[苏]P·B·格姆克列里兹

最优控制理论基础

复旦大学出版社

最优控制理论基础

〔苏〕P·B·格姆克列里兹　著
姚允龙　尤云程　译
金福临　审校

复旦大学出版社

内 容 简 介

本书是继庞特里雅金等著的"最佳过程的数学理论"之后，反映最优控制理论进展的重要著作，作者在书中以自己独创的广义控制为工具，严格地证明了庞特里雅金的最大值原理，进而得到了相当广泛的一类由常微分方程所确定的非线性系统最优解的存在性条件。

本书可作为理工科高校有关专业的高年级学生及研究生的教材，也可供从事自动控制理论的教学、研究与应用的高等院校教师、科学工作者和工程技术人员阅读参考。

最 优 控 制 理 论 基 础

复旦大学出版社出版
（上海国权路579号）

新华书店上海发行所发行　　上海蔡家弄印刷厂印刷

开本 850×1168　1/32　印张 5　字数 144,000
1988年第2月第1版　1988年2月第1次印刷
印数：1—6,000

ISBN 7—309—00124—9/O·035　定价：1.05元

中译本序

本书是继庞特里雅金等著的“最佳过程的数学理论”之后，在最优控制理论方面的一部重要著作。

控制理论中的变分方法首先由美国著名控制理论专家 L. D. Berkovitz 教授在1961年提出。本书作者格姆克列里兹（P. B. Гамкрелидзе）继承并发扬光大了这一方法。1984 年 L. D. Berkovitz 在复旦大学讲学时指出，迄今为止，控制理论中最先进的变分方法已表述在格姆克列里兹所发展的理论和著作中。

作者在本书中应用广义控制的概念给出了庞特里雅金最大值原理的一个新证明，他所采用的方法的优点还体现在最优控制存在性的证明中。近年来，本人与李训经教授在对无限维拟线性带时滞的发展系统的最大值原理所作的研究中，进一步发现了该方法对分布参数控制系统也是有效的。

本译本是根据1978年版的英译本译出的，为了方便读者阅读，中译本中对原书的若干段落的叙述作了更动，有些地方加上了译注。

金福临教授详细地审阅了全稿，并提出许多宝贵的修改意见．复旦大学出版社为本书中译本的出版给予了很大的支持，在此一并表示谢忱。

姚允龙

1984年4月

前　　言

本书是1974年秋我在第比利斯国立大学的讲稿基础上整理而成的。其主要内容是控制理论的一般原理，最大值原理的证明，以及最优控制理论的基本存在定理。

我认为本书所介绍的基本定理的证明虽然未必是最简捷的，但在理论上是完全严格的。我们所引进的概念与方法有一个无可置疑的优点——它们不仅对本书中的定理证明有用，而且对整个控制理论有用.

本书仅限于处理具有固定端点的时间最优问题。这样做既避免了一般性结果表述上的烦赘，同时又保留了最优控制问题所特有的一切实质性难点。所得到的结果推广到具有一般边界条件和更一般的目标泛函的问题通常并无任何特殊困难，并且在多数情形可自动地实现。

本书可作为应用数学系选修或专修最优化理论的学生与研究生一学期课程的教材。假定学生已熟悉最优控制理论的初步知识，包括线性系统理论与凸集理论的基础，并假定读者已掌握一般测度论的最基本的概念。

P·B·格姆克列里兹

目　　录

第一章　时间最优问题与最大值原理

1.1　最优问题的提法

考察 n 维欧氏空间 $\boldsymbol{R}^n$ 中的微分方程

$$\dot{x}=f(t,x,u), \tag{1.1}$$

我们称它为受控方程，其中，

$$x=\begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}\in \boldsymbol{R}^n$$

称为相点，而

$$u=\begin{bmatrix} u^1 \\ \vdots \\ u^r \end{bmatrix}\in \boldsymbol{R}^r$$

称为控制参数[注1]，向量

$$f(t,x,u)=\begin{bmatrix} f^1(t,x,u) \\ \vdots \\ f^n(t,x,u) \end{bmatrix}\in \boldsymbol{R}^n$$

称为相速度。假定 f 是定义在

$$\boldsymbol{R}^{1+n+r}=\{(t,x,u):t\in \boldsymbol{R},x\in \boldsymbol{R}^n,u\in \boldsymbol{R}^r\}$$

上的一个连续函数且关于 x 连续可微：

$$\frac{\partial f(t,x,u)}{\partial x}=\left(\frac{\partial f(t,x,u)}{\partial x^1},\cdots,\frac{\partial f(t,x,u)}{\partial x^n}\right).$$

U 是“参数空间” $\boldsymbol{R}^r$ 中一个任意给定的集合，这个集合是控制参

[注1]　u 是 t 的函数 $u(t)$。——译者注。

数 u 的允许值集. 任意一个定义在 $\boldsymbol{R}$ 上并取值于集合 U 中的有界可测函数称为允许控制。允许控制全体组成的集合 Ω_U 称为允许控制类。

任取一允许控制 $u(t)$ 代入受控方程 (1.1) 得微分方程

$$\dot{x}=f(t,x,u(t))=F(t,x),$$

其右端函数 $F(t,x)$ 关于 x 连续可微且关于 t 可测。在第四、五章将详细讨论这类方程，并证明所有必要的结论。现在我们只须注意以下事实：如果 $x(t)$ 是 $[t_1,t_2]$ 上的绝对连续函数，且对几乎所有的 $t\in[t_1,t_2]$ 满足等式

$$\dot{x}(t)=f(t,x(t),u(t))=F(t,x(t)),$$

则称它是这一方程在区间 $[t_1,t_2]$ 上的一个解。由于 $x(t)$ 的绝对连续性，这等价于 $x(t)$ 对一切 $t\in[t_1,t_2]$ 满足积分方程

$$x(t)=x(t_1)+\int_{t_1}^{t}F(\theta,x(\theta))d\theta.$$

显然，如果 $u_1(t)$ 与 $u_2(t)$ 对几乎所有的 $t\in\boldsymbol{R}$ 相等，则微分方程

$$\dot{x}=f(t,x,u_1(t)) \text{ 与 } \dot{x}=f(t,x,u_2(t))$$

是等价的.

受控对象是由受控方程(1.1)和允许控制类 Ω_U 确定的。这个对象的“状态”可用相点 x 描述，它的相轨迹被选定的允许控制 $u\in\Omega_U$ 所控制。如果在时刻 $t=t_1$ 对象的初态 x_1 已知并且控制 $u(t)$ 已给定，则对象的相轨迹唯一地由方程

$$\dot{x}=f(t,x,u(t)),\qquad x(t_1)=x_1,$$

确定。

由受控方程 (1.1) 与允许控制类 Ω_U 所给定的控制问题将写成如下形式，

$$\dot{x}=f(t,x,u),\qquad u(t)\in\Omega_U. \tag{1.2}$$

具有固定端点的时间最优问题可表述如下：给定相空间 $\boldsymbol{R}^n$ 中的两点 x_1 和 x_2，以及初始时刻 t_1，求一个允许控制 $\tilde{u}\in\Omega_U$，使得方程

$$\dot{x}=f(t,x,\tilde{u}(t))$$

在 $[t_1,t_2]$ 上有一个解 $\tilde{x}(t)$ 满足边界条件

$$\tilde{x}(t_1)=x_1,\qquad \tilde{x}(t_2)=x_2,$$

并且从 x_1 转移到 x_2 的时间为最小：

$$t_2 - t_1 \longrightarrow \min.$$

这样，对任何其他的允许控制 $\hat{u}(t)$，如果在 $[t_1,\hat{t}_2]$ 上，方程 $\dot{x} = f(t,x,\hat{u}(t))$ 也存在解 $\hat{x}(t)$，并且满足同样的边界条件

$$\hat{x}(t_1) = x_1, \qquad \hat{x}(\hat{t}_2) = x_2,$$

则必有 $\hat{t}_2 - t_1 \geqslant t_2 - t_1$，即 $\hat{t}_2 \geqslant t_2$.

控制 $\tilde{u}(t)$ 将称为（时间）最优控制，$\tilde{x}(t)$ 将称为对应的最优轨线，而 $t_2 - t_1$ 称为在初始条件 $x(t_1) = x_1$ 下从 x_1 到 x_2 的最优（转移）时间。

$\{\tilde{u}(t),\tilde{x}(t)\}, t \in [t_1,t_2]$，将称为下列（时间）最优问题的解，或最优解：

$$\begin{aligned} &\dot{x} = f(t,x,u), \qquad u(t) \in \Omega_U, \\ &t = t_1, \qquad x(t_1) = x_1, \\ &x(t_2) = x_2, \qquad t_2 - t_1 \longrightarrow \min. \end{aligned} \tag{1.3}$$

应注意 $t_1, x(t_1)$ 和 $x(t_2)$ 在边界条件中是已经给定的，但是 t_2 不是给定的．在方程（1.1）自治的情形，即它的右端不依赖于 t 的情形，边界条件中初始时刻 t_1 的固定显然是无关紧要的。

也可以研究由同样的受控方程（1.1）但是允许控制类比 Ω_U 更为广泛的受控对象的最优问题。然而，我们仅限于已经给出的 Ω_U。

有时需要缩小控制类 Ω_U，如仅考虑取值在 U 中的逐段连续控制，甚至逐段常值控制．如果控制 $u(t) \in \Omega_U$ 在任一有界区间上至多有有限个不连续点且在每一不连续点处存在左右极限，则称它为逐段连续。如果 t 轴可分为有限个区间，在每一区间上 $u(t)$ 取一常值，则此 $u(t) \in \Omega_U$ 称为逐段常值的。

1.2 带参数的典则方程组与庞特里雅金最大值条件

首先，我们约定若干记号。列向量用罗马字母来标记，而行向量总以希腊字母来标记，如 $\xi=(\xi_1,\ \cdots,\ \xi_n)$。一个行向量与同维列向量的内积可写成矩阵相乘的形式：

$$\xi x=(\xi_1,\cdots,\xi_n)\begin{bmatrix}x^1\\ \vdots\\ x^n\end{bmatrix}=\sum_{i=1}^{n}\xi_i x^i.$$

向量值函数$f(t,x,u)$关于向量 x 的坐标分量的 Jacobi 阵将记为

$$f_x(t,x,u)=\frac{\partial f(t,x,u)}{\partial x}=\left(\frac{\partial f^i(t,x,u)}{\partial x^j}\right),\qquad i,j=1,\cdots,n.$$

任一 $n\times m$ 矩阵 $A=(a_{ij})(i=1,\cdots,n;j=1,\cdots,m)$的范数取欧氏范数

$$|A|=(\sum_{i,j}a_{ij}^2)^{\frac{1}{2}}.$$

定义时间 t 和以下三个向量

$$x=\begin{bmatrix}x^1\\ \vdots\\ x^n\end{bmatrix},\quad \psi=(\psi_1,\cdots,\psi_n),\quad u=\begin{bmatrix}u^1\\ \vdots\\ u^r\end{bmatrix}$$

的数值函数如下：

$$H(t,x,\psi,u)=\psi f(t,x,u)=\sum_{i=1}^{n}\psi_i f^i(t,x,u).$$

函数 H 是向量 ψ 与受控对象的相速度向量 $f(t,x,u)$的内积。它关于ψ是线性的，关于 x 是连续可微的。

记

$$\frac{\partial H}{\partial\psi}=\begin{pmatrix}\dfrac{\partial H}{\partial\psi_1}\\ \vdots\\ \dfrac{\partial H}{\partial\psi_n}\end{pmatrix}=\begin{pmatrix}f^1\\ \vdots\\ f^n\end{pmatrix}=f,$$

$$\frac{\partial H}{\partial x}=\left(\frac{\partial H}{\partial x^1},\cdots,\frac{\partial H}{\partial x^n}\right)=\left(\sum_{i=1}^{n}\psi_i f^i_{x^1}(t,x,u),\cdots,\sum_{i=1}^{n}\psi_i f^i_{x^n}(t,x,u)\right)$$

$$=\psi f_x(t,x,u).$$

向量值函数 x 和 ψ 的方程组：

$$\begin{aligned}\dot{x}&=\frac{\partial H(t,x,\psi,u)}{\partial\psi}=f(t,x,u),\\ \dot{\psi}&=-\frac{\partial H(t,x,\psi,u)}{\partial x}=-\psi f_x(t,x,u),\end{aligned}\tag{1.4}$$

称为 Hamilton 方程组，或称为关于 x 与 ψ 的典则方程组. 函数 H 称为该系统的 Hamilton 函数. 除了变量 x 和 ψ 以外，函数 H 还含有时间 t 及变量 u，u 可以看作为一个参量。

一般地说，任意一个二次连续可微的数值函数 $H(t,x,\psi)$ 可生成对应的 Hamilton 系统或典则系统：

$$\begin{aligned}\dot{x}&=\frac{\partial H(t,x,\psi)}{\partial\psi}=\begin{pmatrix}\dfrac{\partial H}{\partial\psi_1}\\ \vdots\\ \dfrac{\partial H}{\partial\psi_n}\end{pmatrix},\\ \dot{\psi}&=-\frac{\partial H(t,x,\psi)}{\partial x}=-\left(\frac{\partial H}{\partial x^1},\cdots,\frac{\partial H}{\partial x^n}\right),\end{aligned}\tag{1.5}$$

而函数 H 称为这一系统的 Hamilton 函数。

这一系统具有一个值得注意的性质：如果 $x(t)$，$\psi(t)$ 在 $[t_1,t_2]$ 上是该系统的一个解，则下式成立，

$$\frac{d}{dt}H(t,x(t),\psi(t))=\frac{\partial}{\partial t}H(t,x(t),\psi(t)),$$

其中，左侧是关于 t 的全导数，右侧是关于 t 的偏导数。特别，当 H 不显式依赖于 t 时，便有

$$\frac{d}{dt}H(t,x(t),\psi(t))=0.$$

这一事实可由直接计算得到：

$$\frac{d}{dt}H(t,x(t),\psi(t))=\dot{\psi}\frac{\partial H}{\partial \psi}+\frac{\partial H}{\partial x}\dot{x}+\frac{\partial H}{\partial t}$$

$$=-\frac{\partial H}{\partial x}\frac{\partial H}{\partial \psi}+\frac{\partial H}{\partial x}\frac{\partial H}{\partial \psi}+\frac{\partial H}{\partial t}=\frac{\partial H}{\partial t}.$$

不含任何其它参数的系统(1.5)的解 $x(t),\psi(t)$ 是由初始条件

$$t=t_1,\qquad x(t_1)=\hat{x},\qquad \psi(t_1)=\hat{\psi}$$

唯一确定的。相比之下，仅当我们以 t 的一个特定函数代入 H 中的参数 u 时，系统(1.4)的解才由初始条件唯一确定。到现在为止，除了条件 $u(t)\in\Omega_U$ 而外，我们对函数 $u(t)$ 的选取并无其它限制。

为了作出确定的选取，我们假定已给出联系着 t,x,ψ 和 u 的一个关系：

$$\omega(t,x,\psi,u)=0. \tag{1.6}$$

粗糙地讲，这一关系应使我们能用 t,x 和 ψ 表示出 u(我们暂不追究这一函数的具体形式)，因而就能从方程 (1.4) 中消去参数 u。为此，将表示式

$$u=u(t,x,\psi)$$

代入(1.4)中的 u，便得到关于 x 和 ψ 的微分方程组：

$$\dot{x}=f(t,x,u(t,x,\psi)),$$

$$\dot{\psi}=-\psi f_x(t,x,u(t,x,\psi)).$$

在函数 $u(t,x,\psi)$ 充分正则（例如，连续可微）的假定下，只要初始条件给定，这一方程组的解唯一地存在。

但是，方程 (1.6) 关于 u 的正则可解性要加相当严格的限制条件。我们将用如下方法来代替借助于(1.6)从(1.4)中消去 u 的做法。为此，给出下面的精确定义：

设存在绝对连续函数 $x(t)$ 和 $\psi(t)$ ($t\in[t_1,t_2]$) 以及一个可测函数 $u(t)\in\Omega_U$，满足微分方程组(1.4)，方程(1.6)以及初始条件

$$t=t_1,\qquad x(t_1)=x_1,\qquad \psi(t_1)=\psi_1,$$

那么，我们称函数组 $x(t)$，$\psi(t)$ 与 $u(t)$ 为典则方程组(1.4)具有给定初始条件的解，并且称这一解是由关系式(1.6)从典则方程组(1.4)中

消去参数 u 而得到的。这里，函数 $u(t)$ 可测，且等式

$$\omega(t,x(t),\psi(t),u(t))=0$$

对几乎所有 $t\in[t_1,t_2]$ 成立。

古典变分法的正则问题中会遇到方程 (1.6) 关于 u 的正则性与唯一可解性问题。在本章之末将考虑这一情形。

我们主要关心如下形状的条件(1.6)：

$$\boldsymbol{H}(t,x,\psi,u)=M(t,x,\psi), \tag{1.6$'$}$$

这里函数 $M(t,x,\psi)$ 是

$$M(t,x,\psi)=\sup_{u\in U}\boldsymbol{H}(t,x,\psi,u).$$

(1.6)′ 称为庞特里雅金最大值条件。

如果 $x(t),\psi(t),u(t)\ (t\in[t_1,t_2])$ 组成典则方程组(1.4)的一个解并对几乎所有的 $t\in[t_1,t_2]$ 满足庞特里雅金最大值条件，即满足等式

$$\boldsymbol{H}(t,x(t),\psi(t),u(t))=M(t,x(t),\psi(t)),$$

则称这组解是用庞特里雅金最大值条件从典则方程组 (1.4) 中消去参数 u 而得到的解。

1.3 庞特里雅金最大值原理

我们将要叙述的定理 1.1，即庞特里雅金最大值原理，是最优问题 (1.3) 的任一解所满足的必要条件。我们还将讨论运用这一条件实际地寻求最优解的可能性。最大值原理的证明将在第七章里给出。

定理 1.1 设

$$\tilde{u}(t)\in\Omega_U,\qquad \tilde{x}(t),\qquad t\in[t_1,t_2],$$

是最优问题(1.3)的解。那么，必存在非零绝对连续函数

$$\tilde{\psi}(t)=(\tilde{\psi}_1(t),\cdots,\tilde{\psi}_n(t)),\quad t\in[t_1,t_2],$$

使得对几乎所有的 $t\in[t_1,t_2]$，函数组 $\tilde{u}(t),\tilde{x}(t)$ 与 $\tilde{\psi}(t)$ 满足典则方程组

$$\dot{x}=\frac{\partial}{\partial\psi}H(t,x,\psi,u),$$
$$\dot{\psi}=-\frac{\partial}{\partial x}H(t,x,\psi,u),\tag{1.7}$$

其中，Hamilton 函数是

$$H(t,x,\psi,u)=\psi f(t,x,u),\tag{1.8}$$

并且满足庞特里雅金最大值条件

$$H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))=M(t,\tilde{x}(t),\tilde{\psi}(t)).\tag{1.9}$$

函数 $M(t,\tilde{x}(t),\tilde{\psi}(t)$,必是 $t\in[t_1,t_2]$的连续函数，且在 $t=t_2$ 处成立如下不等式，

$$M(t_2,\tilde{x}(t_2),\tilde{\psi}(t_2))\geqslant 0.\tag{1.10}$$

我们将在下一节里讨论最大值条件 (1.9) 和不等式(1.10)的简单几何解释。

关系式组(1.7)—(1.10)满足边界条件

$$t=t_1,\qquad x(t_1)=x_1,\qquad x(t_2)=x_2,\tag{1.11}$$

且 $\psi(t)$非零的任一解 $u(t),x(t),\psi(t)$ 称为最优问题(1.3)的一个极值解。函数

$$H(t,x,\psi,u)=\psi f(t,x,u)$$

称为最优问题(1.3)的 Hamilton 函数。

最大值原理断言，最优问题 (1.3) 的解必包含在该问题的极值解集之中。因此，我们简短地讨论一下这样一个问题：在何种程度上，最大值条件(1.9)与不等式(1.10)连同边界条件 (1.11) 能确定典则方程组(1.7)的解？在这里，最大值条件应被看作 t, x, ψ 和 u 之间的一个关系式(1.6)，由它从方程组(1.7)中消去参数 u.

首先，我们看出函数 $\tilde{\psi}(t)$ 不恒为零的假定是很重要的。事实上，由于 H 关于 ψ 是线性的，故满足

$$\dot{x}=\frac{\partial}{\partial\psi}H(t,x,\psi,u)=f(t,x,u),\quad u(t)\in\Omega_U,$$

$$t=t_1,\qquad x(t_1)=x_1,\qquad x(t_2)=x_2$$

的任一解$u(t), x(t)$连同函数$\psi(t)\equiv 0$必满足关系式组(1.7)—(1.11)。因此，若没有条件$\tilde{\psi}(t)\neq 0$，则最大值原理将成为空洞的断言。有了非零的$\psi(t)$，一般来说，利用最大值原理就能以令人满意的方式从(1.7)中消去参数u，从而得到边值问题(1.7)—(1.11)的一个或几个孤立解。

应看到，若$\tilde{\psi}(t)\neq 0$，即$\tilde{\psi}(t)$不恒等于零，则$\tilde{\psi}(t)\neq 0, \forall t\in[t_1, t_2]$，这是因为$\tilde{\psi}(t)$是线性齐次方程

$$\dot{\psi}=-\psi f_x(t, \tilde{x}(t), \tilde{u}(t))$$

的解的缘故。

对$\psi(t)$未加边界条件是非常自然的。考虑到关系式组(1.7)—(1.11)中的"实质参数"的个数应等于其中"独立条件"的个数，因此这样的系统不是超定的。但是，添加任何新的（独立）条件将导致系统的超定。

实际上，设初始值$t_1, x(t_1), \psi(t_1)$及终止时间t_2为边值问题(1.7)—(1.11)的参数，即一共有$2n+2$个数值参数。如果$t_1, x(t_1), \psi(t_1)$已给定，那么，在最大值条件能够从(1.7)消去u的这一假定下，也即在方程

$$H(t, x(t), \psi(t), u(t))=M(t, x(t), \psi(t))$$

对几乎所有的t可唯一地解出u的假定下，(1.7)—(1.9)之解$u(t)$，$x(t)$，$\psi(t)$将唯一地随时间而变化。但现在t_1与$x(t_1)$在边界条件(1.11)中已被固定，所以$n+1$个数值参数$\psi(t_1), t_2$是自由的。由于有这些自由参数，我们能使其余几个数值条件$x(t_2)=x_2$得以满足。

我们要证明，如果

$$u(t),\quad x(t),\quad \psi(t),\quad t\in[t_1, t_2],$$

是关系式组(1.7)—(1.11)的一个解，则对任一非负常数λ，函数组

$$u(t),\quad x(t),\quad \lambda\psi(t),\quad t\in[t_1, t_2],$$

也是(1.7)—(1.11)的一个解。因此，在向量$\psi(t_1)$的n个坐标中，事实上只有$n-1$个实质参数。再加上参数t_2，共有n个可以变动的实质参数，我们必须用这些参数的适当取值来满足在右端点处的边界条件

$x(t_2)=x_2$。

只要注意到，函数 H 关于 ψ 是线性的，函数 $M(t,x,\psi)$ 关于 ψ 是正齐次的，即

$$M(t,x,\lambda\psi)=\lambda M(t,x,\psi), \qquad \forall\lambda\geqslant 0,$$

不等式(1.10)的左端 $\geqslant 0$，以及边界条件(1.11)并不包含 $\psi(t)$ 的任何值，前一段所述的事实不难得知了。

既然最优问题(1.3)的解必在边值问题 (1.7)—(1.11) 的解集之中，求解这种边值问题的方法就非常重要。在最优化理论中对之应有专门章节来讨论[注1]。

1.4 最大值条件的几何解释

我们以 $P(t,x)$ 记控制问题(1.2)中给定 t 和 x 之值的所有可能的相速度的集合：

$$P(t,x)=\{f(t,x,u):u\in U\}\subset \boldsymbol{R}^n.$$

考虑问题 (1.3) 的一个极值解 $\tilde{u}(t)$，$\tilde{x}(t)$，$\tilde{\psi}(t)$，$t\in[t_1,\ t_2]$，并设 $\Pi_t\subset\boldsymbol{R}^n$ 是在 t 时刻正交于 $\tilde{\psi}(t)$ 的 $(n-1)$ 维子空间。最大值条件(1.9)是说，对几乎所有的 $t\in[t_1,t_2]$，向量 $\tilde{\psi}(t)$ 与相速度向量 $f(t,\tilde{x}(t),\tilde{u}(t))$的内积（其中 $\tilde{x}(t)$，$\tilde{u}(t)$是一极值轨线与控制）不小于这一向量 $\tilde{\psi}(t)$ 与受控对象(1.2)的任一可能的相速度之内积：

$$\tilde{\psi}(t)f(t,\tilde{x}(t),\tilde{u}(t))\geqslant\tilde{\psi}(t)p, \qquad \forall p\in P(t,\tilde{x}(t)).$$

因此，对几乎所有的 $t\in[t_1,t_2]$，集合 $P(t,\tilde{x}(t))$ 将位于 $\boldsymbol{R}^n$ 中由超平面

[注1] 在最优控制选自"无限维"函数族 Ω_U 这一意义上，最优问题 (1.3) 本质上是"无限维"的。以满足微分方程与边界条件：$\dot{x}=f(t,x,u)$，$t=t_1$，$x(t_1)=x_1$，$x(t_2)=x_2$ 的形式对控制 $u(t)\in\Omega_U$ 所加的"约束"，一般而言，并不能使满足这些约束的所有允许控制组成的函数族的"维数"有限化。

然而，考虑到控制最优性同最大值原理的联系，我们已将原来的"无限维"问题归结成为一个复杂的、但毕竟还是有限维的边值问题 (1.7) — (1.11)。

$$f(t,\tilde{x}(t),\tilde{u}(t))+\Pi_t=\{f(t,\tilde{x}(t),\tilde{u}(t))+p;p\in\Pi_t\}$$

所划分的两个闭半空间的一个之中。于是，这一超平面是集合 $P(t,\tilde{x}(t))$ 的支承超平面，而点 $f(t,\tilde{x}(t),\tilde{u}(t))$ 是 $P(t,\tilde{x}(t))$ 的一个边界点。

人们可以想象出极值解运动的"动态"图形：即对几乎所有的 $t\in[t_1,t_2]$，受控对象所选取的相速度 $f(t,\tilde{x}(t),\tilde{u}(t))$ 是使沿向量 $\tilde{\psi}(t)$ 的方向的速度尽可能地大。

按(1.10)，总有可能在终止时刻 $t=t_2$ 选取一个值 $\hat{u}\in U$ 使得相速度 $f(t_2,\tilde{x}(t_2),\hat{u})$ 在 $\tilde{\psi}(t_2)$ 方向上的投影为非负。这样的 $\hat{u}$ 总可取为使得上确界

$$\sup_{u\in U}\tilde{\psi}(t_2)f(t_2,\tilde{x}(t_2),u)=M(t_2,\tilde{x}(t_2),\tilde{\psi}(t_2))\geqslant 0.$$

被达到的那一点。倘若这一上确界不能达到，则我们只能保证内积 $\tilde{\psi}(t_2)f(t_2,\tilde{x}(t_2),u),u\in U$，的值任意接近上确界。

1.5 自治情形的最大值原理

如果最优问题(1.3)是自治的，即函数 $f(t,x,u)$ 不显含 t，则沿问题(1.3)的任一个极值解的 $M(x,\psi)$ 之值将是常数，又按(1.10)它必非负。这一事实往往有重要的应用。我们将以下述命题的形式叙述并证明。

命题 1.1 设函数组 $\tilde{u}(t),\tilde{x}(t),\tilde{\psi}(t),t\in[t_1,t_2]$，满足：

(i) 典则方程组

$$\dot{x}=\frac{\partial}{\partial\psi}H(x,\psi,u),$$

$$\dot{\psi}=-\frac{\partial}{\partial x}H(x,\psi,u),$$

其中，$H(x,\psi,u)=\psi f(x,u)$，

(ii) 对几乎所有的 $t\in[t_1,t_2]$ 的最大值条件

$$H(\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))=\sup_{u\in U}H(\tilde{x}(t),\tilde{\psi}(t),u)=M(\tilde{x}(t),\tilde{\psi}(t)),\tag{1.12}$$

以及

(iii) 不等式

$$M(\tilde{x}(t_2),\tilde{\psi}(t_2))\geqslant 0,$$

则 $M(\tilde{x}(t),\tilde{\psi}(t))$ 作为 t 的函数恒等于一个非负常数:

$$M(\tilde{x}(t),\tilde{\psi}(t))\equiv\text{const}\geqslant 0,\qquad t\in[t_1,t_2].$$

证 我们将利用函数 $M(\tilde{x}(t),\tilde{\psi}(t))$，$t\in[t_1,t_2]$ 的连续性。这将在第七章里证明。

按定义，控制 $\tilde{u}(t)$ 的取值集是有界的[注1]，即集合

$$N=\{\tilde{u}(t),t\in\boldsymbol{R}\}\subset U\subset\boldsymbol{R}^r$$

是有界的，从而闭包 $\overline{N}$ 为紧集。现按如下方式定义一个函数 $v(t)$，$t\in[t_1,t_2]$：如在 t 点等式 $H(\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))=M(\tilde{x}(t),\tilde{\psi}(t))$ 成立，则令 $v(t)=\tilde{u}(t)$，如若不成立，则 $v(t)$ 在 t 点取满足关系式

$$H(\tilde{x}(t),\tilde{\psi}(t),\hat{v})=\max_{u\in\overline{N}}H(\tilde{x}(t),\tilde{\psi}(t),u)$$

的任意一点 $\hat{v}\in\overline{N}$. 记

$$\hat{H}(t)=H(\tilde{x}(t),\tilde{\psi}(t),v(t)),\qquad t\in[t_1,t_2].$$

我们将证明，对任一点 $t\in[t_1,t_2]$，有

$$\hat{H}(t)=\sup_{u\in\overline{U}}H(\tilde{x}(t),\tilde{\psi}(t),u),\tag{1.13}$$

其中 $\overline{U}$ 是 $U\subset\boldsymbol{R}^r$ 的闭包。这一等式对满足下式的那些 t 是显然成立的，

[注1] $\tilde{u}(t)$ 是有界可测函数，见1.1允许控制定义。——译者注

$$\hat{H}(t)=M(\tilde{x}(t),\tilde{\psi}(t))=\sup_{u\in U}H(\tilde{x}(t),\tilde{\psi}(t),u)=\sup_{u\in\overline{U}}H(\tilde{x}(t),\tilde{\psi}(t),u).$$

如果上式不满足，即

$$\hat{H}(t)=\max_{u\in\overline{N}}H(\tilde{x}(t),\tilde{\psi}(t),u)<\sup_{u\in\overline{U}}H(\tilde{x}(t),\tilde{\psi}(t),u),$$

则存在一点 $u'\in U$，使得

$$\hat{H}(t)=\max_{u\in\overline{N}}H(\tilde{x}(t),\tilde{\psi}(t),u)<H(\tilde{x}(t),\tilde{\psi}(t),u').$$

由于 $\overline{N}$ 为紧集，并且 $\tilde{x}(t)$，$\tilde{\psi}(t)$ 和 $H(x,\psi,u)$ 是各自变元的连续函数，因此，对任意一点 $u\in\overline{N}$ 及充分接近于 t 的所有点 τ 成立如下的不等式，

$$H(\tilde{x}(\tau),\tilde{\psi}(\tau),u)<H(\tilde{x}(\tau),\tilde{\psi}(\tau),u'),$$

这同最大值条件(1.12)矛盾。

如果我们能证明 $\hat{H}(t)$ $(t\in[t_1,t_2])$ 是绝对连续的并且它的导数几乎处处为零，则关系式 $M(\tilde{x}(t),\tilde{\psi}(t))=\text{const}\geqslant 0$ 获证。事实上，前者表明 $\hat{H}(t)=\text{const}$，而按定义知道对几乎所有的 $t\in[t_1,t_2]$，连续函数 $M(\tilde{x}(t),\tilde{\psi}(t))=\hat{H}(t)$ 成立。

先证明 $\hat{H}(t)$ 的绝对连续性。记

$$\hat{H}(t',t'')=H(\tilde{x}(t'),\tilde{\psi}(t'),v(t'')).$$

由(1.13)得到

$$\hat{H}(t'',t')-\hat{H}(t')\leqslant\hat{H}(t'')-\hat{H}(t')\leqslant\hat{H}(t'')-\hat{H}(t',t''),$$
$$\forall t',t''\in[t_1,t_2].$$

函数 $\hat{H}(t,\tau)$ 关于它的第一个变元绝对连续（对固定的 τ）．因此，

$$\int_{t'}^{t''}\left(\frac{d\hat{H}(t,t')}{dt}\right)dt\leqslant\hat{H}(t'')-\hat{H}(t')\leqslant\int_{t'}^{t''}\left(\frac{d\hat{H}(t,t'')}{dt}\right)dt,$$
$$\forall t',t''\in[t_1,t_2],t'\leqslant t''.$$

由此得到

$$-\int_{t'}^{t''}\left|\frac{d\hat{H}(t,t')}{dt}\right|dt\leqslant\hat{H}(t'')-\hat{H}(t')\leqslant\int_{t'}^{t''}\left|\frac{d\hat{H}(t,t'')}{dt}\right|dt,\quad(1.14)$$

$$\forall t',t''\in[t_1,t_2],t'\leqslant t''.$$

由于在$[t_1,t_2]$上，函数$\tilde{u}(t)$与$v(t)$有界，因此对几乎所有的$t\in[t_1,t_2]$下式成立：

$$\left|\frac{d\hat{H}(t,\tau)}{dt}\right|=|-\tilde{\psi}(t)f_x(\tilde{x}(t),\tilde{u}(t))f(\tilde{x}(t),v(\tau))+$$

$$+\tilde{\psi}(t)f_x(\tilde{x}(t),v(\tau))f(\tilde{x}(t),\tilde{u}(t))|\leqslant C,$$

$$\forall t,\tau\in[t_1,t_2].$$

根据(1.14)，便得估计式

$$|\hat{H}(t'')-\hat{H}(t')|\leqslant C|t''-t'|,\quad\forall t',t''\in[t_1,t_2].$$

上式立即导出$\hat{H}(t)$的绝对连续性。

为证对几乎所有的$t\in[t_1,t_2]$成立等式

$$\frac{d\hat{H}(t)}{dt}=0,$$

我们先写出对几乎所有的$t\in[t_1,t_2]$成立的如下估计式：

$$\left|\frac{d\hat{H}(t,\tau)}{dt}\right|\leqslant|\tilde{\psi}(t)|\{|f_x(\tilde{x}(t),v(\tau))-f_x(\tilde{x}(t),\tilde{u}(t))|$$

$$\cdot|f(\tilde{x}(t),\tilde{u}(t))|+|f(\tilde{x}(t),\tilde{u}(t))-f(\tilde{x}(t),v(\tau))|$$

$$\cdot|f_x(\tilde{x}(t),\tilde{u}(t))|\}\leqslant D\{|f_x(\tilde{x}(t),\tilde{u}(t))-f_x(\tilde{x}(t),v(\tau))|$$

$$+|f(\tilde{x}(t),\tilde{u}(t))-f(\tilde{x}(t),v(\tau))|\}.$$

记

$$g(t,\tau)=|f_x(\tilde{x}(t),\tilde{u}(t))-f_x(\tilde{x}(t),v(\tau))|+|f(\tilde{x}(t),\tilde{u}(t))$$

$$-f(\tilde{x}(t),v(\tau))|.$$

上面的估计式与(1.14)结合而得

$$-D\int_{t'}^{t''} g(t,t')dt \leqslant \hat{H}(t'') - \hat{H}(t') \leqslant D\int_{t'}^{t''} g(t,t'')dt, \tag{1.15}$$

$$\forall t', t'' \in [t_1, t_2], \quad t' \leqslant t''.$$

设 $\theta \in (t_1, t_2)$ 是函数

$$\hat{H}(t),\ f_x(\tilde{x}(t), \tilde{u}(t)),\ f(\tilde{x}(t), \tilde{u}(t)),\ t \in [t_1, t_2], \tag{1.16}$$

的 Lebesgue 点[注1]，并设 $\tilde{u}(\theta) = v(\theta)$. 这些点在区间 (t_1, t_2) 中构成全测度集。

易见，在上述任意一点 θ，如下关系式成立：

$$\frac{1}{h}\int_{\theta}^{\theta+h} g(t,\theta)dt \longrightarrow 0, \quad \frac{1}{h}\int_{\theta-h}^{\theta} g(t,\theta)dt \longrightarrow 0, \quad (h \to 0).$$

实际上，因为 $\tilde{u}(\theta) = v(\theta)$，我们有

$$\begin{aligned} g(t,\theta) \leqslant & |f_x(\tilde{x}(t), \tilde{u}(t)) - f_x(\tilde{x}(\theta), \tilde{u}(\theta))| + |f(\tilde{x}(t), \tilde{u}(t)) \\ & - f(\tilde{x}(\theta), \tilde{u}(\theta))| + |f_x(\tilde{x}(t), \tilde{u}(\theta)) - f_x(\tilde{x}(\theta), \tilde{u}(\theta))| \\ & + |f(\tilde{x}(t), \tilde{u}(\theta)) - f(\tilde{x}(\theta), \tilde{u}(\theta))|. \end{aligned}$$

由于 θ 是(1.16)中各个函数的 Lebesgue 点，并且 $f_x(\tilde{x}(t), \tilde{u}(\theta))$ 及 $f(\tilde{x}(t), \tilde{u}(\theta))$ 作为 t 的函数对每一个固定的 θ 在 $t=\theta$ 处连续，因此得到如上两个收敛关系式. 再在(1.15)的左边一式中令 $t''=\theta+h$ 且 $t'=\theta$，除以 $h>0$，令 h 趋于零，就得到

$$0 \leqslant \frac{d\hat{H}(\theta)}{dt}.$$

类似地，由(1.15)右边一式产生

$$\frac{d\hat{H}(\theta)}{dt} \leqslant 0.$$

这表明，$d\hat{H}(t)/dt = 0$ 几乎处处成立.

[注1] 参见参考文献[7]第Ⅱ章§4. 此外还要求 θ 是 $\hat{H}(t)$ 的可微点（见下面的证明）. ——译者注.

1.6 开集U的情形．最优控制问题解的典则表述

如果在最优问题（1.3）中，函数$f(t,x,u)$关于x和u二次连续可微，又若控制的允许值集U是$\mathbf{R}^r$中的开集，则由最大值原理可以导出许多重要的结论。这些结论在古典变分法里是熟知的。

首先，由最大值条件可得，对几乎所有的$t\in[t_1,t_2]$，

$$\frac{\partial}{\partial u^i}H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))=0,\quad i=1,\cdots,r.$$

这是因为$H(t,\tilde{x}(t),\tilde{\psi}(t),u)$作为定义在一个开集$U\subset\mathbf{R}^r$上的$u$的函数，在点$\tilde{u}(t)$达到其最大值。我们将上面的方程组写成一个向量方程：

$$\frac{\partial}{\partial u}H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))=0.$$

由最大值条件还可推出$r\times r$矩阵

$$\left(\frac{\partial^2}{\partial u^i\partial u^j}H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))\right),\quad i,j=1,\cdots,r.$$

是非正定的，即二次型不等式

$$\sum_{i,j=1}^{r}\frac{\partial^2}{\partial u^i\partial u^j}H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))v^iv^j\leqslant0$$

对任意的$v^1,\cdots,v^r$值成立。如果该矩阵的行列式满足附加条件：对几乎所有的$t\in[t_1,t_2]$，

$$\left|\det\left(\frac{\partial^2}{\partial u^i\partial u^j}H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))\right)\right|\geqslant\alpha>0,$$

则相应的二次型是负定的：

$$\sum_{i,j=1}^{r}\frac{\partial^2}{\partial u^i\partial u^j}H(t,\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t))v^iv^j\leqslant-\beta|v|^2,\beta>0.\text{[注1]}$$

[注1] β应与t有关。——译者注。

在这一情形，我们称极值解 $\tilde{x}(t),\tilde{\psi}(t),\tilde{u}(t)$，$t\in[t_1,t_2]$，为最优问题(1.3)的一个正则极值解。

矩阵$\left(\frac{\partial^2 H}{\partial u^i \partial u^j}\right)$可看成是向量函数

$$\frac{\partial H}{\partial u}=\left(\frac{\partial H}{\partial u^1},\cdots,\frac{\partial H}{\partial u^r}\right)$$

的 Jacobi 矩阵。这样，由正则性条件

$$\left|\det\left(\frac{\partial^2 H}{\partial u^i \partial u^j}\right)\right|\geqslant a>0$$

与集合$\{\tilde{u}(t);\ t\in[t_1,\ t_2]\}$的有界性，存在一个数 $\varepsilon>0$，不依赖于 $t\in[t_1,t_2]$，使得对几乎所有的 t，关于 u 的方程

$$\frac{\partial H}{\partial u}(t,\tilde{x}(t),\tilde{\psi}(t),u)=0$$

在 $\tilde{u}(t)$ 的 ε-邻域中有唯一解。因此，这个解重合于 $\tilde{u}(t)$. 在$U\subset \boldsymbol{R}^r$是开集的情形，如果最优问题(1.3)的 Hamilton 函数

$$H(t,x,\psi,u)=\psi f(t,x,u)$$

满足条件：

$$\left|\det\left(\frac{\partial^2 H}{\partial u^i \partial u^j}(t,x,\psi,u)\right)\right|\geqslant a>0 \tag{1.17}$$

对任意满足 $|\psi|=1$ 的 ψ 以及位于空间 $\boldsymbol{R}^{1+n+r}$ 的一个有界集合中的一切(t,x,u)都成立（其中的常数 a 与该有界集合有关），我们就称最优问题(1.3)是正则的。

由方才所述可知，当 x 和 ψ 的初始条件给定时，正则最优问题(1.3)的极值解作为典则方程组

$$\dot{x}=\frac{\partial H(t,x,\psi,u)}{\partial \psi},\qquad \dot{\psi}=-\frac{\partial H(t,x,\psi,u)}{\partial x} \tag{1.18}$$

借助于方程

$$\frac{\partial H}{\partial u}(t,x,\psi,u)=0 \tag{1.19}$$

消去 u 所得的解是唯一确定的。在(1.19)任一个解的一个邻域中唯一地确定了 u 为变元 t,x,ψ 的可微函数[注1]

$$u=u(t,x,\psi).$$

消去典则方程组(1.18)中的 u，就把问题(1.3)的"原有变量" x 和 u 变换成为典则共轭变量 x 和 ψ。向量 ψ 是在相差一个常数因子的范围内由(1.19)确定的，对于正则问题（即条件（1.17）满足），从 $\{x,u\}$ 到 $\{x,\lambda\psi\}$ 的变换可以实现。它被称为由函数 $f(t,x,u)$ 派生的 Legendre 变换。

从这一观点看，庞特里雅金最大值条件(1.9)是 Legendre 变换在控制允许值集 U 为任意、且函数 $f(t,x,u)$ 未必对 u 可微之情形的推广。

在1.4节已给出这一推广的 Legendre 变换的几何意义。就是说，为求得对应于给定的 $\{t,x,\psi\}$ 的 u 值，可使 R^n 中正交于 ψ 的超平面 Γ^{n-1} 沿向量 ψ 的方向移动，直至它最后与集合

$$P(t,x)=\{f(t,x,u):u\in U\}$$

相交于点 $\hat{f}=f(t,x,\hat{u})$（假定是唯一的），这个值 $u=\hat{u}$ 就是所求之值（我们假定方程 $\hat{f}=f(t,x,u)$ 是在 $u\in U$ 中唯一可解的）。

本章中所述的利用最大值原理寻求最优解的方法称为典则组法。这一方法可简单地概括为：

构作问题（1.3）的 Hamilton 函数 $H=\psi f(t,x,u)$。利用这一函数写出典则组(1.7)。借助于庞特里雅金最大值条件(1.9)从中消去 u 而得典则组的解。那些同时满足不等式(1.10)和边界条件(1.11)的解组成了极值解集。所要求的最优解含于这些极值解之中。

在许多重要情形，问题（1.3）存在唯一的极值解。这时只要最优解也存在的话，那么极值解必定是最优解。

最优解的存在性问题将在第八章里讨论。

[注1] 关于 t 仅连续。——译者注。

1.7 本章结语

最大值原理是一阶必要条件。它是作为研究某一个唯一地对应于原最优问题的凸控制问题的最优轨线的（一阶）变分锥的结果获得的。

凸控制问题在控制理论中起着重要的作用。为了精确地定义它们，方便的途径是引进广义控制的概念。第二和第三章将用于定义广义控制与凸控制问题，并研究它们的基本性质。

当我们变动受控方程的轨线并构造这些轨线的变分锥时，广义控制较之常义控制的主要优点就变得明显了。相应的构造方法将在第六章阐述。在第七章里，我们基于第六章的结果给出最大值原理的证明。

广义控制优越于常义控制的另一特点是它们的弱紧性。这一重要性质将在第八章里研究，并被用于证明关于最优解的存在性的一些基本事实。

最后，第四与第五章的内容是辅助性的。在这两章里，叙述并证明了在本书中要用到的关于常微分方程解的存在性和对初值及右端函数的连续依赖性的几个一般定理。

第二章 广义控制

在本章中，通过引进广义控制，扩张了允许控制类。当我们研究控制问题中的变分方法并在此基础上证明最大值原理（第六、七章）以及证明最优解的存在性（第八章）时，这一扩张的优越性将显示出来。

这里，我们要介绍广义控制的基本性质，并特别关心广义控制弱收敛的概念。

2.1 广义控制与一个凸控制问题

我们先给出两个定义：设 μ_t 是依赖于参数 $t\in\boldsymbol{R}$ 的 $\boldsymbol{R}^r$ 上的一族 Radon 测度。设 $g(t,u)$ 是变元 $t\in\boldsymbol{R}$ 与 $u\in\boldsymbol{R}^r$ 的（数值或向量值）连续函数，且对每一固定的 $t\in\boldsymbol{R}$，它关于 u 有紧支集（该支集可与 t 有关）。如果对一切具有这种性质的 $g(t,u)$，积分

$$h(t)=\int_{R^r}g(t,u)d\mu_t(u)=\int_{R^r}g(t,u)d\mu_t,\quad t\in\boldsymbol{R},$$

都是 t 的 Lebesgue 可测函数，则称这一族 $\mu_t(t\in\boldsymbol{R})$是弱可测的。

如果存在一个不依赖于 $t\in\boldsymbol{R}$ 的紧集 $K\subset\boldsymbol{R}^r$，使得对几乎所有的 $t\in\boldsymbol{R}$（在 Lebesgue 测度的意义之下），μ_t 的测度集中于K 上，则这族 $\mu_t(t\in\boldsymbol{R})$称为是有限的。

我们将连续函数 $g(t,u)$ 关于测度 μ_t 的积分记为

$$\langle\mu_t,g(t,u)\rangle=\int_{R^r}g(t,u)d\mu_t.$$

一个允许控制 $u(t)\in\Omega_U$ 可以对应 $\boldsymbol{R}^r$ 上一族依赖于时间 $t\in\boldsymbol{R}$ 的 Dirac 测度(Dirac 测度是集中在一点的单位正测度).事实上,控制在时刻 t 的取值 $u(t)$ 对应于集中在 $u(t)\in U$ 这一点的单位正测度 $\delta_{u(t)}$,它作用于任一连续函数 $g(t,u)$ 上便得

$$\langle\delta_{u(t)},g(t,u)\rangle=\int_{R^r}g(t,u)d\delta_{u(t)}=g(t,u(t)).$$

由于我们仅考虑取值于 U 中的有界控制,集合

$$N=\{u(t);t\in\boldsymbol{R}\}\subset U\subset\boldsymbol{R}^r$$

是有界的,因而所有测度 $\delta_{u(t)}$ $(t\in\mathbf{R})$ 都集中于 N 上。换言之,测度族 $\delta_{u(t)}$ 是有限的。

由于函数 $u(t)\in\Omega_U$ 是可测的,则对任一连续函数 $g(t,u)$,

$$h(t)=\langle\delta_{u(t)},g(t,u)\rangle=\int_{R^r}g(t,u)d\delta_{u(t)}=g(t,u(t))$$

是可测函数,即测度族 $\delta_{u(t)}$ 弱可测。这样便知,对任一允许控制 $u(t)\in\Omega_U$,对应的 Dirac 测度族 $\delta_{u(t)}$ 是有限的和弱可测的。

反之,假设我们有任意一族弱可测且有限的 Dirac 测度 $\delta_{v(t)},t\in\boldsymbol{R}$,其在时刻 t 的测度 $\delta_{v(t)}$ 集中于点 $v(t)\in U$。那么,直接由前面所给的定义可知,函数 $v(t)(t\in\boldsymbol{R})$ 是本性有界的。令 $g(t,u)=u$,即得可测函数

$$\langle\delta_{v(t)},u\rangle=v(t)\in\Omega_U.$$

这样,我们建立了允许控制 $u(t)\in\Omega_U$ 与测度集中在集合 U 弱可测且有限的 Dirac 测度族 $\delta_{u(t)}(t\in\boldsymbol{R})$ 之间的自然对应。

现给出如下基本定义:

任意一族弱可测的且有限的概率测度(即测度集中于 $U\subset\mathbf{R}^r$ 上的单位正 Radon 测度族 $\mu_t(t\in\boldsymbol{R})$)可称为广义控制。

我们以 $\mathfrak{M}_U$ 记所有广义控制的集合,称之为广义控制类。此后,$\mu_t(t\in\boldsymbol{R})$ 将表示一个广义控制。

在广义控制的定义中取概率测度而不是任一 Radon 测度的原因是,只有概率测度族具有适用于控制问题的性质,这在逼近引理(见第三章)中将有所表述。

$g(t,u)$ 关于测度 μ_t 积分所得的函数

$$\langle \mu_t, g(t,u) \rangle = \int_{R^r} g(t,u) d\mu_t, \ t \in \boldsymbol{R}.$$

称为广义控制 μ_t 向 $g(t,u)$ 中参数 u 的代入。这与通常的术语是一致的，因为函数 $u(t)$ 代入 $g(t, u)$ 所得结果与对应的一族 Dirac 测度 $\delta_{u(t)}(t \in \boldsymbol{R})$ 代入的结果相同。今后，我们将 $u(t) \in \Omega_U$ 与 $\delta_{u(t)}(t \in \boldsymbol{R})$ 等同看待，即设 $\Omega_U \subset \mathfrak{M}_U$。

