(\tau, \bar{\eta}, 0))$是(4.1)的解$\theta(\cdot, \tau, \bar{\eta}, 0)$的最大定义区间. 设$\eta=\bar{\eta}+\varepsilon\delta w$, 则由引理4.2推得存在一个$\varepsilon_0>0$, 使得函数

$$\theta(\cdot, \tau, \bar{\eta}+\varepsilon\delta w, \beta), \beta=(\varepsilon\alpha, \varepsilon)$$

是(4.1)在区间$[t_0+\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$上的解, 此一事实对$K$中所有$(\delta t_0, \delta t_1, \delta w)$、$P^r$中所有$\alpha$及所有$0\leqslant\varepsilon\leqslant\varepsilon_0$成立. 由于(4.8)关于$\nu$是线性的系统, 故得对$P^r$中所有$\alpha$、$K$中所有$(\delta t_0, \delta t_1, \delta w)$及所有$0\leqslant\varepsilon\leqslant\varepsilon_0$, θ_ε为(4.8)在$[t_0+\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$上的解.

设$\bar{\psi}$为定义在含于$\mathscr{I}_0$内部的区间$[t_0, t_1]$上的$\mathscr{F}$-$\mathscr{N}$极值曲线, 若τ是(t_0, t_1)内的点, 则函数$\bar{\psi}$是微分方程

$$y'=\bar{F}(t, y), y(\tau)=\bar{\psi}(\tau), \tag{4.10}$$

在$[t_0, t_1]$上的唯一解, 其中$\bar{F}\in\mathscr{F}$. $\mathscr{I}$是包含在$\mathscr{I}_0$中的致密区间而$\mathscr{Y}$是含于$\mathscr{Y}_0$中的致密区间, 使得$\bar{\psi}$的图象位于$\mathscr{I}\times\mathscr{Y}$内部.

我们希望用$co\mathscr{F}-\bar{F}$中的元素δF给(4.10)右边施加扰动, 遗憾的是, 这是不可能的, 因为由此得到的元素$\bar{F}+\varepsilon\delta F$不一定属于$\mathscr{F}$. 但由于$\mathscr{F}$是拟凸的, 我们能够对$\bar{F}+\varepsilon\delta F$加上一个"小"的元素$g$使得$\bar{F}+\varepsilon\delta F+g$属于$\mathscr{F}$之中. 我们现在来确切地表达它.

设$F_1, \cdots, F_r$是$\mathscr{F}$中r个确定的函数, 而α是P^r中的元素. 定义

$$\delta F=\sum_{i=1}^{r}\alpha^i F_i-\bar{F}.$$

注意δF依赖于$F_1, \cdots, F_r$和α. 我们将不在记号中标明这个依赖性, 但是读者记住这个依赖性是重要的. 设$0\leqslant\varepsilon\leqslant1$, 则

$$\bar{F}+\varepsilon\delta F=(1-\varepsilon)\bar{F}+\varepsilon\sum_{i=1}^{r}\alpha^i F_i. \tag{4.11}$$

(4.11)的右端是$\mathscr{F}$中$(r+1)$个函数的凸组合. 由于$\mathscr{F}$是拟凸的, 故存在一个定义在$\mathscr{I}_0\times\mathscr{Y}_0$上、依赖于$\alpha$、$\varepsilon$、$\bar{F}$、$F_1, \cdots, F_r$及

$\mathscr{Y}$ 的函数 $g=g(\cdot, \cdot, \alpha, \varepsilon)$，具有定义 2.2 所列的性质，且使

$$\bar{F}+\varepsilon\delta F+g\in\mathscr{F}.$$

由假设 2.1 中的(2.9)得知，我们可以假设定义 2.2 (iii)中的函数 μ 使得(2.13)成立，且对所有 $i=1, \cdots, r$ 和 $\mathscr{I}\times\mathscr{Y}$ 中的 (t, y)，有

$$\begin{aligned} &|\bar{F}(t, y)|\leqslant\mu(t), \quad |F_i(t, y)|\leqslant\mu(t), \\ &|\bar{F}_y(t, y)|\leqslant\mu(t), \quad |F_{iy}(t, y)|\leqslant\mu(t). \end{aligned} \tag{4.12}$$

我们将考虑(4.10)的扰动族

$$\frac{dy}{dt}=\bar{F}(t, y)+\varepsilon\delta F(t, y)+g(t, y, \alpha, \varepsilon), \tag{4.13}$$

其中对某个固定的 r，α 在 P^r 中，且 $0\leqslant\varepsilon\leqslant1$. 又考虑 $\bar{\psi}$ 的定义区间的扰动区间 $[t_0+\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$，其中 δt_0 和 δt_1 的变化遍历一个致密集. 最后，我们将考虑初始条件 $y(\tau)=\bar{\psi}(\tau)$ 的扰动 $\bar{\psi}(\tau)+\varepsilon\delta w$，其中 δw 的变化范围为 E^{n+1} 中的一个致密集. 这样的扰动对轨线 $\bar{\psi}$ 的影响，可概括为下面的引理.

引理 4.3 设 $F_1, \cdots, F_r$ 为 r 个给定在 $\mathscr{F}$ 中的函数，δt_0、δt_1 是纯量，δw 为 E^{n+1} 中的向量，α 为 P^r 中的元素，则存在 $\varepsilon_0>0$，使得对每个 $0\leqslant\varepsilon\leqslant\varepsilon_0$，(4.13)存在一个定义在区间 $[t_0+\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$ 上的解 $\psi=\psi(\cdot, \delta w, \alpha, \varepsilon)$，满足初始条件

$$\psi(\tau)=\bar{\psi}(\tau)+\varepsilon\delta w. \tag{4.14}$$

解 ψ 取形式

$$\psi(t, \delta w, \alpha, \varepsilon)=\bar{\psi}(t)+\varepsilon\delta\bar{\psi}(t)+E(t, \delta w, \alpha, \varepsilon), \tag{4.15}$$

其中 $\delta\bar{\psi}$ 为变分方程

$$\delta y'=\bar{F}_y(t, \bar{\psi}(t))\delta y+\delta F(t, \bar{\psi}(t)) \tag{4.16}$$

满足初始条件

$$\delta\bar{\psi}(\tau)=\delta w \tag{4.17}$$

的解；而且当 $\varepsilon\to0$ 时，

$$|E(t, \delta w, \alpha, \varepsilon)|/\varepsilon\to0 \tag{4.18}$$

对 $[t_0+\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$ 中的 t、致密集 K 中的所有 $(\delta t_0, \delta t_1, \delta w)$ 和

P^r 中的所有 α 一致地成立.

注 4.2 请读者记住, δF 依赖于 $F_1, \cdots, F_r$ 和 α, 所以函数 $\delta\bar{\psi}$ 也依赖于这些量.

证明: 由(4.11), 我们可写出

$$G(t, y, \varepsilon\alpha, \varepsilon) = \bar{F}(t, y) + \varepsilon\delta F(t, y), \tag{4.19}$$

因此, 可把(4.13)写为

$$\frac{dy}{dt} = G(t, y, \varepsilon\alpha, \varepsilon) + g(t, y, \alpha, \varepsilon). \tag{4.20}$$

如果令 $\beta = (\varepsilon\alpha, \varepsilon)$ 并设

$$\hat{G}(t, y, \beta) = G(t, y, \varepsilon\alpha, \varepsilon),$$

则可看出函数 $\hat{G}$ 具有引理 4.1 和 4.2 中加在函数 G 上的那些性质, 所以这些引理及其结论适用于系统

$$\frac{dy}{dt} = G(t, y, \varepsilon\alpha, \varepsilon). \tag{4.21}$$

遗憾的是, 我们没有关于 g 作为 α 和 ε 的函数的正规性的信息, 故不能对(4.20)应用这些结果. 因此, 我们分两个步骤处理. 首先, 我们分析(4.21)及其些有关初始数据和端点的扰动. 然后, 将这些扰动解和(4.20)的那些解加以比较, 因为在适当意义下由 g 的"充分小"性, 这是能够做得到的.

考虑(4.21)具有初始条件

$$y(\tau) = \eta = \bar{\eta} + \varepsilon\delta w, \tag{4.22}$$

其中 $\bar{\eta} = \bar{\psi}(\tau)$. 我们把这个初值问题的解表示为

$$\theta = \theta(\cdot, \eta, \beta), \quad \beta = (\varepsilon\alpha, \varepsilon).$$

由于我们的讨论始终假定 τ 保持固定的, 故在记号中就不标明 θ 对 τ 的依赖性. 由(4.19)可见, 对 $\varepsilon = 0$, 系统退化为(4.10), 而(4.10)在区间 $[t_0, t_1]$ 上的唯一解是 $\bar{\psi}$, 于是

$$\bar{\psi}(t) = \theta(t, \bar{\eta}, o).$$

由引理 4.2 和注 4.1 可知存在 $\varepsilon_0 > 0$, 使得对所有 $0 \leqslant \varepsilon \leqslant \varepsilon_0$ 和对 K 中所有 $(\delta t_0, \delta t_1, \delta w)$ 及 P^r 中的所有 α, 带有 $\beta = (\varepsilon\alpha, \alpha)$ 的函数 $\theta(\cdot, \eta, \beta)$ 是(4.21)在 $[t_0 + \varepsilon\delta t_0, t_1 + \varepsilon\delta t_1]$ 上满足初始条

件(4.22)的解，而且函数 θ 关于 η 和 β 是可微的，其导数 θ_η 和 θ_β 在 θ 的定义区域上是 (t, η, β) 的连续函数. 由 θ_β 和 G 的性质及(4.7)、(4.6)可推得，θ_ε 在其定义域上是 (t, η, ε) 的连续函数. 由推论 4.1 得

$$\theta'_\eta(t, \bar{\eta}, o) = G_y(t, \bar{\psi}(t), o)\theta_\eta(t, \bar{\eta}, o); \tag{4.23}$$

$$\theta'_\varepsilon(t, \bar{\eta}, o) = G_y(t, \bar{\psi}(t), o)\theta_\varepsilon(t, \bar{\eta}, o) + G_\varepsilon(t, \bar{\psi}(t), o), \tag{4.24}$$

其中 θ_ε 和 G_ε 如(4.7)所示，而撇"$'$"表示关于 t 的导数. 注意 θ_η 是矩阵而 θ_ε 是向量. 由(4.19)得

$$G_y(t, \bar{\psi}(t), o) = \bar{F}_y(t, \bar{\psi}(t)),$$

$$G_\varepsilon(t, \bar{\psi}(t), o) = \delta F(t, \bar{\psi}(t)).$$

把这些方程的右端代入(4.23)和(4.24)得

$$\theta'_\eta(t, \bar{\eta}, o) = \bar{F}_y(t, \bar{\psi}(t))\theta_\eta(t, \bar{\eta}, o) \tag{4.25}$$

和
$$\theta'_\varepsilon(t, \bar{\eta}, o) = \bar{F}_y(t, \bar{\psi}(t))\theta_\varepsilon(t, \bar{\eta}, o) + \delta F(t, \bar{\psi}(t)). \tag{4.26}$$

矩阵 θ_η 和向量 θ_ε 还满足初始条件

$$\theta_\eta(\tau, \bar{\eta}, o) = I, \quad \theta_\varepsilon(\tau, \bar{\eta}, o) = 0. \tag{4.27}$$

现比较 $\theta(t, \eta, \beta)$ 与 $\bar{\psi}(t) = \theta(t, \bar{\eta}, o)$. 令

$$\begin{aligned} \Pi(\varepsilon_0) = \{(t, \delta t_0, \delta t_1, \delta w, \alpha, \varepsilon) : t \in [t_0 + \varepsilon\delta t_0, t_1 \\ + \varepsilon\delta t_1]; (\delta t_0, \delta t_1, \delta w) \in K; \alpha \in P^r, 0 < \varepsilon \leqslant \varepsilon_0\}. \end{aligned} \tag{4.28}$$

由于 θ_ε 和 θ_η 存在且对它们的所有变元是连续的，以及由于 $\eta = \bar{\eta} + \varepsilon\delta w$，我们有

$$\begin{aligned} \theta(t, \eta, \beta) - \bar{\psi}(t) = \varepsilon[\theta_\eta(t, \bar{\eta}, o) + o(1)]\delta w \\ + \varepsilon[\theta_\varepsilon(t, \bar{\eta}, o) + o(1)], \end{aligned}$$

其中项 $o(1)$ 当 $\varepsilon \to 0$ 时在 $\Pi(\varepsilon_0)$ 上一致地趋于零. 于是我们可写出

$$\begin{aligned} \theta(t, \eta, \beta) - \bar{\psi}(t) = \varepsilon[\theta_\eta(t, \bar{\eta}, o)\delta w + \theta_\varepsilon(t, \bar{\eta}, o)] \\ + E_2(t, \delta w, \alpha, \varepsilon), \end{aligned}$$

其中当 $\varepsilon \to 0$ 时 $|E_2|/\varepsilon \to 0$ 在 $\Pi(\varepsilon_0)$ 上一致地成立.

