遍 历 理 论 及 其 应 用

叶向东　黄 文　邵 松

2023年5月版本

中国科学技术大学数学科学学院

目 录

引 言

1880 年左右, 玻尔兹曼(Boltzmann)和麦克斯韦尔(Maxwell)等人试图通过力学模型及其基本数学原理来解释热力学现象. 在这种情况下, 玻尔兹曼（Boltzmann）在1885 年创造了“Ergode”一词. 遍历理论是数学的一个分支, 研究动力系统的统计特性. 此处统计特性是指各种函数沿动力系统轨道取时间平均表现的特性. 与概率论一样, 遍历理论也是基于测度理论的一门学科. 它最初的发展受统计物理学问题的推动. 遍历理论不是经典的数学学科之一, 不像数论、概率论等学科, 它的名称没有指出其主题.

遍历理论的中心问题是研究动力系统在长时间运动时的行为. 遍历理论的第一个重要结果是庞加莱（Poincaré）回复定理, 该定理声称: 相空间的任何子集中的几乎所有点在迭代下最终都会重新返回该集合. 各种遍历定理提供了更精确的信息, 这些定理断言在某些条件下, 函数沿着轨道的时间平均值几乎随处都存在, 并且与空间平均值有关. 遍历理论早期最重要的两个定理是 von Neumann 和 Birkhoff 的定理,它们断言了沿轨道时间平均的存在性. 对于所谓的遍历系统, 几乎所有初始点的时间平均都是相同的: 从统计学上讲, 长时间演化的系统“忘记了”其初始状态.

遍历性的概念和遍历假设是遍历理论应用的核心. 基本思想是, 对于某些系统, 其对于时间的平均值等于整个空间的平均值. 遍历理论在数学其他分支上的应用通常涉及为所研究特殊类型的系统建立其遍历性. 例如, 在几何学中使用遍历理论的方法研究黎曼流形上的测地流；在概率论中研究马尔可夫链等. 遍历理论与概率论、调和分析、Lie 群理论、数论等有着紧密的联系.

发展至今, 遍历理论已经成为数学中一门非常成熟的分支之一. 从 Halmos 撰写了第一本遍历论的专著[92]起, 到现今为止关于它的英文著作不下40 余种. 因为遍历论涉及的内容实在广博精深, 很难有一本书能够将其方方面面完整的介绍. 受限于作者对于遍历理论知识的理解和研究, 我们没有能力也没有计划写一本包罗万象的教材. 本书的目的是从我们的理解和偏爱出发编写一本遍历论的入门教材, 向国内数学工作者介绍这门学科. 我们的目标是: 此教材包含了遍历理论的基本知识, 并有若干章节涉及研究的前沿和应用. 在参考文献中我们罗列了一些遍历理论的国内外优秀专著[1, 13, 21, 43, 47, 54, 59, 64, 76, 92, 128, 149, 153, 155, 161, 162, 175, 184, 191, 195, 206], 感兴趣的读者可以自己查阅.

遍历理论包括遍历定理、熵理论、谱理论、不交性理论、结构定理、遍历 Ramsey 理论、轨道等价等众多内容. 遍历定理是遍历论的核心内容之一. 在本引言中, 我们不准备面面俱到, 下面仅从遍历定理入手初步领略这门学科的风貌. 首先我们给出一些必要的数学概念, 更详细的论述请参见本书正文. 保测系统是指四元组 $(X, \mathcal{X}, \mu, T)$, 其中 $(X, \mathcal{X}, \mu)$ 为测度空间, $T : (X, \mathcal{X}, \mu) \to (X, \mathcal{X}, \mu)$ 为可测变换, 并且它是保测的, 即 $\mu(T^{-1}E) = \mu(E)$ 对任何 $E \in \mathcal{X}$ 成立. 为简单起见, 我们一般假设 $\mu(X) = 1$, 即 $(X, \mathcal{X}, \mu)$ 为概率空间. 称 T 为遍历的, 是指对于任何不变可测集要么它为零测集, 要么其补集为零测集, 即 $T^{-1}E = E$ 蕴含 $\mu(E) = 0$ 或 $\mu(X \setminus E) = 0$. 对 $f \in L^1(\mu)$, 称 $A_nf(x) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ 为**遍历平均**.

第一个遍历定理是由著名数学家 von Neumann 给出的[157]. Koopman 注意到对于可逆保测系统 $(X,\mathcal{X},\mu,T)$, $U_T: L^2(\mu)\to L^2(\mu), f\mapsto f\circ T$ 定义了一个酉算子. 这启发 von Neumann 给出了著名的平均遍历定理：设 $\mathcal{H}$ 为 Hilbert 空间, $U:\mathcal{H}\to\mathcal{H}$ 为酉算子, P 为 $\mathcal{H}$ 到全体 U 不变向量组成空间上的投射, 那么对于任何 $x\in\mathcal{H}$, 我们有

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1}U^ix-Px\right\|\to 0,\ n\to\infty.$$

上面的 von Neumann 遍历定理是目前文献中常用的表述, 但是需要注意的是 von Neumann 原始文章处理的是连续流而不是离散系统. 根据上述结论, 容易得到 von Neumann 遍历定理的 L^p 形式：设 $(X,\mathcal{X},\mu,T)$ 为保测系统, $1\le p<\infty$. 如果 $f\in L^p(\mu)$, 那么存在 T 不变函数 $f^*\in L^p(\mu)$（即 $f^*\circ T=f^*$ a.e.）使得

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x))-f^*(x)\right\|_{L^p}\to 0,\ n\to\infty.$$

受到 von Neumann 工作的启发, Birkhoff 在1931年得到了他的逐点收敛遍历定理[23]. 历史上 von Neumann 遍历定理先出现, 但 Birkhoff 工作发表的时间在前. 下面给出 Birkhoff 遍历定理的陈述. 设 $(X,\mathcal{X},\mu,T)$ 为保测系统, $f\in L^1(\mu)$, 则存在 T 不变函数 $f^*\in L^1(\mu)$ 使得

$$A_nf(x)=\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x))\to f^*(x),\mu-a.e.,\ n\to\infty.$$

并且 $\int_E f^*d\mu=\int_E fd\mu$ 对于任何 T 不变可测集 E 成立；当 T 遍历的时候, $f^*=\int_X fd\mu$. 上述表达方式也不是 Birkhoff 遍历定理原始的陈述形式, Birkhoff 原始定理是针对流形上光滑流给出的, 但是其证明实际上对一般保测系统是成立的.

设 $A=\{a_1<a_2<\ldots\}$ 为非负整数序列, 令沿序列 A 的遍历平均为

$$M_n(A,f)(x)=\frac{1}{n}\sum_{1\le i\le n}f(T^{a_i}x).$$

在遍历理论中一个非常重要的问题是对于给定序列 A, 研究对于函数 $f\in L^p(\mu)$ 平均 $M_n(A,f)$ 是否平均收敛？逐点收敛？设 $(X,\mathcal{X},\mu,T)$ 为非周期保测映射, $1\le p\le\infty$. 称 $A=\{a_n\}_{n\in\mathbb{N}}$ 相对系统 $(X,\mathcal{X},\mu,T)$ 为**平均（逐点） L^p 好的**是指对于每个 $f\in L^p(\mu)$, 极限

$$\lim_{t\to\infty}M_n(A,f)(x)$$

对于 L^p 模下(几乎处处)存在. 称序列 $A=\{a_n\}_{n\in\mathbb{N}}$ 为**普遍平均(逐点) L^p 好的**,是指他相对任何非周期系统都是平均(逐点) L^p 好的. 类似定义**平均(逐点) L^p 坏的**, 例如 A 为普遍逐点 L^1 坏的, 是指对于任何系统 $(X,\mathcal{X},\mu)$, 存在 $f\in L^1(\mu)$ 使得 $\lim_{t\to\infty}M_n(A,f)(x)$ 对于某个正测度集上的 x 不存在. von Neumann L^p 形式遍历定理等价于 $\mathbb{N}$ 为普遍平均 L^p

好的；Birkhoff 遍历定理等价于 $\mathbb{N}$ 为普遍逐点 L^1 好的. 1971 年 Krengel 构造了第一个逐点普遍坏的序列；Bellow 于1983 年证明了任何间歇(lacunary)序列（指存在 $c>1$ 使得对于任何 n 有 $a_{n+1}>ca_n$）为普遍逐点 L^1 坏的.

设 q 为整数值多项式, 那么不难证明序列 $\{q(n)\}_{n\in\mathbb{N}}$ 为普遍平均 L^2 好的. 但是对于逐点就困难很多. 1987 年 Bourgain 证明了他著名的多项式逐点收敛定理：设 q 为整数值多项式, 那么序列 $\{q(n)\}_n$ 为普遍逐点 L^2 好的[34]. 之后他推广了自己结果到任何 L^p, $p>1$, 并且他证明了对于任何实值多项式 q, $\{[q(n)]\}_{n\in\mathbb{N}}$ ($[\cdot]$ 为取整符号) 为普遍逐点 L^p 好的, $p>1$. Bourgain 还证明了素数集合为普遍逐点 L^p 好的, $p>\frac{1+\sqrt{3}}{2}$. 很快 Wierdl 于 1988 年推广了这个结果, 证明了对于任何 $p>1$, 素数集合为普遍逐点 L^p 好的. 在很长的一段时间里面, 相关结果在 $p=1$ 的情况如何是此方向上最重要的问题. 直到2005 年 Buczolich 和 Mauldin 证明了整数值多项式是普遍逐点 L^1 坏的[37].

目前多重遍历平均收敛问题是遍历理论研究的热门之一. 为了回答 Erdös 的一个著名猜测, 1975 年 Szemerédi 证明了：具有正上 Banach 密度的自然数子集包含任意有限长的等差数列[186]. Furstenberg 在 1977 年用遍历理论的方法给出了 Szemerédi 定理的新证明, 他的工作开创了遍历 Ramsey 理论这一崭新的数学分支[63]. 这也是他获 2007 年 Wolf 奖和 2020 年 Abel 奖的主要工作之一. 结合 Furstenberg 证明的思想, 2008 年 Green 和 Tao (陶哲轩) 解决了一个古老的数论问题：素数集合包含任意长的等差数列[85]. 这项工作是陶哲轩获Fields奖的主要工作之一.

Furstenberg 证明了对于任何保测系统 $(X,\mathcal{X},\mu,T)$, $k\in\mathbb{N}$, 以及满足条件 $f\geq 0$和$\int fd\mu>0$ 的 $f\in L^\infty(\mu)$, 我们有

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f(T^nx)f(T^{2n}x)\cdots f(T^{kn}x)>0.$$

他的结果直接引发了如下问题: 给定任何保测系统$(X,\mathcal{X},\mu,T)$, $k\in\mathbb{N}$, X 上可测函数 $f_1,\ldots,f_k\in L^\infty(\mu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1}f_1(T^nx)f_2(T^{2n}x)\cdots f_k(T^{kn}x)$$

是否在 L^2 或逐点收敛意义下收敛? 上述平均称为**多重遍历平均**, $k=1$ 的情形就是标准的遍历平均.

L^2 收敛意义下的多重遍历平均问题由 Conze-Lesigne[40, 41, 42]、Furstenberg-Weiss[72]、Host-Kra[100]、陶哲轩[188]、Ziegler[202] 等人逐步用了几十年的时间得到解决. L^2 多重遍历平均问题最终在 2005 年由 Host 与 Kra 解决[100](Ziegler 稍后独立给出另一个证明[202]). 2008 年陶哲轩将之推广到有限个相互交换保测映射的多重遍历平均收敛[188]；2012 年 Walsh 将陶哲轩的结果推广到幂零群作用的情况[194].

逐点收敛意义下多重遍历平均问题是一个极其困难的问题. 目前逐点收敛意义下多重遍历平均的成果不多. 一个重要的进展是 Bourgain 给出的. Bourgain 给出多项式时

间下 1 重的逐点收敛遍历定理[34]：对整数值多项式 $p(n) \in \mathbb{Z}[n]$ 和 $f \in L^\infty(\mu)$,平均 $\frac{1}{N}\sum_{n=0}^{N-1} f(T^{p(n)}x)$ 几乎处处收敛；他还证明了单个映射 2 重的逐点收敛多重遍历定理[35]：对 $a \neq b \in \mathbb{N}, f_1, f_2 \in L^\infty(\mu)$, 平均 $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x)f_2(T^{bn}x)$ 几乎处处收敛. 这些工作是 Bourgain 获 1994 年 Fields 奖的主要工作之一. 最近黄文、邵松和叶向东运用拓扑模型理论, 结合拓扑与测度方法对于遍历 distal 系统 $(X, \mathcal{X}, \mu, T)$ 证明了对于任意 $k \in \mathbb{N}$ 以及 f_1, $\ldots$, $f_k \in L^\infty(\mu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x)\cdots f_k(T^{kn}x)$$

几乎处处收敛[113].

遍历定理是一个很广泛的课题, 以上仅是罗列了一小部分例子. 更多相关结论可以参看相关参考文献, 例如[136]. 希望通过上面介绍大家对遍历论有所认识.

本书的目的是介绍遍历理论的基本概念、一些基本方法和应用. 我们撰写的时候期望尽量能做到自封闭. 在第零章, 我们简要的总结了在本书中会用到的基本知识. 这部分内容主要涉及测度论、泛函、拓扑群理论等.

在第一章中, 我们介绍保测系统的概念和基本性质. 这一章的重点是众多的例子, 这些例子都是动力系统中常见的系统,它们会贯穿整本书. 在本章最后我们介绍遍历论中著名的 Poincaré 定理.

在第二章我们给出遍历性的定义和刻画,并且介绍 von Neumann 遍历定理和 Birkhoff 遍历定理. von Neumann 遍历定理和 Birkhoff 遍历定理是遍历论最基本的定理. 在介绍完遍历性后, 我们给出各种混合的定义, 其中最为重要的是弱混合. 弱混合的谱刻画是指它等价于连续谱, 我们在本章最后给出动力系统谱理论的一些基本知识.

第一章和第二章是遍历论最基本的内容. 在第三章我们介绍连分数, 希望通过这个具体例子更好地解释遍历论如何运用到其它数学分支中.

在第四章中, 我们讨论一个基本的问题：什么时候两个动力系统是“一样的”? 为此我们需要讨论如何定义动力系统的同构, 这里有几个不同的定义, 我们在本章的目的就是讨论这些定义的异同. 在本章中我们还将介绍 Kolmogorov 系统、Rohlin斜积定理等. 最后我们介绍遍历分解定理, 由此可以看出遍历性在保测系统研究中的重要性. 我们在§4.4介绍最简单的动力系统：离散谱系统. 最后我们证明 Halmos-von Neumann 定理：遍历的离散谱系统实际上是群旋转.

在第五章我们简单介绍拓扑动力系统, 给出拓扑动力系统的基本概念和结论. 拓扑动力系统和遍历论是动力系统的两个密切关联的分支. 自学科产生之初, 遍历理论和拓扑动力系统二者就有着不可分割的联系. 一方面, 拓扑动力系统可以自然地视为一个保测系统；另一方面任何遍历系统有其拓扑实现. 两套理论中有许多相对应的概念：遍历—-极小性, 离散谱—-等度连续, 测度熵—-拓扑熵 等等. 这些因素导致了这两套理论有着惊人的平行性, 二者许多结论有着极为相似的陈述, 但各自的证明方法却可能完全不同.

在第六章中我们介绍拓扑动力系统与保测系统的关联, 即主要介绍拓扑动力系统上的不变测度. 我们将看到对于任何拓扑动力系统, 总是可以在Borel σ-代数上找到不变测度使其成为一个保测系统.

第七章专门介绍熵理论的初步知识. 测度空间上的熵的概念是由 Kolmogorov 在 1958 年给出的, 之后 Adler 等人在拓扑空间上引入了拓扑熵的定义. 它是目前为止发现的一个重要的共扼不变量, 并得到广泛、深入地研究. 本章主要涉及熵的经典理论, 具体来说我们先介绍测度熵, 给出其基本性质, 而后我们给出拓扑熵的定义并研究它们的基本属性. 在此处我们会重点介绍测度熵与拓扑熵的计算, 熵的变分原理等. 之后我们将介绍测度 Pinsker σ-代数和熵的 Pinsker公式; 重新介绍测度 Kolmogorov 系统, 然后研究其基本属性并证明 Rohlin-Sinai 定理. 我们在本章最后简要介绍局部熵理论. 涉及局部熵的具体内容可以参见[207].

不交性是由 Furstenberg 在 1967 年引入的, 现在已经是遍历论的核心理论之一. 我们在第八章给出交的初步知识.

在第九章中我们重点介绍多重遍历定理以及遍历论在 Ramsey 型组合数论问题中的应用. 在本章中我们介绍多重遍历定理和多重回复定理. 首先我们从 van der Waerden 定理讲起, 给出它的动力系统证明. 之后我们介绍 Furstenberg 对应原则, 说明如何将 Ramsey 型的组合问题与动力系统关联在一起. 本章的主要目的之一是给出 Furstenberg 关于 Szemerédi 定理的动力系统证明, 为此目的我们先介绍几种特殊情况再给出 Furstenberg-Zimmer 结构定理来完成整个证明. 事实上, 我们对于这些特殊情况给出的不止是回复性质, 还证明了它们的遍历平均收敛性质, 例如 Furstenberg 弱混合多重遍历定理、Furstenberg-Sárközy定理、Roth定理等. 虽然我们不能介绍 Host-Kra 定理的证明, 但是我们将证明陶哲轩关于有限个交换变换的模多重遍历平均定理. 最后我们介绍多重遍历和多重回复研究的最新进展.

本书可以作为高年级的本科生以及动力系统方向的研究生教材, 也可以作为对动力系统感兴趣的数学工作者的参考书. 对遍历理论与拓扑动力系统的关联感兴趣的读者可以参考我们的专著《拓扑动力系统概论》([207]). 本书是作者在中国科学技术大学多年的研究和教学中逐渐发展的, 初稿在中国科学技术大学的研究生课程上多次使用, 同时也在南京大学、南京师范大学等学校的相关课程中使用. 感谢听课的同学指出初稿中的笔误, 感谢...对初稿提出的修改意见。感谢中共科学技术大学出版社的韩继伟编辑为本书的出版所付出的努力。

最后, 衷心感谢家人、朋友、国内外的同事对我们长期的支持和帮助.

作者
2023年5月
于中国科学技术大学

符号约定

- 在本书中, 我们用 $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ 分别表示自然数$\{1,2,\ldots\}$, 整数, 实数, 复数集合. 而记 $\mathbb{Z}_+$ 和 $\mathbb{Z}_-$ 为非负整数全体和非正整数全体, $\mathbb{R}_+$ 及 $\mathbb{R}_-$ 为非负实数和非正实数全体.

- 我们以 $\emptyset$ 记空集. 设 A, B 为集合 X 子集, 定义差集 $B \setminus A$ 为 $\{x \in X : x \in B, x \notin A\}$, 而定义对称差集$A\Delta B = (A \setminus B) \cup (B \setminus A)$. 记 A 的补集为A^c. 对于集合 A, 我们记它的势为 $|A|$ 或 $\mathrm{Card}A$.

- 设 X 为拓扑空间, A 为 X的子集. A 的闭包记为 $\overline{A}$ 或 $\mathrm{cl}(A)$, A 的内点集记为 $\mathring{A}$ 或 $\mathrm{int}(A)$. 如果 d 为 X 上的度量, 对 $x \in X$ 和 $\varepsilon > 0$, 令 $B_\varepsilon(x) = B(x,\varepsilon) = \{y \in X : d(y,x) < \varepsilon\}$.

- 子集 A 的直径用符号 $\mathrm{diam}A$ 来表示. 对于 X 的有限子集族 α, 它的直径 $\mathrm{diam}\ \alpha$ 是指其元素直径的最大值, 即 $\mathrm{diam}\ \alpha = \max\limits_{A\in\alpha} \mathrm{diam}A$.

- 设 $(X, \mathcal{X}, \mu)$ 为测度空间, $f = g\ a.e.$ 表示函数 f 与 g 为 μ 几乎处处一样的, 即$\mu(\{x \in X : f(x) \neq g(x)\}) = 0$. 此时我们也记之为 $f \underset{\mu}{=} g$ 或 $f = g \pmod{\mu}$.

  $A = B\ a.e.$ 或 $A \underset{\mu}{=} B$, 我们指 $\mu(A\Delta B) = 0$. $\mathcal{X}$ 的子集 $\mathcal{A} \underset{\mu}{\subset} \mathcal{B}$ 指对于任何 $A \in \mathcal{A}$ 存在 $B \in \mathcal{B}$ 使得 $\mu(A\Delta B) = 0$. 类似地定义 $\mathcal{A} \underset{\mu}{\supset} \mathcal{B}$ 和 $\mathcal{A} \underset{\mu}{=} \mathcal{B}$.

- $f = const.$ 表示 f 为常值函数.

- 我们也经常用命题 $P \pmod{\mu}$ 来表示命题 P 在忽略 μ 零测集的意义下成立.

- $\mathcal{N}_x$ 表示 x 的邻域系. 设 $A \subseteq X$, 那么用 $\mathcal{N}_A$ 表示 A 的全体邻域.

- 设 $\{X_i\}_{i\in I}$ 为一族拓扑空间, 其中 I 为指标集. 记这族空间的乘积空间为 $\prod_{i\in I} X_i$, 尤其当 $I = \mathbb{N}$ 时记为 $\prod_{i=1}^{\infty} X_i$. 如果对任意 i 都有 $X_i = X$, 那么直接记 $\prod_{i\in I} X_i$ 为 $\prod_{i\in I} X$ 或 X^I. 特别的, 空间 X的n 次乘积空间记为 $X^n = \underbrace{X \times X \times \cdots \times X}_{n次}$, 其中$n \in \mathbb{N}$.

- 设 $f : X \to X$ 为映射 $n \in \mathbb{N}$, 我们记 $f^{(n)} = \underbrace{f \times f \times \cdots \times f}_{n次}$. 对自然数 n, 我们定义 f 的 n 次迭代为 $f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n次}$, 并约定 $f^0 = \mathrm{id}$ (其中 id 表示恒同映射).

- 设 G_1, G_2为群, 如果 G_1 为 G_2 子群, 我们记为 $G_1 \leq G_2$, 如果为正规子群, 那么记为 $G_1 \lhd G_2$. 记号$\langle A\rangle$ 表示由 A 生成的群.

- “$\exists$”表示“存在”; “$\forall$”表示“对任意的”; “$s.t.$”表示“使得”; “$\Rightarrow$”表示“推出”.

- “$\square$”表示证明结束. 在文中我们会用“[[$\cdots\cdots$]]”表示对前面内容的解释.

第零章　预备知识

在本章我们简要的总结在本书中会用到的一些基本知识. 这部分内容主要涉及测度论、泛函、拓扑群理论等.

§0.1　测度与积分

§0.1.1　测度空间

定义 0.1.1 *设 X 是一个非空集合且 $\mathcal{X}$ 是由 X 的一些子集构成的集合. 如果 $\mathcal{X}$ 满足下面三个条件*

1. $X \in \mathcal{X}$;
2. *如果*$B \in \mathcal{X}$, *那么*$X \setminus B \in \mathcal{X}$;
3. *如果*$B_n \in \mathcal{X}$, $n = 1, 2, \dots$, *那么*$\bigcup_{n=1}^{\infty} B_n \in \mathcal{X}$,

则称$\mathcal{X}$*是*X*上的一个*σ**-代数**. *此时称*$(X, \mathcal{X})$*是一个***可测空间***(measurable space).*

若$B \in \mathcal{X}$, *则称*B*是一个*$\mathcal{X}$**可测集**. *在不会引起混淆的情况下, 简称*B*是可测的.*

设X是一个非空集合. 记X所有子集构成的集合为$\mathcal{P}(X)$. 则$\mathcal{P}(X)$是X上的一个σ-代数. 它是X上在包含关系下最大的σ-代数. $\mathcal{N} = \{X, \emptyset\}$ 也为σ-代数, 它是在包含关系下X上最小的σ-代数, 称为**平凡σ-代数**. $(X, \mathcal{N})$ 称为**平凡可测空间**.

任意一族σ-代数的交仍为σ-代数, 但是一族σ-代数的并不一定为σ-代数.

对于由X的一些子集构成的集合$\mathcal{A}$, 存在一个包含$\mathcal{A}$的最小的σ-代数, 记为$\sigma(\mathcal{A})$, 称为**由$\mathcal{A}$ 生成的σ-代数**. 此σ-代数是所有包含$\mathcal{A}$的σ-代数的交.

定义 0.1.2 *设*X*是一个拓扑空间. 包含*X*所有开集的最小*σ*-代数被称为*X*上的***Borel σ-代数**, *记为*$\mathcal{B}(X)$ *或者*$\mathcal{B}_X$.

设X是一个集合, $\mathcal{X}$是由X的一些子集构成的集合且$\emptyset \in \mathcal{X}$. 设$\mu\colon \mathcal{X} \to \mathbb{R} \cup \{\infty\}$是一个映射. 如果$\mu$ 满足下面两个条件

(1) $\mu(\emptyset) = 0$;

(2) 如果$A, B, A \cup B \in \mathcal{X}$ 且$A \cap B = \emptyset$, 那么$\mu(A \cup B) = \mu(A) + \mu(B)$,

则称μ是**有限可加的**. 如果进一步μ满足条件：

(3) 对$\mathcal{X}$中任意一个互不相交的序列$\{B_n\}_{n=1}^{\infty}$, 如果$\bigcup_{n=1}^{\infty} B_n \in \mathcal{X}$, 那么

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n),$$

则称μ是**可数可加的**.

定义 0.1.3 设$(X,\mathcal{X})$ 是一个可测空间. 如果$\mu\colon \mathcal{X}\to[0,\infty]$ 是一个可数可加的映射, 则称μ是$(X,\mathcal{X})$ 上的一个测度. 此时称三元组$(X,\mathcal{X},\mu)$ 是一个**测度空间**_(measure space)_.

如果$\mu(X)<\infty$, 那么称$(X,\mathcal{X},\mu)$ 是一个**有限测度空间**; 如果$\mu(X)=1$, 则称μ 是$(X,\mathcal{X})$ 上的一个概率测度, 此时称$(X,\mathcal{X},\mu)$是一个**概率空间**. 如果存在$\{A_n\}_{n\in\mathbb{N}}\subseteq\mathcal{X}$ 使得$X\subseteq\bigcup_{n=1}^{\infty}A_n$ 且$\mu(A_n)<\infty,\forall n\in\mathbb{N}$, 那么称$(X,\mathcal{X},\mu)$ 为σ **有限的**.

设$(X,\mathcal{X})$ 是一个可测空间. 如果$\mu\colon\mathcal{X}\to\mathbb{R}\cup\{\infty\}$ 是一个可数可加的映射, 则称μ是$(X,\mathcal{X})$ 上的一个**有符号测度**(signed measure), 或**带号测度**.

定理 0.1.4 (Jordan 分解定理) 设$(X,\mathcal{X})$是一个可测空间. 如果μ是$(X,\mathcal{X})$上的一个有限的有符号测度, 那么存在两个有限测度μ_1和μ_2使得μ有下面的分解:

$$\mu=\mu_1-\mu_2.$$

并且μ_1,μ_2由μ唯一决定.

例 0.1.5 1. $(X,\mathcal{N},\mu)$ 称为平凡概率空间, 其中$\mathcal{N}=\{X,\emptyset\}$, $\mu(X)=1$并且$\mu(\emptyset)=0$.

2. 设$n\in\mathbb{N}$, $X=\{0,1,\ldots,n-1\}$. 则$(X,\mathcal{P}(X),\mu_n)$是一个概率空间, 其中

$$\mu_n(\{0\})=\mu_n(\{1\})=\cdots=\mu_n(\{n-1\})=\frac{1}{n}.$$

称μ_n是$\mathbb{Z}_n$上的**等分布测度**.

3. 设X是一个非空集合. 对任意$A\subseteq X$, 定义$\mu(A)$是集合A中元素的个数. 则μ是$(X,\mathcal{P}(X))$上的一个测度, 称之为**计数测度**.

4. 设$(X,\mathcal{X})$是一个可测空间. 对于取定的一个点$x\in X$, 令$\delta_x\colon\mathcal{X}\to\{0,1\}$使得对于任意$B\in\mathcal{X}$,

$$\delta_x(B)=\begin{cases}1, & x\in B;\\ 0, & x\notin B.\end{cases}$$

那么$(X,\mathcal{X},\delta_x)$是一个概率空间. 此时称δ_x为由点x决定的**Dirac 测度**.

5. 设$\mathcal{L}([0,1])$ 为$[0,1]$ 上所有的_Lebesgue_ 可测集的全体, 记m 为$[0,1]$ 上的_Lebesgue_ 测度, 则$([0,1],\mathcal{L}([0,1]),m)$ 是一个概率空间.

6. 设$\mathbb{T}=\mathbb{R}/\mathbb{Z}$为一维环面, $\mathcal{L}(\mathbb{T})$为$\mathbb{T}$ 上所有_Lebesgue_ 可测集的全体, 记$m_{\mathbb{T}}$ 为$\mathbb{T}$ 上的_Lebesgue_ 测度, 则$(\mathbb{T},\mathcal{L}(\mathbb{T}),m_{\mathbb{T}})$ 是一个概率空间.

定义 0.1.6 设X是一个非空集合且$\mathcal{S}\subseteq\mathcal{P}(X)$. 如果$\mathcal{S}$满足下面三个条件

1. $\emptyset\in\mathcal{S}$;

2. 若$A, B \in \mathcal{S}$, 则$A \cap B \in \mathcal{S}$;

3. 若$A \in \mathcal{S}$, 则存在$n \in \mathbb{N}$和$\mathcal{S}$ 中互不相交的集合$B_1, B_2, \dots, B_n$ 使得$X \setminus A = \bigcup_{i=1}^n B_i$,

则称$\mathcal{S}$是X上的一个**半代数**(*semi-algebra*).

例如, 对于闭区间$[0,1]$, 其全体子区间为一个半代数; $\mathcal{S} = \{[0,b], (a,b], : a, b \in [0,1]\}$ 也是一个半代数.

定义 0.1.7 设X是一个非空集合且$\mathcal{A} \subseteq \mathcal{P}(X)$. 如果$\mathcal{A}$满足下面三个条件

1. $\emptyset \in \mathcal{A}$;

2. 若$A, B \in \mathcal{A}$, 则$A \cap B \in \mathcal{A}$;

3. 如果$A \in \mathcal{A}$, 则$X \setminus A \in \mathcal{A}$,

则称$\mathcal{A}$是X上的一个**代数**(*algebra*).

上面的第2条可以换为:
2': 如果$A_1, \dots, A_n \in \mathcal{A}$, 则$A_1 \cup \dots \cup A_n \in \mathcal{A}$.

命题 0.1.8 设X是一个非空集合且$\mathcal{S}$是X上的一个半代数. 则包含$\mathcal{S}$最小的代数恰由X中能被表示成$\mathcal{S}$中有限个互不相交元素并的子集组成. 即

$$\left\{C \subseteq X \colon \text{存在} n \in \mathbb{N} \text{和存在互不相交的集合} C_1, \dots, C_i \in \mathcal{S} \text{使得} C = \bigcup_{i=1}^n C_i\right\}$$

是由$\mathcal{S}$生成的代数$\mathcal{A}(\mathcal{S})$.

定理 0.1.9 设X是一个非空集合且$\mathcal{S}$是X上的一个半代数. 若$\mu \colon \mathcal{S} \to \mathbb{R}_+$ 是有限可加的, 那么存在唯一的扩充$\mu_1 : \mathcal{A}(\mathcal{S}) \to \mathbb{R}_+$ (*i.e.* $\mu_1|_{\mathcal{S}} = \mu$). 如果$\mu$ 为可数可加的, 那么μ_1 亦然.

注意有可能$X \notin \mathcal{S}$, 但是存在$\mathcal{S}$ 中互不相交的集合$B_1, B_2, \dots, B_n$ 使得$X = \bigcup_{i=1}^n B_n$. 于是如果$\sum_{i=1}^n \mu(B_n) = 1$, 那么在上面定理中我们有$\mu_1(X) = 1$.

定理 0.1.10 设X是一个非空集合且$\mathcal{A}$是X上的一个代数. 若$\mu_1 \colon \mathcal{A} \to \mathbb{R}_+$ 是可数可加的, 那么存在唯一的扩充$\mu_2 : \sigma(\mathcal{A}) \to \mathbb{R}_+$ (*i.e.* $\mu_2|_{\mathcal{A}} = \mu_1$).

综合上面定理, 我们可以从半代数的可数可加测度扩充到一个σ代数上的测度. 即

定理 0.1.11 (Carathéodory 扩充定理(Carathéodory Extension Theorem)) 设X 是一个非空集合且$\mathcal{S}$ 是X 上的一个半代数. 若$\mu \colon \mathcal{S} \to \mathbb{R}_+$ 是可数可加的, 则存在测度$\widetilde{\mu} \colon \sigma(\mathcal{S}) \to \mathbb{R}_+$ 使得$\widetilde{\mu}|_{\mathcal{S}} = \mu$, 即$\widetilde{\mu}(A) = \mu(A)$, $\forall A \in \mathcal{S}$. 另外如果$\mathcal{S}$ 中存在可数子集列$\{B_n\}_{n \in \mathbb{N}}$ 使得$X = \bigcup_{n \in \mathbb{N}} B_n$ 且$\mu(B_n) < \infty, \forall n \in \mathbb{N}$, 那么$\widetilde{\mu}$ 是唯一的.

推论 0.1.12 设X 是一个非空集合且$\mathcal{S}$是X上的一个半代数. 若$\mu\colon \mathcal{S}\to[0,1]$是可数可加的, 并且存在$n\in\mathbb{N}$ 和$\mathcal{S}$ 中互不相交的集合$B_1,B_2,\dots,B_n$ 使得$X=\bigcup_{i=1}^n B_n$和$\sum_{i=1}^n\mu(B_n)=1$, 则存在唯一的一个概率测度$\widetilde{\mu}\colon\sigma(\mathcal{S})\to[0,1]$ 满足$\widetilde{\mu}|_{\mathcal{S}}=\mu$, 即$\widetilde{\mu}(A)=\mu(A)$, $\forall A\in\mathcal{S}$.

例如, 对于闭区间$[0,1]$, $\mathcal{S}=\{[0,b],(a,b],:a,b\in[0,1]\}$ 是一个半代数. 从区间长度出发可以定义$\mathcal{S}$ 上的可数可加函数, 由上结论, 我们可以得到$[0,1]$ 上的Lebesgue测度.

定义 0.1.13 设X是一个集合且$\mathcal{M}\subseteq\mathcal{P}(X)$. 如果$\mathcal{M}$满足下面两个条件

1. 如果$A_1\subseteq A_2\subseteq\cdots$, $A_n\in\mathcal{M}$, $n\in\mathbb{N}$, 那么$\bigcup_{n=1}^\infty A_n\in\mathcal{M}$;
2. 如果$B_1\supseteq B_2\supseteq\cdots$, $B_n\in\mathcal{M}$, $n\in\mathbb{N}$, 那么$\bigcap_{n=1}^\infty B_n\in\mathcal{M}$,

则称$\mathcal{M}$是X上的一个**单调类***(monotone class)*.

定理 0.1.14 设X是一个集合且$\mathcal{A}$是X上的一个代数. 则$\sigma(\mathcal{A})$是包含$\mathcal{A}$ 的最小单调类.

下面命题表明, σ-代数中的元素总是可以由其生成代数中的元素逼近.

命题 0.1.15 (近似引理(Approximation Lemma)) 设$(X,\mathcal{X},\mu)$是一个概率空间且$\mathcal{A}\subseteq\mathcal{X}$是一个代数. 若$\sigma(\mathcal{A})=\mathcal{X}$, 则对任意$\varepsilon>0$ 和$B\in\mathcal{X}$, 存在$A\in\mathcal{A}$ 使得$\mu(A\Delta B)<\varepsilon$.

定义 0.1.16 设$\mathcal{X}$是X上的一个σ代数. 如果存在X上的一个可数子集列$\{A_i\}_{i=1}^\infty$ 使得$\mathcal{X}=\sigma(\{A_i\colon i\in\mathbb{N}\})$, 则称$\mathcal{X}$是**可数生成的**.

命题 0.1.17 设$(X,\mathcal{X},\mu)$是一个概率空间且$\mathcal{X}$是可数生成的. 则存在X上的一个可数子集列$\{A_i\}_{i=1}^\infty$ 满足: 对任意$\varepsilon>0$ 和$B\in\mathcal{X}$, 存在$i\in\mathbb{N}$使得$\mu(A_i\Delta B)<\varepsilon$.

设$(X,\mathcal{X},\mu)$为概率空间. 记

$$\overline{\mathcal{X}}=\{B\Delta F:B\in\mathcal{X},\mu(F)=0\},$$

定义$\overline{\mu}:\overline{\mathcal{X}}\to[0,1]$ 使得$\overline{\mu}(B\Delta F)=\mu(B)$. $(X,\overline{\mathcal{X}},\overline{\mu})$ 称为$(X,\mathcal{X},\mu)$的**完备化**. 如果$(X,\mathcal{X},\mu)=(X,\overline{\mathcal{X}},\overline{\mu})$, 那么称$(X,\mathcal{X},\mu)$ 为完备的.

设$(X,\mathcal{X},\mu)$ 是一个概率空间. 设$A\in\mathcal{X}$. 若$\mu(A)>0$, 令$\mathcal{X}_A=\{A\cap B\colon B\in\mathcal{X}\}$ 以及$\mu|_A$为μ 在$\mathcal{X}_A$ 上的限制(即$\mu|_A(A\cap B)=\frac{\mu(A\cap B)}{\mu(B)},\forall B\in\mathcal{X}$), 则$(A,\mathcal{X}_A,\mu|_A)$ 也是一个概率空间. 称之为$(X,\mathcal{X},\mu)$ 在A**上的限制**.

设$(X,\mathcal{X},\mu)$和$(Y,\mathcal{Y},\nu)$ 是两个概率空间. 令$\mathcal{C}=\{A\times B\colon A\in\mathcal{X},\ B\in\mathcal{Y}\}$. 则$\mathcal{C}$ 是$X\times Y$上的一个半代数. 定义$\mu\times\nu\colon\mathcal{C}\to[0,1]$, $A\times B\mapsto\mu(A)\nu(B)$. 则$\mu\times\nu$可以唯一延拓到$\sigma(\mathcal{C})$上成为一个概率测度, 称之为$\mu$ 和ν 的**乘积测度**. 记为$(X\times Y,\mathcal{X}\times\mathcal{Y},\mu\times\nu)$. 上述定义可以推广如下.

设$(X_i, \mathcal{X}_i, \mu_i)$, $i \in \mathbb{Z}$ 是一列概率空间. 令

$$X = \prod_{i=-\infty}^{\infty} X_i = \{(x_i)_{i=-\infty}^{\infty} : x_i \in X_i, i \in \mathbb{Z}\}.$$

令

$$\mathcal{S} = \left\{ \prod_{i=-\infty}^{-(n+1)} X_i \times \prod_{j=-n}^{n} A_j \times \prod_{i=n+1}^{\infty} X_i : A_j \in \mathcal{X}_j,\ -n \le j \le n,\ n \in \mathbb{Z}_+ \right\}.$$

则$\mathcal{S}$是$X = \prod_{i=-\infty}^{\infty} X_i$ 上的一个半代数, 其中元素称为**可测方体**(measurable rectangle). 记$\mathcal{X} = \bigotimes_{i=-\infty}^{\infty} \mathcal{X}_i = \sigma(\mathcal{S})$, 则$(X, \mathcal{X})$ 为可测空间.

定义

$$\mu' : \mathcal{S} \to [0,1],\quad \prod_{i=-\infty}^{-(n+1)} X_i \times \prod_{j=-n}^{n} A_j \times \prod_{i=n+1}^{\infty} X_i \mapsto \prod_{j=-n}^{n} \mu_j(A_j).$$

则μ'可以唯一延拓到$\mathcal{X}$上成为一个概率测度μ. 称$(X, \mathcal{X}, \mu)$是$\{(X_i, \mathcal{X}_i, \mu_i)\}_{i=-\infty}^{\infty}$ 的**乘积概率空间**, 记为

$$\left(\prod_{i=-\infty}^{\infty} X_i, \bigotimes_{i=-\infty}^{\infty} \mathcal{X}_i, \mu \right) = \prod_{i=-\infty}^{\infty} (X_i, \mathcal{X}_i, \mu_i).$$

乘积空间$\prod_{i=0}^{\infty}(X_i, \mathcal{X}_i, \mu_i)$, $\prod_{i=1}^{\infty}(X_i, \mathcal{X}_i, \mu_i)$ 等可以类似定义, 我们不再重复描述. 当存在$n \in \mathbb{N}$使得对任意$|i| \ge n$, $(X_i, \mathcal{X}_i, \mu_i)$ 是平凡时, 上面的乘积测度退化成有限多个测度的乘积.

下面我们考虑一种特殊情况. 设$k \ge 2$ 为自然数且记$Y = \{0, 1, \dots, k-1\}$. 设$\vec{p} = (p_0, p_1, \dots, p_{k-1})$ 是一个**概率向量**, 即$p_i \ge 0$ 且$\sum_{i=0}^{k-1} p_i = 1$. 定义$(Y, \mathcal{P}(Y))$ 上的概率测度$\nu_{\vec{p}}$使得

$$\nu_{\vec{p}}(\{i\}) = p_i,\ i = 0, 1, \dots, k-1.$$

令$(X_i, \mathcal{X}_i, \mu_i) = (Y, \mathcal{P}(Y), \nu_{\vec{p}}), 0 \le i \le k-1$. 由此得到乘积空间

$$(\Sigma_k, \mathcal{B}, \mu_{\vec{p}}) = \prod_{i=-\infty}^{\infty} (Y, \mathcal{P}(Y), \nu_{\vec{p}}).$$

我们也经常记$\Sigma_k = Y^{\mathbb{Z}} = \{0, 1, \dots, k-1\}^{\mathbb{Z}}$.

设$h \le l \in \mathbb{Z}$, $a_h, a_{h+1}, \dots, a_l \in Y$. 定义**柱形集**

$${}_h[a_h, a_{h+1}, \dots, a_l]_l = \{(x_i)_{i=-\infty}^{\infty} \in \Sigma_k : x_i = a_i, h \le i \le l\}.$$

柱形集生成了Σ_k的乘积σ代数, 其测度由下式决定:

$$\mu_{\vec{p}}({}_h[a_h, a_{h+1}, \dots, a_l]_l) = \prod_{j=h}^{l} p_{a_j}.$$

一个重要的特殊情况是等分布的时候, 即$\vec{p} = (\frac{1}{k}, \frac{1}{k}, \dots, \frac{1}{k})$. 一般的, 我们有:

定理 0.1.18 (Daniell-Kolmogorov定理) 设$k \geq 2$ 为自然数, $Y = \{0, 1, \dots, k-1\}$, 令$(\Sigma_k, \mathcal{B}) = \prod_{i=-\infty}^{\infty}(Y, \mathcal{P}(Y))$. 对于每个$n \geq 0$ 以及$a_0, a_1, \dots, a_n \in Y$, 给定了某个非负实值函数$p_n(a_0, \dots, a_n)$ 满足条件

$$\sum_{a_0 \in Y} p_0(a_0) = 1, \quad p_n(a_0, \dots, a_n) = \sum_{a_{n+1} \in Y} p_{n+1}(a_0, \dots, a_n, a_{n+1}).$$

那么在$(\Sigma_k, \mathcal{B})$ 上存在唯一的概率测度μ 使得对任意的$h \leq l$ 有

$$\mu({}_h[a_h, a_{h+1}, \dots, a_l]_l) = p_{l-h}(a_h, \dots, a_l), \forall a_i \in Y, h \leq i \leq l.$$

§0.1.2 可测函数与积分

定义 0.1.19 设$(X, \mathcal{X})$ 是一个可测空间且$f\colon X \to \mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$. 如果对任意的$c \in \mathbb{R}$, 我们有$\{x \in X\colon f(x) > c\} \in \mathcal{X}$, 则称$f$ 是一个**可测函数**.

注记 0.1.20

1. $f\colon X \to \mathbb{R}^*$ 是可测的当且仅当对任意的$c \in \mathbb{R}$, $f^{-1}((c, \infty)) \in \mathcal{X}$ 以及$f^{-1}(+\infty) \in \mathcal{X}$.
2. $f\colon X \to \mathbb{R}$ 是可测的当且仅当对任意的$B \in \mathcal{B}(\mathbb{R})$, $f^{-1}(B) \in \mathcal{X}$.
3. 设X 是一个拓扑空间. 若$f\colon X \to \mathbb{R}$ 连续, 则它是$(X, \mathcal{B}(X))$ 上的一个可测函数.

命题 0.1.21 设$f_n\colon X \to \mathbb{R}^*$, $n = 1, 2, \dots$, 是一列可测函数. 则$\inf_n f_n$, $\sup_n f_n$, $\liminf_{n\to\infty} f_n$, $\limsup_{n\to\infty} f_n$ 都是可测函数.

设$(X, \mathcal{X}, \mu)$ 是一个概率空间且P 是关于X 中元素x 的一个命题. 如果$\{x \in X\colon P(x)\text{不成立}\}$ 是可测的并且测度为0, 则称命题P 对于μ 几乎处处的点$x \in X$ 成立. 当不会引起混淆的情况下, 简称命题P **几乎处处成立**.

定义 0.1.22 设$(X, \mathcal{X}, \mu)$是一个概率空间, f, g, f_n 都为可测函数.

1. 如果$\mu(\{x \in X\colon f(x) = \infty \text{ 或 } -\infty\}) = 0$, 则称$f$是几乎处处有限的.
2. 如果存在$M > 0$ 使得$\mu(\{x \in X\colon |f(x)| \geq M\}) = 0$, 则称$f$ 是几乎处处有界的.
3. 如果$\mu(\{x \in X\colon f(x) < 0\}) = 0$, 则称$f$ 是几乎处处非负的.
4. 如果$\mu(\{x \in X\colon f(x) \neq g(x)\}) = 0$, 则称$f$ 和g 是几乎处处相等的.
5. 如果$\mu(\{x \in X\colon f_n(x) \not\to f(x), n \to \infty\}) = 0$, 则称$f_n$ 几乎处处收敛到f, 记为$f_n \to f$, μ-a.e. 或者$f_n \overset{a.e.}{\to} f, n \to \infty$.

注记 0.1.23 在测度论中, 零测集往往不起作用, 我们经常可以在忽略一个零测集上讨论问题, 甚至一个函数可以在一个零测集上没有定义.

定义 0.1.24 设$A \subseteq X$. 定义A 的**特征函数** *(characteristic function)* $\mathbf{1}_A\colon X \to \mathbb{R}$ 如下:

$$\mathbf{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

则$\mathbf{1}_A$是可测的当且仅当$A \in \mathcal{X}$. 有时候也把$\mathbf{1}_A$ 记为χ_A.

在许多文献中，特征函数也称为**指示函数** *(indicator function)*.

定义 0.1.25 设$(X, \mathcal{X}, \mu)$是一个概率空间且$f\colon X \to \mathbb{R}$是一个可测函数. 如果存在$n \in \mathbb{N}$, 可测集A_i和实数a_i, $i = 1, \ldots, n$, 使得$f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, 则称$f$是一个**简单函数**. 此时$f$的**积分**定义为

$$\int_X f d\mu = \sum_{i=1}^n a_i \mu(A_i).$$

定义 0.1.26 设$f\colon X \to \mathbb{R}^*$是一个非负可测函数. f的积分定义为

$$\int_X f d\mu = \sup\left\{ \int_X g d\mu\colon 0 \le g(x) \le f(x),\ g \text{ 是一个简单函数} \right\}.$$

注意$\displaystyle\int_X f d\mu$可以取$\infty$. 如果$\displaystyle\int_X f d\mu < \infty$, 则称$f$是**可积的**.

定义 0.1.27 设$f\colon X \to \mathbb{R}^*$是一个可测函数. 定义

$$f^+(x) = \max\{f(x), 0\}, \quad f^-(x) = \max\{-f(x), 0\}.$$

则f^+和f^-都是非负可测函数且$f(x) = f^+(x) - f^-(x)$.

如果$\int_X f^+ d\mu < \infty$ 和$\int_X f^- d\mu < \infty$, 则称f是**可积的**, 此时它的积分定义为

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$

若$A \in \mathcal{X}$, f在A上的积分定义为

$$\int_A f d\mu = \int_X f \cdot \mathbf{1}_A d\mu.$$

定理 0.1.28 (单调收敛定理) 设$f_1 \le f_2 \le \ldots$ 为概率空间$(X, \mathcal{X}, \mu)$ 上一列递增实值可积函数. 如果$\left\{ \displaystyle\int_X f_n d\mu \right\}_{n \in \mathbb{N}}$ 有界, 那么$\lim_{n\to\infty} f_n$ 几乎处处存在, 可积并有

$$\int_X (\lim_{n\to\infty} f_n) d\mu = \lim_{n\to\infty} \int_X f_n d\mu.$$

如果$\left\{ \displaystyle\int_X f_n d\mu \right\}_{n \in \mathbb{N}}$为无界的, 那么要么$\displaystyle\lim_{n\to\infty} f_n$ 在一个正测集上无限, 要么$\displaystyle\lim_{n\to\infty} f_n$不可积.

定理 0.1.29 (Fatou引理) 设$\{f_n\}_{n\in\mathbb{N}}$ 为概率空间$(X,\mathcal{X},\mu)$ 上一列可测实值函数且下界被一个可积函数控制. 如果$\liminf\limits_{n\to\infty}\int_X f_n d\mu<\infty$, 则$\liminf_{n\to\infty} f_n$ 为可积的且

$$\int_X \liminf_{n\to\infty} f_n d\mu \le \liminf_{n\to\infty}\int_X f_n d\mu.$$

定理 0.1.30 (Lebesgue控制收敛定理) 如果$g: X\to\mathbb{R}$ 为可积的且函数列$\{f_n\}$ 满足$|f_n|\le g\ a.e.\ (n\ge 1)$ 以及$\lim_{n\to\infty} f_n=f\ a.e.$, 那么$f$ 是可积的并且

$$\lim_{n\to\infty}\int_X f_n d\mu=\int_X f d\mu.$$

对于复数值函数也有类似结论, 我们不再重复论述. 设$1\le p<\infty$. 令$\mathcal{L}^p(X,\mu)$ 是满足$\int_X |f|^p d\mu<\infty$ 的所有可测函数$f\colon X\to\mathbb{C}$ 的全体. 在$\mathcal{L}^p(X,\mu)$中定义一个等价关系: $f\sim g$ 当且仅当$f=g\ \mu-a.e.$. 令

$$L^p(X,\mu)=\mathcal{L}^p(X,\mu)/\sim$$

是由等价类构成的集合. 注意到为了简便, 对于集合$L^p(X,\mu)$, 我们仍采用$\mathcal{L}^p(X,\mu)$ 上关于函数的各种运算. 在$L^p(X,\mu)$ 上定义范数

$$\|f\|_p=\left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}.$$

则$(L^p(X,\mu),\|\cdot\|_p)$ 是一个Banach 空间. $L^p(X,\mu)$ 也经常记为$L^p(X,\mathcal{X},\mu)$ 或者$L^p(\mu)$ 等, 而但需要区分不同空间的范数时，$\|f\|_p$ 也记为$\|f\|_{L^p(X,\mu)}$ 等.

当$p=\infty$ 的时候定义稍微不同. 对于可测函数f, 定义它的本性上界为

$$\|f\|_\infty=\inf\big\{M\ge 0\colon \mu(\{x\in X\colon |f(x)|>M\})=0\big\}.$$

注意空集的上确界定义为∞. 令$\mathcal{L}^\infty(X,\mu)$ 是满足$\|f\|_\infty<\infty$ 的所有可测函数$f\colon X\to\mathbb{R}^*$ 的全体. 事实上$\mathcal{L}^\infty(X,\mu)$ 是几乎处处有限的可测函数的全体. 在$\mathcal{L}^\infty(X,\mu)$ 定义一个等价关系, $f\sim g$ 当且仅当f 和g 几乎处处相等. 令$L^\infty(X,\mu)=\mathcal{L}^\infty(X,\mu)/\sim$ 是由等价类构成的集合. $(L^\infty(X,\mu),\|\cdot\|_\infty)$ 是一个Banach空间.

命题 0.1.31 设$(X,\mathcal{X},\mu)$ 是一个概率空间. 对任意的$1\le p<q\le\infty$, $L^q(X,\mu)\subseteq L^p(X,\mu)$.

定理 0.1.32 设$(X,\mathcal{X},\mu)$ 是一个概率空间且$1\le p\le\infty$. 如果$L^p(X,\mu)$ 中的序列$\{f_n\}$ 依范数收敛到f, 则存在一个子序列$\{f_{n_k}\}$ 几乎处处收敛到f.

命题 0.1.33 如果概率空间$(X,\mathcal{X},\mu)$ 是可数生成的, 则对任意的$1\le p<\infty$, $\big(L^p(X,\mu),\|\cdot\|_p\big)$ 是一个可分*Banach*空间. 特别地$L^2(X,\mu)$ 是一个可分*Hilbert* 空间, 其内积定义为

$$\langle f,g\rangle=\int_X f\cdot\bar{g}d\mu.$$

§0.2 一般拓扑

§0.2.1 拓扑空间

定义 0.2.1 设X是一个非空集合且$\mathcal{T}$是X的一个子集族. 如果$\mathcal{T}$满足下面三个性质:

1. $X, \emptyset \in \mathcal{T}$;
2. 若$\mathcal{U} \subseteq \mathcal{T}$, 则$\bigcup_{A \in \mathcal{U}} A \in \mathcal{T}$;
3. 若$U, V \in \mathcal{T}$, 则$U \cap V \in \mathcal{T}$,

则称$\mathcal{T}$ 是X 上的一个**拓扑**. 此时称$(X, \mathcal{T})$ 是一个**拓扑空间**.

在拓扑$\mathcal{T}$ 不会混淆的情况下, 直接称X 是一个拓扑空间. 设$(X, \mathcal{T})$ 是一个拓扑空间且$A \subseteq X$. 若$A \in \mathcal{T}$, 则称A 是一个**开集**. 若$X \setminus A$ 是一个开集, 则称A 是一个**闭集**. 由定义易见X 和$\emptyset$ 既是开集又是闭集.

定义 0.2.2 设X是一个非空集合.

1. 令$\mathcal{T} = \{X, \emptyset\}$. 则$\mathcal{T}$是$X$上的一个拓扑, 称之为$X$上的**平凡拓扑**.
2. 令$\mathcal{P}(X)$是由X的所有子集构成的集族. 则$\mathcal{P}(X)$是X上的一个拓扑, 称之为X上的**离散拓扑**.