我们现在回到受控方程

$$\dot{x} = f(t,x,u).$$

广义控制 μ_t 代入这一方程右端函数中的 u，即得微分方程

$$\dot{x} = \langle \mu_t, f(t,x,u) \rangle = \int_{R^r} f(t,x,u) d\mu_t = F(t,x).$$

这与方程(1.2)类似。如果初始条件 $x(t_1) = x_1$ 给定，则此方程等价于积分方程

$$x(t) = x_1 + \int_{t_1}^{t} \langle \mu_\theta, f(\theta, x(\theta), u) \rangle d\theta = x_1 + \int_{t_1}^{t} F(\theta, x(\theta)) d\theta.$$

它在 t_1 点的一个邻域中有唯一确定的解（见第四章）。现在，我们不仅对常义控制而且对广义控制的如下形式

$$\dot{x} = \langle \mu_t, f(t,x,u) \rangle, \quad \mu_t \in \mathfrak{M}_U,$$

提出下述边值问题：

寻求一个广义控制 μ_t，使得方程

$$\dot{x} = \langle \mu_t, f(t,x,u) \rangle$$

的解满足边界条件

$$x(t_1) = x_1, \qquad x(t_2) = x_2.$$

进一步，若对此边值问题添加条件

$$t_2 - t_1 \longrightarrow \min,$$

则就得关于广义控制类的一个最优问题。

控制问题

$$\dot{x} = \langle \mu_t, f(t,x,u) \rangle, \qquad \mu_t \in \mathfrak{M}_U, \tag{2.1}$$

称为原控制问题

$$\dot{x} = f(t,x,u), \qquad u(t) \in \Omega_U, \tag{2.2}$$

的凸化，或称为后者所对应的凸控制问题。

相应地，最优问题

$$\dot{x}=\langle\mu_t,f(t,x,u)\rangle,\quad \mu_t\in\mathfrak{M}_U$$
$$t=t_1,\quad x(t_1)=x_1,\quad x(t_2)=x_2, \tag{2.3}$$
$$t_2-t_1\longrightarrow\min$$

将称为原最优问题(1.3)所对应的凸最优问题。

因为 Ω_U 包含在 $\mathfrak{M}_U$ 之中，广义控制类 $\mathfrak{M}_U$ 与方程(2.1)所产生的受控对象比常义控制类 Ω_U 与 (2.2) 的受控对象具有更广的控制可能性。一般而言，原最优问题(1.3)可能无解，而凸最优问题(2.3)倒是有解的。我们将在研究最优解存在性的第八章里探讨这一问题.现在，我们要指出拓广为凸控制问题的几何意义，它将解释这一名称的来源。

首先，所有广义控制组成的集合 $\mathfrak{M}_U$ 是凸的。若 μ_t 与 ν_t 是两个广义控制，则对任意常数 α 与 β，测度族 $\alpha\mu_t+\beta\nu_t$ 是弱可测的，有限的，对任一 t 也是集中在 U 上的，因为 μ_t 与 ν_t 本身即是如此。如常数 α 与 β 还满足条件 $\alpha\geqslant0,\beta\geqslant0$ 及 $\alpha+\beta=1$，则测度 $\alpha\mu_t+\beta\nu_t$ 对每一 t 是概率测度。

其次，凸控制问题 (2.1) 的方程右端所成的集合也是凸的。我们来解释这一论断。

原控制问题(2.2)的右端集合可写成如下形式，

$$\{F(t,x):F(t,x)=f(t,x,u(t)),u(t)\in\Omega_U\}.$$

这是依赖于允许控制取值集合 U 的一个复杂的函数族。在所有关于 t 可测且关于 x 连续可微的 n 维向量值函数 $F(t,x)$ 组成的线性空间中，这族函数未必是凸的。但对应的凸问题的右端集合具有如下形式，

$$\{F(t,x):F(t,x)=\langle\mu_t,f(t,x,u)\rangle,\ \mu_t\in\mathfrak{M}_U\}. \tag{2.4}$$

不管集合 U 如何，这一右端集合总是凸的。事实上，若

$$F_1(t,x)=\langle\mu_t,f(t,x,u)\rangle,\qquad F_2(t,x)=\langle\nu_t,f(t,x,u)\rangle$$

属于(2.4)所表示的集合，则函数

$$\alpha F_1(t,x)+\beta F_2(t,x)=\langle\alpha\mu_t+\beta\nu_t,f(t,x,u)\rangle,\alpha,\beta\geqslant0,\alpha+\beta=1,$$

也必属于这一集合，这是因为$\mathfrak{M}_U$是凸集。特别，当t和x固定时，凸问题(2.1)的所有可能的相速度的集合也是$\boldsymbol{R}^n$中的凸集，我们记它为$P_{\mathfrak{M}}(t,x)$. 易见

$$P_{\mathfrak{M}}(t,x)=\{p:p=\langle\mu,f(t,x,u)\rangle\},$$

其中，μ是测度集中于U的一个有界部分上的任一概率测度，因为测度族$\mu_t\equiv\mu(t\in\boldsymbol{R})$也是广义控制。

命题 2.1 对任意的t和x，凸控制问题(2.1)的所有可能的相速度之集合$P_{\mathfrak{M}}(t,x)$与原控制问题(2.2)的所有可能的相速度（对所考虑的t和x值）之集合$P(t,x)$的凸包重合，

$$P_{\mathfrak{M}}(t,x)=\operatorname{conv}P(t,x)=\operatorname{conv}\{p:p=f(t,x,u),u\in U\},$$

其中，conv{·}表示括号内集合的凸包。

证 因为集合$P_{\mathfrak{M}}(t,x)$是凸的，并且

$$P_{\mathfrak{M}}(t,x)\supset P(t,x),$$

因此，只要证明相反的包含关系即可：

$$P_{\mathfrak{M}}(t,x)\subset\operatorname{conv}P(t,x).$$

设有一点$\hat{p}\notin\operatorname{conv}P(t,x)$，但$\hat{p}\in P_{\mathfrak{M}}(t,x)$，即存在集中于$U$上的一个概率测度$\hat{\mu}$，使得

$$\langle\hat{\mu},f(t,x,u)\rangle=\hat{p}.$$

我们在$\boldsymbol{R}^n$中过$\hat{p}$点作一个$(n-1)$维超平面Γ^{n-1}，使得凸集$\operatorname{conv}P(t,x)$位于它的一侧。记超平面$\Gamma^{n-1}$在映射

$$f(t,x,\cdot):\boldsymbol{R}^r\longrightarrow\boldsymbol{R}^n \tag{2.5}$$

下的逆象的特征函数为$\chi_1(u)$，而该逆象在$\boldsymbol{R}^r$中的补集的特征函数为$\chi_2(u)$：

$$\chi_1(u)=\begin{cases}1, 如\ f(t,x,u)\in\Gamma^{n-1},\\0, 如\ f(t,x,u)\notin\Gamma^{n-1},\end{cases}$$

$$\chi_2(u)=1-\chi_1(u).$$

由于$\hat{\mu}$是概率测度，因此

$$\hat{p} = \langle \hat{\mu}, \hat{p} \rangle,$$

上式等号右边出现的 $\hat{p}$ 意为在 U 上取常值 $\hat{p}$ 的函数。从而，

$$\langle \hat{\mu}, f(t,x,u) - \hat{p} \rangle = \langle \hat{\mu}, \chi_1(u)[f(t,x,u) - \hat{p}] \rangle + \langle \hat{\mu}, \chi_2(u)[f(t,x,u) - \hat{p}] \rangle = 0.$$

用 ξ 表示 Γ^{n-1} 的指向集合 conv $P(t,x)$ 所在一侧的法向量。将上式与 ξ 作内积便得

$$\langle \hat{\mu}, \chi_1(u)\xi[f(t,x,u) - \hat{p}] \rangle + \langle \hat{\mu}, \chi_2(u)\xi\cdot[f(t,x,u) - \hat{p}] \rangle = 0.$$

由于使得 $\chi_1(u) \neq 0$ 的 u 必使 $\xi\cdot[f(t,x,u) - \hat{p}] = 0$，故上式左边的第一项为零。因此，

$$\langle \hat{\mu}, \chi_2(u)\xi\cdot[f(t,x,u) - \hat{p}] \rangle = 0.$$

但测度 $\hat{\mu}$ 是正测度，且当 $u \in U$ 使 $f(t,x,u) \notin \Gamma^{n-1}$ 时，有

$$\chi_2(u)\xi\cdot[f(t,x,u) - \hat{p}] > 0,$$

由此可得出这样的结论：测度 $\hat{\mu}$ 集中在 Γ^{n-1} 关于映射 (2.5) 的逆象之上。

现将 $\boldsymbol{R}^n$ 改为 Γ^{n-1}，将 conv $P(t,x)$ 改为 $\Gamma^{n-1} \cap$ conv $P(t,x)$[注1]，由同样的论证，得到测度 $\hat{\mu}$ 集中在一个 $(n-2)$ 维超平面 $\Gamma^{n-2} \subset \Gamma^{n-1} \subset \boldsymbol{R}^n$ 关于映射(2.5)的逆象之上，这里 $\hat{p} \in \Gamma^{n-2}$。以这种方式逐次降维，最后发现测度 $\hat{\mu}$ 是集中在点 $\hat{p}$ 关于映射 (2.5)的逆象之上。然而，这同假设正测度 $\hat{\mu}$ 集中在 U 上并且

$$\hat{p} \notin \text{conv } P(t,x)$$

矛盾，因为上式表明 U 与 $\hat{p}$ 按映射(2.5)的逆象不相交。证毕。

[注1] 从上可知 $\hat{\mu}$ 集中在 Γ^{n-1} 关于映射(2.5)的逆象与 U 的交集中，但 $\hat{\mu}$ 是正测度，故该交集不空，从而 $\Gamma^{n-1} \cap$ conv $P(t, x)$ 也不空。——译者注。

下面给出一个广义控制的典型例子（参见第三章的逼近引理。）通过第八章的存在性定理的证明，我们就会体会到它的重要性。这一控制称为颤振控制，或滑动制式控制（可参见第三章末尾关于这一术语的注）。

设 $u_1(t),\cdots,u_q(t)$ 是任意有限个取自 Ω_U 的允许控制，且设 $\mu_1(t),\cdots,\mu_q(t)\ (t\in\boldsymbol{R})$ 是同样多个实值 Lebesgue 可测函数，并满足条件：对几乎所有的 t，

$$\mu_i(t)\geqslant 0 \quad 且 \quad \sum_{i=1}^{q}\mu_i(t)=1,$$

这些函数确定的广义控制

$$\mu_t=\sum_{i=1}^{q}\mu_i(t)\delta_{u_i(t)},\quad t\in\boldsymbol{R},\tag{2.6}$$

称为颤振控制。在每一瞬时，μ_t 的测度集中在点 $u_1(t),\cdots,u_q(t)$ 上并带有相应的"加权系数" $\mu_1(t),\cdots,\mu_q(t)$。可以说，控制 μ_t 以 $\mu_1(t),\cdots,\mu_q(t)$ 的比例在 q 个点 $u_1(t),\cdots,u_q(t)$ 之间"分布"。当 $q=1$ 时，我们得到的是一个常义控制

$$\mu_t=\delta_{u_1(t)}\in\Omega_U.$$

利用颤振控制可将受控方程写成如下形式，

$$\dot{x}=\left\langle\sum_{i=1}^{q}\mu_i(t)\delta_{u_i(t)},f(t,x,u)\right\rangle=\sum_{i=1}^{q}\mu_i(t)f(t,x,u_i(t)).$$

2.2 广义控制的弱收敛

在定义于$\boldsymbol{R}^m$（其中 $m\geqslant 1$）上的广义 Radon 测度全体构成的线性空间中已有弱收敛的概念。现在我们回顾一下这个概念，并用它来定义广义控制类中的弱收敛。

在$\boldsymbol{R}^m$ 上给定的一列（广义）Radon 测度 $\nu^{(i)}$, $i=1,2,\cdots$，称为当 $i\to\infty$ 时弱收敛于测度 ν，如果对任一在$\boldsymbol{R}^m$ 上有紧支集的（数值或向量值）连续函数 $g(z)$，序列

$$\langle\nu^{(i)},g(z)\rangle=\int_{\boldsymbol{R}^m}g(z)d\nu^{(i)}(z)$$

当 $i\to\infty$ 时收敛于$\langle \nu, g(z)\rangle$，即

$$\langle \nu^{(i)}, g(z)\rangle \longrightarrow \langle \nu, g(z)\rangle,\ (i\to\infty)\text{[注1]}.$$

对一固定的 t，任一广义控制 μ_t 是$\boldsymbol{R}^r$上的一个概率测度。如果我们定义测度$\nu=\{\mu_t; t\in\boldsymbol{R}\}$在具有紧支集的连续函数$g(t,u)$上的作用为

$$\langle \nu, g(t,u)\rangle = \int_R dt\int_{R^r} g(t,u)d\mu_t(u) = \int_R \langle \mu_t, g(t,u)\rangle dt,$$

其中最后的积分是 Lebesgue 积分，“内积” $\langle \mu_t, g(t,u)\rangle$如前所述是对固定的 t 关于测度 μ_t 的积分，则 $\nu=\{\mu_t; t\in\boldsymbol{R}\}$ 可视为在点(t,u)所属的$(1+r)$维空间 $\boldsymbol{R}^{1+r}$ 上的一个 Radon 测度。

如果一列广义控制 $\mu_t^{(i)}$ 满足

$$\int_R \langle \mu_t^{(i)}, g(t,u)\rangle dt \longrightarrow \int_R \langle \mu_t, g(t,\mu)\rangle dt,\ (i\to\infty), \tag{2.7}$$

其中，μ_t 是广义控制，$g(t,u)$是任一具有紧支集的连续函数，就称广义控制 $\mu_t^{(i)}$ 弱收敛于广义控制 $\mu_t(i\to\infty)$.

应注意下面的注记。设 $v_i(t)(t\in[t_1,t_2])$ 是一列 r 维 Lebesgue 可积函数，按熟知的弱收敛定义，称序列 $v_i(t)$ 弱收敛于 $v(t)(t\in[t_1,t_2])$ 是指对任意一个 r 维连续的行向量函数 $\varphi(t), t\in[t_1,t_2]$，成立

$$\int_{t_1}^{t_2}\varphi(t)v_i(t)dt \longrightarrow \int_{t_1}^{t_2}\varphi(t)v(t)dt,\quad (i\to\infty). \tag{2.8}$$

如果考虑一列常义控制

$$\mu_t^{(i)} = \delta_{u_i(t)},\quad t\in\boldsymbol{R},$$

其中假定函数 $u_i(t), t\in\boldsymbol{R}$，是对 i 一致有界的，则我们既可讨论 $u_i(t)$ 在定义(2.7)意义下的收敛极限，即广义控制 $\delta_{u_i(t)}$ 的弱收敛，也可以讨论这同一列 $u_i(t)$ （在任一段区间$[t_1, t_2]$上）按定义(2.8)意义下的收敛性。但这是两种完全不同的收敛概念，下面的典型例子中将清楚地显示出这一点。

为简单计，考虑 $r=1$ 的标量情形。将 t 轴划分成等长度的区间，区间长度随划分的次数 $i\to\infty$而趋于零。对第 i 次划分定义函数

[注1] 如果在 $\boldsymbol{R}^n$上的所有广义 Radon 测度的线性空间中借助一族半范 $\|\nu\|_g^* = |\langle \nu, g(z)\rangle|$（其中$g(z)$是在$\boldsymbol{R}^n$中具有紧支集的连续函数）引进局部凸拓扑，则弱收敛显然按这一拓扑收敛，可以证明，这样定义的拓扑不具有可列的原点邻域基，因此，它是不可度量化的。

$u_i(t)$ $(t\in \boldsymbol{R})$如下：该函数在每个划分区间上依次交替地取值 1 或 -1. 这样，对于较大的 i,$u_i(t)$是一个在 $+1$ 与 -1 之间"快速振荡"的函数。容易看出，对任一连续函数 $\varphi(t)$ $(t\in[t_1,t_2])$，有

$$\int_{t_1}^{t_2}\varphi(t)u_i(t)dt\longrightarrow 0\quad (i\to\infty)\text{[注1]},$$

即在(2.8)的意义下 $u_i(t)\longrightarrow 0$ $(i\to\infty)$.

但是，序列 $\delta_{u_i(t)}$ 在(2.7)的意义下却根本不收敛于零。事实上，稍后可知，对任一连续函数 $g(t,u)$，有极限：

$$\begin{aligned}\int_{t_1}^{t_2}\langle\delta_{u_i(t)},g(t,u)\rangle dt&=\int_{t_1}^{t_2}g(t,u_i(t))dt\\ &\longrightarrow\frac{1}{2}\int_{t_1}^{t_2}g(t,1)dt+\frac{1}{2}\int_{t_1}^{t_2}g(t,-1)dt\quad(i\to\infty).\end{aligned}\tag{2.9}$$

因此，若 $g(t,u)$在 $\boldsymbol{R}^{1+r}$ 中有紧支集，

$$\begin{aligned}\int_R\langle\delta_{u_i(t)},g(t,u)\rangle dt&\longrightarrow\frac{1}{2}\int_R g(t,1)dt+\frac{1}{2}\int_R g(t,-1)dt\\ &=\int_R\langle\mu_t,g(t,u)\rangle dt\quad(i\to\infty),\end{aligned}\tag{2.10}$$

其中，

$$\mu_t=\frac{1}{2}\delta_1+\frac{1}{2}\delta_{-1}.$$

这里暂不谈(2.9)式的证明，因为它可由更为一般的关系式直接导出，而后者将在下一章证明逼近引理中得到。

关系式(2.10)指出，广义控制 $\delta_{u_i(t)}$ 按(2.7)的意义弱收敛于一个广义控制：

$$\delta_{u_i(t)}\longrightarrow\frac{1}{2}\delta_1+\frac{1}{2}\delta_{-1}\quad(i\to\infty).$$

序标i越大，控制 $u_i(t)$ "颤振"于 1 与 -1 值之间的频率也越大，这一

[注1] 当$\varphi(t)$是连续可微函数时容易得到这一关系式：

$$\int_{t_1}^{t_2}\varphi(t)u_i(t)dt=\left\{\varphi(t_2)\int_{t_1}^{t_2}u_i(t)dt-\int_{t_1}^{t_2}\varphi'(t)\int_{t_1}^{t_2}(u_i(\theta)d\theta)dt\right\}\to 0,\quad(i\to\infty).$$

利用连续可微函数以任意精度逼近一连续函数（如用多项式逼近的Weierstrass定理），即得所述的结果。

控制在弱收敛(2.7)的意义下越"接近"于颤振控制 $\frac{1}{2}\delta_1+\frac{1}{2}\delta_{-1}$，后者是以等加权系数 $\frac{1}{2}$ 在两点 1 与 -1 之间分布。这个例子解释了我们为(2.6)形式的控制所引进的名称（也可参见第三章之末）。广义控制的弱收敛问题在第八章里还要更详细地研究．用于最优解的存在的讨论。

现在来讨论广义 Radon 测度的强收敛概念．在$\boldsymbol{R}^m$上具有紧支集的数值连续函数 $g(z)$ 全体组成的赋范空间记为$C^0(\boldsymbol{R}^m)$，范数定义为

$$\|g(z)\|=\sup_{z\in\boldsymbol{R}^m}|g(z)|.$$

显然，按这一范数的收敛相当于一致收敛。

$\boldsymbol{R}^m$ 上的广义 Radon 测度 ν 的范数或全变差 $\|\nu\|$ 定义为：

$$\|\nu\|=\sup\{\langle\nu,g(z)\rangle;\ \|g(z)\|\leqslant 1,g(z)\in C^0(\boldsymbol{R}^m)\}.$$

显然，概率测度的线性组合或有紧支集的测度都有有限的范数。

不难看出，如果 $\nu_t(t\in\boldsymbol{R})$是$\boldsymbol{R}^m$ 上的一族弱可测的 Radon 测度，则函数

$$h(t)=\|\nu_t\|,\quad t\in\boldsymbol{R},$$

是 Lebesgue 可测的。在第三章里将证明（见命题 3.3）：在$C^0(\boldsymbol{R}^m)$中存在函数列 $g_1(z),g_2(z),\cdots$，满足如下条件：

(1) $\|g_i(z)\|\leqslant 1,i=1,2,\cdots$；

(2) 对范数$\|g(z)\|\leqslant 1$ 的每一个 $g(z)\in C^0(\boldsymbol{R}^m)$，可取到一个子列 $g_{i_k}(z)$，使得

$$\|g_{i_k}(z)-g(z)\|\longrightarrow 0\quad(k\to\infty),$$

其中所有函数 $g_{i_k}(z)$的支集位于$\boldsymbol{R}^m$ 的一个与下标无关的有界集之中．

由于所有函数 $g(z)$ 与 $g_{i_k}(z)$的支集都在$\boldsymbol{R}^m$ 的一个有界集内，我们得到

$$\langle\nu_t,g_{i_k}(z)-g(z)\rangle\longrightarrow 0\quad(k\to\infty),\quad\forall t\in\boldsymbol{R}.$$

因此，

$$\langle\nu_t,g(z)\rangle\leqslant\sup_i\langle\nu_t,g_i(z)\rangle,\quad\forall t\in\boldsymbol{R}.$$

从而，根据$\|g(z)\|\leqslant 1$，$\|g_i(z)\|\leqslant 1$ 及 $g(z)\in C^0(\boldsymbol{R}^m)$的任意性，可得

$$\|\nu_t\|\leqslant\sup_i\langle\nu_t,g_i(z)\rangle\leqslant\|\nu_t\|,\ t\in\boldsymbol{R},$$

即

$$h(t)=\|\nu_t\|=\sup_i\langle\nu_t,g_i(z)\rangle,\quad \forall t\in\boldsymbol{R}.$$

因此，作为可列个可测函数

$$h_i(t)=\langle\nu_t,\ g_i(z)\rangle,\quad t\in\boldsymbol{R},$$

的上确界的函数$\|\nu_t\|$也可测。

我们称在$\boldsymbol{R}^m$上给定的一列 Radon 测度$\nu^{(i)}(i=1,2,\cdots)$，当$i\to\infty$时强收敛于测度ν，如果

$$\|\nu^{(i)}-\nu\|\longrightarrow 0\quad(i\to\infty).$$

称广义控制列$\mu_t^{(i)}(t\in\boldsymbol{R})$强收敛于广义控制$\mu_t(t\in\boldsymbol{R})$，如果

$$\int_R\|\mu_t^{(i)}-\mu_t\|dt\longrightarrow 0\quad(i\to\infty),$$

其中积分号下的范数表示 Radon 测度的范数。

很明显，广义控制列$\mu_t^{(i)}$强收敛于一个广义控制蕴涵弱收敛。但其逆不真。

这里给出一个广义控制列弱收敛但非强收敛的简单例子：设一列点$u_i\in\boldsymbol{R}^r$收敛于一点$\hat{u}$，其中$u_i\neq\hat{u}$。我们定义

$$\mu_t^{(i)}=\delta_{u_i},\quad 且\quad \mu_t=\delta_{\hat{u}},\quad \forall t\in\boldsymbol{R}.$$

我们有

$$\langle\delta_{u_i},g(t,u)\rangle=g(t,u_i)\longrightarrow g(t,\hat{u})=\langle\delta_{\hat{u}},g(t,u)\rangle.$$

因此，

$$\int_R\langle\delta_{u_i},g(t,u)\rangle dt\longrightarrow\int_R\langle\delta_{\hat{u}},g(t,u)\rangle dt\quad i\to\infty.$$

即当$i\to\infty$时，$\delta_{u_i}\to\delta_{\hat{u}}$弱收敛。但是，因为$u_i\neq\hat{u}$蕴涵$\|\delta_{u_i}-\delta_{\hat{u}}\|=2$，所以不是强收敛。

最后，我们要证明在第六章里将要用到的一个命题。

命题 2.2 设$\mu_t^{(i)}(\sigma)$，$i=1,2,\cdots$，是一列带参数$\sigma\in\Sigma$的广义控制。又设该序列当$i\to\infty$时关于$\sigma\in\Sigma$一致地弱收敛于一个广义控制$\mu_t(\sigma)(\sigma\in\Sigma)$。此外，设

$$\mu_t(\sigma)\quad 与\quad \mu_t^{(i)}(\sigma),\quad t\in\boldsymbol{R},\quad \sigma\in\Sigma,\quad i=1,2,\cdots,$$

的测度都集中在一个固定有界集

$$N \subset U \subset \boldsymbol{R}^r$$

之中，那么，对任一连续函数 $F(t,u)$，($(t,u)\in \boldsymbol{R}\times\boldsymbol{R}^r$)及任意的数 t' 与 t''，关于 $\sigma\in\Sigma$ 一致地有

$$\int_{t'}^{t''}\langle\mu_t^{(i)}(\sigma)-\mu_t(\sigma),F(t,u)\rangle dt\longrightarrow 0\ (i\to\infty).\tag{2.11}$$

证 在$\boldsymbol{R}^r$中，取以原点为中心、包含集合N的一个闭球S. 设$\beta(u)$是定义在$\boldsymbol{R}^r$上、在S上恒等于1、取值在0与1之间的具紧支集的连续函数. 事实上，$\beta(u)$可以这样来作出：取一个与S同心的半径更大的球S'，令

$$\beta(u)=\begin{cases}1, & \forall u\in S;\\ 0, & \forall u\notin S'.\end{cases}$$

在S与S'之间的环层中，令函数$\beta(u)$沿径向线性地从1递减为0.

类似地，令$\alpha(t)$ ($t\in\boldsymbol{R}$) 为这样一个连续函数：它在区间$[t',t'']$ ($t'\leqslant t''$)上为1，在某一区间$[t'-\eta,t''+\eta]$之外为0，这里$0<\eta\leqslant 1$，并且在区间$[t'-\eta,t']$与$[t'',t''+\eta]$上是线性的。

由于$\mu_t^{(i)}(\sigma)$与$\mu_t(\sigma)$的测度都集中在集合N上，而$u\in N$时函数$\beta(u)=1$，故

$$\begin{aligned}h^{(i)}(t;\sigma)&=\langle\mu_t^{(i)}(\sigma)-\mu_t(\sigma),\alpha(t)\beta(u)F(t,u)\rangle\\&=\alpha(t)\int_N\beta(u)F(t,u)d(\mu_t^{(i)}(\sigma)-\mu_t(\sigma))\\&=\alpha(t)\int_N F(t,u)d(\mu_t^{(i)}(\sigma)-\mu_t(\sigma))\\&=\langle\mu_t^{(i)}(\sigma)-\mu_t(\sigma),\alpha(t)F(t,u)\rangle.\end{aligned}$$

但函数$\alpha(t)\beta(u)F(t,u)$有紧支集，因此，按假定知

$$\begin{aligned}&\int_R\langle\mu_t^{(i)}(\sigma)-\mu_t(\sigma),\alpha(t)\beta(u)F(t,u)\rangle dt\\&=\int_R\langle\mu_t^{(i)}(\sigma)-\mu_t(\sigma),\alpha(t)F(t,u)\rangle dt\\&=\int_R h^{(i)}(t,\sigma)dt\longrightarrow 0\quad(i\to\infty).\end{aligned}\tag{2.12}$$

且上式的收敛关于 $\sigma\in\Sigma$ 是一致的。

从 $\|\mu_t^{(i)}(\sigma)-\mu_t(\sigma)\|\leqslant 2$ 且 $0\leqslant a(t)\leqslant 1$，可推知

$$
\begin{aligned}
|h^{(i)}(t,\sigma)| &\leqslant \int_N |F(t,u)|\,d|\mu_t^{(i)}(\sigma)-\mu_t(\sigma)| \\
&\leqslant \max_{(t,u)\in[t'-1,t''+1]\times S}\left\{|F(t,u)|\int_N d|\mu_t^{(i)}(\sigma)-\mu_t(\sigma)|\right\} \\
&\leqslant 2\max_{(t,u)\in[t'-1,t''+1]\times S}|F(t,u)| = C,\quad \forall t\in[t'-1,t''+1].
\end{aligned}
$$

这样，从

$$
\begin{aligned}
&\int_R \langle \mu_t^{(i)}(\sigma)-\mu_t(\sigma), a(t)F(t,u)\rangle dt \\
&\quad = \int_R h^{(i)}(t,\sigma)dt \\
&= \int_{t_1-\eta}^{t_1} h^{(i)}(t,\sigma)dt + \int_{t_2}^{t_2+\eta} h^{(i)}(t,\sigma)dt + \int_{t_1}^{t_2} \langle \mu_t^{(i)}(\sigma)-\mu_t(\sigma), F(t,u)\rangle dt
\end{aligned}
$$

得到

$$
\left|\int_{t_1}^{t_2}\langle \mu_t^{(i)}(\sigma)-\mu_t(\sigma), F(t,u)\rangle dt\right| \leqslant \left|\int_R h^{(i)}(t,\sigma)dt\right| + 2C\eta.
$$

这一估计式就完成了本命题的证明．这是因为其中的常数 C 并不依赖于 η，而 $\eta>0$ 可取得任意小，此外已证(2.12)的积分

$$
\int_R h^{(i)}(t,\sigma)dt
$$

当 $i\to\infty$ 时关于 $\sigma\in\Sigma$ 一致地收敛于零。证毕。

第三章　逼近引理

在这一章里，Σ 表示任一距离空间。

设给出从空间 Σ 到广义控制集合 $\mathfrak{M}_U$ 中的一个映射：

$$\sigma \longmapsto \mu_t(\sigma), \qquad \sigma \in \Sigma.$$

也就是说，有一族广义控制

$$\{\mu_t(\sigma) : \sigma \in \Sigma\} \subset \mathfrak{M}_U.$$

如果对任一具有紧支集的连续函数 $g(t,u)$，当 $\sigma_i \to \hat{\sigma}\ (i\to\infty)$ 时，

$$\int_R \langle \mu_t(\sigma_i) - \mu_t(\hat{\sigma}), g(t,u)\rangle dt \longrightarrow 0 \quad (i\to\infty),$$

就称这族广义控制是弱连续的，或关于参数 σ 是弱连续的。若将条件改为

$$\int_R \|\mu_t(\sigma_i) - \mu_t(\hat{\sigma})\| dt \longrightarrow 0 \quad (i\to\infty),$$

则称这族广义控制是强连续的。

显然，任一族强连续广义控制必是弱连续的。但其逆不真。

逼近引理是说，任给一族强连续的广义控制 $\mu_t(\sigma)$ $(\sigma\in\Sigma)$，可以用某一族强连续的常义控制 $\delta_{u(t,\sigma)}$ 在弱收敛的意义下以任意精度逼近，并且这一逼近关于参数 $\sigma\in\Sigma$ 是一致的。因此，常义控制不仅（在弱收敛的意义下）在所有广义控制集 $\mathfrak{M}_U$ 中稠密，而且可以一致逼近任一族强连续广义控制到任意的精度。

逼近定理的确切叙述与证明将在 3.2 节中给出。在此之前，我们先介绍函数的"光滑化"以及单位分解（见定理 3.1），它们不仅在本书定理的证明中要用到，而且对数学的其他分支也是重要的。

3.1 单位分解

设 $u\in \boldsymbol{R}^r$,定义非负函数 $s(u)$ 为:

$$s(u)=\begin{cases}\gamma\exp\left(\dfrac{1}{|u|^2-1}\right), & |u|<1;\\ 0, & |u|\geqslant 1,\end{cases}$$

其中，正值常数 γ 的选取使得下面的等式成立:

$$\int_{R^r} s(u)du=1,$$

即

$$\frac{1}{\gamma}=\int_{|u|<1}\exp\left(\frac{1}{|u|^2-1}\right)du.$$

显然，函数 $s(u)$ 无限次可微且关于原点对称:

$$s(u)=s(-u),\qquad \forall u\in \boldsymbol{R}^r.$$

此外，$s(u)$ 在原点达到最大值 $\gamma\exp(-1)$,在单位球 $|u|\leqslant 1$ 之外为零，而单位球是它的支集。

定义变量 $u,v\in \boldsymbol{R}^r$ 的函数族

$$s_\varepsilon(u,v)=\frac{1}{\varepsilon^r}\ s\left(\frac{u-v}{\varepsilon}\right),$$

这里 $\varepsilon>0$ 是参数。显然 $s_\varepsilon(u,v)$ 与 $s(u)$ 一样也是无限次可微的。这族函数将称为光滑化函数族。从下面的命题 3.1，我们可体会到如此命名的根据。

对任意固定的 ε 和 u,作为 v 的函数 $s_\varepsilon(u,v)$ 以点 u 为对称中心，且在该点达到最大值 $(1/\varepsilon^r)\gamma\exp(-1)$，它的支集是球 $|u-v|\leqslant\varepsilon$，此外还满足如下等式，

$$\int_{R^r} s_\varepsilon(u,v)dv=\int_{|u-v|<\varepsilon} s_\varepsilon(u,v)dv=1,\ \forall u\in \boldsymbol{R}^r,\ \forall\varepsilon>0.$$

这一等式只须通过变量代换 $w=(u-v)/\varepsilon$，并注意到相应的 Jacobi 行列式之绝对值是 ε^r，即可得到。

设 $g(u)$ 是 $\boldsymbol{R}^r$ 上的局部可积函数。所谓局部可积函数是指在任一

有界可测子集上 Lebesgue 可积的函数。对每个 $\varepsilon>0$，由下式：

$$g(u)\longmapsto g_\varepsilon(u)=S_\varepsilon g(u)=\int_{R^r}s_\varepsilon(u,v)g(v)dv=\int_{|u-v|<\varepsilon}s_\varepsilon(u,v)g(v)dv$$

所确定的积分算子族 S_ε 称为光滑算子族，函数 $g_\varepsilon(u)$ 称为 $g(u)$ 的 ε-光滑函数。

命题 3.1 对任一局部可积函数 $g(u)$ 和一切 $\varepsilon>0$，它的 ε-光滑函数 $g_\varepsilon(u)$ 必定无限次可微，且它的各阶偏导数由下式给出，

$$\left(\frac{\partial}{\partial u^1}\right)^{k_1}\cdots\left(\frac{\partial}{\partial u^r}\right)^{k_r}g_\varepsilon(u)=\int_{R^r}\left(\frac{\partial}{\partial u^1}\right)^{k_1}\cdots\left(\frac{\partial}{\partial u^r}\right)^{k_r}s_\varepsilon(u,v)g(v)dv,$$

$$\forall k_1\geqslant 0,\cdots,k_r\geqslant 0.$$

证 设函数 $\omega(u,v)$ 满足如下条件：

(1) 无限次可微；

(2) 当 u 属于任一事先给定的有界集时，它作为 v 的函数在一固定的球 S 之外为零。

我们将证明，在上述条件下，函数

$$h(u)=\int_{R^r}\omega(u,v)g(v)dv$$

关于 u 无限次可微，并且

$$\frac{\partial h}{\partial u^i}=\int_{R^r}\frac{\partial}{\partial u^i}\omega(u,v)g(v)dv,\quad i=1,\cdots,r.$$

由此可直接得出本命题的证明。这是因为所述函数 $s_\varepsilon(u,v)$ 及其各阶偏导数都满足上述对 $\omega(u,v)$ 所给出的条件。

将增量 $h(\hat{u}+\delta u)-h(\hat{u})$ 表示成如下形式，

$$\begin{aligned}h(\hat{u}+\delta u)-h(\hat{u})&=\int_{R^r}[\omega(\hat{u}+\delta u,v)-\omega(\hat{u},v)]g(v)dv\\&=\int_{R^r}dv\int_0^1\frac{\partial}{\partial u}\omega(\hat{u}+\theta\delta u,v)g(v)\,d\theta\,\delta u\\&=\int_S dv\int_0^1\frac{\partial}{\partial u}\omega(\hat{u}+\theta\delta u,v)g(v)d\theta\delta u.\end{aligned}$$

我们有

$$h(\hat{u}+\delta u)-h(\hat{u})-\int_{E^r}\frac{\partial}{\partial u}\omega(\hat{u},v)g(v)dv\,\delta u$$
$$=\int_{E^r}dv\int_0^1\left(\frac{\partial}{\partial u}\omega(\hat{u}+\theta\delta u,v)-\frac{\partial}{\partial u}\omega(\hat{u},v)\right)g(v)d\theta\,\delta u$$
$$=\int_S\int_0^1\left(\frac{\partial}{\partial u}\omega(\hat{u}+\theta\delta u,v)-\frac{\partial}{\partial u}\omega(\hat{u},v)\right)g(v)d\theta dv\,\delta u.$$

因此

$$\frac{1}{|\delta u|}\left|h(\hat{u}+\delta u)-h(\hat{u})-\int_{E^r}\frac{\partial}{\partial u}\omega(\hat{u},v)g(v)dv\delta u\right|$$
$$\leqslant\int_S\int_0^1\left|\frac{\partial}{\partial u}\omega(\hat{u}+\theta\delta u,v)-\frac{\partial}{\partial u}\omega(\hat{u},v)\right|\cdot|g(v)|d\theta\,dv.$$

不等号右侧被积式的第一个因子对任意固定的 θ 和 v 随 $\delta u\to 0$ 而收敛于零，并且它受控于如下常数，

$$2\max_{(u,v)\in V\times S}\left|\frac{\partial}{\partial u}\omega(u,v)\right|\leqslant 2C,$$

其中，V 是点 $\hat{u}$ 的一个有界邻域。因此，根据 Lebesgue 控制收敛定理，便知上述不等式的左侧随 $\delta u\to 0$ 而收敛于零。证毕。

下面一个命题为证明单位分解的定理 3.1 所需，在后面的第四、五章中还将多次用到。

命题 3.2 对任一给定的紧集 $K\subset \boldsymbol{R}^r$ 及任一包含 K 的开集 O，必能作出一个支集在 O 内的无限次可微函数 $\alpha(u)$，满足如下条件，

$$0\leqslant\alpha(u)\leqslant 1,\quad \forall u\in\boldsymbol{R}^r \quad 且 \quad \alpha(u)=1,\quad \forall u\in K.$$

证 设 K 到闭集 $\boldsymbol{R}^r\backslash O$ 的距离为 d。显然，因为 K 是紧集且与闭集 $\boldsymbol{R}^r\backslash O$ 不相交，故必有 $d>0$。我们用 V 表示紧集 K 的 $d/3$ 邻域（如果 $d=\infty$，则 V 可取为 K 的任一有界邻域）。令 $\chi(u)$ 为集 V 的特征函数。它的 ε-光滑函数

$$\chi_\varepsilon(u)=\int_E s_\varepsilon(u,v)\chi(v)dv=\int_{|u-v|<\varepsilon}s_\varepsilon(u,v)\chi(v)dv$$

是无限次可微的。我们要证对 $\varepsilon\leqslant d/3$，所取的函数 $\chi_\varepsilon(u)$ 满足本命题的要求。首先，有如下的估计：

$$0 \leqslant \chi_\varepsilon(u) \leqslant \int_{R^r} s_\varepsilon(u,v)dv = 1.$$

其次，在 $u \in K$ 时，从 $|u-v| \leqslant \varepsilon \leqslant d/3$ 可知 $v \in V$，即 $\chi(v)=1$。因此，$u \in K$ 且 $\varepsilon \leqslant d/3$ 时，有

$$\chi_\varepsilon(u) = \int_{|u-v|\leqslant\varepsilon} s_\varepsilon(u,v)\chi(v)dv = \int_{|u-v|\leqslant\varepsilon} s_\varepsilon(u,v)dv = 1.$$

最后，如 u 到 K 之距离大于 $2d/3$，则 $|u-v| \leqslant \varepsilon \leqslant d/3$ 表明 $v \notin V$，即 $\chi(v)=0$。因此，

$$\chi_\varepsilon(u) = \int_{|u-v|\leqslant\varepsilon} s_\varepsilon(u,v)\chi(v)dv = 0.$$

这样，可取

$$\alpha(u) = \chi_\varepsilon(u), \quad 0 < \varepsilon \leqslant d/3.$$

从证明过程可以看出，在 O 内部任意一个包含紧集 K 的小邻域均可取作 V。任意充分小的正数都可作为 ε。因此，我们可以设想一个函数 $\alpha(u)$，它在 K 上恒为 1 并且在 K 之外急剧地降为 0，又保持为无限次可微。也就是说，有可微函数 $\alpha(u)$ "任意地接近" 于紧集 K 的特征函数。

定理 3.1 设一紧集 $K \subset \boldsymbol{R}^r$ 被有限个开集 $O_1, \cdots, O_p$ 所覆盖，则必存在无限次可微函数 $\alpha_1(u), \cdots, \alpha_p(u)$，满足

(1) $0 \leqslant \alpha_i(u) \leqslant 1, \quad \forall u \in \boldsymbol{R}^r, \quad i=1,2,\cdots,p$;

(2) 函数 $\alpha_i(u)$ 有紧支集且含于 O_i 内，$i=1,2,\cdots,p$;

(3) $\sum_{i=1}^{p} \alpha_i(u) = 1, \quad \forall u \in K$.

这组函数 $\alpha_1(u), \cdots, \alpha_p(u)$ 称为在紧集 K 上关于覆盖 $O_1, \cdots, O_p$ 的一个单位分解。

证 必存在有界开集 $O'_1, \cdots, O'_p$，它们覆盖紧集 K，并满足

$$\overline{O'_i} \subset O_i, \quad i=1,\cdots,p.$$

事实上，我们可以先用有限个开球 $S_1, \cdots, S_q$ 覆盖紧集 K，其中每一个闭球 $\overline{S_j}$ 被包含在某一个 O_i 之内。再记 O'_1 为所有那些闭包属于 O_1 内的 S_j 之并集。若没有这样的球，则令 O'_1 为空集。显然 $\overline{O'_1} \subset O_1$。同样，记 O'_2 为闭包属于 O_2 内的那些球 S_j 之并集，等等。于是，这组开

集 $O'_1,\cdots,O'_p$ 即为所求。

按命题 3.2，若 O'_i 非空，则可构作一个非负的无限次可微函数 $\beta_i(u)$，在 $\overline{O'_i}$ 上等于 1，且有在 O_i 内的紧支集。若 O'_i 是空集，则令 $\beta_i(u)\equiv 0$. 由于对每点 $u\in\overline{O'_i}$ 必有对应的函数 $\beta_i(u)=1$，故有

$$\beta_1(u)+\cdots+\beta_p(u)>0,\quad \forall u\in\bigcup_{i=1}^{p}\overline{O'_i}\supset K.$$

又存在一个无限次可微函数 $\beta(u)$，它在 K 上为 1，且它的支集含于开集 $\bigcup\limits_{i=1}^{p}O'_i$ 之内。我们按下式定义函数 $\alpha_i(u)$，

$$\alpha_i(u)=\begin{cases}\beta(u)\dfrac{\beta_i(u)}{\beta_1(u)+\cdots+\beta_p(u)}, & u\in\bigcup\limits_{j=1}^{p}\overline{O'_j};\\ 0, & u\notin\bigcup\limits_{j=1}^{p}\overline{O'_j}.\end{cases}$$

直接可看出，函数 $\alpha_i(u),i=1,\cdots,p$，组成了紧集 K 上关于覆盖 $O_1,\cdots,O_p$ 的一个单位分解。证毕。

我们将看到，在逼近引理中仅需要 K 上的单位分解函数 $\alpha_i(u)$ 的连续性，并不需要这些函数的可微性。

现在我们来证明已在第一章里用过的如下命题。

命题 3.3 设 $C^0(\boldsymbol{R}^m)$ 是 $\boldsymbol{R}^m$ 上的有紧支集的所有连续函数所组成的赋范空间，其范数定义为

$$\|g\|=\sup_{z\in\boldsymbol{R}^m}\|g(z)\|.$$

在 $C^0(\boldsymbol{R}^m)$ 中必存在一列函数 $g_1(z)$，$g_2(z)$，…，使得对每一个 $g(z)\in C^0(\boldsymbol{R}^m)$，必可从该列函数中选取一个子列 $g_{i_k}(z)$，满足如下条件：

(1) $\|g_{i_k}-g\|\longrightarrow 0,\quad (k\to\infty)$；

(2) $\|g_{i_k}\|\leqslant\|g\|,\quad \forall k=1,2,\cdots$；

(3) 所有函数 $g_{i_k}(z)(k=1,2,\cdots)$ 的支集都落在空间 $\boldsymbol{R}^m$ 的一个有界子集内。

证 l,k 是任意两个固定的自然数，$\overline{S_l}(0)$ 是 $\boldsymbol{R}^m$ 中以原点为中心，l 为半径的闭球。设有限开集族

$$\{O_{l,k}^{(1)},\cdots,O_{l,k}^{(p_{l,k})}\}$$

中每个 $O_{l,k}^{(i)}$ 的直径 $\leqslant 1/k$，且覆盖 $\overline{S_l}(0)$. 用

$$\{a_{l,k}^{(i)}(z), i=1,\cdots,p_{l,k}\}$$

表示$\overline{S_l}(0)$上关于该有限开集族的一个单位分解。当l，k取遍一切自然数，r_i取遍一切有理数时，我们得到下面可列个函数

$$\sum_{i=1}^{p_{l,k}} r_i a_{l,k}^{(i)}(z), l,k=1,2,\cdots.$$

如果将这可列个函数以任意方式排成一个函数列：$g_1(z), g_2(z), \cdots$，这个函数列就满足本命题的三点要求。下面来证明这个论断。

任取$C^0(\boldsymbol{R}^m)$中的一个函数$g(z)$。$\hat{l}$取得充分大，以致函数$g(z)$的支集完全包含在$S_{\hat{l}}(0)$之中。由于函数$g(z)$在全空间是一致连续的，又开集$O_{\hat{l},k}^{(i)}$的直径$\leqslant 1/k$随$k\to\infty$而关于i一致地收敛于零。因此，对每个$k=1,2,\cdots$，可选取有理数$r_k^{(i)}(i=1,\cdots,p_{\hat{l},k})$，使得如下的关系式成立：

$$|r_k^{(i)}|\leqslant\sup\{|g(z)|: z\in O_{\hat{l},k}^{(i)}\}, \ \forall i=1,\cdots,p_{\hat{l},k},$$

$$\sup\{|r_k^{(i)}-g(z)|: z\in O_{\hat{l},k}^{(i)}\}\to 0 \quad (k\to\infty),$$

其中第二式关于i还是一致收敛的。换言之，

$$a_k=\max_i \sup\{|r_k^{(i)}-g(z)|: z\in O_{\hat{l},k}^{(i)}\}\longrightarrow 0(k\to\infty).$$

我们要证明这样取出的函数列

$$\sum_{i=1}^{p_{\hat{l},k}} r_k^{(i)} a_{\hat{l},k}^{(i)}(z), \quad k=1,2,\cdots,$$

对这一$g(z)$满足所述的三点要求。首先，条件(1)由下式可得到满足，

$$\left\|\sum_{i=1}^{p_{\hat{l},k}} r_k^{(i)} a_{\hat{l},k}^{(i)}-g\right\|=\left\|\sum_{i=1}^{p_{\hat{l},k}}(r_k^{(i)}-g)a_{\hat{l},k}^{(i)}\right\|$$

$$\leqslant\sup_{z\in\boldsymbol{R}^m}\left(\sum_i |r_k^{(i)}-g(z)|a_{\hat{l},k}^{(i)}(z)\right)$$

$$\leqslant\sup_{z\in\boldsymbol{R}^m}\left(\sum_i a_k a_{\hat{l},k}^{(i)}(z)\right)$$

$$\leqslant a_k\sup_{z\in\boldsymbol{R}^m}\sum_i a_{\hat{l},k}^{(i)}(z)\leqslant a_k\longrightarrow 0 \quad (k\to\infty).$$

条件(2)由下式得到，

$$\left\|\sum_{i=1}^{p_{\hat{l},k}} r_k^{(i)} \alpha_{\hat{l};k}^{(i)}\right\| = \sup_{x\in R^m}\left|\sum_i r_k^{(i)} \alpha_{\hat{l};k}^{(i)}(z)\right|$$

$$\leqslant \sup_{x\in R^m}\left(\sum_i |r_k^{(i)}| \alpha_{\hat{l};k}^{(i)}(z)\right)$$

$$\leqslant \sup_{x\in R^m}\left(\sum_i \|g\| \alpha_{\hat{l};k}^{(i)}(z)\right)$$

$$\leqslant \|g\| \sup_{x\in R^m} \sum_i \alpha_{\hat{l};k}^{(i)}(z) \leqslant \|g\|.$$

最后，条件(3)的满足是显然的。证毕。

3.2 逼近引理

定理 3.2 如果 $\mu_t(\sigma)$ 是一族带参数 $\sigma\in\Sigma$ 的强连续广义控制，并对一切 $t\in \boldsymbol{R}$ 和 $\sigma\in\Sigma$，$\mu_t(\sigma)$ 的测度都集中在一个固定的有界集 $N\subset U\subset \boldsymbol{R}^r$ 中，则必存在一列逐段常值控制

$$\delta_{u^{(i)}(t;\sigma)}\in\Omega_U,\ i=1,2,\cdots,$$

它关于 $\sigma\in\Sigma$ 强连续，并且满足

(i) 所有 $\delta_{u^{(i)}(t;\sigma)}$ 的测度集中在 N 上。等价地说，对一切 $t\in\boldsymbol{R}$ 及 $\sigma\in\Sigma, i=1,2,\cdots$，都有 $u^{(i)}(t;\sigma)\in N$；

(ii) 当 $i\to\infty$ 时，序列 $\delta_{u^{(i)}(t;\sigma)}$ 关于 $\sigma\in\Sigma$ 一致地弱收敛于 $\mu_t(\sigma)$. 换言之，对任一有紧支集的连续函数 $g(t,u)$，有

$$\int_R \langle \mu_t(\sigma)-\delta_{u^{(i)}(t;\sigma)}, g(t,u)\rangle dt \longrightarrow 0\quad (i\to\infty),$$

并且这个收敛关于 $\sigma\in\Sigma$ 是一致的。

证 设对每个自然数 i，有界开集族

$$\{O_j^{(i)}, j=1,\cdots,p_i\}$$

覆盖了闭集 $\overline{N}$，并设集合 $O_1^{(i)},\cdots,O_{p_i}^{(i)}$ 的最大直径随 $i\to\infty$ 而趋于零。又

$$\{\alpha_j^{(i)}(u), j=1,\cdots,p_i\}$$

是紧集$\overline{N}$上的关于子覆盖

$$O_1^{(i)},\cdots,O_{p_i}^{(i)}$$

的一个单位分解。对任一$\sigma\in\boldsymbol{\Sigma}$，函数

$$\lambda_j^{(i)}(t;\sigma)=\langle\mu_t(\sigma),\alpha_j^{(i)}(u)\rangle, t\in\boldsymbol{R}, \tag{3.1}$$