如果令

$$\delta\bar{\psi}(t)=\theta_\eta(t,\ \bar{\eta},\ o)\delta w+\theta_\varepsilon(t,\ \bar{\eta},\ o),$$

则上述方程可写成

$$\theta(t,\ \eta,\ \boldsymbol{\beta})=\bar{\psi}(t)+\varepsilon\delta\bar{\psi}(t)+E_2(t,\ \delta w,\ \boldsymbol{\alpha},\ \varepsilon). \quad (4.29)$$

从(4.25)、(4.26)及(4.27)可见 $\delta\bar{\psi}$ 为系统(4.16)满足初始条件(4.17)的解，所以，由参数变易公式，

$$\delta\bar{\psi}(t)=\Psi(t,\ \tau)\left[\delta w+\int_\tau^t\Psi^{-1}(s,\ \tau)\delta F(s,\ \bar{\psi}(s))ds\right], \quad (4.30)$$

其中 Ψ 是齐线性系统

$$\frac{dy}{dt}=F_y(t,\ \bar{\psi}(t))y$$

满足初始条件 $\Psi(\tau,\ \tau)=I$ 的基本解矩阵.

回顾我们以 $\mathscr{I}\times\mathscr{Y}$ 表示 $E^1\times E^{n+1}$ 中的致密区间，使得 $\bar{\psi}$ 的图象位于 $\mathscr{I}\times\mathscr{Y}$ 内部. 由(4.29)、E_2 的性质及(4.30)推得，存在 $\varepsilon_0'<\varepsilon_0$ 使得对 $\varepsilon\leqslant\varepsilon_0'$，$\theta$ 的图象对于 $\Pi(\varepsilon_0')$ 中所有的$(t,\ \delta t_0,\ \delta t_1,\ \delta w,\ \alpha,\ \varepsilon)$都位于 $\mathscr{I}\times\mathscr{Y}$ 内部. 现把 ε_0' 重新记作 ε_0，所有以前关于 $\Pi(\varepsilon_0)$ 的论述仍然成立. 关于取 $\varepsilon_0'<\varepsilon_0$ 并把 ε_0' 重新记作 ε_0，使得所有以前关于 $\Pi(\varepsilon_0)$ 或 $\varepsilon\leqslant\varepsilon_0$ 的论述对新的 ε_0 的值成立，这种方法将在证明中多次出现. 今后，我们这样做时，将只说取 $\varepsilon_0'\leqslant\varepsilon_0$，并把 ε_0' 重新记作 ε_0.

现在回到扰动微分方程(4.13)或者它的等价方程(4.20). 考虑(4.20)及初始条件(4.14). 由于(4.20)的右端对固定的 t、α、ε 关于 y 是 $C^{(1)}$ 类的，故存在一个 ε_0' 使得对 $0\leqslant\varepsilon\leqslant\varepsilon_0'$、$P^r$ 中所有 α 及致密集 K' 中所有 δw(K' 是 K 在 δw-空间 E^{n+1} 中的投影). 下面的论断是正确的：微分方程(4.20)满足初始条件(4.14)有唯一解

$$\psi=\psi(\cdot,\ \delta w,\ \alpha,\ \varepsilon), \quad (4.31)$$

它定义在最大区间$(\omega_-(\delta w,\ \alpha,\ \varepsilon),\ \omega_+(\delta w,\ \alpha,\ \varepsilon))$上，使得 ψ 的图象位于 $\mathscr{I}\times\mathscr{Y}$ 内部.

现取 ε_0' 小于 ε_0 并把 ε_0' 重新记作 ε_0. 由(2.13)和(4.12)得到,对包含 $\bar{\psi}$ 的图象在其内部的致密集 $\mathscr{I}\times\mathscr{Y}$ 中的 (t, y), (4.20)的右端函数对 P^r 中的所有 α 及所有 $0<\varepsilon\leqslant\varepsilon_0$, 以 $L_1[\mathscr{I}]$ 中一个可积函数为其界限. 从这个观察结果得到,由(4.31)给出的解 ψ 的集合是 $\mathscr{I}$ 上的等度连续族,这里,在区间 $(\omega_-(\delta w, \alpha, \varepsilon), \omega_+(\delta w, \alpha, \varepsilon))$ 的外部我们是把解 $\psi(\cdot, \delta w, \alpha, \varepsilon)$ 定义为适当常数的方法来延展 $\psi(\cdot, \delta w, \alpha, \varepsilon)$, 使它在整个 $\mathscr{I}$ 上是连续的.

现在比较 $\psi(t, \delta w, \alpha, \varepsilon)$ 和 $\theta(t, \bar{\eta}+\varepsilon\delta w, \beta)$, 其中 $\eta=\bar{\psi}(\tau)$, $\beta=(\varepsilon\alpha, \varepsilon)$. 为了简化记号, 我们将不标明对 $(\delta w, \alpha, \varepsilon)$ 的依赖性,而写成

$$\psi(t)=\psi(t, \delta w, \alpha, \varepsilon),$$
$$\theta(t)=\theta(t, \bar{\eta}+\varepsilon\delta w, \beta), \quad \beta=\varepsilon(\alpha, 1),$$

这里两个函数中的变元 $(\delta w, \alpha, \varepsilon)$ 取相同的值. 由于 ψ 是(4.20)的解, 而 θ 是(4.21)的解, 它们都有相同的初始值(4.22), 故对 $(\omega_-(\delta w, \alpha, \varepsilon), \omega_+(\delta w, \alpha, \varepsilon))$ 中所有的 t, 有

$$|\theta(t)-\psi(t)|\leqslant\int_\tau^t|G(s, \theta(s), \varepsilon\alpha, \varepsilon)-G(s, \psi(s), \varepsilon\alpha, \varepsilon)|ds$$
$$+\left|\int_\tau^t g(s, \psi(s), \alpha, \varepsilon)ds\right|. \tag{4.32}$$

因为解族 ψ 是等度连续族,由(2.14)推得,(4.32)中右端第二项小于 ε^2. 又由(4.19)、(4.12)及 $0<\varepsilon\leqslant\varepsilon_0'$ 这个事实推得, 存在常数 C, 使得

$$|G(s, \theta(s), \varepsilon\alpha, \varepsilon)-G(s, \psi(s), \varepsilon\alpha, \varepsilon)|$$
$$\leqslant C\mu(s)|\theta(s)-\psi(s)|,$$

因此 $\quad |\theta(t)-\psi(t)|\leqslant C\int_\tau^t\mu(s)|\theta(s)-\psi(s)|ds+\varepsilon^2.$

再根据 Gronwall 不等式(引理 IV.4.2)得到,对于 $(\omega_-(\delta w, \alpha, \varepsilon), \omega_+(\delta w, \alpha, \varepsilon))$ 中一切 t, 有

$$|\theta(t)-\psi(t)|\leqslant\varepsilon^2\exp\{C\int_{\mathscr{I}}\mu(s)ds\}. \tag{4.33}$$

我们指出,具有 $\eta=\bar{\psi}(t)+\varepsilon\delta w$ 及 $\beta=(\varepsilon\alpha, \varepsilon)$ 的 $\theta(t, \eta, \beta)$ 是

定义在整个 $\Pi(\varepsilon_0)$ 上，且每个函数 $\theta(\cdot, \eta, \beta)$ 的图象位于 $\mathscr{I}\times\mathscr{Y}$ 内部. 所以由(4.33)得到，对某个 $\varepsilon_0'<\varepsilon_0$，$\psi(t, \delta w, \alpha, \varepsilon)$ 在整个 $\Pi(\varepsilon_0')$ 上有定义，且每个函数 $\psi(\cdot, \delta w, \alpha, \varepsilon)$ 的图象位于 $\mathscr{I}\times\mathscr{Y}$ 内部. 特别要指出的是，$\psi(\cdot, \delta w, \alpha, \varepsilon)$ 对 $0<\varepsilon\leqslant\varepsilon_0'$ 在 $[t_0-\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$ 上有定义. 我们现再次把 ε_0' 记作 ε_0.

由(4.33)还可以得到

$$\psi(t)=\theta(t)+E_1(t, \delta w, \alpha, \varepsilon); \tag{4.34}$$

其中当 $\varepsilon\to 0$ 时 $|E_1(t, \delta w, \alpha, \varepsilon)|/\varepsilon\to 0$ 在 $\Pi(\varepsilon_0)$ 上一致地成立.

由等式

$$\psi(t)-\bar{\psi}(t)=(\psi(t)-\theta(t))+(\theta(t)-\bar{\psi}(t))$$

及(4.29)、(4.34)可得

$$\psi(t)=\bar{\psi}(t)+\varepsilon\delta\bar{\psi}(t)+E_1(t, \delta w, \alpha, \varepsilon)+E_2(t, \delta w, \alpha, \varepsilon).$$

如果在上式中令

$$E(t, \delta w, \alpha, \varepsilon)=E_1(t, \delta w, \alpha, \varepsilon)+E_2(t, \delta w, \alpha, \varepsilon),$$

便得(4.15). 由于 E_1 和 E_2 在 $\Pi(\varepsilon_0)$ 上一致地满足(4.18)，所以对 E 也同样成立. 由于我们已经证明了 $\delta\bar{\psi}$ 是(4.16)满足(4.17)的解，这就证明了引理.

我们列举引理 4.3 的一些推论来结束本节. 设 $\bar{F}$ 及 $\bar{\psi}$ 使映射

$$t\to\bar{F}(t, \bar{\psi}(t)) \tag{4.35}$$

在 $t=t_i(i=0, 1)$ 连续，令

$$\bar{F}_i=\bar{F}(t_i, \bar{\psi}(t_i)), \quad i=0, 1.$$

由(4.35)在 $t=t_i(i=0, 1)$ 的连续性和关系式

$$\bar{\psi}(t_i+\varepsilon\delta t_i)=\bar{\psi}(t_i)+\int_{t_i}^{t_i+\varepsilon\delta t_i}\bar{F}(s, \bar{\psi}(s))ds$$

得到：
$$\bar{\psi}(t_i+\varepsilon\delta t_i)=\bar{\psi}(t_i)+\varepsilon\delta t_i\bar{F}_i+o(\varepsilon).$$

由 $\delta\bar{\psi}$ 的连续性，于是有

$$\delta\bar{\psi}(t_i+\varepsilon\delta t_i)=\delta\bar{\psi}(t_i)+o(1),$$

其中当 $\varepsilon\to 0$ 时 $o(1)\to 0$. 由最后两个关系式及(4.15)得

$$\psi(t_i+\varepsilon\delta t_i)=\bar{\psi}(t_i)+\varepsilon[\delta\bar{\psi}(t_i)+\delta t_i\bar{F}_i]+o(\varepsilon), \tag{4.36}$$

其中 $o(\varepsilon)$ 项依赖于 $(t,\ \delta w,\ \alpha,\ \varepsilon)$ 且在 $\Pi(\varepsilon_0)$ 上一致地是 ε 的高阶无穷小量.

令

$$\delta y_i=\delta\overline{\psi}(t_i)+\delta t_i\overline{F}_i,\quad i=0,\ 1,\tag{4.37}$$

则由(4.30)得

$$\delta y_i=\Psi(t_i,\ \tau)\left[\delta w+\int_\tau^{t_i}\Psi^{-1}(s,\tau)\delta F(s,\overline{\psi}(s))ds\right]+\delta t_i\overline{F}_i,$$
$$i=0,\ 1.\tag{4.38}$$

由(4.36)及(4.37)得到对下文非常重要的关系式

$$e(\psi)-e(\overline{\psi})=\varepsilon(\delta t_0,\ \delta y_0,\ \delta t_1,\ \delta y_1)$$
$$+\rho^*(\delta t_0,\ \delta t_1,\ \delta w,\ \alpha,\ \varepsilon),\tag{4.39}$$

其中当 $\varepsilon\to 0$ 时 $\rho^*/\varepsilon\to 0$ 对 K 中的 $(\delta t_0,\ \delta t_1,\ \delta w)$ 及 P^r 中的 α 一致地成立.

5. 变分凸集

定义 5.1 所谓变分指的是四元组 $(\delta t_0,\ \delta t_1,\ \delta w,\ \delta F)$，其中 δt_0 及 δt_1 是实数，δw 为 E^{n+1} 中的向量，δF 为 $\mathrm{co}(\mathscr{F})-\overline{F}$ 中的元素.