定义 0.2.3 设$(X, \mathcal{T})$是一个拓扑空间且$Y \subseteq X$. 则$\mathcal{T}|_Y = \{U \cap Y : U \in \mathcal{T}\}$是$Y$上的一个拓扑. 此时称$(Y, \mathcal{T}|_Y)$是拓扑空间$(X, \mathcal{T})$的一个子空间.

定义 0.2.4 设$(X, \mathcal{T})$是一个拓扑空间且$\mathcal{A} \subseteq \mathcal{T}$. 如果对任意的$U \in \mathcal{T}$,

$$U = \bigcup \{A \in \mathcal{A} : A \subseteq U\},$$

则称$\mathcal{A}$是$\mathcal{T}$的一个**拓扑基**.

设$\mathcal{C} \subseteq \mathcal{T}$. 如果

$$\{C_1 \cap C_2 \cap \cdots \cap C_n : C_i \in \mathcal{C},\ i = 1, 2, \ldots, n,\ n \in \mathbb{N}\} \cup \{X\}$$

是$\mathcal{T}$的一个拓扑基, 则称$\mathcal{C}$是$\mathcal{T}$的一个**拓扑子基**.

如果拓扑空间$(X, \mathcal{T})$存在一个可数的拓扑基, 则称它是**第二可数的**.

命题 0.2.5 设X是一个非空集合且$\mathcal{C}$是X上的一个子集族. 则存在X上的唯一一个拓扑$\mathcal{T}$使得$\mathcal{C}$是$\mathcal{T}$的一个拓扑子基.

定义 0.2.6 设$(X,\mathcal{T})$是一个拓扑空间, $x\in X$且$U\subseteq X$. 如果存在$V\in\mathcal{T}$ 使得$x\in V\subseteq U$, 则称U是点x的一个**邻域**. 点x的所有邻域构成的X的子集族称为x的**邻域系**. 如果U是包含x的开集, 则U是点x的一个邻域, 此时称U是点x的一个开邻域.

设$\mathcal{N}_x$是点x的的一个邻域族. 如果对点x的每个邻域U, 存在$V\in\mathcal{N}_x$使得$V\subseteq U$, 则称$\mathcal{N}_x$是点x的**邻域系的一个基**.

定义 0.2.7 设$(X,\mathcal{T})$是一个拓扑空间, $x\in X$和$A\subseteq X$.

1. 如果对x的任意一个邻域U, $U\cap(A\setminus\{x\})\neq\emptyset$, 则称$x$是$A$的一个**聚点**或**极限点**. 集合$A$的所有聚点构成的集合称为$A$的导集, 记为$A'$.
2. 如果存在x的一个邻域U使得$U\subseteq A$, 则称x是A的一个**内点**. 集合A的所有内点构成的集合称为A的内部, 记为$\mathrm{int}(A)$ 或$\mathring{A}$.
3. 如果$x\in A$且存在x的一个邻域U使得$U\cap A=\{x\}$, 则称x是A的一个**孤立点**.
4. 如果对x的任意一个邻域U, 有$U\cap A\neq\emptyset$以及$U\cap(X\setminus A)\neq\emptyset$, 则称$x$是$A$的一个**边界点**. 集合$A$的所有边界点构成的集合称为$A$的边界, 记为$\partial(A)$ 或$Bd(A)$.
5. 称集合$A\cup A'$是A的闭包, 记为$\overline{A}$ 或$\mathrm{cl}(A)$.

命题 0.2.8 设$(X,\mathcal{T})$是一个拓扑空间和$A\subseteq X$.

1. A' 是一个闭集.
2. $\mathrm{int}(A)$是包含于A的最大开集.
3. $\mathrm{cl}(A)$是包含A的最小闭集.
4. A是闭集当且仅当$A=\mathrm{cl}(A)$.
5. $\mathrm{cl}(A)=\mathrm{int}(A)\cup\partial(A)$.

定义 0.2.9 设$(X,\mathcal{T})$是一个拓扑空间和$A\subseteq X$. 如果$\mathrm{cl}(A)=X$, 则称A在X中稠密. 若$(X,\mathcal{T})$有一个至多可数的稠密子集, 则称它为**可分的**.

定义 0.2.10 设$(X,\mathcal{T})$是一个拓扑空间. 如果对X中任意互异的两点x和y, 存在x的一个邻域U和y的一个邻域V 使得$U\cap V=\emptyset$, 则称$(X,\mathcal{T})$是一个**Hausdorff** 空间, 或者称$(X,\mathcal{T})$为T_2的.

设$(X,\mathcal{T})$是一个拓扑空间且$G\subseteq X$. 如果存在X中一列开集$\{U_n\}_{n=1}^{\infty}$使得$\bigcap_{n=1}^{\infty}U_n=G$, 则称$G$是$X$的一个$G_\delta$**子集**.

定义 0.2.11 设$(X,\mathcal{T})$是一个拓扑空间.

1. 设$A \subseteq X$. 如果对X中任意一个非空开集U, 存在非空开集V使得$V \subseteq U$ 以及$V \cap A = \emptyset$, 则称A在X中是**无处稠密的***(nowhere dense)*.

2. 设$B \subseteq X$. 如果存在X中一列无处稠密集$\{A_n\}_{n=1}^{\infty}$使得$B = \bigcup_{n=1}^{\infty} A_n$, 则称$B$是一个**第一纲集***(a set of first category)*或**疏朗集***(a meager set)*.

3. 设$C \subseteq X$. 如果C不是一个第一纲集, 则称C是一个**第二纲集***(a set of second category)*.

4. 设$D \subseteq X$. 如果$X \setminus D$是一个第一纲集, 则称D是一个**剩余集***(a residual set or a comeagre set)*.

定理 0.2.12 设X 为拓扑空间, 那么以下等价:

1. X 中每个非空开集为第二纲集;

2. X 的每个剩余集为稠密的;

3. X 中可数个稠密开集的交仍为稠密的.

满足上面条件的空间称为**Baire 空间**.

定理 0.2.13 每个完备度量空间为*Baire*空间; 每个局部紧*Hausdorff*空间为*Baire*空间.

设X是一个集合且$\mathcal{U}$是X中一些子集构成的集合. 如果$\bigcup_{A\in\mathcal{U}} A = X$, 则称$\mathcal{U}$是$X$的一个**覆盖**. 如果$\mathcal{V} \subseteq \mathcal{U}$且$\mathcal{V}$也是$X$的一个覆盖, 则称$\mathcal{V}$是$\mathcal{U}$的一个**子覆盖**.

设X是一个拓扑空间且$\mathcal{U}$是X的一个覆盖. 如果$\mathcal{U}$中的每个元素都是X中的开集, 则称它是一个**开覆盖**. 如果覆盖个数有限, 那么称之为**有限覆盖**.

定义 0.2.14 如果拓扑空间X中的任何开覆盖都有有限子覆盖, 则称X是**紧致的***(compact)*.

设$\mathcal{A}$是X上的一个子集族. 如果$\mathcal{A}$的每个有限子集族都有非空的交, 则称$\mathcal{A}$具有**有限交性质**.

命题 0.2.15 设X是一个拓扑空间. 则X是紧致的当且仅当X中的每个具有有限交性质的闭集族都有非空的交.

关于紧致性一个常用的性质是Lebesgue 覆盖引理:

定理 0.2.16 (Lebesgue 覆盖引理) 设(X, d) 为一个紧致度量空间, α 为开覆盖. 那么存在$\delta > 0$ 使得任何直径小于等于δ 的子集都包含在α 的某个元素中. 此δ 称为覆盖α 的***Lebesgue* 数**.

定义 0.2.17 设$(X,\mathcal{T})$ 和$(Y,\mathcal{S})$ 是两个拓扑空间且$f\colon X\to Y$ 是一个映射. 设$x\in X$. 如果对任意Y 中$f(x)$ 的邻域V, $f^{-1}(V)$ 是x 的一个邻域, 则称x 是f 的一个连续点, 或f 在点x处**连续**.

如果f 在X 中的任何点处都连续, 则称f 是一个**连续映射**. 如果f 是一个连续的一一映射并且它的逆也连续, 则称f 是一个**同胚**.

命题 0.2.18 设$(X,\mathcal{T})$ 和$(Y,\mathcal{S})$ 是两个拓扑空间且$f\colon X\to Y$ 是一个映射. 则下列命题等价:

1. f 连续;
2. $f^{-1}(\mathcal{S})\subseteq\mathcal{T}$, 即对$Y$ 中任意一个开集V, $f^{-1}(V)$ 是X 中的一个开集;
3. 对Y 中任意一个闭集A, $f^{-1}(A)$ 是X 中的一个闭集;
4. 对X 中任意一个子集B, $f(\mathrm{cl}(B))\subseteq \mathrm{cl}(f(B))$;
5. 对Y 中任意一个子集C, $f^{-1}(\mathrm{cl}(C))\supseteq \mathrm{cl}(f^{-1}(C))$.

定义 0.2.19 设$(X,\mathcal{T})$是一个拓扑空间且$A\subseteq X$. 如果A作为子空间是紧致的, 则称A是X的一个**紧致子集**. 如果X中的每个点都有一个紧致的邻域, 则称X是**局部紧的**.

定义 0.2.20 设X和Y是两个拓扑空间且$f\colon X\to Y$是一个映射. 如果f是一个单射且f是从X到$f(X)$的一个同胚, 则称f是一个**嵌入映射**. 易见存在一个X到Y的嵌入映射当且仅当X同胚于Y的一个子空间.

设$(X_1,\mathcal{T}_1)$和$(X_2,\mathcal{T}_2)$是两个拓扑空间. 则$X_1\times X_2$上有唯一一个以$\{U_1\times U_2\colon U_1\in\mathcal{T}_1, U_2\in\mathcal{T}_2\}$ 为基的拓扑$\mathcal{T}$. 此时称$(X_1\times X_2,\mathcal{T})$是$(X_1,\mathcal{T}_1)$和$(X_2,\mathcal{T}_2)$的**乘积拓扑空间**.

一般地, 设$\{(X_i,\mathcal{T}_i)\}_{i\in I}$是一族拓扑空间. 则$\prod_{i\in I}X_i$上有唯一一个以

$$\{p_i^{-1}(U_i)\colon U_i\in\mathcal{T}_i,\ i\in I\}$$

为子基的拓扑$\mathcal{T}$, 其中$p_i\colon\prod_{i\in I}X_i\to X_i$ 是到第i 个坐标的投射. 此时称

$$(\prod_{i\in I}X_i,\mathcal{T})$$

是$\{(X_i,\mathcal{T}_i)\}_{i\in I}$的**乘积拓扑空间**, $\mathcal{T}$为**乘积拓扑**.

当$I=\mathbb{Z}$时, 令

$$\mathcal{S}=\Big\{\prod_{i=-\infty}^{-(n+1)}X_i\times\prod_{j=-n}^{n}U_j\times\prod_{i=n+1}^{\infty}X_i : U_j\in\mathcal{T}_j, -n\le j\le n,\ n\in\mathbb{Z}_+\Big\}.$$

则$\mathcal{S}$是$X=\prod_{i=-\infty}^{\infty}X_i$ 上的一个基.

定理 0.2.21 (Tychonoff) 设$\{(X_i,\mathcal{T}_i)\}_{i\in I}$是一族紧致*Hausdorff*空间. 则$(\prod_{i\in I}X_i,\mathcal{T})$是一个紧致*Hausdorff*空间.

定义 0.2.22 设$(X,\mathcal{T})$是一个拓扑空间, Y是一个集合且$f\colon X\to Y$ 是一个满射. 则$\mathcal{S}=\{U\subseteq Y\colon f^{-1}(U)\in\mathcal{T}\}$ 是Y上的一个拓扑. 此时称$\mathcal{S}$是相对于f的**商拓扑**.

设R是X上的一个等价关系. 商集X/R相对于商映射$p\colon X\to X/R$, $x\mapsto[x]_R$ 的商拓扑为$\mathcal{T}_R$. 称$(X/R,\mathcal{T}_R)$是相对于R的商空间.

定理 0.2.23 设X是一个紧致*Hausdorff*空间且R是X上的一个闭等价关系. 则商空间X/R也是一个紧致*Hausdorff*空间.

定义 0.2.24 设$(X,\mathcal{T})$是一个拓扑空间. 如果X中存在一个非空的真子集是既开又闭的, 则称X是不连通, 否则称X是**连通的**. 设$A\subseteq X$. 如果A作为子空间是连通的, 则称A是X的一个**连通子集**.

命题 0.2.25 设$(X,\mathcal{T})$是一个拓扑空间.

1. 若$A\subseteq X$是连通的, 则$\mathrm{cl}(A)$也是连通的.
2. 设$A,B\subseteq X$都是连通的且$A\cap B\neq\emptyset$, 则$A\cup B$也是连通的.

命题 0.2.26 设X和Y是两个*Hausdorff*空间且$f\colon X\to Y$是一个连续满射.

1. 如果X是紧致的, 则Y也是紧致的.
2. 如果X是连通的, 则Y也是连通的.
3. 如果X是紧致的并且f是一个一一映射, 则f是一个同胚.

§0.2.2　Cantor集

定义 0.2.27 设$(X,\mathcal{T})$是一个拓扑空间且$x\in X$. 定义x的**连通分支**是包含x的所有连通子集的并. 易见x的连通分支是包含x的最大连通子集并且是闭的. X的所有连通分支构成了X上的一个闭剖分, 从而确定X上的一个连通关系.

定义 0.2.28 设$(X,\mathcal{T})$是一个拓扑空间. 如果每个点的连通分支是单点集, 则称$(X,\mathcal{T})$是**完全不连通的**.

设$(X,\mathcal{T})$是一个*Hausdorff*空间. 如果任意的$x\in X$和x的任意一个邻域U, 存在一个既开又闭的集V使得$x\in V\subseteq U$, 则称$(X,\mathcal{T})$是一个0**维空间**.

定理 0.2.29 设X是一个紧致*Hausdorff*空间. 则X是完全不连通的当且仅当它是一个0维空间.

定理 0.2.30 设X是一个紧致*Hausdorff*空间且R是X上的连通关系. 则商空间X/R是一个0维紧致*Hausdorff*空间.

定义 0.2.31 设$C_0=[0,1]$. 对任意的$n\in\mathbb{N}$, 令

$$C_n=\frac{C_{n-1}}{3}\cup\left(\frac{2}{3}+\frac{C_{n-1}}{3}\right).$$

标准Cantor三分集定义为

$$C=\bigcap_{n=0}^{\infty}C_n.$$

定义 0.2.32 设X是一个拓扑空间且$A\subseteq X$. 如果A作为子空间是紧致可度量化, 0-维和无孤立点的, 则称A是X中的一个**Cantor集**.

定理 0.2.33 (Brouwer) 每个*Cantor*集合都同胚于标准三分*Cantor*集C.

定理 0.2.34 (Alexandroff) 每个紧致可度量化空间是C的连续像.

定理 0.2.35 设(X,d)是一个完备可分度量空间. 若$A\subseteq X$是一个不可数的*Borel*集, 则A包含一个*Cantor*集.

命题 0.2.36 设$A=\{0,1,\dots,k-1\}$ $(k\geq 2)$, 并赋予其离散拓扑. 则乘积空间$A^{\mathbb{Z}}$ 同胚于标准*Cantor*三分集.

§0.2.3 函数空间

定义 0.2.37 设X是一个拓扑空间和I是一个集合. 从I到X的所有映射构成的集合记作X^I. 由笛卡尔积的定义, X^I是集族$\{X\}_{i\in I}$的笛卡尔积$\prod_{i\in I}X$, 因此它有乘积拓扑. 习惯上称这个拓扑为X^I的**点态收敛拓扑**.

设X和Y是两个拓扑空间. 记$C(X,Y)$是从X到Y的所有连续映射构成的集合. 因此$C(X,Y)\subseteq Y^X$. $C(X,Y)$作为Y^X的子空间称为从X到Y的具有**点态收敛拓扑**的连续映射空间, 此时也记$C(X,Y)$ 为$C_p(X,Y)$.

定义 0.2.38 设$(X,\mathcal{T})$和$(Y,\mathcal{S})$是两个拓扑空间. 记$\mathcal{C}$是X的全体紧子集构成的集族. 则Y^X上有唯一一个以

$$\{W(E,U)\subseteq Y^X\colon E\in\mathcal{C},U\in\mathcal{S}\}$$

为子基的拓扑$\mathcal{T}_{\mathcal{C}}$, 其中

$$W(E,U)=\{f\in Y^X\colon f(E)\subseteq U\}.$$

称$\mathcal{T}_{\mathcal{C}}$是$Y^X$上的**紧开拓扑**.

$C(X,Y)$作为$(Y^X,\mathcal{T}_{\mathcal{C}})$的子空间称为从$X$到$Y$的具有**紧开拓扑**的连续映射空间, 此时也记$C(X,Y)$ 为$C_c(X,Y)$.

定义 0.2.39 设X是一个集合且(Y,d)是一个度量空间. 定义$d_u\colon Y^X\times Y^X\to\mathbb{R}$ 使得对任意的$f,g\in Y^X$,

$$d_u(f,g)=\begin{cases}1,\ \text{存在}x\in X\text{使得}d(f(x),f(x))\ge 1,\\ \sup\{d(f(x),f(x))\colon x\in X\},\ \text{其它情况}.\end{cases}$$

则d_u是Y^X上的一个度量, 称之为Y^X上的一致收敛度量. 由一致收敛度量d_u诱导的拓扑称为Y^X上的**一致收敛拓扑**$\mathcal{T}_u$.

设X是一个拓扑空间且(Y,d)是一个度量空间. $C(X,Y)$作为$(Y^X,\mathcal{T}_u)$的子空间称为从X到Y的具有**一致收敛拓扑**的连续映射空间, 此时也记$C(X,Y)$ 为$C_u(X,Y)$.

若X是紧致的, 则Y^X上的紧开拓扑和一致收敛拓扑相等.

设X是一个紧致度量空间. 记$C(X)$ 或$C(X,\mathbb{C})$ 为X 上全体复值连续函数空间, 而记$C(X,\mathbb{R})$ 是X 上实值连续函数空间. 对任意$f\in C(X)$, 定义

$$\|f\|_{\sup}=\sup\{|f(x)|\colon x\in X\}.$$

则$\|\cdot\|_{\sup}$ 是一个范数并且$(C(X),\|\cdot\|_{\sup})$ 是一个Banach 空间. 由范数$\|\cdot\|_{\sup}$诱导的拓扑与一致收敛拓扑相等.

对任意$f,g\in C(X)$, 定义f 与g 的乘积为$f\cdot g(x)=f(x)g(x)$, $\forall\ x\in X$. 则$f\cdot g\in C(X)$. 这个乘法下有单位元$\mathbf{e}$ ($\mathbf{e}(x)\equiv 1,\forall\ x\in X$). 故$(C(X),\|\cdot\|_{\sup})$ 是一个Banach 代数, 且$(C(X,\mathbb{R}),\|\cdot\|_{\sup})$ 是它的一个子Banach 代数.

设$\mathcal{A}\subseteq C(X)$. 定义$\mathcal{A}$ 线性张成的子空间为

$$\operatorname{span}(\mathcal{A})=\left\{\sum_{i=1}^{n}a_if_i\colon a_i\in\mathbb{C},f_i\in\mathcal{A},n\in\mathbb{N}\right\}.$$

定义 0.2.40 设X 为拓扑空间, $\mathcal{A}\subseteq C(X,\mathbb{R})$. 如果

1. $\mathbf{e}\in\mathcal{A}$,
2. 对任意$f\in\mathcal{A}$和$c\in\mathbb{R}$, $cf\in\mathcal{A}$,
3. 对任意$f,g\in\mathcal{A}$, $f+g$, $f\cdot g\in\mathcal{A}$,

则称$\mathcal{A}$是一个**子代数**. 将$\mathbb{R}$ 换为$\mathbb{C}$ 便得到$C(X)$ 子代数的定义. 如果对任意互异的$x,y\in X$, 存在$f\in\mathcal{A}$使得$f(x)\neq f(y)$, 则称$\mathcal{A}$**分离点**.

定理 0.2.41 (Stone – Weierstrass) 设X是一个紧致可度量化空间且$\mathcal{A}\subseteq C(X,\mathbb{R})$是一个分离点的子代数, 则$\mathcal{A}$在$C(X,\mathbb{R})$中稠密.

下面是Stone – Weierstrass定理在复值函数空间上的版本.

定理 0.2.42 设X是一个紧致可度量化空间. 设$\mathcal{A}\subseteq C(X)$ 满足

1. $\mathbf{e} \in \mathcal{A}$;
2. 对任意$f \in \mathcal{A}$和$a, b \in \mathbb{R}$, 有$(a+bi)f \in \mathcal{A}$以及$\bar{f} \in \mathcal{A}$, 其中i是虚数单位, $\bar{f}$表示f的共轭;
3. 对任意$f, g \in \mathcal{A}$, 有$f+g$, $f \cdot g \in \mathcal{A}$;
4. 对任意互异的$x, y \in X$, 存在$f \in \mathcal{A}$使得$f(x) \neq f(y)$.

则$\mathcal{A}$在$C(X)$中稠密.

定义 0.2.43 设(X,d) 是一个紧致度量空间且$\mathcal{A} \subseteq C(X)$.

1. 如果存在$M > 0$ 使得对任意$f \in \mathcal{A}$ 和$x \in X$, $|f(x)| \le M$, 则称$\mathcal{A}$是**一致有界的**.
2. 如果对任意$\varepsilon > 0$, 存在$\delta > 0$使得对任意的$f \in \mathcal{A}$和满足$d(x,y) < \delta$的$x, y \in X$,
$$|f(x) - f(y)| < \varepsilon,$$
则称$\mathcal{A}$是**等度连续的**.

下面是Arzelá - Ascoli定理的经典形式.

定理 0.2.44 (Arzelá - Ascoli) 设(X,d) 是一个紧致度量空间且$\mathcal{A} \subseteq C(X)$. 则$\mathcal{A}$ 在$C(X)$中的闭包是紧的当且仅当$\mathcal{A}$ 是一致有界和等度连续的.

下面是Arzelá - Ascoli 定理的一般形式.

定理 0.2.45 (Arzelá - Ascoli) 设X是一个紧致*Hausdorff*空间且Y是一个度量空间. 则一个闭子集$\mathcal{A} \subseteq C(X,Y)$在紧开拓扑下是紧致的当且仅当它是等度连续和逐点相对紧, 即对任意的$x \in X$, $\{f(x) \colon f \in \mathcal{A}\}$在$Y$中的闭包是紧的.

推论 0.2.46 设X是一个紧致度量空间和$\mathcal{A} \subseteq C(X,X)$. 则$\mathcal{A}$在$C(X,X)$上的一致收敛拓扑下的闭包是紧的当且仅当$\mathcal{A}$是等度连续的.

一个关于常见的度量化定理为:

定理 0.2.47 设X为紧致*Hausdorff*空间, 那么以下命题等价:

1. X 可度量;
2. X有可数基;
3. $C(X)$ 有可数稠密子集.

§0.3　条件期望与测度分解

§0.3.1　Radon-Nikodym导数

定义 0.3.1 *设μ 和ν 是可测空间$(X,\mathcal{X})$ 上的两个概率测度.*

1. *如果对任意$A\in\mathcal{X}$, $\mu(A)=0$ 蕴含$\nu(A)=0$, 则称ν 相对于μ 是***绝对连续的***, 记为$\nu\ll\mu$.*

2. *若$\nu\ll\mu$ 和$\mu\ll\nu$, 则称μ 和ν 是等价的, 记为$\mu\sim\nu$.*

3. *如果存在$\mathcal{X}$ 中互不相交的集合A 和B 使得$X=A\cup B$ 并且$\mu(A)=\nu(B)=0$, 则称μ 和ν 是相互***奇异***的, 记为$\mu\perp\nu$.*

定理 0.3.2 (Lebesgue 分解定理) *设μ,m 为$(X,\mathcal{X})$ 上两个概率测度, 则存在$p\in[0,1]$ 以及概率测度μ_1,μ_2 使得$\mu=p\mu_1+(1-p)\mu_2$ 且$\mu_1\ll m$, $\mu_2\perp m$. 其中数p 以及测度μ_1,μ_2 是由μ,m 唯一决定的.*

定理 0.3.3 (Radon-Nikodym定理) *设μ 和ν 是可测空间$(X,\mathcal{X})$ 上的两个概率测度. 如果$\nu\ll\mu$, 则存在$(X,\mathcal{X})$ 上唯一的非负可积函数f 使得对任意的$A\in\mathcal{X}$,*

$$\nu(A)=\int_A f d\mu.$$

*此时称f 是ν 相对于μ 的***Radon-Nikodym导数***, 记为$\dfrac{d\nu}{d\mu}$.*

Radon–Nikodym导数满足链式法则：如果$\eta\ll\nu\ll\mu$, 那么

$$\frac{d\eta}{d\mu}=\frac{d\eta}{d\nu}\cdot\frac{d\nu}{d\mu}.$$

定义 0.3.4 *设$(X,\mathcal{X},m)$为概率空间, $A\in\mathcal{X}$ 且$\mu(A)>0$. 定义*

$$\mu_A\colon\mathcal{X}\to[0,1],\quad B\mapsto\frac{\mu(A\cap B)}{\mu(A)}.$$

*则μ_A 为概率测度, 称μ_A是μ在A上的***条件概率测度***.*

注记 0.3.5　*1. $\mu_A\ll\mu$且$\frac{d\mu_A}{d\mu}=\frac{1}{\mu(A)}\chi_A$*

2. $\mu=\mu(A)\mu_A+(1-\mu(A))\mu_{X\setminus A}$.

§0.3.2 条件期望

条件期望有许多种定义方式, 我们下面首先运用Radon-Nikodym定理定义条件期望, 然后用另一种更为直观的方式给出不同的定义方法.

设$(X,\mathcal{X},\mu)$ 为概率空间, 且$\mathcal{A}$ 为$\mathcal{X}$ 的子σ代数. 下面我们定义**条件期望**算子

$$\mathbb{E}(\cdot|\mathcal{A}):L^1(X,\mathcal{X},\mu)\to L^1(X,\mathcal{A},\mu).$$

如果$f\equiv 0$, 则定义$\mathbb{E}(f|\mathcal{A})\equiv 0$. 如果$f\in L^1(X,\mathcal{X},\mu)$为非负实值函数且$a=\int_X fd\mu>0$, 则

$$\mu_f(C)=a^{-1}\int_C fd\mu,\ \forall\ C\in\mathcal{A}$$

定义了$(X,\mathcal{A},\mu)$ 上的一个概率测度, 并且μ_f 相对于μ 为绝对连续的. 由Radon-Nikodym 定理, 存在函数$\mathbb{E}(f|\mathcal{A})\in L^1(X,\mathcal{A},\mu)$ 使得$\mathbb{E}(f|\mathcal{A})\geq 0$ 且

$$\int_C\mathbb{E}(f|\mathcal{A})d\mu=\int_C fd\mu,\ \ \forall C\in\mathcal{A}.$$

并且$\mathbb{E}(f|\mathcal{A})$ 在几乎处处意义下是唯一的. 如果f 为任意的实值函数, 则分别考虑其正值与负值部分且线性地定义$\mathbb{E}(f|\mathcal{A})$. 对于$f$ 为复值函数的时候, 将它按实部与虚部类似处理之. 这样, 对任意$f\in L^1(X,\mathcal{X},\mu)$, 我们找到了唯一的一个$\mathcal{A}$ 可测函数$\mathbb{E}(f|\mathcal{A})\in L^1(X,\mathcal{A},\mu)$ 满足

$$\int_C\mathbb{E}(f|\mathcal{A})d\mu=\int_C fd\mu,\ \forall\ C\in\mathcal{A}.$$

如此我们就定义了条件期望.

下面我们介绍第二种定义条件期望的方法. 因为$L^2(X,\mathcal{A},\mu)$ 为$L^2(X,\mathcal{X},\mu)$ 的闭子空间, 所以存在投影算子

$$P:L^2(X,\mathcal{X},\mu)\to L^2(X,\mathcal{A},\mu)$$

具有性质: 对于任意$f\in L^2(X,\mathcal{X},\mu)$

$$\int_A fd\mu=\int_X\mathbf{1}_A fd\mu=\int_X\mathbf{1}_A Pfd\mu=\int_A Pfd\mu,\ \forall A\in\mathcal{A}. \tag{0.3.1}$$

下面我们将$P:L^2(X,\mathcal{X},\mu)\to L^2(X,\mathcal{A},\mu)$ 扩充到$L^1(X,\mathcal{X},\mu)$, 这个扩充就定义为条件期望

$$\mathbb{E}(\,\cdot\,|\mathcal{A}):\ L^1(X,\mathcal{X},\mu)\to L^1(X,\mathcal{A},\mu).$$

注意到, $L^2(X,\mathcal{X},\mu)\subseteq L^1(X,\mathcal{X},\mu)$ 为稠密的, 并且对于$f\in L^2(X,\mathcal{X},\mu)$, 有

$$\{x\in X:Pf(x)>0\}\in\mathcal{A},\ \{x\in X:Pf(x)<0\}\in\mathcal{A}.$$

于是根据(0.3.1),

$$\|Pf\|_1\leq\|f\|_1.$$

对于复数值函数, 取实部虚部同样讨论, 就得到

$$\|Pf\|_1 \leq 2\|f\|_1.$$

于是我们可以将$P: L^2(X,\mathcal{X},\mu) \to L^2(X,\mathcal{A},\mu)$ 扩充到$L^1(X,\mathcal{X},\mu)$,

$$\mathbb{E}(\,\cdot\,|\mathcal{A})\colon L^1(X,\mathcal{X},\mu) \to L^1(X,\mathcal{A},\mu),$$

并且(0.3.1) 保持.

下面我们根据第一种方式给出具体性质.

定理 0.3.6 设$(X,\mathcal{X},\mu)$为概率空间, $\mathcal{A}$为$\mathcal{X}$的子σ代数. 则存在一个条件期望映射

$$\mathbb{E}(\,\cdot\,|\mathcal{A})\colon L^1(X,\mathcal{X},\mu) \to L^1(X,\mathcal{A},\mu),$$

满足下面的性质:

1. 对任意的$f \in L^1(X,\mathcal{X},\mu)$, $\mathbb{E}(f|\mathcal{A})$几乎处处由下面两个性质刻画
   (a) $\mathbb{E}(f|\mathcal{A})$是$\mathcal{A}$-可测的;
   (b) 对任意的$A \in \mathcal{A}$, $\int_A \mathbb{E}(f|\mathcal{A})d\mu = \int_A f d\mu$.
2. $\mathbb{E}(\,\cdot\,|\mathcal{A})$是一个具有范数为1的线性算子, 并且$\mathbb{E}(\,\cdot\,|\mathcal{A})$是一个正算子, 即若$f \in L^1(X,\mathcal{X},\mu)$几乎处处非负, 则$\mathbb{E}(f|\mathcal{A})$也是几乎处处非负的.
3. 对任意$f \in L^1(X,\mathcal{X},\mu)$和$g \in L^\infty(X,\mathcal{A},\mu)$, $\mathbb{E}(gf|\mathcal{A}) = g\mathbb{E}(f|\mathcal{A})$ 几乎处处成立.
4. 如果$\mathcal{C} \subseteq \mathcal{A}$是一个子$\sigma$-代数, 则$\mathbb{E}(\mathbb{E}(f|\mathcal{A})|\mathcal{C}) = \mathbb{E}(f|\mathcal{C})$ 几乎处处成立.
5. 如果$f \in L^1(X,\mathcal{A},\mu)$, 则$\mathbb{E}(f|\mathcal{A}) = f$ 几乎处处成立.
6. 对任意$f \in L^1(X,\mathcal{X},\mu)$, $|\mathbb{E}(f|\mathcal{A})| \leq \mathbb{E}(|f||\mathcal{A})$ 几乎处处成立.

证明. (1) 设$f \in L^1(X,\mathcal{X},\mu)$. 首先假设$f \geq 0$且$\int_X f d\mu > 0$. 定义

$$\mu_f\colon \mathcal{X} \to [0,1], \quad B \mapsto \frac{\int_B f d\mu}{\int_X f d\mu}.$$

则μ_f 为概率测度. 注意到$\mu_f|_\mathcal{A} \ll \mu|_\mathcal{A}$. 由Radon–Nikodym导数定理(0.3.3), 存在唯一的一个$g \in L^1(X,\mathcal{A},\mu)$使得对任意的$A \in \mathcal{A}$,

$$\int_A g d\mu = \mu_f(A) = \frac{\int_B f d\mu}{\int_X f d\mu}.$$

令$\mathbb{E}(f|\mathcal{A}) = g \cdot \int_X f d\mu$. 当$f = 0$时, 令$\mathbb{E}(f|\mathcal{A}) = 0$. 一般的, 若$f = f^+ - f^-$, 则令$\mathbb{E}(f|\mathcal{A}) = \mathbb{E}(f^+|\mathcal{A}) - \mathbb{E}(f^-|\mathcal{A})$; 对于复值函数, 考虑实部与虚部. 这就证明了存在性.

设g_1和g_2都满足(a)和(b). 令$A=\{x\in X\colon g_1(x)<g_2(x)\}$. 则$A\in\mathcal{A}$. 根据(b)

$$\int_A g_1 d\mu=\int_A f d\mu=\int_A g_2 d\mu,$$

故$\mu(A)=0$. 类似地可以证明$\mu(\{x\in X\colon g_1(x)>g_2(x)\})=0$. 所以$g_1$与$g_2$几乎处处相等.

(2) 由于$\mathbb{E}(f|\mathcal{A})$由(a)和(b)决定, 易见$\mathbb{E}(\cdot|\mathcal{A})$是一个有界线性算子并且范数为1.

设$f\in L^1(X,\mathcal{X},\mu)$且$f\geq 0$. 令$A=\{x\in X\colon \mathbb{E}(f|\mathcal{A})<0\}$. 则$A\in\mathcal{A}$ 且

$$0\leq\int_A f d\mu=\int_A \mathbb{E}(f|\mathcal{A})d\mu.$$

这蕴含$\mu(A)=0$.

(3) 设$f\in L^1(X,\mathcal{X},\mu)$和$g\in L^\infty(X,\mathcal{A},\mu)$. 则$g\mathbb{E}(f|\mathcal{A})\in L^1(X,\mathcal{A},\mu)$. 如果存在$B\in\mathcal{A}$使得$g=\mathbf{1}_B$, 那么对任意的$A\in\mathcal{A}$,

$$\int_A g\mathbb{E}(f|\mathcal{A})d\mu=\int_{A\cap B}\mathbb{E}(f|\mathcal{A})d\mu=\int_{A\cap B} f d\mu=\int_A gf d\mu.$$

由(1), $\mathbb{E}(gf|\mathcal{A})=g\mathbb{E}(f|\mathcal{A})$. 由于$\mathbb{E}(\cdot|\mathcal{A})$线性, 当$g$是简单函数的时候也成立. 对于一般的$g\in L^\infty(X,\mathcal{A},\mu)$ 存在$(X,\mathcal{A})$上的一致有界的简单函数列$\{g_n\}$几乎处处收敛到g. 利用控制收敛定理即得结论.

(4) 对任意得$C\in\mathcal{C}$, 有$C\in\mathcal{A}$, 从而

$$\int_C \mathbb{E}(\mathbb{E}(f|\mathcal{A})|\mathcal{C})d\mu=\int_C \mathbb{E}(f|\mathcal{A})d\mu=\int_C f d\mu=\int_C \mathbb{E}(f|\mathcal{C})d\mu.$$

由(1), $\mathbb{E}(\mathbb{E}(f|\mathcal{A})|\mathcal{C})=\mathbb{E}(f|\mathcal{C})$.

(5)是(1)得简单推论.

(6) 设$f\in L^1(X,\mathcal{X},\mu)$. 那么存在$g\in L^\infty(X,\mathcal{A},\mu)$使得$|g(x)|=1$, $x\in X$并且

$$|\mathbb{E}(f|\mathcal{A})|=g\mathbb{E}(f|\mathcal{A}).$$

由(3),

$$|\mathbb{E}(f|\mathcal{A})|=\mathbb{E}(gf|\mathcal{A}).$$

对任意得$A\in\mathcal{A}$,

$$\begin{aligned}\int_A|\mathbb{E}(f|\mathcal{A})|d\mu&=\int_A\mathbb{E}(gf|\mathcal{A})d\mu=\int_A gf d\mu\\ &\leq\int_A|gf|d\mu=\int_A|f|d\mu=\int_A\mathbb{E}(|f||\mathcal{A})d\mu.\end{aligned}$$

所以$|\mathbb{E}(f|\mathcal{A})|\leq\mathbb{E}(|f||\mathcal{A})$. □

例 0.3.7 *1.* 若$\mathcal{N}=\{\emptyset,X\}$, 则对任意的$f\in L^1(X,\mathcal{X},\mu)$,

$$\mathbb{E}(f|\mathcal{N})=\int_X f d\mu.$$

2. 若$\mathcal{A}=\{\emptyset, A, X\setminus A, X\}$ (其中$0<\mu(A)<1$), 则对任意的$f\in L^1(X,\mathcal{X},\mu)$,

$$\mathbb{E}(f|\mathcal{A})=\frac{1}{\mu(A)}\int_A f d\mu\cdot\chi_A+\frac{1}{1-\mu(A)}\int_{X\setminus A} f d\mu\cdot\chi_{X\setminus A}.$$

3. 设$X=[0,1]^2$是单位正方形并赋予2-维的*Lebesgue* 测度. 令$\mathcal{A}=\mathcal{X}\times\{\emptyset,[0,1]\}$, 即它是由形如$B\times[0,1]$这样的集合生成的$\sigma$代数, 其中$B$是$[0,1]$上的可测集. 则

$$\mathbb{E}(f|\mathcal{A})(x_1,x_2)=\int_{[0,1]} f(x_1,t)dt.$$

我们也可以通过保测映射将条件期望定义到不同空间上. 设$(X,\mathcal{X},\mu)$ 和$(Y,\mathcal{Y},\nu)$ 为测度空间, $\phi:(X,\mathcal{X},\mu)\to(Y,\mathcal{Y},\nu)$ 为保测映射. 我们定义

$$\mathbb{E}(\cdot|Y):L^1(X,\mathcal{X},\mu)\to L^1(Y,\mathcal{Y},\nu)$$

为

$$\mathbb{E}(f|Y)\circ\phi=\mathbb{E}(f|\phi^{-1}\mathcal{Y}),\forall f\in L^1(X,\mathcal{X},\mu).$$

如果$\mathcal{Y}$ 为$\mathcal{X}$的子σ 代数, 那么对于$\mathrm{id}:(X,\mathcal{X},\mu)\to(X,\mathcal{Y},\mu)$, 上面得到的就是条件期望. 这种条件期望的推广可以参见Furtenberg 的专著[64].

下面我们简要给出[64] 中的处理方式. 设$\phi:(X,\mathcal{X},\mu)\to(Y,\mathcal{Y},\nu)$ 为一个保测映射, 我们可以将$(Y,\mathcal{Y},\nu)$ 上的可测函数通过$f\to f\circ\phi=f^\phi$ 提升为$(X,\mathcal{X},\mu)$ 上的可测函数. 通过映射$f\to f^\phi$, 我们就把$L^2(Y)$ 等同为闭子空间

$$L^2(Y)^\phi\subseteq L^2(X).$$

如果记P 为从$L^2(X)$到$L^2(Y)^\phi$ 的正交投影, 那么对$f\in L^2(X)$ 我们定义$\mathbb{E}(f|Y)$ 为

$$\mathbb{E}(f|Y)\in L^2(Y),\ \ \mathbb{E}(f|Y)^\phi=Pf.$$

定理 0.3.8 在$L^2(X)$上定义的条件期望算子$f\to\mathbb{E}(f|Y)$可以延拓到$L^1(X)$上且满足下面的性质:

(i) $f\to\mathbb{E}(f|Y)$为从$L^1(X,\mathcal{X},\mu)$到$L^1(Y,\mathcal{Y},\nu)$ 的线性算子.

(ii) 如果$f\geq 0$, 那么$\mathbb{E}(f|Y)\geq 0$.

(iii) 如果$f\in L^2(Y)$, 那么$\mathbb{E}(f^\phi|Y)=f$.尤其$\mathbb{E}(\mathbf{1}|Y)=\mathbf{1}$.

(iv) 如果$g\in L^\infty(Y)$, 那么$\mathbb{E}(g^\phi f|Y)=g(\mathbb{E}(f|Y))$.

(v) $\int_X f d\mu=\int_Y \mathbb{E}(f|Y)d\nu$.

§0.3.3 测度分解

对于概率空间$(X,\mathcal{X},\mu)$ 以及子σ-代数$\mathcal{A}\subseteq\mathcal{X}$, 根据条件期望, 我们可以得到$\mu$ 相对于$\mathcal{A}$的条件测度. 设$(X,\mathcal{X})$是一个可测空间, 如果存在Polish 空间Y 使得$(X,\mathcal{X})$ 同构于$(Y,\mathcal{B}(Y))$, 那么称$(X,\mathcal{X})$ 是一个标准Borel 空间. 标准Borel 空间上的概率空间称为标准Borel 概率空间. 关于标准Borel 概率空间具体讨论参见第四章.

定理 0.3.9 *[54, 定理5.14] 设$(X,\mathcal{X},\mu)$ 是一个标准Borel 概率空间, $\mathcal{A}\subseteq\mathcal{X}$ 为子σ代数. 则存在一个$\mathcal{A}$ 可测的μ 满测集X_0 和一族X 上的测度$\{\mu_x^{\mathcal{A}}\colon x\in X_0\}$（称为***条件测度***）满足下面的性质:*

1. *每个$\mu_x^{\mathcal{A}}$ 为概率测度并且满足*
$$\mathbb{E}(f|\mathcal{A})(x)=\int_X f(y)d\mu_x^{\mathcal{A}}(y),\ \mu-a.e.\ \forall f\in L^1(X,\mathcal{X},\mu).$$
事实上, 对于每个$f\in\mathcal{L}^1(X,\mathcal{X},\mu)$（注意不是$L^1(X,\mathcal{X},\mu)$）, $\int_X f(y)d\mu_x^{\mathcal{A}}(y)$ 对于一个$\mathcal{A}$可测的满测集中的任意x 存在, 映射
$$x\mapsto\int_X f(y)d\mu_x^{\mathcal{A}}(y)$$
为$\mathcal{A}$-可测的, 并且对任意的$A\in\mathcal{A}$,
$$\int_A fd\mu=\int_A\Big(\int f(y)d\mu_x^{\mathcal{A}}(y)\Big)d\mu(x).$$
2. *如果$\mathcal{A}$是可数生成的, 则对任意的$x\in X_0$, $\mu_x^{\mathcal{A}}([x]_{\mathcal{A}})=1$, 其中*
$$[x]_{\mathcal{A}}=\bigcap_{x\in A\in\mathcal{A}}A$$
为包含x 的$\mathcal{A}$原子. 进一步, 对于$x,y\in X_0$, 若$[x]_{\mathcal{A}}=[y]_{\mathcal{A}}$, 则$\mu_x^{\mathcal{A}}=\mu_y^{\mathcal{A}}$.
3. *设$\widetilde{\mathcal{A}}$ 为满足$\widetilde{\mathcal{A}}\underset{\mu}{=}\mathcal{A}$ 的子σ代数, 那么$\mu_x^{\widetilde{\mathcal{A}}}=\mu_x^{\mathcal{A}}$.*

注记 0.3.10 设$\mathcal{A}$ 是可数生成的, 即如果存在$\mathcal{A}$ 的一个至多可数子集$\{A_1,A_2,\dots\}$ 使得$\mathcal{A}=\sigma(\{A_1,A_2,\dots\})$. 此时对任意的$x\in X$,
$$[x]_{\mathcal{A}}=\bigcap_{x\in A_i}A_i\cap\bigcap_{x\notin A_i}X\setminus A_i.$$
所以$[x]_{\mathcal{A}}$是$\mathcal{A}$中包含x的最小集合.

对于紧致度量空间, $\mathcal{X}(X)$ 是可数生成的且$[x]_{\mathcal{X}(X)}=\{x\}$.

设$(X,\mathcal{X},\mu)$和$(Y,\mathcal{A},\nu)$ 为测度空间, $\phi:(X,\mathcal{X},\mu)\to(Y,\mathcal{A},\nu)$ 为保测映射. 类似于$\mathbb{E}(\cdot|\mathcal{A})$与$\mathbb{E}(\cdot|Y)$ 的关系, 条件测度也可以有类似的对应. 对于紧致度量空间, 记$\mathcal{M}(X)$ 为$(X,\mathcal{X}(X))$上Borel概率测度全体.

定理 0.3.11 (测度分解(measure disintegration), [64]) *设$(X,\mathcal{B}(X))$ 和$(Y,\mathcal{B}(Y))$ 为标准Borel概率空间, ϕ: $X\to Y$ 为Borel 映射. 设$\mu\in\mathcal{M}(X)$, 令$\nu=\phi_*\mu=\mu\circ\phi^{-1}\in\mathcal{M}(Y)$. 则存在一个Borel 映射*

$$Y\to\mathcal{M}(X),\quad y\mapsto\mu_y$$

使得:

1. *对ν 几乎处处的$y\in Y$, $\mu_y(\pi^{-1}(y))=1$;*

2. *对于任何$A\in\mathcal{B}(X)$, 有*
$$\mu(A)=\int_Y\mu_y(A)d\nu(y),$$
等价的, 对于任何有界Borel 函数$f:X\to\mathbb{R}$ 有
$$\int_X fd\mu=\int_Y\left(\int_X f(x)d\mu_y(x)\right)d\nu(y).$$
把上面陈述的事实简记为
$$\mu=\int_Y\mu_y d\nu(y),$$
*称之为μ 相对于ν 的*测度分解*. 事实上, 对任意的$f\in L^1(X,\mathcal{B}(X),\mu)$, $A\in\mathcal{B}(Y)$,*
$$\int_{\pi^{-1}(A)}fd\mu=\int_A\left(\int f(x)d\mu_y(x)\right)d\nu(y).$$

3. *(2) 中测度分解是唯一的, 即如果$y\mapsto\mu_y'$ 为满足条件的另一分解, 那么$\mu_y=\mu_y'$, $\nu-a.e.$*

4. *对于任何$f\in L^1(X,\mu)$, 那么对于$\nu-a.e.\ y\in Y$, 我们有$f\in L^1(X,\mu_y)$ 并且*
$$\mathbb{E}(f|Y)(y)=\int_X fd\mu_y.$$

下面结论解释了两种定义之间的关联.

定理 0.3.12 *[54, 推论5.22] 设$(X,\mathcal{X},\mu)$ 是一个标准Borel 概率空间, $\mathcal{A}\subseteq\mathcal{X}$ 为可数生成的子σ 代数. 则存在一个$\mathcal{A}$ 可测的μ 满测集X_0 和一个紧致度量空间Y, 可测映射$\phi:X_0\to Y$ 使得*

$$\mathcal{A}\cap X_0=\phi^{-1}(\mathcal{B}(Y)),$$

并且

$$[x]_{\mathcal{A}}=\phi^{-1}(\phi(x)),\forall x\in X_0.$$

记$\mu_{\phi(x)}=\mu_x^{\mathcal{A}}$, 那么

$$Y\to\mathcal{M}(X),\ y\mapsto\mu_y$$

为定义在$\nu=\phi_*\mu$ 满测集上的一个可测映射.

对于动力系统, 我们还有一个非常重要的性质：

定理 0.3.13 *[64]* 设$(X,\mathcal{B}(X),\mu,T)$, $(Y,\mathcal{B}(Y),\nu,S)$ 为标准$Borel$空间上的系统, $\mu=\int_Y\mu_y d\nu(y)$ 为测度分解. 那么对于$\nu-a.e.\ y\in Y$, 我们有

$$T_*\mu_y=\mu_{Sy}.$$

§0.4 拓扑群

本节中我们约定所有的拓扑空间都是Hausdorff的

§0.4.1 拓扑群与Haar测度

定义 0.4.1 设$(G,\cdot)$是一个群且$\mathcal{T}$是G上的一个拓扑. 如果群运算$(a,b)\mapsto a\cdot b$和$a\mapsto a^{-1}$ 相对于拓扑$\mathcal{T}$是连续的, 则称$(G,\cdot,\mathcal{T})$是一个**拓扑群**.

当群运算和拓扑不会引起混淆的情况下, 直接称G是一个拓扑群.

设X是一个拓扑空间, $\mathcal{B}(X)$为其上的Borel σ-代数. 设μ为$\mathcal{B}(X)$上的测度, $E\in\mathcal{B}(X)$. 称μ在E上为**外正则的**(outer regular)是指

$$\mu(E)=\inf\{\mu(U):U\supseteq E,\ U\text{为开集}\}.$$

称μ在E上为**内正则的**(inner regular)是指

$$\mu(E)=\sup\{\mu(K):K\subseteq E,\ K\text{为紧集}\}.$$

如果μ在任何Borel子集上面是外正则且内正则的, 那么称μ为**正则的**. 任何度量空间上的Borel 测度都是正则的.

对于σ有限的测度, 要求它正则会比较苛刻. 一个测度μ称为**Radon测度**, 是指μ在任何紧致子集上取值有限, 在任何Borel子集上外正则, 在任何开集上内正则. 可证Radon测度在其任何σ 有限的Borel 子集上也是内正则的.

设G是一个拓扑群, $\mathcal{B}(G)$为其上的Borel σ 代数. 称其上测度m为**左不变的**是指对任意$x\in G$和$E\in\mathcal{B}(G)$, 有

$$m(xE)=m(E).$$

如果上式换为$m(Ex)=m(E)$, 则称μ为**右不变的**. 易见左不变性等价于下式:

$$\int_G f(xy)dm(y)=\int_G f(y)dm(y),\ \forall f\in L^1(G,m),\ \forall x\in G.$$

对右不变有类似结果.

定义 0.4.2 设G是一个拓扑群, $\mathcal{B}(G)$为其上的$Borel$ σ-代数. G 上非零的左（右）不变的$Radon$测度称为**左Haar测度**(右$Haar$测度).

定理 0.4.3 设G是一个局部紧拓扑群, 那么G上存在左$Haar$测度. 设μ,ν为G的两个左$Haar$测度, 那么存在$c>0$ 使得$\mu=c\nu$.

在上定理中把“左”换为“右”结果仍成立. 一般而言, 左Haar测度与右Haar测度是不一样, 但是对于交换群二者一致.

设G_i为局部紧拓扑群, m_i为其上左Haar测度($i\in\mathbb{Z}$), 那么乘积测度$m=\prod_{i\in\mathbb{Z}}m_i$为乘积群$\prod_{i\in\mathbb{Z}}G_i$上的左Haar测度.

一个拓扑群可度量化当且仅当它具有可数基. 在此情况下我们有更强的结果:

定理 0.4.4 (Birkhoff - Kakutani) 设G是一个局部紧可度量化的拓扑群. 则存在G上的左不变度量d, 即对任意的$x,y,z\in G$, $d(zx,zy)=d(x,y)$.

注记 0.4.5 设$\{V_n\}_{n\in\mathbb{N}}$ 为单位元处的可数基, V_1紧致, $\{V_n\}_{n\in\mathbb{N}}$单调递减, 那么左不变的度量可以取

$$d(x,y)=\sup_{n\in\mathbb{N}}m(xV_n\Delta yV_n),$$

其中m为左$Haar$测度.

对于紧致拓扑群, 有更多好的性质.

注记 0.4.6 设G是一个紧致拓扑群.

1. 根据前面的定理, 存在G上的一个$Borel$概率测度m 使得对任意$x\in G,E\in\mathcal{B}(G)$, $m(xE)=m(E)$, 且m 为正则的. 由于此时要求了$m(X)=1$, 满足这种条件的测度是唯一的.

2. G上的左$Haar$概率测度和右$Haar$概率测度吻合. 证明如下: 设m为左$Haar$概率测度, 对于任何$x\in G$, 令$m_x(E)=m(Ex)$. 则m_x也是左不变正则的, 根据唯一性$m_x=m$. 所以m右不变.

   在本书中, 紧致拓扑群G 上的**Haar 测度**特指这个唯一的概率测度m.

3. 局部紧群上的左$Haar$测度在非空开集上取值都是大于0的, 这点对于紧致群很容易说明. 设U 为G 的非空开集, 则U 的有限次平移覆盖G, 于是$m(U)>0$.

4. 设G是一个紧致可度量化的拓扑群. 则存在G上左不变且右不变度量d, 即对任意的$x,y,z\in G$, $d(xz,yz)=d(x,y)=d(zx,zy)$. 设ρ为G任何相容的度量, 那么

   $$d(x,y)=\int_G\left(\int_G\rho(gxh,gyh)dm(g)\right)dm(h)$$

   为满足条件的度量.

下面是一些具体的例子.

例 0.4.7 *1.* 设$n\in\mathbb{N}$, $\mathbb{Z}_n$ 为模n 加法群. 其上*Haar* 测度为$\nu_{\vec{p}}$, $\vec{p}=(1/n,\dots,1/n)$, 即

$$\nu_{\vec{p}}(\{\bar{0}\})=\cdots=\nu_{\vec{p}}(\overline{n-1})=\frac{1}{n}.$$

2. 设$\mathbb{Z}_2=\{\bar{0},\bar{1}\}$, $\nu(\{\bar{0}\})=\nu(\{\bar{1}\})=\frac{1}{2}$, 则$\nu$为$Y$上的*Haar*测度, $\Sigma_2=\{\bar{0},\bar{1}\}^{\mathbb{Z}}$ 的乘积测度为其*Haar*测度.

3. $\mathbb{R}^n$按照通常的加法和拓扑是一个局部紧可度量化的交换群. *Lebesgue*测度就是其一个*Haar*测度.

4. $\mathbb{Z}$按照通常的加法和作为$\mathbb{R}$的子空间是一个离散交换群, 计数测度为*Haar*测度.

5. 令$\mathbb{S}^1=\{z\in\mathbb{C}\colon |z|=1\}$. 则$\mathbb{S}^1$ 按照复数的乘法和$\mathbb{C}$ 的子拓扑是一个紧致可度量化的交换群. *Haar* 测度为其上的*Lebesgue*测度.

6. 令$\mathbb{T}=\mathbb{R}/\mathbb{Z}$. 按照$\mathbb{R}$ 遗传的加法和商拓扑, $\mathbb{T}$ 是一个紧致可度量化的交换群. 区间$[0,1)$到$\mathbb{T}=\mathbb{R}/\mathbb{Z}$ 与有一个自然的一一对应$x\mapsto x+\mathbb{Z}$. 设$x,y\in[0,1)$,

$$d(x,y)=\min(|x-y|,|1-x+y|)$$

是$\mathbb{T}$ 上的一个相容度量.

映射$e\colon[0,1)\to\mathbb{S}^1$, $x\mapsto e^{2\pi ix}$ 是一个同胚并且为群同构. 所以作为拓扑群, $\mathbb{T}$和$\mathbb{S}^1$ 是同构的.