关于t可测且满足条件

$$0\leqslant\lambda_j^{(i)}(t;\sigma)\leqslant1,\quad \sum_{j=1}^{p_i}\lambda_j^{(i)}(t;\sigma)\equiv1,\ \forall\sigma\in\boldsymbol{\Sigma},\ i=1,2,\cdots.$$

这是因为，对每一$\sigma\in\boldsymbol{\Sigma}$，我们有

$$0\leqslant\lambda_j^{(i)}(t;\sigma)=\int_N\alpha_j^{(i)}(u)d\mu_t(\sigma)\leqslant\int_N d\mu_t(\sigma)=1,$$

$$\sum_{j=1}^{p_i}\lambda_j^{(i)}(t;\sigma)=\langle\mu_t(\sigma),\sum_{j=1}^{p_i}\alpha_j^{(i)}(u)\rangle=\int_N\sum_{j=1}^{p_i}\alpha_j^{(i)}(u)d\mu_t(\sigma)$$
$$=\int_N d\mu_t(\sigma)=1.$$

现来证明，当$\sigma\to\hat{\sigma}$时，

$$\sum_{j=1}^{p_i}\int_R|\lambda_j^{(i)}(t;\sigma)-\lambda_j^{(i)}(t;\hat{\sigma})|dt\longrightarrow0,\quad i=1,2,\cdots. \tag{3.2}$$

事实上，

$$\sum_{j=1}^{p_i}\int_R|\lambda_j^{(i)}(t;\sigma)-\lambda_j^{(i)}(t;\hat{\sigma})|dt=\sum_{j=1}^{p_i}\int_R|\langle\mu_t(\sigma)-\mu_t(\hat{\sigma}),\alpha_j^{(i)}(u)\rangle|dt.$$

但由于$|\alpha_j^{(i)}(u)|\leqslant1$，且$\alpha_j^{(i)}$有紧支集，根据Radon测度范数的定义，上面这个不等式的右侧积分不大于

$$\sum_{j=1}^{p_i}\int_R\|\mu_t(\sigma)-\mu_t(\hat{\sigma})\|dt=p_i\int_R\|\mu_t(\sigma)-\mu_t(\hat{\sigma})\|dt.$$

由假设，$\mu_t(\sigma)$强连续地依赖于参数$\sigma\in\boldsymbol{\Sigma}$，即

$$p_i\int_R\|\mu_t(\hat{\sigma})-\mu_t(\sigma)\|dt\longrightarrow0,\quad(\sigma\to\hat{\sigma}).$$

于是，(3.2) 成立。

对每个$i=1,2,\cdots$，在交集$O_j^{(i)}\cap N$中任意选一点$u_j^{(i)}$，$j=1,\cdots,p_i$.如果该交集是空集，就令$u_j^{(i)}=\hat{u}$，这里$\hat{u}$是N中的一个固定点。

这样，就得到了一个点集

$$\{u_1^{(i)},\cdots,u_{p_i}^{(i)}\}\subset N, i=1,2,\cdots.$$

直观地说，当(3.1)式中的函数 $a_j^{(i)}(u)$ 充分地"接近"于集合 $O_j^{(i)}$ 的特征函数时，$\lambda_j^{(i)}(t;\sigma)$将近似地表示出$\mu_t(\sigma)$于 $O_j^{(i)}\cap N$ 上的测度。但对充分大的 i，所有集合 $O_j^{(i)}$ 的直径必充分小，结果 $\mu_t(\sigma)$可近似地视为按量值 $\lambda_1^{(i)}(t;\sigma),\cdots,\lambda_{p_i}^{(i)}(t;\sigma)$ 集中在点 $u_1^{(i)},\cdots,u_{p_i}^{(i)}$ 上。随着 i 的增大，按弱收敛的意义，两者的误差越来越小。

现在来将上述直观上显然的事实加以严格化，即证明广义控制列（这是一列颤振控制）

$$\Delta_t^{(i)}(\sigma)=\sum_{j=1}^{p_i}\lambda_j^{(i)}(t;\sigma)\delta_{u_j^{(i)}} \tag{3.3}$$

当 $i\to\infty$ 时关于 $\sigma\in\Sigma$ 一致地弱收敛于广义控制$\mu_t(\sigma)$。

任取一具有紧支集的连续函数 $g(t,u)$，并令 I_g 为支集在 t 轴上的投影。记

$$\eta^{(i)}=\max_{1\leqslant j\leqslant p_i}\sup_{(t,u)\in I_g\times O_j^{(i)}}\{|g(t,u)-g(t,u_j^{(i)})|\}.$$

显然，

$$\eta^{(i)}\to 0\qquad(i\to\infty).$$

因此，

$$\begin{aligned}
&\left|\int_R\langle\mu_t(\sigma)-\Delta_t^{(i)}(\sigma),g(t,u)\rangle dt\right|\\
&=\left|\int_R\langle\mu_t(\sigma),\sum_{j=1}^{p_i}a_j^{(i)}(u)g(t,u)\rangle dt-\sum_{j=1}^{p_i}\int_R\lambda_j^{(i)}(t;\sigma)g(t,u_j^{(i)})dt\right|\\
&=\left|\int_R\langle\mu_t(\sigma),\sum_{j=1}^{p_i}a_j^{(i)}(u)[g(t,u)-g(t,u_j^{(i)})]\rangle dt\right|\\
&\leqslant\sum_{j=1}^{p_i}\int_{I_g}\langle\mu_t(\sigma),a_j^{(i)}(u)|g(t,u)-g(t,u_j^{(i)})|\rangle dt\\
&\leqslant\eta^{(i)}\sum_{j=1}^{p_i}\int_{I_g}\langle\mu_t(\sigma),a_j^{(i)}(u)\rangle dt\\
&=\eta^{(i)}\int_{I_g}dt\int_N\sum_{j=1}^{p_i}a_j^{(i)}(u)d\mu_t(\sigma)
\end{aligned}$$

$$=\eta^{(i)}\int_{I_g}dt\int_N d\mu_t(\sigma)=\eta^{(i)}\int_{I_g}dt=\eta^{(i)}|I_g|\longrightarrow 0\quad(i\to\infty),$$

且这一收敛关于 $\sigma\in\Sigma$ 是一致的。

现在来构造一列逐段常值控制

$$\delta_{u^{(i)}(t;\sigma)}\in\Omega_U,\quad i=1,2,\cdots,$$

它们强连续地依赖于 $\sigma(i=1,2,\cdots)$，并且对任一有紧支集的连续函数 $g(t,u)$，关于 $\sigma\in\Sigma$ 一致地成立

$$\int_R\langle\Delta_t^{(i)}(\sigma)-\delta_{u^{(i)}(t;\sigma)},g(t,u)\rangle dt\longrightarrow 0\quad(i\to\infty).\tag{3.4}$$

对每一 $i=1,2,\cdots$，将区间 $[-i,i]$ 等分成 $2i^2$ 个长度为 $1/i$ 的小区间 $I_k^{(i)}(k=1,2,\cdots,2i^2)$. 再将每个子区间 $I_k^{(i)}$ 分成 p_i 个子子区间 $I_{k_j}^{(i)}(\sigma)$ $(j=1,\cdots,p_i)$，其中 $I_{k_j}^{(i)}(\sigma)$ 的长度（依赖于 $\sigma\in\Sigma$）分别是

$$|I_{k_j}^{(i)}(\sigma)|=\int_{I_k^{(i)}}\lambda_j^{(i)}(t;\sigma)dt.\tag{3.5}$$

由于对一切 $\sigma\in\Sigma(i=1,2,\cdots)$，都有 $\sum_{j=1}^{p_i}\lambda_j^{(i)}(t;\sigma)\equiv 1$，这是可以做到的。

设这些区间 $I_{k_j}^{(i)}(\sigma)$ 已按这样的规则来排列：若 $j'\leqslant j''$，便有 $I_{k_{j'}}^{(i)}(\sigma)$ 位于 $I_{k_{j''}}^{(i)}(\sigma)$ 之左。如果对一切 $t\in I_k^{(i)}$ 有 $\lambda_j^{(i)}(t;\sigma)=0$，则对应的区间 $I_{k_j}^{(i)}(\sigma)$ 退化成一点。

定义函数 $u^{(i)}(t;\sigma)$（除有限个点上的值不定外）为

$$\begin{aligned}&u^{(i)}(t;\sigma)=\hat{u},\quad \forall t\notin[-i,i],\\&u^{(i)}(t;\sigma)=u_j^{(i)},\quad \forall t\in I_{k_j}^{(i)}\text{ 且 }\forall k=1,2,\cdots,2i^2.\end{aligned}\tag{3.6}$$

为证明结论(3.4)，设指标 i 充分大，以致包含关系 $I_g\subset[-i,i]$ 成立（记住，I_g 是 $g(t,u)$ 之支集在 t 轴上的投影）。在每一区间 $I_k^{(i)}$ 中任取一点 $t_k^{(i)}$。那么，

$$\begin{aligned}&\int_R\langle\Delta_t^{(i)}(\sigma)-\delta_{u^{(i)}(t;\sigma)},g(t,u)\rangle dt\\&=\sum_{k=1}^{2i^2}\int_{I_k^{(i)}}\langle\Delta_t^{(i)}(\sigma)-\delta_{u^{(i)}(t;\sigma)},g(t_k^{(i)},u)\rangle dt\\&\quad+\sum_{k=1}^{2i^2}\int_{I_k^{(i)}}\langle\Delta_t^{(i)}(\sigma)-\delta_{u^{(i)}(t;\sigma)},g(t,u)-g(t_k^{(i)},u)\rangle dt\end{aligned}$$

$$= \sum_{k=1}^{2i^2} \int_{I_k^{(i)}} \langle \Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}, g(t,u) - g(t_k^{(i)},u) \rangle dt,$$

这是因为，根据(3.3),(3.6)及(3.5),我们有

$$\sum_{k=1}^{2i^2} \int_{I_k^{(i)}} \langle \Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}, g(t_k^{(i)},u) \rangle dt$$

$$= \sum_{k=1}^{2i^2} \left\{ \int_{I_k^{(i)}} \sum_{j=1}^{p_i} \lambda_j^{(i)}(t;\sigma) g(t_k^{(i)},u_j^{(i)}) dt - \sum_{j=1}^{p_i} \int_{I_{k_j}^{(i)}} g(t_k^{(i)},u_j^{(i)}) dt \right\}$$

$$= \sum_{k=1}^{2i^2} \left\{ \sum_{j=1}^{p_i} g(t_k^{(i)},u_j^{(i)}) \left(\int_{I_k^{(i)}} \lambda_j^{(i)}(t;\sigma)\, dt - \int_{I_{k_j}^{(i)}} dt \right) \right\} = 0.$$

因此，

$$\left| \int_R \langle \Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}, g(t,u) \rangle dt \right|$$

$$\leqslant \sum_{k=1}^{2i^2} \int_{I_k^{(i)}} |\langle \Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}, g(t,u) - g(t_k^{(i)},u) \rangle| dt.$$

现对上式右侧进行估计。

令 I 为 t 轴上内部包含紧集 I_0 的任一区间。对于充分大的 i，有

$$\sum_{k=1}^{2i^2} \int_{I_k^{(i)}} |\langle \Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}, g(t,u) - g(t_k^{(i)},u) \rangle| dt$$

$$\leqslant \sum_{I_k^{(i)} \subset I} \int_{I_k^{(i)}} |\langle \Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}, g(t;u) - g(t_k^{(i)},u) \rangle| dt,$$

上式中，右侧的求和是对所有完全在 I 之内的区间 $I_k^{(i)}$ 进行的. $I_k^{(i)}$ 之长度是 $1/i$. 测度 $\Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}$ 集中在有界集 N 上，而它的范数不大于 2，即

$$\|\Delta_i^{(i)}(\sigma) - \delta_{u^{(i)}(t;\sigma)}\| \leqslant \left\| \sum_{j=1}^{p_i} \lambda_j^{(i)}(t;\sigma) \delta_{u_j^{(i)}} \right\| + \|\delta_{u^{(i)}(t;\sigma)}\|$$

$$\leqslant \sum_{j=1}^{p_i} \lambda_j^{(i)}(t;\sigma) + 1 = 2.$$

记

$$\xi^{(i)} = \sup\left\{ |g(t',u) - g(t'',u)| ; u \in N, t', t'' \in I, |t'-t''| \leqslant \frac{1}{i} \right\},$$

我们得到

$$\sum_{I_k^{(i)}\subset I}\int_{I_k^{(i)}}|\langle\Delta_t^{(i)}(\sigma)-\delta_{u^{(i)}(t,\sigma)},g(t,u)-g(t_k^{(i)},u)\rangle|dt$$

$$\leqslant\int_I 2\xi^{(i)}dt=2|I|\xi^{(i)}\longrightarrow 0,\quad (i\to\infty).$$

这就证明了(3.4)。

余下尚须证明这族控制 $\delta_{u^{(i)}(t,\sigma)}$ 强连续地依赖于 $\sigma\in\Sigma$. 为此，我们注意到，如果 $\hat{t}$ 是区间 $I_{k_i}^{(i)}(\hat{\sigma})$ 的任一内点，而且点 σ 充分接近于 $\hat{\sigma}$，则由(3.5)和(3.2)，可知 $\hat{t}$ 也是 $I_{k_i}^{(i)}(\sigma)$ 的内点。这时，按函数 $u^{(i)}(t,\sigma)$的定义，我们得到

$$u^{(i)}(\hat{t},\sigma)=u^{(i)}(\hat{t},\hat{\sigma})=u_j^{(i)}.$$

因此，对除了子区间 $I_{k_i}^{(i)}(\hat{\sigma})$ 的那些（至多有限个）端点外的所有 $t\in\boldsymbol{R}$ 当 $\sigma\to\hat{\sigma}$ 时差值范数 $\|\delta_{u^{(i)}(t,\sigma)}-\delta_{u^{(i)}(t,\hat{\sigma})}\|$ 趋于零，又因该差值范数受控于 2，故根据 Lebesgue 控制收敛定理，可得所要求的强连续性：[注1]

$$\int_R\|\delta_{u^{(i)}(t,\sigma)}-\delta_{u^{(i)}(t,\hat{\sigma})}\|dt\longrightarrow 0,\quad (\sigma\to\hat{\sigma}),\quad \forall i=1,2,\cdots.$$

定理 3.2 至此证毕。

关于颤振控制的注 从上述证明可见，对给定的广义控制

$$\Delta_t=\sum_{j=1}^{p}\lambda_j(t)\delta_{u_j},\quad \sum_{j=1}^{p}\lambda_j(t)\equiv 1,\ u_j\in U,\ j=1,\cdots,p,\qquad (3.7)$$

构造逐段常值控制（3.6）的过程产生出一列函数 $u^{(i)}(t)$，它的值以频率 i 在区间 $t\in[-i,i]$ 上快速地振荡于 $u_1,\cdots,u_p$ 诸值之间。这就是说，函数 $u^{(i)}(t)$在每个区间

$$\frac{k}{i}\leqslant t\leqslant\frac{k+1}{i},\ k=-i^2,\cdots,i^2,$$

上，取值于 u_j 的时间是

$$\int_{[k/i,(k+1)/i]}\lambda_j(t)dt,\quad j=1,\cdots,p.$$

[注1] 当 $|t|>2i$ 时，$u^{(i)}(t,\sigma)=\hat{u}=u^{(i)}(t,\hat{\sigma})$. 这时，$\|\delta_{u^{(i)}(t,\sigma)}-\delta_{u^{(i)}(t,\hat{\sigma})}\|=0$，——译者注。

这列振荡得愈来愈快的函数弱收敛于广义控制Δ_t。这就是为什么我们在第二章里称 Δ_t 为颤振控制的原因。

实际上，颤振控制的名称也可用于以任意的取值于U且关于t有界可测的函数取代点u_i的情形。这时，在本质上无须改动我们方才的证明，也能构造出一列逐段常值控制弱收敛于这样的Δ_t. 某种意义下，这些控制也可以说以频率i(对序列中的第i个函数)在与时间t有关的各点$u_1(t),\cdots,u_p(t)$之间快速地振荡。

第四章　微分方程解的存在性与连续依赖性定理

这一章的基本结果是关于微分方程解的存在性和对初始数据以及右端函数的连续依赖性的定理 4.3 与定理 4.4。在4.3 节先证明具有固定右端的微分方程解的存在性与连续依赖性定理 4.2，随后作为定理 4.2 简单的推论，在 4.4 与 4.5 节中叙述并证明定理 4.3 与定理 4.4。从定理 4.2 导出定理 4.3 与 4.4 的方法是典型的和常用的。

在 4.1 节与 4.2 节中，我们引述了所需要的辅助知识并介绍定理 4.2 证明中要用到的压缩映射不动点定理。

4.1　预备知识

设 $F(t,x)$ 是定义在$\boldsymbol{R}\times\boldsymbol{R}^n$上具紧支集的 n 维向量值函数，它关于 t 是可测的且受控于一可积函数 $m_F(t)$：

$$|F(t,x)|\leqslant m_F(t),\ \forall(t,x)\in\boldsymbol{R}\times\boldsymbol{R}^n,\ \int_{\boldsymbol{R}}m_F(t)dt<+\infty,\tag{4.1}$$

而关于 x 满足 Lipschitz 条件，即存在有限值可积函数 $L_F(t)$使得

$$\begin{gathered}|F(t,x')-F(t,x'')|\leqslant L_F(t)|x'-x''|,\\ \forall(t,x'),(t,x'')\in\boldsymbol{R}\times\boldsymbol{R}^n.\end{gathered}\tag{4.2}$$

以后将满足以上条件的 $F(t,x)$全体记为 E_{Lip}。显然，E_{Lip} 是一个线性空间。

由 Lipschitz 性质立即可知 $F(t,x)$ 关于 x 是连续的。现要证明，对任一 n 维 Lebesgue 可测函数 $x(t),t\in\boldsymbol{R}$，函数 $F(t,x(t)),t\in\boldsymbol{R}$，也

是 Lebesgue 可测的。

如果 $x(t)$ 仅取有限个不同值 $x_1, \cdots, x_p$，则上述论断是显然的。这时，

$$x(t) = \sum_{i=1}^{p} \chi_i(t) x_i,$$

其中，$\chi_i(t)$ 是 $\boldsymbol{R}$ 的一些互不相交的 Lebesgue 可测子集的特征函数。于是，函数

$$F(t,x(t)) = F(t,\sum_{i=1}^{p} \chi_i(t)x_i) = \sum_{i=1}^{p} \chi_i(t)F(t,x_i),\ t\in \boldsymbol{R},$$

的可测性便由函数 $F(t,x_i)(i=1,\cdots,p)$ 的可测性直接得知。

每一可测函数 $x(t)(t\in \boldsymbol{R})$ 可表为一列有限值可测函数

$$x_1(t), x_2(t), \cdots, t\in \boldsymbol{R},$$

的几乎处处收敛的极限。又由于 $F(t, x)$ 关于 x 连续，可知对几乎所有的 t，成立

$$F(t,x(t)) = \lim_{j\to\infty} F(t,x_j(t)),$$

因此，函数 $F(t,x(t))$ 是可测函数。

又由(4.1) 知函数 $F(t,x(t))$ 在 $\boldsymbol{R}$ 上可积。因此，对任一可测函数 $x(t), t\in \boldsymbol{R}$，积分

$$\int_{t_1}^{t} F(\theta, x(\theta))d\theta,\ t\in \boldsymbol{R},$$

有意义。

考虑微分方程

$$\dot{x} = F(t,x), \qquad F(t,x)\in E_{\text{Lip}}. \tag{4.3}$$

任一定义在整个 t 轴上且对几乎所有的 $t\in \boldsymbol{R}$ 满足等式

$$\dot{x}(t) = F(t,x(t))$$

的绝对连续函数 $x(t)$ 称为这一方程的一个解。将上式两边从 τ 到 t 积分，考虑到函数 $x(t)$ 的绝对连续性，就得关于未知函数 $x(t)(t\in \boldsymbol{R})$ 的如下积分方程：

$$x(t) = x_\tau + \int_{\tau}^{t} F(\theta, x(\theta))d\theta,\ \ x_\tau = x(\tau).$$

这一积分方程的连续解等价于微分方程(4.3)的满足初始条件

$$x(\tau)=x_\tau$$

的解。请注意，如果函数 $x(t)$ 没有绝对连续的条件，我们不能把微分方程转化为积分方程，因为导数 $\dot{x}(t)$ 从 τ 到 t 的积分可能不等于差值 $x(t)-x(\tau)$。这是实函数论中的已知事实（不妨回顾 Cantor 奇异函数）。

由于 $F(t,x)$的紧支集在 t 轴上的投影是有界的，只要区间$[t_1,t_2]$包含该投影，方程(4.3)的任一解在此区间$[t_1,t_2]$之外必取常值。

为了表述微分方程（4.3）解的存在性及关于初始值与右端函数的连续依赖性的定理 4.2，我们必须引进 E_{Lip} 上的半范。借助于这一半范，我们将估计 E_{Lip} 中的任意两个函数之间的距离。此外，我们还须定义 E_{Lip} 中的一致 Lipschitz 子集的概念。

先来定义半范$\|\cdot\|_w$.首先证明，对任一函数 $F(t,x)\in E_{\mathrm{Lip}}$，函数

$$g(t',t'',x)=\int_{t'}^{t''}F(t,x)dt,\quad (t',t'',x)\in \boldsymbol{R}\times\boldsymbol{R}\times\boldsymbol{R}^n,$$

关于所有变元是联合连续的。实际上，当 $\delta t'$, $\delta t''$, $\delta x\to 0$ 时，

$$\begin{aligned}
&|g(t'+\delta t',t''+\delta t'',x+\delta x)-g(t',t'',x)|\\
&\leqslant\int_{t'-|\delta t'|}^{t'+|\delta t'|}|F(t,x+\delta x)|dt+\int_{t''-|\delta t''|}^{t''+|\delta t''|}|F(t,x+\delta x)|dt\\
&\quad+\left|\int_{t'}^{t''}|F(t,x+\delta x)-F(t,x)|dt\right|\\
&\leqslant\int_{t'-|\delta t'|}^{t'+|\delta t'|}m_F(t)\,dt+\int_{t''-|\delta t''|}^{t''+|\delta t''|}m_F(t)dt+|\delta x|\int_{\boldsymbol{R}}L_F(t)dt\longrightarrow 0.
\end{aligned}$$

半范$\|\cdot\|_w$定义为

$$\|F(t,x)\|_w=\max_{t',t'',x}\left|\int_{t'}^{t''}F(t,x)\,dt\right|,$$ [注1]

这里，积分是对固定的 x 进行的，而 max 是关于所有 t', $t''\in\boldsymbol{R}$ 及 $x\in\boldsymbol{R}^n$ 而取的。[注2]

[注1] 事实上，$\|F(t,x)\|_w$ 已与 (t,x) 无关．更确切地应写成$\|F(\cdot,\cdot)\|_w$或$\|F\|_w$．但因原书如此，故仍沿用原书的记号.——译者注.

[注2] 在此定义的$\|\cdot\|_w$显然是半范：

$$\|F\|_w\geqslant 0,\qquad \|\lambda F\|_w=|\lambda|\,\|F\|_w,\qquad \|F_1+F_2\|_w\leqslant\|F_1\|_w+\|F_2\|_w.$$

如果在 E_{Lip} 中对任一固定 x 和几乎所有的 $t\in\boldsymbol{R}$ 取值相同的任意两个函数都视为同一个函数，则$\|\cdot\|_w$成为范数。E_{Lip} 中的准距 $\rho_w(F_1,F_2)=\|F_1-F_2\|_w$ 在此时就成为距离。

E_{Lip} 中的函数列 $F_j(t,x)$，$j=1,2,\cdots$，按半范 $\|\cdot\|_w$ 收敛于 $F(t,x) \in E_{\text{Lip}}$ 是指

$$\|F_j(t,x)-F(t,x)\|_w \longrightarrow 0 \quad (j\to\infty).$$

这样的半范所定义的拓扑是很"粗"的：一个函数 $F(t,x)$ 只要有充分小的半范 $\|F(t,x)\|_w$，即可属于按此拓扑的零点的一个充分小邻域，然而它却可以在空间 $\boldsymbol{R}\times\boldsymbol{R}^n$ 的一个任意给定的立方体内取很大的绝对值。一个典型例子是关于 t 快速振荡的函数。例如，

$$F(t,x)=g(t,x)\cos pt,$$

其中 p 取较大值，$g(t,x)$ 是一个有紧支集的连续可微函数。为估计 $\|g(t,x)\cos pt\|_w$，可进行分部积分：

$$\begin{aligned}\int_{t'}^{t''} g(t,x)\cos pt dt &= g(t'',x)\int_{t'}^{t''}\cos pt dt \\ &\quad -\int_{t'}^{t''}\left(\int_{t'}^{t}\cos p\theta d\theta\right)\left(\frac{\partial}{\partial t}g(t,x)\right)dt \\ &= \frac{1}{p}\Big\{(\sin pt''-\sin pt')g(t'',x) \\ &\quad -\int_{t'}^{t''}(\sin pt-\sin pt')\frac{\partial}{\partial t}g(t,x)dt\Big\}.\end{aligned}$$

因此，

$$\begin{aligned}\left|\int_{t'}^{t''} g(t,x)\cos pt dt\right| &\leqslant \frac{1}{p}\cdot 2\left\{|g(t'',x)|+\left|\int_{t'}^{t''}\left|\frac{\partial}{\partial t}g(t,x)\right|dt\right|\right\} \\ &\leqslant \text{const}\cdot\frac{1}{p}\longrightarrow 0 \qquad (p\to\infty)\end{aligned}$$

考虑线性空间

$$\begin{aligned}E_\sigma &= \boldsymbol{R}\times\boldsymbol{R}^n\times E_{\text{Lip}} \\ &= \{\sigma=(\tau,x_\tau,F(t,x)):\ (\tau,x_\tau)\in\boldsymbol{R}\times\boldsymbol{R}^n, F(t,x)\in E_{\text{Lip}}\}\end{aligned}$$

并赋以半范

$$\|\sigma\|_w=\|(\tau,x_\tau,F(t,x))\|_w=|\tau|+|x_\tau|+\|F(t,x)\|_w$$

这是方程(4.3)的解的初始数据 $(\tau,x_\tau)\in\boldsymbol{R}\times\boldsymbol{R}^n$ 空间与右端函数 $F(t,x)\in E_{\text{Lip}}$ 空间之乘积空间。

将微分方程 (4.3) 的满足初始条件 $x(\tau)=x_\tau$ 的解 $x(t)(t\in\boldsymbol{R})$ 写成如下形式，

$$x(t)=x(t;\tau,x_\tau,F)=x(t;\sigma),\quad t\in\boldsymbol{R}.$$

称这个解在一子集 $M_\sigma\subset E_\sigma$ 上连续依赖于初始数据与右端函数，是指 $\forall\hat{\sigma}=(\hat{\tau},\hat{x}_{\hat{\tau}},\hat{F})\in M_\sigma$，$\forall\varepsilon>0$，存在 $\delta>0$，只要满足

$$\|\sigma-\hat{\sigma}\|_w=|\tau-\hat{\tau}|+|x_\tau-\hat{x}_{\hat{\tau}}|+\|F(t,x)-\hat{F}(t,x)\|_w\leqslant\delta,$$
$$\sigma=(\tau,x_\tau,F(t,x))\in M_\sigma,$$

就有

$$\max_{t\in\boldsymbol{R}}|x(t;\tau,x_\tau,F)-x(t;\hat{\tau},\hat{x}_{\hat{\tau}},\hat{F})|=\|x(t;\sigma)-x(t;\hat{\sigma})\|\leqslant\varepsilon.$$

一个集合 $M\subset E_{\mathrm{Lip}}$ 称为一致 Lipschitz 的，如果对任一函数 $F(t,x)\in M$，存在 $\boldsymbol{R}$ 上的可积函数 $L_F(t)$，满足

$$|F(t,x')-F(t,x'')|\leqslant L_F(t)|x'-x''|,\quad\int_{\boldsymbol{R}}L_F(t)dt\leqslant C,\qquad(4.4)$$

其中，常数 C 不依赖于 $F(t,x)\in M$ 的特殊选取。

下面的命题给出了一致 Lipschitz 集的一个重要性质。

命题 4.1 设 Φ 是一族等度连续且在一个公共的固定区间 $[T_1,T_2]$ 之外取常值的 n 维函数 $x(t)$ $(t\in\boldsymbol{R})$ 的集合，设 M 是 E_{Lip} 的一致 Lipschitz 子集，则对任一 $\varepsilon>0$，必存在正数 $\delta>0$，使得当

$$\|F(t,x)\|_w\leqslant\delta\quad 且\quad F(t,x)\in M$$

时，成立

$$\max_{t',t''\in\boldsymbol{R}}\left|\int_{t'}^{t''}F(t,x(t))dt\right|<\varepsilon,\quad\forall x(t)\in\Phi.$$

证 按函数族 Φ 的等度连续性定义，在 $t\geqslant0$ 上存在单调下降函数 $\omega(t)$，满足

(i) $\omega(|t|)\to0\quad(t\to0)$，且 $\omega(0)=0$；

(ii) $|x(t')-x(t'')|\leqslant\omega(|t'-t''|),\quad\forall x(t)\in\Phi.$

对任一区间 $[t_1,t_2]$ 和数 $\eta>0$，交集 $[t_1,t_2]\cap[T_1,T_2]$ 可分成

$$p\leqslant\frac{T_2-T_1}{\eta}+1$$

个长度不大于 η 的子区间。Φ 中任一函数 $x(t)$ 在这些子区间上的振幅不超过 $\omega(\eta)$。根据假定，$x(t)$ 在 $[t_1,t_2]$ 位于交集之右或之左的部分集合（如果存在的话）上的振幅为零。因而，任意区间 $[t_1, t_2]$ 可以分成

$$p\leqslant\frac{T_2-T_1}{\eta}+3$$

个子区间，在每个子区间上函数 $x(t)\in\Phi$ 的振幅都不超过 $\omega(\eta)$。

设对 $F(t,x)\in M$, $x(t)\in\Phi$, 有

$$\max_{t',t''\in R}\left|\int_{t'}^{t''}F(t,x(t))dt\right|=\left|\int_{t_1}^{t_2}F(t,x(t))dt\right|,$$

其中，t_1 和 t_2 是使最大值达到的两个时间端点。按上述方式将此区间 $[t_1,t_2]$ 划分成

$$p\leqslant\frac{T_2-T_1}{\eta}+3$$

个子区间，记诸分点为

$$t_1=\theta_0\leqslant\theta_1\leqslant\cdots\leqslant\theta_{p-1}\leqslant\theta_p=t_2,$$

就有

$$\begin{aligned}\left|\int_{t_1}^{t_2}F(t,x(t))dt\right|&\leqslant\sum_{i=0}^{p-1}\int_{\theta_i}^{\theta_{i+1}}|F(t,x(t))-F(t,x(\theta_i))|dt\\&\quad+\sum_{i=0}^{p-1}\left|\int_{\theta_i}^{\theta_{i+1}}F(t,x(\theta_i))dt\right|\\&\leqslant\sum_{i=0}^{p-1}\int_{\theta_i}^{\theta_{i+1}}L_F(t)|x(t)-x(\theta_i)|dt+p\|F(t,x)\|_w\\&\leqslant C\omega(\eta)+\left(\frac{T_2-T_1}{\eta}+3\right)\|F(t,x)\|_w,\end{aligned}$$

其中，C 见于(4.4).

如果 $\|F(t,x)\|_w=0$，则

$$\left|\int_{t_1}^{t_2}F(t,x(t))dt\right|\leqslant C\omega(\eta).$$

于是，由 $\omega(\eta)\to0$ $(\eta\to0)$，知上式左侧为零。如果 $\|F(t,x)\|_w>0$，则取

$$\eta=(\|F(t,x)\|_w)^{1/2},$$

从而，

$$\left|\int_{t_1}^{t_2} F(t,x(t))dt\right| \leqslant C\omega(\|F(t,x)\|_w)^{1/2}$$
$$+(T_2-T_1)(\|F(t,x)\|_w)^{1/2}+3\|F(t,x)\|_w.$$
由于所得不等式之右侧随 $\|F(t,x)\|_w$ 趋于零而收敛于零，故对任意的 $\varepsilon>0$，必存在 $\eta_1>0$，使得$\|F(t,x)\|_w\leqslant\eta_1$ 蕴涵
$$\max_{t',t''\in R}\left|\int_{t'}^{t''} F(t,x(t))dt\right|\leqslant\varepsilon, \quad \forall x(t)\in\Phi.$$
因此，取 $\delta=\eta_1$ 即可．证毕．

4.2 压缩映射的不动点定理

设 X 是一距离空间，ρ 是X上的距离函数，φ 为空间 X 到自身的（不要求连续）映射．

定义映射 φ 的迭代序列如下：
$$\varphi^{(1)}=\varphi, \qquad \varphi^{(p)}=\varphi\circ\varphi^{(p-1)}=\varphi^{(p-1)}\circ\varphi.$$
即
$$\varphi^{(1)}(x)=\varphi(x), \quad \varphi^{(p)}(x)=\varphi(\varphi^{(p-1)}(x))=\varphi^{(p-1)}(\varphi(x)), \quad \forall x\in X.$$
如果存在正数 $k<1$，使得
$$\rho(\varphi(x),\varphi(y))\leqslant k\rho(x,y), \quad \forall x,y\in X,$$
则称映射 φ 为压缩映射，或简称压缩，k 称为压缩系数．

设 Σ 为任一拓扑空间，其中的点记为 σ．依赖于 $\sigma\in\Sigma$ 的一族映空间 X 到其自身的映射 $\varphi(x,\sigma)$
$$\varphi(\cdot,\sigma):X\to X, \quad \sigma\in\Sigma$$
称为连续依赖于 $\sigma\in\Sigma$，如果对每一固定的 $x\in X$，映射
$$\varphi(x,\cdot):\ \Sigma\to X$$
是连续的．如果存在不依赖于 $\sigma\in\Sigma$ 的正数 $k<1$，使得对每一固定的 $\sigma\in\Sigma$，映射
$$\varphi(\cdot,\sigma):X\to X$$
都是以 k 为压缩系数的压缩，即
$$\rho(\varphi(x,\sigma),\varphi(y,\sigma))\leqslant k\rho(x,y), \quad \forall x,y\in X \quad 且 \quad \forall\sigma\in\Sigma,$$

则称这族映射为一致压缩映射族。

定理 4.1 设 X 为完备距离空间，Σ 是一拓扑空间，又设 $\varphi(x,\sigma)$ $(\sigma\in\Sigma)$ 是任意一族从 X 到其自身的映射。假定这族映射的某次迭代

$$\varphi^{(p)}(\cdot,\sigma):X\longrightarrow X,\quad \sigma\in\Sigma,$$

连续依赖于 $\sigma\in\Sigma$ 且是以 $k<1$ 为压缩系数的一致压缩，则该族中每一映射

$$\varphi(\cdot,\sigma):X\longrightarrow X,\quad \sigma\in\Sigma,$$

有唯一的不动点 x_σ，

$$\varphi(x_\sigma,\sigma)=x_\sigma,$$

且 x_σ 连续依赖于参数 $\sigma\in\Sigma$，即映射 $\sigma\longmapsto x_\sigma$ 是拓扑空间 Σ 到 X 的连续映射。

证 根据标准形式的压缩映射原理，对每一 $\sigma\in\Sigma$，迭代 $\varphi^{(p)}$ 存在唯一的不动点 x_σ：

$$\varphi^{(p)}(x_\sigma,\sigma)=x_\sigma.$$

对上面的等式两边作用映射 $\varphi(\cdot,\sigma)$，即得

$$\varphi(\varphi^{(p)}(x_\sigma,\sigma),\sigma)=\varphi^{(p)}(\varphi(x_\sigma,\sigma),\sigma)=\varphi(x_\sigma,\sigma).$$

这表明，除 x_σ 外，点 $\varphi(x_\sigma,\sigma)$ 也是映射 $\varphi^{(p)}(\cdot,\sigma)$ 的不动点。由不动点的唯一性可得

$$\varphi(x_\sigma,\sigma)=x_\sigma,$$

即 x_σ 是原映射 $\varphi(\cdot,\sigma)$ 的不动点。

映射 $\varphi(\cdot,\sigma)$ 没有其它的不动点。因为它的每个不动点必定也是各次迭代的不动点，但对每个 $\sigma\in\Sigma$，迭代 $\varphi^{(p)}$ 仅有唯一的不动点。

下面，我们证明 x_σ 对 σ 的连续依赖性。

为此，可写出如下不等式：

$$\begin{aligned}\rho(x_\sigma,x_{\hat\sigma})&=\rho(\varphi^{(p)}(x_\sigma,\sigma),\varphi^{(p)}(x_{\hat\sigma},\hat\sigma))\\&\leqslant\rho(\varphi^{(p)}(x_\sigma,\sigma),\varphi^{(p)}(x_{\hat\sigma},\sigma))+\rho(\varphi^{(p)}(x_{\hat\sigma},\sigma),\varphi^{(p)}(x_{\hat\sigma},\hat\sigma))\\&\leqslant k\rho(x_\sigma,x_{\hat\sigma})+\rho(\varphi^{(p)}(x_{\hat\sigma},\sigma),\varphi^{(p)}(x_{\hat\sigma},\hat\sigma))\end{aligned}$$

因此，

$$\rho(x_\sigma, x_{\hat\sigma}) \leqslant \frac{1}{1-k}\rho(\varphi^{(p)}(x_{\hat\sigma},\sigma),\varphi^{(p)}(x_{\hat\sigma},\hat\sigma)).$$

由 $\varphi^{(p)}$ 对参数的连续依赖性，即对任一 $\varepsilon>0$，必存在 $\hat\sigma$ 的一个邻域 $V_{\hat\sigma}$，使得对任意的 $\sigma\in V_{\hat\sigma}$ 有

$$\rho(\varphi^{(p)}(x_{\hat\sigma},\sigma),\varphi^{(p)}(x_{\hat\sigma},\hat\sigma))\leqslant\varepsilon(1-k).$$

因此，

$$\rho(x_\sigma,x_{\hat\sigma})\leqslant\varepsilon,\quad \forall\sigma\in V_{\hat\sigma}.$$

这就证明了 x_σ 在任一点 $\hat\sigma\in\Sigma$ 的连续性。证毕。

最后，我们说，任一不动点 x_σ 可由逐次逼近法求得：从任意一点 $x_0\in X$ 开始，构作序列

$$x_0, x_1=\varphi(x_0,\sigma),\quad x_2=\varphi(x_1,\sigma)=\varphi^{(2)}(x_0,\sigma),\cdots,$$
$$x_{i+1}=\varphi(x_i,\sigma)=\varphi^{(i+1)}(x_0,\sigma),\cdots,$$

则

$$x_i=\varphi^{(i)}(x_0,\sigma)\longrightarrow x_\sigma\quad(i\to\infty).$$

这里，值得注意的是，该逼近序列 $x_0,x_1,\cdots$是用原映射（它本身可能并非压缩）$\varphi(\cdot,\sigma)$构造出来的，而不是用迭代映射 $\varphi^{(p)}(\cdot,\sigma)$作出的。为了证明这一逼近序列确实收敛于不动点 x_σ，只要注意如下 p 个序列的每一个都收敛，

$$x_0, x_p=\varphi^{(p)}(x_0,\sigma),\cdots,\quad x_{ip}=\varphi^{(p)}(x_{(i-1)p},\sigma),\cdots;$$
$$x_1, x_{p+1}=\varphi^{(p)}(x_1,\sigma),\cdots,\quad x_{ip+1}=\varphi^{(p)}(x_{(i-1)p+1},\sigma),\cdots;$$
$$\cdots\cdots;$$
$$x_{p-1}, x_{2p-1}=\varphi^{(p)}(x_{p-1},\sigma),\cdots,\ x_{ip+p-1}=\varphi^{(p)}(x_{(i-1)p+p-1},\sigma),\cdots;$$

而且，按标准的不动点原理，这 p 个序列都收敛到同一极限 x_σ。因此 $x_i\to x_\sigma$ $(i\to\infty)$.

4.3 方程(4.3)解的存在性与连续依赖性定理

定理 4.2 方程

$$\dot x=F(t,x),\quad F(t,x)\in E_{\mathrm{Lip}},$$

存在唯一的满足任意给定的初始条件

$$x(\tau)=x_\tau,\quad (\tau,x_\tau)\in \boldsymbol{R}\times\boldsymbol{R}^n,$$

的解。此解 $x(t;\tau,x_\tau,F)$ 在参数 $\sigma=(\tau,x_\tau,F)$ 集合

$$\boldsymbol{R}\times\boldsymbol{R}^n\times M=\boldsymbol{R}\times\boldsymbol{R}^n\times E_{\text{Lip}}=E_\sigma$$

上连续地依赖于初始数据 (τ,x_τ) 及右端 $F(t,x)$，这里 M 是 E_{Lip} 中任一固定的一致 Lipschitz 子集，集合 $\boldsymbol{R}\times\boldsymbol{R}^n\times M$ 的拓扑由 E_σ 上的半范 $\|\cdot\|_w$ 给出。

在证明定理 4.2 之前，先对连续依赖性作更详细的说明：M 是 E_{Lip} 中的一致 Lipschitz 子集，

$$\hat{x}(t)=x(t;\hat{\tau},\hat{x}_{\hat{\tau}},\hat{F}),\ t\in\boldsymbol{R},$$

是微分方程

$$\dot{x}=\hat{F}(t,x),\quad \hat{F}(t,x)\in M,$$

以

$$\hat{x}(\hat{\tau})=\hat{x}_{\hat{\tau}}$$

为初值的解。这里的连续依赖性是说，对任一 $\varepsilon>0$，存在数 $\delta>0$，对满足条件

$$|\tau-\hat{\tau}|+|x_\tau-\hat{x}_{\hat{\tau}}|+\|F(t,x)-\hat{F}(t,x)\|_w\leqslant\delta$$

的任一点

$$(\tau,x_\tau,F(t,x))\in\boldsymbol{R}\times\boldsymbol{R}^n\times M,$$

所对应的解 $x(t;\tau,x_\tau,F)$ 满足

$$\max_{t\in \boldsymbol{R}}|x(t;\tau,x_\tau,F)-x(t;\hat{\tau},\hat{x}_{\hat{\tau}},\hat{F})|\leqslant\varepsilon.$$

应当注意，在定理 4.2 中，$F(t,x)$ 是遍取于 E_{Lip} 的某一个（尽管是任意的）给定的一致 Lipschitz 子集，而不是整个 E_{Lip} 空间。

作为本定理的直接推论，我们有如下结果：如果

$$\|F_1(t,x)-F_2(t,x)\|_w=0,$$

则对一切 $(\tau,x_\tau)\in\boldsymbol{R}\times\boldsymbol{R}^n$，

$$\max_{t\in \boldsymbol{R}}|x(t;\tau,x_\tau,F_1)-x(t;\tau,x_\tau,F_2)|=0,$$

即两个解 $x(t;\tau,x_\tau,F_1)$ 与 $x(t;\tau,x_\tau,F_2)$ 重合。

证　我们用简单而典型的处理办法将此定理归结为定理 4.1. 令 C^n 为所有定义在 t 轴上的当 $t\to\pm\infty$ 时存在有限极限的 n 维连续函数组成的线性空间，并在 C^n 中引进范数

$$\|x(\cdot)\| = \sup_{t\in \boldsymbol{R}}|x(t)|,\quad x(t)\in C^n,$$

则 C^n 按距离

$$\rho(x(\cdot),y(\cdot)) = \|x(\cdot)-y(\cdot)\|$$

成为完备距离空间。我们定义一族将 C^n 映入自身的依赖于参数

$$\sigma = (\tau,x_\tau,F(t,x))\in \boldsymbol{R}\times\boldsymbol{R}^n\times E_{\mathrm{Lip}}$$

的映射 $\varphi(x(\cdot);\sigma)$ 为

$$x(\cdot)\longmapsto\varphi(x(\cdot);\sigma) = \varphi(x(\cdot);\tau,x_\tau,F)$$
$$= x_\tau+\int_\tau^t F(\theta,x(\theta))d\theta = y(t;x(\cdot);\sigma),\qquad t\in\boldsymbol{R}$$

这一映射的每一个不动点

$$x_\sigma(t) = x(t;\sigma) = x_\tau+\int_\tau^t F(\theta,x(\theta;\sigma))d\theta,\quad t\in\boldsymbol{R}.$$

就是微分方程

$$\dot{x} = F(t,x)$$

以 $x(\tau;\sigma)=x_\tau$ 为初值的解。因此。若我们能够证明，对任一给定的一致 Lipschitz 子集 $M\subset E_{\mathrm{Lip}}$，当 σ 遍取于集合

$$\Sigma = \boldsymbol{R}\times\boldsymbol{R}^n\times M\subset E_\sigma \tag{4.5}$$

时，某一次迭代 $\varphi^{(p)}(x(\cdot);\sigma)$ 是 C^n 到自身内的一致压缩且连续的映射族，则定理 4.2 将由定理 4.1 获证。这里，按我们的约定记法，收敛性 $\sigma\to\hat{\sigma}$ 意为

$$\|\sigma-\hat{\sigma}\|_W = |\tau-\hat{\tau}| + |x_\tau-\hat{x}_{\hat{\tau}}| + \|F(t,x)-\hat{F}(t,x)\|_W\longrightarrow 0.$$

现在，我们用归纳法证明任何一次迭代 $\varphi^{(p)}(x(\cdot);\sigma)$ $(p=1,2,\cdots)$ 都在集(4.5)上连续依赖于 σ，且当 p 充分大时，$\varphi^{(p)}(x(\cdot);\sigma)$ 是集(4.5)上的一致压缩映射族。

为证连续性，我们利用记号

$$\varphi(x(\cdot),\sigma) = y_1(t;x(\cdot),\sigma) = x_\tau+\int_\tau^t F(\theta,x(\theta))d\theta,$$

$$\varphi^{(p+1)}(x(\cdot),\sigma)=y_{p+1}(t;x(\cdot),\sigma)$$
$$=x_\tau+\int_\tau^t F(\theta,y_p(\theta;x(\cdot),\sigma))d\theta,$$
$$p=1,2,\cdots,$$

并写出如下明显的估计式：

$$|y_1(t;x(\cdot),\sigma)-y_1(t;x(\cdot),\hat{\sigma})|\leqslant|x_\tau-\hat{x}_{\hat{\tau}}|+\left|\int_\tau^{\hat{\tau}}|\hat{F}(\theta,x(\theta))|d\theta\right|$$
$$+\left|\int_\tau^t\{F(\theta,x(\theta))-\hat{F}(\theta,x(\theta))\}d\theta\right|.$$

其中，右侧的前两项显然当 $\sigma\to\hat{\sigma}$ 时趋于零．因此，为了证得 $\varphi(x(\cdot),\sigma)$ 关于 $\sigma\in\Sigma$ 的连续依赖性，只须证明对集合 M 中的 $F(t,x)$，当

$$\|F(t,x)-\hat{F}(t,x)\|_w\longrightarrow 0$$

时，下式成立，

$$J=\max_{\theta_1,\theta_2\in R}\left|\int_{\theta_1}^{\theta_2}\{F(\theta,x(\theta))-\hat{F}(\theta,x(\theta))\}d\theta\right|\longrightarrow 0. \quad (4.6)$$

由于当 $t\to\pm\infty$ 时函数 $x(t)$ 有有限的极限，对任意 $\varepsilon>0$，存在区间 $[t_1,t_2]$，当修改这一函数 $x(t)$ 为 $\hat{x}(t)$ 时，

$$\hat{x}(t)\doteq\begin{cases}x(t_1), & t\leqslant t_1;\\ x(t), & t\in[t_1,t_2];\\ x(t_2), & t\geqslant t_2,\end{cases}$$

有

$$\max_{t\in R}|\hat{x}(t)-x(t)|=\|\hat{x}(t)-x(t)\|\leqslant\varepsilon.$$

因此，

$$J\leqslant\max_{\theta_1,\theta_2\in R}\left|\int_{\theta_1}^{\theta_2}\{F(\theta,\hat{x}(\theta))-\hat{F}(\theta,\hat{x}(\theta))\}d\theta\right|$$
$$+\|\hat{x}(t)-x(t)\|\left\{\int_R(L_F(t)+L_{\hat{F}}(t))dt\right\}$$
$$\leqslant\max_{\theta_1,\theta_2\in R}\left|\int_{\theta_1}^{\theta_2}\{F(\theta,\hat{x}(\theta))-\hat{F}(\theta,\hat{x}(\theta))\}d\theta\right|+2C\varepsilon,$$

其中，C 是在关于相应的一致 Lipschitz 集 M 的(4.4)中提到的常数，这一常数并不依赖于 F 或 $\hat{F}$。根据命题 4.1，当 $\|F(t,x)-\hat{F}(t,x)\|_w\to 0$ 时，上式右端第一项收敛于零。但 $\varepsilon>0$ 是任意的，故当 $F(t,x)$

保持在集 M 内并且 $\|F(t,x)-\hat{F}(t,x)\|_w\to 0$ 时，有 $J\to 0$。这就证明了映射族 $\varphi(x(\cdot),\sigma)$ 在形如（4.5）的任一给定集合上关于 σ 的连续性。

现假定 $\varphi^{(p)}(x(\cdot),\sigma)$ 关于 $\sigma\in\Sigma$ 连续，我们来证明映射族 $\varphi^{(p+1)}(x(\cdot),\sigma)$ 的连续性。可写出如下估计式：

$$
\begin{aligned}
&|y_{p+1}(t;x(\cdot),\sigma)-y_{p+1}(t;x(\cdot),\hat{\sigma})|\\
\leqslant&|x_\tau-\hat{x}_{\hat{\tau}}|+\left|\int_\tau^{\hat{\tau}}|\hat{F}(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))|d\theta\right|\\
&+\left|\int_\tau^t|F(\theta,y_p(\theta;x(\cdot),\sigma))-F(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))|d\theta\right|\\
&+\left|\int_\tau^t\{F(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))-\hat{F}(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))\}d\theta\right|\\
\leqslant&|x_\tau-\hat{x}_{\hat{\tau}}|+\left|\int_\tau^{\hat{\tau}}|\hat{F}(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))|d\theta\right|+\|y_p(t;x(\cdot),\sigma)\\
&-y_p(t;x(\cdot),\hat{\sigma})\|\int_E L_F(t)dt+\left|\int_\tau^t\{F(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))\right.\\
&\left.-\hat{F}(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))\}d\theta\right|.
\end{aligned}
$$

显然，当 $\sigma\to\hat{\sigma}$ 时这一不等式右端前两项收敛于零。按归纳假设

$$\|y_p(t;x(\cdot),\sigma)-y_p(t;x(\cdot),\hat{\sigma})\|\longrightarrow 0$$

以及积分

$$\int_E L_F(t)dt$$

的一致有界性，第三项也收敛于零。最后，可以用证明（4.6）时完全相同的方法证得，对 $F(t,x)\in M$，当 $\|F(t,x)-\hat{F}(t,x)\|_w\to 0$ 时，

$$\max_{\theta_1,\theta_2\in E}\left|\int_{\theta_1}^{\theta_2}\{F(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))-\hat{F}(\theta,y_p(\theta;x(\cdot),\hat{\sigma}))\}d\theta\right|\longrightarrow 0.$$

这就结束了连续性的证明。

再证，对充分大的 p，当参数 σ 取自形如(4.5)的任一给定集合 Σ 时，$\varphi^{(p)}(x(\cdot),\sigma)$ 是空间 C^n 到自身内的一致压缩的映射族。为此，估计差值

$$\|\varphi^{(p)}(x'(\cdot),\sigma)-\varphi^{(p)}(x''(\cdot),\sigma)\|$$
$$=\|y_p(t;x'(\cdot),\sigma)-y_p(t;x''(\cdot),\sigma)\|,\quad x'(t),x''(t)\in C^n,$$
$$\sigma\in R\times R^n\times M,$$

我们有

$$|y_p(t;x'(\cdot),\sigma)-y_p(t;x''(\cdot),\sigma)|$$
$$\leqslant\left|\int_\tau^t|F(\theta,y_{p-1}(\theta;x',\sigma))-F(\theta,y_{p-1}(\theta;x'',\sigma))|d\theta\right|$$
$$\leqslant\left|\int_\tau^t L_F(\theta)|y_{p-1}(\theta;x',\sigma)-y_{p-1}(\theta;x'',\sigma)|d\theta\right|$$
$$=\left|\int_\tau^t L_F(\theta_1)|y'_{p-1}(\theta_1)-y''_{p-1}(\theta_1)|d\theta_1\right|\text{（记}y'_{p-1}(\theta)\equiv y_{p-1}(\theta,x',\sigma)\text{）}$$
$$\leqslant\left|\int_\tau^t L_F(\theta_1)d\theta_1\int_\tau^{\theta_1}L_F(\theta_2)|y'_{p-2}(\theta_2)-y''_{p-2}(\theta_2)|d\theta_2\right|\leqslant\cdots$$
$$\leqslant\left|\int_\tau^t L_F(\theta_1)d\theta_1\int_\tau^{\theta_1}L_F(\theta_2)\int_\tau^{\theta_2}\cdots\int_\tau^{\theta_{p-1}}L_F(\theta_p)|x'(\theta_p)-x''(\theta_p)|d\theta_p\right|$$
$$\leqslant\|x'(t)-x''(t)\|\left|\int_\tau^t L_F(\theta_1)d\theta_1\int_\tau^{\theta_1}L_F(\theta_2)d\theta_2\int_\tau^{\theta_2}\cdots\int_\tau^{\theta_{p-1}}L_F(\theta_p)d\theta_p\right|.$$

易用归纳法证得

$$I_p(t,\tau)=\int_\tau^t L_F(\theta_1)d\theta_1\int_\tau^{\theta_1}L_F(\theta_2)d\theta_2\int_\tau^{\theta_2}\cdots\int_\tau^{\theta_{p-1}}L_F(\theta_p)d\theta_p$$
$$=\frac{1}{p!}\left(\int_\tau^t L_F(\theta)d\theta\right)^p.$$

事实上，由分部积分即得

$$I_{p+1}(t,\tau)=\int_\tau^t L_F(\theta)I_p(\theta,\tau)d\theta$$
$$=\int_\tau^t L_F(\theta)\frac{1}{p!}\left(\int_\tau^\theta L_F(\theta')d\theta'\right)^p d\theta$$
$$=\frac{1}{p!}\left(\int_\tau^t L_F(\theta')d\theta'\right)^p\int_\tau^t L_F(\theta')d\theta'$$
$$-\frac{1}{p!}\int_\tau^t p\left(\int_\tau^\theta L_F(\theta')d\theta'\right)^{p-1}L_F(\theta)\left(\int_\tau^\theta L_F(\theta')d\theta'\right)d\theta$$
$$=\frac{1}{p!}\left(\int_\tau^t L_F(\theta')d\theta'\right)^{p+1}-\frac{1}{(p-1)!}\int_\tau^t L_F(\theta)\left(\int_\tau^\theta L_F(\theta')d\theta'\right)^p d\theta$$

$$= \frac{1}{p!}\left(\int_\tau^t L_F(\theta')d\theta'\right)^{p+1} - \frac{p!}{(p-1)!}\int_\tau^t L_F(\theta)I_p(\theta,\tau)d\theta$$

$$= \frac{1}{p!}\left(\int_\tau^t L_F(\theta)d\theta\right)^{p+1} - pI_{p+1}(t,\tau).$$

因此，

$$I_{p+1}(t,\tau) = \frac{1}{(p+1)!}\left(\int_\tau^t L_F(\theta)d\theta\right)^{p+1}$$

于是，我们可写出估计式

$$\|y_p(t,x',\sigma) - y_p(t,x'',\sigma)\| \leqslant \|x'(t) - x''(t)\| \frac{1}{p!}\left(\int_R L_F(\theta)d\theta\right)^p.$$

因为

$$\int_R L_F(\theta)d\theta \leqslant \text{const},$$

其中的常数不依赖于 $F(t,x)\in M$，故得最后的估计

$$\rho(\varphi^{(p)}(x'(\cdot),\sigma),\ \varphi^{(p)}(x''(\cdot),\sigma)) \leqslant \frac{\text{const}^p}{p!}\rho(x'(\cdot),x''(\cdot)),$$

$$\forall x'(t),x''(t)\in C^n \text{ 且 } \forall\sigma = (\tau,x_\tau,F(t,x))\in \boldsymbol{R}\times\boldsymbol{R}^n\times M.$$

证毕.