我们用 $\mathscr{V}$ 表示所有变分的集合，按通常的方式定义变分的加法及其与实数的乘法，在这些定义下 $\mathscr{V}$ 是一个凸集. 对 $\mathrm{co}(\mathscr{F})-\overline{F}$ 中的每一有限元素组 $\delta F_1,\ \cdots,\ \delta F_m$，定义 $\mathscr{V}$ 中的凸集 $\mathscr{V}(\delta F_1,\ \cdots,\ \delta F_m)$ 如下:

$$\mathscr{V}(\delta F_1,\ \cdots,\ \delta F_m)=\{(\delta t_0,\ \delta t_1,\ \delta w,\ \delta F):\delta t_0\in E^1,\ \delta t_1\in E^1,$$
$$\delta w\in E^{n+1},\ \delta F\in\mathrm{co}[\delta F_1,\ \cdots,\ \delta F_m]\}.$$

我们用明显的方法使 $\mathscr{V}(\delta F_1,\ \cdots,\ \delta F_m)$ 等同于 $E^1\times E^1\times E^{n+1}\times P^m$，并把 $E^{2+(n+1)+m}$ 的度量拓扑赋予 $\mathscr{V}(\delta F_1,\ \cdots,\ \delta F_m)$.

对每个函数 δF_i，$i=1,\ 2,\ \cdots,\ m$，在 $\mathscr{F}$ 中存在函数 F_{i1}，F_{i2}，$\cdots$，F_{ir_i} 和 P^{r_i} 中的向量 α_i，使得

$$\delta F_i=\sum_{j=1}^{r_i}\alpha_i^jF_{ij}-\overline{F}.$$

$\mathrm{co}[\delta F_1, \cdots, \delta F_m]$中任何元素$\delta F$具有形式

$$\delta F=\sum_{i=1}^{m}\beta^i\delta F_i,\ \beta^i\geqslant 0,\ \sum_{i=1}^{m}\beta^i=1, \tag{5.1}$$

因此可写为

$$\delta F=\sum_{i=1}^{m}\beta^i\sum_{j=1}^{r_i}\alpha_i^j F_{ij}-\bar{F}.$$

如果令$r=r_1+\cdots+r_m$，并把$F_{ij}(i=1,\cdots,m,\ j=1,\cdots,r_i)$重新记作$F_1,\cdots,F_r$，便得到$\mathrm{co}[\delta F_1,\cdots,\delta F_m]$中的每个$\delta F$[注]可写为

$$\delta F=\sum_{i=1}^{r}\alpha^i F_i-\bar{F}, \tag{5.2}$$

其中$\alpha\in P^r$且依赖于β. 注意函数$F_i,\cdots,F_r$对所有$\mathrm{co}[\delta F_1,\cdots,\delta F_m]$中的元素和我们以前对这个符号的用法是一致的. 我们亦注意这个事实: 从P^m到P^r的映射$\beta\to\alpha=\alpha(\beta)$是连续的. 因此，如果$\{\beta_n\}$是$P^m$中收敛于$\beta$的点列，则$\alpha_n=\alpha(\beta_n)$收敛于$P^r$中的点$\alpha=\alpha(\beta)$.

由于族$\mathscr{F}$是拟凸的，故对每个致密区间$\mathscr{I}\subset\mathscr{I}_0$、每个致密区间$\mathscr{Y}\subset\mathscr{Y}_0$，每个$0<\varepsilon<1$及每个$\beta\in P^m$，存在定义于$\mathscr{I}\times\mathscr{Y}$上的函数$g(\cdot,\ \cdot,\ \alpha(\beta),\ \varepsilon)$，使得

$$\bar{F}+\varepsilon\delta F+g\in\mathscr{F},$$

其中δF由(5.1)或与其等价的(5.2)给出. 对P^m中一切β函数g满足带$\alpha=\alpha(\beta)$的定义2.2的条件; 特别，如果$\beta_n\to\beta$，则对$\mathscr{Y}$中所有的y及$0<\varepsilon<1$，

$$g(\cdot,\ y,\ \alpha(\beta_n),\ \varepsilon)\to g(\cdot,\ y,\ \alpha(\beta),\ \varepsilon) \tag{5.3}$$

在$\mathscr{I}$上依测度成立.

设给定$\delta F_1,\cdots,\delta F_m$，而$\Sigma$是$\mathscr{V}(\delta F_1,\cdots,\delta F_m)$的致密子集. 因为$\mathscr{V}(\delta F_1,\cdots,\delta F_m)$的拓扑是欧氏拓扑，所以存在一个$E^{2+(n+1)}$中的致密集$K$及$P^m$的致密子集$B$，使得$(\delta t_0,\ \delta t_1,\ \delta w,\ \delta F)$是$\Sigma$的元素当且仅当$(\delta t_0,\ \delta t_1,\ \delta w)$在$K$中且$\beta\in B$，其中$\beta$如(5.1)中所示. 又注意到，由于$\beta\to\alpha(\beta)$是连续的，集合

$$A=\{\alpha:\alpha\in P^r,\ \alpha=\alpha(\beta),\ \beta\in B\}$$

[注] 原书误为F.——译者注.

是 P^r 的致密子集. 还要指出，属于 Σ 的变分具有这样的性质：δF 可用(5.2)给出，其中 $\alpha\in A$.

由引理 4.3 可见，存在 $\varepsilon_0>0$，使得若 $0<\varepsilon<\varepsilon_0$，则对 Σ 中的每个变分，存在具有引理 4.3 所描述的性质的轨线 $\psi=\psi(\cdot, \delta w, \alpha(\beta), \varepsilon)$. 这使得我们能够对每个 $0<\varepsilon<\varepsilon_0$ 定义从 Σ 到 $E^{2+2(n+1)}$ 的映射 h_ε 如下：

$$h_\varepsilon(\delta t_0, \delta t_1, \delta w, \delta F)=(e(\psi)-e(\overline{\psi}))/\varepsilon. \tag{5.4}$$

由(4.39)得到

$$\begin{aligned}&h_\varepsilon(\delta t_0, \delta t_1, \delta w, \delta F)\\&=(\delta t_0, \delta y_0, \delta t_1, \delta y_1)+\rho_\varepsilon(\delta t_0, \delta t_1, \delta w, \alpha(\beta)),\end{aligned}$$

其中当 $\varepsilon\to 0$ 时，$\rho_\varepsilon\to 0$ 对 Σ 中的$(\delta t_0, \delta t_1, \delta w, \beta)$一致地成立. 由(4.38)我们看出

$$(\delta t_0, \delta y_0, \delta t_1, \delta y_1)=L(\delta t_0, \delta t_1, \delta w, \delta F), \tag{5.5}$$

其中 L 是定义在整个 $\mathscr{V}$ 上的线性映射. 于是我们可写成

$$h_\varepsilon=L+\rho_\varepsilon. \tag{5.6}$$

引理 5.1 对区间 $0<\varepsilon<\varepsilon_0$ 中的每个固定的 ε，映射 h_ε 和 ρ_ε 在 Σ 上是连续的，而且

$$\lim_{\varepsilon\to 0}\rho_\varepsilon(\delta t_0, \delta t_1, \delta w, \alpha(\beta))=0$$

对 Σ 中的$(\delta t_0, \delta t_1, \delta w, \beta)$一致地成立.

引理中最后论断的真实性已在 h_ε 的定义式(5.4)之后立即指出了. 由 L 的定义(参看(4.38)和(5.1))直接可得 L 在 Σ 上连续，所以要证明引理，只须证明 h_ε 在 Σ 上连续.

映射 h_ε 的分量 δt_0 和 δt_1 显然是连续的，由于 $\overline{\psi}$ 固定，故必须证明对固定的 ε 和固定的 $i=0, 1$，当 Σ 中的 $(\delta t_0, \delta t_i, \delta w, \beta)\to(\delta t_0', \delta t_i', \delta w', \beta')$时，

$$\psi(t_i+\varepsilon\delta t_i, \delta w, \alpha(\beta), \varepsilon)\to\psi(t_i+\varepsilon\delta t_i', \delta w', \alpha(\beta'), \varepsilon).$$

为了简化记号，令

$$\begin{aligned}\psi(t)&=\psi(t, \delta w, \alpha(\beta), \varepsilon),\\\tilde{\psi}(t)&=\psi(t, \delta w', \alpha(\beta'), \varepsilon).\end{aligned}$$

则

$$|\psi(t_i+\varepsilon\delta t_i)-\tilde{\psi}(t_i+\varepsilon\delta t_i')|\leqslant|\psi(t_i+\varepsilon\delta t_i)-\tilde{\psi}(t_i+\varepsilon\delta t_i)|$$
$$+\int_{\varepsilon\delta t_i}^{\varepsilon\delta t_i'}|\tilde{\psi}'(s)|ds. \tag{5.7}$$

函数 $\tilde{\psi}$ 是方程(4.13)或其等价方程(4.20)的解，这时 δF 由(5.2)给出，而 $\alpha=\alpha(\beta')$. 于是

$$\tilde{\psi}'(t)=(1-\varepsilon)F(t,\tilde{\psi}(t))+\varepsilon\sum_{i=1}^{r}\alpha^i(\beta')F_i(t,\tilde{\psi}(t))$$
$$+g(t,\tilde{\psi}(t),\alpha(\beta'),\varepsilon), \tag{5.8}$$

类似地，

$$\psi'(t)=(1-\varepsilon)F(t,\psi(t))+\varepsilon\sum_{i=1}^{r}\alpha^i(\beta)F_i(t,\psi(t))$$
$$+g(t,\psi(t),\alpha(\beta),\varepsilon). \tag{5.9}$$

由(5.8)、(4.12)及(2.13)得

$$|\tilde{\psi}'(s)|\leqslant 2\mu(s),$$

其中 μ 不依赖于 β、ε 及 δw. 因此，对每个 $0<\varepsilon<\varepsilon_0$, (5.7)式右端的积分当 $(\delta t_0,\delta t_1,\delta w,\beta)\to(\delta t_0',\delta t_1',\delta w',\beta')$ 时趋于零.

为要完成证明，还必须证明对固定的 $0<\varepsilon<\varepsilon_0$, (5.7)右端的第一项当 $(\delta t_0,\delta t_1,\delta w,\beta)\to(\delta t_0',\delta t_1',\delta w',\beta')$ 时趋于零.

由于 $\psi(\tau)=\bar{\psi}(\tau)+\varepsilon\delta w$, $\tilde{\psi}(\tau)=\bar{\psi}(\tau)+\varepsilon\delta w'$,

故知对 $[t_0+\varepsilon\delta t_0, t_1+\varepsilon\delta t_1]$ 中的一切 t, 有

$$|\psi(t)-\tilde{\psi}(t)|\leqslant\varepsilon|\delta w-\delta w'|+\int_\tau^t|\psi'(s)-\tilde{\psi}'(s)|ds, \tag{5.10}$$

令 $\alpha'=\alpha(\beta')$、$\alpha=\alpha(\beta)$ 之后，由(5.8)和(5.9)得

$$\psi'(s)-\tilde{\psi}'(s)=(1-\varepsilon)\{F(s,\psi(s))-F(s,\tilde{\psi}(s))\}$$
$$+\varepsilon\sum_{i=1}^{r}\alpha^i(F_i(s,\psi(s))-F_i(s,\tilde{\psi}(s)))$$
$$+\{g(s,\psi(s),\alpha,\varepsilon)-g(s,\tilde{\psi}(s),\alpha,\varepsilon)\}$$
$$+\varepsilon\sum_{i=1}^{r}(\alpha_i-\alpha_i')F_i(s,\tilde{\psi}(s))$$
$$+\{g(s,\tilde{\psi}(s),\alpha,\varepsilon)-g(s,\tilde{\psi}(s),\alpha',\varepsilon)\}.$$

现在由(4.12)、(2.13)、(5.3)且 $\alpha=\alpha(\beta)$ 收敛于 $\alpha'=\alpha(\beta')$, 可得

$$|\psi'(s)-\tilde{\psi}'(s)|\leqslant\mu(s)|\psi(s)-\tilde{\psi}(s)|+\omega(s,\alpha,\alpha',\varepsilon),$$

其中当 $\alpha\to\alpha'$ 时，$\omega(\cdot, \alpha, \alpha', \varepsilon)$ 在 $\mathscr{I}$ 上依测度收敛于零，且对一切 α、α' 和 ε，$|\omega(s, \alpha, \alpha', \varepsilon)|\leqslant\mu(s)$. 把最后这个不等式代入(5.10)得，

$$|\psi(t)-\tilde{\psi}(t)|\leqslant\varepsilon|\delta w-\delta w'|+\int_{\mathscr{I}}|\omega(s, \alpha, \alpha', \varepsilon)|ds$$
$$+\int_{\tau}^{t}\mu(s)|\psi(s)-\tilde{\psi}(s)|ds.$$

因此，由 Gronwall 不等式(引理 IV.4.2)得

$$|\psi(t_i+\varepsilon\delta t_i)-\tilde{\psi}(t_i+\varepsilon\delta t_i)|$$
$$\leqslant\{\varepsilon|\delta w-\delta w'|+\int_{\mathscr{I}}|\omega(s, \alpha, \alpha', \varepsilon)|ds\}\exp\{\int_{\mathscr{I}}\mu(s)ds\}.$$

由于当 $\alpha\to\alpha'$ 时，$\omega(\cdot, \alpha, \alpha', \varepsilon)\to 0$ 在 $\mathscr{I}$ 上依测度成立，且 $|\omega(s, \alpha, \alpha', \varepsilon')|\leqslant\mu(s)$，则得对所有 $0<\varepsilon<\varepsilon_0$，当 $(\delta t_0, \delta t_1, \delta w, \beta)\to(\delta t_0', \delta t_1', \delta w', \beta')$ 时

$$\psi(t_i+\varepsilon\delta t_i)\to\tilde{\psi}(t_i+\varepsilon\delta t_i), \qquad i=0, 1.$$

这就是我们所要证的.