§0.4.2 特征与对偶群

定义 0.4.8 设G是一个局部紧的交换群. 如果$\gamma\colon G\to\mathbb{S}^1$是一个连续的群同态, 则称$\gamma$是$G$的一个**特征**(*character*). G所有特征的全体记为$\widehat{G}$, 称之为G的**对偶**.

定理 0.4.9 $\widehat{G}$中群运算取其元素的逐点相乘成为一个交换群. 如果取$C(G,\mathbb{S}^1)$上紧开拓扑诱导的拓扑, $\widehat{G}$成为一个局部紧的交换拓扑群.

例 0.4.10 以下"$\cong$"指同构:

1. $\widehat{\mathbb{Z}_n}\cong\mathbb{Z}_n$, $\forall n\in\mathbb{N}$.

2. $\widehat{\mathbb{S}^1}\cong\mathbb{Z}$ 并且 $\widehat{\mathbb{Z}}\cong\mathbb{S}^1$.

3. $\widehat{\mathbb{R}}\cong\mathbb{R}$.

设G是一个局部紧的交换群. 我们罗列其特征群的主要结论如下:

(1) G具有可数基当且仅当$\widehat{G}$具有可数基.

(2) G是紧致的当且仅当$\widehat{G}$是离散的. 于是G为紧致可度量的当且仅当$\widehat{G}$为可数离散群.

(3) （**对偶定理**）定义$\phi : \widehat{\widehat{G}} \to G, \alpha \mapsto a$ 为$\alpha(\gamma) = \gamma(a), \forall \gamma \in \widehat{G}$. 此对应建立了$\widehat{\widehat{G}}$与$G$之间一个自然的拓扑群同构.

(4) 设G为紧致的, 那么G为连通的当且仅当$\widehat{G}$为无桡的(torsion free 的), 即除单位元外没有有限阶的元素.

(5) 设G_1, G_2为局部紧交换群, 则$\widehat{G_1 \times G_2} \cong \widehat{G_1} \times \widehat{G_2}$.

于是对于任何$\lambda \in \widehat{G_1 \times G_2}$, 存在$\gamma \in \widehat{G_1}$和$\delta \in \widehat{G_2}$ 使得

$$\lambda(x, y) = \gamma(x)\delta(y),\ \forall\ x \in G_1, y \in G_2.$$

(6) 设$\Gamma < \widehat{G}$, 则

$$H = \{g \in G : \gamma(g) = 1, \forall \gamma \in \Gamma\}$$

为G的闭子群, 并且

$$\widehat{(G/H)} \cong \Gamma.$$

(7) 对任意G的真闭子群H, 存在$\gamma \in \widehat{G}, \gamma \neq 1$ 使得$\gamma(h) = 1,\ h \in H$.

(8) 设G为紧致的, 那么$\widehat{G}$为$L^2(G, m)$的标准正交基, 其中m为G的Haar测度. 于是对于$f \in L^2(G, m)$, 存在**Fourier 展开：**

$$f = \sum_{\gamma \in \widehat{G}} a_\gamma \gamma,$$

a_γ称为Fourier系数. 上式等号的意义是, 对于任何$\varepsilon > 0$,存在有限集$J_\varepsilon \subseteq \widehat{G}$ 使得对于任何包含J_ε的$\widehat{G}$ 的有限子集J有

$$\|f - \sum_{\gamma \in J} a_\gamma \gamma\|_2 < \varepsilon.$$

例如当$G = \mathbb{S}^1$时, 上式即为经典的Fourier 级数展开:

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n.$$

(9) 设$A : G \to G$为自同态, 我们可以定义对偶自同态：

$$\widehat{A} : \widehat{G} \to \widehat{G},\ \gamma \mapsto \gamma \circ A, \forall \gamma \in \widehat{G}.$$

A为单射当且仅当$\widehat{A}$为满射；$\widehat{A}$为单射当且仅当A 为满射. 于是A为自同构当且仅当$\widehat{A}$为自同构.

§0.4.3 环面的同态

对于n维环面, 我们有两种经典的表述方式:

(1) $(\mathbb{S}^1)^n = \mathbb{S}^1 \times \ldots \times \mathbb{S}^1$;

(2) $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

通过映射

$$(\mathbb{S}^1)^n \to \mathbb{T}^n,\ (e^{2\pi i x_1}, \ldots, e^{2\pi i x_n}) \mapsto (x_1, \ldots, x_n) + \mathbb{Z}^n$$

将二者等价视之.

下面的二个定理是经常要用到的, 为此我们给出详细的证明.

定理 0.4.11 *1. 若H是$\mathbb{S}^1$的闭子群, 则要么$H = \mathbb{S}^1$, 要么H 为有限循环群, 即存在$p \in \mathbb{N}$ 使得$H = H_p = \{1, a, a^2, \ldots, a^{p-1}\}$, 其中$a \in \mathbb{S}^1$ 为p 阶单位根.*

2. 若χ是$\mathbb{S}^1$的一个连续自同构, 则χ是恒同映射或映射$z \mapsto z^{-1}$.

3. 若χ是$\mathbb{S}^1$的一个特征, 则存在$m \in \mathbb{Z}$使得$\chi(z) = z^m$.

4. 设$n \in \mathbb{N}$. 若ϕ是$(\mathbb{S}^1)^n$的一个特征, 则存在$m_1, m_2, \ldots, m_n \in \mathbb{Z}$ 使得

$$\phi(z_1, z_2, \ldots, z_n) = z_1^{m_1} z_2^{m_2} \cdots z_n^{m_n}.$$

证明. (1) 设H为$\mathbb{S}^1$的一个闭子群. 如果H无限, 那么它存在极限点. 于是对于任何$\varepsilon > 0$, 存在两个不同点$a, b \in H$ 使得$d(a, b) < \varepsilon$. 于是$d(b^{-1}a, 1) < \varepsilon$, 其中$d$ 为$\mathbb{S}^1$ 上的度量. 尤其$\langle b^{-1}a \rangle = \{(b^{-1}a)^n : n \in \mathbb{Z}\}$ 在$\mathbb{S}^1$ 中为ε稠密的. 由此H在$\mathbb{S}^1$中ε稠密. 因为ε任意以及H为闭的, 所以$H = \mathbb{S}^1$.

如果H为有限的, 设只有p个元素. 于是有$a^p = 1, \forall a \in H$. 因为H每个元素都是p次单位根, 而且H个数为p, 所以H恰好为所有p次单位根的全体.

(2) 设$\chi : \mathbb{S}^1 \to \mathbb{S}^1$为连续自同构. 首先我们有$\chi(1) = 1$. 因为$-1$为$\mathbb{S}^1$中唯一的阶数恰好为2的元素, 所以$\chi(-1) = -1$. 因为$i, -i$ 为$\mathbb{S}^1$中阶数恰为4的所有元素, 所以有两种情况:

(I) $\chi(i) = i, \chi(-i) = -i$;

(II) $\chi(i) = -i, \chi(-i) = i$.

对于(I): 由连续性, χ将区间映为区间. 记$\overrightarrow{[a, b]}$为从a到b的逆时针区间. 于是χ将$\overrightarrow{[1, i]}$要么映为自身要么映为$\overrightarrow{[i, 1]}$. 因为$\overrightarrow{[1, i]}$不包含-1, 它不能映为$\overrightarrow{[i, 1]}$. 于是$\chi(\overrightarrow{[1, i]}) = \overrightarrow{[1, i]}$. 在$\overrightarrow{[1, i]}$中唯一的阶为8的元素是$e^{\pi i/4}$, 于是$\chi(e^{\pi i/4}) = e^{\pi i/4}$. 由此得到$\chi(\overrightarrow{[1, e^{\pi i/4}]}) = \overrightarrow{[1, e^{\pi i/4}]}$. 根据同样的道理归纳就得到

$$\chi(\overrightarrow{[1, e^{2\pi i/2^k}]}) = \overrightarrow{[1, e^{2\pi i/2^k}]},\ \forall k \in \mathbb{N}.$$

由于χ为同态, 上式说明对于任何$k \in \mathbb{N}$, χ在所有2^k阶的单位根上恒同. 所有2^k阶的单位根的全体是$\mathbb{S}^1$的稠密子集, 根据连续性就得到$\chi = \mathrm{id}$.

对于(II), 类似的分析得到$\chi(e^{2\pi i/2^k}) = e^{-2\pi i/2^k}$, $\forall k \in \mathbb{N}$. 由此得到$\chi(z) = z^{-1}$.

(3) 设$\chi : \mathbb{S}^1 \to \mathbb{S}^1$为连续自同态. 如果$\chi$非平凡, 则$\chi(\mathbb{S}^1)$为$\mathbb{S}^1$的连通闭子群, 于是根据(1)所证$\chi(\mathbb{S}^1) = \mathbb{S}^1$. 先考虑$\mathrm{Ker}\chi$. $\mathrm{Ker}\chi$也是$\mathbb{S}^1$的闭子群, 根据(1), $\mathrm{Ker}\chi = \mathbb{S}^1$或者存在$p \in \mathbb{N}$使得$\mathrm{Ker}\chi = H_p$. 如果$\mathrm{Ker}\chi = \mathbb{S}^1$, 那么$\chi = 1$. 如果$\mathrm{Ker}\chi = H_p$, 令$\chi_1 : \mathbb{S}^1/H_p \to \mathbb{S}^1$ 为由χ诱导的映射, 即$\chi_1(zH_p) = \chi(z)$. 设

$$\alpha_p : \mathbb{S}^1/H_p \to \mathbb{S}^1,\ zH_p \mapsto z^p.$$

于是

$$\chi_1 \circ \alpha_p^{-1} : \mathbb{S}^1 \to \mathbb{S}^1$$

为$\mathbb{S}^1$同构, 根据(2), $\chi_1 \circ \alpha_p^{-1}(z) = z, \forall z \in \mathbb{S}^1$或者$\chi_1 \circ \alpha_p^{-1}(z) = z^{-1}, \forall z \in \mathbb{S}^1$. 这样就有

$$\chi(z) = \chi_1(zH_p) = \chi_1 \circ \alpha_p^{-1}(z^p) = z^p,\ \forall z \in \mathbb{S}^1,$$

或者

$$\chi(z) = z^{-p},\ \forall z \in \mathbb{S}^1.$$

(4) 对$i \in \{1, 2, \ldots, n\}$, 设

$$\gamma_i : \mathbb{S}^1 \to (\mathbb{S}^1)^n,\ z \mapsto (1, \ldots, 1, z, 1, \ldots, 1)$$

为到第i个坐标的嵌入映射. 设$\phi : (\mathbb{S}^1)^n \to \mathbb{S}^1$为同态. 则$\phi \circ \gamma_i : \mathbb{S}^1 \to \mathbb{S}^1$为自同态, 由(3)就有存在$m_i \in \mathbb{Z}$使得$\phi \circ \gamma_i(z) = z^{m_i}$. 综上就得到

$$\begin{aligned}\phi(z_1, \ldots, z_n) &= \phi(\gamma_1(z_1) \cdot \gamma_2(z_2) \cdot \ldots \cdot \gamma_n(z_n)) \\ &= \phi(\gamma_1(z_1))\phi(\gamma_2(z_2)) \ldots \phi(\gamma_n(z_n)) \\ &= z_1^{m_1} z_2^{m_2} \ldots z_n^{m_n}.\end{aligned}$$

定理证毕. □

推论 0.4.12 $\widehat{(\mathbb{S}^1)^n} = \mathbb{Z}^n$.

事实上, 若ϕ是$(\mathbb{S}^1)^n$的一个特征, 则存在$m_1, m_2, \ldots, m_n \in \mathbb{Z}$ 使得

$$\phi(z_1, z_2, \ldots, z_n) = z_1^{m_1} z_2^{m_2} \cdots z_n^{m_n}.$$

令

$$\widehat{(\mathbb{S}^1)^n} \to \mathbb{Z}^n,\ \phi \mapsto \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix}$$

此映射就是同构.

定理 0.4.13 *1.* 每个自同态$A:(\mathbb{S}^1)^n\to(\mathbb{S}^1)^n$ 形如

$$A(z_1,\ldots,z_n)=(z_1^{a_{11}}z_2^{a_{12}}\ldots z_n^{a_{1n}},\ldots,z_1^{a_{n1}}z_2^{a_{n2}}\ldots z_n^{a_{nn}}),$$

其中$a_{ij}\in\mathbb{Z}$. 等价的, $A:\mathbb{T}^n\to\mathbb{T}^n$表达式为

$$A\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+\mathbb{Z}^n\right)=[a_{ij}]\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+\mathbb{Z}^n,$$

其中$[a_{ij}]\in\mathbb{Z}^{n\times n}$为方阵, 第$(i,j)$元素为$a_{ij}$.

2. 自同态$A:(\mathbb{S}^1)^n\to(\mathbb{S}^1)^n$为满射当且仅当$\det[a_{ij}]\neq 0$.

3. 自同态$A:(\mathbb{S}^1)^n\to(\mathbb{S}^1)^n$为自同构当且仅当$\det[a_{ij}]=\pm 1$.

证明. (1) 根据定理0.4.11-4易得. (2)和(3)的证明是标准的, 留作习题. □

于是$\mathbb{T}^n$上的满自同态与$GL(\mathbb{Z},n)$($\mathbb{Z}$上$n\times n$阶方阵的全体)之间一一对应.

在以后的讨论中, 如果$A:\mathbb{T}^n\to\mathbb{T}^n$为自同态, 那么我们用$[A]$表示对应的方阵, 而用$\widetilde{A}:\mathbb{R}^n\to\mathbb{R}^n$表示由矩阵$[A]$定义的线性映射. 设$\pi:\mathbb{R}^n\to\mathbb{T}^n=\mathbb{R}^n/\mathbb{Z}^n,\mathbf{x}\mapsto\mathbf{x}+\mathbb{Z}^n$ 为自然投射, 于是我们有下面的交换图标:

$$\begin{tikzcd}\mathbb{R}^n \arrow[r, "\widetilde{A}"] \arrow[d, "\pi"'] & \mathbb{R}^n \arrow[d, "\pi"] \\ \mathbb{T}^n \arrow[r, "A"'] & \mathbb{T}^n.\end{tikzcd}$$

设$A:(\mathbb{S}^1)^n\to(\mathbb{S}^1)^n$为自同态, 下面我们确定它的对偶自同态

$$\widehat{A}:\widehat{(\mathbb{S}^1)^n}\to\widehat{(\mathbb{S}^1)^n}.$$

如前所述, 若ϕ是$(\mathbb{S}^1)^n$的一个特征, 则存在$m_1,m_2,\ldots,m_n\in\mathbb{Z}$ 使得

$$\phi((z_1,z_2,\ldots,z_n)=z_1^{m_1}z_2^{m_2}\cdots z_n^{m_n}.$$

令

$$\widehat{(\mathbb{S}^1)^n}\to\mathbb{Z}^n,\ \phi\mapsto\begin{pmatrix}m_1\\ m_2\\ \vdots\\ m_n\end{pmatrix}$$

此映射就是同构. 于是

$$\widehat{A}:\mathbb{Z}^n\to\mathbb{Z}^n,\begin{pmatrix} m_1\\ m_2\\ \vdots\\ m_n\end{pmatrix}\mapsto [A]^t\begin{pmatrix} m_1\\ m_2\\ \vdots\\ m_n\end{pmatrix}$$

其中$[A]^t$为转置.

§0.5 Perron-Frobenius 理论

因为在本书中讨论Markov 转移系统的时候需要用到非负矩阵的一些性质, 所以在这里我们介绍非负矩阵的Perron-Frobenius 理论.

设$k\in\mathbb{N}$且$A=[a_{ij}]\in\mathbb{R}^{k\times k}$ 为矩阵. 对于$n\in\mathbb{N}$ 我们记$a_{ij}^{(n)}$ 为A^n 的(i,j) 元素, $\forall i,j\in\{1,\ldots,k\}$.

定义 0.5.1 设$A=[a_{ij}]\in\mathbb{R}^{k\times k}$ 为矩阵.

1. 称A为**非负的**(*non-negative*)是指$a_{ij}\geq 0,\forall i,j\in\{1,\ldots,k\}$.
2. 称A为**不可约的**(*irreducible*)是指对于任何$i,j\in\{1,\ldots,k\}$, 存在$n>0$使得$a_{ij}^{(n)}>0$.
3. 称A为**非周期的**(*aperiodic*)是指存在$n>0$使得对于任何$i,j\in\{1,\ldots,k\}$, 有$a_{ij}^{(n)}>0$.

定理 0.5.2 (Perron-Frobeninus定理) 设$A=[a_{ij}]\in\mathbb{R}^{k\times k}$为非负矩阵, $k\in\mathbb{N}$.

1. 存在非负特征值λ使得A的任何其它特征值的绝对值小于等于λ.
2. $\min\limits_{1\leq i\leq k}(\sum\limits_{j=1}^k a_{ij})\leq\lambda\leq\max\limits_{1\leq i\leq k}(\sum\limits_{j=1}^k a_{ij})$
3. 对应于λ存在非负左行特征向量

$$\vec{u}=(u_1,\ldots,u_k)$$

以及非负右列特征向量

$$\vec{v}=\begin{pmatrix} v_1\\ \vdots\\ v_k\end{pmatrix}.$$

即

$$\vec{u}A=\lambda\vec{u},\ A\vec{v}=\lambda\vec{v}.$$

4. 如果A为不可约的, λ为简单的, 并且*3*中对应的特征向量为严格正的, 即$u_i > 0, v_i > 0,\ \forall i \in \{1, 2, \ldots, k\}$.

5. 如果A为不可约的, λ为唯一具有非负特征向量的特征值.

设A为不可约非负矩阵, $\lambda, \vec{u}, \vec{v}$如上定理所述. 定义$k \times k$矩阵$P = [p_{ij}]$如下：

$$p_{ij} = \frac{a_{ij} v_j}{\lambda v_i},\ \forall i, j \in \{1, \ldots, k\}$$

因为$A\vec{v} = \lambda \vec{v}$, 容易验证$0 \le p_{ij} \le 1, 1 \le i, j \le k$, 且

$$\sum_{j=1}^{k} p_{ij} = \sum_{j=1}^{k} \frac{a_{ij} v_j}{\lambda v_i} = \frac{\sum_{j=1}^{k} a_{ij} v_j}{\lambda v_i} = 1,\ 1 \le i \le k.$$

满足这种条件的P称为**随机矩阵**(stochastic matrix). 令$\vec{p} = (p_1, \ldots, p_k)$为

$$p_i = u_i v_i,\ 1 \le i \le k.$$

正规化$\vec{u}, \vec{v}$, 我们可以假设

$$\sum_{i=1}^{k} p_i = 1.$$

即$\vec{p}$为概率向量. 对于$P, \vec{p}$ 我们有:

$$\sum_{i=1}^{k} p_i p_{ij} = \sum_{i=1}^{k} u_i v_i \frac{a_{ij} v_j}{\lambda v_i} = \sum_{i=1}^{k} \frac{a_{ij} u_i}{\lambda} v_j = u_j v_j = p_j, \forall j \in \{1, 2, \ldots, k\}.$$

即

$$\vec{p} P = \vec{p}.$$

定理 0.5.3 设A为不可约非周期的非负矩阵. 设$\lambda, \vec{u}, \vec{v}$ 同*Perron-Frobeninus*定理所述, 此时

$$\vec{u} = (u_1, \ldots, u_k),\ \vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}.$$

为严格正的向量. 那么对于任何$i, j \in \{1, 2, \ldots, k\}$,

$$\lim_{n \to \infty} \frac{a_{ij}^{(n)}}{\lambda^n} = v_i u_j.$$

即

$$\lim_{n \to \infty} \frac{A^n}{\lambda^n} = \vec{v}\vec{u} = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} (u_1, \ldots, u_k).$$

定理 0.5.4 (更新定理(Renewal Theorem)) *设$\{c_n\}_{n=0}^{\infty}$和$\{d_n\}_{n=0}^{\infty}$为实值有界序列, 其中$0 \le c_n \le 1, d_n \ge 0, \forall n \in \mathbb{Z}_+$. 设使得$c_n > 0$ 的所有n 的最大公约数为1. 令*

$$u_n = d_n + c_0 u_n + c_1 u_{n-1} + \ldots + c_n u_0,\ \forall n \in \mathbb{Z}_+.$$

如果$\sum_{n=0}^{\infty} c_n = 1$ 且$\sum_{n=0}^{\infty} d_n < \infty$, 那么

$$\lim_{n\to\infty} u_n = \frac{\sum_{n=0}^{\infty} d_n}{\sum_{n=0}^{\infty} n c_n},$$

在上式中, 如果$\sum_{n=0}^{\infty} n c_n = \infty$, 那么极限定义为0.

§0.6　Furstenberg 族

族的概念最早可以追述到在一般拓扑学与数理逻辑中滤子的使用. 但这种用族的观点来研究系统的动力学性质的思想首先由Gottschalk 和Hedlund (1955) 引入的[84], 之后有许多数学工作者沿着这一思路进行了有意义的讨论. 在Furstenberg 经典著作[64] 中将这一思想进行了深刻而漂亮的阐述. 他的工作将拓扑动力系统与遍历论的应用深刻的植入组合数论与Ramsey 理论之中, 对相应的数学分支有着广泛而深远的影响. Akin 在拓扑动力系统这一范畴下, 在其专著[4]中系统总结和发展了Furstenberg 族的方法.

由于我们在本文中大多情况下只涉及到离散动力系统, 即我们的作用半群是$\mathbb{Z}_+$, 所以为方便与利于理解我们只讨论$\mathbb{Z}_+$ 上的子集族, 而将$\mathbb{Z}_+$ 换为一般作用群或半群G 时, 以下的讨论均是类似的.

设$\mathcal{P}$ 为$\mathbb{Z}_+$ 的全体子集构成的集合. 如果$\mathcal{P}$ 的子集$\mathcal{F}$ 具有遗传向上性, 即若$F_1 \subset F_2$ 且$F_1 \in \mathcal{F}$ 则$F_2 \in \mathcal{F}$, 那么我们就称$\mathcal{F}$ 为**Furstenberg 族**或直接简称为**族**. 族$\mathcal{F}$ 称为**真族**, 如果它是$\mathcal{P}$ 的真子集, 即它既非空集又不为$\mathcal{P}$. 由遗传向上性, $\mathcal{F}$ 为真族当且仅当$\emptyset \notin \mathcal{P}$ 且$\mathbb{Z}_+ \in \mathcal{F}$.

一个真族$\mathcal{F}$ 如果对集合的交运算封闭, 则称之为**滤子**, 即如$F_1, F_2 \in \mathcal{F}$ 则有$F_1 \cap \mathcal{F}_2 \in \mathcal{F}$. 每个$\mathcal{P}$ 的子集$\mathcal{A}$ 均可生成一个族:

$$[\mathcal{A}] = \{F \in \mathcal{P} : \exists A \in \mathcal{A}\ s.t.\ A \subseteq F\}.$$

易见$[\mathcal{A}]$ 是包含$\mathcal{A}$ 的最小的族. 如果$[\mathcal{A}]$ 为滤子, 那么我们称$\mathcal{A}$ 为**滤子基**.

如果$\mathcal{F}$为族, 则它的**对偶**

$$\mathcal{F}^* = \{F \in \mathcal{P} : F \cap F_1 \neq \emptyset \text{ 对所有的} F_1 \in \mathcal{F}\text{成立}\}$$

也为族, 且若$\mathcal{F}$ 为真子族则$\mathcal{F}^*$ 也为真子族. 不难证明:

$$\mathcal{F}^* = \{F \in \mathcal{P} : \mathbb{Z}_+ \setminus F \notin \mathcal{F}\}.$$

$\mathcal{P}$ 的最大的真子族为由$\mathbb{Z}_+$ 的所有非空子集构成的族$\mathcal{P}_+$, 它的对偶族$\mathcal{P}_+^*$ 为最小的真子族$\{\mathbb{Z}_+\}$.

性质 0.6.1 设$\mathcal{F}$, $\mathcal{F}_\alpha$为真族,则

(1) $(\mathcal{F}^*)^* = \mathcal{F}$;

(2) $\mathcal{F}_1 \subset \mathcal{F}_2 \Rightarrow \mathcal{F}_2^* \subset \mathcal{F}_1^*$;

(3) $(\bigcup_\alpha \mathcal{F}_\alpha)^* = \bigcap_\alpha \mathcal{F}_\alpha^*$; $(\bigcap_\alpha \mathcal{F}_\alpha)^* = \bigcup_\alpha \mathcal{F}_\alpha^*$.

设$\mathcal{F}_1$ 与$\mathcal{F}_2$ 为族, 定义运算

$$\mathcal{F}_1 \cdot \mathcal{F}_2 = \{F_1 \cap F_2 : F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}.$$

由定义易见$\mathcal{F}_1 \cup \mathcal{F}_2 \subset \mathcal{F}_1 \cdot \mathcal{F}_2$; $\mathcal{F}_1 \cdot \mathcal{F}_2 = \mathcal{F}_2 \cdot \mathcal{F}_1$; $\mathcal{F}_1 \subseteq \mathcal{F}_2 \Rightarrow \mathcal{F}_1 \cdot \mathcal{F} \subseteq \mathcal{F}_2 \cdot \mathcal{F}$ 等等.

命题 0.6.2 设$\mathcal{F}$, $\mathcal{F}_1$, $\mathcal{F}_2$ 为族,则

(1) $\mathcal{F}_1 \cdot \mathcal{F}_2$ 为真的当且仅当$\mathcal{F}_2 \subset \mathcal{F}_1^*$. 一般的, 我们有: $\mathcal{F}_1 \cdot \mathcal{F}_2 \subseteq \mathcal{F}$ 当且仅当$\mathcal{F}_1 \cdot \mathcal{F}^* \subseteq \mathcal{F}_2^*$.

(2) 真族$\mathcal{F}$ 是滤子当且仅当$\mathcal{F} = \mathcal{F} \cdot \mathcal{F}$. 且如$\mathcal{F}$ 是滤子, 则$\mathcal{F} \subset \mathcal{F} \cdot \mathcal{F}^* = \mathcal{F}^*$.

(3) 如$\mathcal{F}$ 为真子族, 则$(\mathcal{F} \cdot \mathcal{F}^*)^*$ 为包含在$\mathcal{F} \cap \mathcal{F}^*$ 中的滤子, 且此滤子为满足关系$\mathcal{F}_1 \cdot \mathcal{F} \subseteq \mathcal{F}$ 的所有$\mathcal{F}_1$ 中最大的族.

证明. (1) 首先明显有$\mathbb{Z}_+ \in \mathcal{F}_1 \cdot \mathcal{F}_2$, 于是$\mathcal{F}_1 \cdot \mathcal{F}_2$ 为真的当且仅当$\emptyset \notin \mathcal{F}_1 \cdot \mathcal{F}_2$. 即有, 对任意$F_1 \in \mathcal{F}_1$ 及任意$F_2 \in \mathcal{F}_2$, $F_1 \cap F_2 \neq \emptyset$ 成立. 由此式及对偶的定义马上就有$\mathcal{F}_1 \cdot \mathcal{F}_2$ 为真的当且仅当$\mathcal{F}_1 \subseteq \mathcal{F}_2^*$ 当且仅当$\mathcal{F}_2 \subseteq \mathcal{F}_1^*$.

如果$\mathcal{F}_1 \cdot \mathcal{F}_2 \subseteq \mathcal{F}$, 则

$$(\mathcal{F}_1 \cdot \mathcal{F}^*) \cdot \mathcal{F}_2 = (\mathcal{F}_1 \cdot \mathcal{F}_2) \cdot \mathcal{F}^* \subseteq \mathcal{F} \cdot \mathcal{F}^*.$$

于是$(\mathcal{F}_1 \cdot \mathcal{F}^*) \cdot \mathcal{F}_2$ 为真的, 由上已证结论就有$\mathcal{F}_1 \cdot \mathcal{F}^* \subseteq \mathcal{F}_2^*$. 反之设$\mathcal{F}_1 \cdot \mathcal{F}^* \subseteq \mathcal{F}_2^*$, 由上已证结论就有$(\mathcal{F}_1 \cdot \mathcal{F}_2) \cdot \mathcal{F}^*$ 为真的, 继而再用一次上结论就有$\mathcal{F}_1 \cdot \mathcal{F}_2 \subseteq (\mathcal{F}^*)^* = \mathcal{F}$.

(2) 首先注意到$\mathcal{F} \cdot \mathcal{F} \subseteq \mathcal{F}$ 等价于$\mathcal{F} \cdot \mathcal{F} = \mathcal{F}$. 如$\mathcal{F}$ 为滤子, 则$\mathcal{F} \cdot \mathcal{F} = \mathcal{F}$. 由(1)我们就得到$\mathcal{F} \subseteq \mathcal{F}^*$ 以及$\mathcal{F} \cdot \mathcal{F}^* \subseteq \mathcal{F}^*$.

(3) 令$\tilde{\mathcal{F}} = (\mathcal{F} \cdot \mathcal{F}^*)^*$, 则$\mathcal{F} \cdot \mathcal{F}^* = \tilde{\mathcal{F}}^*$. 由(1), $\mathcal{F} \cdot \tilde{\mathcal{F}} \subseteq \mathcal{F}$.

设族$\mathcal{F}_1$ 满足$\mathcal{F}_1 \cdot \mathcal{F} \subseteq \mathcal{F}$, 则$\mathcal{F}_1 \cdot \mathcal{F} \cdot k\mathcal{F} \subseteq \mathcal{F} \cdot k\mathcal{F}$. 又由(1)我们得到$\mathcal{F}_1 \cdot \tilde{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$. 所以$\mathcal{F}_1 \subseteq \mathcal{F}_1 \cdot \tilde{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$. 尤其当$\mathcal{F}_1 = \tilde{\mathcal{F}}$ 时, 有$\tilde{\mathcal{F}} \cdot \tilde{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$. 所以$\tilde{\mathcal{F}}$ 为滤子.

由于$\mathcal{F} \cup \mathcal{F}^* \subseteq \mathcal{F} \cdot \mathcal{F}^*$, 所以$(\mathcal{F} \cdot \mathcal{F}^*)^* \subseteq (\mathcal{F} \cup \mathcal{F}^*)^* = \mathcal{F}^* \cap \mathcal{F}$. □

令$\mathcal{F}_{\text{inf}}$ 为$\mathbb{Z}_+$ 的全体无限子集组成的族, 那么易见它为滤子并且它的对偶族$\mathcal{F}_{\text{inf}}^* = \mathcal{F}_{\text{cf}}$ 为全体有限余集组成的族.

如果$\mathcal{F}^*$ 为滤子, 那么族$\mathcal{F}$ 称为**滤子对偶**. 对于滤子对偶它的重要性在于它有Ramsey 性质:

命题 0.6.3 族$\mathcal{F}$ 为滤子对偶当且仅当它有*Ramsey* 性质,即如$F_1 \cup F_2 \in \mathcal{F}$, 则有$F_1 \in \mathcal{F}$ 或$F_2 \in \mathcal{F}$ 成立.

证明. 设$\mathcal{F}$ 为滤子对偶. 如果$F_1 \notin \mathcal{F}$ 且$F_2 \notin \mathcal{F}$, 则$F_1^c \in \mathcal{F}^*$ 且$F_2^c \in \mathcal{F}^*$. 于是$F_1^c \cap F_2^c \in \mathcal{F}^*$, 所以$(F_1 \cup F_2)^c \in \mathcal{F}^*$, 继而$F_1 \cup F_2 \notin \mathcal{F}$.

反之, 对任意$F_1, F_2 \in \mathcal{F}^*$, 由定义有$F_1^c \notin \mathcal{F}$ 及$F_2^c \notin \mathcal{F}$. 所以$F_1^c \cup F_2^c \notin \mathcal{F}$, 即$(F_1 \cap F_2)^c \notin \mathcal{F}$. 从而$F_1 \cap F_2 \in \mathcal{F}^*$. □

滤子对偶是一条重要的性质, 正是由于这条性质使得族在Ramsey 理论中有着特殊的作用.

在正文中我们会介绍许多重要的族, 在这里我们作为例子介绍一类与密度有关的族. 设A 为$\mathbb{Z}_+$ 或$\mathbb{Z}$ 的子集. 它的**上Banach密度** 定义为:

$$d^*(A) = \limsup_{|I| \to \infty} \frac{|A \cap I|}{|I|},$$

其中I 为$\mathbb{Z}_+$ 或$\mathbb{Z}$ 的区间段, 而$|\cdot|$ 表示集合的基数. 集合A 的**上密度**定义为:

$$\bar{d}(A) = \limsup_{N \to \infty} \frac{|A \cap \{0, 1, \cdots, N-1\}|}{N}.$$

(如A 为$\mathbb{Z}$ 子集, 则定义改为$\bar{d}(A) = \limsup\limits_{N \to \infty} \dfrac{|A \cap \{-N, -N+1, \cdots, N\}|}{2N+1}$). 同理定义**下Banach密度** $d_*(A)$和**下密度** $\underline{d}(A)$. 如果$\bar{d}(A) = \underline{d}(A)$, 那么称$A$ 有**密度** $d(A)$.

容易验证下面性质.

性质 0.6.4 *设E, F 为$\mathbb{Z}_+$ 或$\mathbb{Z}$ 子集.则*

(1) $d_*(E) \le \underline{d}(E) \le \bar{d}(E) \le d^*(E)$;

(2) $d^*(E \cup F) \le d^*(E) + d^*(F)$; 如果$E \cap F = \emptyset$, 则$d_*(E \cup F) \ge d_*(E) + d_*(F)$. 将$d^*, d_*$换为$\bar{d}, \underline{d}$ 结论仍成立.

(3) $\bar{d}(E) = 1 - \underline{d}(E^c)$; $d^*(E) = 1 - d_*(E^c)$.

我们记$\mathcal{F}_{\mathrm{d1}} = \{A : d(A) = 1\} = \{A : \underline{d}(A) = 1\}$ 及$\mathcal{F}_{\mathrm{lbd1}} = \{A : d_*(A) = 1\}$, 它们为滤子. 并且我们有

$$\mathcal{F}_{\mathrm{d1}}^* = \mathcal{F}_{\mathrm{d1}}^* = \{A : A^c \notin \mathcal{F}_{\mathrm{d1}}\} = \{A : \underline{d}(A^c) < 1\} = \{A : \bar{d}(A) > 0\}.$$

同理得到$\mathcal{F}_{\mathrm{lbd1}}^* = \{A : d^*(A) > 0\}$.

§0.7 注记

在Foland的[58], Royden的[171], Rudin的[172, 173]等中, 可以找到上面涉及测度、积分和拓扑的大部分知识. 拓扑群理论可以参见[205, 94]. 测度论系统知识可以参见[32, 91], 一般拓扑学的知识可以参见[33, 127]. Perron-Frobeninus理论在许多矩阵论的书籍中能找到证明, 也可以参见[182]. 关于族的介绍可以参见[64, 4, 207]等, 这部分内容也可参见[96].

第一章　保测系统

设X为空间, G为群, 如果$\phi : G \times X \to X$满足：

1. 对任意$x \in X$有$\phi(e,x) = x$, 其中e为G的单位元,
2. 对任意$x \in X$和$g_1, g_2 \in G$, $\phi(g_1, \phi(g_2, x)) = \phi(g_1 g_2, x)$成立,

那么我们就称(X, G, ϕ)为一个**动力系统**. 一般地, 我们也直接用(X, G)记一个动力系统. 易见此时对于每个$g \in G$, $\phi(g, \cdot) : X \to X$, $x \mapsto \phi(g, x)$为双射. 为方便计, 大多数情况下我们将$\phi(g, x)$简记为gx. 当X为独点集时, 我们称系统(X, G)为**平凡系统**.

如果$G = \mathbb{R}$为实数加群, 我们也称$(X, \mathbb{R})$为一个**流**. 如果$G = \mathbb{Z}$为整数加群, 那么称$(X, \mathbb{Z})$为一个**离散动力系统**.

设$T : X \to X$为一个双射, 我们可以定义$\phi : \mathbb{Z} \times X \to X$ 使得$\phi(n, x) = T^n(x)$. 于是$(X, \mathbb{Z}, \phi)$成为一个离散动力系统. 反之, 如果$(X, \mathbb{Z}, \phi)$ 为一个离散动力系统, 那么$T : X \to X$, $x \mapsto \phi(1, x)$ 为一个双射且对任意$x \in X$ 及$n \in \mathbb{Z}$ 我们有$\phi(n, x) = T^n(x)$. 正因为如此, 我们一般直接以(X, T) 表示离散动力系统.

类似地, 在上面定义中我们可以取G 为半群. 例如, 在上面的定义中我们以非负整数加法半群$\mathbb{Z}_+$ 来替代G, 那么称$(X, \mathbb{Z}_+, \phi)$ 为一个**半离散动力系统**. 在本书中, 这是我们讨论的主要对象. 同样的对于半离散系统, 我们用(X, T) 表示它, 其中$T : X \to X$ 为映射.

在研究中, 我们会赋予(X, G, ϕ) 更为具体的结构. 如果X 为拓扑空间、G 为拓扑群且ϕ为连续的, 那么称(X, G)为**拓扑动力系统**. 如果X 具有测度结构、G为拓扑群且ϕ为可测的, 那么称(X, G)为**测度动力系统**, 这就是本书需要研究的主要对象. 如果X具有微分结构, 那么(X, G)就是微分动力系统研究的对象等等. 遍历理论就是研究测度动力系统性质的理论.

在本章中, 我们将给出测度动力系统的基本概念和诸多例子, 希望通过这些例子使得读者对测度动力系统或遍历理论有一个初步印象. 最后, 我们将介绍Poincaré 回复定理, 这是遍历论最早的经典结论之一.

§1.1　基本概念

在本小节我们介绍测度动力系统的一些基本概念.

§1.1.1　保测系统

定义 1.1.1 设$(X_1, \mathcal{X}_1, \mu_1)$及$(X_2, \mathcal{X}_2, \mu_2)$为概率空间.

(a) 变换$T : X_1 \to X_2$称为**可测的***(measurable)*是指它满足$T^{-1}(\mathcal{X}_2) \subseteq \mathcal{X}_1$, 即对任意$B_2 \in \mathcal{X}_2$有$T^{-1}(B_2) \in \mathcal{X}_1$.

(b) 变换$T : X_1 \to X_2$称为**保测的***(measure preserving)*是指T 为可测的且对任意$B_2 \in \mathcal{X}_2$有

$$\mu_1(T^{-1}(B_2)) = \mu_2(B_2).$$

(c) 变换$T : X_1 \to X_2$称为**可逆保测变换**(*invertible*)是指T为双射,并且T和T^{-1}都是保测的.

由于我们主要讨论迭代T^n, 所以多数情况下我们仅对$(X_1, \mathcal{X}_1, \mu_1) = (X_2, \mathcal{X}_2, \mu_2)$的情形感兴趣.

定义 1.1.2 设$(X, \mathcal{X}, \mu)$为概率空间且$T : (X, \mathcal{X}, \mu) \to (X, \mathcal{X}, \mu)$为保测变换, 那么我们称四元组$(X, \mathcal{X}, \mu, T)$ 为**保测系统**. 此时称T **保持**μ 或者称μ 为T **不变的**.

在遍历论中有两种类型的问题. 第一种是所谓的"内在问题", 主要研究保测变换本身以及决定何时两个系统是"一样的", 即同构的(分类问题). 第二种问题是它在别的数学分支和其它自然科学中的应用. 现在我们先定义何时两个系统称为"一样的".

定义 1.1.3 设$(X_1, \mathcal{X}_1, \mu_1, T_1)$ 和$(X_2, \mathcal{X}_2, \mu_2, T_2)$ 为两个保测系统. 如果存在$M_i \in \mathcal{X}_i$ 使得$\mu_i(M_i) = 1$, $T_iM_i \subseteq M_i$ $(i = 1, 2)$ 以及存在保测变换$\phi : (M_1, \mathcal{X}_1|_{M_1}, \mu_1|_{M_1}) \to (M_2, \mathcal{X}_2|_{M_2}, \mu_2|_{M_2})$ 满足$\phi \circ T_1(x) = T_2 \circ \phi(x)$, $\forall x \in M_1$,

$$\begin{array}{ccc} M_1 & \xrightarrow{T_1} & M_1 \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ M_2 & \xrightarrow[T_2]{} & M_2 \end{array}$$

那么我们称$(X_2, \mathcal{X}_2, \mu_2, T_2)$ 为$(X_1, \mathcal{X}_1, \mu_1, T_1)$ 的**因子**, 或者称$(X_1, \mathcal{X}_1, \mu_1, T_1)$ 为$(X_2, \mathcal{X}_2, \mu_2, T_2)$的**扩充**.

如果ϕ为一个可逆保测变换, 那么我们就称$(X_1, \mathcal{X}_1, \mu_1, T_1)$ 和$(X_2, \mathcal{X}_2, \mu_2, T_2)$ **同构**, 或者直接称T_1**同构于**T_2.

根据定义1.1.1去判定一个变换是否为保测的并非一件容易的事情. 在具体验证所涉及的例子的时候, 我们经常使用下面的引理.

引理 1.1.4 设$(X_1, \mathcal{X}_1, \mu_1)$和$(X_2, \mathcal{X}_2, \mu_2)$为概率空间, $T : X_1 \to X_2$为变换且$\mathcal{S}_2$为生成$\mathcal{X}_2$的半代数. 如果对每个$A_2 \in \mathcal{S}_2$, 都有$T^{-1}(A_2) \in \mathcal{X}_1$且$\mu_1(T^{-1}(A_2)) = \mu_2(A_2)$, 那么$T$为保测的.

证明. 设$\mathcal{C}_2 = \{B \in \mathcal{X}_2 : T^{-1}(B) \in \mathcal{X}_1, \mu_1(T^{-1}B) = \mu_2(B)\}$. 下证$\mathcal{C}_2 = \mathcal{X}_2$. 因为$\mathcal{S}_2 \subseteq \mathcal{C}_2$并且由$\mathcal{S}_2$ 生成的代数$\mathcal{A}(\mathcal{S}_2)$中的元为有限个$\mathcal{S}_2$中元的无交并, 所以我们有$\mathcal{A}(S_2) \subseteq \mathcal{C}_2$. 易见$\mathcal{C}_2$为单调类, 根据$\mathcal{A}(\mathcal{S}_2)$生成的$\sigma$ 代数为由$\mathcal{A}(\mathcal{S}_2)$生成的单调类的事实, 我们易得结论. □

设$(X_1, \mathcal{X}_1, \mu_1, T_1)$和$(X_2, \mathcal{X}_2, \mu_2, T_2)$为两个保测系统. 设$(X_1 \times X_2, \mathcal{X}_1 \times \mathcal{X}_2, \mu_1 \times \mu_2)$为乘积空间, 定义

$$T_1 \times T_2 : X_1 \times X_2 \to X_1 \times X_2, (x_1, x_2) \mapsto (T_1x_1, T_2x_2).$$

我们易验证$(X_1 \times X_2, \mathcal{X}_1 \times \mathcal{X}_2, \mu_1 \times \mu_2, T_1 \times T_2)$ 仍为一个保测系统, 称之为两个系统的**乘积系统**. 同理可以定义任意多个系统的乘积系统.

§1.1.2 Koopman算子

设$(X,\mathcal{X},\mu)$为概率空间. 在本书中如不特别指出, 我们均假设$\mathcal{X}$ 为可数生成的. 这等价于作为度量空间, Hilbert空间$L^2(X,\mathcal{X},\mu)$为可分的.

对任意概率空间$(X,\mathcal{X},\mu)$和$p\geq 1$, 都可定义相应的Banach空间

$$L^p(\mu)=L^p(X,\mathcal{X},\mu)=\{f:X\to\mathbb{C}: f \text{ 可测且 } \int_X |f|^p d\mu<\infty\}.$$

这些空间上的理论是处理可测空间问题的一类重要工具. 设$L^0(X,\mathcal{X},\mu)$ 为$(X,\mathcal{X},\mu)$上可测复值函数全体的集合(在涉及泛函空间的时候, 两个函数相等是指它们几乎处处相等) 并且以$L^0_{\mathbb{R}}(X,\mathcal{X},\mu)$记$L^0(X,\mathcal{X},\mu)$实值函数全体.

定义 1.1.5 设$(X_1,\mathcal{X}_1,\mu_1)$ 及$(X_2,\mathcal{X}_2,\mu_2)$ 为概率空间, $T:X_1\to X_2$ 为保测映射. 诱导算子$U_T:L^0(X_2,\mathcal{X}_2,\mu_2)\to L^0(X_1,\mathcal{X}_1,\mu_1)$定义为

$$(U_Tf)(x)=f(Tx),\ \forall f\in L^0(X_2,\mathcal{X}_2,\mu_2),\ x\in X_1.$$

称U_T为***Koopman*算子**(*Koopman operator*).

我们首先罗列U_T的一些简单的性质:

(1) U_T 为线性的且 $U_TL^0_{\mathbb{R}}(X_2,\mathcal{X}_2,\mu_2)\subseteq L^0_{\mathbb{R}}(X_1,\mathcal{X}_1,\mu_1)$.

(2) $U_T(f\cdot g)=(U_Tf)(U_Tg), U_Tc=c$, 其中$c$ 为常值函数.

(3) U_T 为正算子, 即如果$f\geq 0$, 则$U_Tf\geq 0$.

(4) $U_T\mathbf{1}_B=\mathbf{1}_{T^{-1}B},\forall B\in\mathcal{X}_2$.

(5) 设$T:X_1\to X_2, S:X_2\to X_3$ 为保测映射, 那么$U_{S\circ T}=U_T\circ U_S$.

于是, 如果$T:X\to X$ 为保测的, 那么就有

$$U_{T^n}=U_T^n,\ \forall n\in\mathbb{Z}_+.$$

引理 1.1.6 假设$(X_1,\mathcal{X}_1,\mu_1)$ 及$(X_2,\mathcal{X}_2,\mu_2)$ 为概率空间, $T:X_1\to X_2$ 为保测映射. 对于$F\in L^0(X_2,\mu_2)$, 我们有

$$\int_{X_1} U_TFd\mu_1=\int_{X_2}Fd\mu_2.$$

在上面式子中, 如果一边积分不存在或者为无穷, 则另一边亦然.

证明. 首先设$F=\mathbf{1}_B, B\in\mathcal{X}_2$, 则

$$\int_{X_1} U_TFd\mu_1=\int_{X_1}U_T1_Bd\mu_1=\int_{X_1}\mathbf{1}_{T^{-1}B}d\mu_1=\mu_1(T^{-1}B)=\mu_2(B)=\int_{X_2}Fd\mu_2.$$

由此命题对简单函数成立. 于是根据标准的逼近方法, 我们可以证明引理, 细节留给读者. □

根据引理1.1.6, 我们有如下常用推论:

推论 1.1.7 假设$(X,\mathcal{X},\mu)$ 为概率空间, $T:X\to X$ 为可测映射. 那么T 为保持测度μ 的当且仅当对于$f\in L^\infty(X,,\mathcal{X},\mu)$, 我们有

$$\int_X U_T f d\mu = \int_X f d\mu.$$

定理 1.1.8 假设$(X_1,\mathcal{X}_1,\mu_1)$ 及$(X_2,\mathcal{X}_2,\mu_2)$ 为概率空间, $T:X_1\to X_2$ 为保测映射. 对任意的$p\ge 1$,

$$U_T L^p(X_2,\mu_2)\subseteq L^p(X_1,\mu_1).$$

并且U_T为等距算子, 即对任意的$f\in L^p(X_2,\mu_2)$,

$$\|U_T f\|_p = \|f\|_p.$$

同时还有

$$U_T L^p_{\mathbb{R}}(X_2,\mu_2)\subseteq L^p_{\mathbb{R}}(X_1,\mu_1).$$

证明. 设$f\in L^p(X_2,\mu_2)$, 在引理1.1.6中令$F(x)=|f(x)|^p$就得到$\|U_T f\|_p=\|f\|_p$. 由此即得证明. $\square$

设$(X,\mathcal{X},\mu,T)$是一个保测系统, 则Koopman算子$U_T:L^p(X,\mu)\to L^p(X,\mu)$为等距算子, 尤其当$p=2$并且$T$ 为可逆保测的时候, U_T 是一个酉算子. 研究U_T的理论称为T的**谱理论**, 后面我们会继续研究它.

§1.1.3 逆极限与自然扩充

定义 1.1.9 (逆极限) 设$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i=0}^\infty$ 为保测系统族, 对于任何$i\ge j$ 存在因子映射$\phi_{ij}:X_i\to X_j$ 使得$\phi_{ii}=\mathrm{id}$ 并且

$$\phi_{jk}\circ\phi_{ij}=\phi_{ik},\ \forall i\ge j\ge k.$$

$$\begin{tikzcd}
X_i \arrow[dr, "\phi_{ij}"] \arrow[dd, "\phi_{ik}"'] & \\
& X_j \arrow[dl, "\phi_{jk}"] \\
X_k &
\end{tikzcd}$$

定义$\prod_{i=0}^\infty X_i$ 的子集X 为

$$X=\left\{x=(x_i)_{i=0}^\infty\in\prod_{i=0}^\infty X_i:\phi_{ij}x_i=x_j,\ \forall i\ge j\right\}.$$

对$i\in\mathbb{Z}_+$, 设$\pi_i:X\to X_i,\pi_i(x)=x_i$ 为投射. 易见它满足$\phi_{ij}\circ\pi_i=\pi_j,\forall i\ge j$.

$$\begin{tikzcd}
& X \arrow[dl, "\pi_i"'] \arrow[dr, "\pi_j"] & \\
X_i \arrow[rr, "\phi_{ij}"] & & X_j
\end{tikzcd}$$

设$\mathcal{X}$ 为由$\{\pi_i^{-1}\mathcal{X}_i\}_{i=0}^{\infty}$ 生成的σ-代数. 定义$\mu:\bigcup_{i=0}^{\infty}\pi_i^{-1}(\mathcal{X}_i)\to\mathbb{R}_+$ 满足$\mu(\pi_i^{-1}E)=\mu_i(E)$, $\forall E\in\mathcal{X}_i, i\in\mathbb{Z}_+$. 再将$\mu$ 延拓到$\mathcal{X}$ 上. 定义$T:X\to X$使得

$$T\big((x_i)_{i=0}^{\infty}\big)=(T_ix_i)_{i=0}^{\infty},\ \forall (x_i)_{i=0}^{\infty}\in X.$$

明显的,

$$\pi_i:(X,\mathcal{X},\mu,T)\to(X_i,\mathcal{X}_i,\mu_i,T_i)$$

为因子映射. 称$(X,\mathcal{X},\mu,T)$ 为$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i=0}^{\infty}$ 的**逆极限系统***(inverse limit)*. 记为

$$(X,\mathcal{X},\mu,T)=\varprojlim_{i\to\infty}(X_i,\mathcal{X}_i,\mu_i,T_i),$$

或

$$(X,\mathcal{X},\mu,T)=\mathrm{inv}\ \lim_{i\to\infty}(X_i,\mathcal{X}_i,\mu_i,T_i).$$

以上构造也可推广到一般的有向集指标的系统族.

将半离散系统与离散系统联系在一起的一个桥梁是所谓的**自然扩充**.

定义 1.1.10 (自然扩充) 设$(X,\mathcal{X},\mu,T)$ 为保测系统, 如果T 不可逆, 我们可以通过自然扩充使其变成可逆系统的因子. 一个系统$(\widetilde{X},\widetilde{\mathcal{X}},\widetilde{\mu},\widetilde{T})$ 称为$(X,\mathcal{X},\mu,T)$ 的**自然扩充***(natural extension)*, 是指$\widetilde{T}$ 可逆, 并且存在因子映射$\phi:(\widetilde{X},\widetilde{\mathcal{X}},\widetilde{\mu},\widetilde{T})\to(X,\mathcal{X},\mu,T)$ 使得$\phi^{-1}\mathcal{X}\underset{\widetilde{\mu}}{=}\widetilde{\mathcal{X}}$.

设$(X_i,\mathcal{X}_i,\mu_i,T_i)=(X,T^{-i}\mathcal{X},\mu,T), \phi_{ij}=T^{i-j}$, 那么$\varprojlim_{i\to\infty}(X_i,\mathcal{X}_i,\mu_i,T_i)$ 为$(X,\mathcal{X},\mu,T)$ 的自然扩充. 具体地, 另

$$\widetilde{X}=\Big\{(x_1,x_2,\dots)\in\prod_{i=1}^{\infty}X:Tx_{i+1}=x_i,\ i\geq 1\Big\}.$$

它为乘积空间$\prod_{i=1}^{\infty}X$ 的子集. 定义

$$\widetilde{T}:\widetilde{X}\to\widetilde{X},\ \widetilde{T}(x_1,x_2,\dots)=(Tx_1,Tx_2,Tx_3,\dots)=(Tx_1,x_1,x_2,\dots)$$

那么$\widetilde{T}$ 为可逆保测变换. 事实上,

$$\widetilde{T}^{-1}(x_0,x_1,x_2,\dots)=(x_1,x_2,\dots),\ \forall (x_0,x_1,x_2,\dots)\in\widetilde{X}.$$

易见对每个$n\in\mathbb{N}$, 向第n 分量的投影映射$p_n:\widetilde{X}\to X$ 为因子映射. 尤其p_1 为$(\widetilde{X},\widetilde{T})$ 到(X,T) 的因子映射.

$$\begin{tikzcd}
\widetilde{X} \arrow[r, "\widetilde{T}"] \arrow[d, "p_1"'] & \widetilde{X} \arrow[d, "p_1"] \\
X \arrow[r, "T"'] & X
\end{tikzcd}$$

习 题

1. 给出引理1.1.4的完整证明.
2. 给出引理1.1.6的完整证明.
3. 设$(X_1,\mathcal{X}_1,\mu_1)$ 和$(X_2,\mathcal{X}_2,\mu_2)$ 为两个概率空间，$T_1: X_1\to X_1, T_2: X_2\to X_2$ 为可测变换. 如果存在$M_i\in\mathcal{X}_i$ 使得$\mu_i(M_i)=1$, $T_iM_i\subseteq M_i$ $(i=1,2)$ 以及存在可逆保测变换$\phi: M_1\to M_2$ 满足$\phi\circ T_1(x)=T_2\circ\phi(x), \forall x\in M_1$, 证明：$(X_1,\mathcal{X}_1,\mu_1,T_1)$ 为保测系统但且仅当$(X_2,\mathcal{X}_2,\mu_2,T_2)$ 为保测系统.

§1.2　一些例子

在这一小节, 我们介绍常见的一些保测动力系统, 它们之中有些自然也为拓扑动力系统. 所谓拓扑动力系统是指二元组(X,T)，其中X 为拓扑空间，$T: X\to X$ 为连续映射. 称(X,T) 为极小的是指任意$x\in X$ 的轨道$\{T^n x: n\in\mathbb{Z}_+\}$ 都在X中稠密. 我们在第五章详细介绍拓扑动力系统.