4.4 空间 $E_{\text{Lip}}(G)$

在叙述一般情形的存在性和连续依赖性定理之前，我们给出几个必要的定义.

设 G 是 $\boldsymbol{R}\times\boldsymbol{R}^n$ 中的任一开子集. 以 $E_{\text{Lip}}(G)$ 记所有定义在 G 上并具有以下性质的函数 $F(t,x)$ 所组成的线性空间：对一切定义在 $\boldsymbol{R}\times\boldsymbol{R}^n$ 上且具紧支集在 G 中的数值连续可微函数 $a(t,x)$，都有

$$a(t,x)F(t,x)\in E_{\text{Lip}}.$$

其中函数 $a(t,x)F(t,x)$ 在 $a(t,x)$ 的支集之外等于零.

对每一紧集 $K\subset G$，类似于 E_{Lip} 上的半范，我们来定义 $E_{\text{Lip}}(G)$ 上

的一个半范$\|\cdot\|_{\omega,K}$.设$\alpha_K(t,x)$是定义在$\boldsymbol{R}\times\boldsymbol{R}^n$上的取值于0和1之间的一个连续可微函数，

$$0\leqslant\alpha_K(t,x)\leqslant1,$$

它在K上取值为1，且有紧支集在G内。对每一个$F(t,x)\in E_{\mathrm{Lip}}(G)$函数$\alpha_K(t,x)F(t,x)\in E_{\mathrm{Lip}}$。因此，半范$\|\alpha_K(t,x)F(t,x)\|_\omega$是有定义的。令

$$\|F(t,x)\|_{\omega,K}=\inf_{\alpha_K}\|\alpha_K(t,x)F(t,x)\|_\omega,$$

这里的下确界是对所有具备所述性质的$\alpha_K(t,x)$来取的。显然，若$K_1\subset K_2\subset G$，则

$$\|F(t,x)\|_{\omega,K_1}\leqslant\|F(t,x)\|_{\omega,K_2},\quad\forall F(t,x)\in E_{\mathrm{Lip}}(G).$$

如果函数$F(t,x)\in E_{\mathrm{Lip}}$有紧支集$K\subset G$，则$F(t,x)$在$G$上的限制必属于$E_{\mathrm{Lip}}(G)$，仍用$F(t,x)$记之，并且

$$\|F(t,x)\|_{\omega,K}=\|F(t,x)\|_\omega.$$

以下总记$\alpha_K(t,x)$为任意一个连续可微的数值函数，它有位于所述集合G内的紧支集，在0与1之间取值，且在紧集K上取值为1。只有当这样的函数有附加的性质时才另作说明。

下面的定义类似于一致Lipschitz集合的定义。一函数集合$M\subset E_{\mathrm{Lip}}(G)$称为在集合$K\subset G$的一个邻域上是一致Lipschitz的，如果存在集K的一个紧邻域$V_K\subset G$，使得对任一支集含于V_K中的数值连续可微函数$\alpha(t,x)$，集合

$$\alpha(t,x)M=\{\alpha(t,x)F(t,x)\mid F(t,x)\in M\}\subset E_{\mathrm{Lip}}$$

是E_{Lip}中的一致Lipschitz子集。

一个集合$M\subset E_{\mathrm{Lip}}(G)$称为在$E_{\mathrm{Lip}}(G)$中是一致Lipschitz的，如果它在任一紧集$K\subset G$的一个邻域上是一致Lipschitz的。如果集合$M\subset E_{\mathrm{Lip}}(G)$中的所有函数都具有位于$G$内的紧支集，则$M\subset E_{\mathrm{Lip}}$。一个集合$M$可以在$E_{\mathrm{Lip}}$内是一致Lipschitz的，但若对$M$没有附加约束，它不一定在$E_{\mathrm{Lip}}(G)$中是一致Lipschitz的。

下面谈到的集合M的积分一致有界性，保证了这一集合在$E_{\mathrm{Lip}}(G)$中的一致Lipschitz性可由它在E_{Lip}中的一致Lipschitz性推出。集合

M 称为积分一致有界，是指对每一函数 $F(t,x)\in M$，总有一控制函数 $m_F(t)$，

$$|F(t,x)|\leqslant m_F(t),\quad \forall(t,x)\in \boldsymbol{R}\times\boldsymbol{R}^n,$$

其中

$$\int_{R} m_F(t)dt\leqslant C,$$

常数 C 不依赖于个别的 $F(t,x)\in M$。

现来证明，集 M 在 E_{Lip} 中的一致 Lipschitz 性和它的积分一致有界性蕴涵它在 $E_{\mathrm{Lip}}(G)$ 中的一致 Lipschitz 性。事实上，对每一个有紧支集落在 G 中且连续可微的数值函数 $\alpha(t,x)$，存在常数 D，使得

$$|\alpha(t,x')-\alpha(t,x'')|\leqslant D|x'-x''|,\ \forall(t,x'),(t,x'')\in G.$$

因此，对每个 $F(t,x)\in M$，有如下的估计：

$$\begin{aligned}&|\alpha(t,x')F(t,x')-\alpha(t,x'')F(t,x'')|\\ &\leqslant|\alpha(t,x')|\,|F(t,x')-F(t,x'')|+|\alpha(t,x')-\alpha(t,x'')|\,|F(t,x'')|\\ &\leqslant\max_{(t,x)}|\alpha(t,x)|\,|x'-x''|L_F(t)+D|x'-x''|m_F(t).\end{aligned}$$

由于

$$\max_{(t,x)}|\alpha(t,x)|\int_R L_F(t)dt+D\int_R m_F(t)dt$$

是一个不依赖于 $F(t,x)\in M$ 的特殊选取的常数，故 M 在 $E_{\mathrm{Lip}}(G)$ 中是一致 Lipschitz 的。

在第六章将给出一个对控制问题很重要的 $E_{\mathrm{Lip}}(G)$ 的一致 Lipschitz 性判据。

定义在区间 $[t_1,t_2]$ 上的一个 n 维函数 $x(t)$ 可等同于曲线

$$\{(t,x(t)):t\in[t_1,t_2]\}\subset\boldsymbol{R}\times\boldsymbol{R}^n,$$

若函数 $x(t)$ 在 $[t_1,t_2]$ 上是绝对连续的，就说这条曲线是绝对连续的。如果

$$(t,x(t))\in G,\quad \forall t\in[t_1,t_2],$$

就说曲线 $x(t)\,(t\in[t_1,t_2])$ 位于 G 内。

任一位于 G 内、绝对连续的曲线 $x(t)\,(t\in[t_1,t_2])$，如果对几乎所

有的 $t\in[t_1,t_2]$满足等式

$$\dot{x}(t)=F(t,x(t)),$$

则称它为微分方程

$$\dot{x}=F(t,x),\qquad F(t,x)\in E_{\text{Lip}}(G),\tag{4.7}$$

的一个解，或一条轨线。

按解的定义，积分方程

$$x(t)=x_\tau+\int_\tau^t F(\theta,x(\theta))d\theta$$

与微分方程(4.7)以及初始条件

$$x(\tau)=x_\tau,\quad (\tau,x_\tau)\in G,$$

是等价的。

设方程（4.7)以及它的一个解 $\tilde{x}(t)(t\in[t_1,t_2])$已给出．每一个形为

$$\dot{x}=F(t,x)+\delta F(t,x),\quad \delta F(t,x)\in E_{\text{Lip}}(G),$$

的方程称为摄动方程，右端的函数 $\delta F(t,x)$ 称为一个摄动。摄动方程的任一解 $x(t)(t\in[t_1,t_2])$称为摄动解。两者之差

$$\Delta x(t)=x(t)-\tilde{x}(t),\qquad t\in[t_1,t_2],$$

称为在区间$[t_1,t_2]$上原解 $\tilde{x}(t)$的一个摄动。

4.5 一般情形微分方程解的存在性与连续依赖性定理

定理 4.3 设 K 是 G 中任一紧集，则必存在正数 $\eta>0$，使得微分方程

$$\dot{x}=F(t,x),\qquad F(t,x)\in E_{\text{Lip}}(G)\tag{4.8}$$

以$(\tau,x_\tau)\in K$ 为初值点的任一解 $x(t)$ 在区间 $\tau-\eta\leqslant t\leqslant\tau+\eta$ 上存在且唯一。

证 设 V_K 是该紧集 K 的任一个位于G 内的紧邻域。取函数 $\alpha_{V_K}(t,x)$,并考虑微分方程

$$\dot{x}=\alpha_{V_K}(t,x)F(t,x).\tag{4.9}$$

这一方程以$(\tau,x_\tau)\in K$为初值点的任一解记为

$$x(t)=x(t;\tau,x_\tau),\quad t\in\boldsymbol{R}.$$

根据定理 4.2，这样的解存在且唯一。

因$(\tau,x_\tau)\in K$，故函数$\alpha_{V_K}(t,x)$在点(τ,x_τ)的某个邻域上取值为 1。从而存在数$\Delta>0$，使得函数$x(t)$对几乎所有的$t\in[\tau-\Delta,\tau+\Delta]$满足等式

$$\dot{x}(t)=F(t,x(t)).$$

我们的目的是要证明可以选出对所有初值点$(\tau,x_\tau)\in K$都是同样的数$\Delta>0$。

设函数$\alpha_{V_K}(t,x)F(t,x)$的控制函数为$m(t)$：

$$|\alpha_{V_K}(t,x)F(t,x)|\leqslant m(t),\quad \forall(t,x)\in\boldsymbol{R}\times\boldsymbol{R}^n.$$

K到V_K之边界的距离记为d。取$\eta>0$充分小，以致$\eta<d/2$，并且

$$\int_{t-\eta}^{t+\eta}m(\theta)d\theta\leqslant\frac{d}{2},\quad \forall t\in\boldsymbol{R}.$$

当$|t-\tau|\leqslant\eta$时，可得估计

$$|x(t;\tau,x_\tau)-x_\tau|=\left|\int_\tau^t \alpha_{V_K}(\theta,x(\theta;\tau,x_\tau))F(\theta,x(\theta,\tau,x_\tau))d\theta\right|$$

$$\leqslant\int_{\tau-\eta}^{\tau+\eta}m(\theta)\,d\theta\leqslant\frac{d}{2}.$$

但K到V_K边界之距离是d，故

$$(t,x(t;\tau,x_\tau))\in V_K,\quad \forall t\in[\tau-\eta,\tau+\eta],$$

这与点$(\tau,x_\tau)\in K$的特殊选取无关。从等式

$$\alpha_{V_K}(t,x)=1,\quad \forall(t,x)\in V_K,$$

我们得到，对一切$(\tau,x_\tau)\in K$，曲线

$$x(t)=x(t;\tau,x_\tau),\quad \tau-\eta\leqslant t\leqslant\tau+\eta,$$

是微分方程(4.8)以$x(\tau)=x_\tau$为初始条件的解。这个解的唯一性由方程(4.9)的解$x(t;\tau,x_\tau)(t\in\boldsymbol{R})$的唯一性直接得到。证毕。

定理 4.4 设

$$\bar{x}(t),\quad t_1\leqslant t\leqslant t_2,$$

是微分方程。

$$\dot{x}=F(t,x),\qquad F(t,x)\in E_{\rm Lip}(G),$$

以$(\tilde{\tau},\tilde{x}_{\tilde{\tau}})$为初值点的解，$V_{\tilde{x}(t)}\equiv K\subset G$是曲线$\tilde{x}(t)$ $(t_1\leqslant t\leqslant t_2)$ 的一个紧邻域，摄动集 $M\subset E_{\rm Lip}(G)$ 是 $K\equiv V_{\tilde{x}(t)}$ 上的一致 Lipschitz 的集合，则对任一 $\varepsilon>0$，存在 $\delta>0$，当初始数据(τ,x_τ)与摄动 $\delta F(t,x)\in M$ 满足条件

$$|\tau-\tilde{\tau}|+|x_\tau-\tilde{x}_{\tilde{\tau}}|+\|\delta F(t,x)\|_{w,K}\leqslant\delta$$

时，摄动方程

$$\dot{x}=F(t,x)+\delta F(t,x)\tag{4.10}$$

过(τ,x_τ)点的解 $x(t;\tau,x_\tau,\delta F)$也必在$[t_1,t_2]$上存在，且

$$\max_{t_1\leqslant t\leqslant t_2}|x(t;\tau,x_\tau,\delta F)-\tilde{x}(t)|\leqslant\varepsilon.$$

证 为方便起见，可假定 $\delta F(t,x)=0\in M$。事实上，对一个集合添加有限个 Lipschitz 函数不会改变它原有的一致 Lipschitz 性。

按条件，曲线$\tilde{x}(t)$ $(t\in[t_1,t_2])$存在一个紧邻域 $\widetilde{V}\subset G$，使得任一支集在$\widetilde{V}$ 内的连续可微数值函数$\alpha(t,x)$所对应的集合

$$\alpha(t,x)M=\{\alpha(t,x)\delta F(t,x):\delta F(t,x)\in M\}\subset E_{\rm Lip}$$

在 $E_{\rm Lip}$ 中是一致 Lipschitz 的。

再取曲线$\tilde{x}(t)$的一个紧邻域 W，W 位于邻域 $V_{\tilde{x}(t)}$ 与 $\widetilde{V}$ 两者内部的交集中。又令函数

$$\alpha_W(t,x)$$

是支集含于邻域 $V_{\tilde{x}(t)}$ 与$\widetilde{V}$的交集中的一个函数。因此。集合

$$\alpha_W(t,x)M=\{\alpha_W(t,x)\delta F(t,x):\delta F(t,x)\in M\}\subset E_{\rm Lip}\tag{4.11}$$

在 $E_{\rm Lip}$ 中是一致 Lipschitz 的。

任取一函数 $\alpha_{V_{\tilde{x}(t)}}(t,x)$。显然，

$$\alpha_W(t,x)=\alpha_{V_{\tilde{x}(t)}}(t,x)\alpha_W(t,x),\quad\forall(t,x)\in\boldsymbol{R}\times\boldsymbol{R}^n.$$

现来证明如下不等式：

$$\|a_W(t,x)\delta F(t,x)\|_w \leqslant \text{const}\|\delta F(t,x)\|_{w,V_{\tilde{x}(t)}},$$

其中的常数不依赖于 $\delta F(t,x)\in M$ 的选取。

我们有

$$\int_{t_1}^{t''} a_w(t,x)\delta F(t,x)dt = \int_{t_1}^{t''} a_W(t,x)a_{V_{\tilde{x}(t)}}(t,x)\,\delta F(t,x)dt$$

$$= a_W(t'',x)\int_{t_1}^{t''} a_{V_{\tilde{x}(t)}}(t,x)\delta F(t,x)dt$$

$$-\int_{t_1}^{t''}\frac{\partial a_W(t,x)}{\partial t}\left(\int_{t_1}^{t} a_{V_{\tilde{x}(t)}}(\theta,x)\delta F(\theta,x)d\theta\right)dt.$$

因此，

$$\left|\int_{t_1}^{t''} a_W(t,x)\delta F(t,x)dt\right| \leqslant \left|\int_{t_1}^{t''} a_{V_{\tilde{x}(t)}}(t,x)\delta F(t,x)dt\right|$$

$$+\left|\int_{t_1}^{t''}\left|\frac{\partial a_W(t,x)}{\partial t}\right|dt\right|\max_{\substack{\theta',\theta''\in R\\ x\in R^n}}\left|\int_{\theta'}^{\theta''} a_{V_{\tilde{x}(t)}}(t,x)\delta F(t,x)dt\right|$$

$$\leqslant\left(1+\int\left|\frac{\partial a_W(t,x)}{\partial t}\right|dt\right)\|a_{V_{\tilde{x}(t)}}(t,x)\delta F(t,x)\|_w$$

$$=\text{const}\|a_{V_{\tilde{x}(t)}}(t,x)\delta F(t,x)\|_w.$$

由于函数 $a_{V_{\tilde{x}(t)}}(t,x)$ 的任意性，按 $\|\cdot\|_{w,V_{\tilde{x}(t)}}$ 的定义便得如下不等式

$$\left|\int_{t_1}^{t''} a_W(t,x)\delta F(t,x)dt\right| \leqslant \text{const}\|\delta F(t,x)\|_{w,V_{\tilde{x}(t)}},$$

这等价于我们所要证的不等式。

将微分方程

$$\dot{x} = a_W(t,x)F(t,x) + a_W(t,x)\delta F(t,x),\quad \delta F(t,x)\in M,$$

满足初值条件 $x(\tau)=x_\tau$ 的解记为

$$x(t)=x(t;\tau,x_\tau,\delta F)\qquad t\in \boldsymbol{R}.$$

由于集合 (4.11) 是含有零函数的一致 Lipschitz 集，因此按定理 4.2，对任一 $\varepsilon>0$，存在 $\Delta(\varepsilon)>0$，当

$$|\tau-\tilde{\tau}| + |x_\tau - \tilde{x}_{\tilde{\tau}}| + \|a_W(t,x)\delta F(t,x)\|_w \leqslant \Delta(\varepsilon),$$

$$\delta F(t,x)\in M,$$

时，成立

$$\|x(t;\tau,x_\tau,\delta F)-x(t;\tilde{\tau},\tilde{x}_{\tilde{\tau}},0)\|\leqslant\varepsilon.$$

但是，对所有的 $t\in[t_1,t_2]$，有 $\alpha_W(t,\tilde{x}(t))=1$。因此，唯一性表明

$$x(t;\tilde{\tau},\tilde{x}_{\tilde{\tau}},0)=\tilde{x}(t),\quad\forall t\in[t_1,t_2],$$

其中，$\tilde{x}(t)$ 如定理条件所述。曲线 $\tilde{x}(t)(t\in[t_1,t_2])$ 到邻域 W 的边界的距离记为 d。只要 $\varepsilon\leqslant d$，就有如下的包含关系，

$$(t,x(t;\tau,x_\tau,\delta F))\in W,\quad\forall t\in[t_1,t_2].$$

既然 $\alpha_W(t,x)=1$，$\forall(t,x)\in W$，曲线

$$x(t;\tau,x_\tau,\delta F),\quad t\in[t_1,t_2],$$

当 $\varepsilon\leqslant d$ 时就是摄动方程(4.10)以 $x(\tau)=x_\tau$ 为初值的解。因此，对任一给定的 $\varepsilon>0$，可取充分小的 $\delta>0$ 使定理成立。证毕。

第五章　微分方程解的变分公式

5.1　空间E_1与$E_1(G)$

设E_1是定义在$\boldsymbol{R}\times\boldsymbol{R}^n$上满足如下条件的$n$维函数的线性空间：其中每一函数$F(t,x)\in E_1$具有紧支集，对固定的$t\in\boldsymbol{R}$关于$x$连续可微；对固定的$x$，函数

$$F(t,x),\qquad F_x(t,x)=\frac{\partial F(t,x)}{\partial x}$$

关于t可测，并且存在优函数$m_F(t)$，在$\boldsymbol{R}$上可积，使得 $|F(t,x)|+$

$$|F_x(t,x)|\leqslant m_F(t),\ \forall(t,x)\in\boldsymbol{R}\times\boldsymbol{R}^n,\int_R m_F(t)dt<\infty.\qquad(5.1)$$

我们要证明E_1中的函数都满足Lipschitz条件(4.2)，从而有包含关系$E_1\subset E_{\text{Lip}}$.事实上，

$$|F(t,x')-F(t,x'')|=\left|\int_0^1\frac{d}{ds}F(t,x''+s(x'-x''))ds\right|$$
$$\leqslant\int_0^1|F_x(t,x''+s(x'-x''))||x'-x''|ds\leqslant m_F(t)|x'-x''|,$$
$$\forall(t,x'),(t,x'')\in\boldsymbol{R}\times\boldsymbol{R}^n$$

设G是$\boldsymbol{R}\times\boldsymbol{R}^n$中的开子集，$E_1(G)$表示定义在$G$上并具有如下性质的函数$F(t,x)$所组成的线性空间：如$\alpha(t,x)$是$\boldsymbol{R}\times\boldsymbol{R}^n$上的连续可微的数量函数并且有紧支集在$G$中，则$\alpha(t,x)F(t,x)\in E_1$.

显然，有如下的包含关系：

$$E_1(G)\subset E_{\text{Lip}}(G).$$

因此，有关E_{Lip}与$E_{\text{Lip}}(G)$的所有记号与性质都可自然地转移于E_1与$E_1(G)$.

在E_1上定义半范$\|\cdot\|_1$为

$$\|F(t,x)\|_1=\int_{R}\max_{x\in R^n}(|F(t,x)|+|F_x(t,x)|)dt.$$

易证其中的被积函数是可积的，从而上述定义是合理的．为此，我们注意到，由于函数$|F(t,x)|+|F_x(t,x)|$关于x连续，因此若$x_1,x_2,\cdots$为R^n中的稠密点列，则

$$\max_{x\in R^n}(|F(t,x)|+|F_x(t,x)|)=\sup_{x_i}(|F(t,x_i)|+|F_x(t,x_i)|)$$

因上式右侧是对可列个关于t可测的函数取的上确界，故知其左侧函数必关于t可测。又由于存在优函数(5.1)，所以被积函数是可积的。

除了$\|\cdot\|_1$而外，在第四章中出现的半范$\|\cdot\|_w$也在E_1上有定义。显然，

$$\|F(t,x)\|_w\leqslant\|F(t,x)\|_1,\qquad \forall F(t,x)\in E_1\text{[注1]}.$$

对任一紧集$K\subset G$，我们定义$E_1(G)$上的半范$\|\cdot\|_{1,K}$为

$$\|F(t,x)\|_{1,K}=\inf\|\alpha_K(t,x)F(t,x)\|_1,$$

其中的下确界遍取所有函数$\alpha_K(t,x)$（该记号见第四章）。如下不等式显然成立：

$$\|F(t,x)\|_{w,K}\leqslant\|F(t,x)\|_{1,K},\qquad \forall F(t,x)\in E_1(G)\qquad \forall K\subset G.$$

如果我们考虑的是微分方程

$$\dot{x}=F(t,x),\qquad F(t,x)\in E_1(G)$$

则其右端函数的所有摄动自然取自$E_1(G)$．

现给出$E_1(G)$的一个子集在一给定的紧集$K\subset G$的一个邻域上为一致Lipschitz的简单判据。

命题5.1 设K是G的紧子集，且设$V_K\subset G$是K的任一紧邻域，则集合

$$M=\{F(t,x):\|F(t,x)\|_{1,V_K}\leqslant\text{const}\}\subset E_1(G)$$

在K的这一邻域上是一致Lipschitz的，且其中的常数是固定的。

证 任取一个支集在V_K中的连续可微的数值函数$\alpha(t,x)$，我们要

[注1] 如同半范$\|\cdot\|_w$，在E_1中如将对每一固定x在几乎所有的$t\in R$处取值相同的任意两个函数等同化，则半范$\|\cdot\|_1$成为范数．其实不难证明，命题$\{\forall x\in R^n, F(t,x)=0$对几乎所有$t$成立$\}$就等价于$\|F(t,x)\|_1=0$。

证明，集合

$$\{a(t,x)F(t,x):F(t,x)\in M\}\subset E_1\subset E_{\mathrm{Lip}}$$

在 E_{Lip} 中是一致 Lipschitz 的。

对任一函数 $a_{V_K}(t,x)$，

$$a(t,x)F(t,x)=a(t,x)a_{V_K}(t,x)F(t,x)=a(t,x)F_{V_K}(t,x),$$

其中，我们用了简化记号

$$F_{V_K}(t,x)=a_{V_K}(t,x)F(t,x).$$

我们写出估计式

$$\begin{aligned}
&|a(t,x')F(t,x')-a(t,x'')F(t,x'')|\\
&=\left|\int_0^1\frac{d}{ds}a(t,x''+s(x'-x''))F_{V_K}(t,x''+s(x'-x''))ds\right|\\
&\leqslant\int_0^1\left|\frac{\partial}{\partial x}\{a(t,x''+s(x'-x''))F_{V_K}(t,x''+s(x'-x''))\}\right|\\
&\qquad|x'-x''|ds\\
&\qquad\leqslant\left\{\max_{x\in R^n}\left|\frac{\partial}{\partial x}a(t,x)F_{V_K}(t,x)\right|\right\}|x'-x''|=L(t)|x'-x''|.
\end{aligned}$$

并有

$$\begin{aligned}
L(t)&\leqslant\left\{\max_{x\in R^n}|a(t,x)|+\max_{x\in R^n}\left|\frac{\partial}{\partial x}a(t,x)\right|\right\}\max_{x\in R^n}\left\{|F_{V_K}(t,x)|\right.\\
&\qquad\left.+\left|\frac{\partial}{\partial x}F_{V_K}(t,x)\right|\right\}\\
&\qquad\leqslant C_1\max_{x\in R^n}\left\{|F_{V_K}(t,x)|+\left|\frac{\partial}{\partial x}F_{V_K}(t,x)\right|\right\}.
\end{aligned}$$

因此，我们可得不等式

$$\int_R L(t)dt\leqslant C_1\int_R\max_{x\in R^n}\left\{|F_{V_K}(t,x)|+\left|\frac{\partial}{\partial x}F_{V_K}(t,x)\right|\right\}dt.$$

因为函数 $a_{V_K}(t,x)$ 的任意性，由本命题对集合 M 所设的条件，可得

$$\int_R L(t)dt\leqslant C_1\cdot\mathrm{const},$$

这就完成了证明。

5.2 变分方程与解的变分公式

考虑微分方程

$$\dot{x}=F(t,x), \qquad F(t,x)\in E_1(G). \tag{5.2}$$

设 $\tilde{x}(t)(t\in[t_1,t_2])$ 是这一方程的解，并设 $V_{\tilde{x}(t)}\subset G$ 是曲线 $\tilde{x}(t)$ 的一个紧邻域。按方才证明的命题 5.1，摄动集

$$\left\{\delta F(t,x):\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\leqslant 1\right\}\subset E_1(G)$$

在曲线 $\tilde{x}(t)$ 的邻域 $V_{\tilde{x}(t)}$ 上是一致 Lipschitz 的。如果初值向量 x_1 与摄动 $\delta F(t,x)$ 满足不等式条件

$$|x_1-\tilde{x}(t_1)|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\leqslant\varepsilon,$$

因此，也满足不等式

$$|x_1-\tilde{x}(t_1)|+\|\delta F(t,x)\|_{w,V_{\tilde{x}(t)}}\leqslant\varepsilon,$$

则根据连续依赖性定理（定理 4.4），对一切充分小的 $\varepsilon\geqslant 0$，摄动方程

$$\dot{x}=F(t,x)+\delta F(t,x)$$

在区间 $[t_1,t_2]$ 上存在满足初始条件 $x(t_1)=x_1$ 的解

$$x(t)=x(t;x_1,\delta F).$$

并且，当 $\varepsilon\to 0$ 时，有

$$\max_{t_1\leqslant t\leqslant t_2}|x(t;x_1,\delta F)-\tilde{x}(t)|=\|x(t;x_1,\delta F)-\tilde{x}(t)\|\to 0.$$

由此可见，当

$$|x_1-\tilde{x}(t_1)|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}$$

充分小时，摄动

$$\Delta x(t)=\Delta x(t;\delta x_1,\delta F)=x(t;x_1,\delta F)-\tilde{x}(t), \quad \delta x_1=x_1-\tilde{x}(t_1),$$

的微分方程

$$\Delta\dot{x}=F(t,\tilde{x}(t)+\Delta x)-F(t,\tilde{x}(t))+\delta F(t,\tilde{x}(t)+\Delta x) \tag{5.3}$$

以 $\Delta x(t_1)=\delta x_1$ 为初值的解在区间 $t\in[t_1,t_2]$ 上存在。并且，当 $|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\to 0$时，

$$\max_{t_1<t<t_2}|\Delta x(t;\delta x_1,\delta F)|=\|\Delta x(t;\delta x_1,\delta F)\|\to 0.$$

关于摄动 $\Delta x(t)$ 的这个论断只用到条件

$$|x_1-\tilde{x}(t_1)|+\|\delta F(t,x)\|_{w,V_{\tilde{x}(t)}}\to 0.$$

然后，在条件

$$|x_1-\tilde{x}(t_1)|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\to 0$$

之下，将有更强的结果。这时，$\Delta x(t)$ 可以进一步表示成

$$\Delta x(t;\delta x_1,\delta F)=\delta x(t;\delta x_1,\delta F)+\Delta_2 x(t;\delta x_1,\delta F),t\in[t_1,t_2], \tag{5.4}$$

其中，$\delta x(t;\delta x_1,\delta F)$满足线性微分方程

$$\dot{\delta x}=F_x(t,\tilde{x}(t))\delta x+\delta F(t,\tilde{x}(t)) \tag{5.5}$$

与初值条件

$$\delta x(t_1;\delta x_1,\delta F)=\delta x_1,$$

而 $\Delta_2 x(t;\delta x_1,\delta F)$ 满足条件

$$\begin{aligned}\max_{t_1<t<t_2}|\Delta_2 x(t;\delta x_1,\delta F)|&=\|\Delta_2 x(t;\delta x_1,\delta F)\|\\&=o(|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}),\end{aligned} \tag{5.6}$$

其中，$o(\varepsilon)/\varepsilon\to 0(\varepsilon\to 0)$.

函数 $\delta x(t;\delta x_1,\delta F)$ 是线性方程 (5.5) 的解，它线性依赖于 $(\delta x_1,\delta F)$. 因此，(5.6) 表明函数 $\delta x(t;\delta x_1,\delta F)$ 是映射

$$(\delta x_1,\delta F)\longrightarrow \Delta x(t;\delta x_1,\delta F),\quad t\in[t_1,t_2],$$

在点 $(\delta x_1,\delta F)=0$ 处的线性主部，如果在 $(\delta x_1,\delta F)$ 所成的线性空间中引进半范

$$\|(\delta x_1,\delta F)\|=|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}, \tag{5.7}$$

并在函数 $\Delta x(t;\delta x_1,\delta F)$ 所在的线性空间中引进区间 $[t_1,t_2]$ 上的一致收敛范数，就可以说函数 $\delta x(t;\delta x_1,\delta F)$ 是上述映射的微分。

函数 $\delta x(t;\delta x_1,\delta F)(t\in[t_1,t_2])$ 称为初值 $\tilde{x}(t_1)$ 受摄动 δx_1 且方

程(5.2)受摄动 $\delta F(t,x)$ 时解 $\bar{x}(t)$ 的变分。

非齐次线性方程 (5.5) 称为 (5.2) 在摄动 $\delta F(t,x)$ 下沿轨线 $\bar{x}(t)$ $(t\in[t_1,t_2])$ 的变分方程。这一方程是摄动 $\Delta x(t)$ 的非线性方程 (5.3) 的"线性化"。

齐次方程

$$\delta\dot{x}=F_x(t,\bar{x}(t))\delta x \tag{5.8}$$

称为 (5.2) 的沿轨线 $\bar{x}(t)(t\in[t_1,t_2])$ 的齐次变分方程。

以上所说可表述为本章的一个基本定理：

定理 5.1 当 (5.7) 充分小时，摄动方程 (5.3) 的满足初值条件 $\Delta x(t_1)=\delta x_1$ 的解

$$\Delta x(t)=\Delta x(t;\delta x_1,\delta F),\quad t\in[t_1,t_2],$$

在 $[t_1,t_2]$ 上存在，且此解可由 (5.4) 表示，其中变分 $\delta x(t;\delta x_1,\delta F)$ 满足变分方程 (5.5) 及初始条件

$$\delta x(t_1;\delta x_1,\delta F)=\delta x_1,$$

而函数 $\Delta_2 x(t;\delta x_1,\delta F)(t\in[t_1,t_2])$ 满足条件 (5.6).

在证明本定理之前，我们先作几点说明。

以 $\Gamma(t)$ 记齐次方程 (5.8) 的基本解矩阵，它在 $t=t_1$ 点等于单位阵。$G(t)$ 表示 $\Gamma(t)$ 的逆阵(参见第 5.5 节关于基本解矩阵的构造).

我们有

$$\dot{\Gamma}(t)=F_x(t,\bar{x}(t))\Gamma(t),$$

$$\dot{G}(t)=-G(t)F_x(t,\bar{x}(t))^{[注1]},\qquad t\in[t_1,t_2].$$

方程 (5.5) 的解即变分 $\delta x(t;\delta x_1,\delta F)$ 可表示为

$$\delta x(t;\delta x_1,\delta F)=\Gamma(t)(\delta x_1+\int_{t_1}^{t}G(\theta)\delta F(\theta,\bar{x}(\theta)d\theta), \tag{5.9}$$

[注1] 为导出第二个方程，我们对单位阵 $G(t)\Gamma(t)$ 微分，

$$\dot{G}(t)\Gamma(t)+G(t)\dot{\Gamma}(t)=\{\dot{G}(t)+G(t)F_x(t,\bar{x}(t))\}\Gamma(t)=0,$$

由于 $\Gamma(t)$ 是非异阵，遂得

$$\dot{G}(t)+G(t)F_x(t,\bar{x}(t))=0$$

上式称为解 $\tilde{x}(t)$ 的变分公式。如果只有初值变动 δx_1 而 $\delta F(t,x)=0$，则 (5.9) 变成：

$$\delta x(t;\delta x_1,0)=\Gamma(t)\delta x_1.$$

设 Π 是 $\boldsymbol{R}^n$ 的一个 $(n-1)$ 维子空间。Π_{t_1} 是如下的过 $\tilde{x}(t_1)$ 点的超平面：

$$\Pi_{t_1}=\{\tilde{x}(t_1)+\delta x_1;\delta x_1\in\Pi\}.$$

超平面族

$$\Pi_t=\{\tilde{x}(t)+\Gamma(t)\delta x_1=\tilde{x}(t)+\delta x(t;\delta x_1,0);\delta x_1\in\Pi\},\quad t\in[t_1,t_2],$$

称为超平面 Π_{t_1} 按齐次变分方程 (5.8) 沿轨道 $\tilde{x}(t)(t\in[t_1,t_2])$ 的推移。由于矩阵 $\Gamma(t)$ 的非异性，Π_t 是一族超平面。每一个超平面 Π_t 可由如下映射推移而得：

$$\tilde{x}(t_1)+\delta x_1\longrightarrow\tilde{x}(t)+\Gamma(t)\delta x_1,\qquad\forall\delta x_1\in\Pi,\qquad t\in[t_1,t_2].$$

如果行向量 ξ 正交于 Π，则向量

$$\psi(t)=\xi G(t),\qquad t\in[t_1,t_2],$$

正交于子空间 $\Pi_t-\tilde{x}(t)$，这是因为，对一切 $\delta x_1\in\Pi$ 有

$$\psi(t)\Gamma(t)\delta x_1=\xi G(t)\Gamma(t)\delta x_1=\xi\delta x_1=0.$$

显然，函数 $\psi(t)$ 满足微分方程

$$\dot{\psi}=-\psi F_x(t;\tilde{x}(t)),$$

以及初始条件

$$\psi(t_1)=\xi.$$

$\psi(t)$ 所适合的微分方程称为齐次变分方程 (5.8) 的共轭方程。共轭方程的特征是，它的任一解 $\chi(t)$ 与方程 (5.8) 的任一解 $\delta x(t)$ 之内积是常数，

$$\chi(t)\delta x(t)=\text{const},$$

其中的常数依赖于这两个解。事实上，等式

$$\dot{\chi}(t)\delta x(t)+\chi(t)\delta\dot{x}(t)=\{\dot{\chi}(t)+\chi(t)F_x(t,\tilde{x}(t))\}\delta x(t)=0$$

等价于

$$\dot{\chi}(t)=-\chi(t)F_x(t,\tilde{x}(t));$$

这是由于其中的向量 δx_1 是任意的，从而向量 $\delta x(t)=\Gamma(t)\delta x_1$ 是任意的 $(t_1\leqslant t\leqslant t_2)$.

5.3 定理5.1的证明

首先，我们证明，对于充分小的 $|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}$，摄动

$$\Delta x(t)=\Delta x(t;\delta x_1,\delta F),\quad t_1\leqslant t\leqslant t_2,$$

满足不等式

$$\|\Delta x(t)\|\leqslant C(|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}).$$

为此，将方程（5.3）改写成如下形式，

$$\begin{aligned}\Delta\dot{x}(t)&=F(t,\tilde{x}(t)+\Delta x(t))-F(t,\tilde{x}(t))+\delta F(t,\tilde{x}(t)+\Delta x(t))\\&=F_x(t,\tilde{x}(t))\Delta x(t)+\delta F(t,\tilde{x}(t))\\&\quad+\int_0^1\delta F_x(t,\tilde{x}(t)+s\Delta x(t))\Delta x(t)ds\\&\quad+\int_0^1\{F_x(t,\tilde{x}(t)+s\Delta x(t))-F_x(t,\tilde{x}(t))\}\Delta x(t)ds\\&=F_x(t,\tilde{x}(t))\Delta x(t)+\delta F(t,\tilde{x}(t))+R_1(t)+R_2(t),\end{aligned}$$

其中，

$$R_1(t)=\int_0^1\delta F_x(t,\tilde{x}(t)+s\Delta x(t))\Delta x(t)ds,$$

$$R_2(t)=\int_0^1\{F_x(t,\tilde{x}(t)+s\Delta x(t))-F_x(t,\tilde{x}(t))\}\Delta x(t)ds.$$

这一方程是以

$$\delta F(t,\tilde{x}(t))+R_1(t)+R_2(t)$$

为非齐次线性方程的非齐次项。由于初值 $\Delta x(t_1)=\delta x_1$，因此

$$\Delta x(t)=\Gamma(t)\Big\{\delta x_1+\int_{t_1}^{t}G(\theta)[\delta F(\theta,\tilde{x}(\theta))$$

$$+R_1(\theta)+R_2(\theta)]d\theta\Big\} \tag{5.10}$$

现来估计下式的绝对值，

$$\Gamma(t)\int_{t_1}^{t}G(\theta)[\delta F(\theta,\tilde{x}(\theta))+R_1(\theta)+R_2(\theta)]d\theta,$$

设

$$|\Gamma(t)|\leqslant C_1,\qquad |G(t)|\leqslant C_1,\qquad t_1\leqslant t\leqslant t_2,$$

则

$$\left|\Gamma(t)\int_{t_1}^{t}G(\theta)\delta F(\theta,\tilde{x}(\theta))d\theta\right|\leqslant C_1^2\int_{t_1}^{t}\max_{x\in V_{\tilde{x}(t)}}|\delta F(\theta,x)|d\theta$$

$$\leqslant C_1^2\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}} \tag{5.11}$$

对充分小的 $\|\Delta x(t)\|$，有

$$\tilde{x}(t)+s\Delta x(t)\in V_{\tilde{x}(t)},\qquad \forall(s,t)\in[0,1]\times[t_1,t_2].$$

因此，

$$\left|\Gamma(t)\int_{t_1}^{t}G(\theta)R_1(\theta)d\theta\right|\leqslant C_1^2\|\Delta x(t)\|\int_{t_1}^{t_2}\max_{x\in V_{\tilde{x}(t)}}|\delta F_x(t,x)|dt$$

$$\leqslant C_1^2\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\|\Delta x(t)\|. \tag{5.12}$$

此外，

$$\left|\Gamma(t)\int_{t_1}^{t}G(\theta)R_2(\theta)d\theta\right|\leqslant C_1^2\|\Delta x(t)\|\int_{t_1}^{t_2}dt\int_0^1|F_x(t,\tilde{x}(t)$$

$$+s\Delta x(t))-F_x(t,\tilde{x}(t))|ds\leqslant C_1^2\|\Delta x(t)\|J(\Delta x(t)), \tag{5.13}$$

其中，$J(\Delta x(t))$ 是 $\Delta x(t)(t\in[t_1,t_2])$ 的泛函，

$$J(\Delta x(t))=\int_{t_1}^{t_2}dt\int_0^1\left|F_x(t,\tilde{x}(t)+s\Delta x(t))-F_x(t,\tilde{x}(t))\right|ds.$$

我们将证明，当 $\|\Delta x(t)\|\longrightarrow 0$ 时 $J(\Delta x(t))$ 趋于零。写成

$$J(\Delta x(t))=\int_0^1\int_{t_1}^{t_2}\left|F_x(t,\tilde{x}(t)+s\Delta x(t))-F_x(t,\tilde{x}(t))\right|dtds.$$

对固定的 $(s,t)\in[0,1]\times[t_1,t_2]$，当 $\|\Delta x(t)\|\to 0$ 时，由 $F_x(t,\cdot)$ 的连续性，知被积式收敛于零，它又受控于如下的可积函数：

$$\begin{aligned}&\left|F_x(t,\tilde{x}(t)+s\Delta x(t))-F_x(t,\tilde{x}(t))\right|\leqslant\left|F_x(t,\tilde{x}(t)\right.\\&\left.+s\Delta x(t))\right|+\left|F_x(t,\tilde{x}(t))\right|\leqslant 2\max_{x\in R^n}\left\{\left|\alpha_{V_{\tilde{x}(t)}}(t,x)F(t,x)\right|\right.\\&\left.+\left|\frac{\partial}{\partial x}\alpha_{V_{\tilde{x}(t)}}(t,x)F(t,x)\right|\right\}=2m(t),\end{aligned}$$

只要注意到 $\alpha_{V_{\tilde{x}(t)}}$ 在 $V_{\tilde{x}(t)}$ 上等于 1，而 $(t,\tilde{x}(t)+s\Delta x(t))$ 与 $(t,\tilde{x}(t))$ 均在 $V_{\tilde{x}(t)}$ 之中，便可得上述不等式。按 Lebesgue 极限定理，得

$$\begin{aligned}\lim_{\|\Delta x(t)\|\to 0}J(\Delta x(t))=\int_0^1\int_{t_1}^{t_2}\lim_{\|\Delta x(t)\|\to 0}&\left|F_x(t,\tilde{x}(t)+s\Delta x(t))\right.\\&\left.-F_x(t,\tilde{x}(t))\right|dtds=0.\end{aligned}$$

根据估计 (5.11)—(5.13)，从 (5.10) 可得

$$\begin{aligned}\|\Delta x(t)\|\leqslant C_1|\delta x_1|+C_1^2\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}+\|\Delta x(t)\|\\ \{C_1^1J(\Delta x(t))+C_1^2\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\}.\end{aligned}$$

当 $|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\longrightarrow 0$ 时，由连续依赖性知 $\|\Delta x(t)\|\longrightarrow 0$，从而 $J(\Delta x(t))\to 0$. 由此可得

$$\begin{aligned}\|\Delta x(t)\|\leqslant\frac{(1+C_1)^2}{1-C_1^1J(\Delta x(t))-C_1^2\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}}(|\delta x_1|\\+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}})\leqslant 2(1+C_1^2)(|\delta x_1|\end{aligned}$$

$$+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}), \tag{5.14}$$

其中，$|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}$ 充分小。(5.14) 正是我们希望的不等式。

有了 (5.14) 及 $J(\Delta x(t))\to 0$ 之后，就不难证明关于 $\Delta x(t)$ 的表示式 (5.4) 了。我们将证明

$$\Delta_2 x(t)=\Delta_2 x(t;\delta x_1,\delta F)=\Delta x(t;\delta x_1,\delta F)-\delta x(t;\delta x_1,\delta F),$$
$$t\in[t_1,t_2],$$

具有下述性质：当 $|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\longrightarrow 0$ 时，

$$\frac{\|\Delta_2 x(t)\|}{|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}}\longrightarrow 0.$$

有

$$\Delta\dot{x}=F_x(t,\tilde{x}(t))\Delta x(t)+\delta F(t,\tilde{x}(t))+R_1(t)+R_2(t),$$

以及

$$\delta\dot{x}=F_x(t,\tilde{x}(t))\delta x(t)+\delta F(t,\tilde{x}(t)).$$

从第一个方程减去第二个方程，并注意 $\Delta_2 x(t)=\Delta x(t)-\delta x(t)$，即得

$$\Delta_2\dot{x}(t)=F_x(t,\tilde{x}(t))\Delta_2 x(t)+R_1(t)+R_2(t).$$

因为 $\delta x(t_1)=\delta x_1=\Delta x(t_1)$，故初值是 $\Delta_2 x(t_1)=0$. 因此，

$$\Delta_2 x(t)=\Gamma(t)\int_{t_1}^{t}G(\theta)(R_1(\theta)+R_2(\theta))d\theta.$$