6. 分 离 引 理

分离引理的证明将涉及到 E^n 中凸集的分离定理、Brouwer 不动点定理的一个推论以及对 E^n 中凸集的一个基本结果. 我们从这些材料的回顾开始.

若 $\mathscr{S}$ 是 E^{n-1} 中集合，用 $a+\mathscr{S}$ 表示形如 $a+s$ 的一切向量 x 的集合，其中 $s\in\mathscr{S}$. 若 $\mathscr{S}_1$ 和 $\mathscr{S}_2$ 是两个集合，它们的线性和(linear span)是指所有形如 $s=s_1+s_2$ 的向量 s 的集合，其中 s_1 在 $\mathscr{S}_1$ 中，s_2 在 $\mathscr{S}_2$ 中. E^{n+1} 中的线性流形是指形如 $a+\mathscr{S}$ 的集合，其中 $\mathscr{S}$ 为 E^{n+1} 的向量子空间. 线性流形的维数是向量子空间 $\mathscr{S}$ 的维数. 一个 n 维子空间亦叫做过原点的超平面. E^{n+1} 中的超平面是 n 维线性流形，它由使 $\langle\nu, y\rangle=\gamma$ 的所有向量 y 组成，其中 ν 是 E^{n+1} 中固定的非零向量，而 γ 为纯量. 若 $\gamma=0$，则超平面通过原点. 方程 $\langle\nu, y\rangle=\gamma$ 叫做超平面方程.

两个集合 A 和 B 叫做被方程为 $\langle\nu, y\rangle=\gamma$ 的超平面分离，如

果对 A 中一切 a, $\langle\nu, a\rangle\leqslant\gamma$, 而对 B 中一切 b, $\langle\nu, b\rangle\geqslant\gamma$.

对集合 A, 所谓 A 的承载平面是指包含 A 的最低维线性流形. 注意到由于每个集合 A 总是包含在 E^{n+1} 中, 故知每个集合 A 有一个承载平面. A 的相对拓扑是由它的承载平面在 A 上诱导的拓扑.

关于凸集的分离性的主要结果如下:

引理 6.1 设 A 和 B 是两个 E^{n+1} 中的凸集. 如果 (i) A 和 B 的承载平面的线性和不是整个的 E^{n+1}, 或者 (ii) A 和 B 的相对内部是不相交的. 则 A 和 B 能被超平面分离.

若 A 和 B 的承载平面的线性和不是整个 E^{n+1}, 则它们的线性和包含在一个超平面(比如 $\langle\nu, y\rangle=\gamma$)中. 这个超平面将作为分离超平面, 因为对 A 中一切 a 和 B 中一切 b, $\langle\nu, a\rangle=\langle\nu, b\rangle=\gamma$. 相对内部不相交的凸集 A 和 B 能用超平面分离这个结论的证明可在 [52] (定理 11.3) 中找到.

设 Z 为欧氏空间 E^d 的真子集, 而 $Z^\perp$ 表示 Z 在 E^d 中的正交补集. 设 $\mathscr{S}$ 为 E^d 的致密凸子集. 则对 $\mathscr{S}$ 中每个 s, 存在 $Z^\perp$ 中唯一的 $y(s)$ 和 Z 中的 $z(s)$, 使得 $s=y(s)+z(s)$. 集合 $(y(s)+Z)\cap\mathscr{S}$ 是非空的且凸和致密的, 所以它有唯一最小范数的元素, 把它记为 y_s. 因此, 可以定义一个把 $\mathscr{S}$ 映入 $\mathscr{S}$ 的边界的映射 Λ 如下:

$$\Lambda(s)=y_s. \tag{6.1}$$

引理 6.2 映射 Λ 在 $\mathscr{S}$ 上是连续的.

首先指出, 若 $\mathscr{S}\subset Z$, 则 $\Lambda(s)$ 在 $\mathscr{S}$ 上是常数, 因而是连续的. 其次设 $\mathscr{S}$ 不含于 Z 中且 Λ 在 $\mathscr{S}$ 上不连续, 则在 $\mathscr{S}$ 中必存在点 s_0 和 $\mathscr{S}$ 中互异的点序列 $\{s_n\}$, 使得 $s_n\to s_0$, 而 $\lim\Lambda(s_n)\neq\Lambda(s_0)$. 因为 $\mathscr{S}$ 是致密的, 则存在 $\{s_n\}$ 的一个子序列, 我们仍把它记作 $\{s_n\}$, 使得 $\Lambda(s_n)$ 在 $\mathscr{S}$ 中收敛于某点 s^*, 而 $s^*\neq\Lambda(s_0)$. 设 $s_n=y_n+z_n$, $s_0=y_0+z_0$, 其中 y_n 和 y_0 在 $Z^\perp$ 中, 而 z_n 和 z_0 在 Z 中. 则 $\Lambda(s_n)=y_n+z_n'$, $\Lambda(s_0)=y_0+z_0'$, 其中 z_n' 和 z_0' 为 Z 中的元素. 又因 $y_n\to y_0$, 所以, $s^*=y_0+z_0^*$, 对某个 $z_0^*\in Z$.

设 $\varepsilon=|s^*|-|\Lambda(s_0)|>0$, 则存在一个整数 k_0, 使得对任何的

$k>k_0$ 和 $j>j_0$, $|\Lambda(s_k)-\Lambda(s_j)|<\varepsilon/4$. 设 k 是大于 k_0 并使 $|\Lambda(s_k)-s^*|<\varepsilon/4$ 的整数. 由于 $\mathscr{S}$ 是凸的, 直线段 $[\Lambda(s_0), \Lambda(s_k)]$ 位于 $\mathscr{S}$ 中, 因为点列 $\{s_n\}$ 是相异的, 故点列 $\{y_n\}$ 亦相异. 因此, $y_k\notin y_0+Z$. 由于 $y_n\to y_0$, 则存在一个整数 $j>k_0$ 使得 y_j+Z 交线段 $[\Lambda(s_0), \Lambda(s_k)]$ 于一点, 此点与 $\Lambda(s_0)$ 的距离小于 $\varepsilon/4$. 因此 $|\Lambda(s_j)|\le|\Lambda(s_0)|+\varepsilon/4$. 所以

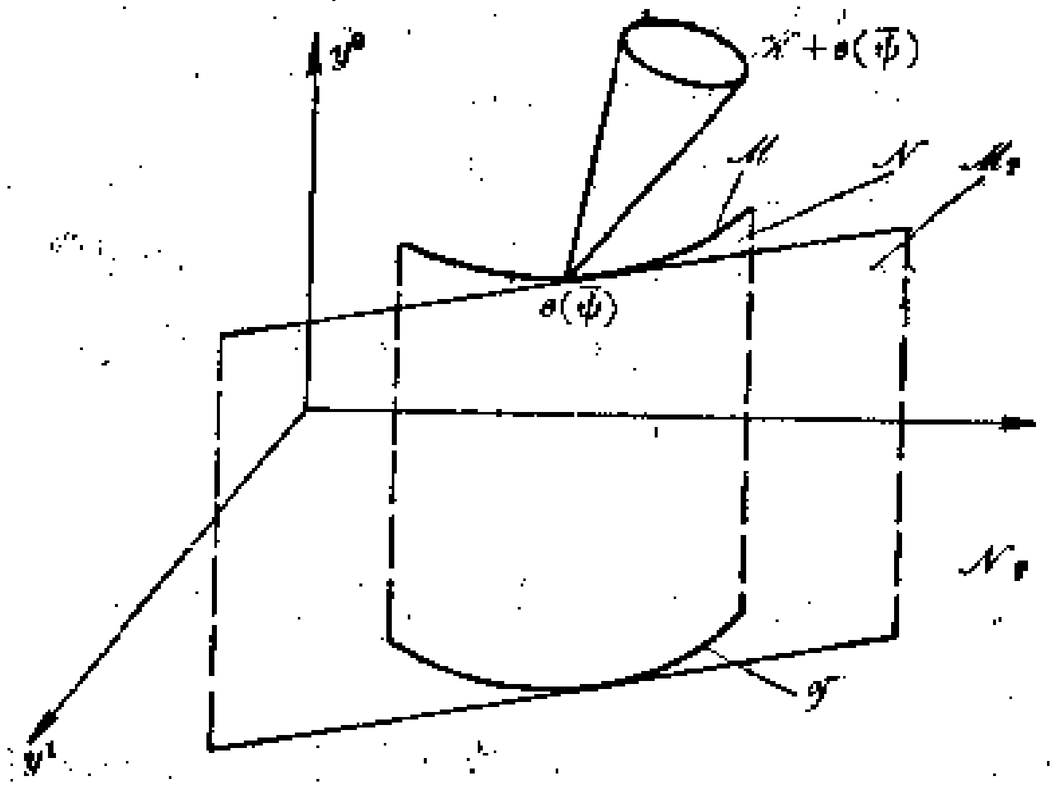

图 2

引理 6.4(分离引理) 集合 $\mathscr{K}$ 和 $\mathscr{N}_T-e(\bar{\psi})$ 可以用过原点的超平面分离.

证明: 不失一般性, 设 $e(\bar{\psi})=0$, 因为这总是能够通过坐标原点的变换来实现的.

设引理不真, 则由于 $\mathscr{K}$ 和 $\mathscr{N}_T$ 都是凸的, 由引理 6.1 知下面的论断是正确的:

(i) $\mathscr{K}$ 的承载平面和 $\mathscr{N}_T$ 的承载平面的线性和为整个空间;

(ii) 存在一点 q, 它为 $\mathscr{K}$ 和 $\mathscr{N}_T$ 二者的相对内点.

由于 q 是 $\mathscr{N}_T$ 的相对内点, 故存在 $\mathscr{N}_T$ 的承载平面中的相对开球 $\mathscr{O}_T$, 使得 $\mathrm{cl}(\mathscr{O}_T)$ 是致密的且含于 $\mathscr{N}_T$ 的相对内部. 所以

$$\mathscr{O}_T\cap\mathscr{M}_T=\emptyset.$$

设 $\mathscr{K}$ 的承载平面的维数为 m, 其中 $0\leqslant m\leqslant 2+2(n+1)$. 则存在 m 维的单纯形 $\mathscr{K}_m$, 使得 $\mathscr{K}_m$ 有 $m+1$ 个顶点, q 在 $\mathscr{K}_m$ 的相对内部之中, 而 $\mathscr{K}_m$ 位于 $\mathscr{K}$ 的相对内部. 令 $q_1, \cdots, q_{m+1}$ 表示 $\mathscr{K}_m$ 的诸顶点, 则在 $\mathscr{V}$ 中存在变分 $P_1, \cdots, P_{m+1}$, 使得 $q_i=L(P_i)$. 设

$$P_i=(\delta t_{0i},\ \delta t_{1i},\ \delta w_i,\ \delta F_i),\ i=1,\ \cdots,\ m+1. \tag{6.2}$$

以 $\mathscr{S}_m$ 表示 $\mathscr{V}(\delta F_1, \cdots, \delta F_{m+1})$ 中的致密单纯形, 它的顶点是 (6.2) 中的变分 P_i. 由于 $\mathscr{S}_m$ 是 $\mathscr{V}(\delta F_1, \cdots, \delta F_{m+1})$ 的致密子集,

由(5.6)确定的映射 h_ε 对一切充分小的 $\varepsilon>0$ 在 $\mathscr{S}_m$ 上有定义，且有引理 5.1 所说的性质.

设 $\mathscr{O}$ 是出现在 $\mathscr{F}$-$\mathscr{N}$ 极值曲线定义中的位于 $E^{2+2(n+1)}$ 空间点 $e(\bar{\psi})$ 处的邻域. 由于 $\mathscr{N}$ 是具有上部边界为 $\mathscr{M}$ 的可微流形，这里 $\mathscr{M}$ 也是可微流形. 因为按 $\mathscr{F}$-$\mathscr{N}$ 极值曲线的定义，$e(\bar{\psi})$ 属于 $\mathscr{M}$，而且由于我们假设 $e(\bar{\psi})=0$，所以下述论断是正确的：存在一个同胚 h^*，它把 $e(\bar{\psi})=0$ 在 $\mathscr{N}_T$ 中的邻域 $\mathscr{O}_T'$ 映满在 $\mathscr{N}$ 中 0 的邻域 $\mathscr{O}'$，其形式为

$$h^*(y)=y+r^*(y), \tag{6.3}$$

其中

$$\frac{|r^*(y)|}{|y|}\to 0,\quad \text{当 } y\to 0,\ y\in\mathscr{O}_T' \text{ 时}. \tag{6.4}$$

不失一般性，假设 $\mathscr{O}_T'$ 和 $\mathscr{O}'$ 使

$$\mathscr{O}_T'\subset\mathscr{O}\quad 和\quad \mathscr{O}'\subset\mathscr{O}. \tag{6.5}$$

图 3 以图解方式说明上述各种定义的集合.