§1.2.1　有限系统

例 1.2.1 (平凡系统) $(X,\mathcal{N},\mu,T)$为保测系统, 其中$\mathcal{N}=\{X,\emptyset\}$, $T=\mathrm{id}: X\to X$. 对于平凡系统, 空间的选取并不重要性, 主要在于σ代数的平凡.

例 1.2.2 (n-周期系统) 设$n\in\mathbb{N}$且$\mathbb{Z}_n=\mathbb{Z}/n\mathbb{Z}=\{0,1,\dots,n-1\}$. 令$T\colon\mathbb{Z}_n\to\mathbb{Z}_n$, $x\mapsto x+1 \pmod n$. 取每个点测度为$\frac{1}{n}$的平均测度μ, 那么$(\mathbb{Z}_n,\mathcal{P}(\mathbb{Z}_n),\mu,T)$ 为保测系统. 此系统也称为n周期系统. $n=1$即为平凡系统.

§1.2.2　仿射系统

例 1.2.3 (圆周旋转) 设$\mathbb{T}=\mathbb{R}/\mathbb{Z}=[0,1) \pmod 1$ 且$\alpha\in\mathbb{T}$. 令

$$R_\alpha\colon\mathbb{T}\to\mathbb{T},\ x\mapsto x+\alpha \pmod 1.$$

因为平移不改变区间长度, 所以$(\mathbb{T},\mathcal{B}(\mathbb{T}),m_{\mathbb{T}},R_\alpha)$为保测系统, 这里$m_{\mathbb{T}}$为$\mathbb{T}$上的*Lebesque*测度.

设$\mathbb{S}^1=\{z\in\mathbb{C}:|z|=1\}$为复平面上的单位圆周, α为无理数, $a=e^{2\pi i\alpha}$. 定义

$$T_\alpha:\mathbb{S}^1\to\mathbb{S}^1, z=e^{2\pi i\theta}\mapsto az=e^{2\pi i(\theta+\alpha)},\ \forall z\in\mathbb{S}^1.$$

同样, $(\mathbb{S}^1,\mathcal{B}(\mathbb{S}^1),m_{\mathbb{S}^1},T_\alpha)$为保测系统, 这里$m_{\mathbb{S}^1}$为$\mathbb{S}^1$上的*Lebesque*测度. 易见, $\phi:\mathbb{T}\to\mathbb{S}^1, x\mapsto e^{i2\pi x}$是同胚, 并且为$(\mathbb{T},\mathcal{B}(\mathbb{T}),m_{\mathbb{T}},R_\alpha)$ 和$(\mathbb{S}^1,\mathcal{B}(\mathbb{S}^1),m_{\mathbb{S}^1},T_\alpha)$之间的同构.

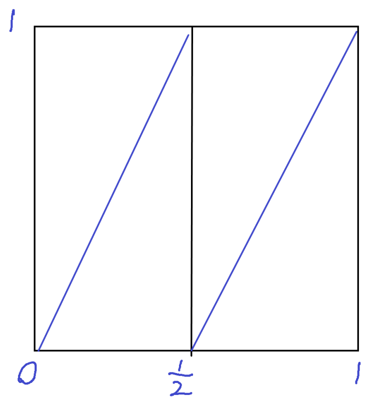
另外圆周无理旋转系统为极小的拓扑动力系统, 即每个点的轨道是稠密的. 因为α 为无理数, 所以$\{a^n: n\in\mathbb{Z}_+\}$ 为无穷集. $\mathbb{S}^1$ 是紧致的, 所以$\{a^n: n\in\mathbb{Z}_+\}$ 存在聚点. 于是对于任何$\varepsilon>0$, 存在$n,k\in\mathbb{N}$ 使得$|a^n-a^{n+k}|<\varepsilon$. 由于$z\mapsto a^kz$ 为等距映射, 我们有$|a^{n+(m+1)k}-a^{n+mk}|=|a^n-a^{n+k}|<\varepsilon, \forall m\in\mathbb{Z}_+$. 由此, $\{a^{n+mk}: m\in\mathbb{Z}_+\}$ 在$\mathbb{S}^1$ 中是ε-

稠密的. 注意$\{a^{n+mk}: m\in\mathbb{Z}_+\}$ 包含在1 的轨道$\{a^n: n\in\mathbb{Z}_+\}$ 中. 综上, 对于任何$\varepsilon>0$, $\{a^n: n\in\mathbb{Z}_+\}$ 在$\mathbb{S}^1$ 中ε- 稠密, 所以在$\mathbb{S}^1$ 中稠密. 因为是旋转作用, 根据此我们得到任何点的轨道稠密, 即圆周无理旋转系统为极小的.

根据这个结论, 我们实际得到如下常用结论的一个证明: $\{n\alpha-[n\alpha]: n\in\mathbb{Z}_+\}$在$[0,1]$中稠密, 其中$[\cdot]$指实数$\cdot$的整数部分.

例 1.2.4 (加倍映射) 定义区间$[0,1]$ 上的映射$\widetilde{T}_2\colon [0,1]\to[0,1]$ 如下:

$$\widetilde{T}_2x=\begin{cases}2x, & 0\le x<1/2;\\ 2x-1, & 1/2\le x\le 1.\end{cases}$$



加倍映射

下证明$([0,1],\mathcal{B}([0,1]),m,\widetilde{T}_2)$为保测系统, 其中$m$ 为*Lebesgue* 测度. 设$B=[a,b)$, 则

$$\widetilde{T}_2^{-1}(B)=[\frac{a}{2},\frac{b}{2})\cup[\frac{a}{2}+\frac{1}{2},\frac{b}{2}+\frac{1}{2})$$

于是

$$m(\widetilde{T}_2^{-1}(B))=\frac{1}{2}(b-a)+\frac{1}{2}(b-a)=b-a=m(B).$$

根据引理*1.1.4*, $([0,1],\mathcal{B}([0,1]),m,\widetilde{T}_2)$ 为保测系统. 因为$\widetilde{T}_2:[0,1]\to[0,1]$ 不连续, 所以它不为拓扑动力系统.

但是如果我们将$[0,1]$ 换为$\mathbb{T}$, 那么我们在不改变测度性质的同时将系统变为拓扑动力系统. 设$T_2\colon\mathbb{T}\to\mathbb{T}$, $x\mapsto 2x \pmod 1$, 以及$m_{\mathbb{T}}$ 为$\mathbb{T}$ 的*Lebesgue* 测度. 易见, 在$[0,1)$ 上, T_2 和$\widetilde{T}_2$ 吻合, 于是$(\mathbb{T},\mathcal{B}(\mathbb{T}),m_{\mathbb{T}},T_2)$ 与$([0,1],\mathcal{B}([0,1]),m,T_2)$ 同构, 从而也为保测系统. 另外作为映射$T_2:\mathbb{T}\to\mathbb{T}$它为连续的, 所以$(\mathbb{T},T_2)$ 为拓扑动力系统.

一般地, 对于$n\ge 2$可以定义$T_n\colon\mathbb{T}\to\mathbb{T}$, $x\mapsto nx \pmod 1$. $(\mathbb{T},\mathcal{B}(\mathbb{T}),m_{\mathbb{T}},T_n)$为保测系统.

例 1.2.5 (圆周自同态) 设$n \in \mathbb{N}$, 令$S_n\colon \mathbb{S}^1 \to \mathbb{S}^1$, $z \mapsto z^n$. 容易验证$\phi : \mathbb{T} \to \mathbb{S}^1, x \mapsto e^{i2\pi x}$是$(\mathbb{T}, \mathcal{B}(\mathbb{T}), m_{\mathbb{T}}, T_n)$ 与$(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), m_{\mathbb{S}^1}, S_n)$ 的同构映射, 尤其$(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), m_{\mathbb{S}^1}, S_n)$ 为保测系统, 其中$m_{\mathbb{S}^1}$ 为$\mathbb{S}^1$ 上的*Haar* 测度.

S_n 为圆周上的自同态，并且根据定理*0.4.11*, $\widehat{\mathbb{S}^1} = \{S_n : n \in \mathbb{Z}\}$.

问题 1.2.6 (Furstengerg $\times_2, \times_3$ 猜测) *1967*年*Furstenberg*提出了如下猜测*(*现称为*Furstengerg* $\times_2, \times_3$猜测[1]*):* 设n, m为两个乘积独立的自然数（即$\log n/\log m$ 为无理数）, 如果μ 是$\mathbb{T}$ 上一个*Borel*概率测度且在$\times_n, \times_m$ 生成的半群作用下遍历*(*遍历定义见第二章*)*, 则μ 要么为$\mathbb{T}$ 上的*Lebesgue* 测度, 要么为等分布在有限个点上的离散测度*[62]*.

例 1.2.7 (环面旋转) 对$k \in \mathbb{N}$, 取$\vec{\theta} = (\theta_1, \ldots, \theta_k) \in \mathbb{R}^k$. 令

$$T_{\vec{\theta}}\colon \mathbb{T}^k \to \mathbb{T}^k, \vec{x} = (x_1, x_2, \ldots, x_k) \mapsto \vec{x} + \vec{\theta} = (x_1 + \theta_1, x_2 + \theta_2, \ldots, x_k + \theta_k).$$

$\mathbb{T}^k$上的*Lebesgue*测度m 是$(\mathbb{T}^k, T_{\vec{\theta}})$的一个不变测度, $(\mathbb{T}^k, \mathcal{B}, m, T_{\vec{\theta}})$ 为保测系统.

例 1.2.8 (加法机器) 设$\Sigma_2 = \{0, 1\}^{\mathbb{Z}_+}$, 我们定义$\Sigma_2$上的一个加法 "$+$"运算. 若$x = (x_0, x_1, \ldots)$和$y = (y_0, y_1, \ldots)$, 我们递归定义$z = x + y = (z_0, z_1, \ldots)$的每一个位置*:* 令$c_0 = 0$. 若$x_n + y_n + c_n \geq 2$, 令$z_n = x_n + y_n + c_n - 2$ 和$c_{n+1} = 1$, 否则令$z_n = x_n + y_n + c_n$ 和$c_{n+1} = 0$. 不难验证$(\Sigma_2, +)$是一个交换群. 进一步, 群的加法运算$(x, y) \mapsto x + y$和取逆运算$x \mapsto -x$都是连续的, 所以$(\Sigma_2, +)$是一个紧致可度量化的交换群. 容易验证$m = \mu_{(\frac{1}{2}, \frac{1}{2})}$ 为其上的*Haar* 测度.

令$\mathbf{1} = (1, 0, 0, \ldots)$ 和

$$R_{\mathbf{1}}\colon \Sigma_2 \to \Sigma_2, x \mapsto x + \mathbf{1}.$$

则$R_{\mathbf{1}}$是一个同胚. 称$(\Sigma_2, R_{\mathbf{1}})$是一个**加法机器***(adding machine* 或 *odometer)*. 因为m 为*Haar*测度，$(\Sigma_2, \mathcal{B}(\Sigma_2), m, R_1)$ 为保测系统.

例 1.2.9 (群旋转) 设G是一个紧致可度量群和$a \in G$. 定义群旋转

$$R_a\colon G \to G, x \mapsto ax.$$

则$(G, \mathcal{B}(G), m, R_a)$ 是一个可逆保测动力系统, 其中m 为*Haar* 测度. 我们一般称$(G, \mathcal{B}(G), m, R_a)$为**Kroneker系统**. 它也是一个拓扑动力系统.

例 1.2.10 (群自同态系统) 设G为紧致可度量群, $A : G \to G$为满的连续自同态映射, m为其上*Haar*测度. 那么$(G, \mathcal{B}(G), m, A)$为保测映射.

[1]该猜测的第一个本质进展是Lyons [151]在1988 年获得的， 他证明了在Furstengerg $\times_2, \times_3$ 猜测的条件下, 如果还要求μ 相对于$\times_n$ 具有Kolmogorov 属性（即$(\mathbb{T}, \mu, \times_n)$ 没有非平凡的零熵因子), 则μ 为上的Lebesgue 测度. 随后, 1990 年Roudolph [174] 证明了：设m, n 为两个互素的自然数， μ 是$\mathbb{T}$ 上一个Borel概率测度且在$\times_n, \times_m$ 生成的半群作用下遍历, 如果μ 相对于$\times_n$ 有正熵, 则μ 为$\mathbb{T}$ 上的Lebesgue 测度.

下面我们证明A为m不变的. 即要证明$m(A^{-1}E) = m(E)$, $\forall E \in \mathcal{B}(G)$. 令$\mu = A_* m$, 即$\mu(E) = m(A^{-1}(E))$. 首先因为$m$正则, μ 也正则. 因为A 为满射, 所以对于任何$y \in G$ 存在$x \in G$ 使得$y = Ax$. 设

$$R_x : G \to G, g \mapsto xg,\ R_y : G \to G, g \mapsto yg$$

为旋转. 那么容易验证$A \circ R_x = R_y \circ A$, 即

$$\begin{tikzcd} G \arrow[r, "R_x"] \arrow[d, "A"'] & G \arrow[d, "A"] \\ G \arrow[r, "R_y"'] & G. \end{tikzcd}$$

尤其, 对于$E \in \mathcal{B}(G)$, 有

$$x^{-1}A^{-1}E = (A \circ R_x)^{-1}(E) = (R_y \circ A)^{-1}(E) = A^{-1}(y^{-1}E).$$

所以

$$\mu(y^{-1}E) = m(A^{-1}(y^{-1}E)) = m(x^{-1}A^{-1}E) = m(A^{-1}E) = \mu(E).$$

所以μ为左不变的. 根据*Haar*测度唯一性, 我们得到$\mu = m$. 所以m为A不变的.

例 1.2.11 (仿射系统) 设G为紧致度量群, $a \in G$,$A : G \to G$为满的连续自同态映射, m 为其上*Haar*测度. 令

$$T : G \to G,\ x \mapsto a \cdot A(x).$$

则称T为**仿射变换***(affine)*. 那么$(G, \mathcal{B}(G), m, T)$为保测映射.

§1.2.3 斜积系统

在这里我们先介绍拓扑斜积系统，我们在后面第四章将详细介绍可测的斜积系统.

定义 1.2.12 设(Y, S)是一个拓扑动力系统. 设Z是一个紧致可度量化空间和令$C(Z, Z)$表示Z到自身所有连续映射的全体并赋予一致收敛拓扑. 设$\phi\colon Y \to C(Z, Z)$ 是一个连续映射. 令$X = Y \times Z$和

$$T\colon X \to X, (y, z) \mapsto (Sy, \phi(y)(z)).$$

则T是连续的, 从而(X, T)成为一个拓扑动力系统. 称之为(Y, S)与Z相对于ϕ的**斜积系统***(skew-product system)*.

设G是一个紧致可度量化的群和$\phi\colon Y \to G$是一个连续映射. 令$X = Y \times G$和

$$T\colon X \to X,\ (y, g) \mapsto (Sy, \phi(y)g).$$

则T是连续的. 从而(X, T)成为一个拓扑动力系统. 称之为(Y, S)相对于ϕ的**G-扩张***(G-extension)*.

易见G-扩张是斜积系统的一种特殊情况.

例 1.2.13 (环面映射) 设$\alpha\in\mathbb{R}$. 设

$$T\colon \mathbb{T}^2\to\mathbb{T}^2, (x,y)\mapsto (x+\alpha, x+y).$$

则$(\mathbb{T}^2,T)$是$(\mathbb{T},R_\alpha)$相对于$\phi\colon \mathbb{T}\to C(\mathbb{T},\mathbb{T})$, $x\mapsto R_x$, 的$\mathbb{T}$-扩张.

一般的, 设$\alpha\in\mathbb{R}$和$k\geq 2$,

$$T_k:\mathbb{T}^k\to\mathbb{T}^k, (x_1,x_2,\dots,x_k)\mapsto (x_1+\alpha, x_2+x_1,\dots,x_k+x_{k-1}).$$

记$\mathbb{T}$上的*Lebesgue*测度为m. 则$\mathbb{T}^k$上的*Lebesgue*测度为m与自身的k次乘积, 记为μ_k. 设A_i是$\mathbb{T}$上的可测集, $i=1,2,\dots,k$. 则

$$\mu_k\Big(\prod_{i=1}^k A_i\Big)=\prod_{i=1}^k m(A_i).$$

和

$$\begin{aligned}
\mu_k\Big(T_k^{-1}\Big(\prod_{i=1}^k A_i\Big)\Big) &= \int_{\mathbb{T}^k}\Big(\prod_{i=1}^k \mathbf{1}_{A_i}\Big)\circ T_k d\mu_k\\
&= \int_{\mathbb{T}^{k-1}}\Big(\prod_{i=1}^{k-1}\mathbf{1}_{A_i}\Big)\circ T_{k-1}\cdot\Big(\int_{\mathbb{T}}\mathbf{1}_{A_k}(x_k+x_{k-1})dm(x_k)\Big)d\mu_{k-1}\\
&= \int_{\mathbb{T}^{k-1}}\Big(\prod_{i=1}^{k-1}\mathbf{1}_{A_i}\Big)\circ T_{k-1}\cdot\Big(\int_{\mathbb{T}}\mathbf{1}_{A_k}(x_k)dm(x_k)\Big)d\mu_{k-1}\\
&= \int_{\mathbb{T}}\int_{\mathbb{T}}\cdots\int_{\mathbb{T}}\prod_{i=1}^k\mathbf{1}_{A_i}(x_i)dm(x_k)\dots dm(x_2)dm(x_1)\\
&= \prod_{i=1}^k m(A_i)=\mu_k\Big(\prod_{i=1}^k A_i\Big).
\end{aligned}$$

由乘积测度的构造, μ_k是T_k-不变的.

§1.2.4 符号系统

例 1.2.14 (单边符号系统) 设$k\geq 2$ 为自然数且记$A=\{0,1,\dots,k-1\}$. 称A为**字母表**和A中元素为**字母**. 令

$$\Sigma_k=A^{\mathbb{Z}_+}=\{(x_i)_{i\in\mathbb{Z}_+}=(x_0,x_1,x_2,\dots)\colon x_i\in A, i\in\mathbb{Z}_+\}.$$

习惯上, 我们也经常把Σ_k 中元素$(x_0,x_1,x_2,\dots)$ 直接写成$x_0x_1x_2\cdots$. 在本书中, 我们经常会混用两种记号.

赋予A离散拓扑和Σ_k乘积拓扑. 则Σ_k是紧致可度量化空间, 其中一个相容的度量可定义为:

$$d(x,y)=\begin{cases}0, & \text{如果}x=y,\\ \frac{1}{n+1}, & \text{如果}x\neq y,\ n=\min\{i\in\mathbb{Z}_+\colon x_i\neq y_i\},\end{cases}$$

其中$x = x_0x_1\dots$, $y = y_0y_1\dots$.

定义转移映射

$$\sigma\colon \Sigma_k \to \Sigma_k,\ x_0x_1x_2\dots \mapsto x_1x_2x_3\dots$$

则σ是一个连续满射, (Σ_k,σ) 为拓扑动力系统. 称(Σ_k,σ) 是k个符号上的**全转移**(*full shift*)或**符号系统**(*shift system*). 如果$X \subseteq \Sigma_k$是一个σ不变非空闭子集(σ不变指$\sigma X \subseteq X$), 则称(X,σ)是**子转移**(*subshift*)或**符号子系统**(*shift subsystem*).

设$n \in \mathbb{N}$, 令

$$A^n = \{x_0x_1\dots x_{n-1}\colon x_i \in A,\ 0 \le i \le n-1\}.$$

称A^n中元素为一个长为n的**词**(*word*)或**块**(*block*). Σ_k 中元素也称为**无穷词**. 令

$$A^* = \bigcup_{n=1}^{\infty} A^n,$$

即A^*是所有有限词构成的集合. 设$u = u_0u_1\dots u_{n-1}$和$v = v_0v_1\dots v_{m-1}$ 是两个有限词, 定义它们的**连接**(*concatenation*)uv为词$u_0u_1\dots u_{n-1}v_0v_1\dots v_{m-1}$. 一个词$u$的**长度**(*length*)记为$|u|$. 若$|u| = n$和$|v| = m$, 则$|uv| = n+m$. 为方便起见, 分别记$uu\cdots u$($m$-个)和$uu\cdots u\cdots$为$u^m$和$u^\infty$.

设$x = x_0x_1\dots$和$0 \le m \le n$, 令$x[m,n] = x_mx_{m+1}\dots x_n$. 若$x[i,i+|u|-1] = u$, 称**词$u$在$x$的第$i$个位置出现**.

设u是一个有限词, 定义**柱形集**(*cylinder*)

$$[u] = \{x \in \Sigma_k\colon x[0,|u|-1] = u\}.$$

易见每个柱形集既是开集又是闭集, 并且不难验证$\{[u]\colon u \in A^*\}$是Σ_k的一组拓扑基.

设$k \ge 2$和$\vec{p} = (p_0,p_1,\dots,p_{k-1})$是一个概率向量. 下面验证$\mu_{\vec{p}}$是$(\Sigma_k,\sigma)$上的一个不变测度. 于是$(\Sigma_k,\mathcal{B},\mu_{\vec{p}},\sigma)$为保测系统. 称为**单边$\vec{p}$-转移系统**.

设$u = u_0u_1\dots u_{n-1}$是一个长为n的有限词. 则

$$\mu_{\vec{p}}([u]) = \prod_{i=0}^{n-1} p_{u_i}$$

和

$$\mu_{\vec{p}}(\sigma^{-1}([u])) = \mu_{\vec{p}}\Big(\bigcup_{j=0}^{k-1}[ju]\Big) = \sum_{j=0}^{k-1}\mu_{\vec{p}}([ju]) = \Big(\sum_{j=0}^{k-1}p_j\Big)\prod_{i=0}^{n-1}p_{u_i} = \prod_{i=0}^{n-1}p_{u_i}.$$

所有柱形集构成一个生成$\mathcal{B}(\Sigma_k)$的半代数, 所以$\mu_{\vec{p}}$是(Σ_k,σ)上的一个不变测度.

定义

$$\phi : \Sigma_2 \to \mathbb{T}, (x_0,x_1,x_2,\dots) \mapsto \sum_{n=0}^{\infty}\frac{x_n}{2^{n+1}}.$$

则ϕ为保测映射, 且图表交换.

$$\begin{tikzcd} \Sigma_2 \arrow[r, "\sigma"] \arrow[d, "\phi"'] & \Sigma_2 \arrow[d, "\phi"] \\ \mathbb{T} \arrow[r, "T_2"'] & \mathbb{T}. \end{tikzcd}$$

ϕ^{-1}仅在可数集合上没有定义,所以它为测度同构. 于是$(\Sigma_2, \mathcal{B}(\Sigma_2), \mu_{(1/2,1/2)}, \sigma)$ 与$(\mathbb{T}, \mathcal{B}(\mathbb{T}), m_{\mathbb{T}}, T_2)$是同构的系统.

例 1.2.15 设

$$\mu([a,b]) = \int_a^b \frac{dx}{\pi(1+x^2)}.$$

令

$$T : \mathbb{R} \setminus \{0\} \to \mathbb{R}, x \mapsto \frac{1}{2}\left(x - \frac{1}{x}\right),$$

以及$T(0) = 0$. 易见

$$\int_{\mathbb{R}} f(Tx) d\mu = \int_{\mathbb{R}} f(x) d\mu, \ \forall f \in L^\infty(\mathbb{R}, \mu).$$

于是$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu, T)$为保测系统.

令

$$\psi : \mathbb{R} \to \mathbb{T}, \ x \mapsto \frac{1}{\pi}\arctan(x) + \frac{1}{2}.$$

则$\psi : (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu, T) \to (\mathbb{T}, \mathcal{B}(\mathbb{T}), m_{\mathbb{T}}, T_2)$为测度同构.

$$\begin{tikzcd} \mathbb{R} \arrow[r, "T"] \arrow[d, "\psi"'] & \mathbb{R} \arrow[d, "\psi"] \\ \mathbb{T} \arrow[r, "T_2"'] & \mathbb{T}. \end{tikzcd}$$

注记 1.2.16 综合之前的例子，我们得到$(\mathbb{T}, \mathcal{B}(\mathbb{T}), m_{\mathbb{T}}, T_2)$, $(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), m_{\mathbb{S}^1}, S_2)$, $(\Sigma_2, \mathcal{B}(\Sigma_2), \mu_{(1/2,1/2)}, \sigma)$以及$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu, T)$ 是同构的系统. 由此可以看出，同构的系统形式上可以相差很大.

例 1.2.17 (双边符号系统) 设$k \geq 2$ 为自然数且记$A = \{0, 1, \ldots, k-1\}$. 仍记

$$\Sigma_k = A^{\mathbb{Z}} = \{(x_i)_{i \in \mathbb{Z}} = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots) \colon x_i \in A, i \in \mathbb{Z}\}.$$

我们经常会把Σ_k 中元素$(\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$ 记为$\ldots x_{-2}x_{-1}\dot{x}_0 x_1 x_3 \ldots$, 其中在一个字符上加“$\cdot$”表示第0个坐标位置.

同样定义转移映射

$$\sigma \colon \Sigma_k \to \Sigma_k, \ \ldots x_{-2}x_{-1}\dot{x}_0 x_1 x_3 \ldots \mapsto \ldots x_{-1}x_0\dot{x}_1 x_2 x_3 \ldots$$

则σ是一个同胚. 称(Σ_k, σ)是k个符号上的**双边全转移**或**双边符号系统**.

给定概率向量$\vec{p}$, 类似于单边符号系统, 可验证双边转移系统$(\Sigma_k, \mathcal{B}, \mu_{\vec{p}}, \sigma)$ 为可逆保测系统, 称为**双边$\vec{p}$- 转移系统**.

例 1.2.18 (Markov转移) 设$k \geq 2$ 为自然数, $A = \{0, 1, \ldots, k-1\}$, 令$(\Sigma_k, \mathcal{B}) = \prod_{i=-\infty}^{\infty}(A, \mathcal{P}(A))$. 对于每个$n \geq 0$ 以及$a_0, a_1, \ldots, a_n \in A$, 给定了某个非负实值函数$p_n(a_0, \ldots, a_n)$ 满足条件

$$\sum_{a_0 \in Y} p_0(a_0) = 1,$$

$$p_n(a_0, \ldots, a_n) = \sum_{a_{n+1} \in Y} p_{n+1}(a_0, \ldots, a_n, a_{n+1}).$$

那么在$(\Sigma_k, \mathcal{B})$上存在唯一的概率测度μ使得对任意的$h \leq l$有

$$\mu({}_h[a_h, a_{h+1}, \ldots, a_l]_l) = p_{l-h}(a_h, \ldots, a_l), \quad \forall a_i \in A, h \leq i \leq l.$$

可以验证此时$(\Sigma_k, \mathcal{B}, \mu, \sigma)$为保测系统.

对于概率向量$\vec{p} = (p_0, \ldots, p_{k-1})$, 令

$$p_n(a_0, a_1, \ldots, a_n) = p_{a_0} p_{a_1} \ldots p_{a_n},$$

我们得到双边$\vec{p}$-转移系统.

设$P = [p_{ij}]$ $(p_{ij} \geq 0, \sum_{j=0}^{k-1} p_{ij} = 1, \forall i, j \in \{0, 1, \ldots, k-1\})$ 为随机矩阵, $\vec{p} = (p_0, \ldots, p_{k-1})$为概率向量, 并且$\vec{p}P = \vec{p}$ (即$\sum_{i=0}^{k-1} p_i p_{ij} = p_j$). 我们令

$$p_n(a_0, \ldots, a_n) = p_{a_0} p_{a_0 a_1} p_{a_1 a_2} \ldots p_{a_{n-1} p_n}.$$

容易验证$\{p_n\}_n$满足上面定理条件, 由此决定了一个测度μ, 此时我们称保测系统为 **双边$(\vec{p}, P)$-*Markov* 转移**.

注意如果根据$p_{ij} = p_j, \forall i, j \in \{0, 1, \ldots, k-1\}$定义随机矩阵, 那么此时双边$(\vec{p}, P)$-*Markov* 转移就是双边$\vec{p}$-转移.

根据上面的双边情况, 类似可以得到单边的系统, 我们不再赘述.

§1.2.5 保测系统与概率论

例 1.2.19 (平稳随机过程) 设$(\Omega, \mathcal{F}, P)$ 为概率空间, $\ldots, f_{-1}, f_0, f_1, f_2, \ldots$ 为Ω 上一列可测函数. 设$\{f_n\}_{n \in \mathbb{Z}}$ 为平稳的*(Stationary)*, 即对于任何$n_1, n_2, \ldots, n_r \in \mathbb{Z}, r \in \mathbb{N}$, 任何*Borel* 集合$B_1, B_2, \ldots, B_r \subseteq \mathbb{R}$, 以及任何$k \in \mathbb{Z}$, 有

$$P\{\omega : f_{n_1}(\omega) \in B_1, \ldots, f_{n_r}(\omega) \in B_r\} = P\{\omega : f_{n_1+k}(\omega) \in B_1, \ldots, f_{n_r+k}(\omega) \in B_r\}.$$

令$\mathbb{R}^{\mathbb{Z}}$为乘积空间,

$$\phi : \Omega \to \mathbb{R}^{\mathbb{Z}}, \omega \mapsto (\ldots, f_{-1}(\omega), f_0(w), f_1(w), f_2(\omega), \ldots),$$

即$(\phi(\omega))_n = f_n(\omega), \forall n \in \mathbb{Z}$. 取$\mathbb{R}^{\mathbb{Z}}$上测度如下:

$$\mu(E) = P(\phi^{-1}E), \ \forall E \in \mathcal{B}(\mathbb{R}^{\mathbb{Z}}).$$

定义$\sigma : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ 为转移变换：

$$(\sigma \mathbf{x})_n = x_{n+1}, \forall \mathbf{x} = (x_n)_{n\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}, n \in \mathbb{Z}.$$

由此得到的系统$(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}(\mathbb{R}^{\mathbb{Z}}), \mu, \sigma)$ 为保测系统. 设$\pi_n : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}, \mathbf{x} = (x_j)_{j\in\mathbb{Z}} \mapsto x_n, n \in \mathbb{Z}$ 为到第n个坐标的投射, 那么$\{\pi_n\}_{n\in\mathbb{Z}}$ 在$\mathbb{R}^{\mathbb{Z}}$上的联合分布与$\{f_n\}_{n\in\mathbb{Z}}$ 在Ω 上联合分布相同. 于是每一个平稳随机过程可以从$\mathbb{R}^{\mathbb{Z}}$的一个转移不变的测度中得到.

习　题

1. 设$f : [0,1] \to [0,1]$, $x \mapsto 1 - |2x-1|$, m 为Lebesgue 测度. 则$([0,1], \mathcal{B}([0,1]), m, T)$ 为保测系统, 也为拓扑动力系统. 此映射称为**帐篷映射**.

2. 称$\vec{\theta} = (\theta_1, \dots, \theta_k) \in \mathbb{R}^k$ 为**有理独立的**, 是指对于任何$n_0, n_1, \dots, n_k \in \mathbb{Z}$,

$$n_0 + n_1\theta_1 + \ldots + n_k\theta_k = 0 \Rightarrow n_0 = n_1 = \ldots = n_k = 0.$$

证明：当$\vec{\theta} = (\theta_1, \dots, \theta_k) \in \mathbb{R}^k$ 为有理独立的, 那么$(\mathbb{T}^k, T_{\vec{\theta}})$ 为极小拓扑动力系统.

3. 对任意的$k \ge 2$, 令

$$T_k : \mathbb{T}^k \to \mathbb{T}^k, (x_1, \dots, x_k) \mapsto (x + \alpha, x_2 + a_{2,1}x_1, \dots, x_k + a_{k,1}x_1 + \dots + a_{k,k-1}x_{k-1}),$$

其中系数$a_{i,j} \in \mathbb{Z}$且$a_{i,i-1} \neq 0$. 则$(\mathbb{T}^k, T_k)$是$(\mathbb{T}^{k-1}, T_{k-1})$的$\mathbb{T}$-扩张. 证明：$(\mathbb{T}^k, \mathcal{B}(\mathbb{T}^k), m, T_k)$ 为保测系统, 其中m 为Lebesgue测度.

4. 设$\beta \in (1,2)$. β **变换**是指映射$T : [0,1) \to [0,1)$:

$$Tx = \beta x \pmod 1 = \begin{cases} \beta x, & 0 \le x < \frac{1}{\beta}; \\ \beta x - 1, & \frac{1}{\beta} \le x < 1. \end{cases}$$

证明：Lebesgue 测度不是T 不变的, 即$([0,1], \mathcal{B}([0,1]), m, T)$ 不为保测系统.

5. 设$T : \mathbb{T}^2 \to \mathbb{T}^2$ 如下：

$$T(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \pmod{\mathbb{Z}^2}$$

证明: $(\mathbb{T}^2, \mathcal{B}(\mathbb{T}^2), m, T)$ 为保测系统, 其中m 为Lebesgue测度. T 称为**Arnold的猫映射**(Arnold's cat map).

6. 设$T : [0,1)^2 \to [0,1)^2$ 定义如下：

$$T(x,y) = \begin{cases} (2x, \frac{y}{2}), & 0 \le x < \frac{1}{2}\ ; \\ (2x-1, \frac{y+1}{2}), & \frac{1}{2} \le x < 1. \end{cases}$$

证明:

(a) $([0,1)^2, \mathcal{B}([0,1)^2), m, T)$ 为保测系统，其中m 为$[0,1)^2$ 上Lebesgue 测度. 此映射称为**Baker变换**.

(b) $([0,1)^2, \mathcal{B}([0,1)^2), m, T)$ 同构于$(\Sigma_2, \mathcal{B}(\Sigma_2), \mu_{(1/2,1/2)}, \sigma)$

7. 设$\Omega = \{-1,1\}^{\mathbb{Z}}$, 其上取测度$\mu_{(1/2,1/2)}$, $\sigma : \Omega \to \Omega$ 为转移映射. 又设$(Y, \mathcal{Y}, \nu, S)$ 为可逆保测映射. 令$X = \Omega \times Y$, 以及
$$T : X \to X, \quad (\omega, y) \mapsto (\sigma\omega, S^{\omega_0} y).$$
证明：$(X, \mathcal{B}(X), \mu_{(1/2,1/2)} \times \nu, T)$ 为可逆保测系统，称之为**随机转移**(random shift).

§1.3 Poincaré回复定理

Poincaré回复定理是一个非常重要的结论, 也是遍历论中最早的主要结果之一[164].

定理 1.3.1 (Poincaré 回复定理(形式1)) *设$(X, \mathcal{X}, \mu, T)$为保测系统. 如果$A \in \mathcal{X}$ 满足$\mu(A) > 0$,那么存在$n \in \mathbb{N}$ 使得*
$$\mu(A \cap T^{-n} A) > 0.$$

证明. 如果存在$A \in \mathcal{X}$, $\mu(A) > 0$ 使得$\mu(A \cap T^{-n} A) = 0$. 那么$A, T^{-1}A, T^{-2}A, \ldots$ 将互不相交(在这里两个集合A, B 不交是指$\mu(A \cap B) = 0$). 这是因为根据μ 为T不变的, 我们有
$$\mu(T^{-i}A \cap T^{-j}A) = \mu(T^{-i}(A \cap T^{-(j-i)}A)) = \mu(A \cap T^{-(j-i)}A) = 0, \ \forall i < j.$$
于是
$$\mu(\bigcup_{n=0}^{\infty} T^{-n}A) = \sum_{n=0}^{\infty} \mu(T^{-n}A) = \sum_{n=0}^{\infty} \mu(A) = \infty.$$
于是存在$n \in \mathbb{N}$ 使得
$$\mu(A \cap T^{-n} A) > 0.$$
证毕! □

根据Poincaré 回复定理, 我们可以定义Poincaré序列：

定义 1.3.2 *$\mathbb{Z}_+$的一个子集S称为***Poincaré 序列***是指对任意保测系统$(X, \mathcal{X}, \mu, T)$以及任意满足$\mu(A) > 0$的$A \in \mathcal{X}$, 存在$0 < n \in S$ 使得$\mu(A \cap T^{-n}A) > 0$.*

于是Poincaré回复定理可以重述为：$\mathbb{Z}_+$为Poincaré序列.

定义 1.3.3 *设$F \subseteq \mathbb{Z}_+$. 称$F - F = \{a - b : a, b \in F, a \geq b\}$ 为F***的差集***或者***Δ 集***.*

*设$H \subseteq \mathbb{Z}_+$. 如果对任意的$\mathbb{Z}_+$的无限子集F, $H \cap (F - F) \neq \emptyset$, 则称$H$***是一个$\Delta^*$ 集***.*

定义 1.3.4 *设$(X, \mathcal{X}, \mu, T)$为保测系统. 对任意的$A, B \in \mathcal{X}$, 定义*
$$N^{\mu}(A, B) = \{n \in \mathbb{Z}_+ : \mu(A \cap T^{-n}B) > 0\}.$$

命题 1.3.5 *设$(X, \mathcal{X}, \mu, T)$为保测系统. 则对任意正测集A, $N^{\mu}(A, A)$是一个Δ^*-集. 即Δ 集为Poincaré 序列.*

证明. 设$F \subseteq \mathbb{Z}_+$是一个无限集. 对任意的$i \in F$, $\mu(A) = \mu(T^{-i}A)$. 因此$\{T^{-i}A : i \in F\}$是X中一列测度值均为$\mu(A)$的集合. 由于$\mu(X) = 1$, 故存在$i < j \in F$使得$\mu(T^{-i}A \cap T^{-j}A) > 0$. 从而

$$\mu(A \cap T^{-(j-i)}A) = \mu(T^{-i}(A \cap T^{-(j-i)}A)) = \mu(T^{-i}A \cap T^{-j}A) > 0,$$

即

$$j - i \in N^{\mu}(A, A) \cap (F - F).$$

证毕! □

定理 1.3.6 (Poincaré回复定理(形式2)) 设$(X, \mathcal{X}, \mu, T)$ 为保测系统, $A \in \mathcal{X}$. 那么A 几乎每个点都无穷次正向回复到本身, 即存在$B \subseteq A$ 使得$\mu(B) = \mu(A)$, 并且对于每个$x \in B$, 存在递增自然数序列$\{n_i\}_{i \in \mathbb{N}}$ 满足$T^{n_i}x \in B, \forall i \in \mathbb{N}$.

证明. 首先我们证明一个断言:

断言: 设$A \in \mathcal{X}$和令

$$A_\infty = \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}A.$$

则$A_\infty \in \mathcal{X}$, $A_\infty = T^{-1}A_\infty$ 和

$$\mu(A_\infty) = \mu\Big(\bigcup_{i=0}^{\infty} T^{-i}A\Big).$$

断言证明: 由于每个$T^{-i}A \in \mathcal{X}$, 所以$A_\infty \in \mathcal{X}$. 由A_∞的定义可知

$$T^{-1}A_\infty = \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-(i+1)}A = \bigcap_{n=0}^{\infty} \bigcup_{i=n+1}^{\infty} T^{-i}A = A_\infty.$$

对任意的$n \geq 0$,

$$\bigcup_{i=n}^{\infty} T^{-i}A \supset \bigcup_{i=n+1}^{\infty} T^{-i}A$$

和

$$\mu\Big(\bigcup_{i=n}^{\infty} T^{-i}A\Big) = \mu\Big(T^{-1}\bigcup_{i=n}^{\infty} T^{-i}A\Big) = \mu\Big(\bigcup_{i=n+1}^{\infty} T^{-i}A\Big).$$

故

$$\mu(A_\infty) = \lim_{n\to\infty} \mu\Big(\bigcup_{i=n}^{\infty} T^{-i}A\Big) = \mu\Big(\bigcup_{i=0}^{\infty} T^{-i}A\Big).$$

断言证毕. □

根据定义, A_∞是所有轨道无穷次进入A的点的全体. 令

$$B = A \cap A_\infty.$$

则B为A中所有轨道无穷次进入A的点的全体. 设$x \in B$, 则存在$0 < n_1 < n_2 < \ldots$使得$T^{n_i}x \in A$. 对于任何取定$i \in \mathbb{N}$, 以及对于任何$j > i$我们有

$$T^{n_j - n_i}(T^{n_i}x) = T^{n_j}x \in A.$$

即$T^{n_i}x \in A$ 并且此点轨道无穷次进入A. 所以$T^{n_i}x \in B$. 综上, 对于任何$x \in B$, 存在$0 < n_1 < n_2 < \ldots$使得$T^{n_i}x \in B$.

由断言, $A_\infty \in \mathcal{X}$, 从而$B \in \mathcal{X}$. 再由断言,

$$A_\infty \subseteq \bigcup_{i=0}^{\infty} T^{-i}A \text{ 和} \mu(A_\infty) = \mu\left(\bigcup_{i=0}^{\infty} T^{-i}A\right).$$

所以

$$\mu(B) = \mu(A \cap A_\infty) = \mu\left(A \cap \bigcup_{i=0}^{\infty} T^{-i}A\right) = \mu(A).$$

证毕! □

注记 1.3.7 如果没有条件$\mu(X) < \infty$, 那么回复定理可能不成立. 例如

$$T : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x + 1.$$

习 题

1. 设$a \in \mathbb{N}$, 证明$a\mathbb{N}$ 为Poincaré序列.
2. 集合$A \subseteq \mathbb{N}$ 称为thick的是指它包含任意长的区间, 即存在$a_i \in \mathbb{N}$ 使得$[a_i, a_i + i] \subseteq A, \forall i \in \mathbb{N}$. 证明任何thick集合为Poincaré序列.

§1.4 注记

我们在书中没有介绍流形上的保测系统、区间交换变换、台球系统、Gauss系统等, 更多的例子可以参见[43, 184, 123, 153, 191]等. 关于符号动力系统可以参见[147, 210].

第二章　遍历性与遍历定理

在本章中我们给出遍历性的定义和刻画, 并且介绍von Neumann 遍历定理和Birkhoff 遍历定理. von Neumann 遍历定理和Birkhoff 遍历定理是遍历论最基本的定理. 在介绍完遍历性后, 我们给出各种混合的定义, 其中最为重要的是弱混合. 弱混合的谱刻画是指它等价于连续谱, 我们在本章最后给出动力系统谱理论的一些基本知识.

§2.1　遍历性

遍历性是遍历论中核心概念和研究内容之一, 遍历性体现的是系统的一种不可分割性. 我们会在后面解释为什么测度动力系统理论也称为遍历论.

§2.1.1　遍历性

定义 2.1.1 *设$(X,\mathcal{X},\mu,T)$为保测系统. 称T 为***遍历**的*, 指对任意满足$T^{-1}B = B$ 的$B \in \mathcal{X}$ 必有$\mu(B) = 0$ 或$\mu(B) = 1$ 成立.*

注记 2.1.2 *对于非概率测度, 我们一般如下定义遍历: 对任意满足$T^{-1}B = B$ 的$B \in \mathcal{X}$ 必有$\mu(B) = 0$ 或$\mu(X \setminus B) = 0$ 成立.*

设$(X,\mathcal{X},\mu,T)$为一个保测系统, 对任意满足$\mu(A)\mu(B) > 0$的$A, B \in \mathcal{X}$, 之前我们定义

$$N^{\mu}(A,B) = \{n \in \mathbb{Z}_+ : \mu(A \cap T^{-n}B) > 0\}.$$

关于遍历性, 我们有众多等价刻画:

定理 2.1.3 *设$(X,\mathcal{X},\mu,T)$为保测系统,则以下命题等价:*

1. *T 为遍历的;*
2. *$\mathcal{X}$ 中元B 如满足$\mu(T^{-1}B\Delta B) = 0$（即$T^{-1}B = B, a.e.$）,则必有$\mu(B) = 0$ 或$\mu(B) = 1$;*
3. *对任意满足$\mu(A) > 0$ 的$A \in \mathcal{X}$ 有$\mu(\bigcup_{n=1}^{\infty} T^{-n}A) = 1$;*
4. *对任意满足$\mu(A)\mu(B) > 0$ 的$A, B \in \mathcal{X}$ 有$N^{\mu}(A,B) \neq \emptyset$, 即存在$n \in \mathbb{N}$ 使得$\mu(A \cap T^{-n}B) > 0$.*

证明. (1) $\Rightarrow$ (2). 设$B \in \mathcal{X}$满足$\mu(T^{-1}B\Delta B) = 0$. 下面我们构造集合$B_\infty \in \mathcal{X}$ 使得$T^{-1}(B_\infty) = B_\infty$ 且$\mu(B\Delta B_\infty) = 0$. 从而根据条件（1）完成证明.

令

$$B_\infty = \bigcap_{n=0}^{\infty}\bigcup_{i=n}^{\infty} T^{-i}B.$$

由Poincaré定理（定理1.3.6）断言证明, $T^{-1}(B_\infty)=B_\infty$ 和$\mu(B_\infty)=\mu(\bigcup_{n=0}^\infty T^{-n}B)$. 由遍历性, $\mu(B_\infty)=0$或1. 对任意的$n\in\mathbb{N}$,

$$\begin{aligned}\mu(T^{-n}B\Delta B)&\le\mu\left(\bigcup_{i=1}^n(T^{-i+1}B\Delta T^{-i}B)\right)\le\sum_{i=1}^n\mu(T^{-i+1}B\Delta T^{-i}B)\\&=\sum_{i=1}^n\mu(T^{-i+1}(B\Delta T^{-1}B))=\sum_{i=1}^n\mu(B\Delta T^{-1}B)=0.\end{aligned}$$

故

$$\mu\left(B\Delta\bigcup_{i=n}^\infty T^{-i}B\right)\le\sum_{i=n}^\infty\mu(B\Delta T^{-i}B)=0,$$

因为对每个$n\in\mathbb{N}$有$(\bigcup_{i=n}^\infty T^{-i}B)\Delta B$测度为0, 我们有$\mu(B_\infty\Delta B)=0$, 于是$\mu(B_\infty)=\mu(B)$. 所以由$\mu(B_\infty)=0$或1, 得到$\mu(B)=0$或1.

(2) $\Rightarrow$ (3). 设$A\in\mathcal{X}$ 满足$\mu(A)>0$. 令$A_1=\bigcup_{n=1}^\infty T^{-n}A$. 因为$T^{-1}A_1\subseteq A_1$及$\mu(T^{-1}A_1)=\mu(A_1)$, 所以

$$\mu(T^{-1}A_1\Delta A_1)=0.$$

由(2), 我们得到$\mu(A_1)=0$ 或1. 因为$\mu(A)>0$, 所以$\mu(A_1)=1$.

(3) $\Rightarrow$ (4). 设$\mu(A)\mu(B)>0$. 由(3), 我们有$\mu(\bigcup_{n=1}^\infty T^{-n}B)=1$, 于是

$$0<\mu(A)=\mu(A\bigcap\bigcup_{n=1}^\infty T^{-n}B)=\mu(\bigcup_{n=1}^\infty(A\cap T^{-n}B)).$$

从而存在$n\ge1$ 使得$\mu(A\cap T^{-n}B)>0$.

(4) $\Rightarrow$ (1). 设$B\in\mathcal{X}$满足$T^{-1}B=B$. 如果$0<\mu(B)<1$, 那么$\mu(X\setminus B)>0$且

$$\mu(T^{-n}B\cap(X\setminus B))=\mu(B\cap(X\setminus B))=0,\ \forall n\ge1.$$

这与(4)矛盾. □

根据证明, 我们知道(3)等价于对于任何满足$\mu(A)>0$的$A\in\mathcal{X}$和$N\in\mathbb{N}$有$\mu(\bigcup_{n=N}^\infty T^{-n}A)=1$. 另外根据证明,（4）中$N^\mu(A,B)$ 为无穷集. 对于$N^\mu(A,B)$ 我们可以说的更多.

定义 2.1.4 集合$A=\{a_1<a_2<\ldots\}\subseteq\mathbb{Z}_+$称为**syndetic的**是指它具有有界的间距,即存在$N>0$使得

$$a_{i+1}-a_i\le N,\ \forall i\in\mathbb{N}.$$

记全体*syndetic*集为$\mathcal{F}_s$.

根据定义，容易验证集合$A\subseteq\mathbb{Z}_+$ 为syndetic 的当且仅当存在$N\in\mathbb{N}$ 使得对于任何$i\in\mathbb{Z}_+$, $A\cap\{i,i+1,\ldots,i+N\}\neq\emptyset$. $\mathbb{Z}$,$\mathbb{N}$ 中的syndetic 子集类似定义.

推论 2.1.5 *设$(X,\mathcal{X},\mu,T)$ 为遍历系统, 那么对任意满足$\mu(A)\mu(B)>0$ 的$A,B\in\mathcal{X}$, $N^\mu(A,B)$为$syndetic$的.*

证明. 因为$\mu(B)>0$, 由T的遍历性我们有$\mu(\bigcup_{i=1}^\infty T^{-i}B)=1$. 取$k$使得

$$\mu(\bigcup_{i=1}^{k}T^{-i}B)>1-\frac{\mu(A)}{2}.$$

根据μ 的不变性, 对每个$l\in\mathbb{Z}_+$ 有

$$\mu(\bigcup_{i=l+1}^{l+k}T^{-i}B)=\mu(\bigcup_{i=1}^{k}T^{-i}B)>1-\frac{\mu(A)}{2}.$$

这意味着对每个$l\in\mathbb{Z}_+$ 有

$$\mu\left(A\cap(\bigcup_{i=l+1}^{l+k}T^{-i}B)\right)>0,$$

亦即对每个$l\in\mathbb{Z}_+$, 存在$j\in[l+1,l+k]$ 使得

$$\mu(A\cap T^{-j}B)>0.$$

从而$N^\mu(A,B)$为syndetic的. □

§2.1.2　不变函数

我们已经定义了算子U_T, 由此自然地定义特征值与特征函数.

定义 2.1.6 *设$(X,\mathcal{X},\mu,T)$为保测系统, 我们称复数λ为T的**特征值**, 如果它为算子U_T的特征值, 即存在非零可测函数f使得*

$$U_Tf=\lambda f.$$

*这个函数f称为相应于特征值λ的**特征函数**.*

下面的定理指出遍历性可以由U_T谱的性质来刻画: 遍历性等价于1为U_T简单特征值.

定理 2.1.7 *设$(X,\mathcal{X},\mu,T)$为保测系统, 则以下命题为等价的:*

1. *T为遍历的;*
2. *如果f可测并且$U_Tf(x)=f(x),\forall x\in X$, 那么$f=const\ a.e.$*
3. *如果f可测并且$U_Tf=f\ a.e.$, 那么$f=const\ a.e.$*
4. *如果$f\in L^2(\mu)$满足$U_Tf(x)=f(x),\forall x\in X$, 那么$f=const\ a.e.$*
5. *如果$f\in L^2(\mu)$满足$U_Tf(x)=f(x)\ a.e.$, 那么$f=const\ a.e.$*

证明. $(3)\Rightarrow(2),(2)\Rightarrow(4),(5)\Rightarrow(4),(3)\Rightarrow(5)$ 是平凡的. 我们只需证明$(1)\Rightarrow(3)$和$(4)\Rightarrow(1)$.

$(1)\Rightarrow(3)$. 假设T为遍历的且f可测. 不妨设f为实值的（对于复值映射,我们可以分实部与虚部讨论）. 对$k\in\mathbb{Z}$及$n>0$令

$$X(k,n)=\{x\in X:k/2^n\le f(x)<(k+1)/2^n\}=f^{-1}([k/2^n,(k+1)/2^{n+1})).$$

我们有

$$T^{-1}X(k,n)\Delta X(k,n)\subseteq\{x\in X:f\circ T(x)\neq f(x)\},$$

继而$\mu(T^{-1}X(k,n)\Delta X(k,n))=0$. 根据定理2.1.3就有$\mu(X(k,n))=0$或1.

对每个$n\in\mathbb{N}$, $\bigcup_{k\in\mathbb{Z}}X(k,n)=X$为无交并, 于是存在唯一一个$k_n$使得$\mu(X(k_n,n))=1$. 令

$$Y=\bigcap_{n=1}^{\infty}X(k_n,n),$$

则$\mu(Y)=1$且f在Y上为常数, 即$f=const$ a.e.

$(4)\Rightarrow(1)$. 假设$B\in\mathcal{X}$及$T^{-1}B=B$. 因为$U_T\mathbf{1}_B=\mathbf{1}_B$, 我们有$\mathbf{1}_B$ 为几乎处处为常值. 亦即, $\mu(B)=0$或1. 所以T为遍历的. □

注记 2.1.8 根据上面定理的证明, 在定理*2.1.7*中将$L^2(\mu)$ 换为$L^p(\mu),p\ge 1$, 结论是一样的.

设$(X,\mathcal{X},\mu,T)$为保测系统令

$$\mathcal{I}=\mathcal{I}(T)=\{A\in\mathcal{X}\colon \mu(A\Delta T^{-1}A)=0\}. \tag{2.1.1}$$

易见$\mathcal{I}$ 是$\mathcal{X}$ 的一个子σ 代数且$T^{-1}\mathcal{I}\subseteq\mathcal{I}$.

引理 2.1.9 设$(X,\mathcal{X},\mu,T)$为保测系统. 则对任意的$f\in L^1(X,\mu)$,

$$\mathbb{E}(f\circ T|\mathcal{I})=\mathbb{E}(f|\mathcal{I}).$$

证明. 设$f\in L^1(X,\mu)$. 对任意的$A\in\mathcal{I}$,

$$\begin{aligned}\int_A\mathbb{E}(f\circ T|\mathcal{I})d\mu&=\int_A f\circ Td\mu=\int_{T^{-1}A}f\circ Td\mu\\&=\int_A fd\mu=\int_A\mathbb{E}(f|\mathcal{I})d\mu.\end{aligned}$$

所以$\mathbb{E}(f\circ T|\mathcal{I})=\mathbb{E}(f|\mathcal{I})$. □

引理 2.1.10 设$(X,\mathcal{X},\mu,T)$为保测系统. 设$f\in L^1(X,\mu)$. 则$f\circ T=f$ *a.e.* 当且仅当f 是$\mathcal{I}$ 可测的.

证明. 设$f\circ T=f$ $a.e.$ 不妨设f 为实值函数. 对任意的$A\in\mathcal{B}(\mathbb{R})$, $\mu(f^{-1}(A)\Delta(f\circ T)^{-1}(A))=0$, 即$\mu(f^{-1}(A)\Delta T^{-1}(f^{-1}(A)))=0$. 所以$f^{-1}(A)\in\mathcal{I}$, 从而$f$是$\mathcal{I}$ 可测的.