利用估计式 (5.12)—(5.14)，可得不等式

$$\|\Delta_2 x(t)\|\leqslant\{C_1^2 J(\Delta x(t))+C_1^2\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\}\|\Delta x(t)\|$$

$$\leqslant 2C_1^2(1+C_1^1)\{J(\Delta x(t))$$

$$+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}}\}(|\delta x_1|+\|\delta F(t,x)\|_{1,V_{\tilde{x}(t)}})$$

这就完成了定理 5.1 的证明。

5.4 一个反例

设一族摄动 $\delta F(t,x;\varepsilon)(0\leqslant\varepsilon\leqslant1)$ 关于 x 二次连续可微，在方程(5.2)的一个解 $\tilde{x}(t)$ $(t_1\leqslant t\leqslant t_2)$ 的一个邻域上一致 Lipschitz，并且满足条件

$$\|\delta F(t,x;\varepsilon)\|_{\omega,V_{\tilde{x}(t)}}+\|\delta F_x(t,x;\varepsilon)\|_{\omega,V_{\tilde{x}(t)}}\leqslant\varepsilon\cdot\text{const}. \tag{5.15}$$

显然，这一条件弱于

$$\|\delta F(t,x;\varepsilon)\|_{1,V_{\tilde{x}(t)}}\leqslant\varepsilon\cdot\text{const}. \tag{5.16}$$

条件(5.15)是说左侧的半范的无穷小阶数当 $\varepsilon\to0$ 时不低于 ε. 按照连续依赖性定理(定理 4.4)，对充分小的一切 $\varepsilon\in[0,1]$，摄动方程

$$\dot{x}=F(t,x)+\delta F(t,x;\varepsilon)$$

有如下的解：

$$x(t;\varepsilon),\quad t_1\leqslant t\leqslant t_2,$$

$$x(t_1;\varepsilon)\equiv\tilde{x}(t_1).$$

当 $\varepsilon\to0$ 时，此解在区间 $[t_1,t_2]$ 上一致地趋于 $\tilde{x}(t)$，即

$$\max_{t_1\leqslant t\leqslant t_2}|x(t;\varepsilon)-\tilde{x}(t)|=\|x(t;\varepsilon)-\tilde{x}(t)\|=\|\Delta x(t;\varepsilon)\|\longrightarrow0.$$

摄动 $\Delta x(t;\varepsilon)$ 之无穷小阶数不低于 ε. 事实上，将方程

$$\Delta\dot{x}(t)=F(t,\tilde{x}(t)+\Delta x(t))-F(t,\tilde{x}(t))+\delta F(t,\tilde{x}(t)+\Delta x(t);\varepsilon)$$

的右侧函数展开成 Δx 的幂级数，我们得到

$$\begin{aligned}\Delta\dot{x}&=F_x(t,\tilde{x}(t))\Delta x+\delta F(t,\tilde{x}(t);\varepsilon)\\&\quad+\delta F_x(t,\tilde{x}(t);\varepsilon)\Delta x+\cdots,\end{aligned} \tag{5.17}$$

$\Delta x(t)$ 可视为以

$$\delta F(t, \tilde{x}(t); \varepsilon) + \delta F_x(t, \tilde{x}(t); \varepsilon)\Delta x + \cdots$$

为非齐次项的线性非齐次方程 (5.17) 的解。于是，

$$\Delta x(t; \varepsilon) = \Gamma(t)\int_{t_1}^{t} G(\theta)\{\delta F(\theta, \tilde{x}(\theta); \varepsilon)$$

$$+ \delta F_x(\theta, \tilde{x}(\theta); \varepsilon)\Delta x(\theta) + \cdots\}d\theta.$$

这里，未写出的是那些 Δx 阶数不小于 2 的项。因此，不难看出，被 "截断" 的（关于 Δx 线性的）方程

$$\Delta_1\dot{x} = F_x(t, \tilde{x}(t))\Delta_1 x + \delta F(t, \tilde{x}(t); \varepsilon)$$

$$+ \delta F_x(t, \tilde{x}(t); \varepsilon)\Delta_1 x \tag{5.18}$$

的解

$$\Delta_1 x(t; \varepsilon), \quad t \in [t_1, t_2],$$

$$\Delta_1 x(t_1; \varepsilon) = 0,$$

与 $\Delta x(t; \varepsilon)$ 之差的无穷小阶数高于 ε：

$$\|\Delta x(t; \varepsilon) - \Delta_1 x(t; \varepsilon)\| = o(\varepsilon).$$

现在我们来比较 "截断" 方程 (5.18) 与变分方程 (5.5)．前者的右端多一项 $\delta F_x(t, \tilde{x}(t); \varepsilon)\Delta_1 x$．如果条件 (5.16) 满足，那么根据定理 5.1，变分方程 (5.5) 的具有零初值的解 $\delta x(t; \varepsilon)$ 与 $\Delta x(t; \varepsilon)$ 之差的阶数高于 ε．然而，较弱的条件 (5.15) 并不保证其解 $\delta x(t; \varepsilon)$ 有此性质。如果 (5.15) 满足，我们仅能保证截断方程 (5.18) 的解 $\Delta_1 x(t; \varepsilon)$ 与 $\Delta x(t; \varepsilon)$ 之差的阶数高于 ε．

我们现给出一个例子，其中，由于函数 $\delta F(t, x; \varepsilon)$ 以及 $\delta F_x(t, x; \varepsilon)$ 沿 t 轴的振荡随 $\varepsilon \to 0$ 而越来越快，致使半范 (5.15) 有无穷小阶为 ε．在此例中，"截断" 方程 (5.18) 之解 $\Delta_1 x(t; \varepsilon)$ 是随 $\varepsilon \to 0$ 与 $\delta F_x(t, x; \varepsilon)$ 一起"共振"的．由此原因，所出现的项 $\delta F_x(t, x; \varepsilon)\Delta_1 x(t; \varepsilon)$

的阶为 ε，尽管它的两个因子的阶也是 ε。因此，"截断"方程中的右端这一项不再是可忽略的了[注1]。

例：原方程是数量方程

$$\dot{x}=0,$$

原始解为

$$\tilde{x}(t)\equiv 0,\quad 0\leqslant t\leqslant 1$$

一族摄动是

$$\delta F(t,x;\varepsilon)=x\sin(t/\varepsilon)+\cos(t/\varepsilon).$$

摄动 $\Delta x(t;\varepsilon)$ 所适合的方程（5.17）与截断方程（5.18）都是

$$\Delta\dot{x}=\cos(t/\varepsilon)+\sin(t/\varepsilon)\Delta x.$$

但变分方程是

$$\delta\dot{x}=\cos(t/\varepsilon),\quad \delta x(t;\varepsilon)=\varepsilon\sin(t/\varepsilon).$$

摄动是

$$\begin{aligned}\Delta x(t;\varepsilon)&=e^{-\varepsilon\cos(t/\varepsilon)}\int_0^t\cos\left(\frac{\theta}{\varepsilon}\right)e^{\varepsilon\cos(\theta/\varepsilon)}d\theta\\&=(1-\varepsilon\cos(t/\varepsilon)+\cdots)\int_0^t\cos(\theta/\varepsilon)(1+\varepsilon\cos(\theta/\varepsilon)+\cdots)d\theta\\&=(1-\varepsilon\cos(t/\varepsilon)+\cdots)(\varepsilon\sin(t/\varepsilon)+\varepsilon\int_0^t\cos^2(\theta/\varepsilon)d\theta+\cdots)\\&=\varepsilon\sin(t/\varepsilon)+\varepsilon\int_0^t\cos^2(\theta/\varepsilon)d\theta+\varepsilon^2\{\cdots\}.\end{aligned}$$

因此，

$$\Delta_2x(t;\varepsilon)=\Delta x(t;\varepsilon)-\delta x(t;\varepsilon)=\varepsilon\int_0^t\cos^2(\theta/\varepsilon)d\theta+\varepsilon^2\{\cdots\}.$$

[注1] 按定义，因子 $\delta F_x(t,\tilde{x}(t);\varepsilon)$ 的阶等于 $\|\delta F_x(t,x;\varepsilon)\|_{*}\,V_{\tilde{x}(t)}$ 关于 ε 的阶；因子 $\Delta_1x(t;\varepsilon)$ 以及乘积 $\delta F_x(t,\tilde{x}(t);\varepsilon)\Delta_1x(t;\varepsilon)$ 的阶分别等于如下两个值的阶，

$$\max_{t_1\leqslant t\leqslant t_2}|\Delta_1x(t;\varepsilon)|\quad 及\quad \max_{t_1\leqslant t\leqslant t_2}|\delta F_x(t,\tilde{x}(t);\varepsilon)\Delta_1x(t;\varepsilon)|.$$

上式右侧第一项是 ε 的一次幂，这是由 $\Delta x(t;\varepsilon)$ 的首项 $\varepsilon\sin(t/\varepsilon)$ 与函数 $\delta F_x(t,\tilde{x}(t);\varepsilon)=\sin(t/\varepsilon)$ 发生共振所致。

5.5 线性矩阵微分方程的解

这一节里，我们将用逐次逼近法来构造矩阵微分方程的解，并研究解对系数矩阵的依赖关系。

设 $\mathscr{T}$ 为 t 轴的任一区间；它可以是 $\boldsymbol{R}$. 每个分量都是可积函数的所有 $n\times n$ 矩阵值函数 $A(t)(t\in\mathscr{T})$ 组成的线性空间记为 E_A. 于是，

$$\int_{\mathscr{T}}|A(t)|dt<\infty,\qquad \forall A(\cdot)\in E_A.$$

对空间 E_A 赋以范数：

$$\|A(t)\|_w=\sup_{t',t''\in\mathscr{T}}\left|\int_{t'}^{t''}A(t)dt\right|$$

如果一族矩阵函数 $D\subset E_A$ 满足以下性质：

$$\left\{\int_{\mathscr{T}}|A(t)|dt;A(t)\in D\right\}.$$

是有界集，那就称 D 是*积分一致有界的*. 这时，n 维函数族

$$\{F(t,x)=A(t)x;A(t)\in D,x\in\boldsymbol{R}^n\}\subset E_1(\mathscr{T}\times\boldsymbol{R}^n)$$

显然在 $\mathscr{T}\times\boldsymbol{R}^n$ 的任一紧子集的一个邻域上是一致 Lipschitz 的。

矩阵函数族 $D\subset E_A$ 称为*积分等度连续*，如果对任一 $\varepsilon>0$，存在正数 $\delta>0$，当 $|t'-t''|\leqslant\delta$，且 $t',t''\in\mathscr{T}$ 时，有

$$\left|\int_{t'}^{t''}|A(t)|dt\right|\leqslant\varepsilon,\qquad \forall A(t)\in D.$$

譬如，若矩阵函数族 D 的矩阵范数一致有界：

$$|A(t)|\leqslant C,\qquad \forall A(t)\in D,$$

那么，它必是积分等度连续的。

我们以 $C_{\mathscr{T}}^{n^2}$ 表示每个分量在 $\mathscr{T}$ 上连续并在 t 趋于 $\mathscr{T}$ 的两个端点时存在有限极限的 $n\times n$ 矩阵 $X(t)(t\in\mathscr{T})$ 全体。在 $C_{\mathscr{T}}^{n^2}$ 上引进一致收敛范数：

$$\|X(t)\| = \sup_{t\in\mathscr{T}} |X(t)|,$$

显然，C^{n^2} 在上述范数下成为一个 Banach 空间。

任一绝对连续的矩阵值函数 $X(t)(t\in\mathscr{T})$，若对几乎所有的 $t\in\mathscr{T}$ 满足微分方程

$$\dot{X} = A(t)X, \qquad A(t)\in E_A,$$

就称它为这个方程的解。该方程的解 $\Gamma(t)$ $(t\in\mathscr{T})$ 如果对所有的 $t\in\mathscr{T}$ 都是非异阵，就称之为基本解矩阵。根据唯一性定理，$\Gamma(t)$ 在 $\mathscr{T}$ 上点点非异当且仅当它在某一点上是非异的。事实上，倘若在任一点 $t_0\in\mathscr{T}$ 处，$\operatorname{rank}\Gamma(t_0)<n$，则存在列向量 $x_0\neq 0$ 使得 $\Gamma(t_0)x_0=0$。因此，$x(t)=\Gamma(t)x_0$ 是线性方程

$$\dot{x} = A(t)x$$

在 $t=t_0$ 点为零的解。因此，$x(t)\equiv 0$，即在每一点 $t\in\mathscr{T}$ 处都有 $\operatorname{rank}\Gamma(t)<n$.

定理5.2 每一矩阵微分方程

$$\dot{X} = A(t)X, \qquad A(t)\in E_A, \tag{5.19}$$

有唯一解

$$X(t) = X(t;\tau, X_\tau, A), \qquad t\in\mathscr{T},$$

它满足任意给定的初始条件

$$X(\tau) = X_\tau, \qquad \tau\in\mathscr{T}.$$

设给定一个积分一致有界集 $D\subset E_A$，且设

$$\hat{X}(t) = X(t;\hat{\tau},\hat{X}_{\hat{\tau}},\hat{A}), \qquad t\in\mathscr{T},$$

是方程

$$\dot{X} = \hat{A}(t)X, \qquad \hat{A}(t)\in D$$

的一个解，则对任一 $\varepsilon>0$，存在正数 $\delta>0$，当初始数据 $(\tau, X_\tau)(\tau\in\mathscr{T})$ 以及矩阵 $A(t)\in D$ 满足条件

$$|\tau-\hat{\tau}|+|X_\tau-\hat{X}_{\hat{\tau}}|+\|A(t)-\hat{A}(t)\|_w\leqslant\delta$$

时，对应的解 $X(t;\tau,X_\tau,A)$ 便满足如下估计，

$$\|X(t;\tau,X_\tau,A)-X(t;\hat{\tau},\hat{X}_{\hat{\tau}},\hat{A})\|\leqslant\varepsilon. \tag{5.20}$$

如果进一步设 D 是一个积分等度连续族，且初值 X_τ 属于一个有界集 $\mathscr{R}$，则解集

$$\{X(t;\tau,X_\tau,A):X_\tau\in\mathscr{R},A(t)\in D\} \tag{5.21}$$

是 Banach 空间 $C_{\mathscr{T}}^{n^2}$ 中的致密集（从而是有界集）。

证　现以记号 E_σ 表示由点

$$\sigma=(\tau,X_\tau,A(t))\in\boldsymbol{R}\times E_{X_\tau}\times E_A$$

赋以范数

$$\|\sigma\|=|\tau|+|X_\tau|+\|A(t)\|_w$$

所组成的线性赋范空间。

定义函数空间 $C_{\mathscr{T}}^{n^2}$ 到其自身内的一族映射 $\varphi(X(\cdot),\sigma)$ 为

$$X(\cdot)\longrightarrow\varphi(X(\cdot),\sigma):=X_\tau+\int_\tau^t A(\theta)X(\theta)d\theta=Y(t;X(\cdot),\sigma),\quad t\in\mathscr{T},$$

其中 σ 是参数，

$$\sigma=(\tau,X_\tau,A(t))\in\mathscr{T}\times E_{X_\tau}\times E_A\subset E_\sigma$$

由于每个函数 $X(t)\in C_{\mathscr{T}}^{n^2}$ 必以常数为界，即 $|X(t)|\leqslant\text{const}$，　$t\in\mathscr{T}$，

$$|Y(t'';X(\cdot),\sigma)-Y(t';X(\cdot),\sigma)|\leqslant\text{const}\left|\int_{t'}^{t''}|A(t)|dt\right|.$$

这表明上述映射确实将 $C_{\mathscr{T}}^{n^2}$ 映入自身中。

映射 φ 的每一不动点是方程 (5.19) 以 (τ,X_τ) 为初始数据的解：

$$X_\sigma(t)=X(t;\sigma)=X_\tau+\int_\tau^t A(\theta)X_\sigma(\theta)d\theta,\qquad t\in\mathscr{T}.$$

如果我们能够证明：第一，对一固定的 $X(\cdot)$，当 σ 遍取于某一集

$$\mathscr{T}\times E_{X_\tau}\times D\subset E_\sigma, \tag{5.22}$$

其中 $D\subset E_A$ 是一积分一致有界的矩阵函数族时，迭代映射 $\varphi^{(p)}(X(\cdot),\sigma)$ 连续依赖于 σ；第二，当 p 充分大时，这些迭代是关于集 (5.22)

的一致压缩映射，那么，作为不动点定理（定理4.1）的推论，就得到解的存在唯一性与连续依赖性关系（5.20）。

这些性质的证明只是定理 4.2 证明的重复(并且更为简单)，因此从略了。现来证明集合 (5.21)的致密性，即证明集合 (5.21) 在$C_{\mathcal{T}}^{n^2}$中的闭包是紧集。这是本定理的最后一个结论。

先证集 (5.21) 的一致有界性。现有下列不等式：

$$
\begin{aligned}
|X(t)| &= |X(t;\tau,X_\tau,A)| \leqslant |X_\tau| + \left|\int_\tau^t |A(\theta_1)|\cdot|X(\theta_1)|d\theta_1\right| \\
&\leqslant |X_\tau|\left(1+\left|\int_\tau^t |A(\theta_1)|d\theta_1\right|\right) + \left|\int_\tau^t |A(\theta_1)|d\theta_1 \int_\tau^{\theta_1}|A(\theta_2)|\right. \\
&\qquad\qquad \left.\cdot|X(\theta_2)|d\theta_2\right| \leqslant \cdots \\
&\leqslant |X_\tau|\left\{1+\left|\int_\tau^t |A(\theta_1)|d\theta_1\right|\right. \\
&\quad + \left|\int_\tau^t |A(\theta_1)|d\theta_1\int_\tau^{\theta_1}|A(\theta_2)|\,d\theta_2\right| + \cdots \\
&\quad \left. + \left|\int_\tau^t |A(\theta_1)|d\theta_1\int_\tau^{\theta_1}|A(\theta_2)|d\theta_2\cdots\int_\tau^{\theta_{p-1}}|A(\theta_p)|d\theta_p\right|\right\} \\
&\quad + \left|\int_\tau^t |A(\theta_1)|d\theta_1\cdots\int_\tau^{\theta_{p-1}}|A(\theta_p)|d\theta_p\right. \\
&\quad \left.\cdot\int_\tau^{\theta_p}|A(\theta_{p+1})|\cdot|X(\theta_{p+1})|d\theta_{p+1}\right|.
\end{aligned}
$$

应用已在第四章里证过的公式：

$$
\begin{aligned}
&\int_\tau^t |A(\theta_1)|d\theta_1\int_\tau^{\theta_1}|A(\theta_2)|d\theta_2\cdots\int_\tau^{\theta_{p-1}}|A(\theta_p)|d\theta_p \\
&= \frac{1}{p!}\left(\int_\tau^t |A(\theta)|d\theta\right)^p,
\end{aligned}
$$

并考虑到 D 的积分一致有界性：

$$
\int_{\mathcal{T}} |A(\theta)|d\theta \leqslant C, \qquad \forall A(t)\in D,
$$

我们得到

$$
|X(t)| \leqslant |X_\tau|\left(1+C+\frac{C^2}{2!}+\cdots+\frac{C^p}{p!}\right)+\|X(t)\|\frac{C^{p+1}}{(p+1)!}.
$$

令 $p\to+\infty$，得

$$\|X(t)\|\leqslant|X_\tau|e^{C}.$$

因为初值 X_τ 的取值集 $\mathfrak{N}$ 是有界集，故这一估计证实了范数 $\|X(t;\tau,X_\tau,A)\|$ 的一致有界性。

再证集(5.21)的等度连续性。这易由 D 的积分等度连续性推出。实际上，根据 $\|X(t;\tau,X_\tau,A)\|$ 的一致有界性，以及 D 族函数的积分等度连续性，与 $X_\tau\in\mathfrak{N}$ 及 $A(t)\in D$ 的选取无关地存在极限：

$$\begin{aligned}
&|X(t';\tau,X_\tau,A)-X(t'';\tau,X_\tau,A)|\\
&\leqslant\left|\int_{t'}^{t''}|A(t)|\,|X(t;\tau,X_\tau,A)|dt\right|\\
&\leqslant\|X(t;\tau,X_\tau,A)\|\left|\int_{t'}^{t''}|A(t)|dt\right|\longrightarrow 0,\ |t'-t''|\to 0.
\end{aligned}$$

再根据 Arzelà 定理，知集合(5.21)是相对紧的。[注1]

[注1] 定理 5.2 及其证明是在区间 $\mathscr{T}$ 有界的情形下作出的。当 $\mathscr{T}$ 是无界区间时，譬如 $\mathscr{T}=[a,+\infty)$（a 是有限数）时，为了使定理 5.2 继续保持正确，必须将集 $D\subset E_A$ 的积分等度连续性的定义作如下的补充：对任一 $\varepsilon>0$，存在正数 $M>0$，当 t'，$t''\geqslant M$ 且 $t',t''\in\mathscr{T}$ 时，对 $A(t)\in D$ 一致地成立

$$\left|\int_{t'}^{t''}|A(t)|dt\right|\leqslant\varepsilon.$$

这样补充定义后，定理 5.2 对 $\mathscr{T}=[a,+\infty)$ 情形仍成立，但证明集合(5.21)的相对紧性时，必须补充说明

$$|X(t';\tau,X_\tau,A)-X(t'';\tau,X_\tau,A)|\to 0,\quad (t',t''\to\infty)$$

关于 $X_\tau\in\mathscr{N}$ 及 $A(t)\in D$ 一致地成立，然后应用无界区间上的 Arzelà 定理获证（参见中译本参考文献 [8]）。——译者注。

第六章　凸控制问题中的轨线变分

前两章的结果使我们有一种很自然的方法去变动凸控制问题中的状态轨线。这种方法有一些性质使其成为研究控制问题的一个重要的工具。本章将致力于论述这一方法并证明它的基本性质。

6.1　广义控制的变分与对应的受控方程的变分

我们回到凸控制问题

$$\dot{x}=\langle\mu_t,f(t,x,u)\rangle,\qquad \mu_t\in\mathfrak{M}_U.$$

设 $\tilde{\mu}_t$ 是任一给定的广义控制，$\tilde{x}(t)$ $(t\in[t_1,t_2])$ 是 $\mu_t=\tilde{\mu}_t$ 时方程

$$\dot{x}=\langle\tilde{\mu}_t,f(t,x,u)\rangle=F(t,x)\tag{6.1}$$

的一条轨线。可以证明，函数 $F(t,x)$ 定义在全空间 $\boldsymbol{R}\times\boldsymbol{R}^n$ 上，关于 x 连续可微，对 t 可测，并且在任一紧集 $K\subset\boldsymbol{R}\times\boldsymbol{R}^n$ 上有界。最后一点可这样来证：由于概率测度 $\tilde{\mu}_t$ 是集中在一个与 $t\in\boldsymbol{R}$ 无关的有界集 $N\subset U\subset\boldsymbol{R}^r$ 之上，因此，

$$\begin{aligned}\sup_{(t,x)\in K}|F(t,x)|&\leqslant\sup_{(t,x)\in K}\int_N|f(t,x,u)|d\tilde{\mu}_t\\&\leqslant\sup_{\substack{(t,x)\in K\\u\in N}}|f(t,x,u)|\int_N d\tilde{\mu}_t=\sup_{\substack{(t,x)\in K\\u\in N}}|f(t,x,u)|.\end{aligned}\tag{6.2}$$

这表示 $F(t,x)$ 在紧集 K 上有界。由此可见，对于任一具有紧支集的连

续可微的数值函数 $\alpha(t,x)$，有 $\alpha(t,x)F(t,x)\in E_1$，即

$$F(t,x)\in E_1(\boldsymbol{R}\times\boldsymbol{R}^n).$$

对 $\mu_t\in\mathfrak{M}_U$，差

$$\delta\mu_t=\mu_t-\tilde{\mu}_t$$

称为广义控制 $\tilde{\mu}_t$ 的一个变分，或摄动。控制 $\tilde{\mu}_t$ 的所有摄动的集合将记为 $\delta\mathfrak{M}_{\tilde{\mu}_t}$. 由 $\mathfrak{M}_U$ 的凸性立即得到 $\delta\mathfrak{M}_{\tilde{\mu}_t}$ 的凸性：

$$\lambda_1(\mu_t^{(1)}-\tilde{\mu}_t)+\lambda_2(\mu_t^{(2)}-\tilde{\mu}_t)=\lambda_1\mu_t^{(1)}+\lambda_2\mu_t^{(2)}-\tilde{\mu}_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},$$
$$\forall\lambda_1,\lambda_2\geqslant 0,\lambda_1+\lambda_2=1.$$

对每一摄动 $\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t}$，有方程(6.1)右端的一个摄动与之对应。现将摄动控制 $\tilde{\mu}_t+\delta\mu_t$ 代替 (6.1)中的 $\tilde{\mu}_t$，我们就得到摄动方程

$$\dot{x}=\langle\tilde{\mu}_t+\delta\mu_t,f(t,x,u)\rangle=\langle\tilde{\mu}_t,f(t,x,u)\rangle+\langle\delta\mu_t,f(t,x,u)\rangle.$$

可见，右端摄动

$$\langle\delta\mu_t,f(t,x,u)\rangle=\delta F(t,x)$$

线性（更确切地说，应是“仿射”）依赖于控制摄动 $\delta\mu_t$. 我们可以把两个摄动 $\delta\mu_t^{(1)}$ 与 $\delta\mu_t^{(2)}$ 以及满足条件 $\lambda_1\geqslant 0,\lambda_2\geqslant 0$ 及 $\lambda_1+\lambda_2=1$ 的系数 λ_1 与 λ_2 结合起来作为摄动，这就使得凸控制问题比原来的常义控制问题更便于研究。特别是，这个事实以及逼近引理在下一章的最大值原理证明中将起到决定性作用。

每一摄动 $\langle\delta\mu_t,f(t,x,u)\rangle$ 都属于 $E_1(\boldsymbol{R}\times\boldsymbol{R}^n)$，这与 $F(t,x)$ 有类似结论的理由相同。

(6.1) 沿轨线 $\tilde{x}(t)\,(t\in[t_1,t_2])$ 的变分方程是

$$\delta\dot{x}=F_x(t,\tilde{x}(t))\delta x+\langle\delta\mu_t,f(t,\tilde{x}(t),u)\rangle.$$

保持第五章的有关记号，我们可以写出解 $\tilde{x}(t)$ 的有零初值的变分如下：

$$\delta x(t;\delta\mu_\theta)=\Gamma(t)\int_{t_1}^{t}G(\theta)\langle\delta\mu_\theta,f(\theta,\tilde{x}(\theta),u)\rangle d\theta$$

设 Σ 是任一指标集，一族摄动

$$\{\delta\mu_t(\sigma);\sigma\in\Sigma\}$$

称为有限的，如果 $\delta\mu_t(\sigma)$ 的测度都集中在一个与 $t\in\boldsymbol{R}$ 和 $\sigma\in\Sigma$ 都无关的有界子集上。

命题 6.1 对任一族有限的摄动 $\delta\mu_t(\sigma)(\sigma\in\Sigma)$，对应的函数族

$$\{\langle\delta\mu_t(\sigma),f(t,x,u)\rangle=\delta F(t,x;\sigma);\sigma\in\Sigma\}$$

在任一紧集 $K\subset\boldsymbol{R}\times\boldsymbol{R}^n$ 的一个邻域上是一致 Lipschitz 的。

证 任取一函数 $\alpha_{V_K}(t,x)$，其中 V_K 是紧集 K 的任意一个紧邻域。设 I 是函数 $\alpha_{V_K}(t,x)$ 的支集在 t 轴上的投影，而 X 是该支集在空间 $\boldsymbol{R}^n$ 上的投影。

与不等式(6.2)一样，可证明以下函数

$$|\delta F(t,x;\sigma)|\text{ 及 }|\delta F_x(t,x;\sigma)|,\sigma\in\Sigma$$

在函数 $\alpha_{V_K}(t,x)$ 的紧支集上以某常数 C 为界。从而，下面的不等式成立：

$$\begin{aligned}\|\delta F(t,x;\sigma)\|_{1,V_K}&\leqslant\int\max_{x\in\boldsymbol{R}^n}\Big\{|\alpha_{V_K}(t,x)\delta F(t,x;\sigma)|\\&\quad+\left|\frac{\partial}{\partial x}\alpha_{V_K}(t,x)\delta F(t,x;\sigma)\right|\Big\}dt\\&\leqslant\int_I\max_{x\in X}\Big\{|\delta F(t,x;\sigma)|+|\delta F_x(t,x;\sigma)|\\&\quad+\left|\frac{\partial}{\partial x}\alpha_{V_K}(t,x)\right||\delta F(t,x;\sigma)|\Big\}dt\\&\leqslant C\int_I\left(2+\max_{x\in\boldsymbol{R}^n}\left|\frac{\partial}{\partial x}\alpha_{V_K}(t,x)\right|\right)dt.\end{aligned}$$

从这个不等式及命题 5.1 直接得本命题。

下一个命题建立了广义控制的弱收敛与半范 $\|\cdot\|_w$ 之间的联系。

命题 6.2 设依赖于参数 $\sigma\in\Sigma$ 的一列摄动

$$\delta\mu_t^{(i)}(\sigma),\quad\sigma\in\Sigma,\quad i=1,2,\cdots,$$

当 $i\to\infty$ 时关于 $\sigma\in\Sigma$ 一致地弱收敛于零，即对任一在 $\boldsymbol{R}\times\boldsymbol{R}^r$ 上有紧支集的连续函数 $g(t,u)$，关于 $\sigma\in\Sigma$ 一致地成立：

$$\int_R \langle \delta\mu_t^{(i)}(\sigma),\ g(t,u)\rangle dt \longrightarrow 0 \quad (i\to\infty).$$

又设 $\delta\mu_t^{(i)}(\sigma)(t\in \boldsymbol{R}, i=1,2,\cdots,\sigma\in\boldsymbol{\Sigma})$的测度都集中在同一有界集

$$N\subset U\subset \boldsymbol{R}^r$$

上，则对任一紧集 $K\subset \boldsymbol{R}\times\boldsymbol{R}^n$，关于 $\sigma\in\boldsymbol{\Sigma}$ 一致地成立

$$\|\langle \delta\mu_t^{(i)}(\sigma), f(t,x,u)\rangle\|_{w,K}\longrightarrow 0 \quad (i\to\infty).$$

证 取一函数 $\alpha_K(t,x)$，设 $\hat{K}$是它的支集。记

$$g_\sigma^{(i)}(t',t'',x)=\left|\int_{t'}^{t''}\langle \delta\mu_t^{(i)}(\sigma), \alpha_K(t,x)f(t,x,u)\rangle dt\right|$$

由于

$$\|\langle \delta\mu_t^{(i)}(\sigma), f(t,x,u)\rangle\|_{w,K}\leqslant \max_{\substack{t',t''\in R\\ x\in R^n}} g_\sigma^{(i)}(t',t'',x), \forall\sigma\in\boldsymbol{\Sigma}.$$

只要证明关于 $t',t''\in\boldsymbol{R}, x\in\boldsymbol{R}^n$ 以及 $\sigma\in\boldsymbol{\Sigma}$ 一致地成立：

$$g_\sigma^{(i)}(t',t'',x)\to 0, (i\to\infty),$$

也就证明了本命题。

假使不然，则能找到一个正数 $\eta>0$，一列单调增加的整数 i_1, $i_2,\cdots,i_k,\cdots$，以及（由于支集 $\hat{K}$的紧性）收敛序列

$$t'^{(k)}\to \hat{t}', t''^{(k)}\to \hat{t}'', x^{(k)}\to\hat{x}, (k\to\infty),$$

和一列点 $\sigma_k\in\boldsymbol{\Sigma}$，使得

$$g_{\sigma_k}^{(i_k)}(t'^{(k)},t''^{(k)},x^{(k)})\geqslant\eta>0, \forall k=1,2,\cdots.$$

考察不等式

$$\begin{aligned}
&\left|g_{\sigma_k}^{(i_k)}(\hat{t}',\hat{t}'',\hat{x})-g_{\sigma_k}^{(i_k)}(t'^{(k)},t''^{(k)},x^{(k)})\right|\\
&\leqslant\left|\int_{\hat{t}'}^{\hat{t}''}|\langle \delta\mu_t^{(i_k)}(\sigma_k), \alpha_K(t,\hat{x})f(t,\hat{x},u)\right.\\
&\quad\left.-\alpha_K(t,x^{(k)})f(t,x^{(k)},u)\rangle|dt\right|\\
&\quad+\left|\int_{\hat{t}''}^{t''(k)}|\langle\delta\mu_t^{(i_k)}(\sigma_k), \alpha_K(t,x^{(k)})f(t,x^{(k)},u)\rangle|dt\right|\\
&\quad+\left|\int_{t'(k)}^{\hat{t}'}|\langle\delta\mu_t^{(i_k)}(\sigma_k), \alpha_K(t,x^{(k)})f(t,x^{(k)},u)\rangle|dt\right|.
\end{aligned}\tag{6.3}$$

我们要证明当 $k\to\infty$ 时不等式 (6.3) 右侧的每一项以及函数 $g_{\sigma_k}^{(i_k)}(\hat{t}',\hat{t}'',\hat{x})$ 都收敛于零。这样，对充分大的 k，将有

$$0<\eta/2<|g_{\sigma_k}^{(i_k)}(\hat{t}',\hat{t}'',\hat{x})-g_{\sigma_k}^{(i_k)}(t'^{(k)},t''^{(k)},x^{(k)})|<\eta/2,$$

产生了矛盾，从而证明了本命题。

现在，$\delta\mu_t^{(i)}(\sigma)$ 的测度都集中在有界集 $N\subset \boldsymbol{R}^r$ 上，且因它们是两个概率测度之差，故其范数 $\leqslant 2$。因此，由命题 2.2，我们得知对任意固定的 t',t'' 和 x，关于 $\sigma\in\Sigma$ 一致地有

$$g_\sigma^{(i)}(t',t'',x)=\int_{t_1}^{t''}\langle\delta\mu_t^{(i)}(\sigma),\alpha_K(t,x)f(t,x,u)\rangle dt\longrightarrow 0\quad(i\to\infty).$$

特别是：

$$g_{\sigma_k}^{(i_k)}(\hat{t}',\hat{t}'',\hat{x})\to 0\quad(k\to\infty).$$

现考察不等式(6.3)的右侧各项，与证明式(6.2)的方法一样，可证对所有的 $(t,x)\in\hat{K}$（从而，也对所有的 $(t,x)\in\boldsymbol{R}\times\boldsymbol{R}^n$）成立不等式：

$$|\langle\delta\mu_t^{(i)}(\sigma),\alpha_K(t,x)f(t,x,u)\rangle|\leqslant\text{const}.$$

因此，当 $k\to\infty$ 时，根据 N 是有界集，可得

$$\left|\int_{t'^{(k)}}^{\hat{t}'}|\langle\delta\mu_t^{(i_k)}(\sigma_k),\alpha_K(t,x^{(k)})f(t,x^{(k)},u)\rangle|dt\right|$$

$$\leqslant\text{const}\left|\int_{t'^{(k)}}^{\hat{t}'}dt\right|\longrightarrow 0,$$

$$\left|\int_{\hat{t}''}^{t''^{(k)}}|\langle\delta\mu_t^{(i_k)}(\sigma_k),\alpha_K(t,x^{(k)})f(t,x^{(k)},u)\rangle|dt\right|$$

$$\leqslant\text{const}\left|\int_{\hat{t}''}^{t''^{(k)}}dt\right|\longrightarrow 0.$$

又因 $x^{(k)}\to\hat{x}$，且函数 $\alpha_K(t,x)f(t,x,u)$ 是连续的，故

$$\sup_{\substack{u\in N\\ t\in[\hat{t}',\hat{t}'']}}|\alpha_K(t,\hat{x})f(t,\hat{x},u)-\alpha_K(t,x^{(k)})f(t,x^{(k)},u)|\longrightarrow 0,(k\to\infty).$$

最后得到

$$\left|\int_{\hat{t}'}^{\hat{t}''}\left|\langle\delta\mu_t^{(i_k)}(\sigma_k),\alpha_K(t,\hat{x})f(t,\hat{x},u)-\alpha_K(t,x^{(k)})f(t,x^{(k)},u)\rangle\right|dt\right|$$

$$\leqslant\left|\int_{\hat{t}'}^{\hat{t}''}dt\int_N\left|\alpha_K(t,\hat{x})f(t,\hat{x},u)-\alpha_K(t,x^{(k)})f(t,x^{(k)},u)\right|d\delta\mu_t^{(i_k)}\right|$$

$$\leqslant 2\left|\int_{\hat{t}'}^{\hat{t}''}\sup_{\substack{u\in N\\ t\in[\hat{t}',\hat{t}'']}}\left|\alpha_K(t,\hat{x})f(t,\hat{x},u)-\alpha_K(t,x^{(k)})f(t,x^{(k)},u)\right|dt\right|\longrightarrow 0$$

证毕.

如果以强收敛取代命题 6.2 中的弱收敛，则得知下命题。

命题 6.3 设一列摄动

$$\delta\mu_t^{(i)}(\sigma),\quad \sigma\in\boldsymbol{\Sigma},\quad i=1,2,\cdots,$$

关于 $\sigma\in\boldsymbol{\Sigma}$ 一致地强收敛于零，即关于 $\sigma\in\boldsymbol{\Sigma}$ 一致地有

$$\int_{\boldsymbol{R}}\|\delta\mu_t^{(i)}(\sigma)\|dt\to 0(i\to\infty),$$

并设 $\delta\mu_t^{(i)}(\sigma)$ 的测度都集中在一个有界集 $N\subset U\subset\boldsymbol{R}^r$ 之上，则对任一紧集 $K\subset\boldsymbol{R}\times\boldsymbol{R}^n$，有

$$\|\langle\delta\mu_t^{(i)}(\sigma),f(t,x,u)\rangle\|_{1,K}\longrightarrow 0\quad(i\to\infty).$$

极限是关于 $\sigma\in\boldsymbol{\Sigma}$ 一致的。

证

我们记 $\langle\delta\mu_t^{(i)}(\sigma),f(t,x,u)\rangle=\delta F^{(i)}(t,x;\sigma)$. 取一函数 $\alpha_K(t,x)$.

设 $\hat{K}$ 是它的支集，下面的估计显然对于 $\sigma\in\boldsymbol{\Sigma}$ 一致地成立：

$$\sup_{x\in R^n}\left\{\left|\alpha_K(t,x)\delta F^{(i)}(t,x;\sigma)\right|+\left|\frac{\partial}{\partial x}\alpha_K(t,x)\delta F^{(i)}(t,x,\sigma)\right|\right\}$$

$$\leqslant\sup_{\substack{(\theta,x)\subset R\times R^n\\ u\in N}}\left\{\left|\alpha_K(\theta,x)f(\theta,x,u)\right|\right.$$

$$\left.+\left|\frac{\partial}{\partial x}\alpha_K(\theta,x)f(\theta,x,u)\right|\right\}\|\delta\mu_t^{(i)}(\sigma)\|$$

于是，我们得到

$$\sup_{\sigma\in\Sigma}\|\delta F^{(i)}(t,x;\sigma)\|_{1,K}$$

$$\leqslant\sup_{\sigma\in\Sigma}\int_R\sup_{x\in R^n}\Big\{|a_K(t,x)\delta F^{(i)}(t,x;\sigma)|$$

$$+\Big|\frac{\partial}{\partial x}a_K(t,x)\delta F^{(i)}(t,x;\sigma)\Big|\Big\}dt$$

$$\leqslant\sup_{\substack{(t,x)\in R\times R^n\\ u\in N}}\Big\{|a_K(t,x)f(t,x,u)|$$

$$+\Big|\frac{\partial}{\partial x}a_K(t,x)f(t,x,u)\Big|\Big\}\sup_{\sigma\in\Sigma}\int_R\|\delta\mu_t^{(i)}(\sigma)\|dt\longrightarrow 0\quad(i\to\infty).$$

这就证明了本命题。

在结束这一节之前，我们要介绍构造广义控制的变分族的一个重要的方法。在下一节中，将借助于这一方法构造一族轨线变分。在第七章里，这样的轨线变分族将用于最大值原理的证明之中。

设给出控制 $\tilde{\mu}_t$ 的 p 个摄动

$$\delta\mu_t^{(1)},\cdots,\delta\mu_t^{(p)}\in\delta\mathfrak{M}_{\tilde{\mu}_t},$$

它们在 t 轴的一个（任意固定的）区间 $I\subset\boldsymbol{R}$之外为零，即

$$\delta\mu_t^{(k)}=0,\forall t\notin I\text{ 且 }\forall k=1,2,\cdots,p.$$

对任一组固定的非负数 $\lambda_1\geqslant0,\cdots,\lambda_p\geqslant0$，当 $\varepsilon\geqslant0$ 充分小时，一族测度

$$\delta\mu_t(\varepsilon,\lambda)=\varepsilon\sum_{k=1}^{p}\lambda_k\delta\mu_t^{(k)},t\in\boldsymbol{R},\lambda=(\lambda_1,\cdots,\lambda_p),\varepsilon\geqslant0.$$

是控制 $\tilde{\mu}_t$ 的一族摄动。事实上，因为集合 $\delta\mathfrak{M}_{\tilde{\mu}_t}$ 是凸的并且含有零摄动 $\delta\mu_t=0\in\delta\mathfrak{M}_{\tilde{\mu}_t}$，所以对 $1-\varepsilon\sum_{k=1}^{p}\lambda_k\geqslant0$，我们有

$$\delta\mu_t(\varepsilon,\lambda)=\sum_{k=1}^{p}\varepsilon\lambda_k\delta\mu_t^{(k)}+(1-\varepsilon\sum_{k=1}^{p}\lambda_k)0\in\delta\mathfrak{M}_{\tilde{\mu}_t}.$$

ε 的上限依赖于 λ. 设 λ 遍取于一个有界集 $\Lambda\subset\boldsymbol{R}^p$，其中

$$\lambda=(\lambda_1,\cdots,\lambda_p)\geqslant0,\forall\lambda\in\Lambda,$$

即 $\lambda_1\geqslant0,\cdots,\lambda_p\geqslant0$，那么，对一切 $\lambda\in\Lambda$，存在同一正值上限 $\varepsilon_0>0$，从而可得一族摄动

$$\Big\{\delta\mu_t(\varepsilon,\lambda)=\varepsilon\sum_{i=1}^{p}\lambda_i\delta\mu_t^{(i)};\lambda\in\Lambda,0\leqslant\varepsilon\leqslant\varepsilon_0\Big\}\subset\delta\mathfrak{M}_{\tilde{\mu}_t}.\tag{6.4}$$

显然，这族摄动是有限的。由于范数 $\|\delta\mu_t^{(k)}\|\leqslant 2$，故在 $(\varepsilon,\lambda)\to(\hat{\varepsilon},\hat{\lambda})$ 时，有

$$\int_R \|\delta\mu_t(\varepsilon,\lambda)-\delta\mu_t(\hat{\varepsilon},\hat{\lambda})\|dt=\int_R\left\|\sum_{k=1}^{p}(\varepsilon\lambda_k-\hat{\varepsilon}\hat{\lambda}_k)\delta\mu_t^{(k)}\right\|dt$$
$$\leqslant\sum_{k=1}^{p}|\varepsilon\lambda_k-\hat{\varepsilon}\hat{\lambda}_k|\int_I\|\delta\mu_t^{(k)}\|dt\longrightarrow 0.$$

这表明，摄动族(6.4)强连续(因此弱连续)地依赖于$(\varepsilon,\lambda)\in[0,\varepsilon_0]\times\Lambda$。

此外，当 $\varepsilon\to 0$ 时，摄动 $\delta\mu_t(\varepsilon,\lambda)$ 关于 $\lambda\in\Lambda$ 一致地强收敛（从而弱收敛）于零。这可由下式得知，

$$\|\delta\mu_t(\varepsilon,\lambda)\|\leqslant\varepsilon\cdot\text{const},\forall t\in\mathbf{R},\forall\lambda\in\Lambda,$$

而上述不等式成立是由于 Λ 的有界性以及 $t\notin I$ 时有 $\delta\mu_t(\varepsilon,\lambda)=0$。

6.2 轨线的变分

下面的命题是命题 6.1、连续依赖性定理 4.4 及解的变分定理 5.1 的一个直接推论。

命题 6.4 设 $\{\delta\mu_t(\sigma):\sigma\in\Sigma\}\subset\delta\mathfrak{M}_{\tilde{\mu}}$，

是控制 $\tilde{\mu}_t$ 的一族有限的摄动，设 $V_{\tilde{x}(t)}$ 是方程

$$\dot{x}=\langle\tilde{\mu}_t,f(t,x,u)\rangle=F(t,x)$$

的解 $\tilde{x}(t)(t\in[t_1,t_2])$的任意一个紧邻域，则存在正数 $\varepsilon_1>0$，使得对所有的 $\varepsilon\in[0,\varepsilon_1]$，以及任一满足条件

$$\|\langle\delta\mu_t(\sigma),f(t,x,u)\rangle\|_{w,V_{\tilde{x}(t)}}\leqslant\varepsilon \tag{6.5}$$

的 $\delta\mu_t(\sigma)$，摄动方程

$$\dot{x}=F(t,x)+\langle\delta\mu_t(\sigma),f(t,x,u)\rangle$$

有解

$$x(t)=x(t;\delta\mu_\theta(\sigma)),\quad t\in[t_1,t_2],$$

$$\sigma\in\Sigma,\quad x(t_1;\delta\mu_\theta(\sigma))=\tilde{x}(t_1),$$

且它满足如下的估计：

$$\max_{t\in[t_1,t_2]}|x(t;\delta\mu_s(\sigma))-\tilde{x}(t)|=\|x(t;\delta\mu_s(\sigma))-\tilde{x}(t)\|\longrightarrow 0$$

$$(\varepsilon\to 0).$$

如果条件(6.5)改成更强的条件

$$\|\langle\delta\mu_t(\sigma),f(t,x,u)\rangle\|_{1,V_{\tilde{x}(t)}}\leqslant\varepsilon,$$

则如下的估计成立：

$$\sup_{t\in[t_1,t_2];\sigma\in\Sigma}|x(t;\delta\mu_s(\sigma))-\tilde{x}(t)$$

$$-\Gamma(t)\int_{t_1}^{t}G(\theta)\langle\delta\mu_s(\sigma),f(\theta,\tilde{x}(\theta),u)\rangle d\theta|=o(\varepsilon),$$

其中，

$$o(\varepsilon)/\varepsilon\to 0(\varepsilon\to 0).$$

命题 6.5 设 Σ 是一距离空间，摄动族

$$\{\delta\mu_t(\sigma);\sigma\in\Sigma\}$$

是有限的，且关于参数 $\sigma\in\Sigma$ 弱连续，即 $\sigma_i\to\hat{\sigma}(i\to\infty)$ 蕴涵摄动差 $\delta\mu_t(\sigma_i)-\delta\mu_t(\hat{\sigma})$ 弱收敛于零．又设对每一 $\sigma\in\Sigma$，方程

$$\dot{x}=F(t,x)+\langle\delta\mu_t(\sigma),f(t,x,u)\rangle$$

有解

$$x(t;\sigma),t\in[t_1,t_2],$$

$$\sigma\in\Sigma,x(t_1;\sigma)\equiv\tilde{x}(t_1).$$

那么，函数 $x(t;\sigma)$ 关于 $(t,\sigma)\in[t_1,t_2]\times\Sigma$ 连续

证 设 $\sigma_i\to\hat{\sigma}(i\to\infty)$．$V_{x(t;\hat{\sigma})}$ 表示曲线 $x(t;\hat{\sigma})(t\in[t_1,t_2])$ 的任一紧邻域。按命题 6.2[注1]，

$$\|\langle\delta\mu_t(\sigma_i)-\delta\mu_t(\hat{\sigma}),f(t,x,u)\rangle\|_{w,V_{x(t;\hat{\sigma})}}\longrightarrow 0\qquad(i\to\infty).$$

又由命题 6.1，函数集合

$$\{\langle\delta\mu_t(\sigma_i)-\delta\mu_t(\hat{\sigma}),f(t,x,u)\rangle;i=1,2,\cdots\}$$

［注1］ 可视命题 6.2 中的 $\delta\mu_t^{(i)}$ 为此处的 $\delta\mu_t^{(\sigma_i)}$ 且 $\sigma=\mathrm{const}$．——译者注．

在 $\boldsymbol{R}\times\boldsymbol{R}^n$ 的任一紧子集的一个邻域上是一致 Lipschitz 的。因此，根据解的连续依赖性定理 4.4，我们得出结论：方程

$$
\begin{aligned}
\dot{x} &= F(t,x)+\langle\delta\mu_t(\sigma_i),f(t,x,u)\rangle\\
&= F(t,x)+\langle\delta\mu_t(\hat{\sigma}),f(t,x,u)\rangle+\langle\delta\mu_t(\sigma_i)\\
&\quad -\delta\mu_t(\hat{\sigma}),f(t,x,u)\rangle, \qquad i=1,2,\cdots,
\end{aligned}
$$

的解序列 $x(t;\sigma_i)(t\in[t_1,t_2])$ 满足如下条件：

$$
\max_{t\in[t_1;t_2]}|x(t;\sigma_i)-x(t;\hat{\sigma})|\longrightarrow 0(i\to\infty).
$$

这就证明了本命题。

下面，我们来证明凸控制问题中有关轨线变分族构造的两条基本定理。它们在证最大值原理时要用到。

在这两条定理中，我们始终保持方程

$$
\dot{x}=\langle\tilde{\mu}_t,f(t,x,u)\rangle=F(t,x)
$$

的一个解 $\tilde{x}(t)(t\in[t_1,t_2])$ 不变。在定理 6.2 中，还假定参数集合 Σ 是距离空间。

定理 6.1 设 $\delta\mu_t(\varepsilon,\lambda)$ 是(6.4)给出的有限的摄动族，则当 $\varepsilon\geqslant 0$ 充分小时（如 $0\leqslant\varepsilon\leqslant\varepsilon_1,\varepsilon_1>0$），摄动方程

$$
\begin{aligned}
\dot{x} &= F(t,x)+\langle\delta\mu_t(\varepsilon,\lambda),f(t,x,u)\rangle\\
&\equiv F(t,x)+\varepsilon\sum_{k=1}^{p}\lambda_k\langle\delta\mu_t^{(k)},f(t,x,u)\rangle,\lambda\in\Lambda
\end{aligned}
$$

满足初值条件 $x(t_1;\varepsilon,\lambda)\equiv\tilde{x}(t_1)$ 的变分解族

$$
x(t;\varepsilon,\lambda)\equiv x(t;\delta\mu_\theta(\varepsilon,\lambda)),t\in[t_1,t_2],\lambda\in\Lambda,
$$

是 $(t,\varepsilon,\lambda)\in[t_1,t_2]\times[0,\varepsilon_1]\times\Lambda$ 上的连续函数，它可表示成如下形式：

$$
x(t;\varepsilon,\lambda)=\tilde{x}(t)+\varepsilon\sum_{k=1}^{p}\lambda_k\delta x(t;\delta\mu_\theta^{(k)})+\Delta_2 x(t;\varepsilon,\lambda),
$$

其中，$\delta x(t;\delta\mu_\theta^{(k)})$ 是变分方程

$$
\delta\dot{x}=F_x(t,\tilde{x}(t))\delta x+\langle\delta\mu_t^{(k)},f(t,\tilde{x}(t),u)\rangle
$$

的具有零初始条件

$$\delta x(t_1;\delta\mu_\theta^{(k)})=0$$

的解，而 $\Delta_2 x(t;\varepsilon,\lambda)$ 满足如下条件，

$$|\Delta_2 x(t;\varepsilon,\lambda)|\leqslant o(\varepsilon),\forall(t,\lambda)\in[t_1,t_2]\times\Lambda,$$

$$o(\varepsilon)/\varepsilon\to 0,(\varepsilon\to 0).$$

证 摄动族 $\delta\mu_t(\varepsilon,\lambda)$ 在 $\varepsilon\to 0$ 时关于 $\lambda\in\Lambda$ 一致地强收敛于零。因此，根据命题 6.3，对任一紧集 $K\subset\boldsymbol{R}\times\boldsymbol{R}^n$，有

$$\|\langle\delta\mu_t(\varepsilon,\lambda),f(t,x,u)\rangle\|_{1,K}\longrightarrow 0(\varepsilon\to 0).$$

于是，从命题 6.4 与 6.5 可直接得到本定理。

定理 6.2 存在正数 $\varepsilon_0>0$，它仅与方程（6.1）本身、与解 $\bar{x}(t)$ $(t\in[t_1,t_2])$、与该解的任一紧邻域 $V_{\tilde{x}(t)}$ 以及 $\boldsymbol{R}^r$ 中的有界集 N 有关，使得对一切满足

$$\|\langle\delta\mu_t(\sigma),f(t,x,u)\rangle\|_{w,V_{\tilde{x}(t)}}\leqslant\varepsilon_0,\qquad\forall\sigma\in\Sigma,\qquad(6.6)$$

的有限的弱连续摄动族 $\delta\mu_t(\sigma)(\sigma\in\Sigma)$，以及一切关于 $\sigma\in\Sigma$ 一致地弱收敛于 $\delta\mu_t(\sigma)$ 的有限的弱连续摄动族序列

$$\delta\mu_t^{(i)}(\sigma),\sigma\in\Sigma,i=1,2,\cdots$$

$(\delta\mu_t^{(i)}(\sigma)$ 测度都集中在同一有界集 $N\subset\boldsymbol{R}^r$ 之中)，只要 i 充分大，摄动方程

$$\dot{x}=F(t,x)+\langle\delta\mu_t(\sigma),f(t,x,u)\rangle$$

与

$$\dot{x}=F(t,x)+\langle\delta\mu_t^{(i)}(\sigma),f(t,x,u)\rangle$$

在 t_1 通过 $\tilde{x}(t_1)$ 的解 $x(t;\sigma)$ 与 $x^{(i)}(t;\sigma)$ 均在 $[t_1,t_2]\times\Sigma$ 上存在且

$$\sup_{\substack{t\in[t_1,t_2]\\ \sigma\in\Sigma}}|x^{(i)}(t;\sigma)-x(t;\sigma)|\longrightarrow 0\qquad(i\to+\infty),\qquad(6.7)$$

此外，$x(t;\sigma)$ 与 $x^{(i)}(t;\sigma)$（i 充分大）在 $[t_1,t_2]\times\Sigma$ 上关于 (t,σ) 连续.