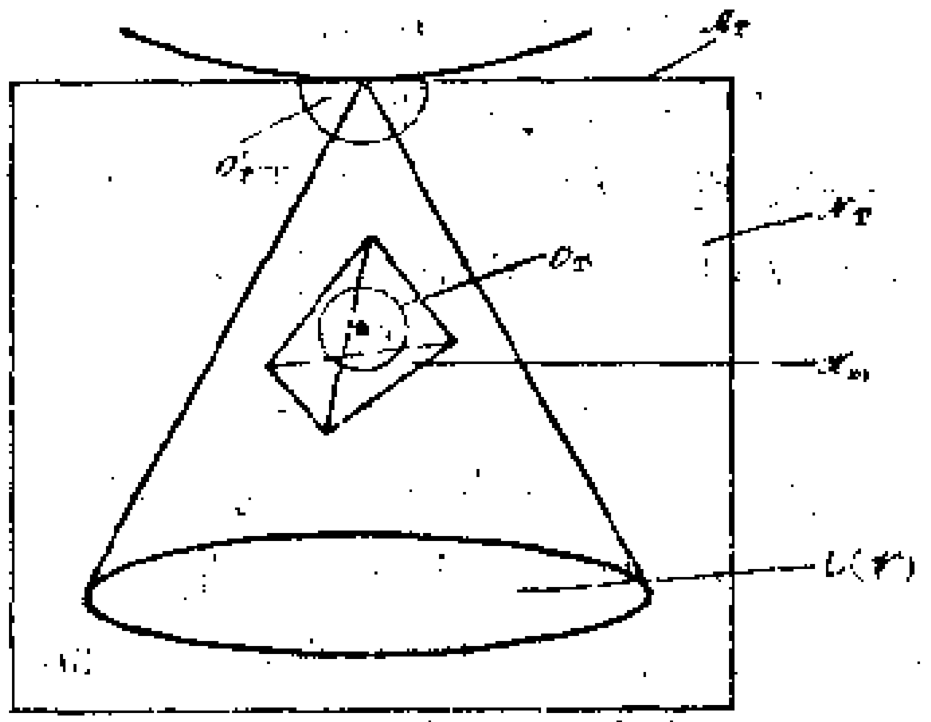

图 3

令 $\mathscr{S}=\mathscr{S}_m\times\mathrm{cl}(\mathscr{O}_T)$，用 σ 表示 $\mathscr{S}_m$ 中的点，η 表示 $\mathrm{cl}(\mathscr{O}_T)$ 中的点. 由于 $\mathrm{cl}(\mathscr{O}_T)$ 是致密的，故得知对充分小的 $\varepsilon>0$，

$$\varepsilon[\mathrm{cl}(\mathscr{O}_T)]\subseteq\mathscr{O}_T'. \tag{6.6}$$

因此，我们可以定义把 $\mathscr{S}$ 映入 $E^{2+2(n+1)}$ 的映射 $\mathscr{V}_\varepsilon$ 如下：

$$\mathscr{V}_\varepsilon(\sigma,\eta)=h_\varepsilon(\sigma)+\frac{1}{\varepsilon}h^*(\varepsilon\eta), \tag{6.7}$$

其中 h_s 由(5.4)定义，h^* 由(6.3)定义．由(6.3)和(5.6)得

$$\mathscr{V}_s(\sigma, \eta)=L(\sigma)-\eta+\rho_s(\sigma)-\frac{1}{\varepsilon}\mathscr{V}^*(\varepsilon\eta).$$

令

$$\begin{cases} L_1(\sigma, \eta)=L(\sigma)-\eta, \\ R(\sigma, \eta, \varepsilon)=\rho_s(\sigma)-\dfrac{1}{\varepsilon}\mathscr{V}^*(\varepsilon\eta), \end{cases} \tag{6.8}$$

则

$$\gamma_s(\sigma, \eta)=L_1(\sigma, \eta)+R(\sigma, \eta, \varepsilon) \tag{6.9}$$

由 $\mathscr{S}_m$ 与 $\mathrm{cl}\mathscr{O}_T$ 的致密性以及(6.4)和引理5.1可得，对 $\mathscr{S}$ 中的每个固定的 (σ, η)，有

$$\lim_{\varepsilon\to 0} R(\sigma, \eta, \varepsilon)=0 \tag{6.10}$$

关于 $\mathscr{S}$ 中的 (σ, η) 一致地成立．

因为 q 属于 $\mathscr{K}_m$ 的相对内部，故在 $\mathscr{S}_m$ 的内部存在一个 σ_0，使得 $L(\sigma_0)=q$．由于 $q\in\mathscr{O}_T$，则 (σ_0, q) 在 $\mathscr{S}$ 内，且

$$L_1(\sigma_0, q)=L(\sigma_0)-q=0.$$

因此，$E^{2+2(n+1)}$ 的原点为在 L_1 下 $\mathscr{S}$ 的象．由于 $\mathscr{K}_m$ 和 $\mathscr{O}_T$ 的承载平面的线性和是整个 $E^{2+2(n+1)}$，故得 $L_1(\mathscr{S})$ 包含一个中心在原点半径为 δ 的球 B_δ 中．由(6.10)可得，存在一个 $\varepsilon_0>0$，使得若 $0<\varepsilon<\varepsilon_0$，则

$$|R(\sigma, \eta, \varepsilon)|<\delta. \tag{6.11}$$

映射 L_1 通常在 $\mathscr{S}$ 上不是一对一的．我们现在确定 $\mathscr{S}$ 的子集 $\mathscr{S}^*$，使得 L_1 对 $\mathscr{S}^*$ 的限制是一对一的且有连续的逆映射．

因为 $\mathrm{cl}(\mathscr{O}_T)\subset\mathscr{N}_T$，而 $\mathscr{N}_T$ 包含在原点的某一流形的切平面内，故知 $\mathrm{cl}(\mathscr{O}_T)$ 的承载平面 Π_1 是一个线性空间．设 Π_2 表示由(6.2)定义的变分 $P_1, \cdots, P_{m+1}$ 所生成的线性空间，于是

$$\Pi_2=\left\{v: v=\sum_{i=1}^{m+1}\alpha^i P_i\right\}.$$

显然，$\mathscr{S}\subseteq\Pi_2\times\Pi_1$．在(6.8)中定义在 $\mathscr{S}$ 上的映射 L_1 可以扩展成在整个 $\Pi_2\times\Pi_1$ 上由公式

$$L_2(v, q)=\sum_{i=1}^{m+1}\alpha^i L(P_i)-q, \quad (v\in\Pi_2,\ q\in\Pi_1)$$

定义的线性映射 L_2.

令 Z 表示 L_2 的核，而 $Z^\perp$ 表示 Z 关于 $\Pi_2\times\Pi_1$ 的正交补集. 设 Λ 相对于 Z 如(6.1)所定义，并且令

$$\mathscr{S}^*=\Lambda(\mathscr{S}),$$

则 $\mathscr{S}^*\subseteq\mathscr{S}$，而 $L_1(\mathscr{S}^*)=L_1(\mathscr{S})$. 以 L^* 表示 L_1 对 $\mathscr{S}^*$ 的限制，则

$$L^*(\mathscr{S}^*)=L_1(\mathscr{S}). \tag{6.12}$$

由 Λ 和 $\mathscr{S}^*$ 的定义得出，若 s_1 和 s_2 为 $\mathscr{S}^*$ 的不同元素，且 $s_1=y_1+z_1$，$s_2=y_2+z_2$，其中 $y_i\in Z^\perp$ 而 $z_i\in Z$，$i=1, 2$，则 $y_1\neq y_2$. 因此，映射 L^* 在 $\mathscr{S}^*$ 上是1-1的，且 L^{*-1} 在 $L^*(\mathscr{S}^*)$ 上定义. 由于 $\mathscr{S}$ 致密且 Λ 连续，故 $\mathscr{S}^*=\Lambda(\mathscr{S})$ 是致密的. 由于 L^* 是连续映射的限制，故它在 $\mathscr{S}^*$ 上连续，因此 L^{*-1} 在 $L^*(\mathscr{S}^*)$ 上连续.

设 ε_0 与(6.11)中的一样. 对每个 $0<\varepsilon<\varepsilon_0$，在 $L^*(\mathscr{S}^*)$ 上定义映射 H_ε 如下：

$$H_\varepsilon(p)=-\gamma_\varepsilon(L^{*-1}(p))+p. \tag{6.13}$$

若令 $s=L^{*-1}(p)$，则 $s\in\mathscr{S}^*$，因而 $s\in\mathscr{S}$. 因此映射 H_ε 就定义好了. 由于 L^{*-1} 是连续的，且对固定的 ε，γ_ε 也是连续的，于是得到，对每个 $0<\varepsilon<\varepsilon_0$，映射 H_ε 是连续的. 借助于(6.9)，就得

$$\gamma_\varepsilon(L^{*-1}(p))=L_1(L^{*-1}(p))+R(s, \varepsilon).$$

但在 $\mathscr{S}^*$ 上 $L_1=L^*$，故

$$\gamma_\varepsilon(L^{*-1}(p))=P+R(s, \varepsilon).$$

因此，由以上讨论和(6.11)得知

$$|H_\varepsilon(p)|=|R(s, \varepsilon)|<\delta. \tag{6.14}$$

前一段推导不等式(6.11)时，我们已指出在 $E^{2+2(n+1)}$ 中半径为 δ、球心在原点的球包含在 $L_1(\mathscr{S})$ 中，根据这个事实，由(6.14)中的不等式和(6.12)可见，H_ε 把 $L^*(\mathscr{S}^*)$ 映到自身. 由于 $\mathscr{S}$ 致密且凸，而 L_1 是线性的，故 $L_1(\mathscr{S})$ 是致密的和凸的. 根据(6.12)，这对于 $L^*(\mathscr{S}^*)$ 同样是正确的. 于是对每个 $0<\varepsilon<\varepsilon_0$，$H_\varepsilon$ 是致密凸集 $L^*(\mathscr{S}^*)$ 到自身的连续映射. 因此，根据引理6.3，映射 H_ε 有一个不动点 $\bar{p}(\varepsilon)$，于是

$$-\gamma_\varepsilon(L^{*-1}(\bar{p}(\varepsilon)))+\bar{p}(\varepsilon)=\bar{p}(\varepsilon),$$

因而

$$\gamma_\varepsilon(L^{*-1}(\bar{p}(\varepsilon)))=0. \tag{6.15}$$

令 $\bar{s}(\varepsilon)=L^{*-1}(\bar{p}(\varepsilon))$. 由于 $\bar{s}(\varepsilon)\in\mathscr{S}^*$, 对 $\mathscr{S}$ 中的某个 $(\bar{\sigma}(\varepsilon), \bar{\eta}(\varepsilon))$ 有 $\bar{s}(\varepsilon)=(\bar{\sigma}(\varepsilon), \bar{\eta}(\varepsilon))$, 于是(6.15)就变为

$$\gamma_\varepsilon(\bar{\sigma}(\varepsilon), \bar{\eta}(\varepsilon))=0.$$

由(6.7)我们还得出

$$h_\varepsilon(\bar{\sigma}(\varepsilon))=\frac{1}{\varepsilon}h^*(\varepsilon\bar{\eta}(\varepsilon)).$$

若我们用 ψ_ε 表示对应于 $\bar{\sigma}(\varepsilon)$ 的轨线, 则由(5.4)及上面最后这个等式得

$$e(\psi_\varepsilon)=h^*(\varepsilon\bar{\eta}(\varepsilon)). \tag{6.16}$$

点 $\bar{\eta}(\varepsilon)$ 在 $\mathscr{O}_T$ 中. 故由(6.6)得, 对充分小的 ε, $\varepsilon\bar{\eta}(\varepsilon)$ 位于 $\mathscr{O}_T$ 中. 因此, 由 h^* 的定义, (6.16)的右端是位于 $\mathscr{N}$ 中的相对开集 $\mathscr{O}'$ 中的一点. 根据(6.5), 它亦位于 $\mathscr{F}-\mathscr{N}$ 极值曲线的定义中出现的邻域 $\mathscr{O}$ 内, 于是

$$e(\psi_\varepsilon)=\mathscr{O}\cap\mathscr{E}\cap\mathscr{N}. \tag{6.17}$$

另一方面, 由于 $\mathrm{cl}\,\mathscr{O}_T$ 包含在 $\mathscr{N}_T$ 的相对内部之中, 且由于 $\bar{\eta}(\varepsilon)\in\mathscr{O}_T$, 故得 $\bar{\eta}(\varepsilon)$ 在 $\mathscr{N}_T$ 的相对内部之中. 因此, $\varepsilon\bar{\eta}(\varepsilon)$ 不可能属于 $\mathscr{M}_T$, 从而 $h^*(\varepsilon\bar{\eta}(\varepsilon))$ 不可能属于 $\mathscr{M}$. 所以 $e(\psi_\varepsilon)$ 不属于 $\mathscr{M}$. 这个结论和(6.17)与 $\bar{\psi}$ 为 $\mathscr{F}-\mathscr{N}$ 极值曲线的假设矛盾. 因此, 我们假设 $\mathscr{K}$ 和 $\mathscr{N}_T$ 不能分离是错误的, 引理证毕.