设f是$\mathcal{I}$ 可测的. 对任意的$A\in\mathcal{B}(\mathbb{R})$, $f^{-1}(A)\in\mathcal{I}$, 即$\mu(f^{-1}(A)\Delta T^{-1}(f^{-1}(A)))=0$, 从而$\mu(f^{-1}(A)\Delta(f\circ T)^{-1}(A))=0$. 特别地, 对任意的$a\in\mathbb{Q}$,

$$\mu\big(f^{-1}((-\infty,a))\cap(f\circ T)^{-1}([a,+\infty))\big)=0$$

和

$$\mu\big(f^{-1}((a,+\infty))\cap(f\circ T)^{-1}((-\infty,a])\big)=0.$$

所以$f\circ T=f$ 几乎处处成立. □

§2.1.3　一些例子

例 2.1.11 (n-周期系统) 设$n\in\mathbb{N}$ 和$\mathbb{Z}_n=\mathbb{Z}/n\mathbb{Z}=\{0,1,\dots,n-1\}$. 令$T\colon\mathbb{Z}_n\to\mathbb{Z}_n$, $x\mapsto x+1\pmod n$. 测度取每个点测度为$\frac{1}{n}$ 的平均测度μ, 那么$(\mathbb{Z}_n,\mathcal{P}(\mathbb{Z}_n),\mu,T)$ 为遍历系统.

例 2.1.12 (圆周旋转) 设$\mathbb{S}^1$为复平面上的单位圆周,$\alpha\in\mathbb{R}$, $a=e^{2\pi i\alpha}$, $m_{\mathbb{S}^1}$ 为$\mathbb{S}^1$上的$Haar$测度. 定义

$$T_a:\mathbb{S}^1\to\mathbb{S}^1, z=e^{2\pi i\theta}\mapsto az=e^{2\pi i(\theta+\alpha)},\ \forall z\in\mathbb{S}^1.$$

则系统$(\mathbb{S}^1,\mathcal{B}(\mathbb{S}^1),m_{\mathbb{S}^1},T_a)$为遍历的当且仅当$\alpha$为无理数, 即$a$ 不为单位根.

证明.设a 为单位根, 那么存在$p\neq 0$ 使得$a^p=1$. 令$f(z)=z^p$. 则$U_{T_a}f=f\circ T=f$, 但是f不是常值函数, 根据定理2.1.7, T_a 不是遍历的.

反之, 设$f\in L^2(\mathbb{S}^1,m_{\mathbb{S}^1})$ 满足$U_{T_a}f=f$. 令$f(z)=\sum_{n=-\infty}^{\infty}b_nz^n$ 为其Fourier 级数. 则

$$U_{T_a}f(z)=f(az)=\sum_{n=-\infty}^{\infty}b_na^nz^n,$$

于是由表示的唯一性,

$$b_n(a^n-1)=0,\quad\forall n\in\mathbb{Z}.$$

如果$n\neq 0$, 那么$b_n=0$, 继而$f=b_0$ 为常值函数. 根据定理2.1.7, T_a为遍历的. □

下面我们给出另外一个证明.

证明. 设$\mathbb{T}=\mathbb{R}/\mathbb{Z}=[0,1)\pmod 1$ 和$\alpha\in\mathbb{T}$. 令

$$R_\alpha\colon\mathbb{T}\to\mathbb{T},\ x\mapsto x+\alpha\pmod 1.$$

那么$(\mathbb{T},\mathcal{B}(\mathbb{T}),m_{\mathbb{T}},R_\alpha)$ 为遍历的当且仅当α 为无理数.

如果$\alpha\in\mathbb{Q}$, 设$\alpha=\frac{p}{q}, p,q\in\mathbb{N}$, $(p,q)=1$. 于是$R_\alpha^q=\mathrm{id}$. 取子集$A\subseteq\mathbb{T}$ 使得$1<m_{\mathbb{T}}(A)<\frac{1}{q}$. 那么

$$B=A\cup R_\alpha A\cup\ldots\cup R_\alpha^{q-1}A$$

满足$R_\alpha^{-1}B=B$, 但是$m_{\mathbb{T}}(B)\in(0,1)$. 所以$R_\alpha$ 不遍历.

如果$\alpha\notin\mathbb{Q}$. 那么R_α 为极小作用. 设$B\subseteq\mathbb{T}$ 为R_α 不变的. 对于任何$\varepsilon>0$, 取连续函数$f\in C(\mathbb{T})$ 使得$\|f-\mathbf{1}_B\|_1<\varepsilon$. 因为$\mathbf{1}_B\circ R_\alpha=\mathbf{1}_B$, 我们有

$$\|f\circ R_\alpha^n-f\|_1<2\varepsilon,\ \forall n\in\mathbb{Z}.$$

因为f 连续以及R_α 极小, 所以根据上式容易得到

$$\|f\circ R_t-f\|_1<2\varepsilon,\ \forall t\in\mathbb{R}.$$

于是根据$m_{\mathbb{T}}$ 的平移不变性,

$$\begin{aligned}\left\|f-\int_{\mathbb{T}}f(t)dt\right\|_1&=\int_{\mathbb{T}}\left|\int_{\mathbb{T}}(f(x)-f(x+t))dt\right|dx\\&\le\int_{\mathbb{T}}\int_{\mathbb{T}}|f(x)-f(x+t)|\,dxdt\le 2\varepsilon.\end{aligned}$$

于是

$$\begin{aligned}\|\mathbf{1}_B-m_{\mathbb{T}}(B)\|_1&=\|\mathbf{1}_B-\int_{\mathbb{T}}\mathbf{1}_Bdt\|_1\\&\le\|\mathbf{1}_B-f\|_1+\left\|f-\int_{\mathbb{T}}f(t)dt\right\|_1+\left\|\int_{\mathbb{T}}f(t)dt-\int_{\mathbb{T}}\mathbf{1}_Bdt\right\|_1\\&<4\varepsilon.\end{aligned}$$

根据ε 的任意性, $\|\mathbf{1}_B-m_{\mathbb{T}}(B)\|_1=0$. 于是$\mathbf{1}_B$ 为常数, $m_{\mathbb{T}}(B)=0$ 或1. 所以R_α 遍历. □

例 2.1.13 (群旋转) 设G是一个紧致可度量化的群和$a\in G$. 定义群旋转

$$R_a\colon G\to G, x\mapsto ax.$$

则$(G,\mathcal{B}(G),m,R_a)$是一个保测动力系统, 其中m为$Haar$测度. R_a 为遍历的当且仅当$\langle a\rangle=\{a^n:n\in\mathbb{Z}\}$在$G$中稠密.

尤其如果$(G,\mathcal{B}(G),m,R_a)$是遍历的, 那么G为交换群.

证明.如果$\langle a\rangle=\{a^n:n\in\mathbb{Z}\}$ 不在G 中稠密. 设$H=\overline{\{a^n:n\in\mathbb{Z}\}}$, 则$H\neq G$ 为G 闭子群. 于是存在$\gamma\in\widehat{G}$ 使得$\gamma\neq 1$并且$\gamma(H)=1$. 由此

$$U_{R_a}\gamma(x)=\gamma(R_ax)=\gamma(ax)=\gamma(a)\gamma(x)=\gamma(x).$$

所以根据定理2.1.7 R_a 不为遍历的.

反之, 设$\langle a\rangle=\{a^n:n\in\mathbb{Z}\}$在$G$中稠密. 于是$G$为交换群. 设$f\in L^2(\mu)$满足$U_{R_a}f=f$. 令$f(x)=\sum_{\gamma\in\widehat{G}}b_\gamma\gamma(x)$为其Fourier 级数. 则

$$U_{R_a}f(x)=f(ax)=\sum_{\gamma\in\widehat{G}}b_\gamma\gamma(a)\gamma(x),$$

于是由表示的唯一性,

$$b_\gamma(\gamma(a)-1)=0, \quad \forall \gamma \in \widehat{G}.$$

如果$b_\gamma \neq 0$, 那么$\gamma(a)=1$. 从而

$$\gamma(a^n)=\gamma(a)^n=1, \forall n \in \mathbb{Z}.$$

由γ 的连续性, $\gamma|_G=\gamma|_{\overline{\{a^n : n\in\mathbb{Z}\}}}=1$, 即$\gamma=1$. 所以$b_\gamma=0, \forall \gamma \neq 1$. 继而$f=b_1$为常值函数. 根据定理2.1.7, R_a为遍历的. □

例 2.1.14 (圆周自同态) 设$n\in\mathbb{N}$, 令$S_n\colon \mathbb{S}^1\to\mathbb{S}^1$, $z\mapsto z^n$. 那么$(\mathbb{S}^1,\mathcal{B}(\mathbb{S}^1),m_{\mathbb{S}^1},S_n)$为保测系统. 下面证明当$n\geq 2$时它为遍历的. 我们以$n=2$为例, 一般情况的证明是完全类似的.

设$f\in L^2(\mu)$满足$U_{S_2}f=f$. 令$f(z)=\sum_{n=-\infty}^{\infty} b_n z^n$为其*Fourier* 级数. 则

$$U_{S_2}f(z)=f(z^2)=\sum_{n=-\infty}^{\infty} b_n z^{2n},$$

于是由表示的唯一性,如果$n\neq 0$, 则

$$b_n=b_{2n}=b_{2^2n}=b_{2^3n}=\ldots.$$

因为$\sum_{j=-\infty}^{\infty}|b_j|^2=\|f\|_2^2<\infty$, 所以$b_n=0$. 综上, 如果$n\neq 0$, 那么$b_n=0$, 继而$f(z)=b_0$为常值函数.根据定理*2.1.7*, S_n为遍历的.

例 2.1.15 (群自同态系统) 设G为紧致度量群, $A:G\to G$为满的连续自同态映射, m为其上*Haar*测度. 那么$(G,\mathcal{B}(G),m,A)$ 为保测映射. $(G,\mathcal{B}(G),m,A)$为遍历的当且仅当平凡的特征$\gamma=1$为唯一满足存在$n\in\mathbb{N}$ 使得$\gamma\circ A^n=\gamma$成立的特征.

证明.设存在非平凡$\gamma\in\widehat{G}$以及$n\in\mathbb{N}$使得$\gamma\circ A^n=\gamma$. 不妨设n为满足上式最小的自然数. 令

$$f=\gamma+\gamma\circ A+\ldots+\gamma\circ A^{n-1}.$$

因为$\gamma,\gamma\circ A,\ldots,\gamma\circ A^{n-1}\in\widehat{G}$互异, 所以互相垂直, 尤其$f\neq \mathrm{const}$. 但是$f\circ A=f$, 所以根据定理2.1.7, A不为遍历的.

反之, 假设如存在$n\in\mathbb{N}$使得$\gamma\circ A^n=\gamma$, 则$\gamma=1$. 设$f\in L^2(\mu)$ 满足$U_Af=f\circ A=f$. 令$f(x)=\sum_{\gamma\in\widehat{G}} b_\gamma\gamma(x)$为其Fourier 级数. 则

$$U_Af(x)=f(Ax)=\sum_{\gamma\in\widehat{G}} b_\gamma\gamma(Ax),$$

于是由表示的唯一性,

$$b_\gamma=b_{\gamma\circ A}=b_{\gamma\circ A^2}=\ldots.$$

根据条件, 如果$\gamma\neq 1$, $\gamma,\gamma\circ A,\gamma\circ A^2,\ldots$ 为互异的特征. 由于$\sum_{\gamma\in\widehat{G}}|b_\gamma|^2<\infty$, 所以只能是$b_\gamma=0$. 继而$f=b_1$为常值函数. 根据定理2.1.7, A为遍历的. □

例 2.1.16 (环面自同态) 根据定理*0.4.13*, 每个自同态$A:(\mathbb{S}^1)^n\to(\mathbb{S}^1)^n$ 形如

$$A(z_1,\ldots,z_n)=(z_1^{a_{11}}z_2^{a_{12}}\ldots z_n^{a_{1n}},\ldots,z_1^{a_{n1}}z_2^{a_{n2}}\ldots z_n^{a_{nn}}),$$

其中$a_{ij}\in\mathbb{Z}$. 等价的, $A:\mathbb{T}^n\to\mathbb{T}^n$表达式为

$$A\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+\mathbb{Z}^n\right)=[a_{ij}]\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+\mathbb{Z}^n,$$

其中$[a_{ij}]\in\mathbb{Z}^{n\times n}$为方阵, 第$(i,j)$元素为$a_{ij}$. 如果$A$为满自同态, 那么$A$为遍历的当且仅当矩阵$[A]$没有单位根作为它的特征值.

证明.根据§0.4小节知识,

$$\widehat{A}:\mathbb{Z}^n\to\mathbb{Z}^n,\begin{pmatrix}m_1\\ m_2\\ \vdots\\ m_n\end{pmatrix}\mapsto[A]^t\begin{pmatrix}m_1\\ m_2\\ \vdots\\ m_n\end{pmatrix}$$

其中$[A]^t$为转置. 根据例子2.1.15, 如果A 不是遍历的, 那么存在$\vec{q}\in\mathbb{Z}^n,\vec{q}\neq\mathbf{0}$ 以及$k\in\mathbb{N}$ 使得$([A]^t)^k\vec{q}=\vec{q}$. 于是1 为$([A]^t)^k$ 的特征值. 所以1 为$[A]^k$ 的特征值, 从而存在单位根为$[A]$的特征值.

反之, 如果存在单位根为$[A]$ 的特征值, 那么存在$k\in\mathbb{N}$ 使得1 为$([A]^t)^k$ 的特征值. 于是存在非零向量$\vec{y}\in\mathbb{R}^n$ 使得

$$\left(([A]^t)^k-I\right)\vec{y}=0.$$

因为$[A]_t$为整数值矩阵, 所以可以设$\vec{y}\in\mathbb{Z}^n$. 于是$([A]^t)^k\vec{y}=\vec{y}$. 根据例子2.1.15, A不是遍历的. □

例 2.1.17 (转移系统) 双边（单边）$\vec{p}$-转移系统$(\Sigma_k,\mathcal{B},\mu_{\vec{p}},\sigma)$为遍历的.

例 2.1.18 (Markov转移) 双边（单边）$(\vec{p},P)$-*Markov*转移系统$(\Sigma_k,\mathcal{B},\mu,\sigma)$ 为遍历的当且仅当P为不可约的.

$\vec{p}$-转移系统是遍历的证明不难, 请读者自己完成. 但是Markov 转移系统的情况要复杂些, 我们把证明留在给出遍历定理之后（定理2.3.15）.

习 题

1. 证明：双边（单边）$\vec{p}$-转移系统$(\Sigma_k,\mathcal{B},\mu_{\vec{p}},\sigma)$为遍历的.

2. 设$T:[0,1]\to[0,1]$ 定义如下:

$$T(x)=\begin{cases}2x, & 0\le x<\frac{1}{4};\\ 2x-\frac{1}{2}, & \frac{1}{4}\le x<\frac{3}{4};\\ 2x-1, & \frac{3}{4}\le x\le 1.\end{cases}$$

证明: $([0,1],\mathcal{B}([0,1]),m,T)$ 为保测但是不遍历的系统, 其中m 为Lebesgue 测度.

3. 设G为紧致连通度量交换群$a\in G$,$A:G\to G$为满的连续自同态映射, m 为其上Haar测度. 令$T:G\to G,\ x\mapsto a\cdot A(x)$ 为仿射变换. 则以下等价：

   (a) $(G,\mathcal{B}(G),m,T)$为遍历的.

   (b) 如存在$n\in\mathbb{N}$使得$\gamma\circ A^n=\gamma$, 则$\gamma\circ A=\gamma$, 并且包含a和BG 的最小闭子群为G, 其中$B:G\to G,x\mapsto x^{-1}\cdot A(x)$.

   (c) 存在$x_0\in G$使得$\overline{\{T^nx_0:n\ge 0\}}=G$.

   (d) $m(\{x:\overline{\{T^nx:n\ge 0\}}=G\})=1$.

4. 对$k\in\mathbb{N}$, 取$\vec{\theta}=(\theta_1,\dots,\theta_k)\in\mathbb{R}^k$. 令

$$T_{\vec{\theta}}:\mathbb{T}^k\to\mathbb{T}^k,\vec{x}=(x_1,x_2,\dots,x_k)\mapsto\vec{x}+\vec{\theta}=(x_1+\theta_1,x_2+\theta_2,\dots,x_k+\theta_k).$$

   $\mathbb{T}^k$上的Lebesgue测度m 是$(\mathbb{T}^k,T_{\vec{\theta}})$的一个不变测度.

   证明：$(\mathbb{T}^k,\mathcal{B},m,T_{\vec{\theta}})$ 为遍历的当且仅当$\vec{\theta}=(\theta_1,\dots,\theta_k)\in\mathbb{R}^k$ 为有理独立的.

5. 设$\alpha\in(0,1)\setminus\mathbb{Q}$和$k\ge 2$, $T_k:\mathbb{T}^k\to\mathbb{T}^k$, $(x_1,x_2,\dots,x_k)\mapsto(x_1+\alpha,x_2+x_1,\dots,x_k+x_{k-1})$. $\mathbb{T}^k$上的Lebesgue测度是$(\mathbb{T}^k,T_k)$的一个遍历测度.

6. 设 $(X,\mathcal{X},\mu,T)$ 为保测系统. 我们称 T 为完全遍历的, 指对于任意的整数 $n\ge 1$, T^n 为遍历的. 给定 $K\ge 1$, 定义空间

$$X^{(K)}=X\times\{1,2,\dots,K\},$$

   $\mathcal{X}^{(K)}$ 为其上乘积 σ 代数, $\mu^{(K)}=\mu\times\nu$ 为乘积测度 (ν 为 $\{1,2,\dots,K\}$ 上计数测度, 即对于任何子集 $A\subseteq\{1,2,\dots,K\}$, $\nu(A)=\frac{|A|}{K}$, $|A|$ 为集合 A 基数). 定义 $T^{(K)}:X^{(K)}\to X^{(K)}$ 为: $\forall(x,i)\in X^{(K)}$.

$$T^{(K)}(x,i)=\begin{cases}(x,i+1), & \text{如果 } 1\le i<K;\\ (Tx,1), & \text{如果 } i=K.\end{cases}$$

   证明:

   (1) $(X^{(K)},\mathcal{X}^{(K)},\mu^{(K)},T^{(K)})$ 为保测系统.

   (2) $(X^{(K)},\mathcal{X}^{(K)},\mu^{(K)},T^{(K)})$ 为遍历的当且仅当 $(X,\mathcal{X},\mu,T)$ 为遍历的. 如果 $K>1$, 那么 $(X^{(K)},\mathcal{X}^{(K)},\mu^{(K)},T^{(K)})$ 不是完全遍历的.

   (3) 如果 $(Y,\mathcal{Y},m,S)$ 不是完全遍历的系统, 那么存在保测系统 $(X,\mathcal{X},\mu,T)$ 以及 $K>1$ 使得$(Y,\mathcal{Y},m,S)$ 同构于 $(X^{(K)},\mathcal{X}^{(K)},\mu^{(K)},T^{(K)})$.

§2.2 von Neumann 遍历定理

设$(X, \mathcal{X}, \mu, T)$为保测系统, 则Koopman算子$U_T : L^2(X, \mu) \to L^2(X, \mu)$, $f \mapsto f \circ T$, 是一个等距算子. 一般我们也经常记

$$Tf = U_T f = f \circ T.$$

对$f \in L^1(\mu)$, 称

$$\mathbb{A}_n f(x) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) = \frac{1}{n}\sum_{i=0}^{n-1} T^i f(x)$$

为 **遍历平均**. von Neumann 遍历定理断言遍历平均在范数下收敛, 而Birkhoff 遍历定理表明遍历平均几乎处处收敛. 这一节我们先证明von Neumann 遍历定理. 首先我们给出von Neumann 遍历定理的Hilbert空间形式.

§2.2.1 von Neumann 遍历定理（Hilbert空间形式）

首先回顾一些基本概念. 设$U : \mathcal{H}_1 \to \mathcal{H}_2$ 为Hilbert 空间上的线性算子, 那么根据

$$\langle Uf, g\rangle = \langle f, U^* g\rangle$$

定义了关联算子$U^* : \mathcal{H}_2 \to \mathcal{H}_1$, 称之为$U$ 的**伴随算子**. U 为等距算子（指$\|Uf\|_{\mathcal{H}_2} = \|f\|_{\mathcal{H}_1}, \forall f \in \mathcal{H}_1$）当且仅当

$$U^* U = \mathrm{id}_{\mathcal{H}_1},$$

且

$$UU^* = P_{\mathrm{Im}U},$$

其中$P_{\mathrm{Im}U}$ 为到U 值域ImU 的投射. U 称为**酉算子**是指它为可逆等距的, 这等价于可逆并且$U^{-1} = U^*$, 也等价于可逆且保内积$\langle Uf, Ug\rangle = \langle f, g\rangle, \forall f, g \in \mathcal{H}_1$.

定理 2.2.1 (von Neumann 遍历定理（Hilbert空间形式）) *设$\mathcal{H}$是一个Hilbert空间和$U : \mathcal{H} \to \mathcal{H}$是一个收缩算子(即$\|Ux\| \le \|x\|, \forall x \in \mathcal{H}$). 令$P$是从$H$到$\mathfrak{I}_U := \{x \in \mathcal{H} : Ux = x\}$的投影算子. 则对任意的$x \in \mathcal{H}$,*

$$\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n x - Px \right\| = 0.$$

证明. 首先我们证明下面的断言

断言:

$$\mathcal{H} = \mathfrak{I}_U \bigoplus \overline{\mathrm{Range}(I - U)}.$$

断言证明. 如果$u \perp \mathrm{Range}(I-U)$, 即对任意的$v \in \mathcal{H}$, $\langle u, (I-U)v\rangle = 0$, 则

$$0 = \langle u, v\rangle - \langle u, Uv\rangle = \langle u, v\rangle - \langle U^*u, v\rangle = \langle u - U^*u, v\rangle, \forall v \in \mathcal{H}.$$

所以$U^*u = u$. 我们需要证明$Uu = u$.

$$\begin{aligned}\|Uu-u\|^2 &= \langle Uu-u, Uu-u\rangle \\ &= \langle Uu, Uu\rangle - \langle Uu, u\rangle - \langle u, Uu\rangle + \langle u, u\rangle \\ &= \|Uu\|^2 - \langle u, U^*u\rangle - \langle U^*u, u\rangle + \|u\|^2 \\ &\le \|u\|^2 - \langle u, u\rangle - \langle u, u\rangle + \|u\|^2 = 0.\end{aligned}$$

故$Uu = u$, 即$u \in \mathfrak{I}_U$.

反之, 设$u \in \mathfrak{I}_U$, 则$Uu = u$. 因为

$$\begin{aligned}\|U^*v\|^2 &= |\langle U^*v, U^*v\rangle| = |\langle UU^*v, v\rangle| \\ &\le \|UU^*v\|\|v\| \le \|U^*v\|\|v\|.\end{aligned}$$

所以U^* 也为收缩算子. 根据上面证明, 根据$Uu = u$ 以及U^* 收缩, 得到$U^*u = u$.

于是我们有

$$\langle u, v - Uv\rangle = \langle u, v\rangle - \langle u, Uv\rangle = \langle u - U^*u, v\rangle = 0,\ \forall v \in \mathcal{H}.$$

所以$u \perp \mathrm{Range}(I-U)$. 断言证毕. □

我们分三种情况讨论：

（1）如果$x \in \mathfrak{I}_U$, 则$Ux = x$和$Px = x$, 从而

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^n x - Px\right\| = \left\|\frac{1}{N}\sum_{n=0}^{N-1} x - x\right\| = 0.$$

（2）如果$x \in \mathrm{Range}(I-U)$, 则存在$y \in \mathcal{H}$ 使得$x = y - Uy$. 从而

$$\begin{aligned}\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^n x\right\| &= \lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^n (y - Uy)\right\| \\ &= \lim_{N\to\infty}\frac{1}{N}\|y - U^N y\| \\ &\le \lim_{N\to\infty}\frac{\|y\| + \|U^N y\|}{N} \le \lim_{N\to\infty}\frac{2\|y\|}{N} = 0.\end{aligned}$$

如果$x \in \overline{\mathrm{Range}(I-U)}$, 则存在$y_i \in \mathcal{H}$使得$x_i = y_i - Uy_i \to x, i \to \infty$. 对于任何$\varepsilon > 0$, 取$i$使得$\|x_i - x\| < \frac{\varepsilon}{2}$, 并且取$M$使得当$N > M$时有

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^n x_i\right\| < \frac{\varepsilon}{2}.$$

于是$N > M$时有

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}U^n x\right\| \le \left\|\frac{1}{N}\sum_{n=0}^{N-1}U^n (x-x_i)\right\| + \left\|\frac{1}{N}\sum_{n=0}^{N-1}U^n x_i\right\| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

（3）一般地, 根据断言, $x=y+z$, 其中$y=Px\in\mathfrak{I}_U, z\in\overline{\mathrm{Range}(I-U)}$. 故根据(1)(2)

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}U^n x - Px\right\| = 0.$$

定理证毕! □

注记 2.2.2 如果U 为等距算子, 那么断言证明会略微简单些:

设$u\in\mathcal{H}$. 如果$u\in\mathrm{Range}(I-U)$, 即存在$v\in\mathcal{H}$使得$u=v-Uv$, 则对任意的$w\in\mathfrak{I}_U$有,

$$\begin{aligned}\langle u,w\rangle &= \langle v-Uv,w\rangle = \langle v,w\rangle - \langle Uv,w\rangle\\ &= \langle v,w\rangle - \langle Uv,Uw\rangle = \langle v,w\rangle - \langle v,w\rangle = 0.\end{aligned}$$

故$u\perp\mathfrak{I}_U$.

如果$u\perp\mathrm{Range}(I-U)$, 即对任意的$v\in\mathcal{H}$, $\langle u,(I-U)v\rangle=0$, 则$\langle u,Uu\rangle=\langle u,u\rangle=\overline{\langle u,u\rangle}=\overline{\langle u,Uu\rangle}=\langle Uu,u\rangle$. 由于$U$是一个等距算子, $\langle Uu,Uu\rangle=\langle u,u\rangle$. 所以

$$\begin{aligned}\langle Uu-u,Uu-u\rangle &= \langle Uu,Uu\rangle - \langle Uu,u\rangle - \langle u,Uu\rangle + \langle u,u\rangle\\ &= \langle u,u\rangle - \langle u,u\rangle - \langle u,u\rangle + \langle u,u\rangle = 0.\end{aligned}$$

故$Uu=u$, 即$u\in\mathfrak{I}_U$.

§2.2.2 von Neumann 遍历定理（L^p空间形式）

设$(X,\mathcal{X},\mu,T)$是一个保测动力系统, 则Koopman算子

$$U_T: L^2(X,\mu)\to L^2(X,\mu),\ f\mapsto f\circ T$$

是一个等距算子. 将定理2.2.1应用到U_T, 我们得到下面的von Neumann平均遍历定理.

定理 2.2.3 (von Neumann平均遍历定理) 设$(X,\mathcal{X},\mu,T)$是一个保测系统, 设P_T 是从$L^2(X,\mu)$到$\{f\in L^2(X,\mu): U_T f = f\}$的投影映射. 则对任意的$f\in L^2(\mu)$.

$$\lim_{n\to\infty}\left\|\mathbb{A}_n f - P_T f\right\|_2 = \lim_{n\to\infty}\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^i f - P_T f\right\|_2 = 0.$$

当μ是遍历时, $P_T f = \int_X f d\mu$.

注记 2.2.4 之前我们定义了$\mathcal{I}=\mathcal{I}(T)=\{A\in\mathcal{X}\colon \mu(A\Delta T^{-1}A)=0\}$. 易见

$$P_T f=\mathbb{E}(f|\mathcal{I}).$$

所以平均遍历定理可以陈述为:

$$\mathbb{A}_n f\xrightarrow{L^2}\mathbb{E}(f|\mathcal{I}), n\to\infty.$$

在给出下面定理之前, 我们先回顾一个关于L^p 空间的一个结果.

定理 2.2.5 设$(X,\mathcal{X},\mu)$为测度空间, $\mu(X)<\infty$, 并且$0<p<q\le\infty$, 那么有$L^q(\mu)\subseteq L^p(\mu)$, 并且

$$\|f\|_p\le\|f\|_q\mu(X)^{\frac{1}{p}-\frac{1}{q}}.$$

定理 2.2.6 (L^1-平均遍历定理) 设$(X,\mathcal{X},\mu,T)$ 是一个保测动力系统. 对任意的$f\in L^1(X,\mu)$, 存在一个T不变的函数$f^*\in L^1(X,\mu)$ 使得$\int fd\mu=\int f^*d\mu$ 且

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^if-f^*\right\|_1=0.$$

当μ是遍历时, $f^*=\int_X fd\mu$.

证明. 设$g\in L^\infty(X,\mu)\subseteq L^2(X,\mu)$, 由von Neumann 平均遍历定理存在$g^*\in L^2(X,\mu)$ 使得

$$\frac{1}{n}\sum_{i=0}^{n-1}T^ig\xrightarrow{L^2}g^*, n\to\infty.$$

因为$\|\cdot\|_1\le\|\cdot\|_2$, 我们得到

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^ig-g^*\right\|_1\to 0, n\to\infty.$$

设$f\in L^1(\mu)$ 和$\varepsilon>0$. 选取$g\in L^\infty(\mu)$ 使得$\|f-g\|_1<\varepsilon$. 则对任意的$n\in\mathbb{N}$,

$$\begin{aligned}\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^if-\frac{1}{n}\sum_{i=0}^{n-1}T^ig\right\|_1&\le\frac{1}{n}\sum_{i=0}^{n-1}\|T^if-T^ig\|_1\\&=\frac{1}{n}\sum_{i=0}^{n-1}\|f-g\|_1<\varepsilon.\end{aligned}$$

另一方面存在$N\in\mathbb{N}$ 使得对任意的$n,m\ge N$,

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^ig-\frac{1}{m}\sum_{i=0}^{m-1}T^ig\right\|_1<\varepsilon.$$

故对任意的$n,m\geq N$,

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^if-\frac{1}{m}\sum_{i=0}^{m-1}T^if\right\|_1\leq 3\varepsilon.$$

这说明$\{\frac{1}{n}\sum_{i=0}^{n-1}T^if\}_n$ 是$L^1(X,\mu)$ 中的一个Cauchy列, 从而存在极限, 记之为$f^*\in L^1(X,\mu)$. 由于对任意的$n\geq 1$,

$$\left\|T\left(\frac{1}{n}\sum_{i=0}^{n-1}T^if\right)-\frac{1}{n}\sum_{i=0}^{n-1}T^if\right\|_1\leq\frac{2}{n}\|f\|_1,$$

故f^* 是T 不变的.

根据

$$\begin{aligned}\left|\int_X f^*d\mu-\int_X fd\mu\right|&=\left|\int_X f^*d\mu-\int_X\frac{1}{n}\sum_{i=0}^{n-1}T^ifd\mu\right|\\&\leq\int_X\left|f^*-\frac{1}{n}\sum_{i=0}^{n-1}T^if\right|d\mu\\&=\left\|f^*-\frac{1}{n}\sum_{i=0}^{n-1}T^if\right\|_1\to 0,\quad n\to\infty,\end{aligned}$$

有

$$\int_X f^*d\mu=\int_X fd\mu.$$

当μ是遍历时, f^*几乎处处为常值, 故$f^*=\int_X fd\mu$. □

注记 2.2.7 实际上对于$g\in L^\infty(X,\mu)$, 可以证明$g^*\in L^\infty(X,\mu)$.

首先易见$\|\frac{1}{n}\sum_{i=0}^{n-1}T^ig\|_\infty\leq\|g\|_\infty$ 和对任意的$B\in\mathcal{X}$,

$$\left|\left\langle\frac{1}{n}\sum_{i=0}^{n-1}T^ig,1_B\right\rangle\right|\leq\|g\|_\infty\mu(B).$$

则对任意的$B\in\mathcal{X}$,

$$|\langle g^*,1_B\rangle|\leq\|g\|_\infty\mu(B).$$

于是$\|g^*\|_\infty\leq\|g\|_\infty$.

也可以如下证明: 因为$\frac{1}{n}\sum_{i=0}^{n-1}T^ig\xrightarrow{L^2}g^*,n\to\infty$, 于是存在子列$\frac{1}{n_k}\sum_{i=0}^{n_k-1}T^ig\to g^*$ 几乎处处成立. 于是根据$\|\frac{1}{n}\sum_{i=0}^{n-1}T^ig\|_\infty\leq\|g\|_\infty$ 得到$\|g^*\|_\infty\leq\|g\|_\infty$.

类似地, 我们可以证明下面定理:

定理 2.2.8 设$(X,\mathcal{X},\mu,T)$是一个保测动力系统和$1\leq p<\infty$. 证明对任意的$f\in L^p(X,\mu)$, 存在一个T 不变的函数$f^*\in L^p(X,\mu)$ 使得

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{i=0}^{n-1}T^if-f^*\right\|_p=0.$$

我们在下一节会给出这个定理的另一个证明.

§2.2.3　Khintchine 定理

根据推论2.1.5 ，如果$(X,\mathcal{X},\mu,T)$ 为遍历系统, 那么对任意满足$\mu(A)>0$ 的$A\in\mathcal{X}$, $N^{\mu}(A,A)=\{n\in\mathbb{N}:\mu(A\cap T^{-n}A)>0\}$ 为syndetic 的. 实际上根据Poincaré 回复定理，上述结论对于非遍历保测系统仍成立(§1.3 习题2). Khintchine 定理加强了上述命题.

定理 2.2.9 (Khintchine, 1934) *设$(X,\mathcal{X},\mu,T)$为保测系统. $A\in\mathcal{X}$, $\mu(A)>0$. 则对于任何$\varepsilon>0$,*

$$\{n\in\mathbb{N}:\mu(A\cap T^{-n}A)>\mu(A)^2-\varepsilon\}$$

为syndetic的.

证明. 设$\varepsilon>0$. 根据von Neunmann 遍历定理, 存在$N\in\mathbb{N}$ 使得

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}T^n1_A-P1_A\right\|_2<\sqrt{\frac{\varepsilon}{2}},$$

其中P 为到不变函数空间的投射. 因为$P1_A$ 为T 不变的, 所以对于任何$l\in\mathbb{N}$ 我们有

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}T^{n+l}1_A-P1_A\right\|_2<\sqrt{\frac{\varepsilon}{2}}.$$

记$S_N=\frac{1}{N}\sum_{n=0}^{N-1}T^n1_A$, 则根据上面两个不等式有

$$\left\|T^lS_N-S_N\right\|_2<2\sqrt{\frac{\varepsilon}{2}},\ \forall l\in\mathbb{N}.$$

平方之,

$$\langle T^lS_N-S_N,T^lS_N-S_N\rangle<4\frac{\varepsilon}{2}=2\varepsilon.$$

展开约化得到：

$$\|S_N\|_2^2-\langle S_N,T^lS_N\rangle<\varepsilon,\ \forall l\in\mathbb{N}.$$

上式结合

$$0\le\|S_N-\mu(A)\|_2^2=\langle S_N-\mu(A),S_N-\mu(A)\rangle=\|S_N\|_2^2-\mu(A)^2,$$

我们得到

$$\langle S_N,T^lS_N\rangle>\|S_N\|_2^2-\varepsilon\ge\mu(A)^2-\varepsilon,\ \forall l\in\mathbb{N}.$$

但是,

$$\begin{aligned}\langle S_N,T^lS_N\rangle&=\left\langle\frac{1}{N}\sum_{n=0}^{N-1}T^n1_A,\frac{1}{N}\sum_{m=0}^{N-1}T^{m+l}1_A\right\rangle\\&=\frac{1}{N^2}\sum_{n,m=0}^{N-1}\mu(T^{-n}A\cap T^{-m-l}A)\\&=\frac{1}{N^2}\sum_{n,m=0}^{N-1}\mu(A\cap T^{-(m-n+l)}A),\end{aligned}$$

所以存在$n,m\in\{0,1,\dots,N-1\}$ 使得

$$\mu(A\cap T^{-(m-n+l)}A)>\mu(A)^2-\varepsilon.$$

由$-N<m-n<N$ 以及$l\in\mathbb{N}$ 任意, 我们得到对于长为$2N$ 的区间内必存在k 使得

$$\mu(A\cap T^{-k}A)>\mu(A)^2-\varepsilon.$$

即$\{n\in\mathbb{N}:\mu(A\cap T^{-n}A)>\mu(A)^2-\varepsilon\}$ 为syndetic的. 证毕! □

注记 2.2.10 *Bergelson-Host-Kra* (*2005*)*[16]* 证明了如下结论: 设$(X,\mathcal{X},\mu,T)$为遍历系统, $A\in\mathcal{X}$, $\mu(A)>0$. 则对于任何$\varepsilon>0$,

$$\{n\in\mathbb{N}:\mu(A\cap T^{-n}A\cap T^{-2n}A)>\mu(A)^3-\varepsilon\}$$

以及

$$\{n\in\mathbb{N}:\mu(A\cap T^{-n}A\cap T^{-2n}A\cap T^{-3n}A)>\mu(A)^4-\varepsilon\}$$

为*syndetic*的.

但是更高阶项没有类似结论: 存在遍历系统$(X,\mathcal{X},\mu,T)$, 对于任何$l\in\mathbb{N}$, 存在$A=A(l)\in\mathcal{X}$, $\mu(A)>0$ 使得

$$\mu(A\cap T^{-n}A\cap T^{-2n}A\cap T^{-3n}A\cap T^{-4n}A)\le\mu(A)^l,\ \forall n\in\mathbb{N}.$$

对于非遍历系统*Khintchine* 是最佳结论: 存在非遍历系统$(X,\mathcal{X},\mu,T)$, 对于任何$l\in\mathbb{N}$, 存在$A=A(l)\in\mathcal{X}$, $\mu(A)>0$ 使得

$$\mu(A\cap T^{-n}A\cap T^{-2n}A)\le\mu(A)^l,\ \forall n\in\mathbb{N}.$$

习 题

1. 证明定理2.2.8.
2. 设$(X,\mathcal{X},\mu,T)$是一个保测系统, 设P_T 是从$L^2(X,\mu)$ 到$\{f\in L^2(X,\mu):U_Tf=f\}$的投影映射. 则对任意的$f\in L^2(\mu)$.
$$\lim_{n-m\to\infty}\left\|\frac{1}{n-m}\sum_{i=m}^{n-1}T^if-P_Tf\right\|_2=0.$$

§2.3 Birkhoff 遍历定理

在本节中我们介绍著名的Birkhoff 遍历定理, 目前关于这个定理的证明很多, 我们在这里给出三个证明, 前两个是目前教材中标准的证明, 第三个是[1]中介绍的简短证明. 之后我们给出定理的一些应用.

§2.3.1 Birkhoff 逐点遍历定理

下面我们给出著名的Birkhoff遍历定理, 注意在表述和证明这个定理的时候我们假设空间是σ 有限的.

定理 2.3.1 (Birkhoff 逐点遍历定理) 设$(X,\mathcal{X},\mu,T)$ 是一个保测系统(此处我们假设$(X,\mathcal{X},\mu)$为σ 有限的). 对任意的$f\in L^1(X,\mu)$, 遍历平均

$$\mathbb{A}_n f(x)=\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)$$

几乎处处收敛于函数$f^*(x)$, 其中$f^*(x)\in L^1(X,\mu)$ 且满足$f^*\circ T=f^*$ *a.e.*

如果$\mu(X)<\infty$, 那么$\int_X f^*d\mu=\int_X fd\mu$; 如果$\mu$ 还为遍历的, 则$f^*(x)=\frac{1}{\mu(X)}\int_X fd\mu, a.e.$.

注记 2.3.2 *1.* 遍历性的原始定义是依据遍历性假设给出的. 所谓遍历假设就是指"时间平均=空间平均". 对于函数f, 时间平均是指$\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)$, 而空间平均是指$\frac{1}{\mu(X)}\int_X fd\mu$. 于是遍历性假设就是指

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)=\frac{1}{\mu(X)}\int_X fd\mu.$$

2. 设$(X,\mathcal{X},\mu,T)$为保测系统, $A\in\mathcal{X}$. 对$x\in X$, 一个自然的问题是: 其轨道$\mathrm{orb}(x,T)$将以何种频率进入集合A? 明显地, $T^ix\in A$ 当且仅当$\mathbf{1}_A(T^ix)=1$, 于是$x,Tx,\dots,T^{n-1}x$在A 中的个数就等于$\sum_{i=0}^{n-1}\mathbf{1}_A(T^ix)$. 取平均即为$\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{1}_A(T^ix)$, 如果此时$T$ 为遍历的, 那么由遍历定理就有$\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{1}_A(T^ix)\to\mu(A)$ *a.e.*.

3. 根据定理容易看出遍历定理比回复定理提供了的更多的信息, 由*Birkhoff* 遍历定理可以直接推导出*Poincaré* 回复定理.

下面我们给出证明, 首先需要几个引理.

定理 2.3.3 设$(X,\mathcal{X},\mu,T)$ 是一个保测系统. 设$U: L^1_{\mathbb{R}}(X,\mu)\to L^1_{\mathbb{R}}(X,\mu)$ 是一个正的线性算子且$\|U\|\le 1$. 设$N\in\mathbb{N}$, $f\in L^1_{\mathbb{R}}(X,\mu)$. 递归定义

$$f_0=0,\ f_1=f,\ f_2=f+Uf,$$
$$f_n=f+Uf+\ldots+U^{n-1}f,\ \forall n\ge 1$$

和$F_N=\max_{0\le n\le N}f_n\ge 0$. 则

$$\int_{\{x\in X:F_N(x)>0\}}fd\mu\ge 0.$$

证明. 首先$F_N \in L^1(X,\mu)$, 且对于任意$0 \le n \le N$, $F_N \ge f_n$. 由于U是一个正线性算子, 我们有

$$UF_N + f \ge Uf_n + f = f_{n+1}.$$

故

$$UF_N + f \ge \max_{1\le n\le N} f_n.$$

令$A = \{x \in X : F_N(x) > 0\}$. 由于$f_0 = 0$, 当$x \in A$时, 有

$$F_N(x) = \max_{0\le n\le N} f_n(x) = \max_{1\le n\le N} f_n(x).$$

因此, 当$x \in A$时, 有

$$UF_N(x) + f(x) \ge F_N(x),$$

即

$$f(x) \ge F_N(x) - UF_N(x).$$

对任意的$x \in X$, 由于$F_N(x) \ge 0$, 所以$UF_N(x) \ge 0$. 由于$\|U\| \le 1$, 有

$$\begin{aligned}\int_A f d\mu &\ge \int_A F_N d\mu - \int_A UF_N d\mu \\ &= \int_X F_N d\mu - \int_A UF_N d\mu \\ &\ge \int_X F_N d\mu - \int_X UF_N d\mu \\ &= \|F_N\|_1 - \|UF_N\|_1 \ge 0.\end{aligned}$$

证毕! □

定理 2.3.4 设$(X, \mathcal{X}, \mu, T)$ 是一个保测系统, 且$g \in L^1_{\mathbb{R}}(X,\mu)$. 对于$\alpha \in \mathbb{R}$, 定义

$$B_\alpha = \Big\{x \in X : \sup_{n\ge 1} \frac{1}{n}\sum_{i=0}^{n-1} g(T^i x) > \alpha\Big\}.$$

则

$$\alpha\mu(B_\alpha) \le \int_{B_\alpha} g d\mu \le \|g\|_1.$$

进一步, 如果$A \in \mathcal{X}$ 满足$T^{-1}A = A$, 则

$$\alpha\mu(B_\alpha \cap A) \le \int_{B_\alpha \cap A} g d\mu.$$

证明. 设$U_T : L^1(X,\mu) \to L^1(X,\mu)$, $f \mapsto f \circ T$. 令$f = g - \alpha$. 容易验证:

$$B_\alpha = \bigcup_{N=0}^{\infty} \{x \in X : F_N(x) > 0\}.$$

由定理2.3.3 $\int_{B_\alpha} f d\mu \geq 0$, 即

$$\int_{B_\alpha} g d\mu(x) \geq \alpha\mu(B_\alpha).$$

如果$T^{-1}A = A$, 我们将前一部分的结论用于系统$(A, \mathcal{X} \cap A, \mu_A, T|_A)$ 即可. □

注记 2.3.5 根据类似的证明, 可以证明对于任何$f \in L^1(X, \mu)$, $\alpha > 0$,

$$\mu(\{x \in X : \limsup_{n\to\infty} \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| > \alpha\}) \leq \mu(\{x \in X : \sup_{n\geq 1} \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| > \alpha\}) \leq \frac{\|f\|_1}{\alpha}.$$

定理2.3.3 或定理2.3.4 经常被称为**最大遍历定理**（Maximal Ergodic Theorem）.

Birkhoff 定理证明. 我们通过分解为实部和虚部, 仅需考虑$f \in L^1_{\mathbb{R}}(X, \mu)$, 即考虑实值函数. 令

$$f^*(x) = \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x), \quad f_*(x) = \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x).$$

易见

$$\frac{n+1}{n}\left(\frac{1}{n+1}\sum_{i=0}^{n} f(T^i x)\right) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(Tx)) + \frac{1}{n}f(x).$$

即

$$\frac{n+1}{n}\mathbb{A}_{n+1}f(x) = \mathbb{A}_n f(Tx) + \frac{f(x)}{n}.$$

由此我们得到

$$f^* = f^* \circ T \text{ 和 } f_* = f_* \circ T.$$

第 *(i)* 步：如果$\mu(X) < \infty$，那么$f^* = f_*$ $a.e$，进而$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) \xrightarrow{a.e} f^*, n \to \infty$.

对于任何两个实数α, β, 令

$$E_{\alpha,\beta} = \{x \in X : f_*(x) < \beta \text{ 并且 } \alpha < f^*(x)\}.$$

于是

$$\{x \in X : f_*(x) < f^*(x)\} = \bigcup_{\beta<\alpha, \alpha,\beta\in\mathbb{Q}} E_{\alpha,\beta}.$$

于是为证$f_* = f^*$ $a.e.$, 我们仅需证明对于$\beta < \alpha$, $\mu(E_{\alpha,\beta}) = 0$.

因为$f^* = f^* \circ T$ 和 $f_* = f_* \circ T$, 所以$T^{-1}E_{\alpha,\beta} = E_{\alpha,\beta}$. 令

$$B_\alpha = \left\{x \in X : \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) > \alpha\right\}.$$

因为$\sup\limits_{n\geq 1}\dfrac{1}{n}\sum\limits_{i=0}^{n-1}f(T^ix) \geq \inf\limits_{k\geq 1}\sup\limits_{n\geq k}\dfrac{1}{n}\sum\limits_{i=0}^{n-1}f(T^ix) = f^*(x)$, 所以$E_{\alpha,\beta} \subseteq B_\alpha$, 即$E_{\alpha,\beta}\cap B_\alpha = E_{\alpha,\beta}$. 由定理2.3.4得到

$$\alpha\mu(E_{\alpha,\beta}) = \alpha\mu(E_{\alpha,\beta}\cap B_\alpha) \leq \int_{E_{\alpha,\beta}\cap B_\alpha} fd\mu = \int_{E_{\alpha,\beta}} fd\mu. \tag{2.3.1}$$

即$\alpha\mu(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} fd\mu$.

将f, α, β 分别替换成$-f$, $-\beta$, $-\alpha$, 此时$(-f)^* = -(f_*)$, $(-f)_* = -(f^*)$ 和

$$E_{\alpha,\beta} = \{x\in X : (-f)^*(x) > -\beta \text{ 并且} (-f)_*(x) < -\alpha\}.$$

再次利用定理2.3.4,

$$-\beta\mu(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}}(-f)d\mu,$$

即

$$\int_{E_{\alpha,\beta}} fd\mu \leq \beta\mu(E_{\alpha,\beta}). \tag{2.3.2}$$

综上，根据(2.3.1)和(2.3.2)我们得到

$$\alpha\mu(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} fd\mu \leq \beta\mu(E_{\alpha,\beta}).$$

因为$\alpha > \beta$ 以及$\mu(E_{\alpha,\beta}) < \mu(X) < \infty$, 故$\mu(E_{\alpha,\beta}) = 0$. 所以

$$\mu(\{x\in X : f_*(x) < f^*(x)\}) = 0,$$

这也就是说$f^*(x) = f_*(x)$ 几乎处处成立, 即

$$\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix) \xrightarrow{a.e} f^*, n\to\infty.$$

第*(ii)*步： $f^*\in L^1(X,\mu)$.

由Fatou引理

$$\begin{aligned}\int_X |f^*(x)|d\mu(x) &= \int_X \lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)\right|d\mu(x)\\ &\leq \liminf_{n\to\infty}\int_X\left|\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)\right|d\mu(x)\\ &\leq \liminf_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\int_X|f(T^ix)|d\mu(x) = \int_X|f(x)|d\mu(x).\end{aligned}$$

故$f^*\in L^1(X,\mu)$.

第(*iii*)步：如果$\mu(X)<\infty$, 那么$\int_X f^* d\mu = \int_X f d\mu$.

下面我们证明$\int_X f^* d\mu = \int_X f d\mu$. 对任意的$k\in\mathbb{Z}$和$n\in\mathbb{N}$, 令

$$D_k^n = \{x\in X : \frac{k}{n} \le f^*(x) < \frac{k+1}{n}\}.$$

对充分小的$\varepsilon>0$, $D_k^n \cap B_{\frac{k}{n}-\varepsilon} = D_k^n$. 因为$f^*$ 是不变函数，所以$T^{-1}D_k^n = D_k^n$. 根据定理2.3.4,

$$\Big(\frac{k}{n}-\varepsilon\Big)\mu(D_k^n) \le \int_{D_k^n} f d\mu,$$

从而由ε 任意性，我们有

$$\frac{k}{n}\mu(D_k^n) \le \int_{D_k^n} f d\mu.$$

于是

$$\int_{D_k^n} f^* d\mu \le \frac{k+1}{n}\mu(D_k^n) \le \frac{1}{n}\mu(D_k^n) + \int_{D_k^n} f d\mu.$$

上式对$k\in\mathbb{Z}$进行求和得到

$$\int_X f^* d\mu \le \frac{\mu(X)}{n} + \int_X f d\mu.$$

令$n\to\infty$ 就有：

$$\int_X f^* d\mu \le \int_X f d\mu.$$

用$-f$ 替代f 重复上面的讨论得$\int_X (-f)^* d\mu \le \int_X (-f) d\mu$, 从而有

$$\int_X f_* d\mu \ge \int_X f d\mu.$$

再根据$f_* = f^*$ 得$\int_X f^* d\mu = \int_X f d\mu$.

第(*iv*)步：当$\mu(X)=\infty$ 时，第(*i*)步的结论仍成立.

当$\mu(X)=\infty$ 时, 一旦我们证明对于$\beta<\alpha$, $\mu(E_{\alpha,\beta})<\infty$, 那么上面第(i)步的证明仍然可行. 于是我们仅需证明：对于$\beta<\alpha$, $\mu(E_{\alpha,\beta})<\infty$.

首先假设$\alpha>0$. 根据μ 为σ有限的, 取$C\in\mathcal{X}$ 使得$C\subseteq E_{\alpha,\beta}$ 并且$\mu(C)<\infty$. 令

$$h = f - \alpha \mathbf{1}_C \in L^1(X,\mu).$$

由定理2.3.3,

$$\int_{\{x\in X: H_N(x)>0\}} (f-\alpha\mathbf{1}_C) d\mu \ge 0, \tag{2.3.3}$$

其中H_N 为如定理2.3.3中h 对应的函数. 设$x \in E_{\alpha,\beta}$, 则$f^*(x) > \alpha$. 于是$\sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) > \alpha$. 继而

$$\sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} h(T^i x) = \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} \big(f-\alpha_C\big)(T^i x) \geq \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} \big(f-\alpha\big)(T^i x) > 0.$$

即$E_{\alpha,\beta} \subseteq \{x \in X : \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} h(T^i x) > 0\}$. 于是根据(2.3.3) 以及

$$C \subseteq E_{\alpha,\beta} \subseteq \{x \in X : \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} h(T^i x) > 0\} = \bigcup_{N=0}^{\infty} \{x \in X : H_N(x) > 0\},$$

我们得到$\alpha\mu(C) \leq \int_X |f| d\mu = \|f\|_1$.

综上，对于任何满足$C \subseteq E_{\alpha,\beta}$ 和$\mu(C) < \infty$ 的$C \in \mathcal{X}$ 我们有

$$\mu(C) \leq \frac{1}{\alpha}\|f\|_1.$$

因为X 为σ有限的, 根据上面事实我们容易得到$\mu(E_{\alpha,\beta}) < \infty$.

如果$\alpha \leq 0$, 则$\beta < 0$, 我们用$-f$, $-\beta$ 和$-\alpha$ 替代上面的f, α 和β, 可以得到: 对于任何满足$C \subseteq E_{\alpha,\beta}$ 和$\mu(C) < \infty$ 的$C \in \mathcal{X}$ 我们有$\mu(C) \leq \frac{1}{-\beta}\|f\|_1$. 进而同样推出$\mu(E_{\alpha,\beta}) < \infty$. 整个证明结束. □

注记 2.3.6 *在定理2.3.1的证明过程中, 我们也可以使用L^1 平均遍历定理(定理2.2.6) 来证明$f^* \in L^1(X,\mu)$ 和$\int_X f^* d\mu = \int_X f d\mu$. 根据定理2.2.6 存在$\hat{f} \in L^1(X,\mu)$ 使得$\int_X \hat{f} d\mu = \int_X f d\mu$ 和*

$$\lim_{n\to\infty} \Big\| \frac{1}{n}\sum_{i=0}^{n-1} f\circ T^i - \hat{f} \Big\|_1 = 0.$$

由定理0.1.32, 存在一个子序列$\{n_k\}$ 使得$\frac{1}{n_k}\sum_{i=0}^{n_k-1} f\circ T^i$几乎处处收敛到$\hat{f}$. 故由唯一性$f^* = \hat{f}$.

§2.3.2 第二个证明

在上面证明中，我们看到困难的地方在于证明$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) \xrightarrow{a.e} f^*, n\to\infty$. 在这里我们给出另外一个证明, 这个证明十分类似于von Neumann 遍历定理的证明.

定理 2.3.7 *设$(X,\mathcal{X},\mu,T)$ 为保测系统, 那么$B = \{f\circ T - f : f \in L^1(X,\mu)\}$ 在*

$$\mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I}) = \{f \in L^1(X,\mu) : \mathbb{E}(f|\mathfrak{I}) = 0\}$$

中稠密.

尤其, 我们有

$$L^1(X,\mu)=\overline{B}+I,$$

其中$\overline{B}\cap I=\{0\}$, $I=\{f\in L^1(X,\mu): f\circ T=f\}$.

证明. 因为

$$\mathbb{E}(f\circ T-f|\mathfrak{I})=\mathbb{E}(f\circ T|\mathfrak{I})-\mathbb{E}(f|\mathfrak{I})=\mathbb{E}(f|\mathfrak{I})\circ T-\mathbb{E}(f|\mathfrak{I})=0,$$

所以$B\subseteq \mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I})$. 为了证明$B$ 在$\mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I})$ 中稠密, 由$L^\infty=(L^1)^*$, 我们仅需证明对任何$h\in L^\infty(X,\mu)$, 如果$\int_X ghd\mu=0,\forall g\in B$, 那么$\int_X ghd\mu=0,\forall g\in \mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I})$.