证 由弱一致收敛性 $\delta\mu_t^{(i)}(\sigma)\to\delta\mu_t(\sigma)$，根据命题 6.2，在(6.6)成立时，必存在自然数 i_1，当 $i\geqslant i_1$ 时下式成立：

$$\|\langle\delta\mu_t^{(i)}(\sigma),f(t,x,u)\rangle\|_{w,V_{\tilde{x}(t)}}\leqslant 2\varepsilon_0.$$

但根据 $\delta\mu_t^{(i)}(\sigma)$ 与 $\delta\mu_t(\sigma)$ 的有限性，应用命题 6.1 及连续依赖性定理

4.4，存在充分小的 $\varepsilon_0>0$，在 $i\geqslant i_1$ 及 (6.6) 之下，解族 $x(t;\sigma)$ 与 $x^{(i)}(t;\sigma)$ 在 $[t_1,t_2]\times\Sigma$ 上存在且在紧邻域 $V_{\tilde{x}(t)}$ 之中。从命题6.5 得在两解族关于 (t,σ) 的连续性。

记

$$\Delta x^{(i)}(t;\sigma)=x^{(i)}(t;\sigma)-x(t;\sigma),\quad t\in[t_1,t_2],\sigma\in\Sigma,$$
$$\delta v_t^{(i)}(\sigma)=\delta\mu_t^{(i)}(\sigma)-\delta\mu_t(\sigma).$$

关于 $\Delta x^{(i)}(t;\sigma)$ 的微分方程是：

$$\begin{aligned}\Delta\dot{x}^{(i)}&=F(t;x(t;\sigma)+\Delta x^{(i)})-F(t,x(t;\sigma))+\langle\delta\mu_t(\sigma),\\&\quad f(t,x(t;\sigma)+\Delta x^{(i)},u)-f(t,x(t;\sigma),u)\rangle\\&\quad+\langle\delta v_t^{(i)}(\sigma),f(t,x^{(i)}(t;\sigma),u)\rangle\\&=\left\{\int_0^1F_x(t,x(t;\sigma)+s\Delta x^{(i)})ds+\int_0^1\langle\delta\mu_t(\sigma),f_x(t,x(t;\sigma)\right.\\&\quad\left.+s\Delta x^{(i)},u)\rangle ds\right\}\Delta x^{(i)}+\langle\delta v_t^{(i)}(\sigma),f(t,x^{(i)}(t;\sigma),u)\rangle\end{aligned}$$

将上式括号内的表示式记为 $A^{(i)}(t;\sigma)$，即

$$\begin{aligned}A^{(i)}(t;\sigma)=\int_0^1[F_x(t,x(t;\sigma)+s\Delta x^{(i)}(t;\sigma))+\langle\delta\mu_t(\sigma),\\f_x(t,x(t;\sigma)+s\Delta x^{(i)}(t;\sigma),u)\rangle]ds,\\t\in[t_1,t_2],\quad\sigma\in\Sigma,\quad i\geqslant i_1.\end{aligned}$$

按假定，$\delta\mu_t(\sigma)$ 与 $\delta\mu_t^{(i)}(\sigma)$ 的测度都集中在 $\boldsymbol{R}^r$ 的同一个有界子集上。我们还可假设所有曲线 $x(t;\sigma)+s\Delta x^{(i)}(t;\sigma)$ $(t\in[t_1,\ t_2],\ 0\leqslant s\leqslant 1,\ \sigma\in\Sigma,i\geqslant i_1)$ 都属于紧邻域 $V_{\tilde{x}(t)}$．因此，绝对值

$$|A^{(i)}(t;\sigma)|,\qquad t\in[t_1,t_2],\qquad \sigma\in\Sigma,\qquad i\geqslant i_1,$$

不超过某一常数 C.

按定理 5.2，齐次方程

$$\delta\dot{x}=A^{(i)}(t;\sigma)\delta x$$

的基本解矩阵 $\Gamma^{(i)}(t;\sigma)(t\in[t_1,t_2])$ 以及其逆阵 $G^{(i)}(t;\sigma)$ 的范数以某一常数 C_1 为界，即

$$|\Gamma^{(i)}(t;\sigma)|\leqslant C_1,\qquad |G^{(i)}(t;\sigma)|\leqslant C_1,\qquad \forall t\in[t_1,t_2],$$
$$\forall\sigma\in\Sigma,\qquad \forall i\geqslant i_1.$$

将$\Delta x^{(i)}(t;\sigma)$表示为如下形式，

$$\Delta x^{(i)}(t;\sigma)=\Gamma^{(i)}(t;\sigma)\int_{t_1}^{t}G^{(i)}(\theta;\sigma)\langle\delta v_{\theta}^{(i)}(\sigma),f(\theta,x^{(i)}(\theta;\sigma),u)\rangle d\theta.$$

分部积分得

$$\Delta x^{(i)}(t;\sigma)=\int_{t_1}^{t}\langle\delta v_{\theta}^{(i)}(\sigma),f(\theta,x^{(i)}(\theta;\sigma),u)\rangle d\theta$$

$$+\Gamma^{(i)}(t;\sigma)\int_{t_1}^{t}G^{(i)}(\theta;\sigma)A^{(i)}(\theta;\sigma)$$

$$\left[\int_{t_1}^{\theta}\langle\delta v_{\theta_1}^{(i)}(\sigma),f(\theta_1,x^{(i)}(\theta_1;\sigma),u)\rangle d\theta_1\right]d\theta,$$

于是，

$$\|\Delta x^{(i)}(t;\sigma)\|\leqslant\{1+CC_1^2(t_2-t_1)\}$$

$$\left\|\int_{t_1}^{t}\langle\delta v_{\theta_1}^{(i)}(\sigma),f(\theta,x^{(i)}(\theta;\sigma),u)\rangle d\theta\right\|,$$

其中

$$\left\|\int_{t_1}^{t}\langle\delta v_{\theta}^{(i)}(\sigma),f(\theta,x^{(i)}(\theta;\sigma),u)\rangle d\theta\right\|$$

$$=\max_{t\in[t_1,t_2]}\left|\int_{t_1}^{t}\langle\delta v_{\theta}^{(i)}(\sigma),f(\theta,x^{(i)}(\theta;\sigma),u)\rangle d\theta\right|.$$

按定理的假设，当$i\to\infty$时序列$\delta v_t^{(i)}(\sigma)$关于$\sigma\in\Sigma$一致地弱收敛于零。因此，命题6.2指出，关于$\sigma\in\Sigma$一致地有

$$\|\langle\delta v_t^{(i)}(\sigma),f(t,x,u)\rangle\|_{\omega,V_{\tilde{x}(t)}}\longrightarrow 0\quad(i\to\infty).$$

另一方面，$x^{(i)}(t;\sigma)$满足Lipschitz条件且其Lipschitz常数L与$\sigma\in\Sigma$及$i\geqslant i_1$无关：

$$|x^{(i)}(t';\sigma)-x^{(i)}(t'';\sigma)|\leqslant\left|\int_{t'}^{t''}|F(t,x^{(i)}(t,\sigma))|dt\right|$$

$$+\left|\int_{t'}^{t''}|\langle\delta\mu_t^{(i)}(\sigma),f(t,x^{(i)}(t;\sigma),u)\rangle|dt\right|$$

$$\leqslant\left|\int_{t'}^{t''}\sup_{(t,x)\in V_{\tilde{x}(t)}}\{|F(t,x)|+|\langle\delta\mu_t^{(i)}(\sigma),f(t,x,u)\rangle|\}dt\right|$$

$$\leqslant L|t'-t''|.$$

因此，函数族$x^{(i)}(t;\sigma)(i\geqslant i_1)$具有高度连续性。

这样，我们证明了关于$\sigma\in\Sigma$一致地成立

$$\|\langle \delta\nu_t^{(i)}(\sigma), f(t,x,u)\rangle\|_{w,V_{\tilde{x}(t)}} \longrightarrow 0 \quad (i\to\infty),$$

以及函数族$x^{(i)}(t;\sigma)$的高度连续性。根据命题 4.1 便可得(注1)

$$\left\| \int_{t_1}^{t} \langle \delta\nu_\theta^{(i)}(\sigma), f(\theta, x^{(i)}(\theta;\sigma), u)\rangle d\theta \right\| \longrightarrow 0, \quad (i\to\infty). \tag{6.8}$$

由此得出$\|\Delta x^{(i)}(t;\sigma)\|$关于$\sigma\in\Sigma$一致地收敛于零．这就证明了(6.7)．定理证完。

[注1] 这里实际上用到命题4.1的以下变型：

设 $[t_1,t_2]$ 上的函数族 $\Phi=\{x(\cdot)\}$ 等度连续且存在一个紧集 $K\subset R\times R^n$，使得一切 $t\in[t_1,t_2]$，$x(\cdot)\in\Phi$ 都有 $(t,x(t))\in K$．又 M 是 $E_1(R\times R^n)$ 的一个子集，它在 K 的一个邻域上一致 Lipschitz，则

$$\sup_{\substack{t',t''\in[t_1,t_2]\\ x(\cdot)\in\Phi}} \left| \int_{t''}^{t'} \delta F(t,x(t))dt \right| \to 0,$$

$$\|\delta F\|_{w,K}\to 0, \delta F\in M.$$

其证明与命题4.1类似，在此从略。——译者注。

第七章　最大值原理的证明

在本章我们将证明凸最优控制问题的积分形式的最大值原理并由此导出第一章中的最大值原理。

当控制类是 Ω_U 与 $\mathfrak{M}_U$ 时，积分最大值原理与庞特里雅金最大值原理是等价的。但对更一般的控制类，两者不一定等价。一般说来，积分最大值原理适用范围更广泛些。

7.1　积分最大值条件与庞特里雅金最大值条件及两者之间的等价性

设

$$H(t,x,\psi,u)=\psi f(t,x,u)$$

是最优问题(1.3)的 Hamilton 函数，相应的典则方程是

$$\dot{x}=\frac{\partial H}{\partial \psi}=f(t,x,u),$$

$$\dot{\psi}=-\frac{\partial H}{\partial x}=-\psi f_x(t,x,u).$$

在第一章中，上式中的 u 取允许控制 $u(t)\in\Omega_U$. 现在取广义控制 μ_t 来代入 u:

$$\dot{x}=\frac{\partial H}{\partial \psi}=\langle\mu_t,f(t,x,u)\rangle,$$

$$\dot{\psi}=-\frac{\partial H}{\partial x}=-\psi\langle\mu_t,f_x(t,x,u)\rangle. \tag{7.1}$$

特别地，当 $\mu_t=\delta_{u(t)}$ 时，便回到前一系统。

μ_t 不是典则方程组(7.1)所能唯一确定的。对 x 和 ψ 给定的初值，要使这一方程组存在唯一解，还需引进一个附加条件以"消去" μ_t. 积分最大值条件与庞特里雅金最大值条件就是这种相互等价的附加条件。

设 $x(t)$ 与 $\psi(t)$ 是 $[t_1,t_2]$ 上的连续函数。如果一个广义控制 $\hat{\mu}_t\in\mathfrak{M}_U$ 满足

$$\int_{t_1}^{t_2}\langle\hat{\mu}_t,H(t,x(t),\psi(t),u)\rangle dt=\sup_{\mu_t\in\mathfrak{M}_U}\int_{t_1}^{t_2}\langle\mu_t,H(t,x(t),\psi(t),u)\rangle dt,\tag{7.2}$$

就称 $\hat{\mu}_t$ 沿 $x(t),\psi(t)(t\in[t_1,t_2])$ 满足积分最大值条件。如果对几乎所有的 $t\in[t_1,t_2]$,

$$\langle\hat{\mu}_t,H(t,x(t),\psi(t),u)\rangle=\sup_{\mu}\langle\mu,H(t,x(t),\psi(t),u)\rangle,$$

其中的上确界是对所有测度集中在 $U\subset\boldsymbol{R}^r$ 上的概率测度 μ 取的，那么就称 $\hat{\mu}_t$ 满足庞特里雅金最大值条件。

我们将证明，对任一 $t\in[t_1,t_2]$,

$$\sup_{\mu}\langle\mu,H(t,x(t),\psi(t),u)\rangle=\sup_{u\in U}H(t,x(t),\psi(t),u)$$
$$=M(t,x(t),\psi(t)),$$

于是，庞特里雅金最大值条件可以改写成如下形式：对几乎所有的 $t\in[t_1,t_2]$,

$$\langle\hat{\mu}_t,H(t,x(t),\psi(t),u)\rangle=\sup_{\mu}\langle\mu,H(t,x(t),\psi(t),u)\rangle$$
$$=M(t,x(t),\psi(t)).\tag{7.3}$$

显然有不等式

$$\sup_{\mu}\langle\mu,H(t,x(t),\psi(t),u)\rangle=\sup_{\mu}\psi(t)\langle\mu,f(t,x(t),u)\rangle$$
$$\geqslant M(t,x(t),\psi(t)).$$

事实上，这只要取 $\mu=\delta_u$，其中 u 是 U 中任意一点。

现来证反向不等式。首先注意到，集中在 U 上的任一概率测度 μ 满足条件（见命题 2.1）

$$\langle\mu,f(t,x(t),u)\rangle\in\mathrm{conv}\{f(t,x(t),u);u\in U\}$$
$$=\mathrm{conv}P(t,x(t)).$$

因此，

$$\psi(t)\langle\mu,f(t,x(t),u)\rangle\leqslant\sup_{p\in\mathrm{conv}P(t,x(t))}\psi(t)p.$$

然而，对任一 n 维行向量 ψ 和任一子集 $P\subset\mathbf{R}^n$，有

$$\sup_{p\in\mathrm{conv}P}\psi p=\sup_{p\in P}\psi p\text{[注1]}.$$

这样，

$$\psi(t)\langle\mu,f(t,x(t),u)\rangle\leqslant\sup_{p\in P(t,x(t))}\psi(t)p=\sup_{u\in U}\psi(t)f(t,x(t),u).$$

因此，(7.3) 式成立。

命题 7.1 积分形式最大值条件(7.2)与庞特里雅金形式最大值条件(7.3)是等价的。

证 先证明(7.2)⇒(7.3).假定 $\theta\in[t_1,t_2]$是函数

$$F(t)=\langle\hat{\mu}_t,H(t,x(t),\psi(t),u)\rangle,t_1\leqslant t\leqslant t_2$$

的任一勒贝格点。对任一充分小的 $h>0$，定义广义控制 $\mu_t^h\in\mathfrak{M}_U$ 如下，

$$\mu_t^h=\hat{\mu}_t,\qquad t_1\leqslant t\leqslant\theta \text{ 或 } \theta+h\leqslant t\leqslant t_2;$$
$$\mu_t^h=\mu,\qquad \theta<t<\theta+h.$$

其中的 μ 是测度集中在 U 中的概率测度。由(7.2)得到

$$\frac{1}{h}\int_\theta^{\theta+h}\langle\hat{\mu}_t,H(t,x(t),\psi(t),u)\rangle dt$$

$$\geqslant\frac{1}{h}\int_\theta^{\theta+h}\langle\mu,H(t,x(t),\psi(t),u)\rangle dt.$$

注意到上式右侧的被积函数关于 t 连续，令 $h\to0$ 取极限，即得不等

[注1] 由于

$$\sup_{p\in P}\psi p\leqslant\sup_{p\in\mathrm{conv}P}\psi p,$$

不妨设不等式左侧为有限值 S.于是，集合 P 在半空间 $\psi P\leqslant S$ 之中，因此 conv P 也在这半空间之中，这表明相反不等式也成立：

$$\sup_{p\in\mathrm{conv}P}\psi p\leqslant S=\sup_{p\in P}\psi p.$$

式

$$\langle \hat{\mu}_\theta, H(\theta, x(\theta), \psi(\theta), u)\rangle \geqslant \langle \mu, H(\theta, x(\theta), \psi(\theta), u)\rangle$$

对所有的勒贝格点 $\theta\in[t_1,t_2]$成立。由于$[t_1,t_2]$中的勒贝格点全体组成的集是全测度的，这就证明了条件(7.3)。

再证明(7.3)⇒(7.2)。将下式（对几乎所有的 $t\in[t_1,t_2]$ 成立）关于 t 从 t_1 到 t_2 积分便可得(7.2)：

$$\begin{aligned}\langle \hat{\mu}_t, H(t, x(t), \psi(t), u)\rangle &= \sup_{\mu}\langle \mu, H(t, x(t), \psi(t), u)\rangle \\ &\geqslant \langle \mu_t, H(t, x(t), \psi(t), u)\rangle,\end{aligned}$$

其中，$\mu_t\in\mathfrak{M}_U$ 是任一广义控制。

若在(7.3)式中作代换

$$\hat{\mu}_t = \delta_{\hat{u}(t)}, \qquad \hat{u}(t)\in\Omega_U$$

则它就回复成最大值条件(1.9)。

7.2 广义控制类中的最大值原理

考虑凸最优控制问题

$$\begin{aligned}&\dot{x} = \langle \mu_t, f(t,x,u)\rangle, \mu_t\in\mathfrak{M}_U,\\ &x(t_1) \quad x_1, x(t_{\_}) = x_2,\\ &\qquad t_2 - t_1 \longrightarrow \min.\end{aligned}\tag{7.4}$$

类似于第一章极值解的定义，如果 $\tilde{\mu}_t$, $\tilde{\psi}(t)\neq 0$, $\bar{x}(t)$ $(t_1\leqslant t\leqslant t_2)$ 满足典则方程(7.1)及边值条件

$$\tilde{x}(t_1) = x_1, \quad \tilde{x}(t_2) = x_2,$$

并且满足最大值条件 (7.2) (或与它等价的(7.3))，则称 $\tilde{\mu}_t$, $\tilde{\psi}(t)$ $\tilde{x}(t)$ $(t_1\leqslant t\leqslant t_2)$为凸最优控制问题(7.4)的一个极值解。

如果 $\tilde{\mu}_t$ 是常义控制，即 $\tilde{\mu}_t=\delta_{\tilde{u}(t)}$，则凸问题 (7.4) 的一个极值解

$$\delta_{\tilde{u}(t)},\ \tilde{x}(t),\tilde{\psi}(t)$$

也是原最优问题(1.3)的极值解.

凸问题 (7.4) 的最大值原理虽然可这样表述：如果 $\tilde{\mu}_t,\tilde{x}(t)$ 是(7.4)的解，则必存在一绝对连续函数 $\tilde{\psi}(t)\neq 0$ 使得函数组 $\tilde{\mu}_t,\tilde{x}(t),\tilde{\psi}(t)$是它的一个极值解。但不难看出，这一命题（虽则成立）不足以推出第一章的最大值原理。这是因为 $\tilde{\mu}_t=\delta_{\tilde{u}(t)}$ 在广义控制类 $\mathfrak{M}_U$ 中是最优时，该命题才能保证存在 $\tilde{\psi}(t)$具有所需的诸性质。而第一章的最大值原理只要求 $\delta_{\tilde{u}(t)}$在较窄一类控制 Ω_U 中是最优时就保证 $\tilde{\psi}(t)$存在。因此，为了导出第一章的最大值原理，我们必须给出并证明凸问题(7.4)的一个更强的必要条件。

在广义控制 $\tilde{\mu}_t$ 下，方程

$$\dot{x}=\langle\tilde{\mu}_t,f(t,x,u)\rangle$$

满足边值条件 $x(t_1)=x_1,x(t_2)=x_2$ 的解记为 $\tilde{x}(t)(t\in[t_1,t_2])$. 如果不存在允许控制 $\delta_{u(t)}\in\Omega_U$ 使得方程

$$\dot{x}=\langle\delta_{u(t)},f(t,x,u)\rangle=f(t,x,u(t))$$

满足初值条件 $x(t_1)=x_1$ 的解 $x(t)$ 在小于 t_2 的时刻到达 x_2，则称 $\tilde{\mu}_t,\ \tilde{x}(t)(t\in[t_1,t_2])$是问题(7.4)的一个弱最优解。

很明显，凸最优问题 (7.4) 的任一解是弱最优的. 因此“弱最优性”的概念实际上弱于最优性的概念。但显然在常义控制类中，这两种最优性是一致的.

下面来叙述凸最优问题的最大值原理。

定理 7.1 （凸最优问题的最大值原理）

设 $\tilde{\mu}_t,\tilde{x}(t)$是凸最优问题 (7.4) 的弱最优解，则必存在非零的绝对连续函数 $\tilde{\psi}(t)$,使得

$$\tilde{\mu}_t,\quad \tilde{\psi}(t),\qquad \tilde{x}(t),$$

是凸最优问题(7.4)的一个极值解。此外，t 的函数

$$M(t,\tilde{x}(t),\tilde{\psi}(t))=\sup_{u\in U} H(t,\tilde{x}(t),\tilde{\psi}(t),u)$$

在整个区间 $t_1\leqslant t\leqslant t_2$ 上是连续的，并且

$$M(t_2,\tilde{x}(t_2),\tilde{\psi}(t_2))\geqslant 0.$$

7.3 变分锥的构造

在证明定理 7.1 之前，我们先来构造所谓的一阶变分锥，并给出它的基本性质。

设 $\tilde{\mu}_t$ 是任意一个（未必最优的）广义控制，并设 $\tilde{x}(t)(t\in[t_1,t_2])$是方程

$$\dot{x}=\langle\tilde{\mu}_t,f(t,x,u)\rangle=F(t,x) \tag{7.5}$$

的满足边值条件

$$\tilde{x}(t_1)=x_1,\qquad \tilde{x}(t_2)=x_2$$

的轨线.

如同第六章，我们将用记号

$$\delta x(t;\delta\mu_s),\quad t\in[t_1,t_2],\quad \delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}},$$

表示变分方程

$$\delta\dot{x}=F_x(t,\tilde{x}(t))\delta x+\langle\delta\mu_t,f(t,\tilde{x}(t),u)\rangle$$

的在零初值条件

$$\delta x(t_1;\delta\mu_s)=0$$

下的解。虽然函数 $\tilde{x}(t)$ 在点 t_2 不一定可微，但是适当地选择收敛于零的数列 η_j,可使下面的极限存在且有限:

$$\lim_{j\to\infty}\frac{1}{\eta_j}[\tilde{x}(t_2)-\tilde{x}(t_2-\eta_j)],\quad \eta_j>0.$$

这是因为对所有的 t, $\tilde{\mu}_t$ 的单位测度集中在同一有界集上，因此,

$$|\tilde{x}(t_2)-\tilde{x}(t_2-\eta)|\leqslant\int_{t_2-\eta}^{t_2}|\dot{\tilde{x}}(t)|dt=\int_{t_2-\eta}^{t_2}|\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle|dt$$
$$\leqslant\eta\cdot\text{const}.$$

任意选定这样的一个极限并记之为 F。借助于向量 F，构造在 $\boldsymbol{R}^n$ 中的集合：

$$Q=\{\gamma\delta x(t_2;\delta\mu_\theta)-\theta F;\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},\gamma\geqslant0,\theta\geqslant0\}.\qquad(7.6)$$

不难看出，Q 是 $\boldsymbol{R}^n$ 中以原点为顶点的凸锥，即对一切 $\delta x',\delta x''\in Q$ 和 $\alpha,\beta\geqslant0$，必有 $\alpha\delta x'+\beta\delta x''\in Q$。事实上，已知 $\delta\mathfrak{M}_{\tilde{\mu}_t}$ 是包含原点的凸集，因此，

$$\alpha(\gamma'\delta x(t_2;\delta\mu'_\theta)-\theta'F)+\beta(\gamma''\delta x(t_2;\delta\mu''_\theta)-\theta''F)$$
$$=\delta x(t_2;\alpha\gamma'\delta\mu'_\theta+\beta\gamma''\delta\mu''_\theta)-(\alpha\theta'+\beta\theta'')F$$
$$=\frac{1}{\varepsilon}\delta x(t_2;\varepsilon\alpha\gamma'\delta\mu'_\theta+\varepsilon\beta\gamma''\delta\mu''_\theta)-(\alpha\theta'+\beta\theta'')F,\quad\varepsilon>0,$$

其中对充分小的 $\varepsilon>0$，有

$$\varepsilon\alpha\gamma'\delta\mu'_t+\varepsilon\beta\gamma''\delta\mu''_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},$$

所以，Q 是凸锥。

Q 是集合

$$\{\delta x(t_2;\delta\mu_\theta),-F;\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t}\}\subset\boldsymbol{R}^n$$

的凸包张射而成的以原点为顶点的锥，即由点集

$$\alpha\delta x(t_2;\delta\mu_\theta)-\beta F,\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},\alpha\geqslant0,\beta\geqslant0,\alpha+\beta=1,$$

与原点连接张射而成的集合。锥 Q 称为凸控制问题

$$\dot{x}=\langle\mu_t,f(t,x,u)\rangle,\mu_t\in\mathfrak{M}_U,$$

在轨线 $\tilde{x}(t)(t\in[t_1,t_2])$ 的右端点 $\tilde{x}(t_2)$ 的一阶变分锥。

如果函数 $\tilde{x}(t)$ 在点 t_2 可微，则向量

$$F=\frac{d\tilde{x}(t_2)}{dt},$$

是唯一的，从而锥 Q 的构造完全确定。例如，当点 t_2 是函数 $F(t,$

$\tilde{x}(t)$)的勒贝格点时便是如此。但在一般情形，Q 将依赖于 F 的具体选取。

变分锥 Q的顶点（即$\boldsymbol{R}^n$ 的原点）可以是集合 $Q\subset\boldsymbol{R}^n$ 的边界点，也可以是内点。在后一情形，集合 Q 显然重合于 $\boldsymbol{R}^n$.

下面一个重要命题表述了变分锥的基本性质。

命题 7.2 如果变分锥$Q=\boldsymbol{R}^n$，那末必存在终点 $\tilde{x}(t_2)$ 的一个充分小的邻域，使得这个邻域中的每一点都可被在时刻 t_1 出发于 x_1，并在某个逐段常值控制 $\delta_{u(t)}\in\Omega_U$ 作用下的轨线在时刻 t_2 之前达到。因此，必存在逐段常值控制 $u(t)$及时刻 $\hat{t}_2<t_2$,使得方程

$$\begin{cases}\dot{x}=\langle\delta_{u(t)},f(t,x,u)\rangle=f(t,x,u(t))\\ x(t_1)=\tilde{x}(t_1), \qquad x(\hat{t}_2)=\tilde{x}(t_2)\end{cases}$$

存在解 $x(t)(t\in[t_1,\hat{t}_2])$.

在下面的证明中，我们将本命题归结为一个n 维单纯形到自身的连续映射的不动点存在性。

证 选取 $n+1$ 点

$$\delta x^{(0)},\cdots,\delta x^{(n)},$$

它们组成包含原点为一内点的某个 n 维单纯形的顶点。我们记此单纯形为 X：

$$X=\left\{\delta x:\delta x=\sum_{k=0}^{n}\lambda_k\delta x^{(k)},\lambda_k\geqslant 0,\sum_{k=0}^{n}\lambda_k=1\right\}.$$

由于 $Q=\boldsymbol{R}^n$，故按(7.6)，必存在一些摄动与非负数

$$\delta\mu_t^{(k)}\in\delta\mathfrak{M}_{\tilde{x}t},\qquad \gamma^{(k)}\geqslant 0,\qquad \theta^{(k)}\geqslant 0,\qquad k=0,1,\cdots,n,$$

使得

$$\delta x^{(k)}=\delta x(t_2;\gamma^{(k)}\delta\mu_\theta^{(k)})-\theta^{(k)}F. \tag{7.7}$$

不失一般性，可假定所有的 $\theta^{(k)}>0,k=0,1,\cdots,n$. 事实上，作为$n$ 维单纯形的顶点，$\delta x^{(k)}(k=0,1,\cdots,n)$，处于最广位置上。因此，所有的 $\theta^{(k)}\geqslant 0$ 可以增加到 $\hat{\theta}^{(k)}>0$,并且增量可以小得使变动后的点

$$\delta\hat{x}^{(k)}=\delta x(t_2;\gamma^{(k)}\delta\mu_\theta^{(k)})-\hat{\theta}^{(k)}F,\qquad k=0,1,\cdots,n,$$

依然处于最广位置。于是，这些变动后的点所张成的单纯形也以原点为内点。

下面仅用到 $\delta\mu_t^{(k)}$ 在 $t\in[t_1,t_2]$ 中的值。不妨再假定 $\delta\mu_t^{(k)}=0$，$t\notin[t_1,t_2]$. 这样假定是为了便于引用第六章的一些命题。

以 $d>0$ 表示从原点到单纯形 X 的边界之距离。不妨假定

$$0<\theta^{(0)}\leqslant\theta^{(1)}\leqslant\cdots\leqslant\theta^{(n)}. \tag{7.8}$$

用 Λ 表示 n 维单纯形

$$\Lambda=\left\{\lambda=(\lambda_0,\cdots,\lambda_n):\lambda_k\geqslant 0,\ \sum_{k=0}^{n}\lambda_k=1\right\}.$$

根据定理 6.1 对一切充分小的 $\varepsilon\geqslant 0$（不妨设 ε 满足 $0\leqslant\varepsilon\leqslant\varepsilon_1$，其中 $\varepsilon_1>0$），摄动控制族

$$\delta\mu_t(\varepsilon,\lambda)=\varepsilon\sum_{k=0}^{n}\lambda_k\gamma^{(k)}\delta\mu_t^{(k)},\qquad \lambda=(\lambda_0,\cdots,\lambda_n)\in\Lambda,$$

所对应的摄动方程

$$\dot{x}=F(t,x)+\langle\delta\mu_t(\varepsilon,\lambda),f(t,x,u)\rangle$$

的解

$$x(t;\varepsilon,\lambda)=\tilde{x}(t)+\varepsilon\sum_{k=0}^{n}\lambda_k\delta x(t;\gamma^{(k)}\delta\mu_\theta^{(k)})+\Delta_2x(t;\varepsilon,\lambda),$$

$$t\in[t_1,t_2],$$

$$x(t_1;\varepsilon,\lambda)=x_1, \tag{7.9}$$

在 $[t_1,t_2]$ 上存在，这里，$\Delta_2x(t;\varepsilon,\lambda)$ 满足估计

$$\|\Delta_2x(t;\varepsilon,\lambda)\|\leqslant o(\varepsilon),\qquad\forall\lambda\in\Lambda,\qquad o(\varepsilon)/\varepsilon\to 0\qquad(\varepsilon\to 0).$$

此外，函数 $x(t;\varepsilon,\lambda)$ 连续依赖于

$$(t,\varepsilon,\lambda)\in[t_1,t_2]\times[0,\varepsilon_1]\times\Lambda.$$

现在证明在 Λ 上存在这样的正值连续函数 $\varepsilon(\lambda)$ 与 $\theta(\lambda)$，满足条件

$$0<\varepsilon(\lambda)\theta(\lambda)\leqslant t_2-t_1,\qquad\varepsilon(\lambda)\leqslant\varepsilon_1,\qquad\forall\lambda\in\Lambda,$$

并且有以下性质：存在正数 $\delta>0$ 与 $\boldsymbol{R}^n$ 的原点的一个邻域 $\hat{V}$，对一切满足不等式

$$|\omega(\lambda)|\leqslant\delta,\qquad\forall\lambda\in\Lambda,$$

的连续函数 $\omega(\lambda)$，λ 的方程

$$S(\lambda)+\omega(\lambda)=\delta x, \qquad \lambda\in\Lambda,$$

对一切 $\delta x\in\hat{V}$ 有解，这里

$$S(\lambda)=x(t_2-\varepsilon(\lambda)\theta(\lambda);\varepsilon(\lambda),\lambda)-\tilde{x}(t_2)$$

是 Λ 到 $\boldsymbol{R}^n$ 的连续映射，即单纯形 Λ 在映射

$$\lambda\longmapsto S(\lambda)+\omega(\lambda)$$

下的象集复盖了整个邻域 $\hat{V}$.

特别当 $\omega(\lambda)\equiv 0$ 时，我们得到以下推论：$\tilde{x}(t_2)$的邻域

$$\tilde{x}(t_2)+\hat{V}=\{x:x=\tilde{x}(t_2)+\delta x,\delta x\in\hat{V}\}$$

中的任一点可被(7.9)的某一轨线在时刻 $t_2-\varepsilon(\lambda)\theta(\lambda)<t_2$ 达到。

定义

$$\theta(\lambda)=\sum_{k=0}^{n}\lambda_k\theta^{(k)}, \qquad \lambda=(\lambda_0,\cdots,\lambda_n)\in\Lambda,$$

其中 $\theta^{(k)}$ 由(7.7)给出。显然，$\theta^{(0)}\leqslant\theta(\lambda)\leqslant\theta^{(n)}(\forall\lambda\in\Lambda)$.事实上，根据(7.8)即得

$$0<\theta^{(0)}\leqslant\theta^{(0)}+\sum_{k=1}^{n}\lambda_k(\theta^{(k)}-\theta^{(0)})=\theta(\lambda)$$

$$=\theta^{(n)}+\sum_{k=1}^{n}\lambda_k(\theta^{(k)}-\theta^{(n)})\leqslant\theta^{(n)}.$$

设

$$\nu_j=F-\frac{\tilde{x}(t_2)-\tilde{x}(t_2-\eta_j)}{\eta_j}, \tag{7.10}$$

则当 $j\to+\infty$ 时，$\eta_j\to+0,\nu_j\to 0$. 记正值连续函数

$$\varepsilon_j(\lambda)=\frac{\eta_j}{\theta(\lambda)}, \qquad \lambda\in\Lambda,$$

则在 j 充分大时，$\varepsilon_j(\lambda)\leqslant\varepsilon_1,\forall\lambda\in\Lambda$,且

$$\frac{\eta_j}{\theta^{(n)}}\leqslant\varepsilon_j(\lambda)\leqslant\frac{\eta_j}{\theta^{(0)}}, \qquad \forall\lambda\in\Lambda. \tag{7.11}$$

对这些充分大的 j，定义

$$z_j(\lambda)=x(t_2-\varepsilon_j(\lambda)\theta(\lambda);\varepsilon_j(\lambda),\lambda)-\tilde{x}(t_2)$$
$$=x(t_2-\eta_j;\varepsilon_j(\lambda),\lambda)-\tilde{x}(t_2),\qquad \lambda\in\Lambda.$$

根据(7.10)，$z_j(\lambda)$可以写成

$$z_j(\lambda)=x(t_2-\eta_j;\varepsilon_j(\lambda),\lambda)-\bar{x}(t_2-\eta_j)-\eta_j F+\eta_j\nu_j.$$

再按(7.9)式（取 $t=t_2-\eta_j,\varepsilon=\varepsilon_j(\lambda)$），并注意到 $\theta(\lambda)$的定义，不难得到

$$z_j(\lambda)=\varepsilon_j(\lambda)\left\{\sum_{k=0}^{s}\lambda_k[\delta x(t_2;\gamma^{(k)}\delta\mu_\theta^{(k)})-\theta^{(k)}F]\right.$$
$$+\sum_{k=0}^{s}\lambda_k[\delta x(t_2-\eta_j;\gamma^{(k)}\delta\mu_\theta^{(k)})-\delta x(t_2;\gamma^{(k)}\delta\mu_\theta^{(k)})]$$
$$\left.+\frac{\Delta_2 x(t_2-\eta_j;\varepsilon_j(\lambda),\lambda)}{\varepsilon_j(\lambda)}+\theta(\lambda)\nu_j\right\},$$
$$=\varepsilon_j(\lambda)\left\{\sum_{k=0}^{s}\lambda_k\delta x^{(k)}+R_j(\lambda)\right\}$$

其中，

$$R_j(\lambda)=\sum_{k=0}^{s}\lambda_k[\delta x(t_2-\eta_j;\gamma^{(k)}\delta\mu_\theta^{(k)})-\delta x(t_2;\gamma^{(k)}\delta\mu_\theta^{(k)})]$$
$$+\frac{\Delta_2 x(t_2-\eta_j;\varepsilon_j(\lambda),\lambda)}{\varepsilon_j(\lambda)}+\theta(\lambda)\nu_j.$$

当 $j\to\infty$ 时，$R_j(\lambda)$关于 $\lambda\in\Lambda$一致地收敛于零：

$$R_j(\lambda)\to 0\ (j\to\infty)$$ [注1]

于是，存在充分大的指标 $\hat{j}$，满足

$$|R_{\hat{j}}(\lambda)|\leqslant\frac{d}{3},\qquad \forall\lambda\in\Lambda.$$

记

$$\varepsilon(\lambda)=\varepsilon_{\hat{j}}(\lambda),\qquad R(\lambda)=R_{\hat{j}}(\lambda),$$

[注1] 这里，$R_j(\lambda)$ 右侧 $\dfrac{\Delta_2 x(t_2-\eta_j;\varepsilon_j(\lambda),\lambda)}{\varepsilon_j(\lambda)}\to 0\ (j\to\infty)$ 的一致收敛性需从命题6.4的证明中看出。——译者注。

$$S(\lambda)=z_i^{\wedge}(\lambda)=x(t_2-\varepsilon(\lambda)\theta(\lambda);\varepsilon(\lambda),\lambda)-\tilde{x}(t_2)$$
$$=\varepsilon(\lambda)\left\{\sum_{k=0}^{n}\lambda_k\delta x^{(k)}+R(\lambda)\right\},\qquad \lambda\in\Lambda.$$

我们来证明，若取

$$\hat{V}=\left\{\delta x:|\delta x|\leqslant\frac{\eta_i^{\wedge}}{\theta^{(n)}}\cdot\frac{d}{3}\right\},$$

并且 $\omega(\lambda)$ 是满足不等式

$$|\omega(\lambda)|\leqslant\frac{\eta_i^{\wedge}}{\theta^{(n)}}\cdot\frac{d}{3}$$

的任一连续函数，则对任一 $\delta x\in\hat{V}$，λ 的方程

$$S(\lambda)+\omega(\lambda)=\delta x,\qquad \lambda\in\Lambda, \tag{7.12}$$

必有解．先将此方程改写成如下形式：

$$\sum_{k=0}^{n}\lambda_k\delta x^{(k)}=\frac{\delta x-\omega(\lambda)}{\varepsilon(\lambda)}-R(\lambda),\qquad \lambda\in\Lambda.$$

上式左侧可看成是将 n 维单纯形 Λ 映到 n 维单纯形 X 的双射[注1] L: $\Lambda\to X$，即

$$\lambda=(\lambda_0,\cdots,\lambda_n)\longmapsto\sum_{k=0}^{n}\lambda_k\delta x^{(k)}.$$

因此，该方程可以写成

$$L(\lambda)=\frac{\delta x-\omega(\lambda)}{\varepsilon(\lambda)}-R(\lambda),\qquad \lambda\in\Lambda.$$

由于（参见(7.11)）

$$\left|\frac{\delta x-\omega(\lambda)}{\varepsilon(\lambda)}-R(\lambda)\right|\leqslant\frac{1}{\varepsilon(\lambda)}\frac{\eta_i^{\wedge}}{\theta^{(n)}}\frac{2d}{3}+\frac{d}{3}\leqslant d,$$

所以，对一切 $\lambda\in\Lambda$，

$$\frac{\delta x-\omega(\lambda)}{\varepsilon(\lambda)}-R(\lambda)\in \text{单纯形 } X.$$

引入 L 的逆映射

$$L^{-1}:X\to\Lambda.$$

[注1] 即 $L:\Lambda\to X$ 是一对一的且到 X 上．——译者注．

则(7.12)最终可写成

$$\lambda = L^{-1}\left(\frac{\delta x - \omega(\lambda)}{\varepsilon(\lambda)} - R(\lambda)\right), \qquad \lambda \in \Lambda.$$

连续映射

$$\lambda \longmapsto L^{-1}\left(\frac{\delta x - \omega(\lambda)}{\varepsilon(\lambda)} - R(\lambda)\right)$$

映 Λ 到自身，按熟知的 Brouwer 定理，它必存在不动点。

这样，方程(7.12)至少有一个解。并且，它的每一个解 $\hat{\lambda}$ 使得下式成立。

$$x(t_2 - \varepsilon(\hat{\lambda})\theta(\hat{\lambda})\,;\,\varepsilon(\hat{\lambda}),\hat{\lambda}) = \tilde{x}(t_2) + \delta x.$$

然而，上式左侧轨线所对应的广义控制

$$\tilde{\mu}_t + \varepsilon(\hat{\lambda})\sum_{k=0}^{n}\hat{\lambda}_k\delta\mu_t^{(k)}$$

一般未必是常义的允许控制，更不用说是逐段常值控制。因此，上面的等式还未证得命题 7.2. 为了最后完成证明，还得借助于逼近引理（定理 3.2）。根据该引理，广义控制族

$$\tilde{\mu}_t + \delta\mu_t(\varepsilon,\lambda), \qquad (\varepsilon,\lambda) \in [0,\varepsilon_1] \times \Lambda$$

可以用逐段常值允许控制在弱收敛的意义下来任意逼近。按我们的需要，这可改述为：存在一列逐段常值允许控制

$$u^{(i)}(t;\varepsilon,\lambda) \in \Omega_U, \qquad i = 1,2,\cdots,$$

它强连续地依赖于参数

$$\sigma = (\varepsilon,\lambda) \in [0,\varepsilon_1] \times \Lambda.$$

并且，

(i) $\delta_{u^{(i)}(t;\varepsilon,\lambda)}$的测度都集中在 U 的同一个有界子集之中；

(ii) 摄动序列

$$\delta\nu_t^{(i)}(\varepsilon,\lambda) = \delta_{u^{(i)}(t;\varepsilon,\lambda)} - (\tilde{\mu}_t + \delta\mu_t(\varepsilon,\lambda)) \in \delta\mathfrak{M}_{\tilde{\mu}_t + \delta\mu_t(\varepsilon,\lambda)}$$

在 $i \to +\infty$ 时，关于 $(\varepsilon,\lambda) \in [0,\varepsilon_1] \times \Lambda$ 一致地弱收敛于零。这样，摄动序列

$$\delta\mu_t(\varepsilon,\lambda) + \sigma\nu_t^{(i)}(\varepsilon,\lambda) \in \delta\mathfrak{M}_{\tilde{\mu}_t},$$

在 $i \to +\infty$ 时，关于$(\varepsilon,\lambda) \in [0,\varepsilon_1] \times \Lambda$ 一致地弱收敛于 $\delta\mu_t(\varepsilon,\lambda)$.

根据定理6.2，存在充分小的 $\varepsilon_1>0$ 及充分大的 i_1，对一切

$$(\varepsilon,\lambda)\in[0,\varepsilon_1]\times\Lambda,\qquad i\geqslant i_1,$$

方程

$$\dot{x}=\langle\delta_{u^{(i)}(t;\varepsilon,\lambda)},f(t,x,u)\rangle=F(t,x)+\langle\delta\mu_t(\varepsilon,\lambda),f(t,x,u)\rangle$$
$$+\langle\delta v_t^{(i)}(\varepsilon,\lambda),f(t,x,u)\rangle,\qquad i\geqslant i_1,$$

的解

$$y^{(i)}(t;\varepsilon,\lambda)=y(t;\delta_{u^{(i)}(t;\varepsilon,\lambda)}),\qquad t\in[t_1,t_2],$$
$$y^{(i)}(t_1;\varepsilon,\lambda)=x_1,$$

在 $[t_1,t_2]$ 上存在且连续依赖于 $(t,\varepsilon,\lambda)\in[t_1,t_2]\times[0,\varepsilon_1]\times\Lambda$，并关于 $(\varepsilon,\lambda)\in[0,\varepsilon_1]\times\Lambda$ 一致地成立极限

$$\|y^{(i)}(t;\varepsilon,\lambda)-x(t;\varepsilon,\lambda)\|\longrightarrow 0\qquad(i\to\infty).$$

这里，$x(t;\varepsilon,\lambda)$ 是解族(7.9)．

记

$$y^{(i)}(t_2-\varepsilon(\lambda)\theta(\lambda);\varepsilon(\lambda),\lambda)-x(t_2-\varepsilon(\lambda)\theta(\lambda);\varepsilon(\lambda),\lambda)=w^{(i)}(\lambda).$$

由前面所证的事实可知，在 i 充分大时，

$$|w^{(i)}(\lambda)|\leqslant\frac{\eta_{\hat{j}}}{\theta^{(n)}}\ \frac{d}{3}.$$

这时，关于 λ 的方程

$$y^{(i)}(t_2-\varepsilon(\lambda)\theta(\lambda);\varepsilon(\lambda),\lambda)-\bar{x}(t_2)=S(\lambda)$$
$$+w^{(i)}(\lambda)=\delta x,\qquad\lambda\in\Lambda,$$

对任一 $\delta x\in\hat{V}$ 有解。

但轨线 $y^{(i)}(t;\varepsilon,\lambda)$ 所对应的是逐段常值允许控制 $u^{(i)}(t;\varepsilon,\lambda)$．于是，命题7.2获证。

7.4 最大值原理的证明

有了变分锥与命题7.2，不难通过有限维凸集的边界点存在支承

超平面的定理来证明最大值原理.