7. 分离引理的解析推论

定理 3.1 的必要条件蕴含在引理 6.4 中, 即分离引理之中. 我们现在来证明这一点, 并由此完成定理 3.1 的证明.

令 Π 表示过原点且分离 $\mathscr{K}$ 和 $\mathscr{N}_T-e(\bar{\psi})$ 的超平面. 令 c 为 Π 在原点的法线, 设它指向包含 $\mathscr{N}_T-e(\bar{\psi})$ 的半空间. 则对一切 $\zeta\in\mathscr{K}$ 和 $\nu\in\mathscr{N}_T-e(\bar{\psi})$, 成立

$$\langle c, \zeta\rangle\leqslant 0\leqslant\langle c, \nu\rangle. \tag{7.1}$$

设 ν' 是 $\mathscr{M}_T-e(\bar{\psi})$ 中的向量, 则由于 $\mathscr{M}_T-e(\bar{\psi})\subset\mathscr{N}_T-e(\bar{\psi})$, 有

$\langle c, \nu'\rangle\geqslant 0$. 但是 $\mathscr{M}_T - e(\bar{\psi})$ 是一个子空间，故 $-\nu'\in\mathscr{M}_T - e(\bar{\psi})$，因而 $\langle c, -\nu'\rangle\geqslant 0$. 因此，对一切 $\nu'\in\mathscr{M}_T-e(\bar{\psi})$，有 $\langle c, \nu'\rangle=0$. 换句话说，c 在 $e(\bar{\psi})$ 处正交于 $\mathscr{M}_T$.

由(5.5)以及 $\mathscr{K}$ 作为 $\mathscr{V}$ 在 L 作用下的象的定义，得到 $\mathscr{K}$ 中的代表性元素 ζ 具有形式

$$\zeta=(\delta t_0, \delta y_0, \delta t_1, \delta y_1).$$

若把 c 写为

$$c=(c_0^0, c_0, c_1^0, c_1),$$

其中 $c_i=(c_i^1, \cdots, c_i^n)$, $i=0, 1$, 则由(7.1)，对每个 $\zeta\in\mathscr{K}$，有

$$c_0^0\delta t_0+\langle c_0, \delta y_0\rangle+c_1^0\delta t_1+\langle c_1, \delta y_1\rangle\leqslant 0.$$

如果把由(4.38)给出的 $\delta y_i(i=0, 1)$ 的表达式代入上式，便得到对于任意的纯量 $\delta t_i(i=0, 1)$，E^{n+1} 中的任意 δw 和 $\mathrm{co}(\mathscr{F})-F$ 中的任意 δF，有

$$\sum_{i=0}^{1}\Big\{(c_i^0+\langle c_i, \bar{F}_i\rangle)\delta t_i+\langle c_i, \Psi(t_i, \tau)\big[\delta w+\int_\tau^{t_i}\Psi^{-1}(s,\tau)\delta F(s, \bar{\psi}(s))ds\big]\rangle\Big\}\leqslant 0, \tag{7.2}$$

其中 Ψ 是齐次线性方程组

$$\frac{dy}{dt}=F_y(t, \bar{\psi}(t))y \tag{7.3}$$

满足初始条件 $\Psi(\tau, \tau)=I$ 的基本解矩阵，而

$$\bar{F}_i=F_i(t_i, \bar{\psi}(t_i)), \qquad i=0, 1. \tag{7.4}$$

若在(7.2)中取 $\delta t_0=0$, $\delta w=0$, $\delta F=0$ 和 δt_1[注]$=\pm 1$，便得

$$c_1^0+\langle c_1, \bar{F}_1\rangle=0. \tag{7.5}$$

然后在(7.2)中取 $\delta t_0=\pm 1$, $\delta w=0$, $\delta F=0$, $\delta t_1=0$，便得

$$c_0^0+\langle c_0, \bar{F}_0\rangle=0. \tag{7.6}$$

现在若取 $\delta F=0$, δw 为任意，并利用(7.5)、(7.6)则得

$$\langle c_0, \Psi(t_0, \tau)\delta w\rangle+\langle c_1, \Psi(t_1, \tau)\delta w\rangle=0.$$

所以，对 E^{n+1} 中一切 δw，有

$$\langle\Psi^*(t_0, \tau)c_0+\Psi^*(t_1, \tau)c_1, \delta w\rangle=0,$$

[注] 原书误为 δt_i.——译者注.

因而

$$\Psi^*(t_0,\ \tau)c_0 = -\Psi^*(t_1,\ \tau)c_1. \tag{7.7}$$

最后，若我们取 $\delta w=0$，则对 $\mathrm{co}(\mathscr{F})-\overline{F}$ 中的任意 δF，得

$$\langle\Psi^*(t_0,\ \tau)c_0,\ \int_\tau^{t_0}\Psi^{-1}\delta\overline{F}ds\rangle+\langle\Psi^*(t_1,\ \tau)c_1,\ \int_\tau^{t_1}\Psi^{-1}\delta\overline{F}ds\rangle\leqslant 0,$$

其中令 $\delta\overline{F}(s)=\delta F(s,\ \overline{\psi}(s))$. 若应用(7.7)，这个不等式可写成

$$\langle\Psi^*(t_1,\ \tau)c_1,\ \int_{t_0}^{t_1}\Psi^{-1}\delta\overline{F}ds\rangle\leqslant 0,$$

它又可改写为

$$\int_{t_0}^{t_1}\langle\Psi^{*-1}(s,\ \tau)\Psi^*(t_1,\ \tau)c_1,\ \delta F(s,\overline{\psi}(s))\rangle ds\leqslant 0. \tag{7.8}$$

令
$$\eta(t)=\Psi^{*-1}(t,\ \tau)\Psi^*(t_1,\ \tau)c_1,$$
$$\eta_1=\Psi^*(t_1,\ \tau)c_1,$$

则得

$$\eta(t)=\Psi^{*-1}(t,\ \tau)\eta_1, \tag{7.9}$$

和

$$\eta(t_1)=c_1,\ \eta_1=\eta(\tau). \tag{7.10}$$

微分恒等式

$$\Psi(t,\ \tau)\Psi^{-1}(t,\ \tau)=I,$$

并利用(7.3)得:

$$\begin{aligned} O&=\Psi'\Psi^{-1}+\Psi(\Psi^{-1})'=(\overline{F}_y\Psi)\Psi^{-1}+\Psi(\Psi^{-1})' \\ &=\overline{F}_y+\Psi(\Psi^{-1})'. \end{aligned}$$

因此 $(\Psi^{-1})'=-\Psi^{-1}\overline{F}_y$，它等价于关系式

$$(\Psi^{*-1})'=-\overline{F}_y^*(\Psi^{*-1}).$$

于是 Ψ^{*-1} 为线性方程组

$$z'=-\overline{F}_y^*(t,\ \overline{\psi}(t))z$$

满足初始条件 $\Psi^{*-1}(\tau,\ \tau)=I$ 的基本解矩阵. 由此及(7.8)、(7.9)可得

$$\eta'(t)=-\overline{F}_y^*(t,\ \overline{\psi}(t))\eta(t)$$

几乎处处成立. 这个关系式即为定理 3.1 的(3.4)[注].

定理 3.1 中的关系式(3.3)可由(3.1)和(3.2)直接推出.

[注] 原书误为两边再取转置. ——译者注.

为了说明对$[t_0, t_1]$中的所有t，$\eta(t)\neq 0$，首先指出，若对$[t_0, t_1]$中的某个t_2，$\eta(t_2)$为零，则由(7.9)我们应有$\eta_1=0$. 但是，根据(7.10)，这将意味着$c_1=0$. 由(7.5)得$c_1^0=0$，又由(7.7)得$c_0=0$. 然后利用(7.6)得$c_0^0=0$，因此$c=0$，但这是不可能的.

利用(7.9)和(7.10)，对$\mathrm{co}(\mathscr{F})-\overline{F}$中的所有$\delta F$，可把(7.8)写为

$$\int_{t_0}^{t_1}\langle\eta(s),\ \delta F(s,\ \overline{\psi}(s))\rangle ds\leqslant 0. \tag{7.11}$$

若在(7.11)中，取$\delta F=F-\overline{F}$，其中F是$\mathscr{F}$的任意元素，则得对$\mathscr{F}$中一切F，有

$$\int_{t_0}^{t_1}\langle\eta(s),\ \overline{F}(s,\ \overline{\psi}(s))\rangle ds\geqslant\int_{t_0}^{t_1}\langle\eta(s),\ F(s,\ \overline{\psi}(s))\rangle ds,$$

这正好是定理3.1的(3.5).

我们已证明了$c=(c_0^0,\ c_0,\ c_1^0,\ c_1)$在$e(\overline{\psi})$处正交于$\mathscr{M}$. 现在我们通过证明向量(3.6)恰好为$c$来证明(3.6)在$e(\overline{\psi})$处正交于

这些概念的某些事实．至于证明和进一步讨论，读者可参考 Natanson [46]，pp.260～262[注]．

设 E 是直线上的可测集、x_0 为任意点、对 $h>0$ 以 $I(h)$ 表示区间 $[x_0-h,\ x_0+h]$．点 x_0 称为 E 的全密点，如果

$$\lim_{h\to 0}\operatorname{meas}(E\cap I(h))/2h=1.$$

对可测集 E 来说几乎所有 E 的点都是全密点．

设 f 为定义在区间 $[a,\ b]$ 上的实值函数，而 x_0 为 $[a,\ b]$ 的内点．我们称 f 在 x_0 是近似连续的，如果存在一个 $[a,\ b]$ 的可测子集 E 使得 x_0 是 E 的全密点，且 f 限制于 $E\cup\{x_0\}$ 上在 x_0 是连续的，也就是

$$\lim_{\substack{x\to x_0\\ x\in E}} f(x)=f(x_0).$$

定义在区间 $[a,\ b]$ 上的实值可测函数 f 在 $[a,\ b]$ 的几乎所有的点上是近似连续的．

近似连续的定义以及可测函数几乎处处近似连续的命题，可以用考虑各实值分量映射的方法使之从直线推广到 E^n 的映射上去．

现在证明推论 V.3.1．由假设 V.(3.1) 的 (ii)，ϕ 的连续性，$\hat{\lambda}$ 的连续性，以及 u 的可测性可得，由

$$\mathscr{H}(t)=H(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t)) \tag{8.1}$$

定义的映射 $\mathscr{H}$ 在 $[t_0,\ t_1]$ 上是可测的．因此，(8.1) 在 $[t_0,\ t_1]$ 上近似连续．

假设推论 V.3.1 的结论不真，则应当有一个正测度集合 E，在 E 上 V.(3.5) 不成立．设 t_2 为 E 的点，在此点上由 (8.1) 定义的函数 $\mathscr{H}$ 是近似连续的．由于 V.(3.5) 在 t_2 不成立，则存在 $\mathscr{C}$ 中的点 z，使得

$$H(t_2,\ \phi(t_2),\ u(t_2),\ \hat{\lambda}(t_2))-H(t_2,\ \phi(t_2),\ z,\ \hat{\lambda}(t_2))<0. \tag{8.2}$$

由于 $\mathscr{H}$ 在 $[t_0,\ t_1]$ 上近似连续，故由

[注] 中译本见《实变函数论》下册 pp. 321～324．——译者注．

$$\gamma(t)=H(t,\phi(t),u(t),\hat{\lambda}(t))-H(t,\phi(t),z,\hat{\lambda}(t)) \tag{8.3}$$

定义的映射γ在$[t_0,t_1]$上也近似连续．而且t_2是γ的近似连续点．所以，由(8.2)推得，存在一个可测集E_1，$\text{meas}(E_1)>0$，使得对E_1中的t，

$$\gamma(t)<0.$$

现定义控制v如下：

$$v(t)=\begin{cases}u(t), & \text{当 } t\notin E_1,\\ z, & \text{当 } t\in E_1.\end{cases}$$

则
$$\int_{t_0}^{t_1}[H(t,\phi(t),u(t),\hat{\lambda}(t))-H(t,\phi(t),v(t),\hat{\lambda}(t))]dt$$
$$=\int_{E_1}\gamma(t)dt<0,$$

它与V.(3.3)矛盾，这就证明了推论V.3.1.