设$h\in L^\infty(X,\mu)$, 并且

$$\int_X (f\circ T-f)hd\mu=0,\ \forall f\in L^1(X,\mu).$$

因为$h\in L^1(X,\mu)$, 所以在上式中令$f=h$ 得到

$$\int_X (h\circ T-h)hd\mu=0.$$

由此式容易得到$\int_X (h\circ T-h)h\circ Td\mu=0$, 于是二式相减就有

$$\int_X (h\circ T-h)^2d\mu=0.$$

即$h\circ T=h$ $a.e.$ 由此对于任何$g\in \mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I})$ (即$\mathbb{E}(g|\mathfrak{I})=0$), 我们有

$$\int_X ghd\mu=\int_X \mathbb{E}(gh|\mathfrak{I})d\mu=\int_X h\mathbb{E}(g|\mathfrak{I})d\mu=0.$$

综上我们证明了B 在$\mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I})$ 中稠密. 根据

$$f=(f-\mathbb{E}(f|\mathfrak{I}))+\mathbb{E}(f|\mathfrak{I}),\forall f\in L^1(X,\mu)$$

我们得到$L^1(X,\mu)=\overline{B}+I$. 而$\overline{B}\cap I=\{0\}$ 是显然的, 于是整个证明完毕! □

Birkhoff 定理证明. 设

$$B_0=\{g\circ T-g: g\in L^\infty(X,\mu)\}.$$

易见$\overline{B_0}=\overline{B}$, 闭包在$L^1(X,\mu)$ 中取. 首先易验证对于任何$g\in B_0+I$, $\frac{1}{n}\sum_{i=0}^{n-1}g(T^ix)$ 几乎处处收敛.

根据定理2.3.7, $L^1(X,\mu)=\overline{B}+I=\overline{B_0}+I$, 我们仅需证明对于任何$f\in\overline{B_0}=\mathrm{Ker}\mathbb{E}(\cdot|\mathfrak{I})$,

$$\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)\xrightarrow{a.e}0,\ n\to\infty.$$

设$f \in L^1(X,\mu)$ 满足$\mathbb{E}(f|\mathfrak{I}) = 0$, 令

$$R(f,x) = \limsup_{n\to\infty} \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \right|.$$

因为对于任何$g \in B_0$, $R(g,x) = 0$, 容易验证$R(f,x) = R(f-g,x), \forall g \in B_0$. 设$\varepsilon > 0$, 取$g \in B_0$使得$\|f-g\|_1 < \varepsilon^2$. 根据注记2.3.5,

$$\mu(\{x \in X : R(f,x) > \varepsilon\}) = \mu(\{x \in X : R(f-g,x) > \varepsilon\}) \le \frac{\|f-g\|_1}{\varepsilon} < \varepsilon.$$

由此$R(f,x) = 0$, $a.e.$ 证毕! □

§2.3.3 一个简短证明

下面我们引用[1]中给出的Birkhoff 定理简短证明. 我们证明：设$(X, \mathcal{X}, \mu, T)$ 是一个保测系统, 对任意的$f \in L^1(X,\mu)$,

$$\frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \xrightarrow{a.e} f^*, n \to \infty.$$

Birkhoff 定理的简短证明: 设$\mathbb{A}_n f(x) = \dfrac{1}{n} \sum\limits_{i=0}^{n-1} f(T^i(x))$. 对$\phi \in L^1(X,\mu)$, 令

$$M_n\phi = \max\left\{ \sum_{j=0}^{k-1} \phi \circ T^j : 1 \le k \le n \right\},$$

此时易见,$\mathbb{A}_n\phi \le \frac{1}{n} M_n\phi$.设

$$A = A(\phi) = \{x \in X : \sup_n M_n\phi(x) = \infty\} \in \mathcal{I}_T.$$

对于任意$x \in A^c$, 有

$$\limsup_{n\to\infty} \mathbb{A}_n\phi(x) \le 0.$$

注意$M_n\phi$为递增函数列, 并且

$$M_n\phi(Tx) = \max\left\{ \sum_{j=1}^{k-1} \phi(T^j x) : 2 \le k \le n+1 \right\}.$$

于是有

$$M_{n+1}\phi(x) - M_n\phi(Tx) = \phi(x) - \min\{0, M_n\phi(Tx)\} \ge \phi(x).$$

这样, $M_n\phi(Tx)$与$M_{n+1}\phi(x)$在同一k值取到时, 此时它们相差一个$\phi(x)$; 或者

$$M_{n+1}\phi(x) = \phi(x) > \phi(x) + M_n\phi(Tx),$$

后者当且仅当$M_n\phi(Tx)<0$时成立.

于是在集合A上,序列$M_{n+1}\phi(x)-M_n\phi(Tx)$递减趋于$\phi$, 且由Lebesgue 控制收敛定理有

$$0\leq\int_A(M_{n+1}\phi-M_n\phi)d\mu\leq\int_A(M_{n+1}\phi-M_n\phi\circ T)d\mu\to\int_A\phi d\mu.$$

设$f\in L^1(\mu)$ 且$\epsilon>0$取定, 且令$\phi=f-f^*-\varepsilon$. 因为$A=A(\phi)\in\mathcal{I}_T$, 我们有$\int_A fd\mu=\int_A f^*d\mu$, 于是

$$\int_A\phi d\mu=\int_A(f-f^*-\varepsilon)d\mu=-\varepsilon\mu(A)\leq 0,$$

这样$\int_A\phi d\mu=0$, 从而$\mu(A(\phi))=0$. 所以

$$\limsup_{n\to\infty}\mathbb{A}_n\phi(x)\leq 0,\forall x\in X\ a.e..$$

由于f^*为T不变的, $\mathbb{A}_n\phi=\mathbb{A}_nf-f^*-\varepsilon$, 我们有

$$\limsup_{n\to\infty}\mathbb{A}_nf(x)\leq f^*+\varepsilon.$$

最后, 对$-f$同样讨论可得

$$\liminf_{n\to\infty}\mathbb{A}_nf(x)\geq f^*-\varepsilon.$$

这样我们完成了整个证明. □

§2.3.4　一些应用

根据定理2.3.1, 我们也可以证明von Neumann 的L^p遍历定理

定理 2.3.8 (von Neumann的L^p遍历定理) 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $1\leq p<\infty$. 如果$f\in L^p(X,\mu)$, 那么存在$f^*\in L^p(X,\mu)$ 满足$f^*\circ T=f^*$ $a.e.$ 且

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x))-f^*(x)\right\|_p\to 0,\ n\to\infty.$$

证明. 设$f\in L^p(X,\mu)$. 我们首先证明$\{\mathbb{A}_nf\}_{n\in\mathbb{N}}$ 为$L^p(X,\mu)$的Cauchy 列, 从而存在$f^*\in L^p(X,\mu)$ 使得$\|\lim_n\mathbb{A}_nf-f^*\|_p=0$. 再根据

$$\frac{n+1}{n}\mathbb{A}_{n+1}f(x)=\mathbb{A}_nf(Tx)+\frac{f(x)}{n}$$

得到$f^*\circ T=f^*$.

下证明$\{\mathbb{A}_nf\}_{n\in\mathbb{N}}$ 为$L^p(X,\mu)$的Cauchy 列. 首先注意$\|\mathbb{A}_nf\|_p\leq\|f\|_p$. 对于任何$\varepsilon>0$, 取$g\in L^\infty(X,\mu)$ 使得

$$\|f-g\|_p<\varepsilon/3.$$

由$g \in L^\infty(X,\mu)$，知$g \in L^1(X,\mu)$. 根据Birkhoff 遍历定理,

$$\frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) \xrightarrow{a.e.} g^*(x), n \to \infty.$$

易见$g^* \in L^\infty(X,\mu)$, 从而$g^* \in L^p(X,\mu)$. 另外我们有

$$\left|\frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) - g^*(x)\right|^p \xrightarrow{a.e.} 0, n \to \infty.$$

于是

$$\left\|\frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) - g^*(x)\right\|_p \to 0, n \to \infty.$$

下面的证明与定理2.2.6证明类似. 取$N = N(\varepsilon, g) \in \mathbb{N}$ 使得$n > N$时, 对于任何$k \in \mathbb{N}$有

$$\|\mathbb{A}_n g - \mathbb{A}_{n+k} g\|_p = \left\|\frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) - \frac{1}{n+k}\sum_{i=0}^{n+k-1} g(T^i(x))\right\|_p < \varepsilon/3.$$

于是对于任何$n > N$以及$k \in \mathbb{N}$,

$$\begin{aligned}\|\mathbb{A}_n f - \mathbb{A}_{n+k} f\|_p &\le \|\mathbb{A}_n f - \mathbb{A}_n g\|_p + \|\mathbb{A}_n g - \mathbb{A}_{n+k} g\|_p + \|\mathbb{A}_{n+k} g - \mathbb{A}_{n+k} f\|_p \\ &\le \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.\end{aligned}$$

即$\{\mathbb{A}_n f\}_{n\in\mathbb{N}}$ 为$L^p(X,\mu)$ 的Cauchy 列. □

回顾非负整数子集密度的定义. 设$F \subseteq \mathbb{Z}_+$. F的**上密度**(upper density)和**下密度**(lower density)定义为

$$\overline{d}(F) = \limsup_{n\to\infty} \frac{|F \cap \{0,1,\dots,n-1\}|}{n}; \quad \underline{d}(F) = \liminf_{n\to\infty} \frac{|F \cap \{0,1,\dots,n-1\}|}{n}.$$

如果$\overline{d}(F) = \underline{d}(F)$, 则称$F$的**密度**(density)存在, 它的密度记为$d(F) = \overline{d}(F) = \underline{d}(F)$.

推论 2.3.9 设$(X,\mathcal{X},\mu,T)$为遍历系统, $A \in \mathcal{X}$. 则对几乎所有点$x \in X$,

$$N(x,A) = \{n \in \mathbb{Z}_+ : T^n x \in A\}$$

的密度为$\mu(A)$.

证明. 设$A \in \mathcal{X}$. 对A 的特征函数$\mathbf{1}_A$ 应用逐点遍历定理2.3.1, 对于几乎所有点$x \in X$,

$$\mu(A) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbf{1}_A(T^i x) = \lim_{n\to\infty} \frac{|N(x,A) \cap \{0,1,\dots,n-1\}|}{n}.$$

证毕! □

定理 2.3.10 (Borel正则数定理) *Lebsgue*测度下几乎处处所有的$x \in (0,1)$, 它的二进制展开中数字0 和1 出现的频率均为$\frac{1}{2}$.

证明. 根据例子2.1.14, 加倍映射

$$T_2 : \mathbb{T} \to \mathbb{T}, x \mapsto 2x \pmod 1$$

关于的Lebesgue 测度$m_{\mathbb{T}}$ 是遍历的. 设Y 是$(0,1)$中所有具有唯一二进制展开点的全体. 则$(0,1) \setminus Y$是一个可数集, 从而$m_{\mathbb{T}}(Y) = 1$. 对任意$x \in Y$, 令

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}.$$

则

$$T_2 x = \sum_{n=1}^{\infty} \frac{a_{n+1}}{2^n}.$$

设$f(x) = \mathbf{1}_{[\frac{1}{2},1)}(x)$. 则

$$f(T_2^i x) = \begin{cases} 1, & \text{如果} a_{i+1} = 1 \\ 0, & \text{如果} a_{i+1} = 0. \end{cases}$$

对$f(x)$应用逐点遍历定理2.3.1, 对于$m_{\mathbb{T}}$几乎所有点$x \in Y$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T_2^i x) = \int f(x) dm_{\mathbb{T}}(x) = \frac{1}{2},$$

即几乎处处得点二进制展示中1出现得频率为$\frac{1}{2}$. □

下面定理给出了一个遍历性的等价刻画.

定理 2.3.11 设$(X, \mathcal{X}, \mu, T)$为保测系统. T 为遍历的当且仅当对任意$A, B \in \mathcal{X}$ 有

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(A \cap T^{-i}B) = \mu(A)\mu(B).$$

证明. 设T为遍历的. 令$f = \mathbf{1}_B$, 运用定理2.3.1, 我们得到

$$\frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_B(T^i(x)) \xrightarrow{a.e.} \mu(B), n \to \infty.$$

两边乘上$\mathbf{1}_A$即为

$$\frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_B(T^i(x))\mathbf{1}_A \xrightarrow{a.e.} \mu(B)\mathbf{1}_A, n \to \infty.$$

由控制收敛定理得到

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}B)=\mu(A)\mu(B).$$

反之, 设$T^{-1}E=E$, $E\in\mathcal{X}$. 令$A=B=E$, 由条件即有

$$\frac{1}{n}\sum_{i=0}^{n-1}\mu(E)\to\mu(E)^2.$$

于是$\mu(E)=\mu(E)^2$. 这样就有$\mu(E)=0$或1. 所以T为遍历的. □

为方便应用上面结论, 我们可以约化到仅验证半代数, 即：

定理 2.3.12 设$(X,\mathcal{X},\mu,T)$为保测系统, $\mathcal{S}$为生成$\mathcal{X}$的半代数. 则T 为遍历的当且仅当对任意$A,B\in\mathcal{S}$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}B)=\mu(A)\mu(B).$$

证明. 仅需证明上面右边推出左边即可. 因为$\mathcal{S}$生成的代数$\mathcal{A}(\mathcal{S})$的元素由$\mathcal{S}$中有限个元素无交并组成, 所以对任意$A,B\in\mathcal{A}(\mathcal{S})$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}B)=\mu(A)\mu(B).$$

下面我们需要把上式中$\mathcal{A}(\mathcal{S})$换为$\mathcal{X}$.

对于任何$\varepsilon>0$, 任何$A,B\in\mathcal{X}$, 存在$A_0,B_0\in\mathcal{A}(\mathcal{S})$ 使得

$$\mu(A\Delta A_0)<\varepsilon,\ \text{且}\ \mu(B\Delta B_0)<\varepsilon.$$

对于任何$i\ge 0$,

$$(A\cap T^{-i}B)\Delta(A_0\cap T^{-i}B_0)\subseteq(A\Delta A_0)\cup(T^{-i}B\Delta T^{-i}B_0).$$

于是$\mu\big((A\cap T^{-i}B)\Delta(A_0\cap T^{-i}B_0)\big)<2\varepsilon$. 从而

$$|\mu(A\cap T^{-i}B)-\mu(A_0\cap T^{-i}B_0)|<2\varepsilon.$$

另外,

$$\begin{aligned}&\mu(A\cap T^{-i}B)-\mu(A)\mu(B)\\=&\big(\mu(A\cap T^{-i}B)-\mu(A_0\cap T^{-i}B_0)\big)+\big(\mu(A_0\cap T^{-i}B_0)-\mu(A_0)\mu(B_0)\big)\\&+(\mu(A_0)\mu(B_0)-\mu(A)\mu(B_0))+(\mu(A)\mu(B_0)-\mu(A)\mu(B)).\end{aligned}$$

于是易得

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}B)-\mu(A)\mu(B)\right| \\
=&\left|\frac{1}{n}\sum_{i=0}^{n-1}\big(\mu(A\cap T^{-i}B)-\mu(A_0\cap T^{-i}B_0)\big)\right|+\left|\frac{1}{n}\sum_{i=0}^{n-1}\big(\mu(A_0\cap T^{-i}B_0)-\mu(A_0)\mu(B_0)\big)\right| \\
&+\left|\frac{1}{n}\sum_{i=0}^{n-1}(\mu(A_0)\mu(B_0)-\mu(A)\mu(B_0))\right|+\left|\frac{1}{n}\sum_{i=0}^{n-1}(\mu(A)\mu(B_0)-\mu(A)\mu(B))\right| \\
<&\left|\frac{1}{n}\sum_{i=0}^{n-1}\mu(A_0\cap T^{-i}B_0)-\mu(A_0)\mu(B_0)\right|+4\varepsilon.
\end{aligned}
$$

由此容易完成整个证明. □

下面我们运用Birkhoff 定理证明独立等分布的Kolmogorov 强大数定理.

定理 2.3.13 (Kolmogorov 强大数定理) 设$(\Omega,\mathcal{F},P)$ 为概率空间，$X_1,X_2,\dots$ 为相互独立且等分布的随机变量序列. 那么

$$\frac{1}{n}\Big(X_1+X_2+\ldots+X_n\Big)\xrightarrow{a.e}\mathbb{E}(X_1),\ n\to\infty. \tag{2.3.4}$$

证明. 由于$X_1,X_2,\dots$ 为等分布的, 设ν 为其分布, 即$\nu=(X_j)_*P=P\circ X_j^{-1}$ 为$\mathbb{R}$ 上的Borel概率测度(ν 与$j\in\mathbb{N}$ 无关). 另外

$$\mathbb{E}(X_j)=\int_{\mathbb{R}}td\nu(t)$$

与j 选取无关. 设

$$(X,\mathcal{X},\mu)=(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^{\mathbb{N}}),\nu^{\mathbb{N}})$$

以及$\sigma:X\to X$ 为转移映射. 则容易验证$(X,\mathcal{X},\mu,\sigma)$ 为遍历系统.

对于$n\in\mathbb{N}$, 设$Y_n:X=\mathbb{R}^N\to\mathbb{R}$ 为到n 坐标的投射. 记$g=Y_1$, 那么我们有$Y_{j+1}=g\circ\sigma^j,\forall j\ge 0$. 根据Birkhoff 遍历定理,

$$\frac{1}{n}\Big(Y_1+Y_2+\ldots+Y_n\Big)=\frac{1}{n}\sum_{j=0}^{n-1}g\circ\sigma^j\xrightarrow{a.e}\int_X gd\mu,\ n\to\infty. \tag{2.3.5}$$

注意到$g_*\mu=\nu$, 我们有

$$\int_X gd\mu=\int_{\mathbb{R}}td\nu(t)=\mathbb{E}(X_1).$$

设

$$\phi:\Omega\to X,\quad \phi(\omega)=(X_n(\omega))_{n\in\mathbb{N}},\ \forall\omega\in\Omega.$$

因为$X_1,X_2,\dots$ 为相互独立的，所以$\phi_*P=\mu$. 于是ϕ 诱导了Koopman 算子

$$\Phi:L^0(X,\mu)\to L^0(\Omega,P),\quad f\mapsto f\circ\phi.$$

根据定义，我们有对于任何$n\in\mathbb{N}$, $\Phi Y_n=Y_n\circ\phi=X_n$. 容易验证：

$$\lim_{n\to\infty}f_n=f,\ \mu-a.e.\implies\lim_{n\to\infty}\Phi(f_n)=\Phi(f),\ P-a.e.$$

令$f_n=\frac{1}{n}\Big(Y_1+Y_2+\ldots+Y_n\Big)$. 那么根据(2.3.5), 我们得到(2.3.4). 证明完毕! $\square$

§2.3.5 Markov转移的遍历性条件

下面我们运用定理2.3.12来给出Markov转移遍历的条件.

引理 2.3.14 *设P为随机矩阵, $\vec{p}$为严格正的概率向量, 且$\vec{p}P=\vec{p}$. 则*

$$Q=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}P^n$$

存在. 并且

1. *Q 也为随机矩阵;*
2. *$QP=PQ=Q$;*
3. *P的特征值1对应的特征向量也为Q的特征向量;*
4. *$Q^2=Q$.*

证明. 设$(\Sigma_k,\mathcal{B},\mu,\sigma)$为双边$(\vec{p},P)$-Markov系统. 令

$$\chi_i=\mathbf{1}_{{}_0[i]_0}=\{(x_j)_{j\in\mathbb{Z}}:x_0=i\},i\in\{0,1,\ldots,k-1\}.$$

根据Birkhoff遍历定理, 对于$j\in\{0,1,\ldots,k-1\}$

$$\frac{1}{N}\sum_{n=0}^{N-1}\chi_j(\sigma^nx)\xrightarrow{a.e.}\chi_j^*(x),N\to\infty.$$

两边同时乘以χ_i, $i\in\{0,1,\ldots,k-1\}$, 再由控制收敛定理, 得到

$$\frac{1}{N}\sum_{n=0}^{N-1}\int_{\Sigma_k}\chi_j(\sigma^nx)\chi_i(x)d\mu(x)\longrightarrow\int_{\Sigma_k}\chi_j^*(x)\chi_i(x)d\mu(x),N\to\infty.$$

注意到

$$\begin{aligned}\int_{\Sigma_k}\chi_j(\sigma^nx)\chi_i(x)d\mu(x)&=\mu(\sigma^{-n}{}_0[j]_0\cap{}_0[i]_0)\\&=\mu\left(\{(x_k)_{k\in\mathbb{Z}}\in\Sigma_k:x_0=i,x_n=j\}\right)\\&=\sum_{0\le a_1,\ldots,a_{n-1}\le k-1}p_ip_{ia_1}p_{a_1a_2}\ldots p_{a_{n-2}a_{n-1}}p_{a_{n-1}j}\\&=p_ip_{ij}^{(n)}.\end{aligned}$$

所以对$i,j\in\{0,1,\ldots,k-1\}$

$$\frac{1}{N}\sum_{n=0}^{N-1}p_ip_{ij}^{(n)}\longrightarrow\int_{\Sigma_k}\chi_j^*(x)\chi_i(x)d\mu(x),N\to\infty.$$

令$q_{ij}=\dfrac{1}{p_i}\displaystyle\int_{\Sigma_k}\chi_j^*(x)\chi_i(x)d\mu(x),Q=[q_{ij}]\in\mathbb{R}^{k\times k}$. 则我们得到

$$\frac{1}{N}\sum_{n=0}^{N-1}p_{ij}^{(n)}\to q_{ij},N\to\infty,$$

即

$$Q=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}P^n.$$

其余性质是容易验证的. □

定理 2.3.15 *设$(\Sigma_k,\mathcal{B},\mu,\sigma)$ 为（双边或者单边）$(\vec{p},P)$-Markov 转移. 设$\vec{p}$ 为严格正概率向量, Q 同引理2.3.14. 则以下等价*

1. *σ 是遍历的.*
2. *Q 任何两行是一样的.*
3. *$Q>0$, 即Q 每个值都是正的.*
4. *P 是不可约的.*
5. *1 为P 的简单特征值.*

证明. 所有符号同引理2.3.14.

(1) $\Rightarrow$ (2). 如果σ为遍历的, 那么根据定理2.3.12, 我们得到对于$i,j\in\{0,1,\ldots,k-1\}$

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(\sigma^{-n}{}_0[j]_0\cap{}_0[i]_0)\to\mu({}_0[i]_0)\mu({}_0[j]_0),\ \mathbb{N}\to\infty.$$

注意到

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(\sigma^{-n}{}_0[j]_0\cap{}_0[i]_0)=p_iq_{ij},$$

以及

$$\mu({}_0[i]_0)\mu({}_0[j]_0)=p_ip_j,$$

我们得到$p_iq_{ij}=p_ip_j$. 所以

$$q_{ij}=p_j,\forall i,j\in\{0,1,\ldots,k-1\}.$$

即$Q = \begin{pmatrix} p_0 & p_1 & \cdots & p_{k-1} \\ p_0 & p_1 & \cdots & p_{k-1} \\ \vdots & \vdots & \vdots & \vdots \\ p_0 & p_1 & \cdots & p_{k-1} \end{pmatrix}$任何两行是一样的.

(2) $\Rightarrow$ (3). 如果Q每行是一样的, 那么$\vec{p}Q = \vec{p}$推出$q_{ij} = p_j$, 即

$$Q = \begin{pmatrix} p_0 & p_1 & \cdots & p_{k-1} \\ p_0 & p_1 & \cdots & p_{k-1} \\ \vdots & \vdots & \vdots & \vdots \\ p_0 & p_1 & \cdots & p_{k-1} \end{pmatrix}.$$

尤其$Q > 0$.

(3) $\Rightarrow$ (4). 因为$Q > 0$以及对于$i, j \in \{0, 1, \ldots, k-1\}$

$$\frac{1}{N} \sum_{n=0}^{N-1} p_{ij}^{(n)} \to q_{ij}, N \to \infty,$$

对于任何取定$i, j \in \{0, 1, \ldots, k-1\}$, n充分大时就有$p_{ij}^{(n)} > 0$, 即P为不可约的.

(4) $\Rightarrow$ (3). 因为$Q = QP$, 所以$Q = QP^n, \forall n \in \mathbb{N}$. 于是

$$q_{ij} = \sum_{s=0}^{k-1} q_{is} p_{sj}^{(n)} \geq q_{it} p_{tj}^{(n)}, \forall t \in \{0, 1, \ldots, k-1\}, \forall n \in \mathbb{N}.$$

对$i \in \{0, 1, \ldots, k-1\}$, 令

$$E_i = \{j \in \{0, 1, \ldots, k-1\} : q_{ij} > 0\}.$$

因为$\sum_j q_{ij} = 1$, 所以$E_i \neq \emptyset$. 设$t \in E_i$. 对于任何$j \in \{0, 1, \ldots, k-1\}$, 因为$P$为不可约的, 存在$n$ 使得$p_{tj}^{(n)} > 0$. 于是

$$q_{ij} \geq q_{it} p_{tj}^{(n)} > 0.$$

即$j \in E_i$, $E_i = \{0, 1, \ldots, k-1\}$. 所以$q_{ij} > 0, \forall i, j \in \{0, 1, \ldots, k-1\}$.

(3) $\Rightarrow$ (2). 对于固定的$j \in \{0, 1, \ldots, k-1\}$, 令

$$q_j = \max_{0 \leq i \leq k-1} q_{ij}.$$

因为$Q^2 = Q$, 如果存在$i \in \{0, 1, \ldots, k-1\}$使得$q_{ij} < q_j$, 那么对于任何$t \in \{0, 1, \ldots, k-1\}$

$$q_{tj} = \sum_{h=0}^{k-1} q_{th} q_{hj} < \sum_{h=0}^{k-1} q_{th} q_j = q_j \sum_{h=0}^{k-1} q_{th} = q_j.$$

此与q_j定义矛盾! 所以任何$i \in \{0, 1, \ldots, k-1\}$都有$q_{ij} = q_j$, 即$Q$每一行都是一样的. 因为$\vec{p}Q = \vec{p}$, 我们有

$$Q = \begin{pmatrix} p_0 & p_1 & \cdots & p_{k-1} \\ p_0 & p_1 & \cdots & p_{k-1} \\ \vdots & \vdots & \vdots & \vdots \\ p_0 & p_1 & \cdots & p_{k-1} \end{pmatrix}.$$

(2) ⇒ (1). 根据定理2.3.12, 我们仅需对于

$$A = {}_a[i_0, \ldots, i_r]_{a+r}, B = {}_b[j_0, \ldots, j_s]_{b+s}, a, b \in \mathbb{Z}, r, s \in \mathbb{Z}_+,$$

证明

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap \sigma^{-n} B) = \mu(A)\mu(B).$$

对于$n > a + r - b$ 且$n > 0$, 我们有

$$\mu(A \cap \sigma^{-n} B) = p_{i_0} p_{i_0 i_1} \ldots p_{i_{r-1} i_r} p_{i_r j_0}^{(n-(a+r-b))} p_{j_0 j_1} p_{j_1 j_2} \ldots p_{j_{s-1} j_s}.$$

根据(2), $q_{ij} = p_j, \forall i, j \in \{0, 1, \ldots, k-1\}$. 所以

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap \sigma^{-n} B) \\
&= \lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} p_{i_0} p_{i_0 i_1} \ldots p_{i_{r-1} i_r} p_{i_r j_0}^{(n-(a+r-b))} p_{j_0 j_1} p_{j_1 j_2} \ldots p_{j_{s-1} j_s} \\
&= p_{i_0} p_{i_0 i_1} \ldots p_{i_{r-1} i_r} \lim_{N\to\infty} \left(\frac{1}{N} \sum_{n=0}^{N-1} p_{i_r j_0}^{(n-(a+r-b))} \right) p_{j_0 j_1} p_{j_1 j_2} \ldots p_{j_{s-1} j_s} \\
&= p_{i_0} p_{i_0 i_1} \ldots p_{i_{r-1} i_r} q_{i_r j_0} p_{j_0 j_1} p_{j_1 j_2} \ldots p_{j_{s-1} j_s} \\
&= p_{i_0} p_{i_0 i_1} \ldots p_{i_{r-1} i_r} p_{j_0} p_{j_0 j_1} p_{j_1 j_2} \ldots p_{j_{s-1} j_s} \\
&= \mu({}_a[i_0, \ldots, i_r]_{a+r}) \mu({}_b[j_0, \ldots, j_s]_{b+s}) \\
&= \mu(A)\mu(B).
\end{aligned}$$

(2) ⇒ (5). 因为$Q = \begin{pmatrix} p_0 & p_1 & \cdots & p_{k-1} \\ p_0 & p_1 & \cdots & p_{k-1} \\ \vdots & \vdots & \vdots & \vdots \\ p_0 & p_1 & \cdots & p_{k-1} \end{pmatrix}$, 所以对应于特征值1 的$Q$ 的唯一左特征向量为$\vec{p}$ 的数乘, 由引理2.3.14, 它们也是P 对应于特征值1 的所有左特征向量.

(5) ⇒ (2). 设1 为P 简单特征值, 那么由于$Q = QP$, Q 的每行对应的向量是P 特征向量, 所以它们是一样的. □

习 题

1. 设$(X,\mathcal{X},\mu,T)$ 是一个可逆保测系统. 证明对任意的$f\in L^1(X,\mu)$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}f(T^{-i}x)=\lim_{n\to\infty}\frac{1}{2n+1}\sum_{i=-n}^{n}f(T^ix)$$

几乎处处成立.

2. 给出Borel正则数定理对于十进制数的版本.

3. 设$(X,\mathcal{X},\mu,T)$ 是一个保测系统. 证明对任意的$f\in L^1(X,\mu)$,

$$\lim_{n\to\infty}\frac{f(T^nx)}{n}=0$$

几乎处处成立.

4. 设$(X,\mathcal{X},\mu,T)$是一个遍历动力系统. 设$A\in\mathcal{X}$且$\mu(A)>0$. 证明A的第一返回时间函数

$$\begin{aligned}n_A: A&\to\mathbb{N}\cup\{\infty\}\\ x&\mapsto\min\{n>0:T^nx\in A\},\end{aligned}$$

是几乎处处有限的并且满足

$$\int_A n_A(x)d\mu_A(x)=\frac{1}{\mu(A)}.$$

§2.4 混合性

我们已经看到遍历系统为“不可分割”的, 并且对任意正测集其轨道将覆盖几乎整个空间. 本节将介绍一些比遍历性更强的性质, 具有这些性质的系统将有更为复杂的动力学性状.

§2.4.1 混合性的定义

首先根据定理2.3.11, 保测系统$(X,\mathcal{X},\mu,T)$为遍历的当且仅当对任意$A,B\in\mathcal{X}$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}(\mu(A\cap T^{-i}B)-\mu(A)\mu(B))=0.$$

定义 2.4.1 设$(X,\mathcal{X},\mu,T)$为一个保测系统.

1 如果对任意$A,B\in\mathcal{X}$有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}|\mu(A\cap T^{-i}B)-\mu(A)\mu(B)|=0,$$

那么称T为**(测度)弱混合的**;

2 如果对任意$A, B \in \mathcal{X}$有
$$\lim_{n\to\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B),$$
那么称T为**(测度)强混合的**.

3 设$k \in \mathbb{N}$, 如果对于任何$A_0, A_1, \ldots, A_k \in \mathcal{X}$, 有
$$\lim_{n_1,\ldots,n_k\to\infty} \mu(A_0 \cap T^{-n_1}A_1 \cap T^{-(n_1+n_2)}A_2 \cap \ldots \cap T^{-(n_1+n_2+\ldots+n_k)}A_k) = \mu(A_0)\mu(A_1)\ldots\mu(A_k),$$
那么称T为**k-阶混合的**.

4 设$\mathcal{F}(B_1, \ldots, B_k; n)$ 为由$\{T^{-j}B_i : j \geq n, i = 1, 2, \ldots, k\}$ 生成的σ-代数, 如果对于任何A ,B_1, …, $B_k \in \mathcal{X}$, 有
$$\lim_{n\to\infty} \sup_{C\in\mathcal{F}(B_1,\ldots,B_k;n)} |\mu(A \cap C) - \mu(A)\mu(C)| = 0,$$
那么称T为**一致混合的**.

注记 2.4.2　*1.* 遍历性、弱混合、混合的比较: 对于任何数列$\{a_n\}_{n\in\mathbb{N}}$, 我们有
$$\lim_{n\to\infty} a_n = 0 \ \Rightarrow\ \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} |a_n| = 0 \ \Rightarrow\ \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} a_n = 0.$$
于是根据定义我们得出
$$\text{强混合} \ \Rightarrow\ \text{弱混合} \ \Rightarrow\ \text{遍历性}.$$

2. 后面我们会证明圆周无理旋转是遍历但不是弱混合的.

3. 设$(X, \mathcal{X}, \mu)$为概率空间, 记$\mathrm{Aut}(X, \mu)$为$(X, \mathcal{X}, \mu)$ 上全体可逆保测映射, 赋予$\mathrm{Aut}(X, \mu)$**弱拓扑**: $T_n, T \in \mathrm{Aut}(X, \mu)$,
$$T_n \to T, n \to \infty \Leftrightarrow \mu(T_nA\Delta TA) \to 0, \forall A \in \mathcal{X}.$$
Halmos (1944) 证明了在赋予弱拓扑的$\mathrm{Aut}(X, \mu)$ 中, 弱混合系统的全体包含了稠密G_δ集; *Rohlin (1948)* 证明了强混合系统全体是第一纲集[92]. 于是一定存在弱混合但是不是强混合的例子.

4. 根据定义, 强混合等价于1阶混合. 如果$n < m$, 那么m阶混合蕴含n阶混合. 我们不知道反过来是否成立, 参见下面的*Rohlin*问题.

5. 容易看出一致混合蕴含任意阶的混合. 在定义中取$A = A_0, B_1 = A_1$, 那么就得到强混合, 即*1*-阶混合. 下面证明*2*-阶混合. 设$A_0, A_1, A_2 \in \mathcal{X}$, 由*1*- 阶混合, 存在$n_2^{(0)}$ 使得$n_2 \geq n_2^{(0)}$ 时
$$|\mu(A_1 \cap T^{-n_2}A_2) - \mu(A_1)\mu(A_2)| < \varepsilon.$$

在一致混合定义中取$A=A_0, B_1=A_1, B_2=A_2$, 取$n_1^{(0)}$ 使得$n_1\ge n_1^{(0)}$ 时有

$$\left|\mu\big(A_0\cap T^{-n_1}(A_1\cap T^{-n_2}A_2)\big)-\mu(A_0)\mu\big(A_1\cap T^{-n_2}A_2\big)\right|<\varepsilon.$$

于是

$$\left|\mu\big(A_0\cap T^{-n_1}A_1\cap T^{-n_1-n_2}A_2\big)-\mu(A_0)\mu(A_1)\mu(A_2)\right|<2\varepsilon.$$

归纳地, 可以证明一致混合混合蕴含任意阶的混合.

但是反之不对. 因为可以证明T为一致混合等价于T为K-系统(定理7.11.8). 可以找到非K- 系统的强混合系统, 例如Horocycle 流.

问题 2.4.3 (Rohlin, 1949) 设$(X,\mathcal{X},\mu,T)$为一个保测系统, 强混合是否蕴含2阶混合? 强混合是否蕴含了任何阶的混合?

注记 2.4.4 这个问题的高维群作用情况是不正确的, *Ledrappier* 指出在$\mathbb{Z}^2$ 的作用下, 存在1阶混合但不是2 混合的例子[141]. *Rohlin* 问题的相对化问题也是不正确的, 最近*de La Rue* 指出了此点[176].

§2.4.2 混合性的刻画

类似于定理2.3.12, 下面的定理提供了一种验证一个保测系统为弱混合、强混合的方法.

定理 2.4.5 设$(X,\mathcal{X},\mu,T)$为保测系统, $\mathcal{S}$ 为生成$\mathcal{X}$ 的半代数. 则

1. T为弱混合的当且仅当对任意$A,B\in\mathcal{S}$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}|\mu(A\cap T^{-i}B)-\mu(A)\mu(B)|=0.$$

2. T为强混合的当且仅当对任意$A,B\in\mathcal{S}$

$$\lim_{n\to\infty}\mu(A\cap T^{-n}B)=\mu(A)\mu(B).$$

证明. 完全类似于定理2.3.12的证明, 请读者自己完成. □

在刻画弱混合前, 我们回顾一些概念. 设S为$\mathbb{Z}_+$的子集, 如果S的上密度和下密度

$$\overline{d}(S)=\limsup_{n\to\infty}\frac{1}{n}|S\cap\{0,1,\dots,n-1\}|$$

及

$$\underline{d}(S)=\liminf_{n\to\infty}\frac{1}{n}|S\cap\{0,1,\dots,n-1\}|.$$

相等且等于a, 那么我们称S具有**密度**a, 记为$d(S)=a$. 记$\mathcal{F}_{\rm d1}$ 为$\mathbb{Z}_+$中密度为1的子集全体组成的集合.

引理 2.4.6 (Koopman-von Neumann, 1932) 设$\{a_n\}_{n=0}^{\infty}$为有界序列,则以下等价:

1. $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{i=0}^{n-1}|a_i|=0$.
2. 存在$\mathbb{Z}_+$的零密度集J使得$\lim\limits_{J\not\ni n\to\infty}a_n=0$.
3. $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{i=0}^{n-1}|a_i|^2=0$.

证明. 对于$m<n$, 我们记$[m,n]=\{m,m+1,\ldots,n\}$ 为整数区间.

(1) $\Rightarrow$ (2). 设

$$J_k=\{n\in\mathbb{Z}_+:|a_n|\geq\frac{1}{k}\},\ k\in\mathbb{N}.$$

则

$$J_1\subseteq J_2\subseteq J_3\subseteq\ldots.$$

因为

$$\frac{1}{n}\sum_{i=0}^{n-1}|a_i|\geq\frac{1}{n}\frac{1}{k}\left|J_k\cap[0,n-1]\right|,$$

所以

$$\lim_{n\to\infty}\frac{|J_k\cap[0,n-1]|}{n}=0,$$

即$d(J_k)=0,\forall k\in\mathbb{N}$. 于是存在整数列$0=l_0<l_1<l_2<\ldots$ 使得$n\geq l_k$ 时有

$$\frac{|J_{k+1}\cap[0,n-1]|}{n}<\frac{1}{k+1}.$$

令

$$J=\bigcup_{k=0}^{\infty}J_{k+1}\cap[l_k,l_{k+1}-1].$$

我们下面证明$d(J)=0$.

因为$J_1\subseteq J_2\subseteq J_3\subseteq\ldots$, 如果$l_k\leq n\leq l_{k+1}-1$, 则

$$J\cap[0,n-1]=\big(J\cap[0,l_k-1]\big)\cup\big(J\cap[l_k,n-1]\big)\subseteq\big(J_k\cap[0,l_k-1]\big)\cup\big(J_{k+1}\cap[0,n-1]\big).$$

于是

$$\begin{aligned}\frac{|J\cap[0,n-1]|}{n}&\leq\frac{|J_k\cap[0,l_k-1]|+|J_{k+1}\cap[0,n-1]|}{n}\\&\leq\frac{|J_k\cap[0,n-1]|+|J_{k+1}\cap[0,n-1]|}{n}\\&<\frac{1}{k}+\frac{1}{k+1}.\end{aligned}$$

于是
$$\frac{|J\cap[0,n-1]|}{n}\to 0, n\to\infty,$$
即$d(J)=0$.

如果$n>l_k$ 并且$n\notin J$, 则$n\notin J_{k+1}$, 于是$|a_n|<\frac{1}{k+1}$. 由此得到
$$\lim_{J\not\ni n\to\infty}|a_n|=0.$$

(2) $\Rightarrow$ (1). 因为$\{a_n\}_{n\in\mathbb{Z}_+}$ 有界, 存在$M>0$ 使得$|a_n|\le M,\forall n\in\mathbb{Z}_+$. 根据条件, 对于任何$\varepsilon>0$, 存在$N_\varepsilon\in\mathbb{N}$ 使得
$$|a_n|<\varepsilon,\ \forall n>N_\varepsilon \text{ 且 } n\notin J.$$
再取$N\ge N_\varepsilon$ 使得$n\ge N$时有
$$\frac{|J\cap[0,n-1]|}{n}<\varepsilon.$$
于是$n\ge\frac{1}{\varepsilon}N$ 时, 我们有
$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}|a_i| &= \frac{1}{n}\left[\sum_{i=0}^{N-1}|a_i|+\sum_{i\in J\cap[N,n-1]}|a_i|+\sum_{i\in[N,n-1]\setminus J}|a_i|\right]\\ &< \frac{1}{n}\left(MN+M|J\cap[0,n-1]|+n\varepsilon\right)\\ &< M\varepsilon+M\varepsilon+\varepsilon=(2M+1)\varepsilon.\end{aligned}$$

最后, 因为对于零密度集合J,
$$\lim_{J\not\ni n\to\infty}|a_n|=0$$
当且仅当
$$\lim_{J\not\ni n\to\infty}|a_n|^2=0.$$
我们得到(3)与(1), (2)等价. □

因为零密度集的补集为密度为1的集合，所以引理2.4.6-2 等价于：存在$\mathbb{Z}_+$ 密度为1 的子集I 使得$\lim_{I\ni n\to\infty}a_n=0$. 根据引理2.4.6, 我们马上得到下面定理：

定理 2.4.7 设 $(X,\mathcal{X},\mu,T)$ 为保测系统.则下列命题等价:

1. T为弱混合的.

2. 对任意$A,B\in\mathcal{X}$, 存在$J\in\mathcal{F}_{\mathrm{d1}}$ 使得
$$\lim_{J\ni n\to\infty}\mu(A\cap T^{-n}B)=\mu(A)\mu(B).$$

3. 对任意$A, B \in \mathcal{X}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(A \cap T^{-i}B) - \mu(A)\mu(B)|^2 = 0.$$

我们把下面定理证明留作习题:

定理 2.4.8 设$(X, \mathcal{X}, \mu, T)$为保测系统. 那么我们有:

1. 以下等价
   (a) T遍历.
   (b) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \langle U_T^i f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle, \forall f, g \in L^2(X, \mu)$.
   (c) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \langle U_T^i f, f\rangle = \langle f, 1\rangle\langle 1, f\rangle, \forall f \in L^2(X, \mu)$.
2. 以下等价
   (a) T弱混合.
   (b) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left|\langle U_T^i f, g\rangle - \langle f, 1\rangle\langle 1, g\rangle\right| = 0, \forall f, g \in L^2(X, \mu)$.
   (c) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left|\langle U_T^i f, f\rangle - \langle f, 1\rangle\langle 1, f\rangle\right| = 0, \forall f \in L^2(X, \mu)$.
   (d) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left|\langle U_T^i f, f\rangle - \langle f, 1\rangle\langle 1, f\rangle\right|^2 = 0, \forall f \in L^2(X, \mu)$.
3. 以下等价
   (a) T强混合.
   (b) $\lim_{n\to\infty} \langle U_T^n f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle, \forall f, g \in L^2(X, \mu)$.
   (c) $\lim_{n\to\infty} \langle U_T^n f, f\rangle = \langle f, 1\rangle\langle 1, f\rangle, \forall f \in L^2(X, \mu)$.

定义 2.4.9 两个保测系统$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 称为**弱不交的**是指它们的乘积系统$(X \times Y, \mathcal{X} \times \mathcal{Y}, \mu \times \nu, T \times S)$ 为遍历的.

现在我们有

定理 2.4.10 设$(X, \mathcal{X}, \mu, T)$为保测系统. 则下列命题等价:

1. $(X, \mathcal{X}, \mu, T)$ 为弱混合的.
2. $(X \times X, \mathcal{X} \times \mathcal{X}, \mu \times \mu, T \times T)$ 为遍历的.

3. $(X\times X,\mathcal{X}\times\mathcal{X},\mu\times\mu,T\times T)$ 为弱混合的.

4. T 弱不交于任意遍历系统.

证明. (1) $\Rightarrow$ (3) 对正测集$A,B,C,D\in\mathcal{X}$,存在密度为1的集合$J_1,J_2\subseteq\mathbb{Z}_+$使得

$$\begin{aligned}\lim_{J_1\ni n\to\infty}\mu(A\cap T^{-n}B)&=\mu(A)\mu(B)>0\\ \lim_{J_2\ni n\to\infty}\mu(C\cap T^{-n}D)&=\mu(C)\mu(D)>0.\end{aligned}$$

于是

$$\begin{aligned}&\lim_{J_1\cap J_2\ni n\to\infty}(\mu\times\mu)\big((A\times C)\cap(T\times T)^{-n}(B\times D)\big)\\ =&\lim_{J_1\cap J_2\ni n\to\infty}\mu(A\cap T^{-n}B)\mu(C\cap T^{-n}D)\\ =&\mu(A)\mu(B)\mu(C)\mu(D)\\ =&(\mu\times\mu)(A\times C)(\mu\times\mu)(B\times D).\end{aligned}$$

由引理2.4.6, 我们有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|(\mu\times\mu)\big((A\times C)\cap(T\times T)^{-n}(B\times D)\big)-(\mu\times\mu)(A\times C)(\mu\times\mu)(B\times D)\right|=0.$$

因为可测方体组成$\mathcal{X}\times\mathcal{X}$的半代数, 由定理2.4.5, 这意味着$T\times T$ 为弱混合的.

(3) $\Rightarrow$ (2) 是显然的.

(2) $\Rightarrow$ (1) 设$T\times T$为遍历的, 我们要证明对任意$A,B\in\mathcal{X}$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}|\mu(A\cap T^{-i}B)-\mu(A)\mu(B)|^2=0$$

成立.

根据$T\times T$遍历, 我们有

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap T^{-i}B)&=\frac{1}{n}\sum_{i=0}^{n-1}(\mu\times\mu)\big((A\times X)\cap(T\times T)^{-i}(B\times X)\big)\\ &\to(\mu\times\mu)(A\times X)(\mu\times\mu)(B\times X)\\ &=\mu(A)\mu(B),\ n\to\infty.\end{aligned}$$

并且

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}\big(\mu(A\cap T^{-i}B)\big)^2&=\frac{1}{n}\sum_{i=0}^{n-1}(\mu\times\mu)\big((A\times A)\cap(T\times T)^{-i}(B\times B)\big)\\ &\to(\mu\times\mu)(A\times A)(\mu\times\mu)(B\times B)\\ &=\mu(A)^2\mu(B)^2,\ n\to\infty.\end{aligned}$$

这样

$$\begin{aligned}
& \frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(A\cap T^{-i}B)-\mu(A)\mu(B)\right|^2 \\
= \ & \frac{1}{n}\sum_{i=0}^{n-1}\left(\mu(A\cap T^{-i}B)^2-2\mu(A\cap T^{-i}B)\mu(A)\mu(B)+\mu(A)^2\mu(B)^2\right) \\
\to \ & 2\mu(A)^2\mu(B)^2-2\mu(A)^2\mu(B)^2=0, n\to\infty.
\end{aligned}$$

于是由定理2.4.7, T为弱混合的.

(4) $\Rightarrow$ (2) 是显然的, 我们证明(1) $\Rightarrow$ (4). 设$(Y,\mathcal{Y},\nu,S)$ 为遍历的, 我们要证明$(X\times Y,\mathcal{X}\times\mathcal{Y},\mu\times\nu,T\times S)$ 为遍历的. 为此我们证明, 对于任何$A_1,B_1\in\mathcal{X}, A_2,B_2\in\mathcal{Y}$, 成立

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu\times\nu\left((A_1\times A_2)\cap(T\times S)^{-i}(B_1\times B_2)\right)=\mu\times\nu(A_1\times A_2)\mu\times\nu(B_1\times B_2). \quad (2.4.1)$$

$$\begin{aligned}
& \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu\times\nu\left((A_1\times A_2)\cap(T\times S)^{-i}(B_1\times B_2)\right) \\
= & \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A_1\cap T^{-i}B_1)\nu(A_2\cap S^{-i}B_2) \\
= & \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A_1)\mu(B_1)\nu(A_2\cap S^{-i}B_2)+\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left[\mu(A_1\cap T^{-i}B_1)-\mu(A_1)\mu(B_1)\right]\nu(A_2\cap S^{-i}B_2).
\end{aligned}$$

因为S 遍历,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A_1)\mu(B_1)\nu(A_2\cap S^{-i}B_2)=\mu(A_1)\mu(B_1)\nu(A_2)\nu(B_2).$$

因为T 弱混合,

$$\begin{aligned}
& \lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=0}^{n-1}\left[\mu(A_1\cap T^{-i}B_1)-\mu(A_1)\mu(B_1)\right]\nu(A_2\cap S^{-i}B_2)\right| \\
\le & \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(A_1\cap T^{-i}B_1)-\mu(A_1)\mu(B_1)\right|=0.
\end{aligned}$$

综合上面两个式子, 我们得到(2.4.1). □

下面我们讨论弱混合系统谱方面的性质.

定义 2.4.11 *我们称保测系统$(X,\mathcal{X},\mu,T)$ 具有***连续谱***是指1为U_T唯一的特征值, 而它唯一的特征函数是常值函数.*

注记 2.4.12 *1. 因为对于任何常值函数c, 我们有$U_Tc=c$. 所以对于U_T, $\lambda=1$总是它的特征值, 常值函数总是它的特征函数.*

2. 易见, T为遍历的当且仅当1为U_T简单特征值.

3. 保测系统$(X,\mathcal{X},\mu,T)$ 具有连续谱当且仅当1为U_T唯一特征值并且T为遍历的.

对弱混合系统,我们有

定理 2.4.13 如果$(X,\mathcal{X},\mu,T)$为可逆保测系统,则T为弱混合的当且仅当T有连续谱.

这个定理的证明我们会在§2.5中给出.

下面介绍一些例子

命题 2.4.14 紧致度量交换群上旋转不是弱混合的, 尤其圆周旋转不是弱混合的.

证明. 设$T_a: G\to G, x\mapsto ax$ 为紧致度量交换群G上旋转. 令$\gamma\in\widehat{G}$. 则

$$U_{T_a}\gamma(x)=\gamma(ax)=\gamma(a)\gamma(x).$$

所以γ为T_a 以$\gamma(a)$为特征值的特征函数. 根据定理2.4.13, U_{T_a} 不是连续谱的, 进而非弱混合的.

另一个证明方法如下：取$\gamma\in\widehat{G}\setminus\{1\}$. 则

$$U_{T_a\times T_a}\gamma(x)\overline{\gamma(y)}=\gamma(ax)\overline{\gamma(ay)}=\gamma(x)\overline{\gamma(y)}.$$

即$\gamma(x)\overline{\gamma(y)}$ 为非平凡的$T_a\times T_a$ 不变函数, 所以$T_a\times T_a$ 不遍历, 即非弱混合的. □

定理 2.4.15 对于紧致交换度量群上的连续满自同态, 遍历、弱混合和强混合等价.

证明. 我们仅需要证明, 如果紧致交换度量群G上的满自同态$A: G\to G$ 为遍历的, 那么它为强混合的. 注意到$\widehat{G}$中元素组成了$L^2(G,m)$ 的标准正交基, 对于$\gamma,\delta\in\widehat{G}$, 如果不是$\gamma=\delta=1$, 那么对于充分大$n$, 有$\langle U_A^n\gamma,\delta\rangle=0$. 于是我们总是有

$$\lim_{n\to\infty}\langle U_A^n\gamma,\delta\rangle=\langle\gamma,1\rangle\langle 1,\delta\rangle.$$

取定$\delta\in\widehat{G}$, 令

$$H_\delta=\{f\in L^2(G,m):\lim_{n\to\infty}\langle U_A^n f,\delta\rangle=\langle f,1\rangle\langle 1,\delta\rangle\}.$$

则H_δ 为$L^2(G,m)$ 的子空间. 因为$\widehat{G}\subseteq H_\delta$, 如果$H_\delta$ 为闭的, 那么$H_\delta=L^2(G,m)$.

下面证明H_δ 为闭的. 如果$\delta=1$, 那么易见$H_\delta=L^2(G,m)$, 所以自然为闭的. 下设$\delta\neq 1$. 此时$\langle 1,\delta\rangle=0$, 于是$H_\delta=\{f\in L^2(G,m):\lim\limits_{n\to\infty}\langle U_A^n f,\delta\rangle=0\}$. 设$\{f_k\}_k\subseteq H_\delta$, $f_k\to f\in L^2(G,m), k\to\infty$. 于是

$$\begin{aligned}|\langle U_A^n f,\delta\rangle| &\le |\langle U_A^n f,\delta\rangle-\langle U_A^n f_k,\delta\rangle|+|\langle U_A^n f_k,\delta\rangle|\\ &\le \|f-f_k\|_2\|\delta\|_2+|\langle U_A^n f_k,\delta\rangle|\\ &=\|f-f_k\|_2+|\langle U_A^n f_k,\delta\rangle|\end{aligned}$$

由此易得$\lim\limits_{n\to\infty}\langle U_A^n f,\delta\rangle=0$, 从而$f\in H_\delta$.

现在取定$f\in L^2(G,m)$, 令

$$F_f=\{g\in L^2(G,m):\lim_{n\to\infty}\langle U_A^n f,g\rangle=\langle f,1\rangle\langle 1,g\rangle\}.$$

易见F_f为$L^2(G,m)$闭子空间. 由上结论, $\widehat{G}\subseteq F_f$, 所以$F_f=L^2(G,m),\forall f\in L^2(G,m)$.

综上, 我们得到对于任何$f,g\in L^2(G,m)$,

$$\lim_{n\to\infty}\langle U_A^n f,g\rangle=\langle f,1\rangle\langle 1,g\rangle.$$

所以A为强混合的. □

定理 2.4.16 双边（单边）$\vec{p}$-转移为强混合的.

证明. 根据定理2.4.5易证. □

定理 2.4.17 设$(\sigma_k,\mathcal{B},\mu,\sigma)$为$(\vec{p},P)$-*Markov*转移, 则以下等价

1. T弱混合.
2. T强混合.
3. P为不可约且非周期的.
4. $\lim_{n\to\infty}p_{ij}^{(n)}=p_j,\forall i,j$.

证明. 根据定理0.5.4, (3) 与(4) 等价.

(1) ⇒ (3). 因为

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|\mu({}_0[i]_0\cap T^{-n}{}_0[j]_0)-\mu({}_0[i]_0)\mu({}_0[j]_0)\right|\to 0,N\to\infty,$$

我们得到

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|p_{ij}^{(n)}-p_j\right|\to 0,N\to\infty.$$

根据引理2.4.6, 存在零密度集$J\subseteq\mathbb{Z}_+$ 使得

$$\lim_{J\not\ni\to\infty}p_{ij}^{(n)}=p_j,\forall i,j.$$

于是P 为不可约且非周期的.