设 $\tilde{\mu}_t$ 和 $\tilde{x}(t)(t\in[t_1,t_2])$为弱最优。根据命题 7.2,变分锥(7.6)的顶点(即原点)必是它的边界点。因此，存在过原点的以单位向量 ξ 为外法向的超平面，使下面不等式成立：

$$\xi\cdot(\delta x(t_2;\delta\mu_\theta)-\theta F)\leqslant 0,\quad \forall\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},\quad \forall\theta\geqslant 0.$$

但 $\delta\mu_t$ 与 θ 互相独立，故可得

$$\begin{aligned}&\xi\cdot\delta x(t_2;\delta\mu_\theta)=\xi\Gamma(t_2)\int_{t_1}^{t_2}G(t)\langle\delta\mu_t,f(t,\tilde{x}(t),u)\rangle dt\leqslant 0,\\&\qquad\forall\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},\\&\qquad\xi\cdot F\geqslant 0.\end{aligned}\tag{7.13}$$

其中，矩阵 $G(t)$ 满足方程（见第五章）：

$$\dot{G}=-G(t)F_x(t,\tilde{x}(t)).$$

记 n 维行向量

$$\tilde{\psi}(t)=\xi\cdot\Gamma(t_2)G(t),t\in[t_1,t_2],$$

它是方程

$$\dot{\tilde{\psi}}(t)=-\tilde{\psi}(t)F_x(t,\tilde{x}(t))=-\tilde{\psi}(t)\langle\tilde{\mu}_t,f_x(t,\tilde{x}(t),u)\rangle$$

在条件

$$\tilde{\psi}(t_2)=\xi\Gamma(t_2)G(t_2)=\xi\neq 0$$

下的非零解。

引进 $\tilde{\psi}(t)$后，(7.13)式可改写成如下形式：

$$\int_{t_1}^{t_2}\langle\delta\mu_t,\tilde{\psi}(t)f(t,\tilde{x}(t),u)\rangle dt\leqslant 0,\quad\forall\delta\mu_t\in\delta\mathfrak{M}_{\tilde{\mu}_t}.$$

这一不等式实际上就是积分最大值原理。事实上，代入

$$\delta\mu_t=\mu_t-\tilde{\mu}_t\in\delta\mathfrak{M}_{\tilde{\mu}_t},$$

就得

$$\int_{t_1}^{t_2}\langle\tilde{\mu}_t,H(t,\tilde{x}(t),\tilde{\psi}(t),u)\rangle dt\geqslant\int_{t_1}^{t_2}\langle\mu_t,H(t,\tilde{x}(t),\tilde{\psi}(t),u)\rangle dt,$$

$$\forall\mu_t\in\mathfrak{M}_U.$$

这样，函数 $\tilde{\psi}(t)\neq 0(t\in[t_1,t_2])$连同 $\tilde{\mu}_t$ 与 $\tilde{x}(t)$一起，给出了凸最优

问题(7.4)的极值解.

$$\dot{\tilde{x}} = \langle\tilde{\mu}_t, f(t,\tilde{x}(t),u)\rangle = \frac{\partial}{\partial\psi}\langle\tilde{\mu}_t, H(t,\tilde{x}(t),\tilde{\psi}(t),u)\rangle,$$

$$\begin{aligned}\dot{\tilde{\psi}} &= -\tilde{\psi}(t)\langle\tilde{\mu}_t, f(t,\tilde{x}(t),u)\rangle \\ &= -\frac{\partial}{\partial x}\langle\tilde{\mu}_t, H(t,\tilde{x}(t),\tilde{\psi}(t),u)\rangle,\end{aligned} \tag{7.14}$$

$$\begin{aligned}&\int_{t_1}^{t_2}\langle\tilde{\mu}_t, H(t,\tilde{x}(t),\tilde{\psi}(t),u)\rangle dt \\ &\geqslant \int_{t}^{t_2}\langle\mu_t, H(t,\tilde{x}(t),\tilde{\psi}(t),u)\rangle dt, \forall\mu_t\in\mathfrak{M}_U.\end{aligned}$$

还须证明，函数

$$\sup_{u\in U} H(t,\tilde{x}(t),\tilde{\psi}(t),u) = M(t,\tilde{x}(t),\tilde{\psi}(t)), \quad t\in[t_1,t_2],$$

是整个区间 $t\in[t_1,t_2]$ 上的连续函数，以及

$$M(t_2,\tilde{x}(t_2),\tilde{\psi}(t_2))\geqslant 0.$$

连续性可作为最大值条件的推论得到。事实上，从(7.2)推出(7.3)的过程中，我们已经证明了，如果 $\tilde{\mu}_t$ 满足积分形式的最大值条件，则等式

$$\langle\tilde{\mu}_\theta, H(\theta,\tilde{x}(\theta),\tilde{\psi}(\theta),u)\rangle = M(\theta,\tilde{x}(\theta),\tilde{\psi}(\theta)) \tag{7.15}$$

在函数

$$F(t,\tilde{x}(t)) = \langle\tilde{\mu}_t, f(t,\tilde{x}(t),u)\rangle$$

的任一勒贝格点 θ 处成立。按假设，$\tilde{\mu}_t(t\in\boldsymbol{R})$ 的测度都集中在同一个有界集

$$N\subset U\subset\boldsymbol{R}^r$$

之中。因此，

$$\begin{aligned}\tilde{\psi}(\theta)\langle\tilde{\mu}_\theta, f(\theta,\tilde{x}(\theta),u)\rangle &= \langle\tilde{\mu}_\theta, H(\theta,\tilde{x}(\theta),\tilde{\psi}(\theta),u)\rangle \\ &\leqslant \sup_{u\in\bar{N}} H(\theta,\tilde{x}(\theta),\tilde{\psi}(\theta),u) = \max_{u\in\bar{N}} H(\theta,\tilde{x}(\theta),\tilde{\psi}(\theta),u),\end{aligned}$$

这里，闭包 $\bar{N}$ 是紧集。但是，

$$\max_{u\in\bar N} H(\theta,\tilde x(\theta),\tilde\psi(\theta),u)\leqslant\sup_{u\in U} H(\theta,\tilde x(\theta),\tilde\psi(\theta),u)$$
$$=M(\theta,\tilde x(\theta),\tilde\psi(\theta)).$$

再由(7.15)知，在函数 $F(t,\tilde x(t))$ 的每一勒贝格点 θ 成立等式

$$M(\theta,\tilde x(\theta),\tilde\psi(\theta))=\max_{u\in\bar N} H(\theta,\tilde x(\theta),\tilde\psi(\theta),u). \qquad (7.16)$$

函数

$$M_{\bar N}(t,x,\psi)=\max_{u\in\bar N} H(t,x,\psi,u)$$

连续依赖于 (t,x,ψ)[注1]。于是，(7.16)可以写成

$$M(\theta,\tilde x(\theta),\tilde\psi(\theta))=M_{\bar N}(\theta,\tilde x(\theta),\tilde\psi(\theta)),$$

上式事实上在区间 $[t_1,t_2]$ 中点点成立。因为，倘若有某一点 $\hat t$，使得

$$M_{\bar N}(\hat t,\tilde x(\hat t),\tilde\psi(\hat t))<M(\hat t,\tilde x(\hat t),\tilde\psi(\hat t))$$
$$=\sup_{u\in U} H(\hat t,\tilde x(\hat t),\tilde\psi(\hat t),u),$$

则必有一点 $\hat u\in U$，使得

$$M_{\bar N}(\hat t,\tilde x(\hat t),\tilde\psi(\hat t))<H(\hat t,\tilde x(\hat t),\tilde\psi(\hat t),\hat u).$$

但因 $M_{\bar N}(t,\tilde x(t),\tilde\psi(t))$ 与 $H(t,\tilde x(t),\tilde\psi(t),\hat u)$（$\hat u$ 固定）都连续依赖于 t，这样，我们就得到一个对所有接近于 $\hat t$ 的勒贝格点 θ 都成立的与

[注1] 我们证明函数 $M_{\bar N}(t,x,\psi)$ 在任一点 (t,x,ψ) 是上半连续的。假设不然，则由 $H(t,x,\psi,u)$ 的连续性与 $\bar N$ 的紧性，知存在收敛于零的序列 $\delta t_i,\delta x_i$ 和 $\delta\psi_i$，收敛点列 $u_i\in\bar N, u_i\to\hat u, (i\to\infty)$，以及一个正数 $\varepsilon>0$，使得

$$M_{\bar N}(t+\delta t_i,x+\delta x_i,\psi+\delta\psi_i)=H(t+\delta t_i,x+\delta x_i,\psi+\delta\psi_i,u_i)$$
$$\geqslant M(t,x,\psi)+\varepsilon\geqslant H(t,x,\psi,\hat u)+\varepsilon.$$

在上式中，令 $i\to\infty$，便得到矛盾：

$$H(t,x,\psi,\hat u)\geqslant H(t,x,\psi,\hat u)+\varepsilon.$$

下半连续性可类似地证明（不需要 N 紧的条件）。

等式(7.16)相矛盾的不等式。这说明(7.16)对一切 $t\in[t_1,t_2]$ 都成立，从而证明了 $M(t,\tilde{x}(t),\tilde{\psi}(t))$ 在区间 $t\in[t_1,t_2]$ 上的连续性。

在推导关系式(7.14)以及证明 $M(t,\tilde{x}(t),\tilde{\psi}(t))$ 的连续性的过程中，没有用过(7.13)的第二个不等式。这相当于仅用到变分锥(7.6)中 $\theta=0$ 的那部分。

为了证明本定理的最后一个关系式，即

$$M(t_2,\tilde{x}(t_2),\tilde{\psi}(t_2))\geqslant 0,$$

我们必须用到整个变分锥 Q 和(7.13)的第二个不等式，

$$\xi\cdot F\geqslant 0.$$

由于 $\tilde{\psi}(t_2)=\xi$，上式可写成

$$\tilde{\psi}(t_2)\cdot F\geqslant 0.$$

设

$$F=\frac{\tilde{x}(t_2)-\tilde{x}(t_2-\eta_j)}{\eta_j}+\nu_j,\quad \eta_j\to 0,\quad \nu_j\to 0,\quad (j\to\infty).$$

它可以写成

$$\begin{aligned}F&=\frac{1}{\eta_j}\int_{t_2-\eta_j}^{t_2}\frac{d\tilde{x}(t)}{dt}\,dt+\nu_j\\&=\frac{1}{\eta_j}\int_{t_2-\eta_j}^{t_2}\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle dt+\nu_j.\end{aligned}$$

因此，

$$\begin{aligned}\tilde{\psi}(t_2)\cdot F=&\frac{1}{\eta_j}\int_{t_2-\eta_j}^{t_2}\tilde{\psi}(t)\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle dt\\&+\frac{1}{\eta_j}\int_{t_2-\eta_j}^{t_2}(\tilde{\psi}(t_2)-\tilde{\psi}(t))\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle dt\\&+\tilde{\psi}(t_2)\cdot\nu_j.\end{aligned}$$

利用最大值条件(7.2)和(7.3)的等价性，并将几乎处处等于被积函数 $\tilde{\psi}(t)\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle$ 的函数 $M(t,\tilde{x}(t),\tilde{\psi}(t))$ 代入，得

$$\left|\tilde{\psi}(t_2)\cdot F-\frac{1}{\eta_j}\int_{t_2-\eta_j}^{t_2}M(t,\tilde{x}(t),\tilde{\psi}(t))dt\right|$$

$$\leqslant \max_{t_2-\eta_j\leqslant t\leqslant t_2}|\tilde{\psi}(t_2)-\tilde{\psi}(t)|\cdot\int_{t_2-\eta_j}^{t_2}\frac{1}{\eta_j}|\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle|dt$$

$$+|\tilde{\psi}(t_2)||\nu_j|.$$

由于 $\tilde{\mu}_t$ 的测度集中在同一有界集上，且函数 $f(t,x,u)$连续，便有

$$\max_{t_2-\eta_j\leqslant t\leqslant t_2}|\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle|\leqslant \text{const},$$

遂得估计

$$\left|\tilde{\psi}(t_2)\cdot F-\frac{1}{\eta_j}\int_{t_2-\eta_j}^{t_2}M(t,\tilde{x}(t),\tilde{\psi}(t))dt\right|\leqslant \max_{t_2-\eta_j\leqslant t\leqslant t_2}|\tilde{\psi}(t_2)$$

$$-\tilde{\psi}(t)|\text{const}+|\tilde{\psi}(t_2)||\nu_j|.$$

当 $j\to\infty$ 时，由于 $M(t,\tilde{x}(t),\tilde{\psi}(t))$的连续性，就得

$$M(t_2,\tilde{x}(t_2),\tilde{\psi}(t_2))=\tilde{\psi}(t_2)\cdot F\geqslant 0.$$

最后，从证明中可见如下事实：函数 $\bar{\psi}(t),t\in[t_1,t_2]$，可以取为微分方程

$$\dot{\psi}=-\psi F_x(t,\bar{x}(t))$$

的满足如下边值条件的任一非零解，只要向量 $\bar{\psi}(t_2)$与过原点的 Q 的一个支承超平面正交且指向 Q 的外侧。

须注意，对于 $\bar{\psi}(t)$的边值条件是在终止时刻 t_2 给出的，而不是在初始时刻 $t=t_1$ 给出的。

第八章　最优解的存在性

广义控制及相应的凸控制问题之优点已在第六章的轨线变分及第七章的最大值原理的证明中得到了体现。

广义控制另一个重要的优点将在最优解存在性的研究中显示出来。在很广泛又很自然的假定下，将证明广义控制集合是弱紧的（确切定义在后面给出）。这种弱紧性保证了相当大一类凸最优问题解的存在性，但常义最优问题(1.3)只有在对 $f(t,x,u)$ 加上相当的限制后才有解（见8.3节）。

首先，我们要证明命题 8.1 及其推论。该命题是判别 Radon 测度空间中的集合弱紧性的一般命题。特别地，从该命题可得出在$U\subset \boldsymbol{R}^r$是紧集时，广义控制类 $\mathfrak{M}_U$ 是弱紧集这一推论（定理 8.1）。

从定理 8.1 易导出凸最优问题解的存在性（定理 8.2）。用定理 8.2 来证明常义控制类中的最优解的存在性，即定理 8.3（又称 A.F. Fillippov 定理）。

作为定理 8.3 的应用，我们研究了滑动最优制式(sliding optimal regimes)，并将证明古典变分法中的 Hilbert-Tonelli 存在性定理（定理 8.4）。

8.1　广义控制类的弱紧性

在本章中，$C^0(\boldsymbol{R}^m)$表示所有定义在 $\boldsymbol{R}^m$ 上并有紧支集的连续函数构成的线性空间，$\|\cdot\|_M$表示$C^0(\boldsymbol{R}^m)$上由子集 $M\subset\boldsymbol{R}^m$按下式确定的半范：

$$\|g(z)\|_M=\sup_{z\in M}|g(z)|,\ g(z)\in C^0(\boldsymbol{R}^m),$$

而 $C^*(\boldsymbol{R}^m)$表示 $\boldsymbol{R}^m$ 上所有实的（未必正）Radon 测度的集合。

与第二章给出的定义相一致，一列测度 $\nu^{(i)}\in C^*(\boldsymbol{R}^m)\,(i=1,2,\cdots)$ 称为弱收敛于测度 $\nu\in C^*(\boldsymbol{R}^m)$，如果对任一函数 $g(z)\in C^0(\boldsymbol{R}^m)$，

$$\langle \nu^{(i)},g(z)\rangle=\int_{R^m} g(z)d\nu^{(i)}(z)\longrightarrow\langle \nu,g(z)\rangle=\int_{R^m} g(z)d\nu(z)\,(i\to\infty).$$

我们用 $C^*(M)$ 表示空间 $C^*(\boldsymbol{R}^m)$ 中测度集中在集合 $M\subset\boldsymbol{R}^m$ 上的 Radon 测度全体所组成的子空间。

易见，如果集合M是闭集，那么子空间 $C^*(M)$ 关于弱收敛是弱列闭的。也就是说，如果序列 $\nu^{(i)}\in C^*(M)$，$i=1,2\cdots$，弱收敛于 ν，则 $\nu\in C^*(M)$.

这是因为，补集 $\boldsymbol{R}^m\backslash M$ 是开集，因此，对任一支集含于 $\boldsymbol{R}^m\backslash M$中的函数 $g(z)\in C^0(\boldsymbol{R}^m)$，等式$\langle\nu,g(z)\rangle=0$ 成立。事实上，我们有 $\langle\nu^{(i)},g(z)\rangle=0\,(i=1,2,\cdots)$，故

$$\langle\nu,g(z)\rangle=\langle\nu,g(z)\rangle-\lim_{i\to\infty}\langle\nu^{(i)},g(z)\rangle=\lim_{i\to\infty}\langle\nu-\nu^{(i)},g(z)\rangle=0.$$

如果M不是闭集，那么$C^*(M)$未必弱列闭。这一点很容易从下面的简单例子中看出：取一点列$\{z_i\}$收敛于 z，且 $z\neq z_i$，$\forall i$. 那么，Dirac 测度序列 δ_{z_i}，弱收敛于测度 δ_z. 取$M=\{z_1,z_2,\cdots\}$，则 $\delta_z\notin C^*(M)$，即 $C^*(M)$不是弱列闭。

显然有不等式：

$$|\langle\nu,g(z)\rangle|\leqslant\|\nu\|\,\|g(z)\|_M,\ \forall\nu\in C^*(M),$$

这里，$\|\nu\|$表示测度 ν 的范数（即全变差）.

设集合 $\mathfrak{N}\subset C^*(\boldsymbol{R}^m)$，如果从 $\mathfrak{N}$中的任一序列$\nu^{(i)}\,(i=1,2,\cdots)$总可选出一个子列弱收敛于 $\mathfrak{N}$中的某一测度，则称 $\mathfrak{N}$为弱列紧.

命题8.1 对任一闭集 $M\subset\boldsymbol{R}^m$ 及任一非负常数 C，对应的 Radon 测度集合

$$\mathfrak{N}=\mathfrak{N}(M,C)=\{\nu:\ \nu\in C^*(M),\|\nu\|\leqslant C\}\subset C^*(\boldsymbol{R}^m)$$

是弱列紧集。

证 设$\{g_1(z),g_2(z),\cdots\}$是 $C^0(\boldsymbol{R}^m)$中关于半范 $\|\cdot\|_M$ 的稠密集，即对任一正数 $\varepsilon>0$及任一$g(z)\in C^0(\boldsymbol{R}^m)$，存在 $g_i(z)$使得

$$\|g(z)-g_i(z)\|_M\leqslant\varepsilon.$$

从命题 3.3 与不等式

$$\|g(z)\|_M \leqslant \|g(z)\| = \sup_{z\in R^m} |g(z)|$$

可知这样的序列$\{g_i(z)\}$是存在的。此结果对任一集M成立。

设$\nu^{(i)}\in\mathfrak{N}, i=1,2,\cdots$，由于

$$|\langle \nu^{(i)}, g_1(z)\rangle| \leqslant \|\nu^{(i)}\| \|g_1(z)\|_M \leqslant C\|g_1(z)\|_M, i=1,2,\cdots,$$

是有界的，故存在子序列$\nu^{(i_k)}=\nu_1{}^{(k)}$使得

$$\lim_{i\to\infty}\langle \nu_1^{(k)}, g_1(z)\rangle = \lambda_1.$$

类似地，从序列$\{\nu_1^{(i)}\}$可选出子列

$$\nu_2^{(k)} = \nu_1^{(i_k)}, \ k=1,2,\cdots,$$

使得下面的极限存在

$$\lim_{i\to\infty}\langle \nu_2^{(i)}, g_2(z)\rangle = \lambda_2.$$

逐步进行下去可得

$$\{\nu_1^{(i)}\}_1^\infty \supset \{\nu_2^{(i)}\}_1^\infty \supset \cdots \supset \{\nu_k^{(i)}\} \supset \cdots,$$

其中对$g_j(z)(j=1,2,\cdots,k)$，满足

$$\lim_{i\to\infty}\langle \nu_k^{(i)}, g_j(z)\rangle = \lambda_j, \forall j=1,\cdots,k;$$

$$k=1,2,\cdots.$$

取对角线序列

$$\hat{\nu}^{(i)} = \nu_i^{(i)}, \ i=1,2,\cdots,$$

则对一切$g_j(z)$成立：

$$\lim_{i\to\infty}\langle \hat{\nu}^{(i)}, g_j(z)\rangle = \lambda_j, \forall j=1,2,\cdots.$$

我们进一步证明，对任一函数$g(z)\in C^0(R^m)$，数列$\{\langle\hat{\nu}^{(i)}, g(z)\rangle\}_1^\infty$都收敛。由于函数列$g_j(z)(j=1,2,\cdots)$在$C^0(R^m)$中按半范$\|\cdot\|_M$稠密，故对任一$\varepsilon>0$，存在$j=j_\varepsilon$，使得

$$\|g(z)-g_{j_\varepsilon}(z)\|_M \leqslant \varepsilon/4C.$$

但当$i\to\infty$时数列$\langle\hat{\nu}^{(i)}, g_j(z)\rangle$收敛，因此对充分大的$i', i''$，有

$$|\langle \hat{\nu}^{(i')} - \hat{\nu}^{(i'')}, g_j(z)\rangle| \leqslant \varepsilon/2.$$

又因

$$\|\hat{\nu}^{(i')}-\hat{\nu}^{(i'')}\|\leqslant\|\hat{\nu}^{(i')}\|+\|\hat{\nu}^{(i'')}\|\leqslant 2C,$$

于是，在 i',i'' 充分大时，下式成立，

$$\begin{aligned}|\langle\hat{\nu}^{(i')},g(z)\rangle-\langle\hat{\nu}^{(i'')},g(z)\rangle|&\leqslant|\langle\hat{\nu}^{(i')}-\hat{\nu}^{(i'')},g(z)-g_{j_0}(z)\rangle|\\&\quad+|\langle\hat{\nu}^{(i')}-\hat{\nu}^{(i'')},g_{j_0}(z)\rangle|\\&\leqslant\|\hat{\nu}^{(i')}-\hat{\nu}^{(i'')}\|\frac{\varepsilon}{4C}+\frac{\varepsilon}{2}\leqslant\varepsilon.\end{aligned}$$

这就证明了，对每一 $g(z)\in C^0(\boldsymbol{R}^m)$，数列 $\langle\hat{\nu}^{(i)},g(z)\rangle(i=1,2,\cdots)$ 都收敛．定义 $C^0(\boldsymbol{R}^m)$ 上的泛函 $\nu(g(z))$ 如下，

$$g(z)\longmapsto\nu(g(z))=\lim_{i\to\infty}\langle\hat{\nu}^{(i)},g(z)\rangle.$$

显然，对任意的 $g_1(z),g_2(z)\in C^0(\boldsymbol{R}^m)$，以及任意的实数 α 和 β，有

$$\begin{aligned}\nu(\alpha g_1(z)+\beta g_2(z))&=\lim_{i\to\infty}\langle\hat{\nu}^{(i)},\alpha g_1(z)+\beta g_2(z)\rangle\\&=\alpha\lim_{i\to\infty}\langle\hat{\nu}^{(i)},g_1(z)\rangle+\beta\lim_{i\to\infty}\langle\hat{\nu}^{(i)},g_2(z)\rangle\\&=\alpha\nu(g_1(z))+\beta\nu(g_2(z)),\end{aligned}$$

并且，对一切 $g(z)\in C^0(\boldsymbol{R}^m)$，

$$|\nu(g(z))|=\lim_{i\to\infty}|\langle\hat{\nu}^{(i)},g(z)\rangle|\leqslant\overline{\lim_{i\to\infty}}\|\hat{\nu}^{(i)}\|\|g(z)\|_M\leqslant C\|g(z)\|_M.$$

因此，$\nu(\cdot)$ 是 $C^0(\boldsymbol{R}^m)$ 上的线性泛函，并且它在 $C^0(\boldsymbol{R}^m)$ 中的支集含于一个给定紧集 $K\subset\boldsymbol{R}^m$ 内的函数组成的子空间上的限制是按 $\|\cdot\|_K$ 为连续的泛函。这就是说，ν 是一个 Radon 测度[注1]。

于是可记

$$\nu(g(z))=\langle\nu,g(z)\rangle.$$

按 ν 的定义，有

$$\lim_{i\to\infty}\langle\nu-\hat{\nu}^{(i)},g(z)\rangle=0,\qquad\forall g(z)\in C^0(\boldsymbol{R}^m),$$

即序列 $\nu_i\in\mathfrak{M}(M,C)$ 的子列 $\hat{\nu}^{(i)}$ 弱收敛于 Radon 测度 ν.

因为 M 是闭集，所以 $C^*(M)$ 是弱闭的。这样从 $\hat{\nu}^{(i)}\in C^*(M)$ 可知

$$\nu\in C^*(M),$$

[注1] 可参见中译本参考文献[9]．——译者注．

并由估计

$$|\nu(g(z))| = |\langle\nu, g(z)\rangle| \leqslant C\|g(z)\|_M, \quad \forall g(z)\in C^0(\boldsymbol{R}^m),$$

得

$$\|\nu\| \leqslant C.$$

这表明，$\nu\in\mathfrak{N}(M,C)$.

这就完成了命题的证明。

注 在闭集M非紧时，即使测度 $\hat{\nu}^{(i)}(i=1,2,\cdots)$都是概率测度，但极限测度却可能是零。

例如，对$M=\boldsymbol{R}^m$，有一点列$z_i\in\boldsymbol{R}^m$趋于无穷，则 Dirac 测度序列$\delta_{z_i}(i=1,2,\cdots)$将弱收敛于零测度。这是因为当$z_i$位于函数$g(z)$的支集之外时，有

$$\langle\delta_{z_i}, g(z)\rangle = 0.$$

但当M是紧集时，则这种“测度消失于无穷远”的现象不可能发生。这时不难证明，如果一列 $\hat{\nu}^{(i)}$都是概率测度时，极限测度ν也将是概率测度。

现在我们来讨论广义控制 $\mu_t\in\mathfrak{M}_U$. 如第二章中所述，广义控制可看成按下式定义的$\boldsymbol{R}^{1+r}$上的 Radon 测度，

$$\langle\nu, g(t,u)\rangle = \int_R\langle\mu_t, g(t,u)\rangle dt = \int_R dt\int_{R^r} g(t,u)d\mu_t(u).$$

我们将证明下述基本定理。

定理 8.1 如果控制取值集$U\subset\boldsymbol{R}^r$是紧集，则广义控制全体是$\mathfrak{M}_U$弱列紧集，即对任一广义控制序列$\mu_t^{(i)}\in\mathfrak{M}_U\ (i=1,2,\cdots)$必可选取一个子列弱收敛于广义控制$\mu_t\in\mathfrak{M}_U$.[注1]

证 设$\{\mu_t^{(i)}\}\ (i=1,2,\cdots)$是任一广义控制列。现定义带参数$\tau\in\boldsymbol{R}$的一列$\boldsymbol{R}^{1+r}$上的 Radon 测度 $\nu_\tau^{(i)}(i=1,2,\cdots)$为

$$[\nu_\tau^{(i)}, g(t,u)] \equiv \int_{R^{1+r}} g(t,u)d\nu_\tau^{(i)}(t,u)$$

$$= \int_0^\tau\langle\mu_t^{(i)}, g(t,u)\rangle dt = \int_0^\tau dt\int_U g(t,u)d\mu_t^{(i)}(u), \tag{8.1}$$

[注1] 由此可见，当U是紧集时，控制类$\Omega_U=\{u(\cdot)|u(t)\in U$ 且于$[t_1,t_2]$可测$\}$是相对弱紧集。

其中 $g\in C^0(\boldsymbol{R}^{1+r})$。这里，方括号 $[\cdot,\cdot]$ 用来表示 $C^0(\boldsymbol{R}^{1+r})$ 中的函数 $g(t,u)$ 关于测度 $\nu_\tau^{(i)}$ 的积分。

由于对一切支集含在 $\boldsymbol{R}^{1+r}\setminus([0,\tau]\times U)$ 之中的函数 $g(t,u)\in C^0(\boldsymbol{R}^{1+r})$，有

$$[\nu_\tau^{(i)},g(t,u)]=\int_0^\tau dt\int_U g(t,u)d\mu_t^{(i)}=0,$$

因此，$\nu_\tau^{(i)}$ 的测度集中在紧集 $[0,\tau]\times U\subset\boldsymbol{R}^{1+r}$ 之中。又不难看出

$$\|\nu_\tau{}^{(i)}\|=\left|\int_0^\tau\|\mu_t^{(i)}\|dt\right|=|\tau|.$$

由于(8.1)所示的积分形式，可将测度 $\nu_\tau^{(i)}$ 记为

$$d\nu_\tau^{(i)}=d\mu_t^{(i)}dt,\qquad t\in[0,\tau],\tau\in\boldsymbol{R}.$$

这样，根据命题 8.1（取 $M=[0,\tau]\times U, C=|\tau|$），对每个 $\tau\in\boldsymbol{R}$，测度序列 $\nu_\tau^{(i)}(i=1,2,\cdots)$ 存在弱收敛于 $\mathfrak{R}(M,C)$ 中某一元的子列。

任取 $\boldsymbol{R}$ 中的一个稠密数列 $\tau_1,\tau_2,\cdots$. 由于对每个 τ_j 序列 $\nu^{(i)}{}_{\tau_j}$ $(i=1,2,\cdots)$ 存在弱收敛的子列，因此，用证明命题 8.1 已用过的取对角线子列的办法可证明存在一整数列 $i_1<i_2<\cdots$，使得对每个 τ_j $(j=1,2,\cdots)$，有

$$\nu_{\tau_j}^{(i_k)}\xrightarrow{\text{弱}}\nu_{\tau_j}(k\to\infty),$$

其中，$\nu_{\tau_j}\in\mathfrak{R}([0,\tau_j]\times U, |\tau_j|)$.

现在来证明，测度序列 $\nu_\tau^{(i_k)}(k=1,2,\cdots)$ 不仅对 $\tau=\tau_j(j=1,2,\cdots)$ 是弱收敛的，而且对每一个 $\tau\in\boldsymbol{R}$ 都是弱收敛的：

$$\nu_\tau^{(i_k)}\xrightarrow{\text{弱}}\nu_\tau\ (k\to\infty),$$

其中，$\nu_\tau\in\mathfrak{R}([0,\tau]\times U,|\tau|)$. 事实上，对任意的 $\tau',\tau''\in\boldsymbol{R}$ 及任意的 $g(t,u)\in C^0(\boldsymbol{R}^{1+r})$，有不等式

$$|[\nu_{\tau'}^{(i_k)},g(t,u)]-[\nu_{(\tau'')}^{(i_k)},g(t,u)]|$$

$$=\left|\int_{\tau''}^{\tau'}\langle\mu_t^{(i_k)},g(t,u)\rangle dt\right|\leqslant|\tau'-\tau''|\,\|g(t,u)\|_{(\tau',\tau'')\times U}. \tag{8.2}$$

于是，对任一$\tau\in\boldsymbol{R}$，任意的$\varepsilon>0$，存在τ_j满足

$$|\tau-\tau_j|\leqslant\varepsilon.$$

当k',k''充分大时，

$$|[\nu_\tau^{(i_{k'})}-\nu_\tau^{(i_{k''})},g(t,u)]|\leqslant|[\nu_\tau^{(i_{k'})}-\nu_{\tau_j}^{(i_{k'})},g(t,u)]|$$

$$+|[\nu_{\tau_j}^{(i_{k'})}-\nu_{\tau_j}^{(i_{k''})},g(t,u)]|+|[\nu_\tau^{(i_{k''})}-\nu_{\tau_j}^{(i_{k''})},g(t,u)]|$$

$$\leqslant 2\varepsilon\|g(t,u)\|_{(\tau-\varepsilon,\tau+\varepsilon)\times U}+\varepsilon.$$

这表明，$\nu_\tau^{(i_k)}$是弱基本序列，从而存在弱极限$\nu_\tau\in\mathfrak{R}([0,\tau]\times U,|\tau|)$.

为了完成定理的证明，我们还须证明存在一广义控制$\mu_t\in\mathfrak{M}_U$使得

$$[\nu_\tau,g(t,u)]=\int_0^\tau\langle\mu_t,g(t,u)\rangle dt,\ \ \forall\tau\in\boldsymbol{R},\ \forall g\in C^0(\boldsymbol{R}^{1+r}).\quad(8.3)$$

从(8.3)很易推出

$$\mu_t^{(i_k)}\xrightarrow{\text{弱}}\mu_t(k\to\infty).$$

事实上，取$\tau>0$充分大，以至$g(t,u)\in C^0(\boldsymbol{R}^{1+r})$的支集在$t$轴上的投影包含在$[-\tau,\tau]$之中，则

$$\int_R\langle\mu_t^{(i_k)}-\mu_t,\ g(t,u)\rangle\ dt=\int_{-\tau}^{\tau}\langle\mu_t^{(i_k)}-\mu_t,g(t,u)\rangle dt$$

$$=[\nu_\tau^{(i_k)}-\nu_\tau,g(t,u)]-[\nu_{-\tau}^{(i_k)}-\nu_{-\tau},g(t,u)]\longrightarrow 0\quad(k\to\infty).$$

现来证明(8.3).在(8.2)中令$k\to\infty$，得

$$|[\nu_{\tau'}-\nu_{\tau''},g(t,u)]|\leqslant|\tau',\tau''|\cdot\|g(t,u)\|_{(\tau',\tau'')\times U}.$$

这表示依赖于$g\in C^0(\boldsymbol{R}^{1+r})$的函数

$$h_g(\tau)=[\nu_\tau,g(t,u)],\tau\in\boldsymbol{R},$$

满足 Lipschitz 条件，从而是绝对连续函数。

这样，存在全测度集$T_g\subset\boldsymbol{R}$，使函数$h_g(\tau)$在T_g上点点可微。用$\mu_\tau(g)$表示函数$h_g(\tau)$在T_g上的导数：

$$\mu_\tau(g)=\frac{d}{d\tau}[\nu_\tau,g(t,u)],\tau\in T_g.$$

显然，集合T_g与$g(t,u)$有关。现要证明，存在一个与$g\in C^0(\mathrm{R}^{1+r})$的具体选取无关的全测度集$T\subset\boldsymbol{R}$，使得函数

$$h_g(\tau)=[\nu_\tau,g(t,u)],\quad \tau\in\boldsymbol{R}$$

在集合 T 上点点可微。为此，我们选取 $C^0(\boldsymbol{R}^{1+r})$ 的一个可列稠密集 $\{g_1,g_2,\cdots\}$. 按稠密集的定义，对每一个 $g(t,u)\in C^0(\boldsymbol{R}^{1+r})$ 及任意的 $\varepsilon>0$，必存在 g_i，使得

$$\|g(t,u)-g_i(t,u)\|_{R\times U}\leqslant\varepsilon.$$

记交集为

$$T=\bigcap_{i=1}^{\infty}T_{g_i}\subset\boldsymbol{R}.$$

由于每个 T_{g_i} 是全测度集，故交集 T 也是 $\boldsymbol{R}$ 中的全测度集，并且每个函数 $h_{g_i}(\tau)$ 在 T 上是可微的.

下面来证明对每个 $g\in C^0(\boldsymbol{R}^{1+r})$，函数 $h_g(\tau)(\tau\in\boldsymbol{R})$ 在 T 上也是可微的。为此，只须证明对所有的 $\hat{\tau}\in T$，以及 $\tau',\tau''\to\hat{\tau}$，成立

$$\frac{h_g(\tau')-h_g(\hat{\tau})}{\tau'-\hat{\tau}}-\frac{h_g(\tau'')-h_g(\hat{\tau})}{\tau''-\hat{\tau}}\longrightarrow 0.\tag{8.4}$$

记

$$\Delta_{\tau_1,\tau_2}(g)=\frac{h_g(\tau_1)-h_g(\tau_2)}{\tau_1-\tau_2}=\frac{[\nu_{\tau_1}-\nu_{\tau_2},g(t,u)]}{\tau_1-\tau_2}.$$

由(8.2)的极限关系式可得

$$|\Delta_{\tau',\hat{\tau}}(g)-\Delta_{\tau'',\hat{\tau}}(g)|\leqslant|\Delta_{\tau',\hat{\tau}}(g-g_i)|+|\Delta_{\tau',\hat{\tau}}(g_i)-\Delta_{\tau'',\hat{\tau}}(g_i)|$$
$$+|\Delta_{\tau'',\hat{\tau}}(g-g_i)|$$
$$\leqslant 2\|g(t,u)-g_i(t,u)\|_{R\times U}+|\Delta_{\tau',\hat{\tau}}(g_i)-\Delta_{\tau'',\hat{\tau}}(g_i)|.$$

但函数 $h_{g_i}(\tau)$ 在 $\hat{\tau}$ 点可微，故

$$\overline{\lim_{\tau',\tau''\to\hat{\tau}}}|\Delta_{\tau',\hat{\tau}}(g)-\Delta_{\tau'',\hat{\tau}}(g)|\leqslant 2\|g(t,u)-g_i(t,u)\|_{R\times U}.$$

上式左侧与 $g(t,u)$ 无关，而右侧可取适当的 $g_i(t,u)$ 使之变得任意小。这就表明(8.4)成立。于是，对所有的 $g(t,u)\in C^0(\boldsymbol{R}^{1+r})$，导数

$$\mu_\tau(g)=\frac{d[\nu_\tau,g(t,u)]}{d\tau}$$

在全测度集 $T\subset\boldsymbol{R}$ 的每一点都存在。

显然，泛函 $g\longmapsto\mu_\tau(g)$：

$$\mu_\tau:C^0(\boldsymbol{R}^{1+r})\longrightarrow\boldsymbol{R},\quad \tau\in T,$$

是线性的：

$$\mu_\tau(\alpha g_1(t,u)+\beta g_2(t,u))=\alpha\mu_\tau(g_1(t,u))+\beta\mu_\tau(g_2(t,u)),$$

并且，

$$|\mu_\tau(g)|\leqslant\max_{u\in U}|g(\tau,u)|,\forall\tau\in T. \tag{8.5}$$

(8.5)可从(8.2)的极限关系式推出：

$$|\mu_\tau(g)|=\left|\lim_{\tau'\to\tau}\frac{[\nu_{\tau'}-\nu_\tau,g(t,u)]}{\tau'-\tau}\right|\leqslant\overline{\lim_{\tau'\to\tau}}\|g(t,u)\|_{[\tau';\tau]\times U}$$

$$=\max_{u\in U}|g(\tau,u)|.$$

因此，对每个 $\tau\in T$，μ_τ 是 $\boldsymbol{R}^{1+r}$ 上的 Radon 测度。上式还指出，它的测度集中在集合$\{\tau\}\times U\subset\boldsymbol{R}^{1+r}$上。所以 μ_τ 可视为 $\boldsymbol{R}^r$ 上的一个Radon 测度，其测度集中在紧集$U\subset\boldsymbol{R}^r$上，且对任一函数$g(u)\in C^0(\boldsymbol{R}^r)$，有

$$\langle\mu_\tau,g(u)\rangle=\frac{d}{d\tau}[\nu_\tau,h(t)g(u)],$$

其中$h(t)$是 $\boldsymbol{R}$ 上具紧支集且 $h(\tau)=1$ 的连续函数[注1]．

集 $T\subset\boldsymbol{R}$ 有全测度，且对任一连续函数 $g(t,u)\in C^0(\boldsymbol{R}^{1+r})$，作为 t 的函数

$$\mu_t(g)=\langle\mu_t,g(t,u)\rangle,t\in T,$$

是可测的。因此，适当地在 $\boldsymbol{R}\backslash T$上补充定义 μ_t，可使 μ_t 成为$\boldsymbol{R}^r$ 上对所有 $t\in\boldsymbol{R}$ 都确定的一族可测的 Radon 测度，并具有：

$$[\nu_\tau,g(t,u)]=\int_0^\tau\mu_t(g)dt=\int_0^\tau\langle\mu_t,g(t,u)\rangle dt.$$

余下只须证对所有的$t\in T$，μ_t 都是概率测度．在 $g(t,u)\in C^0(\boldsymbol{R}^{1+r})$ 是非负函数时，$\langle\mu_t^{(i)},g(t,u)\rangle\geqslant0$，从而

$$[\nu_\tau^{(i_k)},g(t,u)]=\int_0^\tau\langle\mu_t^{(i_k)},g(t,u)\rangle dt,\ k=1,2,\cdots,$$

是 $\tau\in\boldsymbol{R}$ 的不减函数。因此，作为 $k\to\infty$ 的极限，$[\nu_\tau,g(t,u)]$ 也是 $\tau\in\boldsymbol{R}$ 的不减函数。这样，当 $\tau\in T$ 时，导数

[注1] 从(8.5)知，$\frac{d}{d\tau}[\nu_\tau,h(t)g(u)]$与$h(\cdot)$的具体选取无关．——译者注．

$$\langle\mu_\tau,g(t,u)\rangle=\frac{d}{d\tau}[\nu_\tau,g(t,u)]\geqslant 0,$$

即 μ_τ 是非负测度。

至此，集 U 的紧性尚未用到。事实上，只要 U 是闭集（于是命题 8.1 适用），上面的证明都适用。但 U 仅仅是闭集时，前已述及，测度可能在无穷远处消失以至出现 $\mu_t=0$ 的现象。现在来证明，在 U 是紧集时，$\|\mu_t\|=1(t\in T)$.

任意取一点 $\hat{t}\in T$.设 $g(t,u)\in C^0(\boldsymbol{R}^{1+r}),0\leqslant g(t,u)\leqslant 1$,且在集合 $\{\hat{t}\}\times U\subset\boldsymbol{R}^{1+r}$ 的某一个领域上恒为 1．由于 U 是紧集，故具有这些性质的函数 $g(t,u)$是存在的。这样，当 τ 充分接近 $\hat{t}$ 时，对一切 i 都有

$$\langle\mu_\tau^{(i)},g(t,u)\rangle=1,$$

从而，

$$[\nu_\tau^{(i)},g(t,u)]=\int_0^\tau\langle\mu_t^{(i)},g(t,u)\rangle dt=\int_0^{\hat{t}}+\int_{\hat{t}}^\tau$$
$$=[\nu_{\hat{t}}^{(i)},g(t,u)]+(\tau-\hat{t}).$$

令 $i=i_k\to\infty$,即得

$$[\nu_\tau,g(t,u)]=[\nu_{\hat{t}},g(t,u)]+(\tau-\hat{t}).$$

上式两边对 τ 求导而得

$$\langle\mu_{\hat{t}},g(t,u)\rangle=\frac{d}{d\tau}[\nu_\tau,g(t,u)]\Big|_{\tau=\hat{t}}=1.$$

这表示，$\|\mu_{\hat{t}}\|\geqslant 1$.又从(8.5)可知 $\|\mu_{\hat{t}}\|\leqslant 1$.所以，$\|\mu_{\hat{t}}\|=1$,即 $\mu_{\hat{t}}$ 是概率测度。

定理 8.1 至此证完。

8.2 凸最优问题的存在性定理

称受控方程

$$\dot{x}=f(t,x,u),u\in U,$$

具有解的无限延展性，如果对一切广义控制 $\mu_t\in\mathfrak{M}_U$ 及初始数据 τ,x_τ,微分方程

$$\dot{x}=\langle\mu_t,f(t,x,u)\rangle$$
$$x(\tau)=x_\tau$$

的解在$(-\infty,+\infty)$上存在。

用第四章的方法，可以证明如果$f(t,x,u)$适合如下条件，则受控方程将具有解的无限延展性：

$$|f(t,x,u)|\leqslant m(t)g(|x|),\quad \forall(t,x,u)\in \boldsymbol{R}\times\boldsymbol{R}^n\times U,$$

其中，$m(t)$是局部可积函数，$g(\lambda)$是$[0,+\infty)$上满足

$$\overline{\lim_{\lambda\to+\infty}}\frac{g(\lambda)}{\lambda}<+\infty$$

的连续函数。

定理8.2 设允许值集$U\subset\boldsymbol{R}^r$是紧集，$\dot{x}=f(t,x,u)$具有解的无限延展性，并且凸控制问题

$$\begin{cases}\dot{x}=\langle\mu_t,f(t,x,u)\rangle,\ \mu_t\in\mathfrak{M}_U\\ x(t_1)=x_1,x(t_2)=x_2\end{cases}\tag{8.6}$$

至少对某一个$t_2=\hat{t}_2>t_1$与$\hat{\mu}_t\in\mathfrak{M}_U$存在解$\hat{x}(t)(t\in[t_1,\hat{t}_2])$（未必最优），则凸控制问题(8.6)必存在最优解

$$\tilde{\mu}_t,\tilde{x}(t),t\in[t_1,t_2],$$

使得$t_2-t_1\longrightarrow\min$。

证 我们记

$\mathfrak{M}_{x_2}=\{\mu_t:\mu_t\in\mathfrak{M}_U$，解$x(t;\mu_\theta),t\geqslant t_1,x(t_1;\mu_\theta)=x_1$，

在$[t_1,\hat{t}_2]$中的某一时刻取值$x_2\}$.

直观地说，$\mu_t\in\mathfrak{M}_{x_2}$所对应的曲线

$$\{(t,x(t;\mu_\theta)):t\in[t_1,\hat{t}_2]\}\subset\boldsymbol{R}^{1+n}$$

与线段

$$\{(t,x_2):t\in[t_1,\hat{t}_2]\}\subset\boldsymbol{R}^{1+n}$$

有交点。

按定理的假设，有$\hat{\mu}_t\in\mathfrak{M}_{x_2}$，故集合$\mathfrak{M}_{x_2}$非空。对每个$\mu_t\in\mathfrak{M}_{x_2}$，用$\tau(\mu_\theta)$表示满足$x(t;\mu_\theta)=x_2$的时刻$t\in[t_1,\hat{t}_2]$的下确界。由于$x(t;\mu_\theta)$是$t$的连续函数，解$x(t;\mu_\theta)$在$t=\tau(\mu_\theta)$的取值为$x_2$。

显然，如果数集

$$\{\tau(\mu_\theta):\mu_t\in\mathfrak{M}_{x_2}\}\tag{8.7}$$

存在最小值，则本定理获证。取到最小值 $\tilde{\tau}=\tau(\tilde{\mu}_\theta)$的广义控制$\tilde{\mu}_t$，以及对应的轨线

$$\tilde{x}(t)=x(t;\tilde{\mu}_\theta),\quad t\in[t_1,\tilde{\tau}],$$

就是凸控制问题(8.6)的最优解，而 $\tilde{\tau}-t_1$ 是最优转移时间。

为了证明最优控制 $\tilde{\mu}_t$ 的存在性，首先要证集合 $\mathfrak{M}_{x_2}$ 是弱闭的，这样，由于定理 8.1 已证得的 $\mathfrak{M}_U$的弱紧性，就知道 $\mathfrak{M}_{x_2}$ 关于广义控制弱收敛是列紧集。其次，还要证函数 $\tau(\mu_\theta)$ 关于这样的收敛是下半弱连续的，即如果一列广义控制 $\mu_t^{(i)}\in\mathfrak{M}_{x_2}$ 当 $i\to\infty$ 时弱收敛于广义控制 $\mu_t\in\mathfrak{M}_{x_2}$，则

$$\tau(\mu_t)\leqslant\varliminf_{i\to\infty}\tau(\mu_t^{(i)}).$$

现在先证 $\mathfrak{M}_{x_2}$的弱闭性。设

$$\mu_t^{(i)}\in\mathfrak{M}_{x_2},\ \mu_t^{(i)}\xrightarrow{\text{弱}}\mu_t(i\to\infty).$$

根据解的无限延展性，存在解

$$x(t)=x(t;\mu_\theta),\quad t\in[t_1,\ \hat{t}_2].$$

由命题 6.2 及解的连续依赖性定理 4.4，解列

$$x^{(i)}(t)=x(t;\mu_\theta^{(i)}),\quad t\in[t_1,\hat{t}_2],$$

当 $i\to\infty$时一致收敛于$x(t)$.

如果 $\mu_t\notin\mathfrak{M}_{x_2}$，则必存在 $\varepsilon_0>0$，使得

$$|x(t)-x_2|\geqslant\varepsilon_0>0,\ \forall t\in[t_1,\ \hat{t}_2].$$

但这是不可能的，因为每个$x^{(i)}(t)$在$[t_1,\hat{t}_2]$中的某一点 t 处取值x_2. 于是证明了集 $\mathfrak{M}_{x_2}$ 是弱列闭的。

再证 $\tau(\mu_\theta)$的下半弱连续性。设

$$\mu_t^{(i)}\xrightarrow{\text{弱}}\hat{\mu}_t(i\to\infty),$$

则轨线

$$x^{(i)}(t)=x(t;\mu_\theta^{(i)}),t\in[t_1,\ \hat{t}_2],$$

在区间$[t_1, \hat{t}_2]$上一致收敛于轨线 $\hat{x}(t)=x(t;\mu_\theta)$.

由于$\hat{\tau}=\hat{\tau}(\mu_t)$，$\hat{\mu}_t\in\mathfrak{M}_{x_2}$是曲线$\hat{x}(t)=x(t;\hat{\mu}_\theta)$，$t\in[t_1,\hat{t}_2]$，与线段$x(t)\equiv x_2(t\in[t_1,\hat{t}_2])$相交的那些$t\in[t_1,\hat{t}_2]$之最小值，因此，对任一$\varepsilon>0$，当$i$充分大时，$x^{(i)}(t)$与$x(t)\equiv x_2$线段的初次相交时刻不可能小于$\hat{\tau}-\varepsilon$. 这就证明了$\tau(\mu_\theta)$的下半弱连续性。

有了$\mathfrak{M}_{x_2}$的弱列紧性以及函数$\tau(\mu_\theta)$的下半弱连续性，我们可推出存在$\tilde{\mu}_t\in\mathfrak{M}_{x_2}$，使得

$$\tau(\tilde{\mu}_t)\leqslant\tau(\mu_t),\ \forall\mu_t\in\mathfrak{M}_{x_2}.$$

实际上，只要取广义控制的最小化序列$\mu_t^{(i)}\in\mathfrak{M}_{x_2}\ (i=1,2,\cdots)$使得

$$\tau(\mu_t^{(i)})\longrightarrow\inf_{\mu_t\in\mathfrak{M}_{x_2}}\tau(\mu_t)=\tilde{\tau}\geqslant t_1\quad(i\to\infty).$$

从这一序列可选出子列

$$\mu_t^{(i_1)},\cdots,\mu_t^{(i_k)},\cdots,$$

弱收敛于广义控制$\tilde{\mu}_t\in\mathfrak{M}_{x_2}$. 这是因为，$\mathfrak{M}_{x_2}$是弱列紧的。函数$\tau(\mu_\theta)$的下半弱连续性蕴涵了

$$\tau(\tilde{\mu}_\theta)\leqslant\varliminf_{k\to\infty}\tau(\mu_\theta^{(i_k)})$$

但

$$\varliminf_{k\to\infty}\tau(\mu_\theta^{(i_k)})=\lim_{k\to\infty}\tau(\mu_\theta^{(i_k)})=\inf_{\mu_t\in\mathfrak{M}_{x_2}}\tau(\mu_\theta).$$

因此，

$$\tau(\tilde{\mu}_\theta)=\inf_{\mu_t\in\mathfrak{M}_{x_2}}\tau(\mu_\theta)=\tilde{\tau}.$$

至此定理 8.2 证完。

注 解的无限延展性条件保证了映射

$$\mu_t\longmapsto x(t;\mu_\theta),\ t\in[t_1,t_2],$$

对任一区间$[t_1,t_2]$的t存在。若无此条件则定理 8.2 不一定成立。因为如果解没有无限延展性，则可能存在一列广义控制$\mu_t^{(i)}$，对应的轨线

$$x^{(i)}(t)=x(t;\mu_{\theta}^{(i)}),\quad t\in[t_1,t^{(i)}],$$
$$x^{(i)}(t_1)=x_1,\quad x^{(i)}(t^{(i)})=x_2,$$

中，$t^{(1)}\geqslant t^{(2)}\geqslant\cdots$，且 $i\to\infty$ 时 $t^{(i)}-t_1\longrightarrow\inf$，但是，极限控制 μ_t 所对应的轨线

$$x(t;\mu_{\theta}),x(t_1;\mu_{\theta})=x_1,$$

却不能延展到整个闭区间 $t_1\leqslant t\leqslant t_1+\inf$，它可能当 t 趋于 $t_1+\inf$ 时趋于无穷。

8.3 常义控制类中的存在性定理

集合 $U\subset\boldsymbol{R}^r$ 的紧性不足以保证最优问题的解在常义控制类中的存在性：

$$\dot{x}=f(t,x,u),u(t)\in\Omega_U,$$
$$x(t_1)=x_1,x(t_2)=x_2,t_2-t_1\longrightarrow\min. \tag{8.8}$$

（在此，总假定对应的控制问题

$$\dot{x}=f(t,x,u),u(t)\in\Omega_U,$$
$$x(t_1)=x_1,\quad x(t_2)=x_2. \tag{8.9}$$

至少有一个解存在，并且具有解的无限延展性。）

这是因为，即使 U 是紧集，集合 Ω_U 在广义控制类中非弱列紧，另一方面，Ω_U 的弱闭包将重合于（根据逼近定理）整个集合 $\mathfrak{M}_U$.