现在证明推论V.3.2. 以T表示$[t_0,t_1]$中使V.(3.5)成立的诸点t的集合，在T上u是近似连续的，且λ、ϕ是可微的．则$\text{meas}\,T=t_1-t_0$.

令t和t_2为T的点，设$\Delta t=t-t_2$，$\Delta\phi=\phi(t)-\phi(t_2)$，$\Delta\hat{\lambda}=\hat{\lambda}(t)-\hat{\lambda}(t_2)$，而对$0\leqslant s\leqslant 1$，令

$$p(s;t_2,t)=(t_2+s\Delta t,\ \phi(t_2)+s\Delta\phi,\ u(t),\ \hat{\lambda}(t_2)+s\Delta\hat{\lambda}). \tag{8.4}$$

由(8.1)和V.(3.5)得

$$\begin{aligned}\mathscr{H}(t)-\mathscr{H}(t_2)&=H(t,\phi(t),u(t),\hat{\lambda}(t))\\&\qquad-H(t_2,\phi(t_2),u(t_2),\hat{\lambda}(t_2))\\&\leqslant H(t,\phi(t),u(t),\hat{\lambda}(t))\\&\qquad-H(t_2,\phi(t_2),u(t),\hat{\lambda}(t_2))\\&=H(p(1;t_2,t))-H(p(0;t_2,t)).\end{aligned}$$

如果我们对$0\leqslant s\leqslant 1$上定义的函数$s\to H(p(s;t_2,t))$应用中值定理，并将$p(s;t_2,t)$写成$p(s)$，则得到在开区间$(0,1)$中存在θ使

$$\begin{aligned}\mathscr{H}(t)-\mathscr{H}(t_2)\leqslant H_t(p(\theta))\Delta t&+\langle H_x(p(\theta)),\Delta\phi\rangle\\&+\langle H_p(p(\theta)),\Delta\hat{\lambda}(t)\rangle.\end{aligned} \tag{8.5}$$

由于ϕ和λ在$[t_0, t_1]$上连续，它们在$[t_0, t_1]$上有界．又由于$\hat{\lambda}=(\lambda^0, \lambda)$，其中$\lambda^0$是常数，故$\hat{\lambda}$在$[t_0, t_1]$上有界．根据假设，$u$在$[t_0, t_1]$上有界．因此，在$(E^1\times E^n\times E^m\times E^{n+1})$中存在一个闭球$B$，使得对$[t_0, t_1]$中所有$t$、$t_2$及所有$0\leqslant s\leqslant 1$，点$p(s; t_2, t)$在$B$中．则由$\hat{f}_t$和$\hat{f}_x$在$\mathscr{R}\times\mathscr{U}$上的连续性及$\hat{\lambda}$的连续性得到，存在一个常数$K_1>0$，使得对$[t_0, t_1]$中所有$t$、$t_2$及所有$0\leqslant s\leqslant 1$，有

$$\begin{gathered}|H_t(p(s; t_2, t))|\leqslant K_1, \quad |H_x(p(s; t_2, t))|\leqslant K_1,\\ |H_p(p(s; t_2, t))|\leqslant K_1.\end{gathered}\tag{8.6}$$

由 V.(3.2)我们得知

$$\begin{cases}\Delta\phi=\displaystyle\int_{t_2}^{t} H_p(s, \phi(s), u(s), \hat{\lambda}(s))ds,\\ \Delta\lambda=-\displaystyle\int_{t_2}^{t} H_x(s, \phi(s), u(s), \hat{\lambda}(s))ds.\end{cases}$$

由于ϕ、$\hat{\lambda}$和u在$[t_0, t_1]$上有界，以及H_x和H_p连续，故得存在常数K_2，使得对$[t_0, t_1]$中所有t、t_2，

$$|\Delta\phi|\leqslant K_2\Delta t, \quad |\Delta\hat{\lambda}|\leqslant K_2\Delta t.\tag{8.7}$$

从(8.5)、(8.6)及(8.7)得知，存在常数K，使得对T中所有t、t_2，有

$$\mathscr{H}(t)-\mathscr{H}(t_2)\leqslant K|t_2-t|.\tag{8.8}$$

运用类似于推出(8.5)的那些论证，可得

$$\begin{aligned}\mathscr{H}(t)-\mathscr{H}(t_2)&\geqslant H(t, \phi(t), u(t_2), \hat{\lambda}(t))\\ &\quad -H(t_2, \phi(t_2), u(t_2), \hat{\lambda}(t_2))\\ &=H_t(p'(\theta'))\Delta t+\langle H_x(p'(\theta')), \Delta\phi\rangle+\langle H_p(p'(\theta')), \Delta\hat{\lambda}\rangle,\end{aligned}\tag{8.9}$$

其中$0<\theta'<1$，而

$$p'(s)=p'(s; t_2, t)=(t_2+s\Delta t, \phi(t_2)+s\Delta\phi, u(t_2), \hat{\lambda}(t_2)+\Delta\hat{\lambda}).$$

由此运用建立(8.8)的论证方法，我们得出结论

$$\mathscr{H}(t)-\mathscr{H}(t_2)\geqslant -K|t_2-t|.$$

把最后这个不等式与(8.8)合并，就可得出$\mathscr{H}$对于T中的t、t_2满足 Lipshitz 条件：

$$|\mathscr{H}(t)-\mathscr{H}(t_2)|\leqslant K|t-t_2|.$$

由于 $\operatorname{meas} T = t_1 - t_0$, 集合 T 在 $[t_0, t_1]$ 中稠密. 因此, 存在函数 h, 它在整个 $[t_0, t_1]$ 上连续, 且使得对 T 中所有 t, $h(t) = \mathscr{H}(t)$. 于是如同推论 V.3.2 中所断言的, 有

$$h(t) = H(t, \phi(t), u(t), \hat{\lambda}(t))$$

几乎处处成立. 函数 h 亦满足具有与 $\mathscr{H}$ 相同常数的 Lipschitz 条件.

由于 h 满足 Lipschitz 条件, 故它绝对连续. 以 T_1 表示 T 中那些使 h 在其上可微的点, 则 $\operatorname{meas} T_1 = t_1 - t_0$. 今设 t_2 为 T_1 的点, 且 $t_2 \neq t_1$, 则由于 $h'(t_2)$ 存在, 我们有

$$h'(t_2) = \lim_{t \to t_2} \frac{h(t) - h(t_2)}{t - t_2},$$

其中极限是对 $t > t_2$ 且使得 $t \in T_1$ 的那些值取的. 由于 $\operatorname{meas} T_1 = t_1 - t_0$, 这样的点 t 的集合是非空的, 且以 t_2 为极限点. 由于对这样的 t, 有

$$\frac{h(t) - h(t_2)}{t - t_2} = \frac{\mathscr{H}(t) - \mathscr{H}(t_2)}{t - t_2}, \tag{8.10}$$

以及由于 $\Delta t = t - t_2 > 0$, 从(8.5)得到

$$\frac{h(t) - h(t_2)}{\Delta t} \leqslant H_t(p(\theta)) + \left\langle H_x(p(\theta)), \frac{\Delta\phi}{\Delta t} \right\rangle + \left\langle H_p(p(\theta)), \frac{\Delta\hat{\lambda}}{\Delta t} \right\rangle.$$

现令 $t \to t_2$, 从而 $\Delta t \to 0$, 由 ϕ 和 λ 的连续性、(8.4)以及 t_2 是 u 的全密点这个事实, 就得当 $t \to t_2$ 时,

$$p(\theta) \longrightarrow \Pi(t_2) = (t_2, \phi(t_2), u(t_2), \hat{\lambda}(t_2)).$$

从 $\phi'(t_2)$ 和 $\lambda'(t_2)$ 存在这个事实及 H_t、H_x 和 H_p 的连续性, 我们得

$$h'(t_2) \leqslant H_t(\Pi(t_2)) + \langle H_x(\Pi(t_2)), \phi'(t_2) \rangle + \langle H_p(\Pi(t_2)), \lambda'(t_2) \rangle.$$

若应用 V.(3.2)便得

$$h'(t_2) \leqslant H_t(\Pi(t_2)). \tag{8.11}$$

由(8.10)和(8.9)得

$$\frac{h(t)-h(t_2)}{\Delta t} \geqslant H_t(p'(\theta'))+\left\langle H_x(p'(\theta')),\ \frac{\Delta\phi}{\Delta t}\right\rangle$$
$$+\left\langle H_p(p'(\theta')),\ \frac{\Delta\hat{\lambda}}{\Delta t}\right\rangle.$$

运用类似上节用过的那些论证，我们得

$$h'(t_2) \geqslant H_t(H(t_2)).$$

把这个结果和(8.11)合并便得

$$h'(t_2)=H_t(t_2,\ \phi(t_2),\ u(t_2),\ \hat{\lambda}(t_2)).$$

由于 t_2 是 T_1 的任意点和 meas $T_1=t_1-t_0$，我们就证得

$$h'(t)=H_t(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t))$$

几乎处处成立.

若函数 u 分段连续，则除了那些可能使 u 不连续的点外，$\mathscr{H}$ 对所有的点连续．由于 u 是分段连续的，在其不连续的点 τ 处

$$u(\tau+0)=\lim_{t\to\tau} u(t)$$

和

$$u(\tau-0)=\lim_{t\to\tau} u(t)$$

存在且为有限．由于 $\mathscr{C}$ 是闭的，则 $u(\tau+0)$ 和 $u(\tau-0)$ 属于 $\mathscr{C}$. 若 $t<\tau$，则

$$H(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t)) \geqslant H(t,\ \phi(t),\ u(\tau+0),\ \hat{\lambda}(t)),$$

现令 $t\to\tau$，就得

$$H(\tau,\ \phi(\tau),\ u(\tau-0),\ \hat{\lambda}(\tau))$$
$$\geqslant H(\tau,\ \phi(\tau),\ u(\tau+0),\ \hat{\lambda}(\tau)).$$

类似的论证可得

$$H(\tau,\ \phi(\tau),\ u(\tau+0),\ \hat{\lambda}(\tau))$$
$$\geqslant H(\tau,\ \phi(\tau),\ u(\tau-0),\ \hat{\lambda}(\tau)).$$

故得 $\mathscr{H}$ 在所有的点上连续，于是 $\mathscr{H}(t)=h(t)$ 处处成立，因此 $\mathscr{H}$ 是绝对连续的，而且，$\mathscr{H}'(t)=h'(t)$，故

$$\mathscr{H}'(t)=H_t(t,\ \phi(t),\ u(t),\ \hat{\lambda}(t))$$

几乎处处成立.

文 献 评 注

第 一 章

1．第2节的生产计划问题是 Arrow 和 Karlin 在[1]的第四章中提出和解决的，他们给出有关离散时间型问题的早期工作．

2．对于飞行力学问题的早期论述，参看 Leitmann[34]，其中还提供了更早期工作的参考文献．

3．在工程方面的文献里，引起有关例5的工作以及随之发生在这个领域里对最优控制理论兴趣的文章是 McDonald 的[37]．Bushaw 在[14]中假定最优控制是 bang-bang 型的前提下解决了[37]中提出的时间最优问题．

第 三 章

1．例2.2是用于说明变分法中极小函数不存在的一个例子的改进，原来的例子属于 Weierstrass (参看[12] pp. 418～419)．

2．例2.4属于 Roxin [53]．

3．把定理5.1关于每个集合 $\mathscr{Q}^+(t, x)$为凸的假设，换为每个集合 $\mathscr{Q}(t, x)$为凸的这个更严格的假设，本质上是属于 Fillipov [22]和 Roxin [53]的．

4．有关定理5.1的大量应用的例子可在 Athans 和 Falb [2]中找到．

5．在[16]、[17]中，Cesari 引进了集合 $\mathscr{Q}^+(t,x)$和 Cesari 性质，此性质他称之为"性质(Q)"．与推论4.1的形式等价的定理4.1本质上是属于 Cesari 的工作[17]和[18]．在本书的证明中与 Cesari 的证明不同之处在于利用弱的 Cesari 性质，它首先由我们在[7]中给出．

6．关于等度绝对连续的条件(6.2)是 Cesari [18]得到的，然而，推论6.1可追溯到 de la Vallee Poussin (参看 Natanson [46]，p. 159)．

7．定理7.1是由 Meshane 和 Warfield [44] 对 Filippov 在他的文章[22]中引进的关于可测函数的隐函数定理所做的推广．后一结果是熟知的 Filippov 引理．Filippov 引理的另一推广是由 Castaing [15]给出的．