(4) ⇒ (2). 根据定理2.4.5, 我们仅需对于$A={}_a[i_0,\dots,i_r]_{a+r},B={}_b[j_0,\dots,j_s]_{b+s}$ 验证$\lim_{n\to\infty}\mu(A\cap T^{-n}B)=\mu(A)\mu(B)$. 这与定理2.3.15 类似, 从略. □

§2.4.3 mild混合

定义 2.4.18 设X为拓扑空间, $\mathcal{F}$为$\mathbb{Z}_+$的子集族. 序列$\{x_n\}_n \subseteq X$ **依$\mathcal{F}$ 收敛于**$x \in X$ 是指对x的任意邻域U有

$$\{i : x_i \in U\} \in \mathcal{F}$$

成立.

如果序列$\{x_n\}$ 依$\mathcal{F}$ 收敛于$x \in X$, 那么记之为

$$\mathcal{F} - \lim x_n = x.$$

设$\mathcal{F}_{cf}$ 为全体有限余集合的全体, 即

$$\mathcal{F}_{cf} = \{A \subseteq \mathbb{Z}_+ : |\mathbb{Z}_+ \setminus A| < \infty\}.$$

那么容易得到

$$\lim_{n\to\infty} x_n = x \text{ 当且仅当 } \mathcal{F}_{cf} - \lim x_n = x.$$

于是族收敛是一般收敛的推广.

定理 2.4.19 设$(X, \mathcal{X}, \mu, T)$为保测系统. 那么

1. T 为弱混合的当且仅当

$$\mathcal{F}_{d1} - \lim \mu(A \cap T^{-n}B) = \mu(A)\mu(B), \forall A, B \in \mathcal{X}.$$

2. T 为强混合的当且仅当

$$\mathcal{F}_{cf} - \lim \mu(A \cap T^{-n}B) = \mu(A)\mu(B), \forall A, B \in \mathcal{X}.$$

最后我们讨论mild混合. 我们需要如下概念:

定义 2.4.20 $A \subseteq \mathbb{Z}_+$为**IP集**是指存在正整数序列$p_1, p_2, \ldots$使得

$$A = \{p_{i_1} + \ldots + p_{i_k} : i_1 < \ldots < i_k,\ k \in \mathbb{N}\}.$$

此时称A 为由$p_1, p_2, \ldots$ 生成的, 记为$\mathrm{FS}(\{p_i\})$. 记集合$\{A \subseteq \mathbb{Z}_+ : A \text{ 包含一个}IP\text{集}\}$ 为$\mathcal{F}_{\mathrm{ip}}$.

$A \subseteq \mathbb{Z}_+$ 称为IP^***集**是指A 与任何IP 集相交非空, 全体IP^*集的集合记为$\mathcal{F}_{ip}^*$.

定义 2.4.21 设$(X, \mathcal{X}, \mu, T)$为保测系统, 称T 为**mild 混合的**是指对任意$A,\ B \in \mathcal{X}$有

$$\mathcal{F}_{ip}{}^* - \lim \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

上式经常也记为

$$\mathrm{IP}^* - \lim \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

我们称可测函数f为**刚性的** 是指存在序列$n_k \geq 1$ 使得$T^{n_k}f \to f$ (在$L^2(X)$中) .

可以证明：保测系统$(X,\mathcal{X},\mu,T)$ 为mild混合的但且仅当对于任何遍历系统$(Y,\mathcal{Y},\nu,S)$, $\nu(Y)\leq\infty$, 我们有$(X\times Y,\mathcal{X}\times\mathcal{Y},\mu\times\nu,T\times S)$ 为遍历的[70]. 由此易见mild混合介于强弱混合之间. 类似于系统为弱混合当且仅当它没有非常值特征函数, 可以证明, 一个保测系统为mild 混合的当且仅当它在$L^2(X,\mu)$中没有非常值刚性函数[64, 65].

习　题

1. 完成定理2.4.5的证明.
2. 完成定理2.4.8的证明.
3. 设$(X,\mathcal{X},\mu,T)$为弱混合的保测系统. 如果$(X,\mathcal{X},\mu)$具有可数基, 那么存在$J\subseteq\mathbb{Z}_+, d(J)=0$, 使得
$$\lim_{J\not\ni n\to\infty}\mu(A\cap T^{-n}B)=\mu(A)\mu(B),$$
对任意$A,B\in\mathcal{X}$ 成立.
4. 设$(X,\mathcal{X},\mu,T)$为保测系统. 证明：
   (a) 如果T 为弱混合, 根据上面证明可以证明对于任何k, $(X^k,\mathcal{X}^k,\mu^k,T^{(k)})$ 为弱混合的.
   (b) 如果T 为弱混合, 对于任何k, $(X,\mathcal{X},\mu,T^k)$ 为弱混合的.
5. 设$(X,\mathcal{X},\mu,T)$为保测系统. 证明：T 为弱混合的当且仅当对于任何满足$\mu(A)\mu(B)>0$ 的$A,B\in\mathcal{X}$, 我们有
$$N^\mu(A,B)\cap N^\mu(A,A)\neq\emptyset.$$
6. 设G为紧致交换度量群, $T(x)=a\cdot A(x)$ 为仿射.　证明以下等价：
   (a) T为强混合.
   (b) T为弱混合.
   (c) A是遍历的.

§2.5　Koopman-von Neumann谱混合定理

§2.5.1　特征值与特征函数

首先回顾一些基本概念. 设$(X,\mathcal{X},\mu,T)$为保测系统,
$$U_T: L^2(X,\mu)\to L^2(X,\mu).$$
如果存在非零可测函数f 使得
$$U_Tf=\lambda f,$$
复数λ 称为T 的特征值, 这个函数f 称为相应于特征值λ 的特征函数.

定理 2.5.1 *设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 我们有如下结论:*

1. 如果$U_T f=\lambda f, f\in L^2(X,\mu), f\neq 0$, 那么$|\lambda|=1$ 并且$|f|=const$, a.e.

2. 对应于不同特征值的特征函数互相垂直.

3. 如f,g 为同一特征值对应的特征函数, 那么存在$c\in\mathbb{C}$ 使得$f=cg$, 即特征值为简单的.

4. U_T 所有特征值组成了$\mathbb{S}^1$ 的一个子群, 记为Λ_T.

证明. (1) 由$\|U_T f\|=|\lambda|\|f\|$ 以及U_T等距性, 我们有$\|f\|=|\lambda|\|f\|$. 因为$\|f\|\neq 0$, 所以$|\lambda|=1$. 由此又有

$$|U_T f|=|\lambda f|=|f|.$$

根据遍历性, $|f|$ 几乎处处为常数.

(2) 设$\lambda\neq\xi\in\mathbb{S}^1$, $U_T f=\lambda f, U_T g=\xi g$. 于是

$$\langle f,g\rangle=\langle U_T f,U_T g\rangle=\langle\lambda f,\xi g\rangle=\lambda\bar{\xi}\langle f,g\rangle.$$

因为$\lambda\bar{\xi}\neq 1$, 所以$\langle f,g\rangle=0$.

(3) 设$U_T f=\lambda f, U_T g=\lambda g$. 由(1), $|g|=const, a.e.$ 尤其$g\neq 0, a.e.$ 由此f/g 可定义, 并且有$U_T f/g=f/g$. 根据遍历性, 我们得到$f/g=const$.

(4) 设$\lambda\neq\xi\in\mathbb{S}^1$, $U_T f=\lambda f, U_T g=\xi g$. 由$f,g$ 非零, 我们得到

$$U_T(f\bar{g})=U_T f U_T\bar{g}=(\lambda\bar{\xi})(f\bar{g}).$$

即$\lambda\xi^{-1}=\lambda\bar{\xi}$ 为U_T 的特征值. 所以特征值全体组成一个$\mathbb{S}^1$ 的子群. □

注记 2.5.2 *1. 因为我们总是假设$(X,\mathcal{X},\mu)$ 是可数生成的, 所以Hilbert 空间$\mathcal{H}=L^2(X,\mathcal{X},\mu)$ 为可分的. 于是根据 (2) (3) 每个特征空间是一维的且互相垂直, 于是特征值组成的群Λ_T 为可数群.*

*2. 设U 为一般Hilbert 空间$\mathcal{H}$ 上的酉算子. 如果存在$\lambda\in\mathbb{C}$ 满足$Ux=\lambda x$, 那么$x\in\mathcal{H}\setminus\{0\}$ 称为***特征向量***, 而λ 称为***相对于x的特征值***. 不难看出定理2.5.1 的结论仍然是成立的.*

§2.5.2 Koopman-von Neumann谱混合定理

设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统. 如果$f\in\mathcal{H}=L^2(X,\mu)$ 为系统$(X,\mathcal{X},\mu,T)$ 的特征函数, 则容易知道$\overline{\{U_T^n f:n\in\mathbb{Z}\}}$ 为$\mathcal{H}$ 的紧子集. 一般地,

定义 2.5.3 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统. 我们称使得

$$\overline{\{U_T^n f:n\in\mathbb{Z}\}}$$

为$\mathcal{H}$的紧子集的函数f 为**几乎周期函数** 或者**紧函数**.

$\mathcal{H}$ 的子集$\mathcal{H}_1$ 称为**代数**, 如果$\mathcal{H}_1$ 为$L^\infty(X,\mathcal{X},\mu)$ 的线性子空间且对任意$f,g\in\mathcal{H}_1$, 我们有$fg\in\mathcal{H}_1$. 显然系统$(X,\mathcal{X},\mu,T)$ 所有有界的几乎周期函数全体(定义为$\mathcal{A}_d$) 形成了$\mathcal{H}$ 的一个U_T 不变和复共轭不变的代数, 即$\mathcal{A}_d$ 为$\mathcal{H}$ 的线性子空间且对任意$f,g\in\mathcal{A}_d$, 我们有U_Tf, fg 和$\overline{f}$ 均属于$\mathcal{A}_d$. 对实的几乎周期函数f 和$M>0$, f 的截断函数

$$f_M(x)=\begin{cases} f(x), & \text{如果}|f(x)|<M\\ \mathrm{sign}(f(x))\cdot M, & \text{如果}|f(x)|\ge M\end{cases}$$

为有界的几乎周期函数. 这说明$\mathcal{A}_d$ 在$\mathcal{H}$ 中的闭包包含了全体实的几乎周期函数. 进而由于几乎周期函数的实部和虚部仍为几乎周期函数, $\mathcal{A}_d$ 在$\mathcal{H}$ 中的闭包(定义为$\mathcal{H}_d$) 恰为几乎周期函数全体.

显然$\mathcal{H}$ 中特征函数的线性组合全体的闭包包含于$\mathcal{H}_d$, 事实上$\mathcal{H}_d$ 恰为特征函数的线性组合全体的闭包且我们可以将$\mathcal{H}_d^\perp$ 的元素刻划出来. 这就是著名的Koopman-von Neumann谱混合定理.

定理 2.5.4 (Koopman-von Neumann谱混合定理) 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统, $\mathcal{H}=L^2(X,\mathcal{X},\mu)$. 则

$$\mathcal{H}=\mathcal{H}_d\bigoplus\mathcal{H}_c,$$

其中

$$\begin{aligned}\mathcal{H}_d&=\{f\in\mathcal{H}:f\text{是几乎周期函数}\}=\overline{\mathrm{span}\{f\in\mathcal{H}:\exists\lambda\in\mathbb{C},\text{使得 }U_T(f)=\lambda f\}},\\ \mathcal{H}_c&=\{f\in\mathcal{H}:\exists S\subseteq\mathbb{N},d(S)=1\text{使得对任意}g\in\mathcal{H}\text{有}\lim_{S\ni n\to+\infty}\langle U_T^nf,g\rangle=0\},\end{aligned}\tag{2.5.1}$$

这里$d(S)$为S的密度以及$\langle\cdot,\cdot\rangle$为$\mathcal{H}$的内积.

在Koopman-von Neumann 谱混合定理中, 如果$\mathcal{H}=\mathcal{H}_d$, 则称系统$(X,\mathcal{X},\mu,T)$ 或T 为**离散谱的**, 此时系统的特征函数的线性组合张成了整个Hilbert 空间$\mathcal{H}$. 由于非零常值函数为T的特征函数, 所以$\mathcal{H}_d$ 至少为$\mathcal{H}$的一维线性子空间. 我们也用复数集合$\mathbb{C}$ 来表示所有常值函数组成的空间. 如果$\mathcal{H}_d=\mathbb{C}$, 那么称系统$(X,\mathcal{X},\mu,T)$ 或T 为**连续谱的**. 由Koopman-von Neumann 谱混合定理, 系统$(X,\mathcal{X},\mu,T)$ 为弱混合的当且仅当它为连续谱的（定理2.4.13）.

事实上，对于一般Hilbert空间上的酉算子，Koopman-von Neumann谱混合定理仍然成立，我们在下面给出的就是一般意义下Koopman-von Neumann谱混合定理的证明.

§2.5.3 Herglotz定理以及谱测度

下面给出Koopman-von Neumann谱混合定理的证明, 为此我们首先简要介绍一些谱理论的基本知识. 在本节中出现的Hilbert 空间我们均假设为可分的.

定义 2.5.5 函数$\phi:\mathbb{Z}\to\mathbb{C}$ 称为**正定的**, 如果对每个有限的复数序列$\{a_n\}_{n=0}^N$ 有

$$\sum_{m=0}^N\sum_{n=0}^N a_m\overline{a_n}\phi(m-n)\ge 0.$$

我们给出两个正定的函数的例子：

1. 设U 为Hilbert 空间$\mathcal{H}$ 上的酉算子和$x \in \mathcal{H}$. 因对每个有限的复数序列$\{a_n\}_{n=0}^N$ 有
$$\sum_{n,m=0}^{N} \langle U^{m-n}x, x\rangle a_m \overline{a_n} = \langle \sum_{n=0}^{N} a_n U^n x, \sum_{m=0}^{N} a_m U^m x\rangle \geq 0,$$
函数$\phi(n) = \langle U^n x, x\rangle$ 是正定的.

2. 设μ 为$\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$ 上的非负Borel 有限测度, 该测度的Fourier 变换: 函数
$$\widehat{\mu}(n) = \int_{\mathbb{S}^1} z^n d\mu(z)$$
是正定的.

上面例(2)的逆命题也是成立的, 这就是著名的Herglotz定理. 为了完整起见，我们给出这个定理的证明，对证明不感兴趣的读者可以跳过这个证明.

定理 2.5.6 (Herglotz定理) *对给定正定函数$\phi : \mathbb{Z} \to \mathbb{C}$, 存在$\mathbb{S}^1$ 上唯一的非负Borel 有限测度μ 使得*
$$\widehat{\mu}(n) = \phi(n), \quad \forall n \in \mathbb{Z}.$$

证明. 由于ϕ为正定函数, 我们不难看出$\phi(0) \geq 0$以及对每个$\lambda \in \mathbb{C}$,
$$(1 + |\lambda|^2)\phi(0) + \phi(n)\lambda + \phi(-n)\overline{\lambda} \geq 0.$$
因此对每个$\lambda \in \mathbb{C}$我们有
$$\phi(n)\lambda + \phi(-n)\overline{\lambda}$$
为实数, 这说明$\phi(-n) = \overline{\phi(n)}$. 令$\lambda = \theta\overline{\phi(n)}$ 于是
$$(1 + |\theta|^2|\phi(n)|^2)\phi(0) + \theta|\phi(n)|^2 + \overline{\theta}|\phi(n)|^2 \geq 0$$
对每个$\theta \in \mathbb{C}$成立.

特别地, 当θ为实数时, 上式表明
$$(1 + \theta^2|\phi(n)|^2)\phi(0) + 2\theta|\phi(n)|^2 \geq 0$$
对全体实数θ成立, 由此可推得
$$|\phi(n)| \leq \phi(0), \forall n \in \mathbb{Z}.$$
这说明函数$\phi(n)$是有界的. 如果$\phi(0) = 0$, 则$\phi(n) \equiv 0$, 此时唯一非负测度为在每个Borel集上取零值的测度.

以下设$\phi(0) > 0$, 不失一般性, 不妨设$\phi(0) = 1$. 对$s \in (0,1)$, 从正定性可以推出对所有$|z| = 1$有

$$f_s(z) = \sum_{n,m=0}^{\infty} \phi(n-m) s^{n+m} z^{m-n} \geq 0.$$

注意到

$$\begin{aligned}\sum_{n,m=0}^{\infty} \phi(m-n) s^{n+m} z^{n-m} &= \sum_{n=-\infty}^{+\infty} \phi(n) z^{-n} \sum_{m=0}^{\infty} s^{|n|+2m} \\ &= \sum_{n=-\infty}^{+\infty} \phi(n) z^{-n} s^{|n|} \frac{1}{1-s^2}.\end{aligned}$$

因此

$$\int_{\mathbb{S}^1} f_s(z) z^{-n} dz = \frac{\phi(-n) s^{|n|}}{1-s^2}. \tag{2.5.2}$$

现在定义$\mathbb{S}^1$上的非负Borel测度μ_s使得

$$\frac{d\mu_s}{dz} = (1-s^2) f_s(z) \geq 0.$$

由等式(2.5.2)知

$$\int_{\mathbb{S}^1} z^{-n} d\mu_s = \phi(-n) s^{|n|},\ \mu_s(\mathbb{S}^1) = \phi(0) = 1, \tag{2.5.3}$$

这说明$\mu_s \in \mathcal{M}(\mathbb{S}^1)$, 这里$\mathcal{M}(\mathbb{S}^1)$为$\mathbb{S}^1$上Borel概率测度全体.

选择序列$s_m \to 1 (0 < s_m < 1)$, 使得$\mu_{s_m} \to \mu$在$\mathcal{M}(\mathbb{S}^1)$的弱*拓扑意义下成立. 在等式(2.5.3)中取$s = s_m$, 再让$m \to \infty$我们有$\widehat{\mu}(n) = \phi(n), \forall n \in \mathbb{Z}$. 至此, 定理的存在性部分得证.

以下说明唯一性, 设ν为$\mathbb{S}^1$上使得

$$\widehat{\nu}(n) = \phi(n), \forall n \in \mathbb{Z}$$

非负的有限Borel测度.则对$\{z^k\}_{k=0}^{\infty}$的任意有限的线性组合$p(z)$, 我们有

$$\int_{\mathbb{S}^1} p(z) d\nu = \int_{\mathbb{S}^1} p(z) d\mu.$$

因如此的函数在$C(\mathbb{S}^1)$中稠密, 所以

$$\int_{\mathbb{S}^1} f(z) d\nu = \int_{\mathbb{S}^1} f(z) d\mu$$

对所有$f \in C(\mathbb{S}^1)$成立, 这就说明$\nu = \mu$. □

设U 为Hilbert空间$\mathcal{H}$ 上的酉算子和$x \in \mathcal{H}$. 我们用$Z(x)$ 表示由x 生成的循环子空间, 即$Z(x)$ 是$\mathcal{H}$ 的包含$\{U^n x : n \in \mathbb{Z}\}$ 的最小的闭子空间. 因函数$\phi(n) = \langle U^n x, x\rangle$ 是正定的, 由Herglotz 定理知存在$\mathbb{S}^1$ 上唯一的非负Borel 测度σ_x 使得

$$\widehat{\sigma_x}(n) = \langle U^n x, x\rangle.$$

特别地,

$$\sigma_x(\mathbb{S}^1)=\widehat{\sigma_x}(0)=\int_{\mathbb{S}^1}1d\sigma_x=\langle x,x\rangle=||x||^2.$$

我们称σ_x 为x (相对于U)的**谱测度**. 不难看出$x=0$ 当且仅当σ_x 为零测度.

命题 2.5.7 *设U为Hilbert空间$\mathcal{H}$上的酉算子和$x\in\mathcal{H}\setminus\{0\}$. 定义*

$$V:L^2(\mathbb{S}^1,\sigma_x)\to L^2(\mathbb{S}^1,\sigma_x)\text{ 使得 }Vf(z)=zf(z)\ (z\in\mathbb{S}^1),$$

则V为酉算子且$W_x^{-1}VW_x=U$, 其中$W_x:Z(x)\to L^2(\mathbb{S}^1,\sigma_x)$为映射$U^nx\mapsto z^n\in L^2(\mathbb{S}^1,\sigma_x)$的唯一线性等距扩张.

$$\begin{CD} Z(x) @>U>> Z(x)\\ @VW_xVV @VVW_xV\\ L^2(\mathbb{S}^1,\sigma_x) @>>V> L^2(\mathbb{S}^1,\sigma_x). \end{CD}$$

证明. 如果$p(z)=\sum_{j=-m}^m a_jz^j$ 和$q(z)=\sum_{l=-n}^n b_lz^l$ 为三角多项式, 则

$$\begin{aligned}&\langle p(U)x,q(U)x\rangle_{Z(x)}\\ =&\sum_{j=-m}^m\sum_{l=-n}^n a_j\overline{b_l}\langle U^{j-l}x,x\rangle_{Z(x)}=\sum_{j=-m}^m\sum_{l=-n}^n a_j\overline{b_l}\widehat{\sigma_x}(j-l)\\ =&\sum_{j=-m}^m\sum_{l=-n}^n a_j\overline{b_l}\int_{\mathbb{S}^1}z^{j-l}d\sigma_x=\langle p(z),q(z)\rangle_{L^2(\mathbb{S}^1,\sigma_x)}.\end{aligned}$$

因为形如$p(U)x$ 的向量在$Z(x)$ 中稠密, 与此同时三角多项式$p(z)$ 在$L^2(\mathbb{S}^1,\sigma_x)$ 中稠密, 这就完成了证明. □

命题 2.5.8 *设U 为Hilbert 空间$\mathcal{H}$ 上的酉算子和$x\in\mathcal{H}$.*

1. *设μ 为$\mathbb{S}^1$ 上非负的有限Borel 测度, 如果$\mu\ll\sigma_x$, 则存在$y\in Z(x)$ 使得$\sigma_y=\mu$.*
2. *设$y\in\mathcal{H}$, 如果$Z(y)\perp Z(x)$, 则$\sigma_{x+y}=\sigma_x+\sigma_y$.*

证明. (1). 如果σ_x 为零测度, 这是明显地. 以下假设σ_x 不为零测度, 即$x\neq 0$. 设$f=\sqrt{\frac{d\mu}{d\sigma_x}}$, 则$f\in L^2(\mathbb{S}^1,\sigma_x)$. 取$y=W_x^{-1}f$, 其中$W_x$ 如命题2.5.7所定义. 则$y\in Z(x)$ 且

$$\begin{aligned}\widehat{\sigma_y}(n)&=\langle U^ny,y\rangle_{Z(x)}=\langle W_xU^ny,W_xy\rangle_{L^2(\mathbb{S}^1,\sigma_x)}\\ &=\langle V^nf,f\rangle_{L^2(\mathbb{S}^1,\sigma_x)}=\int_{\mathbb{S}^1}z^n\frac{d\mu}{d\sigma_x}d\sigma_x\\ &=\widehat{\mu}(n)\end{aligned}$$

这说明$\sigma_y=\mu$.

(2). 设$y\in\mathcal{H}$, 如果$Z(y)\perp Z(x)$, 则

$$\widehat{\sigma_{x+y}}(n)=\langle U^n(x+y),x+y\rangle=\langle U^nx,x\rangle+\langle U^ny,y\rangle=\widehat{\sigma_x+\sigma_y}(n).$$

于是$\sigma_{x+y}=\sigma_x+\sigma_y$. □

§2.5.4 离散谱与连续谱

定义 2.5.9 设U 为*Hilbert* 空间$\mathcal{H}$ 上的酉算子. 我们用$\mathcal{H}_d$表示$\mathcal{H}$的所有特征向量线性组合的闭包, $\mathcal{H}_d$ 为$\mathcal{H}$的子空间, 称之为$\mathcal{H}$的**离散谱空间**. 称$\mathcal{H}_c \triangleq \mathcal{H}_d^{\perp}$ 为**连续谱空间**.

显然$U\mathcal{H}_d = \mathcal{H}_d$ 和$U\mathcal{H}_c = U\mathcal{H}_d^{\perp} = \mathcal{H}_d^{\perp} = \mathcal{H}_c$.

设μ 为$\mathbb{S}^1$ 上的正的Borel 测度, 如果存在$\mathbb{S}^1$ 的可数子集A 满足$\mu(\mathbb{T} \setminus A) = 0$, 则称$\mu$ 是**离散的**或者**纯原子的**; 如果对每个点$z \in \mathbb{S}^1$ 有$\mu(\{z\}) = 0$, 则称μ 是**连续的**. 易见，测度μ为离散的当且仅当存在$a_i \geq 0, x_i \in \mathbb{S}^1, i \in \mathbb{N}$ 使得$\mu = \sum_{i\in\mathbb{N}} a_i\delta_{x_i}$, $\sum_{i\in\mathbb{N}} a_i < \infty$.

命题 2.5.10 设U 为*Hilbert* 空间$\mathcal{H}$ 上的酉算子, $\mathcal{H}_d$ 为$\mathcal{H}$ 的离散谱空间. 则

1. 对$x \in \mathcal{H} \setminus \{0\}$, $x \in \mathcal{H}_d$ 当且仅当σ_x 为离散的.

2. 对$x \in \mathcal{H} \setminus \{0\}$, $x \in \mathcal{H}_c$ 当且仅当σ_x 为连续的.

证明. (i) 首先,我们说明如果$x \in \mathcal{H}_d$, 则σ_x为离散的. 设$x \in \mathcal{H}_d$,则x可以写为

$$x = \sum_{i\in I} a(i)x_i,$$

其中I为可数集, $a(i) \in \mathbb{C}$, x_i为具有特征值λ_i的特征向量且$\langle x_i, x_j\rangle = 0$当$i \neq j \in I$, 以及

$$\sum_{i\in I} |a(i)|^2||x_i||^2 < +\infty.$$

因此$U^n x = \sum\limits_{i\in I} a(i)\lambda_i^n x_i$以及

$$\int_{\mathbb{S}^1} z^n d\sigma_x(z) = \langle U^n x, x\rangle = \sum_{i\in I} |a(i)|^2\lambda_i^n||x_i||^2.$$

这样由Herglotz定理可得

$$\sigma_x = \sum_{i\in I} |a(i)|^2||x_i||^2\delta_{\lambda_i},$$

其中$\delta_z, z \in \mathbb{S}^1$ 表示集中在z上的点测度(即对$\mathbb{S}^1$ 的每个Borel集A, 如果$z \in A$则$\delta_z(A) = 1$, 否则$\delta_z(A) = 0$). 这样σ_x是为离散的.

(ii) 其次, 如果$x \in \mathcal{H}_c = \mathcal{H}_d^{\perp}$, 则$\sigma_x$ 为连续的. 如若不然, 存在$\lambda \in \mathbb{S}^1$ 使得$\sigma_x(\{\lambda\}) > 0$, 这说明$\delta_\lambda \ll \sigma_x$, 因此由命题2.5.8知存在$y \in Z(x)$ 使得$\sigma_y = \delta_\lambda$. 于是

$$\langle U^n y, y\rangle = \int_{\mathbb{S}^1} z^n d\delta_\lambda(z) = \lambda^n, \quad \forall n \in \mathbb{Z}.$$

尤其$\|y\|^2 = \langle y, y\rangle = 1, \langle Uy, y\rangle = \lambda$. 进而

$$\begin{aligned}\langle Uy - \lambda y, Uy - \lambda y\rangle &= \langle Uy, Uy\rangle - \lambda\overline{\langle Uy, y\rangle} - \overline{\lambda}\langle Uy, y\rangle + \langle y, y\rangle \\ &= 2\langle y, y\rangle - 2|\lambda|^2 = 2 - 2 = 0.\end{aligned}$$

即$Uy=\lambda y$. 于是

$$y\in\mathcal{H}_d\cap Z(x)\subseteq\mathcal{H}_d\cap\mathcal{H}_d^{\perp}=\{0\}.$$

因此$y=0$, 这与$\sigma_y=\delta_\lambda$不为零测度相矛盾.

(iii) 设$x\in\mathcal{H}$, 则存在$x_d\in\mathcal{H}_d, x_c\in\mathcal{H}_d^{\perp}$ 使得$x=x_d+x_c$. 由于$Z(x_d)\perp Z(x_c)$, 从命题2.5.8 的性质(2)知$\sigma_x=\sigma_{x_d}+\sigma_{x_c}$. 现在, 由上面的(i)和(ii)知：如果$x_d\neq 0$ 则σ_{x_d} 为离散的; 如果$x_c\neq 0$ 则σ_{x_c} 为连续的.

因此如果σ_x 为离散的, 则σ_{x_c} 为零测度, 即$x_c=0$ 和$x=x_d\in\mathcal{H}_d$; 同理, 如果σ_x 为连续的, 则$x\in\mathcal{H}_d^{\perp}$. 这就完成了命题的证明. □

注记 2.5.11 从上面定理证明可以看出, x 为以λ 为特征值的特征向量当且仅当$\sigma_x=||x||^2\delta_\lambda$.

定理 2.5.12 (Wiener定理) *设U 为Hilbert空间$\mathcal{H}$ 上的酉算子, $x\in\mathcal{H}$. 则以下陈述彼此等价*

1. $x\in\mathcal{H}_c$.
2. $\lim\limits_{N\to\infty}\frac{1}{N}\sum\limits_{n=0}^{N-1}|\langle U^nx,x\rangle|^2=0$.
3. 对每个$y\in\mathcal{H}$, $\lim\limits_{N\to\infty}\frac{1}{N}\sum\limits_{n=0}^{N-1}|\langle U^nx,y\rangle|^2=0$.
4. 存在$S\subseteq\mathbb{Z}_+, d(S)=1$ 使得对每个$y\in\mathcal{H}$, $\lim\limits_{S\ni n\to+\infty}\langle U^nx,y\rangle=0$.

证明. (4) ⇒ (3) ⇒ (2) 是明显成立的.

(2) ⇔ (1) 首先, 对每个$z\in\mathbb{T}$我们有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N}z^n=\delta_1(\{z\}).$$

因此

$$\begin{aligned}&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|\langle U^nx,x\rangle|^2\\&=\lim_{N\to\infty}\int_{\mathbb{S}^1\times\mathbb{S}^1}\frac{1}{N}\sum_{n=0}^{N-1}(z_1\overline{z_2})^n d\sigma_x\times\sigma_x(z_1,z_2)\\&=\int_{\mathbb{S}^1\times\mathbb{S}^1}\delta_1(\{z_1\overline{z_2}\})d\sigma_x\times\sigma_x(z_1,z_2)\\&=\int_{\mathbb{S}^1}\int_{\mathbb{S}^1}\delta_1(\{z_1\overline{z_2}\})d\sigma_x(z_1)d\sigma_x(z_2)\\&=\int_{\mathbb{S}^1}\sigma_x(\{z_2\})d\sigma_x(z_2).\end{aligned}\tag{2.5.4}$$

这就说明

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|\langle U^n x,x\rangle|^2=0$$

当且仅当$\sigma_x(\{z_2\})=0$对每个$z_2\in\mathbb{S}^1$成立, 即(1) $\Leftrightarrow$ (2).

(2) $\Rightarrow$ (4) 假设(2)成立.令

$$\mathcal{H}_x=\{y\in\mathcal{H}:\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|\langle U^n x,y\rangle|^2=0\}$$

则容易验证$\mathcal{H}_x$ 为$\mathcal{H}$的U不变的闭子空间. 因$x\in\mathcal{H}_x$, $Z(x)\subseteq\mathcal{H}_x$. 显然$Z(x)^\perp\subseteq\mathcal{H}_x$. 这说明$\mathcal{H}_x=\mathcal{H}$.

由于$\mathcal{H}$ 为可分空间, 取$\mathcal{H}$ 的可数稠密子集$\{y'_j\}_{j\in\mathbb{N}}$. 令$y_j=\frac{y'_j}{||y'_j||}$, 如果$y'_j\neq 0$, 否则令$y_j=0$. 设

$$a_n=\sum_{j=1}^{\infty}\frac{1}{2^j}|\langle U^n x,y_j\rangle|^2,\ \forall n\in\mathbb{Z}_+.$$

因$\mathcal{H}_x=\mathcal{H}$, $\lim\limits_{N\to+\infty}\frac{1}{N}\sum_{n=1}^{N}|a_n|=0$. 现在利用引理2.4.6知存在$S\subseteq\mathbb{N},d(S)=1$ 使得

$$\lim_{S\not\ni n\to\infty}a_n=0.$$

通过简单的逼近讨论, 便可获得(4). □

§2.5.5 几乎周期向量

定义 2.5.13 设U为*Hilbert*空间$\mathcal{H}$上的酉算子, 我们把满足$\overline{\{U^n x:n\in\mathbb{Z}\}}$为$\mathcal{H}$的紧子集的向量$x$ 称为**几乎周期向量**或者**紧向量**.

我们用$\mathcal{H}_{ap}$ 记全体的几乎周期向量组成的集合. 注意到两个几乎周期向量的线性组合仍为几乎周期向量, 我们不难说明$\mathcal{H}_{ap}$ 为$\mathcal{H}$的闭子空间. 显然特征向量为几乎周期向量, 进而$\mathcal{H}_d\subseteq\mathcal{H}_{ap}$. 事实上, 以下定理说明相反的包含关系也是成立的.

定理 2.5.14 设U为*Hilbert*空间$\mathcal{H}$上的酉算子.则$\mathcal{H}_{ap}=\mathcal{H}_d$.

证明. 我们只需证明$\mathcal{H}_{ap}\subseteq\mathcal{H}_d$. 任取$x\in\mathcal{H}_{ap}$, 将$x$ 分解为$x=x_1+x_2$ 满足$x_1\in\mathcal{H}_d\subseteq\mathcal{H}_{ap}$ 和$x_2\in\mathcal{H}_d^\perp$. 显然也有$x_2=x-x_1\in\mathcal{H}_{ap}$. 如果我们能说明$x_2=0$, 则$x=x_1\in\mathcal{H}_d$, 这就完成了证明.

如果$x_2\neq 0$, 则$\epsilon=\frac{||x_2||}{2}>0$. 因$\overline{\{U^n x_2:n\in\mathbb{Z}\}}$ 为$\mathcal{H}$ 的紧子集, 存在$N\in\mathbb{N}$ 使得对每个$n\in\mathbb{N}$ 有

$$\min_{k\in\{0,1,\cdots,N\}}||U^n x_2-U^k x_2||<\epsilon. \tag{2.5.5}$$

进而, 我们能够找到$k\in\{0,1,\cdots,N\}$ 使得集合

$$F=\{n\in\mathbb{Z}_+:||U^n x_2-U^k x_2||<\epsilon\}$$

满足$\underline{d}(F)=\liminf\limits_{m\to\infty}\dfrac{|F\cap[0,m-1]|}{m}>0$.

注意到当$n\in F$,

$$\begin{aligned}
&|\langle U^n x_2, U^k x_2\rangle|\\
=&|\langle U^k x_2, U^k x_2\rangle+\langle U^n x_2-U^k x_2, U^k x_2\rangle|\\
=&|||x_2||^2+\langle U^n x_2-U^k x_2, U^k x_2\rangle|\\
\geq&||x_2||^2-|\langle U^n x_2-U^n x_2, U^k x_2\rangle|\\
\geq&||x_2||^2-||U^n x_2-U^k x_2||\cdot||U^k x_2||\\
\geq&||x_2||^2-\epsilon||x_2||=\epsilon||x_2||,
\end{aligned}\tag{2.5.6}$$

我们有

$$\begin{aligned}
&\limsup_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m-1}|\langle U^n x_2, U^k x_2\rangle|^2\\
\geq&\liminf_{m\to\infty}\frac{1}{m}\sum_{n\in F\cap[0,m-1]}|\langle U^n x_2, U^k x_2\rangle|^2\\
\geq&\liminf_{m\to\infty}\frac{1}{m}\sum_{n\in F\cap[0,m-1]}\epsilon^2||x_2||^2\quad(\text{由}(2.5.6))\\
=&\underline{d}(F)\epsilon^2||x_2||^2>0.
\end{aligned}\tag{2.5.7}$$

这与$x_2\in\mathcal{H}_d^{\perp}$相矛盾(参见定理2.5.12的性质(3)). □

下面我们回到Koopman-von Neumann谱混合定理（定理2.5.4）. 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统. 在可分的复Hilbert空间$\mathcal{H}=L^2(X,\mathcal{X},\mu)$ 上, 对酉算子$U_T:\mathcal{H}\to\mathcal{H}, f\mapsto f\circ T$ 运用定理2.5.12和定理2.5.14即可得Koopman-von Neumann谱混合定理.

§2.5.6 Kronecker 代数

下面我们进一步研究$\mathcal{H}_d$. 首先我们需要一个经典结论(参见[203, 定理1.2]):

定理 2.5.15 *设$(X,\mathcal{X},\mu,T)$ 为保测系统, $\mathcal{H}_1$是由$L^2(X,\mathcal{X},\mu)$中有界函数构成的复共轭不变的代数.则存在$\mathcal{X}$ 的子σ代数$\mathcal{A}$ 使得$\overline{\mathcal{H}_1}=L^2(X,\mathcal{A},\mu)$. 进而如果$T$可逆且$\mathcal{H}_1$ 为U_T 不变的, 则$\mathcal{A}$ 为T 不变的.*

证明. 设$\mathcal{A}$为$\mathcal{X}$中由

$$\mathcal{D}=\{f^{-1}(B):f\in\mathcal{H}_1, B\text{为}\mathbb{R}\text{ 的Borel 子集}\}$$

生成的最小的σ代数. 显然, $\mathcal{H}_1\subseteq L^2(X,\mathcal{A},\mu)$. 由于$L^2(X,\mathcal{A},\mu)$为$L^2(X,\mathcal{X},\mu)$的闭线性子空间, $\mathrm{cl}(\mathcal{H}_1)\subseteq L^2(X,\mathcal{A},\mu)$.

对于反向的包含关系$L^2(X,\mathcal{A},\mu)\subseteq\overline{\mathcal{H}_1}$, 由于$\overline{\mathcal{H}_1}$为$L^2(X,\mathcal{X},\mu)$的闭线性子空间以及$\mathcal{A}$可测的特征函数的线性组合在$L^2(X,\mathcal{A},\mu)$中稠密, 我们只需证明：如果$A\in\mathcal{A}$, 则特征函数$1_A\in\overline{\mathcal{H}_1}$. 设

$$\mathcal{D}_1=\{A_1\cap\cdots\cap A_n : A_1,\cdots,A_n\in\mathcal{D}\}$$

和$\mathcal{D}_0$为$\mathcal{D}_1$中有限个互不相交元素的并构成的集族.则$\mathcal{D}_0$为由$\mathcal{D}$生成的代数.

因代数$\mathcal{D}_0$生成了σ代数$\mathcal{A}$, 我们只需证明：如果$A\in\mathcal{D}_0$,则特征函数$1_A\in\overline{\mathcal{H}_1}$即可. 进而, 因$\mathcal{D}_0$的每个元素可以表示为$\mathcal{D}_1$中有限个互不相交元素的并, 我们只需证明：如果$A\in\mathcal{D}_1$, 则特征函数$\mathbf{1}_A\in\overline{\mathcal{H}_1}$ 即可.

设$f_i\in\mathcal{H}_1$, $R_i=||f_i||_\infty$以及$B_i\subseteq[-R_i,R_i]$为$\mathbb{R}$的Borel子集, $i=1,2,\cdots,k$. 设$g_i(x)=\mathbf{1}_{B_i}(x), i=1,2,\cdots,k$. 则对每个$g_i$, 存在$\mathbb{R}$ 上的多项式序列$P_{i,n}$使得

$$\lim_{n\to+\infty}P_{i,n}(x)=\mathbf{1}_{B_i}(x)$$

对$x\in[-R_i,R_i]$成立. 现在, 因$\mathcal{H}_1$为代数,易见$\prod_{i=1}^k P_{i,n}\circ f_i\in\mathcal{H}_1$. 令$n\to+\infty$, 我们有

$$\prod_{i=1}^k P_{i,n}\circ f_i(x)\to\prod_{i=1}^k g_i\circ f_i(x)$$

对μ-a.e. $x\in X$. 当然上述收敛也在$L^2(X,\mathcal{X},\mu)$ 意义下成立, 因此$\prod_{i=1}^k g_i\circ f_i\in\overline{\mathcal{H}_1}$. 即

$$\mathbf{1}_{\cap_{i=1}^k f_i^{-1}(B_i)}=\prod_{i=1}^k\mathbf{1}_{f_i^{-1}(B_i)}=\prod_{i=1}^k\mathbf{1}_{B_i}\circ f_i\in\overline{\mathcal{H}_1},$$

这就完成了证明. □

注意到$\mathcal{H}_d$ 为$L^2(X,\mathcal{X},\mu)$ 的U_T 不变和共轭不变的代数$\mathcal{A}_c$ 的闭包, 从引理2.5.15 我们知道存在$\mathcal{X}$ 的T 不变的子σ 代数$\mathcal{K}_\mu$ 使得$\mathcal{H}_d=L^2(X,\mathcal{K}_\mu,\mu)$.

定义 2.5.16 我们称$\mathcal{K}_\mu$ 为$(X,\mathcal{X},\mu,T)$ 的**Kronecker 代数**. 对应的系统称为**Kronecker 因子**.

由Koopman-von Neumann 谱混合定理, 我们知系统$(X,\mathcal{X},\mu,T)$ 为离散谱的当且仅当$\mathcal{K}_\mu=\mathcal{X}$; 系统$(X,\mathcal{X},\mu,T)$ 为弱混合的当且仅当$\mathcal{K}_\mu=\mathcal{N}=\{X,\emptyset\}$.

§2.5.7　一些补充

设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为概率空间, $H(x,y)\in L^2(X\times Y,\mu\times\nu)$. 定义算子

$$L_H: L^2(Y,\nu)\to L^2(X,\mu),\ \phi\mapsto H*\phi$$

$$H*\phi=\int_Y H(x,y)\phi(y)d\nu(y).$$

这个算子是紧算子. 如果$H(x,y)=\overline{H(y,x)}$, 那么算子L_H为*Hermitian* 的. 一般把这个算子称为***Hilbert-Schmidt*** 算子.

在最后一章我们会证明下面定理的推广，在本处我们不再给出证明.

定理 2.5.17 设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 那么

1. 如果$H\in L^2(X\times X,\mu\times\mu)$ 为非常值$T\times T$ 的函数, 那么L_H 的值域有一组T的特征函数组成的基, 至少有一个是非常值的.
2. 如果f为非零的紧函数, 且设$f\otimes\overline{f}\in L^2(X\times X.\mu\times\mu)$. 那么$\mathbb{E}(f\otimes\overline{f}|\mathcal{I}(T\times T))\neq 0$.
3. 如果紧函数全体是稠密的, 那么$\{H*\phi: H\in L^\infty(X\times X,\mu\times\mu),\phi\in L^\infty(X,\mu)\}$ 也是稠密的.

根据上面定理可以得到

定理 2.5.18 设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 那么

- $(X,\mathcal{X},\mu,T)$是紧的当且仅当$(X,\mathcal{X},\mu,T)$具有离散谱.
- $(X,\mathcal{X},\mu,T)$为弱混合当且仅当T 有连续谱.

习 题

1. 运用本节知识证明von Neumann 遍历定理.
2. 完成定理2.5.17的证明.

§2.6 动力系统谱理论简介

§2.6.1 $\mathcal{M}(\mathbb{S}^1)$概念回顾

设$\mathcal{M}(\mathbb{S}^1)$ 为$\mathbb{S}^1$ 上全体Borel（复）测度全体. 其上卷积定义为:

$$\mu*\nu(E)=\int_{\mathbb{S}^1}\mu(E-t)d\nu(t)m,\quad \mu,\nu\in\mathcal{M}(\mathbb{S}^1),E\in\mathcal{B}(\mathbb{S}^1),$$

模定义为$\|\mu\|=\int_{\mathbb{S}^1}d|\mu|$, 由此$\big(\mathcal{M}(\mathbb{S}^1),*,\|\cdot\|\big)$ 成为一个代数. 对于μ, 其Fourier 系数定义为

$$\hat{\mu}(n)=\int_{\mathbb{S}^1}z^n d\mu(z),\ \forall n\in\mathbb{Z}.$$

一个重要的性质是:

$$\widehat{\mu*\nu}(n)=\hat{\mu}(n)\hat{\nu}(n),\ \forall n\in\mathbb{Z}.$$

把全体离散测度集合记为$\mathcal{M}_d(\mathbb{S}^1)$; 全体连续测度集合记为$\mathcal{M}_c(\mathbb{S}^1)$. 对于任何$\mu\in\mathcal{M}(\mathbb{S}^1)$, 存在唯一的$\mu_d\in\mathcal{M}_d(\mathbb{S}^1)$ 和$\mu_c\in\mathcal{M}_c(\mathbb{S}^1)$ 使得

$$\mu=\mu_d+\mu_c.$$

定义 2.6.1 设m 为$\mathbb{S}^1$ 上的*Lebesgue*测度, 如果$\mu\in\mathcal{M}(\mathbb{S}^1)$ 与m 为奇异的, 那么就称μ 为**奇异的**; 如果$\mu\ll m$, 那么就称μ 为**绝对连续的**.

令$\mu:=\mu*\mu*\ldots*\mu$ (n 次), 如果$\mu^n\perp\mu^m,\forall n\neq m\in\mathbb{N}$, 那么就称$\mu$ 具有**独立幂** *(independent powers)*.

根据Riemann-Lebesgue 引理, 我们有：

命题 2.6.2 *如果$\mu \in \mathcal{M}(\mathbb{S}^1)$ 为绝对连续的, 那么$\lim_{n\to\infty} \hat{\mu}(n) = 0$.*

定理 2.6.3 (Wiener) *设$\mu \in \mathcal{M}(\mathbb{S}^1)$ 为有限Borel测度. 如果H 为$L^2(\mathbb{S}^1, \mu)$ 的V 不变闭子空间($VH = H$):*

$$V : L^2(\mathbb{S}^1, \mu) \to L^2(\mathbb{S}^1, \mu),\ f(z) \mapsto zf(z).$$

那么存在$B \in \mathcal{B}(\mathbb{S}^1)$ 使得

$$H = 1_B L^2(\mathbb{S}^1, \mu) = \{f \in L^2(\mathbb{S}^1, \mu) : f|_{B^c} = 0\}.$$

证明. 设

$$1 = k + h, k \in H^\perp, h \in H.$$

因为H 为V 不变的, 那么$k \perp V^n h, \forall n \in \mathbb{Z}$, 即

$$\int_{\mathbb{S}^1} \overline{k}(z)h(z)z^n d\mu(z) = 0,\ \forall n \in \mathbb{Z}.$$

于是$\overline{k}h = 0, a.e.$. 根据$1 = k + h$, 我们有$1 = |k|^2 + |h|^2, a.e.$ 由此二式, 存在$B \in \mathcal{B}(\mathbb{S}^1)$ 使得$|k| = 1_{B^c}, |h| = 1_B$. 但是$1 = k + h$, 我们又得到$k = 1_{B^c}, h = 1_B$.

由$1_B \in H$, 我们得到$z^n 1_B(z) \in H, \forall n \in \mathbb{Z}$. 由此我们就有

$$1_B L^2(\mathbb{S}^1, \mu) \subseteq H.$$

同理

$$1_{B^c} L^2(\mathbb{S}^1, \mu) \subseteq H^\perp.$$

于是$H = 1_B L^2(\mathbb{S}^1, \mu)$. 证毕! □

§2.6.2 酉算子的谱分解定理

设$\mathcal{H}$ 为（可分）Hilbert空间, $U : \mathcal{H} \to \mathcal{H}$上的酉算子. 对于任何$x \in \mathcal{H}$, 因函数$\phi(n) = \langle U^n x, x\rangle$ 是正定的, 由Herglotz定理知存在$\mathbb{S}^1$ 上唯一的非负Borel测度σ_x 使得

$$\widehat{\sigma_x}(n) = \langle U^n x, x\rangle.$$

我们称σ_x为x(相对于U)的**谱测度**. 注意

$$\|\sigma_x\| = \sigma_x(\mathbb{S}^1) = \widehat{\sigma_x}(0) = \int_{\mathbb{S}^1} 1 d\sigma_x = \langle x, x\rangle = ||x||^2.$$

对于$x \in \mathcal{H}$, 我们用$Z(x)$表示由x生成的循环子空间, 即$Z(x)$ 是$\mathcal{H}$ 的包含$\{U^n x : n \in \mathbb{Z}\}$ 的最小的闭子空间.

定义 2.6.4 设$U_i:\mathcal{H}_i\to\mathcal{H}_i, i=1,2$ 为*Hilbert*空间上的酉算子, 如果存在等距满算子$W:\mathcal{H}_1\to\mathcal{H}_2$ 使得$WU_1=U_2W$,

$$\begin{tikzcd} \mathcal{H}_1 \arrow[r, "U_1"] \arrow[d, "W"'] & \mathcal{H}_1 \arrow[d, "W"] \\ \mathcal{H}_2 \arrow[r, "U_2"'] & \mathcal{H}_2. \end{tikzcd}$$

那么我们就称U_1,U_2 为**酉等价的***(unitarily equivalent)*, 记为$U_1\simeq U_2$.

根据命题2.5.7, 设U 为Hilbert空间$\mathcal{H}$ 上的酉算子和$x\in\mathcal{H}\setminus\{0\}$. 定义

$$V_x:L^2(\mathbb{S}^1,\sigma_x)\to L^2(\mathbb{S}^1,\sigma_x) \text{ 使得 } V_xf(z)=zf(z)\ (z\in\mathbb{S}^1),$$

则V_x 为酉算子且$W_x^{-1}V_xW_x=U$, 其中$W_x:Z(x)\to L^2(\mathbb{S}^1,\sigma_x)$ 为映射$U^nx\mapsto z^n\in L^2(\mathbb{S}^1,\sigma_x)$ 的唯一线性等距扩张.

$$\begin{tikzcd} Z(x) \arrow[r, "U"] \arrow[d, "W_x"'] & Z(x) \arrow[d, "W_x"] \\ L^2(\mathbb{S}^1,\sigma_x) \arrow[r, "V_x"'] & L^2(\mathbb{S}^1,\sigma_x). \end{tikzcd}$$

即

$$U|_{Z(x)}\simeq V_x.$$

在命题2.5.8 中，我们证明了：

- 设$x\in\mathcal{H}$, 如果$\mu\in\mathcal{M}(\mathbb{S}^1)$ 满足$\mu\ll\sigma_x$ 的有限非负Borel 测度, 那么存在$y\in Z(x)$ 使得$\mu=\sigma_y$.
- 设$y\in\mathcal{H}$, 如果$Z(y)\perp Z(x)$, 则$\sigma_{x+y}=\sigma_x+\sigma_y$.

下面我们给出更多性质.

命题 2.6.5 设U 为*Hilbert* 空间$\mathcal{H}$ 上酉算子.

1. 设$x,y\in\mathcal{H}$, 那么$U|_{Z(x)}\simeq U|_{Z(y)}$ 当且仅当$\sigma_x\sim\sigma_y$.
2. 设$x,y\in\mathcal{H}$, 如果$y\in Z(x)$, 那么$\sigma_y\ll\sigma_x$. 其中$\sigma_y\sim\sigma_x$ 当且仅当$Z(y)=Z(x)$ 成立.
3. 设$x,y\in\mathcal{H}$, 那么$\sigma_x\perp\sigma_y$ 蕴含$Z(x)\perp Z(y)$.
4. 设$x,y\in\mathcal{H}$, 那么$\sigma_x\perp\sigma_y$ 蕴含$\sigma_{x+y}=\sigma_x+\sigma_y$, 并且
$$Z(x+y)=Z(x)\bigoplus Z(y).$$
5. 设$x,y,z\in\mathcal{H}$. 如果$x,y\in Z(z)$, 那么$Z(x)\perp Z(y)$ 蕴含$\sigma_x\perp\sigma_y$.

6. 设U_i 为$\mathcal{H}_i$ 上酉算子（$i=1,2$）, $x_i\in\mathcal{H}_i$. 如果$U_1\simeq U_2$ 并且$U_1|_{Z(x_1)}\simeq U_2|_{Z(x_2)}$, 那么

$$U_1|_{Z(x_1)^\perp}\simeq U_2|_{Z(x_2)^\perp}.$$

证明. (1) 由$U|_{Z(x)}\simeq V_x$ 以及$U|_{Z(y)}\simeq V_y$. 我们仅需证明$V_x\simeq V_y$ 当且仅当$\sigma_x\sim\sigma_y$. 设

$$W:(L^2(\mathbb{S}^1,\sigma_x),V_x)\to(L^2(\mathbb{S}^1,\sigma_y),V_y)$$

为酉同构的等距算子. 令$f(z)=W(1)$, 那么$WV_x^n1=V_y^nf$, 即

$$W(z^n)=f(z)z^n,\ \forall n\in\mathbb{Z}.$$

由此我们得到

$$W(g)=fg,\ \forall g\in L^2(\mathbb{S}^1,\sigma_x).$$

尤其对于任何$A\in\mathcal{B}(\mathbb{S}^1)$, $W(1_A)=f1_A$. 因为W 等距, 所以

$$\sigma_x(A)=\int_A|f|^2d\sigma_y.$$

这样就有$\sigma_x\ll\sigma_y$. 同理$\sigma_y\ll\sigma_x$. 于是就有$\sigma_x\sim\sigma_y$.

反之, 设$\sigma_x\sim\sigma_y$. 令

$$W:(L^2(\mathbb{S}^1,\sigma_x),V_x)\to(L^2(\mathbb{S}^1,\sigma_y),V_y),g\mapsto g\sqrt{\frac{d\sigma_x}{d\sigma_y}}.$$

容易验证在W 下, $V_x\simeq V_y$.

(2) 设$W_x:(Z(x),U)\to(L^2(\mathbb{S}^1,\sigma_x),V_x)$ 为前面定义的酉等价, 即$U^nx\mapsto z^n$. 我们将问题转换到$(L^2(\mathbb{S}^1,\sigma_x),V_x)$ 上处理. 设$f=W_x(y)$. 因为$1=W_x(x)$, 于是命题等价于证明：$\sigma_f\ll\sigma_x$, 并且$\sigma_f\sim\sigma_x$ 当且仅当$Z(f)=Z(1)$ (在空间$L^2(\mathbb{S}^1,\sigma_x)$ 中). 因为

$$\int_{\mathbb{S}^1}z^nd\sigma_f=\widehat{\sigma_f}(n)=\langle V_x^nf,f\rangle=\int_{\mathbb{S}^1}z^n|f|^2d\sigma_x,$$

所以$d\sigma_f=|f|^2d\sigma_x\ll d\sigma_x$.