问题(8.8)无解的简单而典型的例子稍后给出。现对(8.8)附加一个条件：对所有 (t,x)，如下集合 $P(t,x)$ 是凸的：

$$P(t,x)=\{f(t,x,u):u\in U\}=\operatorname{conv}P(t,x),\forall(t,x)\in\boldsymbol{R}\times\boldsymbol{R}^n.$$

在这些假设下我们将证明问题(8.8)的解的存在性定理，即定理 8.3，又称为 A. F. Fillipov 定理。它的证明建立在定理 8.2 和下面要证的命题 8.2（即 Fillipov 引理）的基础上。

定理 8.3 设允许控制取值集 $U\subset\boldsymbol{R}^r$ 是紧集，最优问题(8.8)具有解的无限延展性，且方程右端 $f(t,x,u)$ 满足条件

$$P(t,x)=\{f(t,x,u):u\in U\}=\mathrm{conv}P(t,x),\forall(t,x)\subset \boldsymbol{R}^{1+n}.$$

并设控制问题(8.9)至少存在一个解$\hat{u}(t)$, $\hat{x}(t)$($t\in[t_1,\hat{t}_2]$). 那么，控制问题(8.9)存在最优解

$$\tilde{u}(t),\quad \tilde{x}(t),\quad t\in[t_1,t_2],$$

即最优问题(8.8)的解存在。

证 由定理 8.2，对应的凸最优控制问题有解

$$\tilde{\mu}_t,\tilde{x}(t),t_1\leqslant t\leqslant t_2.$$

因此，对几乎所有的$t\in[t_1,t_2]$,

$$\dot{\tilde{x}}(t)=\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle,\tag{8.10}$$

并且t_2-t_1是在广义控制类$\mathfrak{M}_U$中满足给定的边值条件的最小转移时间。

我们将构造出一个允许控制$\tilde{u}(t)\in\Omega_U$,使得广义控制$\tilde{\mu}_t$所对应的轨线$\tilde{x}(t)$($t_1\leqslant t\leqslant t_2$)也可由$\tilde{u}(t)$实现，即对几乎所有的$t\in[t_1,t_2]$，满足如下方程

$$\dot{\tilde{x}}(t)=f(t,\tilde{x}(t),\tilde{u}(t)),$$

这样，我们将得到最优问题(8.8)的一个解

$$\tilde{u}(t),\tilde{x}(t),\quad t_1\leqslant t\leqslant t_2.$$

记

$$F(t)=\dot{\tilde{x}}(t)=\langle\tilde{\mu}_t,f(t,\tilde{x}(t),u)\rangle,\quad t_1\leqslant t\leqslant t_2,$$

由条件

$$P(t,x)=\mathrm{conv}P(t,x)$$

以及命题2.1,可得

$$F(t)\in P(t,\tilde{x}(t))=\{f(t,\tilde{x}(t),u):u\in U\}.$$

因此，我们总可选取函数$\tilde{u}(t)$($t_1\leqslant t\leqslant t_2$),它取值于$U$并且满足条件

$$F(t)=\dot{\tilde{x}}(t)=f(t,\tilde{x}(t),\tilde{u}(t)).$$

主要的困难在于要求这样选出的函数须是 $t_1\leqslant t\leqslant t_2$ 上的可测函数。如果我们能够满足这一要求，则将此函数 $\tilde{u}(t)$ 在区间 $[t_1, t_2]$ 之外以任意容许的方式延拓，便得所希求的最优控制 $\tilde{u}(t)\in\Omega_U$。

记

$$f(t,u)=f(t,\tilde{x}(t),u)-F(t),\quad (t,u)\in[t_1,t_2]\times U.$$

我们把寻求可测函数 $\tilde{u}(t)(t_1\leqslant t\leqslant t_2)$ 的问题归结为如下命题。

命题 8.2 (Fillipov 引理)

设 $f(t,u)$ 是集合 $(t,u)\in I\times U$ 上定义的一个 n 维向量值函数，其中 I 是 t 轴上的区间，而 U 是 $\boldsymbol{R}^r$ 的一个紧子集。设对固定的 $u\in U$，$f(t,u)$ 关于 t 是 Lebesgue 可测的；当 $t\in I$ 固定时，$f(t,u)$ 关于 u 连续。又设对每一 $t\in I$，关于 u 的方程

$$f(t,u)=0,\qquad u\in U,\tag{8.11}$$

至少有一个解。那么，存在 Lebesgue 可测函数 $u(t)(t\in I)$，它取值于 U 且满足方程

$$f(t,u(t))=0,\forall t\in I.$$

证 对每个 $t\in I$，用 U_t 表示方程 (8.11) 的所有解的集合。按假定，集合 U_t 非空，并且对每个 $t\in I$ 都是紧集（因为 $f(t,u)$ 关于 u 连续且 U 是紧集）。

首先证明，对任一紧子集 $K\subset\boldsymbol{R}^r$，集合

$$T=\{t,U_t\cap K\neq\phi\}\subset I\tag{8.12}$$

是(勒贝格)可测集。在 U 中选取稠密点列 $u_1, u_2,\cdots$。用 V_ε 记集 $K\subset\boldsymbol{R}^r$ 的 ε 邻域，W_ε 记 $\boldsymbol{R}^n$ 中原点的 ε 邻域。

按假定，诸函数 $f(t,u_i)(t\in I,i=1,2,\cdots)$ 是可测函数。从而，

$$T_{i,\varepsilon}=\{t:t\in I,f(t,u_i)\in W_\varepsilon\}=[f(\cdot,u_i)]^{-1}(W_\varepsilon)\subset I$$

是可测集。对固定的 ε，我们得到如下的可测集

$$T_\varepsilon=\bigcup_i\{T_{i,\varepsilon}:u_i\in V_\varepsilon\}.$$

显然，若 $\varepsilon'\leqslant\varepsilon''$，则 $T_{\varepsilon'}\subset T_{\varepsilon''}$。因此，所有 $T_\varepsilon(\varepsilon>0)$ 之交集是可测集，因为它可表为一列可测集 $T_{\varepsilon_i}(\varepsilon_i\to0,i\to\infty)$ 之交集。

不难看出对任意的 $\varepsilon>0$，$T\subset T_\varepsilon$。事实上，设 $\hat{t}\in T$。对于任意一点

$\hat{u}\in U_{\hat{t}}\cap K$，必存在子列$\{u_{i_k}\}(k=1,2,\cdots)$，当$k\to\infty$时它收敛于$\hat{u}$. 于是，由$f(t,u)$关于$u$的连续性，我们有：

$$f(\hat{t},u_{i_k})\to f(\hat{t},\hat{u})=0\quad(k\to\infty)。$$

因此，对当k充分大时，有

$$u_{i_k}\in V_\varepsilon,f(\hat{t},u_{i_k})\in W_\varepsilon.$$

即$\hat{t}\in T_\varepsilon$，这就证明了包含关系

$$T\subset\bigcap_{\varepsilon>0}T_\varepsilon.$$

相反的包含关系

$$T\supset\bigcap_{\varepsilon>0}T_\varepsilon$$

也是很容易验证的。如果$\hat{t}\in\bigcap\limits_{\varepsilon>0}T_\varepsilon$，则对每一$\varepsilon>0$，存在$u_{i_\varepsilon}\in V_\varepsilon$，使得

$$f(\hat{t},u_{i_\varepsilon})\in W_\varepsilon.$$

这些点u_{i_ε}中可选取一个点列收敛于某一点$\hat{u}\in K$。又由$f(t,u)$关于u的连续性知$f(\hat{t},\hat{u})=0$，即$\hat{u}\in U_{\hat{t}}\cap K\neq\phi$，从而$\hat{t}\in T$. 这样，我们得到等式

$$T=\bigcup_{\varepsilon>0}T_\varepsilon.$$

从而证明了集合T是可测集。

作了以上这些准备之后，我们来构造所要求的函数$u(t)$ $(t\in I)$。取一正数a充分大，使得r维立方体

$$Q=\left\{u=\begin{pmatrix}u^1\\ \vdots\\ u^r\end{pmatrix}:-a\leqslant u^i\leqslant a,i=1,\cdots,r\right\}\subset\boldsymbol{R}^r$$

包含紧集U.

区间$[-a,a]$的第i级划分是将其按相同的长度$2a/2^i=a/2^{i-1}$分成2^i个子区间

$$I_i^{(j)}=\left\{\theta:-a+(j-1)\frac{a}{2^{i-1}}\leqslant\theta\leqslant-a+j\cdot\frac{a}{2^{i-1}}\right\},j=1,2,\cdots,2^i.$$

在第 i 级划分中，任意 r 个区间 $I_{j_1}^{(i)},\cdots,I_{j_r}^{(i)}$的直积

$$I_{j_1}^{(i)}\times\cdots\times I_{j_r}^{(i)}\subset Q$$

称为第 i 级立方体。显然，这些立方体的总数是 2^{ri}.

每个第$(i-1)$级立方体必可由一些第 i 级立方体之并来表示。特别，对每一个 i，所有第 i 级立方体的并集恰等于 Q. 对 $i=1,2,\cdots$，将所有第 i 级立方体按一定的顺序进行编号，排成一列，

$$Q_1^{(i)},\cdots,Q_{2^{ri}}^{(i)},$$

使得当 $k>j$ 时，第$(i-1)$级立方体 $Q_j^{(i-1)}$ 中所包含的所有第 i 级立方体的下标编号均小于 $Q_k^{(i-1)}$ 中所包含的所有第 i 级立方体的下标编号，即

$$Q_{k'}^{(i)}\subset Q_{l'}^{(i-1)},Q_{k''}^{(i)}\subset Q_{l''}^{(i-1)},l'<l''.$$

应蕴涵不等式 $k'<k''$。利用这一编序，我们来构造取值于 U 的可测函数列

$$u^{(i)}(t),t\in I,\ i=1,2,\cdots.$$

首先，在每一立方体 $Q_j^{(i)}$的内部任取一点$u_j^{(i)},j=1,2,\cdots,2^{ri}$.

记

$$T_j^{(i)}=\{t:U_t\cap Q_{(i)}^{j}\neq\phi\},\ j=1,2,\cdots,2^{ri};\ i=1,2,\cdots.$$

由前面的论述知，集合 $T_j^{(i)}$ 是可测集。由于对每个 $t\in I$，集合 U_t 不空，并且

$$\bigcup_{j=1}^{2^{ri}}Q_j^{(i)}=Q\supset U\supset U_t,$$

故

$$\bigcup_{j=1}^{2^{ri}}T_j^{(i)}=I,\ i=1,2,\cdots.$$

令

$$u^{(i)}(t)=u_1^{(i)},\forall t\in T_1^{(i)},$$

$u^{(i)}(t)=u_k^{(i)},\quad \forall t\in T_k^{(i)}\setminus\bigcup_{j=1}^{k-1}T_j^{(i)},k=2,\cdots,2^{ri};i=1,2,\cdots$. 也就是说，对每一个 $t\in I$，如果与 U_t 相交的第 i 级立方体的最小编号为 k，则定义$u^{(i)}(t)=u_k^{(i)}$.

这样构造出来的 $u^{(i)}(t)$ 在有限个可测集上取有限个常值（不超过 2^{r_i}），因此，它们是可测函数。

我们来证明，对任一 $t\in I, i=1,2,\cdots$，如果 $u^{(i)}(t)\in Q_k^{(i)}$，则对一切 $p=1,2,\cdots$，也有 $u^{(i+p)}(t)\in Q_k^{(i)}$。设 $u^{(i+p)}(t)\in Q_l^{(i+p)}$。因为 $U_t\cap Q_k^{(i)}\neq\phi$，因此在 $Q_k^{(i)}$ 之中的与 U_t 相交的第 $(i+p)$ 级立方体中必存在最小下标编号 l' 的立方体 $Q_{l'}^{(i+p)}$。因此，$l\leqslant l'$. 如果是 $l<l'$，则表明立方体 $Q_l^{(i+p)}$ 含于具有更小的下标 $k'<k$ 的立方体 $Q_{k'}^{(i)}$ 之中。于是，$U_t\cap Q_{k'}^{(i)}\neq\phi$. 这同函数 $u^{(i)}(t)$ 的定义相矛盾. 因此只能是 $l=l'$，从而，$u^{(i+p)}(t)\in Q_{l'}^{(i+p)}\subset Q_k^{(i)}$.

因为 $Q_j^{(i)}$ 的直径随 $i\to\infty$ 而趋于零，注意到 U 的紧性，我们知可测函数列 $u^{(i)}(t)$ 点点收敛于一个取值于 U 的可测函数 $u(t)$.

最后，可测函数 $u(t)$ 满足方程

$$f(t,u(t))=0, \forall t\in I.$$

事实上，因 $u^{(i)}(t)\in Q_j^{(i)}$ 对某个 j 成立，且 $U_t\cap Q_j^{(i)}\neq\phi$，则 $u^{(i)}(t)$ 到闭集 U_t 的距离不大于 $Q_j^{(i)}$ 的直径，后者随 $i\to\infty$ 而趋于零。因此，$\lim\limits_{i\to\infty}u^{(i)}(t)=u(t)\in U_t$，从而 $f(t,u(t))=0\,(t\in I)$. 这就完成了本命题的证明。同时，也完成了定理 8.3 的证明。

注 若控制问题是

$$\dot{x}=f(t,x)+B(t,x)u, u(t)\in\Omega_U, x(t_1)=x_1, x(t_2)=x_2.$$

其中，U 是 $\boldsymbol{R}^r$ 中的一个凸紧集，函数 $f(t,x)$ 及 $B(t,x)$ 都对 t 可测而关于 x 连续可微，则按本节定理，只要方程具有解的无限延展性，并且至少存在给定边值问题的一个解。那么该问题的最优解存在。

实际上，在此情形集合

$$P(t,x)=\{f(t,x)+B(t,x)u; u\in U\}$$

是凸集。

8.4 滑动最优制式

我们先给出控制问题(8.9)的一个简单例子，它满足定理8.3除了

凸性条件之外的所有其他条件，但是没有最优解。

考虑二维自治控制问题：

$$\begin{aligned}&\dot{x}=1-y^2,\\&\dot{y}=u,\\&u(t)\in U=\{1,-1\},\\&x(0)=y(0)=0,\\&x(\tau)=1,y(\tau)=0,(\tau>0).\end{aligned}\tag{8.13}$$

根速度 $\dot{x},\dot{y}$ 的可能取值是由依赖于 y 的两个向量组成的：

$$\begin{pmatrix}1-y^2\\1\end{pmatrix},\ \begin{pmatrix}1-y^2\\-1\end{pmatrix}.$$

这个集合显然非凸。

易见，该问题满足给定边值条件的解必有 $\tau>1$，且存在转移时间 τ 任意接近于 1 的解。

取弱收敛于广义控制 $(1/2)\delta_1+(1/2)\delta_{-1}$ 的控制列 $u^{(i)}(t)$，其中，每个 $u^{(i)}(t)$ 频繁地交替取值 ±1. $u^{(i)}(t)$ 对应的轨线记为

$$x^{(i)}(t),\quad y^{(i)}(t),\quad t\geqslant0,$$

在任一有限区间 $0\leqslant t\leqslant\tau$ 上，$x^{(i)}(t),y^{(i)}(t)$ 一致收敛于极限曲线

$$x(t)=t,\quad y(t)\equiv0,\quad t\geqslant0.$$

显然，这条曲线是广义控制 $(1/2)\delta_1+(1/2)\delta_{-1}$ 所对应的轨线，沿此轨线从 $(0,0)$ 转移到 $(1,0)$ 的时间恰等于 1，这也是问题 (8.13) 的转移时间的（不可达到的）下确界。为此缘故，称极限曲线为控制问题 (8.13) 的滑动最优解，或滑动最优制式。

在定理 8.1 的基础上，可得出如下结论：在 U 是紧集的一般情形，每一个滑动解是问题

$$\dot{x}=f(t,x,u),\ u(t)\in\Omega_U,$$
$$x(t_1)=x_1,\ x(t_2)=x_2$$

的轨线的一个对应于广义控制的一致极限。

实际上，还可得到更强的命题，即每一滑动解是对应于颤振控制的一条轨线。颤振控制是如下形式的广义控制（参见(2.6)式），

$$\mu_t = \sum_{i=1}^{p} \mu_i(t)\delta_{u_i}(t), \quad u_i(t) \in \Omega_U, \quad \mu_i(t) \geqslant 0,$$

$$\sum_{i=1}^{p} \mu_i(t) \equiv 1.$$

这种形式的控制因此也称为滑动控制。

这个事实由下述命题直接获得。

命题 8.3 如果凸控制问题

$$\dot{x} = \langle \mu_t, f(t,x,u) \rangle, \quad \mu_t \in \mathfrak{M}_U,$$

$$x(t_1) = x_1, \quad x(t_2) = x_2$$

满足定理 8.2 的假设，则对它的任一解 $\hat{\mu}_t$, $\hat{x}(t)$ $(t \in [t_1, t_2])$，可选取一个颤振控制

$$\lambda_t = \sum_{i=0}^{n} \lambda_i(t)\delta_{u_i}(t), \quad u_i(t) \in \Omega_U, \quad \lambda_i(t) \geqslant 0, \quad \sum_{i=0}^{n} \lambda_i(t) \equiv 1,$$

其中，n 是列向量 $f(t,x,u)$ 之维数，使得曲线 $\hat{x}(t)$ 满足方程

$$\dot{\hat{x}}(t) = \langle \sum_{i=0}^{n} \lambda_i(t)\delta_{u_i(t)}, f(t, \hat{x}(t), u) \rangle = \sum_{i=0}^{n} \lambda_i(t) f(t, \hat{x}(t), u_i(t)),$$

即 $\hat{x}(t)$ 是该问题对应于颤振控制 λ_t 的轨线。

证 我们利用后面要给出证明的 Carathéodory 引理：

任一集合 $M \subset \boldsymbol{R}^m$ 的凸包重合于集 M 中所有的 $m+1$ 个点的凸包之并集。

设 Λ 是 n 维单纯形：

$$\Lambda = \left\{ \lambda = (\lambda_0, \lambda_1, \cdots, \lambda_n) : \lambda_i \geqslant 0, \sum_{i=0}^{n} \lambda_i = 1 \right\},$$

W 是集 U 的 $(n+1)$ 重直积：

$$W = \overbrace{U \times U \times \cdots U}^{n+1} = \{ w = (u_0, \cdots, u_n) : u_i \in U, i = 0, 1, \cdots, n \}.$$

定义 $g(t,\lambda,w)$ 为如下函数

$$g(t,\lambda,w) = \sum_{i=0}^{n} \lambda_i f(t, \hat{x}(t), u_i) - \langle \hat{\mu}_t, f(t, \hat{x}(t), u) \rangle.$$

其中，$t\in[t_1,t_2]$，$\lambda=(\lambda_0,\lambda_1,\cdots,\lambda_n)\in\Lambda$，$w=(u_0,u_1,\cdots,u_n)\in W$．$f(t,\hat{x}(t),u_i)(i=0,1,\cdots,n)$是$P(t,\hat{x}(t))$中的$n+1$个点，这$n+1$点的凸包由所有如下形式的点组成，

$$\sum_{i=0}^{n}\lambda_i f(t,\hat{x}(t),u_i).$$

根据命题 2.1，

$$\langle\hat{\mu}_t,f(t,\hat{x}(t),u)\rangle\in\operatorname{conv}P(t,\hat{x}(t)).$$

因此，按 Carathéodory 引理，对所有的$t\in[t_1,t_2]$，关于(λ,w)的方程

$$g(t,\lambda,w)=0,\qquad(\lambda,w)\in\Lambda\times W$$

有解．利用命题 8.2，可选取可测函数

$$\lambda(t)\in\Lambda,\quad w(t)\in W,\quad t\in[t_1,t_2],$$

满足方程

$$g(t,\lambda(t),w(t))=0,\quad t\in[t_1,t_2].$$

因此，对几乎所有的$t\in[t_1,t_2]$，等式

$$\dot{\hat{x}}(t)=\langle\hat{\mu}_t,f(t,\hat{x}(t),u)\rangle=\sum_{i=0}^{n}\lambda_i(t)f(t,\hat{x}(t),u_i(t))$$

$$=\langle\sum_{i=0}^{n}\lambda_i(t)\delta_{u_i(t)},f(t,\hat{x}(t),u)\rangle$$

成立．本命题获证．

Carathéodory 引理的证明　我们要证，如果

$$z=\sum_{i=0}^{p}\lambda_i z_i,\ z_i\in M,\ \lambda_i>0,\ \sum_{i=0}^{p}\lambda_i=1,\ p>m,$$

则点z必可表示为点z_i的个数更少的类似的线性和，这就相当于证明了这一引理．

因向量$z_1-z_0,\cdots,z_p-z_0\in\boldsymbol{R}^m$线性相关$(p>m)$，故存在不全为零的数$\nu_1,\cdots,\nu_p$使得

$$\sum_{i=1}^{p}\nu_i(z_i-z_0)=-\sum_{i=1}^{p}\nu_i z_0+\sum_{i=1}^{p}\nu_i z_i=0.$$

记$\nu_0=-\sum_{i=1}^{p}\nu_i$，即得

$$\sum_{i=0}^{p}\nu_i z_i=0,\quad\sum_{i=0}^{p}\nu_i=0.$$

可取 $\varepsilon>0$，使得所有 $\lambda_i+\varepsilon\nu_i$非负 $(i=0,1,\cdots,p)$，并使这些数中至少一个，设为 $\lambda_j+\varepsilon\nu_j$，等于零，即

$$\lambda_i+\varepsilon\nu_i\geqslant 0,\qquad i=0,1,\cdots,p,\qquad \lambda_j+\varepsilon\nu_j=0.$$

设重排各点下标致使 $j=0$ 对应于 $\lambda_0+\varepsilon\nu_0=0$，则得

$$z=\sum_{i=0}^{p}(\lambda_i+\varepsilon\nu_i)\ z_i=\sum_{i=1}^{p}(\lambda_i+\varepsilon\nu_i)z_i,\quad \lambda_i+\varepsilon\nu_i\geqslant 0,\quad i=1,\cdots,p,$$

$$\sum_{i=1}^{p}(\lambda_i+\varepsilon\nu_i)=\sum_{i=0}^{p}(\lambda_i+\varepsilon\nu_i)=\sum_{i=0}^{p}\lambda_i=1.$$

引理获证。

借助于最优滑动解，可以给出比定理 8.3 更为一般的关于常义控制类的最优解存在定理。这一推广往往是有用的（见定理 8.4），它由如下命题给出。

命题 8.4 设定理 8.3 的所有条件除了

$$P(t,x)=\{f(t,x,u):u\in U\}=\mathrm{conv}P(t,x)$$

外都满足，并设上面的条件被下面的更一般的假定所代替：集 $P(t,x)$ 包含（特别情形是重合于）它的凸包 $\mathrm{conv}P(t,x)\subset\boldsymbol{R}^n$ 的边界。那么，控制问题(8.9)存在最优解

$$\tilde{u}(t),\ \tilde{x}(t),\ t_1\leqslant t\leqslant t_2,$$

即最优问题(8.8)有解。

证 根据定理 8.2 及命题 8.3，存在最优解

$$\tilde{\mu}_t,\ \tilde{x}(t),\ t_1\leqslant t\leqslant t_2,$$

其中 $\tilde{\mu}_t$ 是一颤振控制，

$$\tilde{\mu}_t=\tilde{\lambda}_t=\sum_{i=0}^{n}\tilde{\lambda}_i(t)\delta_{\tilde{u}_i(t)},\quad \tilde{u}_i(t)\in\boldsymbol{\Omega}_U,\quad \tilde{\lambda}_i(t)\geqslant 0,\quad \sum_{i=0}^{n}\tilde{\lambda}_i(t)\equiv 1.$$

既然

$$(\tilde{\lambda}(t),\tilde{w}(t))=(\tilde{\lambda}_0(t),\cdots,\tilde{\lambda}_n(t);\tilde{u}_0(t),\cdots,\tilde{u}_n(t)),$$

$$\tilde{x}(t),t\in[t_1,t_2],$$

是控制问题

$$\dot{x}=\sum_{i=0}^{n}\lambda_i f(t,x,u_i),\quad (\lambda(t),w(t))\in\Omega_{A\times W},$$

$$x(t_1)=x_1,\ \ x(t_2)=x_2,$$

的一个最优解，按最大值原理，必存在一个绝对连续函数$\tilde{\psi}(t)\neq 0$ $(t_1\leqslant t\leqslant t_2)$，使得对几乎所有的$t\in[t_1,t_2]$，有

$$\tilde{\psi}(t)\sum_{i=0}^{n}\tilde{\lambda}_i(t)f(t,\tilde{x}(t),\tilde{u}_i(t))=\max_{(\lambda,w)\in\Lambda\times w}\tilde{\psi}(t)\sum_{i=0}^{n}\lambda_i f(t,\tilde{x}(t),u_i),$$

$$(\lambda,w)=(\lambda_0,\cdots,\lambda_n;u_0,\cdots,u_n).$$

这表明，对几乎所有的$t\in[t_1,t_2]$，相速度向量

$$\dot{\tilde{x}}(t)=\sum_{i=0}^{n}\tilde{\lambda}_i(t)f(t,\tilde{x}(t),\tilde{u}_i(t))$$

属于集$\operatorname{conv}P(t,\tilde{x}(t))$的边界，从而按假定属于$P(t,\tilde{x}(t))$自身。

由命题8.2，可取到一个容许控制

$$\tilde{u}(t)\in\Omega_U,$$

使得几乎处处成立

$$f(t,\tilde{x}(t),\tilde{u}(t))=\dot{\tilde{x}}(t).$$

因此，

$$\tilde{u}(t),\tilde{x}(t),\qquad t\in[t_1,t_2],$$

是所要求的最优解。

注 如果函数$f(t,x,u)$不依赖于时间，

$$f(t,x,u)=f(x,u),$$

（参见8.5节）则命题8.4的证明可不依赖于最大值原理。

为此，只须证明，使得相速度向量

$$\sum_{i=0}^{n}\tilde{\lambda}_i(t)f(\tilde{x}(t),\tilde{u}_i(t))=f(t)$$

不属于集$\operatorname{conv}P(\tilde{x}(t))$的边界的那些$t\in[t_1,t_2]$所组成的集合$T$的勒贝格测度为零。

设T_k是满足

$$\left(1+\frac{1}{k}\right)f(t)\in\operatorname{conv}P(\tilde{x}(t))$$

的那些时刻$t\in[t_1,t_2]$之集合。可见$T\subset\bigcup_{k=1}^{\infty}T_k$. 因此，若 meas $T>0$，则有某个$\hat{k}$，meas $T_{\hat{k}}>0$.

我们将证明，如果存在这样一个$\hat{k}$，则从$\tilde{x}(t_1)$到$\tilde{x}(t_2)$的转移时间可缩短。这同$\tilde{\mu}_t$，$\tilde{x}(t)$ $(t\in[t_1,t_2])$是最优解的假设矛盾。

我们记

$$\omega(t)=\begin{cases}1, & t\in[t_1,t_2]\backslash T_{\hat{k}};\\[2ex] 1+\dfrac{1}{\hat{k}}, & t\in T_{\hat{k}}.\end{cases}$$

引进新的时间变量τ如下：

$$\frac{d\tau}{dt}=\frac{1}{\omega(t)},\qquad \tau(t_1)=0.$$

当t从t_1变到t_2时，τ从0变到

$$\hat{\tau}=\int_{t_1}^{t_2}\frac{dt}{\omega(t)}=t_2-t_1-\text{meas}T_{\hat{k}}+\frac{\text{meas}T_{\hat{k}}}{1+(1/\hat{k})}<t_2-t_1.$$

函数

$$\tilde{\mu}_{t(\tau)},\ \tilde{x}(t(\tau)),\qquad 0\leqslant\tau\leqslant\hat{\tau}$$

满足方程

$$\frac{d\tilde{x}}{d\tau}=\omega(t(\tau))\sum_{i=0}^{n}\tilde{\lambda}_i(t(\tau))f(\tilde{x}(t(\tau)),\tilde{u}_i(t(\tau)))\in\text{conv}P(\tilde{x}(t(\tau))),$$

以及边值条件

$$\tilde{x}(t(0))=\tilde{x}(t_1),\qquad \tilde{x}(t(\hat{\tau}))=\tilde{x}(t_2).$$

因此，这个解使转移时间小于t_2-t_1.

8.5 正则变分问题的存在性定理

为了阐明所得结果，现来证明古典变分法中的一个定理，即正则

变分问题解的存在性定理。该定理的实质性部分是 Hilbert 与 Tonelli 给出的

设

$$f(\tau,z,u),\quad (\tau,z,u)\in \boldsymbol{R}\times \boldsymbol{R}^n\times \boldsymbol{R}^n$$

是连续可微函数。若对任意固定的 τ 和z，成立

$$f(\tau,z,\alpha u_1+\beta u_2)\leqslant \alpha f(\tau,z,u_1)+\beta f(\tau,z,u_2),$$

$$\forall \alpha,\beta\geqslant 0,\ \alpha+\beta=1,\ \forall u_1,u_2\in \boldsymbol{R}^n,$$

则称这一函数关于 u 是凸的。

简单变分法问题是在满足给定的边界条件

$$z(\tau_1)=z_1,\qquad z(\tau_2)=z_2$$

的所有绝对连续函数

$$z(\tau),\qquad \tau\in[\tau_1,\tau_2],$$

的集合中求积分泛函

$$J(z(\tau))=\int_{\tau_1}^{\tau_2} f\left(\tau,z(\tau),\ \frac{dz(\tau)}{d\tau}\right)\ d\tau$$

达到最小值的 $z(\tau)$.

如果函数 $f(\tau,z,u)>0$ 且关于 u 是凸的，则变分问题

$$\min\int_{\tau_1}^{\tau_2} f\left(\tau,z(\tau),\frac{dz(\tau)}{d\tau}\right)d\tau$$

称为正则的[注1]。

定理 8.4 设 $f>0$且满足附加条件

$$\frac{|u|}{f(\tau,z,u)}\longrightarrow 0\ \ (|u|\to\infty),\tag{8.14}$$

其中，τ 和 z 是任意固定的，则简单变分法问题

$$\min\int_{\tau_1}^{\tau_2} f\left(\tau,z,\frac{dz}{d\tau}\right)d\tau,\ z(\tau_1)=z_1,z(\tau_2)=z_2,$$

[注1] 在这一定义中，f 大于零的要求可用 f 有下界代替，因为两个有固定端点的变分问题只要被积函数之差是常数，就是等价的。

必有解$z(\tau)(\tau\in[\tau_1,\tau_2])$.

证明是将此变分法问题归结为具有固定端点的时间最优问题，再应用命题 8.4. 由于$f(\theta,z(\theta),dz(\theta)/d\theta)>0$，故对任一绝对连续曲线$z(\tau),\tau\in[\tau',\tau'']$,不定积分

$$t=\int_{\tau'}^{\tau} f(\theta,z(\theta),\frac{dz(\theta)}{d\theta})d\theta,\qquad \tau\in[\tau',\tau''],$$

是变量τ的严格单调函数。取这一积分为新的独立变量——时间，将z和τ作为相变量，而导数dz/dt作为控制变量u，即可写出如下形式的等价的最优问题：

$$\begin{cases}\dfrac{d\tau}{dt}=\dfrac{1}{f(\tau,z,u)} & \tau(0)=\tau_1,\tau(T)=\tau_2,\\ \dfrac{dz}{dt}=\dfrac{dz}{d\tau}\dfrac{d\tau}{dt}=\dfrac{u}{f(\tau,z,u)}, & z(0)=z_1,z(T)=z_2,\\ \min T.\end{cases}$$

现在我们考虑以$(\tau,x)\in\boldsymbol{R}\times\boldsymbol{R}^n$为相点,$u\in\boldsymbol{R}^n$为控制的自治控制问题：

$$\begin{cases}\dfrac{d\tau}{dt}=\dfrac{1}{f(\tau,x,u)},\quad \dfrac{dx}{dt}=\dfrac{u}{f(\tau,x,u)},\quad u(t)\in\Omega_R{}^n,\\ \tau(0)=\tau_1,\qquad \tau(T)=\tau_2,\qquad x(0)=z_1,\qquad x(T)=z_2,\end{cases}\tag{8.15}$$

若将点$(0,0)\in\boldsymbol{R}\times\boldsymbol{R}^n$添加到问题(8.15)中任意$(\tau,x)$点的所有可能的相速度的集合之中,则由条件(8.14)，这个相速度集合成为$\boldsymbol{R}^{1+n}$中的紧集。记它为$P(\tau,x)$,即

$$P(\tau,x)=\{(0,0)\}\cup\left\{\left(\frac{1}{f(\tau,x,u)},\ \frac{u}{f(\tau,x,u)}\right):u\in\boldsymbol{R}^n\right\}\subset\boldsymbol{R}^{1+n}.$$

现来证明，集$P(\tau,x)$可表为n维球面$S^n\subset\boldsymbol{R}^{n+1}$的一个同胚像。为此，我们固定一点$\hat{w}\in S^n$,并且先给出$S^n\setminus\{\hat{w}\}$到$\boldsymbol{R}^n$的一个同胚映射(例如，球极投影)

$$u:S^n\setminus\{\hat{w}\}\longrightarrow\boldsymbol{R}^n,\tag{8.16}$$

使得在此映射下，任一收敛于 $\hat{w}$ 的点列 $w_i \in S^n \setminus \{\hat{w}\}$ 被映成点列 $u(w_i) \in \boldsymbol{R}^n$，后者趋于无穷。考虑定义在 $S^n \setminus \{\hat{w}\}$ 上关于 w 的连续函数

$$\frac{1}{f(\tau, x, u(w))} \quad \text{和} \quad \frac{u(w)}{f(\tau, x, u(w))},$$

扩充定义这两个函数在 $\hat{w}$ 处等于零。由条件(8.14)，在扩充定义后，它们在整个球面 S^n 上有定义且连续。

直接可得，这样定义的连续映射是 1 对 1 的。由于 S^n 是紧集，因此该映射是同胚。

于是，定义如下的自治控制问题：

$$\begin{cases} \dfrac{d\tau}{dt} = \dfrac{1}{f(\tau, x, u(w))}, & \dfrac{dx}{dt} = \dfrac{u(w)}{f(\tau, x, u(w))}, \quad w(t) \in \Omega_{S^n}, \\ \tau(0) = \tau_1, \quad \tau(T) = \tau_2, & x(0) = z_1, \quad x(T) = z_2 \end{cases} \tag{8.17}$$

如果要求转移时间 T 为最小：

$$\min T, \tag{8.18}$$

则此问题成为形如(1.3)的具有允许控制取值紧集 $U = S^n \subset \boldsymbol{R}^{n+1}$ 的一个最优问题。

对每一绝对连续曲线

$$z(\tau), \quad \tau \in [\tau_1, \tau_2], \quad z(\tau_1) = z_1, \quad z(\tau_2) = z_2, \tag{8.19}$$

可以证明必存在控制问题(8.17)的解

$$w(t), \quad \tau(t), \quad x(t), \quad 0 \leqslant t \leqslant T. \tag{8.20}$$

使得解(8.20)的转移时间 T 恰等于积分 $J(z(\tau))$：

$$T = J(z(\tau)) = \int_{\tau_1}^{\tau_2} f\left(\tau, z(\tau), \frac{dz(\tau)}{d\tau}\right) d\tau, \tag{8.21}$$

这里，对几乎所有的 $t \in [0, T]$，要求。$w(t) \neq \hat{w}$。反之，每一个这样的解(8.20)都可由这一对应得到。有了这个事实，定理的证明就可归结为最优问题(8.17)—(8.18)的解(8.20)的存在性。

现在先来证明这个事实。对给定的(8.19)引进绝对连续函数

$$t = t(\tau) = \int_{\tau_1}^{\tau} f\left(\theta, z(\theta), \frac{dz(\theta)}{d\theta}\right) d\theta, \quad \tau \in [\tau_1, \tau_2].$$

由于对几乎所有的 $\tau\in[\tau_1,\tau_2]$，函数 $t=t(\tau)$ 的导数大于零， $dt/d\tau=f(\tau,z(\tau),dz(\tau)/d\tau)>0$，因此，逆函数

$$\tau=\tau(t),\qquad 0=t(\tau_1)\leqslant t\leqslant T=t(\tau_2)=\int_{\tau_1}^{\tau_2}f(\theta,z(\theta),\frac{dz(\theta)}{d\theta})d\theta$$

存在，并且也是绝对连续的。

我们来定义解(8.20)．令

$$w(t)=u^{-1}\left(\frac{dz(\tau(t))}{d\tau}\right),\qquad \tau(t),\qquad x(t)=z(\tau(t)),$$

$$0\leqslant t\leqslant T=J(z(\tau)).$$

其中 u^{-1} 是(8.16)之逆映射。函数

$$\tau(t)\ \text{与}\ x(t),\qquad 0\leqslant t\leqslant T,$$

是满足给定的边界条件的绝对连续函数，其中复合函数 $x(t)=z(\tau(t))$ 的绝对连续性由函数 $z(\tau)$ 和 $\tau(t)$ 各自的绝对连续性以及 $d\tau/dt>0$ 几乎处处成立的事实得知．又因 u^{-1} 是连续函数，$dz(\tau)/d\tau$ 是可测的，并且 $\tau(t)$ 绝对连续，导数几乎处处大于零，故知函数 $w(t)$ 是可测函数。

有了这些注述，由直接微分可以验证函数 $w(t),z(t),x(t)$ 满足微分方程(8.17)。这就证实了由曲线（8.19）对应于满足（8.21）的解(8.20)．

再证明反向的对应关系．已有解(8.20)，要定义曲线(8.19)这可借助于前一段已述的方法实现。

由于对几乎所有的 $t\in[0,T],w(t)\neq\hat{w}$,故

$$\frac{d\tau}{dt}=\frac{1}{f(\tau(t),x(t),u(w(t)))}>0,$$

因此，函数 $\tau(t)$，$0\leqslant t\leqslant T$ 的逆函数

$$t(\tau),\qquad \tau_1\leqslant\tau\leqslant\tau_2,$$

是绝对连续的。我们置

$$z(\tau)=x(t(\tau)),\qquad \tau_1\leqslant\tau\leqslant\tau_2$$

由于 $x(t)$ 和 $t(\tau)$ 都是绝对连续的且 $t(\tau)$ 的导数几乎处处为正，这样定义的 $z(\tau)$ 是绝对连续函数且满足边界条件

$$z(\tau_1)=z_1,\qquad z(\tau_2)=z_2.$$

我们要证，对几乎所有的 $t\in[0,T]$，有

$$\frac{dz(\tau(t))}{d\tau}=u(w(t)),$$

且

$$J(z(\tau))=\int_{\tau_1}^{\tau_2}f\left(\tau,z(\tau),\frac{dz(\tau)}{d\tau}\right)d\tau=T.$$

前者可以直接求导进行验证。后者将

$$\frac{dt}{d\tau}=f(\tau,x(t(\tau)),u(w(t(\tau))))$$

从 τ_1 到 τ_2 关于 $d\tau$ 积分可得，其中的函数 $x(t(\tau))$ 与 $u(w(t(\tau)))$ 可分别用函数 $z(\tau)$ 及 $\frac{dz(\tau)}{d\tau}$ 代入。这就证明了两者的对应关系。

现在尚待证明的是最优问题(8.17)—(8.18)有解(8.20)。

如果我们能够证明，对任意的 τ 和 x，集合 $P(\tau,x)\subset\boldsymbol{R}^{1+n}$ 是某一个有界开凸集

$$D(\tau,x)\subset\boldsymbol{R}^{1+n}$$

的边界，那么由命题 8.4 和最大值原理可推出解的存在性。我们先来论证这一点。这样，$P(\tau,x)$ 重合于凸包 $\mathrm{conv}P(\tau,x)$ 的边界。又因控制问题(8.17)总有一个解，且(8.17)的方程具有解的无限延展性，根据命题8.4，最优问题(8.17)—(8.18)有解

$$\tilde{w}(t),\qquad\tilde{\tau}(t),\qquad\tilde{x}(t),\qquad 0\leqslant t\leqslant T.$$

利用最大值原理，我们证明最优解必对几乎所有的 $t\in[0,T]$，

$$\tilde{w}(t)\neq\hat{w}.$$

由于所考虑的系统（指转化后的系统(8.17)）是自治的，按最大值原理，存在一个绝对连续函数

$$\tilde{\psi}(t)=(\tilde{\chi}_0(t),\tilde{\chi}(t))\neq 0,\qquad 0\leqslant t\leqslant T,$$

使得对几乎所有的 $t\in[0,T]$，成立

$$\frac{\tilde{\chi}_0(t)}{f(\tilde{\tau}(t),\tilde{x}(t),u(\tilde{w}(t)))}+\frac{\tilde{\chi}(t)u(\tilde{w}(t))}{f(\tilde{\tau}(t),\tilde{x}(t),u(\tilde{w}(t)))}$$

$$=\sup_{u\in R^n}\frac{\tilde{\chi}_0(t)+\tilde{\chi}(t)u}{f(\tilde{\tau}(t),\tilde{x}(t),u)}=C\geqslant 0.$$

如果$C>0$,则因在$\hat{w}$点将有

$$\frac{1}{f(\tau,x,u(\hat{w}))}=\frac{u(\hat{w})}{f(\tau,x,u(\hat{w}))}=0,$$

表明$\tilde{w}(t)\doteq\hat{w}$对几乎所有的$t\in[0,T]$成立。如果$C=0$，则必有$\tilde{\chi}(t)\equiv 0,\tilde{\chi}_0(t)<0$，因此，对几乎所有的$t\in[0,T]$,

$$\tilde{w}(t)=\hat{w}.$$

但是,如果$\tau_1\neq\tau_2$,这种情形是不会出现的。因此一旦$\tilde{w}(t)\doteq\hat{w}$,则相点$(\tilde{\tau}(t),\tilde{x}(t))(0\leqslant t\leqslant T)$将驻定，产生矛盾。

现在，我们还须构作上述的有界开凸集$D(\tau,x)$.利用集合

$$Q(\tau,x)=P(\tau,x)\setminus\{(0,0)\}$$

的参数表示

$$Q(\tau,x)=\left\{(v_0,v)=\left(\frac{1}{f(\tau,x,u)},\ \frac{u}{f(\tau,x,u)}\right):u\in \boldsymbol{R}^n\right\}\subset\boldsymbol{R}^{1+n},$$

即$v_0=1/f(\tau,x,u),v=u/f(\tau,x,u),u\in\boldsymbol{R}^n$. 由此可见，$u=v/v_0$. 因此，

$$Q(\tau,x)=\{(v_0,v)\in\boldsymbol{R}^{1+n}:v_0f(\tau,x,v/v_0)=1,v_0>0\}.$$

我们定义$D(\tau,x)$为$\boldsymbol{R}^{1+n}$中的如下点集，

$$D(\tau,x)=\{(v_0,v):v_0>0,v_0f(\tau,x,v/v_0)<1\}\subset\boldsymbol{R}^{1+n}.$$

由于函数f的连续性，这个集合是开集，我们证明它是有界集。若$|(v_0,v)|\to\infty$,那么或是$|v_0|\to\infty$,或是$|v|\to\infty$.对$|v_0|\to\infty$,我们有

$$|v_0f(\tau,x,v/v_0)|\to\infty,$$

这是因为条件(8.14)及连续性保证了函数$f(\tau,x,u)$以一正值常数为下界。对$|v|\to\infty$,由条件(8.14),有

$$|v_0 f(\tau,x,v/v_0)| = |v|\frac{f(\tau,x,v/v_0)}{|v/v_0|}\to\infty.$$

这就表明$D(\tau,x)$是有界集。

再则，易见$P(\tau,x)$包含集合$D(\tau,x)$的边界。事实上，如果$(v_0,v)\in D(\tau,x)$且$(v_0,v)\to(\hat{v}_0,\hat{v})$，$\hat{v}_0\neq 0$，则

$$v_0 f(\tau,x,v/v_0)\to\hat{v}_0 f(\tau,x,\hat{v}/\hat{v}_0)\leqslant 1;$$

如果$(\hat{v}_0,\hat{v})\notin D(\tau,x)$，则

$$\hat{v}_0 f(\tau,x,\hat{v}/\hat{v}_0)=1,$$

即$(\hat{v}_0,\hat{v})\in P(\tau,x)$．另一方面，如果$\hat{v}_0=0$，则也必有$\hat{v}=0$，否则$\hat{v}\neq 0$，可设$v$充分接近于$\hat{v}$时$|v|\neq 0$，从而

$$v_0 f(\tau,x,v/v_0)=|v|\frac{f(\tau,x,v/v_0)}{|v|/v_0}\to\infty,\ (v_0\to\hat{v}_0,v\to\hat{v}).$$

这同$(v_0,v)\in D(\tau,x)$相矛盾。所以，$(\hat{v}_0,\hat{v})=(0,0)\in P(\tau,x)$.

反之，每一点$(\hat{v}_0,\hat{v})\in P(\tau,x)$都可表为点$(v_0,v)\in D(\tau,x)$的极限。我们来证明这一点。若$\hat{v}_0\neq 0$，则$\hat{v}>0$，对$0<\lambda<1$，有$(\lambda\hat{v}_0,\lambda\hat{v})\in D(\tau,x)$，因为

$$\lambda\hat{v}_0 f(\tau,x,\lambda\hat{v}/\lambda\hat{v}_0)=\lambda\hat{v}_0 f(\tau,x,\hat{v}/\hat{v}_0)\leqslant\lambda<1.$$

但$\lambda\to 1$时，$(\lambda\hat{v}_0,\lambda\hat{v})\to(\hat{v}_0,\hat{v})$，即知$(\hat{v}_0,\hat{v})$是$D(\tau,x)$中点的极限。若$\hat{v}_0=0$，则$\hat{v}=0$，我们可得，当$v_0\to 0(v_0>0)$时，

$$(v_0,0)\to(\hat{v}_0,\hat{v})=(0,0),$$

并且，对$v_0>0$充分小，$v_0 f(\tau,x,0)<1$，表明$(v_0,0)\in D(\tau,x)$.

最后，集合$D(\tau,x)$的凸性由$f(\tau,x,u)$关于u的凸性易得。设

$$v'_0 f(\tau,x,v'/v'_0)<1,v''_0 f(\tau,x,v''/v''_0)<1,v'_0,v''_0>0,\alpha,\beta\geqslant 0,\alpha+\beta=1.$$

我们要证

$$(\alpha v_0' + \beta v_0'')f(\tau, x, \frac{\alpha v' + \beta v''}{\alpha v_0' + \beta v_0''}) < 1.$$

记 $\Delta = \alpha v_0' + \beta v_0''$，利用 f 关于 u 的凸性，我们最后获得了

$$\Delta f(\tau, x, \frac{\alpha v' + \beta v''}{\Delta}) = \Delta f\left(\tau, x, \frac{\alpha v_0'}{\Delta} \cdot \frac{v'}{v_0'} + \frac{\beta v_0''}{\Delta} \cdot \frac{v''}{v''}\right)$$

$$\leqslant \Delta \left\{ \frac{\alpha v_0'}{\Delta} f(\tau, x, \frac{v'}{v_0'}) + \frac{\beta v_0''}{\Delta} f(\tau, x, \frac{v''}{v_0''}) \right\}$$

$$= \alpha v_0' f\left(\tau, x, \frac{v'}{v'}\right) + \beta v_0'' f\left(\tau, x, \frac{v''}{v_0''}\right) < \alpha + \beta = 1.$$

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍。

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962 (有中译本：最佳过程的数学理论).

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonnė, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] И.П. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonnė, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962 (有中译本：最佳过程的数学理论).

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍。

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] И.П. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962 (有中译本：最佳过程的数学理论).

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著:

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962 (有中译本: 最佳过程的数学理论).

最优控制现代理论可参见下列专著, 它们互为补充:

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中, 我们推荐下列两本:

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著, 徐瑞云译, 实变函数论(上册), 人民教育出版社, 1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直, 泛函分析讲义, 高等教育出版社, 1958.

参 考 文 献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962（有中译本：最佳过程的数学理论）.

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍.

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.

参考文献

书中用到的数学结果的详细叙述可参见专著：

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, J. Wiley, New York, 1962 (有中译本：最佳过程的数学理论).

最优控制现代理论可参见下列专著，它们互为补充：

[2] H. Hermes and J. P. LaSalle, *Functional Analysis and Time-Optimal Control*, Academic Press, New York, 1969.

[3] L. D. Berkovitz, *Optimal Control Theory*, Springer Verlag, New York, 1974.

[4] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

Warga 的书中有测度论、凸集理论等辅助性内容的详细介绍。

在专门介绍凸集与凸函数理论以及 Radon 测度理论的专著中，我们推荐下列两本：

[5] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.

[6] J. Dieudonné, *Treatise on Analysis*, Vol. 2, Academic Press, New York, 1970.

本书中译本参考文献

[7] П.И. 那汤松著，徐瑞云译，实变函数论(上册)，人民教育出版社，1958.

[8] Б. М. Левитан, Усп., 1947(5), 133—192.

[9] 关肇直，泛函分析讲义，高等教育出版社，1958.