8．Cesari 和他的学生们极其详细地研究了为保证 Cesari 性质而加在 f^0

和 f 上的条件，并非常完整地建立变分法中各种古典条件与 Cesari 性质之间的关系。这些结果概述于[19]中，其中还提供了有关的其他工作的文献。

9．定理 8.1 是我们在[8]中给出的。

第 四 章

1．定理 2.1 的一种推广是在[6]中给出的。

2．松弛轨线的概念可追溯到 L. C. Young [60]，他对变分法问题在"广义曲线"的名义下引进这个概念。包含 Bolza 问题的广义曲线的更详尽的研究的论述，是后来由 E. J. Mcshane 在其连续三篇文章[40]、[41]、[42]中给出的。松弛控制和松弛轨线已由 Warga [58]和 Gamkrelidze [24]独立地且以不同的形式引入控制理论中。在[43]中，Mcshane 改进了他关于广义曲线的较早期结果，并把它应用于约束不必是致密的控制问题中。关于这一点，亦可参阅 Cesari [18]，Warga [19]和 L. C. Young 的书[60]。[60]中关于松弛问题的论述与本文所引用的不同，我们的定义是 Gamkrelidze [24]的。

3．我们关于颤振引理的论述是根据 Gamkrelidze [25]的。

4．关于控制集合是 E^m 中的立方体的线性系统的可达集的性质已由 LaSalle [32]给出。在类似于定理 5.3 那些假设下非线性系统的可达集的性质首先由 Roxin [53]给出。

5．引理 6.1 和定理 6.1 证明的根本思想取自 Lindenstrauss [36]关于 Liapunov 定理的精致证明。引理 6.1 取自 Herms 和 LaSalle（[27]，定理 8.2）。

6．定理 6.3 首先由 Neustadt [47]叙述。推广到非致密情形则是由 Olech [49]和 Jacobs [30]作出的，本书的证明不同于他们的证明。

7．关于线性时间最优问题的 bang-bang 控制的第一个结果可在 LaSalle [31]和[32]、Bellman、Glicksberg 和 Gross [3]以及 Gamkrelidze [23]中找到。上述作者，特别是 Bushaw [14]和 McDonald [37]已意识到最优控制必须是 bang-bang 的。bang-bang 原理的一个比定理 6.3 更漂亮的说法，参见 Sonneborn 和 Van Vleck [56]。

第 五 章

1．按照第 1 节给出的线索，对 W 作较弱假设的最大值原理的推导，可参阅 Berkovitz [5]和 Mirică [45]。

2．例 3.1 可追溯到 Bolza（[12]，pp. 116～117）。

3．关于具有微分方程边界条件的变分问题的乘子法则可追溯到 Euler

和 Lagrange. 然而，直到 1900 年初，乘数法则完美无缺的证明才最后由 Kneser 和 Hilbert 给出. 乘子法则一直到 1909 年的简短发展历史在 Bolza 文章([12], pp. 566~568)中给出，往后的历史述评可在 Bliss[9]中找到.

4. 这里给出的一般性的 Weierstrass 条件首先由 McShane [39] 中证明的. 在[39]以前，Weierstrass 条件是在乘子法则对具有 $\psi^0=-1$ 的唯一一组乘子(ψ^0, ψ)成立的假设下建立的. McShane 去掉了这个要求，在[39]中 McShane 引进了变分凸集. 这个思想以后由 Понтрягин 及其同事们在他们的最大值原理[50]、[51]的证明中进一步改进和发展了.

5. 练习 5.2 是 Berkovitz [4]论述的，有关其他的工作可在那里找到.

第 六 章

1. 这里给出的最大值原理的证明实质上是 Gamkrelide [25]的.

参 考 文 献

[1] K. J. Arrow, S. Karlin and H. Scarf, «Studies in the Mathematical Theory of Inventory and Productions», Stanford University Press, Stanford, California, 1958.

[2] M. Athans and P. Falb, «Optimal Control», McGraw-Hill, New York, 1966.

[3] R. Bellman, I. Glicksberg, and O. Gross, "On the "bang-bang" control problem", Quart. Appl. Math., 14 (1956), 11~18.

[4] L. D. Berkovitz, "Variational methods in problems of control and programming", J. Math. Anal. Appl., 3 (1961), 145~169.

[5] L. D. Berkovitz, "Necessary conditions for optimal strategies in a class of differential games and control problems", SIAM J. Control 5 (1967), 1~24.

[6] L. D. Berkovitz, "An existence theorem for optimal controls", J. Optimization Theory Appl., 6 (1969), 77~86.

[7] L. D. Berkovitz, "Existence theorems in problems of optimal control", Studia Math., 44 (1972), 275~285.

[8] L. D. Berkovitz, "Existence theorems in problems of optimal control without property (Q)", in «Techniques of Optimization», A. V. Balakrishnan ed., Academic Press, New York and London, 1972, 197~209.

[9] G. A. Bliss, "The problem of Lagrange in the calculus of variations", Amer. J. Math., 52 (1930), 673~741.

[10] G. A. Bliss, «Lectures on the Calculus of Variations», The University of Chicago Press, Chicago, 1946.

[11] V. G. Boltyanskii, R. V. Gamkrelidze and L. S. Pontryagin, "The theory of optimal processes I, the maximum principle." Izv. Akad. Nauk SSSR. Ser Mat., 24 (1960), 3~42. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 341~382.

[12] O. Bolza, «Vorlesungen uber Variationsrechnung», Reprint of 1909 edition, Chelsea Publishing Co., New York.

[13] A. E. Bryson and Y. C. Ho, «Applied Optimal Control», Blaisdell 1969, Waltham, Toronto, London.

[14] D. Bushaw, "Optimal discontinuous forcing terms", «Contributions to the Theory of Nonlinear Oscillations IV», Annals of Math Study 41, S. Lefschetz

ed., Princeton University Press, Princeton, (1958), 23~52.

[15] C. Castaing, "Sur les multi-applications mesurables", Rev. Française Automat. Informat. Recherche Opérationnelle 1 (1967), 91~126.

[16] L. Cesari, "Existence theorems for optimal solutions in Pontryagin and Lagrange problems", SIAM J. Control 3 (1966), 475~498.

[17] L. Cesari, "Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints I", Trans. Amer. Math. Soc., 124 (1966), 369~412.

[18] L. Cesari, "Existence theorems for optimal controls of the Mayer type", SIAM J. Control 6 (1968), 517~552.

[19] L. Cesari, "Closure, lower closure, and semicontinuity theorems in optimal control", SIAM J. Control 9 (1971), 287~315.

[20] N. Dunford and J. T. Schwartz, «Linear Operators Part I: General Theory», Interscience, New York, 1958.

[21] H. G. Eggleston, «Convexity», Cambridge University Press, Cambridge, 1958.

[22] A. F. Filippov, "On certain questions in the theory of optimal control", SIAM J. Control 1 (1962), 76~89. Orig. Russ. article in Vestnik Moskov. Univ. Ser. Mat. Mech. Astr., 2 (1959), 25~32.

[23] R. V Gamkrelidze, "Theory of time-optimal processes for linear systems", Izv. Adad. Nauk. SSSR. Ser Mat., 22 (1958), 449~474 (Russian).

[24] R. V. Gamkrelidze, "On sliding optimal regimes", Dokl. Adad. Nauk SSSR. 143 (1962), 1243~1245. Translated as Soviet Math. Dokl., 3 (1962), 390~395.

[25] R. V. Gamkrelidze, "On some extremal problems in the theory of differential equations with applications to the theory of optimal control", SIAM J. Control 3 (1965), 106~128.

[26] E. Hille and R. S. Phillips, «Functional Analysis and Semi-Groups», Revised Ed., American Mathematical Society, Providence, 1957.

[27] H. Hermes and J. P. LaSalle, «Functional Analysis and Time Optimal Control», Academic Press, New York-London, 1969.

[28] M. R. Hestenes, «Calculus of Variations and Optimal Control Theory», John Wiley, New York-London-Sydney, 1966.

[29] J. G. Hocking and G. S. Young, «Topology», Addison-Wesley, Reading, Mass., 1961.

[30] M. Q. Jacobs, "Attainable sets in systems with unbounded controls", J. Differential Equations 4 (1968), 408~423.

[31] J. P. LaSalle, "Study of the Basic Principle Underlying the Bang-Bang

Servo", Goodyear Aircraft Corp. Report GER-5518 (July 1953). Abstract 247t. Bull. Amer. Math. Soc, 60 (1954), 154.

[32] J. P. LaSalle, "The time optimal control problem", «Contributions to the Theory of Nonlinear Oscillations Vol. 5», Annals of Math. Study No. 45 Princeton University Press, Princeton (1960), 1~24.

[33] E. B. Lee and L. Markus, «Foundations of Optimal Control Theory», John Wiley, New York, London, Sydney, 1967.

[34] G. Leitmann, "On a Class of Variational problems in rocket flight", Jour. Acro/Space Sciences 26 (1959), 586~591.

[35] G. Leitmann, «An Introduction to Optimal Control», McGraw-Hill, New York, 1966.

[36] J. Lindenstrauss, "A short proof of Liapounoff's convexity theorem", J. Math. Mech., 15 (1966), 971~972.

[37] D. McDonald, "Non linear techniques for improving servo performance", National Electronics Conference 6 (1950), 400~421.

[38] E. J. McShane, «Integration», Princeton University Press, Princeton, 1944.

[39] E. J. McShane, "On multipliers for Lagrange problems", Amer. J. Math., 61 (1939), 809~819.

[40] E. J. McShane, "Necessary conditions in generalized-curve problems in the calculus of variations", Duke Math. J., 7 (1940), 1~27.

[41] E. J. McShane, "Existence theorem for Bolza problems in the calculus of variations", Duke Math. J., 7 (1940), 28~61.

[42] E. J. McShane, "Generalized curves", Duke Math. J., 6 (1940), 513~536.

[43] E. J. McShane, "Relaxed controls and variational problems", SIAM J. Control 5 (1967), 438~485.

[44] E. J. McShane and R. B. Warfield, Jr., "On Filippov's implicit functions lemma", Proc. Amer. Math. Soc., 18 (1967), 41~47.

[45] S. Mirică, "On the admissible synthesis in optimal control theory and differential games", SIAM J. Control 7 (1969), 292~316.

[46] I. P. Natanson, "Theory of Functions of a Real Variable", Eng. Trans. by Leo Boron, revised ed., F. Ungar, New York, 1961.(有中译本)

[47] L. W. Neustadt, "The existence of optimal controls in the absence of convexity conditions", J. Math. Anal. Appl., 7 (1963), 110~117.

[48] L. W. Neustadt, «Optimization: A Theory of Necessary Conditions», Princeton University Press, Princeton, New Jersey, To Appear.

[49] C. Olech, "Extremal Solutions of a control system", J. Differential Equations 2 (1966), 74~101.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press, Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padeva 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本, 上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本，上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press, Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本，上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本, 上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本，上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本, 上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad unae misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press, Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本,上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3~20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321~339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109~119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102~106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151~159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111~128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212~234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.

[50] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko. «The Mathematical Theory of Optimal Processes», (Translated by K. N. Trirogoff, L. W. Neustadt, editor) John Wiley, 1962. (有中译本, 上海科学技术出版社出版)

[51] L. S. Pontryagin, "Optimal regulation processes", Uspehi Mat. Nauk (N. S.) 14 (1959), 3～20. English translation in Amer. Math. Soc. Transl. Ser. 2, 18 (1961), 321～339.

[52] R. T. Rockafellar, «Convex Analysis», Princeton University Press. Princeton, 1970.

[53] E. Roxin, "The existence of optimal controls", Mich. Math. J., 9 (1962), 109～119.

[54] W. Rudin, «Real and Complex Analysis», McGraw-Hill, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1966.

[55] G. Scorza-Dragoni, "Un teorema sulle funzioni continue rispetto ad una e misurabile rispetto ad un'altra variabile", Rend. Sem. Mat. Univ. Padova 17 (1948), 102～106.

[56] L. M. Sonneborn and F. S. Van Vleck, "The bang-bang principle for linear control systems", SIAM J. Control 2 (1964), 151～159.

[57] M. M. Vainberg, «Variational Methods for the Study of Nonlinear Operators», Eng. Trans. by A. Feinstein, Holden-Day, San Francisco, London, Amsterdam, 1964.

[58] J. Warga, "Relaxed variational problems", J. Math. Anal. Appl., 4 (1962), 111～128.

[59] J. Warga, «Optimal Control of Differential and Functional Equations», Academic Press, New York, 1972.

[60] L. C. Young, "Generalized curves and the existence of an attained absolute minimum in the calculus of variations", Compt. Rend. Soc. Sci. et Lettres. Varsovie, Cl III 30 (1937), 212～234.

[61] L. C. Young, «Lectures on the Calculus of Variations and Optimal Control Theory», W. B. Saunders Co., Philadelphia, London, Toronto, 1969.