如果$Z(f)=Z(1)$, 那么根据（1）$\sigma_f\sim\sigma_x$. 如果$Z(f)$ 为$Z(1)=L^2(\mathbb{S}^1,\sigma_x)$ 的真子空间, 由它为V_x 不变的, 根据Wiener 定理存在$B\in\mathcal{B}(\mathbb{S}^1)$ 使得

$$Z(f)=1_BL^2(\mathbb{S}^1,\sigma_x),\ \text{且}\ \sigma_x(B)<\sigma_x(\mathbb{S}^1).$$

于是有$\sigma_x(\mathbb{S}^1\setminus B)>0$ 且$\sigma_f(\mathbb{S}^1\setminus B)=0$, 即$\sigma_f\not\sim\sigma_x$.

(3) 设$y=y_1+y_2$, 其中$y_2\in Z(x),y_1\perp Z(x)$. 根据$y_1\perp Z(x)$, 易见$Z(y_1)\perp Z(x)$. 由此易得

$$\langle U^ny,y\rangle=\langle U^ny_1,y_1\rangle+\langle U^ny_2,y_2\rangle,\ \forall n\in\mathbb{Z},$$

即
$$\int_{\mathbb{S}^1} z^n d\sigma_y = \int_{\mathbb{S}^1} z^n d\sigma_{y_1} + \int_{\mathbb{S}^1} z^n d\sigma_{y_2},\ \forall n \in \mathbb{Z}.$$
于是$\sigma_y = (\sigma_{y_1} + \sigma_{y_2}) \perp \sigma_x$. 由$y_2 \in Z(x)$ 以及(3), $\sigma_{y_2} \ll \sigma_x$. 于是$\sigma_{y_2} = 0$. 所以$y_2 = 0$, $y = y_1$ 且$Z(y) \perp Z(x)$.

(4) 根据(3), $Z(x) \perp Z(y)$. 于是
$$\langle U^n(x+y), (x+y)\rangle = \langle U^n x, x\rangle + \langle U^n y, y\rangle,\ \forall n \in \mathbb{Z}.$$
即$\int_{\mathbb{S}^1} z^n d\sigma_{x+y} = \int_{\mathbb{S}^1} z^n d\sigma_x + \int_{\mathbb{S}^1} z^n d\sigma_y, \forall n \in \mathbb{Z}$. 所以
$$\sigma_{x+y} = \sigma_x + \sigma_y.$$

下证明$Z(x+y) = Z(x) \bigoplus Z(y)$. 因为$\sigma_{x+y} = \sigma_x + \sigma_y$ 以及$\sigma_x \perp \sigma_y$, 我们有$\frac{d\sigma_x}{d\sigma_{x+y}} \in L^1(\mathbb{S}^1, \sigma_{x+y})$ 为特征函数，并且易见$\displaystyle\int_{\mathbb{S}^1} \frac{d\sigma_x}{d\sigma_{x+y}} d\sigma_x = \int_{\mathbb{S}^1} 1 d\sigma_x$.

对于任何$\varepsilon > 0$, 取多项式$p(z)$ 使得
$$\left\| \frac{d\sigma_x}{d\sigma_{x+y}} - p(z) \right\|_2^2 = \int_{\mathbb{S}^1} \left| \frac{d\sigma_x}{d\sigma_{x+y}} - p(z) \right|^2 d\sigma_{x+y} < \varepsilon.$$
注意
$$\begin{aligned}
&\int_{\mathbb{S}^1} \left| \frac{d\sigma_x}{d\sigma_{x+y}} - p(z) \right|^2 d\sigma_{x+y} \\
=&\int_{\mathbb{S}^1} (\frac{d\sigma_x}{d\sigma_{x+y}})^2 d\sigma_{x+y} - 2\mathrm{Re}\int_{\mathbb{S}^1} p(z)\frac{d\sigma_x}{d\sigma_{x+y}} d\sigma_{x+y} + \int_{\mathbb{S}^1} |p(z)|^2 d\sigma_{x+y} \\
=&\int_{\mathbb{S}^1} 1 d\sigma_x - 2\mathrm{Re}\int_{\mathbb{S}^1} p(z) d\sigma_x + \int_{\mathbb{S}^1} |p(z)|^2 d\sigma_{x+y}.
\end{aligned}$$
于是
$$\begin{aligned}
&\|x - p(U)(x+y)\|^2 \\
=&\langle x, x\rangle - 2\mathrm{Re}\langle x, p(U)(x+y)\rangle + \|p(U)(x+y)\|^2 \\
=&\int_{\mathbb{S}^1} 1 d\sigma_x - 2\mathrm{Re}\langle x, p(U)x\rangle + \int_{\mathbb{S}^1} |p(z)|^2 d\sigma_{x+y} \\
=&\int_{\mathbb{S}^1} 1 d\sigma_x - 2\mathrm{Re}\int_{\mathbb{S}^1} p(z) d\sigma_x + \int_{\mathbb{S}^1} |p(z)|^2 d\sigma_{x+y} \\
=&\int_{\mathbb{S}^1} \left| \frac{d\sigma_x}{d\sigma_{x+y}} - p(z) \right|^2 d\sigma_{x+y} < \varepsilon.
\end{aligned}$$
由于ε 任意，所以$x \in Z(x+y)$. 同理得到$y \in Z(x+y)$. 于是$Z(x) \bigoplus Z(y) \subseteq Z(x+y)$.

反之, 如果存在$v \in Z(x+y)$ 且$v \perp \Big(Z(x) \bigoplus Z(y)\Big)$, 那么对于任何$\varepsilon > 0$, 取多项式$p(z)$ 使得$\|v - p(U)(x+y)\|^2 < \varepsilon$. 根据$v \perp \Big(Z(x) \bigoplus Z(y)\Big)$, 上式化为
$$\|v\|^2 + \|p(U)(x+y)\|^2 < \varepsilon.$$

尤其$\|v\|^2 < \varepsilon$. 因为ε 任意, 所以$v = 0$. 所以$Z(x+y) = Z(x) \bigoplus Z(y)$.

(5) 设$W_z : (Z(z), U) \to (L^2(\mathbb{S}^1, \sigma_z), V_z)$ 为酉等价的同构, 设$f = W_z(x), g = W_z(y)$. 等价地, 我们需要证明：$Z(f) \perp Z(g)$, 那么$\sigma_f \perp \sigma_g$. 根据Wiener 定理, 存在$A, B \in \mathcal{B}(\mathbb{S}^1)$ 使得

$$Z(f) = 1_A L^2(\mathbb{S}^2, \sigma_z),\ Z(g) = 1_B L^2(\mathbb{S}^2, \sigma_z).$$

因为$Z(f) \perp Z(g)$, $\sigma_z(A \cap B) = 0$. 因为$d\sigma_f = |f|^2 d\sigma_z, d\sigma_g = |g|^2 d\sigma_z$, 我们就得到$\sigma_f \perp \sigma_g$.

(6) 因为酉等价为等价关系, 我们不妨设$\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$且$U_1 = U_2 = U$. 于是我们需要证明：设$x, y \in \mathcal{H}$, 如果$U|_{Z(x)} \simeq U|_{Z(y)}$, 那么$U|_{Z(x)^\perp} \simeq U|_{Z(y)^\perp}$

因为$U|_{\overline{Z(x)+Z(y)}} \simeq U|_{\overline{Z(x)+Z(y)}}$, 我们可以假设$\mathcal{H} = \overline{Z(x) + Z(y)}$. 取$y_0, y_1 \in Z(y)$ 使得

$$\sigma_y = \sigma_{y_0} + \sigma_{y_1}, \sigma_{y_0} \ll \sigma_x, \sigma_{y_1} \perp \sigma_x.$$

于是$Z(y) = Z(y_0) \bigoplus Z(y_1)$.

接着我们分解σ_x. 如果$\sigma_x = \mu + \nu$ 使得$\mu \ll \sigma_{y_0}$, $\nu \perp \sigma_{y_0}$，那么由$\sigma_{y_0} \ll \sigma_x$ 有$\sigma_{y_0} \ll \mu$, 继而$\mu \sim \sigma_{y_0}$. 这就意味着存在$\phi \in L^2(\sigma_{y_0})$ 使得$\mu = |\phi|^2 \sigma_{y_0} = \sigma_{\phi(U)y_0}$. 令$y_0' = \phi(U)y_0$. 那么我们有$\sigma_{y_0'} \ll \sigma_x$ 并且$Z(y_0') = Z(y_0)$. 由此分析，我们可以取$y_0 \in Z(y)$ 使得存在$x_0 \in Z(x)$ 满足

$$\sigma_x = \sigma_{y_0} + \sigma_{x_0}, \sigma_{x_0} \perp \sigma_{y_0},$$

并且$Z(x) = Z(y_0) \bigoplus Z(x_0)$. 于是我们得到分解

$$\mathcal{H} = Z(x_0) \bigoplus Z(y_0) \bigoplus Z(y_1).$$

由假设以及(1)我们有$U|_{Z(x)} \simeq U|_{Z(y)} \Leftrightarrow \sigma_x \sim \sigma_y$. 根据构造

$$\sigma_x \sim \sigma_y \Leftrightarrow \sigma_{x_0} \sim \sigma_{y_1}.$$

根据(1), 这意味着

$$U|_{Z(x_0)} \simeq U|_{Z(y_1)},$$

此即为所需证明的. □

注记 2.6.6 $Z(x) \perp Z(y)$ 能蕴含$\sigma_{x+y} = \sigma_x + \sigma_y$, 但是推不出$\sigma_x \perp \sigma_y$.

定义 2.6.7 设$\mu \in \mathcal{M}(\mathbb{S}^1)$, 称$\{\nu \in \mathcal{M}(\mathbb{S}^1) : \nu \sim \mu\}$ 为μ 的**型** *(type of μ)*.

定理 2.6.8 (谱分解定理形式I) 设U 为*Hilbert* 空间$\mathcal{H}$ 上的酉算子, 那么存在$\{x_n\}_{n=1}^\infty \subseteq \mathcal{H}$ 使得

1. $\mathcal{H} = \bigoplus_{n=1}^\infty Z(x_n)$, 并且$Z(x_i) \perp Z(x_j), \forall i \neq j \in \mathbb{N}$;

2. $\sigma_{x_1} \gg \sigma_{x_2} \gg \sigma_{x_3} \gg \dots$.

对于任何满足条件的另外一组$\{x'_n\}_{n=1}^{\infty}$, 我们有$\sigma_{x_n} \sim \sigma_{x'_n}, \forall n \in \mathbb{N}$.

注记 2.6.9 在定理表述中, 为了简单起见我们直接使用可数序列$\{x_n\}_{n=1}^{\infty}$, 并且要求$x_i \neq x_j, i \neq j \in \mathbb{N}$. 事实上, 可能出现序列是有限的, 此时我们不再引入新的记号, 直接按照有限序列理解定理.

证明.一个循环子空间$Z(x)$ 称为最大的, 是指它不真包含在更大的循环子空间中. 易见

(1) 根据命题2.6.5-(5), $Z(x)$ 为最大的当且仅当对于任何$y \in \mathcal{H}$ 成立$\sigma_y \ll \sigma_x$.

(2) 根据Zorn 引理, 对于任何$x \in \mathcal{H}$ 存在一个最大循环子空间包含它.

下面开始证明定理. 首先取$\mathcal{H}$ 的可数集稠密子集$\{y_n\}_{n\in\mathbb{N}}$. 取最大循环子空间$Z(x_1)$ 包含y_1. 如果$\mathcal{H} = Z(x_1)$, 则证明结束, 否则设

$$n_2 = \min\{n \in \mathbb{N} : y_n \notin Z(x_1)\} \geq 2.$$

设$y_{n_2}^{\perp}$ 为y_{n_1} 在$Z(x_1)^{\perp}$ 上投射, 取$(Z(x_1)^{\perp}, U|_{Z(x_1)^{\perp}})$ 中包含$y_{n_2}^{\perp}$ 的最大循环子空间$Z(x_2)$. 则

$$y_1, \dots, y_{n_2} \in Z(x_1) \bigoplus Z(x_2),\ \sigma_{x_1} \gg \sigma_{x_2}.$$

如果$\mathcal{H} = Z(x_1) \bigoplus Z(x_2)$ 则证明结束, 否则取

$$n_3 = \min\{n \in \mathbb{N} : y_n \notin Z(x_1) \bigoplus Z(x_2)\} \geq 3.$$

设$y_{n_3}^{\perp}$ 为y_{n_3} 在$\big(Z(x_1) \bigoplus Z(x_2)\big)^{\perp}$ 上投射, 取$\Big(\big(Z(x_1) \bigoplus Z(x_2)\big)^{\perp}, U|_{\left(Z(x_1) \bigoplus Z(x_2)\right)^{\perp}}\Big)$ 中包含$y_{n_3}^{\perp}$ 的最大循环子空间$Z(x_3)$. 于是

$$y_1, \dots, y_{n_3} \in Z(x_1) \bigoplus Z(x_2) \bigoplus Z(x_3),\ \sigma_{x_1} \gg \sigma_{x_2} \gg \sigma_{x_3}.$$

归纳的, 设$x_1, x_2, \dots, x_k$, $n_k \geq k$ 已经确定使得

$$y_1, \dots, y_{n_k} \in Z(x_1) \bigoplus \dots \bigoplus Z(x_k),\ \sigma_{x_1} \gg \sigma_{x_2} \gg \dots \gg \sigma_{x_k}.$$

如果$\mathcal{H} = Z(x_1) \bigoplus \dots \bigoplus Z(x_k)$ 则证明结束, 否则取

$$n_{k+1} = \min\{n \in \mathbb{N} : y_n \notin Z(x_1) \bigoplus \dots \bigoplus Z(x_k)\} \geq k+1.$$

设$y_{n_{k+1}}^{\perp}$ 为$y_{n_{k+1}}$ 在$\big(Z(x_1) \bigoplus \dots \bigoplus Z(x_k)\big)^{\perp}$ 上投射, 取$\Big(\big(\bigoplus_{j=1}^k Z(x_j)\big)^{\perp}, U|_{\left(\bigoplus_{j=1}^k Z(x_j)\right)^{\perp}}\Big)$ 中包含$y_{n_{k+1}}^{\perp}$ 的最大循环子空间$Z(x_{k+1})$. 于是

$$y_1, \dots, y_{n_{k+1}} \in Z(x_1) \bigoplus \dots \bigoplus Z(x_{k+1}),\ \sigma_{x_1} \gg \sigma_{x_2} \gg \dots \gg \sigma_{x_{k+1}}.$$

根据归纳法, 我们得到序列$\{x_n\}_{n=1}^{\infty}$ (有可能只有有限项), 使得

$$\{y_n\}_{n=1}^{\infty} \subseteq \bigoplus_{n=1}^{\infty} Z(x_n), \quad \sigma_{x_1} \gg \sigma_{x_2} \gg \sigma_{x_3} \gg \dots,$$

并且$Z(x_i) \perp Z(x_j), \forall i \neq j \in \mathbb{N}$. 因为$\{y_n\}_{n=1}^{\infty}$ 为$\mathcal{H}$ 稠密子集, 所以$\mathcal{H} = \bigoplus_{n=1}^{\infty} Z(x_n)$.

下证唯一性. 设$\{x_n'\}_{n=1}^{\infty}$ 为满足条件的另外一组向量. 因为$Z(x_1)$ 与$Z(x_2)$ 都是最大循环子空间, 所以$\sigma_{x_1} \sim \sigma_{x_1'}$. 于是根据命题2.6.5-(1), $U|_{Z(x_1)} \simeq U|_{Z(x_1')}$. 根据命题2.6.5-(7), $U|_{Z(x_1)^{\perp}} \simeq U|_{Z(x_1')^{\perp}}$. 对于$(Z(x_1)^{\perp}, U|_{Z(x_1)^{\perp}})$ 和$(Z(x_1')^{\perp}, U|_{Z(x_1')^{\perp}})$ 重复上面论证, 可以得到$\sigma_{x_2} \sim \sigma_{x_2'}$, 以及$U|_{(Z(x_1)\oplus Z(x_2))^{\perp}} \simeq U|_{(Z(x_1')\oplus Z(x_2'))^{\perp}}$. 归纳地就可以证明$\{x_n\}$ 与$\{x_n'\}$ 个数一致, 且$\sigma_{x_n} \sim \sigma_{x_n'}, \forall n \in \mathbb{N}$. □

注记 2.6.10 根据定理*2.6.8*, $(\mathcal{H}, U)$ 酉等价于$(\bigoplus_{n=1}^{\infty} L^2(\mathbb{S}^1, \sigma_{x_n}), V)$:

$$V(f_1, f_2, \ldots)(z_1, z_2, \ldots) = (z_1 f_1(z_1), z_2 f_2(z_2), \ldots),\ \forall (f_n)_n \in \bigoplus_{n=1}^{\infty} L^2(\mathbb{S}^1, \sigma_{x_n}).$$

定义 2.6.11 设U 为*Hilbert* 空间上的酉算子, 定理*2.6.8* 中σ_{x_1} 的谱型$[\sigma_{x_1}]$ 称为算子U 的**最大谱型** (*maximal spectral type*), 也把它记为$[\sigma_U]$. $\{[\sigma_{x_n}]\}_{n=1}^{\infty}$ 称为U 的**谱序列** *(spectral sequence)*.

设$\{[\sigma_{x_n}]\}_{n=1}^{\infty}$ 为U 的谱序列, 令

$$A_n = \operatorname{supp} \frac{d\sigma_{x_n}}{d\sigma_U}.$$

那么我们得到Borel集序列:

$$A_1 \supset A_2 \supset \ldots.$$

令

$$M_U = \sum_{n=1}^{\infty} 1_{A_n} : \mathbb{S}^1 \to \mathbb{N} \cup \{\infty\},$$

称之为**重数函数** (multiplicity function).

定理 2.6.12 在酉等价意义下, $U : \mathcal{H} \to \mathcal{H}$ 由它的最大谱型$[\sigma_U]$ 和重数函数M_U 唯一决定.

证明. 已知σ_U, M_U, 令

$$A_n = \{x \in \mathbb{S}^1 : M_U(x) \geq n\}.$$

则

$$\mu_n = \int_{\mathbb{S}^1} 1_{A_n} d\sigma_U \sim \sigma_{x_n}.$$

于是$(\mathcal{H}, U)$ 酉等价于$(\bigoplus_{n=1}^{\infty} L^2(\mathbb{S}^1, \mu_n), V)$, V 的定义见注记2.6.10. □

定义 2.6.13 称算子U 为**离散/ 连续/ 奇异/ 绝对连续/ Lebesgue 的**, 是指σ_U 为离散/ 连续/ $\sigma_U \perp m$ / $\sigma_U \ll m$ / $\sigma_U \sim m$, 其中m 为*Lebsgue* 测度.

定义 2.6.14 设U 为酉算子, 称$\operatorname{esssup} M_U$ 为U 的**谱重数**(*spectral multiplicity of* U). 如果谱重数有限, 那么称U 具有**有限谱重数**; 如果存在$N \in \mathbb{N} \cup \{\infty\}$ 使得$M_U \equiv N$, 那么称U 具有N **重齐性谱**(*homogeneous spectrum of multiplicity* N). 特别地, 如果$M_U \equiv 1$, 那么称U 具有**简单谱** (*simple spectrum*).

U 具有**可数Lebesgue 谱**是指它具有无穷重齐性谱, 并且$\sigma_U \sim m$.

具有有限谱重数当且仅当$\mathcal{H} = \bigoplus_{n=1}^N Z(x_n)$; 具有$N$ 重齐性谱当且仅当

$$\mathcal{H} = \bigoplus_{n=1}^N Z(x_n), \sigma_{x_1} \sim \ldots \sim \sigma_{x_N}.$$

定理 2.6.15 (谱分解定理形式II) 设U 为*Hilbert* 空间$\mathcal{H}$ 上的酉算子, 那么存在$B_n \in \mathcal{B}(\mathbb{S}^1)$ 以及序列$x_n^{(k)} \in \mathcal{H}$, $k = 1, 2, \ldots, n, n \in \mathbb{N} \cup \{\infty\}$, 使得

1. $\{B_n\}_{n\in\mathbb{N}}$ 为$\mathbb{S}^1$ 的剖分;
2. $\mathcal{H} = \bigoplus_{n=1}^\infty \bigoplus_{k=1}^n Z(x_n^{(k)})$, 其中$Z(x_n^{(k)}) \perp Z(x_{n'}^{k'}), (n, k) \neq (n', k')$;
3. $\sigma_{x_n^{(1)}} = \sigma_{x_n^{(2)}} = \ldots = \sigma_{x_n^{(n)}} := \sigma^{(n)}$, 并且$\sigma^{(n)}(\mathbb{S}^1 \setminus B_n) = 0$;
4. $\sum_{n=1}^\infty \|\sigma^{(n)}\| < \infty$.

对于任何满足条件的另外一组$\{y_n^{(k)}\}$, 我们有$\sigma_{x_n^{(k)}} \sim \sigma_{y_n^{(k)}}, \forall k = 1, 2, \ldots, n, n \in \mathbb{N}$.

证明. 如果定理成立, 令

$$x_n = \sum_{k=n}^\infty x_n^{(k)}.$$

那么就得到定理2.6.8.

反之, 设定理2.6.8 成立. 设$A_n = \operatorname{supp} \frac{d\sigma_{x_n}}{d\sigma_U}$ 则存在$f_n \in \mathcal{H}$ 使得

$$\mathcal{H} = \bigoplus_{n=1}^\infty Z(f_n), \quad \sigma_{f_n} = 1_{A_n} \sigma_U.$$

令

$$B_n = A_n \setminus A_{n+1}, \quad B_\infty = \bigcap_{n=1}^\infty A_n.$$

则$\{B_n\}_{n\in\mathbb{N}\cup\{\infty\}}$ 为$\mathbb{S}^1$ 的剖分.

设$n \in \mathbb{N} \cup \{\infty\}$. 对于$1 \le k \le n$ 取$x_n^{(k)} \in Z(f_k)$ 使得$\sigma_{x_n^{(k)}} = 1_{B_n} \sigma_{f_k}$. 因为$A_k \supset A_n$,

$$\sigma_{x_n^{(k)}} = 1_{A_n \setminus A_{n+1}} 1_{A_k} \sigma_U = 1_{A_n \setminus A_{n+1}} \sigma_U,$$

并且$\sigma_{x_n^{(k)}}$ 不依赖于k. 从而设$\sigma^{(n)} = \sigma_{x_n^{(k)}}, 1 \le k \le n$.

根据构造, 对于$k \in \mathbb{N}$
$$Z(x_k) = \bigoplus_{n \geq k} Z(x_n^{(k)}).$$
于是
$$\mathcal{H} = \bigoplus_{n=1}^{\infty} Z(f_n) = \bigoplus_{n=1}^{\infty} \bigoplus_{k=1}^{n} Z(x_n^{(k)}).$$
证毕! □

根据定理2.6.15, 重数函数为: 当$t \in B_n$, 则$M(t) = n$. U 具有有限谱重数当且仅当存在$N \in \mathbb{N}$ 使得$\sigma^{(n)} = 0, \forall n \geq N$; U 具有N 重齐性谱当且仅当$\sigma^{(n)} = 0, \forall n \neq N$.

§2.6.3 动力系统的谱

设$(X, \mathcal{X}, \mu, T)$ 为可逆保测系统, 那么$U_T : L^2(X, \mu) \to L^2(X, \mu)$ 为酉算子. 因为常值函数总是U_T 关于1 的特征函数, 所以U_T 总有离散谱部分, 为此一般讨论U_T 的谱性质主要针对下面空间
$$L_0^2(X, \mu) := \mathbb{C}^{\perp} = \{f \in L^2(X, \mu) : \langle f, 1 \rangle = \int_X f d\mu = 0\}.$$

定理 2.6.16 *1. $(X, \mathcal{X}, \mu, T)$ 为遍历的当且仅当1 为U_T 简单特征值, 当且仅当$\sigma_f(\{0\}) = 0, \forall f \in L_0^2(X, \mu)$, 即$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \hat{\sigma}_f(n) = 0, \forall f \in L_0^2(X, \mu)$.*

2. $(X, \mathcal{X}, \mu, T)$ 为弱混合的当且仅当$\sigma_f \in \mathcal{M}_c(\mathbb{S}^1), \forall f \in L_0^2(X, \mu)$, 即$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\hat{\sigma_f}(n)| = 0, \forall f \in L_0^2(X, \mu)$.

3. $(X, \mathcal{X}, \mu, T)$ 为强混合的当且仅当$\lim_{n\to\infty} |\hat{\sigma_f}(n)| = 0, \forall f \in L_0^2(X, \mu)$.

一个著名的问题是:

问题 2.6.17 (Banach) *是否存在具有简单Lebesgue谱的遍历系统?*

§2.7 注记

本章介绍遍历论中最基本的概念和性质, 在各种专著中都可以找到相关内容. 对于遍历性、混合性等内容我们采取了[195] 中的处理方式. Birkhoff 定理的简短证明来自[1], 原始证明参见[22, 23], 更多关于Birkhoff 定理的证明参见[54, 125, 136]等. 一般群作用下Birkhoff 定理参见Lindenstrauss的工作[148]. Mild 混合是由Furstenberg 和Weiss 引入的, 关于这个性质的更多性质参见[64, 65, 70, 76]. 关于谱理论参见[82, 145, 161, 156, 165] 等. 更多关于$\mathcal{M}(\mathbb{S}^1)$ 的性质参见[124, 165]等.

第三章　连分数简介

在这一章中, 我们介绍连分数, 希望通过这个具体例子更好地解释遍历论如何运用到其它数学分支中. 在本章中为了简单起, **我们仅考虑非负实数**.

§3.1　连分数基本概念

在这一节我们先介绍连分数的基本概念, 在这一节不涉及其上的动力系统性质.

§3.1.1　连分数的定义

如下形式的数称为**连分数**(continued fraction):

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}} \tag{3.1.1}$$

我们将(3.1.1) 记为

$$[a_0; a_1, a_2, a_3, \dots],$$

其中$a_n \in \mathbb{N}, \forall n \geq 1$, $a_0 \in \mathbb{Z}_+$. 对于有限的情况, 我们用

$$[a_0; a_1, a_2, \dots, a_n]$$

来记

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}.$$

于是, 我们有

$$[a_0; a_1, a_2, \dots, a_n] = a_0 + \frac{1}{[a_1; a_2, \dots, a_n]}.$$

注意如果$a_n \geq 2$, 那么

$$[a_0; a_1, a_2, \dots, a_n] = [a_0; a_1, a_2, \dots, a_n - 1, 1].$$

在本节中, 我们主要要说明非负实数总是可以连分数展开的. 尤其要证明：非负实数为有理数当且仅当它的连分数展开是有限的；实数为无理数当且仅当它有(3.1.1) 的展开. 在本节中, 我们还要说明连分数展开提供了很好的有理数逼近无理数的方法.

§3.1.2　基本性质

首先我们有下面重要的命题：

命题 3.1.1 设$\{a_n\}_{n\geq 0}$ 为满足$a_0 \in \mathbb{Z}_+, a_n \in \mathbb{N}, \forall n \geq 1$. 设

$$\frac{p_n}{q_n} = [a_0; a_1, a_2, \ldots, a_n],\ n \geq 0, \tag{3.1.2}$$

其中$p_n \geq 1, q_n \geq 1, (p_n, q_n) = 1$ 可以根据下式归纳得到：

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}, \forall n \geq 0. \tag{3.1.3}$$

我们约定$p_{-1} = 1, q_{-1} = 0, p_0 = a_0, q_0 = 1$.

一般称$\dfrac{p_n}{q_n}$ 为**第n 个渐近分数** 或**第n 个渐近值**

证明. 对n进行归纳证明. $n = 0$ 是显然的. 假设(3.1.3) 对于$0 \leq n \leq k-1$ 和任何序列$\{a_j\}_{j\geq 0}$ 都成立. 对于序列$\{a_j\}_{j\geq 1}$ 用归纳假设$k-1$ 的情况, 我们有

$$\frac{x}{y} = [a_1; a_2, a_3, \ldots, a_k],$$

以及

$$\begin{pmatrix} x & x' \\ y & y' \end{pmatrix} = \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix}.$$

于是

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x & x' \\ y & y' \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_k & p_{k-1} \\ q_k & q_{k-1} \end{pmatrix}.$$

即

$$\begin{pmatrix} a_0x + y & a_0x' + y' \\ x & x' \end{pmatrix} = \begin{pmatrix} p_k & p_{k-1} \\ q_k & q_{k-1} \end{pmatrix}.$$

于是

$$\frac{p_k}{q_k} = \frac{a_0x + y}{x} = a_0 + \frac{y}{x} = a_0 + \frac{1}{[a_1; a_2, \ldots, a_k]} = [a_0; a_1, a_2, \ldots, a_k].$$

所以结论对于$n = k$ 成立. 根据数学归纳法, 结论证毕. □

根据命题3.1.1,

$$\begin{pmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \begin{pmatrix} a_{n+1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_{n+1}p_n + p_{n-1} & p_n \\ a_{n+1}q_n + q_{n-1} & q_n \end{pmatrix},$$

我们得到

$$\begin{aligned} p_{n+1} &= a_{n+1}p_n + p_{n-1} \\ q_{n+1} &= a_{n+1}q_n + q_{n-1}, \forall n \geq 1. \end{aligned} \tag{3.1.4}$$

因为$a_n \geq 1, \forall n \geq 1$, 所以

$$1 = q_0 \leq q_1 < q_2 < q_3 < \dots. \tag{3.1.5}$$

由此我们得到

$$q_n \geq 2^{\frac{n-2}{2}},\ \forall n \geq 1. \tag{3.1.6}$$

类似得到

$$p_n \geq 2^{\frac{n-2}{2}},\ \forall n \geq 1. \tag{3.1.7}$$

在(3.1.3)中行列式, 就有

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1}. \tag{3.1.8}$$

于是得到

$$\frac{p_1}{q_1} = a_0 + \frac{1}{q_0 q_1},$$

$$\frac{p_2}{q_2} = \frac{p_1}{q_1} - \frac{1}{q_1 q_2} = a_0 + \frac{1}{q_0 q_1} - \frac{1}{q_1 q_2},$$

以及

$$\begin{aligned} \frac{p_n}{q_n} &= \frac{p_{n-1}}{q_{n-1}} + (-1)^{n+1}\frac{1}{q_{n-1}q_n} \\ &= a_0 + \frac{1}{q_0 q_1} - \frac{1}{q_1 q_2} + \dots + (-1)^{n+1}\frac{1}{q_{n-1}q_n},\ \forall n \geq 1. \end{aligned} \tag{3.1.9}$$

根据(3.1.9)以及(3.1.6) 我们得到连分数总是绝对收敛到一个实数的：

$$\begin{aligned} u = [a_0; a_1, a_2, \dots] &= \lim_{n\to\infty}[a_0; a_1, a_2, \dots, a_n] \\ &= \lim_{n\to\infty}\frac{p_n}{q_n} \\ &= a_0 + \sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{q_{n-1}q_n},\ \forall n \geq 1. \end{aligned} \tag{3.1.10}$$

根据(3.1.5), 我们观察到:

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \dots < \frac{p_{2n}}{q_{2n}} < \dots < u < \dots < \frac{p_{2n+1}}{q_{2n+1}} < \dots < \frac{p_3}{q_3} < \frac{p_1}{q_1}. \tag{3.1.11}$$

定义 3.1.2 我们称$[a_0; a_1, a_2, \dots]$ 为u 的**连分数展开**.

我们马上会说明任何实数都可以连分数展开, 对于无理数连分数展开是唯一的.

定理 3.1.3 如果$a_n \in \mathbb{N}, \forall n \geq 1, a_0 \in \mathbb{Z}_+$, 那么

$$u = [a_0; a_1, a_2, \ldots] = \lim_{n\to\infty} [a_0; a_1, a_2, \ldots, a_n] = \lim_{n\to\infty} \frac{p_n}{q_n}$$

为无理数.

证明. 否则, 设$u = \frac{a}{b} \in \mathbb{Q}$. 根据(3.1.14),

$$|q_n a - b p_n| < \frac{b}{a_{n+1} q_n} \leq \frac{b}{q_n}.$$

因为$q_n \to \infty, n \to \infty$, 以及$q_n a - b p_n \in \mathbb{Z}$, 所以对于充分大的$n$ 有

$$q_n a - b p_n = 0.$$

即对于充分大的n,

$$u = \frac{a}{b} = \frac{p_n}{q_n}.$$

但是根据引理3.1.1, $(p_n, q_n) = 1$, 以及$\lim_{n\to\infty} q_n = \infty$, 这是不可能的.

所以$u \notin \mathbb{Q}$. □

命题 3.1.4 任何非负无理数u 具有连分数展开$[a_0; a_1, a_2, \ldots]$, 其中$a_0 \in \mathbb{Z}_+, a_n \in \mathbb{N}, \forall n \in \mathbb{N}$. 并且映射

$$\phi : \mathbb{Z}_+ \times \mathbb{N}^{\mathbb{N}} \to \mathbb{R},\ (a_0, a_1, \ldots) \mapsto u = [a_0; a_1, a_2, \ldots]$$

是单射，尤其非负无理数连分数展开唯一.

证明. 设$(a_0, a_1, \ldots) \in \mathbb{Z}_+ \times \mathbb{N}^{\mathbb{N}}$, 那么$u = [a_0; a_1, a_2, \ldots] > 0$. 因为

$$u = a_0 + \frac{1}{[a_1; a_2, a_3, \ldots]},$$

我们得到

$$u \in (a_0, a_0 + \frac{1}{a_1}) \subseteq (a_0, a_0 + 1).$$

于是a_0 由u 唯一决定. 令$u_1 = \dfrac{1}{u - a_0}$, 则

$$u_1 = a_1 + \frac{1}{[a_2; a_3, a_4, \ldots]}.$$

所以

$$u_1 \in (a_1, a_1 + \frac{1}{a_2}) \subseteq (a_1, a_1 + 1).$$

于是a_1 由$u_1 = \dfrac{1}{u - a_0}$, 即由u 唯一决定. 根据此归纳下去, 我们得到$a_0, a_1, \ldots$ 由u 唯一决定, 即映射是单射.

根据上面的做法，对于任何非负无理数u 均可以连分数展开. 我们在下一小节引理3.2.2 中将给出更为严格的证明. □

注记 3.1.5 *1.* 根据命题*3.1.4*的证明, 实际提供了一种方法去给出一个数u 的连分数展开. 即首先取u 的整数部分$[u]$, 即为a_0; 再取$a_1 = \left[\frac{1}{\{u\}}\right]$, 其中$[\cdot]$ 表示取整, $\{\cdot\}$ 表示取小数部分. 接着往下做就可以得到$a_2, a_3, \ldots$. 虽然命题*3.1.4*是对非负实数进行陈述, 但是根据其证明可以知道任何实数可以连分数展开.

2. 设u 为有理数, 那么它的连分数必为有限项的, 否则与定理*3.1.3* 矛盾. 于是一个数是有理数当且仅当连分数展开是有限项的; 一个数为无理数当且仅当连分数展开是无限的. 注意有理数的连分数展开不一定为唯一的; 而命题*3.1.4* 告诉我们无理数的连分数展开是唯一的.

例 3.1.6 设
$$u = \frac{\sqrt{5}-1}{2}.$$
注意到
$$\frac{2}{\sqrt{5}-1} = \frac{\sqrt{5}+1}{2} \in (1,2),$$
$$\frac{\sqrt{5}+1}{2} - 1 = \frac{\sqrt{5}-1}{2}.$$
于是如果设
$$u = \frac{\sqrt{5}-1}{2} = [a_0; a_1, a_2, \ldots]$$
则$a_0 = 0$, 并且由$\dfrac{2}{\sqrt{5}-1} = \dfrac{\sqrt{5}+1}{2} \in (1,2)$, 知$a_1 = 1$.

于是
$$[0; a_2, a_3, \ldots] = \frac{\sqrt{5}+1}{2} - 1 = \frac{\sqrt{5}-1}{2} = [0; a_1, a_2, \ldots].$$
于是$1 = a_1 = a_2 = \ldots$, 即
$$u = \frac{\sqrt{5}-1}{2} = [0; 1, 1, 1, \ldots].$$

对于无理数$u = [a_0; a_1, a_2, \ldots]$, 有理数$\frac{p_n}{q_n}$ 提供了一个快速的有理逼近:
$$u - \frac{p_n}{q_n} = (-1)^n \left[\frac{1}{q_n q_{n+1}} - \frac{1}{q_{n+1} q_{n+2}} + \ldots\right], \tag{3.1.12}$$
于是结合(3.1.5),
$$\left| u - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}. \tag{3.1.13}$$
由(3.1.4),
$$\left| u - \frac{p_n}{q_n} \right| < \frac{1}{a_{n+1} q_n^2} \leq \frac{1}{q_n^2}. \tag{3.1.14}$$

对于$t \in \mathbb{R}$, 我们记
$$\langle t \rangle = \min_{n \in \mathbb{Z}} |t - n|.$$
根据(3.1.14), 我们得到:

命题 3.1.7 对于任何$u \in \mathbb{R}$, 存在序列$\{q_n\}_{n\in\mathbb{N}}$ 使得$q_n \to \infty$, 并且

$$q_n\langle q_n u\rangle < 1.$$

一般而言我们没有

$$\liminf_{n\to\infty} n\langle nu\rangle = 0.$$

问题 3.1.8 (Littlewood猜测) 对于任何$u, v \in \mathbb{R}$,

$$\liminf_{n\to\infty} n\langle nu\rangle\langle nv\rangle = 0.$$

注记 3.1.9 关于*Littlewood*猜测最好的结果是*2003*年*Einsiedler, Katok* 和*Lindenstrauss* 给出的:

$$\Theta = \{(u,v) \in \mathbb{R}^2 : \liminf_{n\to\infty} n\langle nu\rangle\langle nv\rangle > 0\}$$

的*Hausdorff* 维数为*0*.

下面命题指出$\left\{\dfrac{p_n}{q_n}\right\}_{n\in\mathbb{N}}$ 给出了u 很好的逼近.

命题 3.1.10 设$a_n \in \mathbb{N}, \forall n \geq 0$, $u = [a_0; a_1, a_2, \ldots]$. 对于任何$n > 1$ 以及$p, q \in \mathbb{Z}$, $0 < q \leq q_n$, 如果$\dfrac{p}{q} \neq \dfrac{p_n}{q_n}$, 那么

$$|p_n - q_n u| < |p - qu|. \tag{3.1.15}$$

尤其

$$\left|\frac{p_n}{q_n} - u\right| < \left|\frac{p}{q} - u\right|. \tag{3.1.16}$$

即分母不大于q_n 的所有分数中, $\frac{p_n}{q_n}$ 与u 最接近.

证明. 首先如果已证(3.1.15), 那么就有

$$\frac{1}{q}\left|\frac{p_n}{q_n} - u\right| < \frac{1}{q_n}\left|\frac{p}{q} - u\right| \leq \frac{1}{q}\left|\frac{p}{q} - u\right|,$$

就得到(3.1.16). 于是仅需要证明(3.1.15).

根据(3.1.13),

$$\left|u - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}},$$

以及

$$\left|u - \frac{p_{n+1}}{q_{n+1}}\right| < \frac{1}{q_{n+1}q_{n+2}}.$$

因为(3.1.11), 下面小括号中的值都是正的, 或者都是负的:

$$\left(u - \frac{p_n}{q_n}\right) = \left(\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right) - \left(\frac{p_{n+1}}{q_{n+1}} - u\right),$$

所以有

$$\left|u-\frac{p_n}{q_n}\right|=\left|\frac{p_{n+1}}{q_{n+1}}-\frac{p_n}{q_n}\right|-\left|\frac{p_{n+1}}{q_{n+1}}-u\right|.$$

由(3.1.4)(3.1.14),

$$\left|u-\frac{p_n}{q_n}\right|>\frac{1}{q_nq_{n+1}}-\frac{1}{q_{n+1}q_{n+2}}=\frac{q_{n+2}-q_n}{q_nq_{n+1}q_{n+2}}=\frac{a_{n+2}}{q_nq_{n+2}}.$$

我们就有

$$\frac{1}{q_{n+2}}<|p_n-q_nu|<\frac{1}{q_{n+1}},\ \forall n\geq 1. \tag{3.1.17}$$

根据上式,

$$|p_n-q_nu|<\frac{1}{q_{n+1}}<|q_{n-1}u-p_{n-1}|,$$

我们可以不妨设$q_{n-1}<q\leq q_n$.

如果$q=q_n$, 那么$\left|\dfrac{p_n}{q_n}-\dfrac{p}{q}\right|\geq\frac{1}{q_n}$, 以及

$$\left|u-\frac{p_n}{q_n}\right|<\frac{1}{q_nq_{n+1}}\leq\frac{1}{2q_n},\ (q_n\geq 2,n\geq 2).$$

由此

$$\left|u-\frac{p}{q}\right|\geq\frac{1}{2q_n}=\frac{1}{2q},$$

以及

$$|q_nu-p_n|<\frac{1}{2}\leq|qu-p|.$$

现在设$q_{n-1}<q<q_n$, 记

$$\begin{pmatrix}p_n & p_{n-1}\\ q_n & q_{n-1}\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}=\begin{pmatrix}p\\ q\end{pmatrix}.$$

由(3.1.8), $a,b\in\mathbb{Z}$. 明显$ab\neq 0$, 否则$q=q_{n-1}$ 或者$q=q_n$. 现在$q=aq_n+bq_{n-1}<q_n$, 所以$ab<0$. 根据(3.1.11), p_n-q_nu 和$p_{n-1}-q_{n-1}u$ 符号相反, 所以$a(p_n-q_nu)$ 和$b(p_{n-1}-q_{n-1}u)$ 符号相同. 这样式子

$$p-qu=a(p_n-q_nu)+b(p_{n-1}-q_{n-1}u)$$

蕴含

$$|p-qu|>|p_{n-1}-q_{n-1}u|>|p_n-q_nu|.$$

证毕! □

例如

$$\pi=[3;7,15,1,292,1,1,1,21,31,14,2,1,2,2,2,2,1,84,2,1,1,15,\ldots]$$

的渐近分数分别为

$$\frac{3}{1}, \frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \frac{103993}{33102}, \frac{104348}{33215}, \ldots.$$

祖冲之得到的渐近值为$\frac{22}{7}, \frac{355}{113}$ 都在上面最佳渐近数之列.

习　题

1. 设$u = [0; a_1, a_2, \ldots] \in (0,1), n \in \mathbb{N}$. 令

$$u_n = \frac{1}{[0; a_n, a_{n+1}, \ldots]}, \quad s_n = \frac{q_{n-1}}{q_n}.$$

证明:

(a) $s_n = [0; a_n, a_{n-1}, \ldots, a_1]$;

(b) $u = \dfrac{p_n u_{n+1} + p_{n-1}}{q_n u_{n+1} + q_{n-1}}$.

2. 满足$b_1 = b_2 = 1, b_{n+1} = b_{n-1} + b_n, n \geq 2$ 的数列称为Fibonacci 数列. 证明：

(a) $\frac{1+\sqrt{5}}{2}$ 的第n 个渐近分数为$\frac{b_{n+2}}{b_{n+1}}$.

(b) 设连分数$[a_0; a_1, a_2, \ldots]$, $i \in N$. 如果$a_i = 2, a_n = 1, n \neq i$, 那么证明：$m > i$

$$\frac{p_m}{q_m} = \frac{b_{i+1} b_{m-i+3} + b_i b_{m-i+1}}{b_i b_{m-i+3} + b_{i-1} b_{m-i+1}}.$$

§3.2　连分数映射与高斯测度

在本节我们将定义连分数上的动力系统, 以此为出发点来研究连分数. 因为我们在本节仅考虑$x = [a_0; a_1, a_2, \ldots] \in (0,1)$, 此时$a_0 = 0$, 所以约定

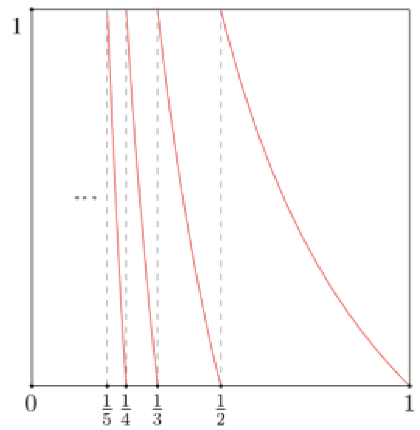
$$[a_1, a_2, \ldots] = [0; a_1, a_2, \ldots].$$

§3.2.1　高斯映射与高斯测度

定义区间$[0,1]$ 上映射$T : [0,1] \to [0,1]$ 如下

$$T(x) = \begin{cases} \frac{1}{x} - \left[\frac{1}{x}\right] = \left\{\frac{1}{x}\right\}, & 0 < x \leq 1; \\ 0, & x = 0, \end{cases} \tag{3.2.1}$$

其中$[\cdot]$ 表示取整, $\{\cdot\}$ 表示取小数部分. T 称为**连分数映射**(continued fraction) 或**高斯映射**(Gauss map).



高斯映射图像

根据定义，我们马上得到如下观察：设$x=[a_1,a_2,\ldots]\in[0,1]$ 为连分数表示，如果x 为无理数，那么$a_n\geq 1,\forall n\in N$, 那么

$$T(x)=T([a_1,a_2,\ldots])=[a_2,a_3,\ldots].$$

如果x 为有理数, 那么$x=[a_1,a_2,\ldots,a_n]$. 此时如果$n\geq 2$, 那么$T(x)=T([a_1,a_2,\ldots,a_n])=[a_2,a_3,\ldots,a_n]$; 如果$n=1$，那么$T(x)=T([a_1])=0$. 一般我们对有理数的连分数展开兴趣不大，所以后面我们主要考虑无理数集合$Y=[0,1]\setminus\mathbb{Q}$ 上的连分数展开.

对于$x\in Y$, 根据定义我们有：

$$x=\frac{1}{a_1+T(x)}=\cfrac{1}{a_1+\cfrac{1}{a_2+T^2(x)}}=\ldots=\cfrac{1}{a_1+\cfrac{1}{a_2+\cdots\cfrac{1}{a_{n-1}+\cfrac{1}{a_n+T^n(x)}}}}.$$

定义**高斯测度**如下：

$$\mu(A)=\frac{1}{\log 2}\int_A\frac{1}{1+x}dx,\ \forall A\in\mathcal{B}([0,1]).$$

因为如此定义的测度为无原子的, 所以$\mu([0,1]\cap\mathbb{Q})=0$.

命题 3.2.1 连分数系统$([0,1],\mathcal{B}([0,1]),\mu,T)$ 为保测系统.

证明. 显然T 为可测的, 我们仅需要证明

$$\mu(T^{-1}[0,s])=\mu([0,s]),\ \forall s>0.$$

首先易见

$$T^{-1}[0,s]=\{x:0\leq Tx\leq s\}=\bigsqcup_{n=1}^{\infty}\left[\frac{1}{s+n},\frac{1}{n}\right],$$

其中$\bigsqcup$ 表示无交并. 注意到恒等式

$$\frac{1+\frac{s}{n}}{1+\frac{s}{n+1}}=\frac{1+\frac{1}{n}}{1+\frac{1}{s+n}},$$

我们得到

$$\begin{aligned}\mu(T^{-1}[0,s])&=\frac{1}{\log 2}\sum_{n=1}^{\infty}\int_{\frac{1}{s+n}}^{\frac{1}{n}}\frac{1}{1+x}dx\\&=\frac{1}{\log 2}\sum_{n=1}^{\infty}\left(\log(1+\frac{1}{n})-\log(1+\frac{1}{s+n})\right)\\&=\frac{1}{\log 2}\sum_{n=1}^{\infty}\left(\log(1+\frac{s}{n})-\log(1+\frac{s}{1+n})\right)\\&=\frac{1}{\log 2}\sum_{n=1}^{\infty}\int_{\frac{s}{n+1}}^{\frac{s}{n}}\frac{1}{1+x}dx\\&=\mu([0,s]).\end{aligned}\tag{3.2.2}$$

即μ 为T 不变的. 所以连分数系统$([0,1],\mathcal{B}([0,1]),\mu,T)$ 为保测系统. □

结合注记3.1.5, 我们可以发现运用连分数映射可以对实数进行连分数展开. 设$x\in Y=[0,1]\setminus\mathbb{Q}$, 定义数列$\{a_n\}_{n\in\mathbb{N}}=\{a_n(x)\}_{n\in\mathbb{N}}$ 如下

$$\frac{1}{1+a_n}<T^{n-1}(x)<\frac{1}{a_n},\tag{3.2.3}$$

即

$$a_n(x)=\left[\frac{1}{T^{n-1}x}\right]\in\mathbb{N}.\tag{3.2.4}$$

于是得到连分数

$$[a_1,a_2,\ldots].$$

或者如下刻画：设$I_k=(\frac{1}{k+1},\frac{1}{k}]$, 于是根据$T^{n-1}x\in I_{a_n}$ 确定a_n.

引理 3.2.2 对于任何$x\in Y=[0,1]\setminus\mathbb{Q}$, 设$\{a_n\}_{n\in\mathbb{N}}=\{a_n(x)\}_{n\in\mathbb{N}}$ 如(3.2.4) 定义，那么

$$x=[a_1(x),a_2(x),\ldots].$$

证明. 根据注记3.1.5, 我们实际已经知道$x=[a_1(x),a_2(x),\ldots]$. 下面给出严格证明.

设$\{a_n\}_{n\in\mathbb{N}}=\{a_n(x)\}_{n\in\mathbb{N}}$ 如(3.2.4) 定义, 设

$$u=[a_1,a_2,\ldots].$$

下证$x=u$. 根据(3.1.11),

$$\frac{p_{2n}}{q_{2n}}<u<\frac{p_{2n+1}}{q_{2n+1}}.$$

由(3.1.8)(3.1.6), 我们有

$$\frac{p_{2n+1}}{q_{2n+1}}-\frac{p_{2n}}{q_{2n}}=\frac{1}{q_{2n}q_{2n-1}}\leq\frac{1}{2^{2n-2}}\to 0, n\to\infty.$$

下面我们归纳证明

$$[a_1,a_2,\ldots,a_{2n}]=\frac{p_{2n}}{q_{2n}}<x<\frac{p_{2n+1}}{q_{2n+1}}=[a_1,a_2,\ldots,a_{2n+1}]. \tag{3.2.5}$$

由此即有$u=x$.

$n=0$时, $\dfrac{p_0}{q_0}=0,\dfrac{p_1}{q_1}=\dfrac{1}{a_1}$, 所以根据$a_1$ 定义(3.2.5) 成立. 下面我们假设(3.2.5) 对n 和对于任何$x\in Y$ 成立. 尤其对于$Tx\in Y$ 有

$$[a_2,a_3,\ldots,a_{2n+1}]<Tx<[a_2,a_3,\ldots,a_{2n+2}].$$

因为$Tx=\frac{1}{x}-a_1$, 带入上式

$$a_1+[a_2,a_3,\ldots,a_{2n+1}]<\frac{1}{x}<a_1+[a_2,a_3,\ldots,a_{2n+2}].$$

于是

$$[a_1,a_2,\ldots,a_{2n+2}]=\frac{1}{a_1+[a_2,a_3,\ldots,a_{2n+2}]}<x<\frac{1}{a_1+[a_2,a_3,\ldots,a_{2n+1}]}=[a_1,a_2,\ldots,a_{2n+1}].$$

所以(3.2.5)对$n+1$ 成立. 证毕! □

注记 3.2.3 设

$$\phi:\mathbb{N}^{\mathbb{N}}\to[0,1],\quad (a_1,a_2,\ldots)\mapsto u=[a_1,a_2,\ldots]$$

σ 为$\mathbb{N}^{\mathbb{N}}$ 上的转移映射, 那么我们得到交换图表

$$\begin{tikzcd}
\mathbb{N}^{\mathbb{N}} \arrow[r, "\sigma"] \arrow[d, "\phi"'] & \mathbb{N}^{\mathbb{N}} \arrow[d, "\phi"] \\
{[0,1]} \arrow[r, "T"'] & {[0,1]}.
\end{tikzcd}$$

尤其对于$(a_1,a_2,\ldots)\in\mathbb{N}^{\mathbb{N}}$, $T([a_1,a_2,\ldots])=[a_2,a_3,\ldots]$.

注意前面对于二进制展开我们也得到类似交换图表:

$$\begin{tikzcd}
\Sigma_2 \arrow[r, "\sigma"] \arrow[d, "\psi"'] & \Sigma_2 \arrow[d, "\psi"] \\
{[0,1]} \arrow[r, "T_2"'] & {[0,1]}.
\end{tikzcd}$$

其中

$$\psi:\Sigma_2=\{0,1\}^{\mathbb{N}}\to[0,1],\quad (a_1,a_2,\ldots)\mapsto\sum_{n=1}^{\infty}\frac{a_n}{2^n}.$$

§3.2.2　高斯测度的遍历性

下面我们将证明高斯测度μ 为遍历的. 我们需要一些准备. 因为经常会对不同的x 用到命题3.1.1 中的p_n, q_n, 所以有时为了避免混淆我们记

$$\frac{p_n(x)}{q_n(x)} = [a_0(x); a_1(x), \ldots, a_n(x)] = a_0(x) + \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cdots \cfrac{1}{a_{n-1}(x) + \cfrac{1}{a_n(x)}.}}},\ n \geq 0. \tag{3.2.6}$$

设$u \notin \mathbb{Q}$, 其连分数展开为

$$u = [a_0; a_1, a_2, \ldots].$$

令

$$u_n = [a_n; a_{n+1}, a_{n+2}, \ldots]. \tag{3.2.7}$$

即

$$u = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots \cfrac{1}{a_{n-1} + \cfrac{1}{u_n}.}}}, \quad \forall n \in \mathbb{N}.$$

u_n 称为$[a_0; a_1, a_2, \ldots]$ 的**第$n+1$ 个完全商**（Complete quotient）. 对照高斯映射的定义，容易看出

$$u_n = \frac{1}{T^{n-1}(u)}, \quad \forall n \geq 2. \tag{3.2.8}$$

用两次命题3.1.1,

$$\begin{aligned}\begin{pmatrix} p_{n+k} \\ q_{n+k} \end{pmatrix} &= \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n+k} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \begin{pmatrix} a_{n+1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+2} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n+k} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}\end{aligned}$$

运用(3.2.6)和(3.2.7)的符号约定, 我们有

$$\begin{pmatrix} p_{n+k} \\ q_{n+k} \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \begin{pmatrix} p_{k-1}(u_{n+1}) & p_{k-2}(u_{n+1}) \\ q_{k-1}(u_{n+1}) & q_{k-2}(u_{n+1}) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \forall n, k \in \mathbb{N}.$$

于是有

$$\frac{p_{n+k}}{q_{n+k}}=\frac{p_n\dfrac{p_{k-1}(u_{n+1})}{q_{k-1}(u_{n+1})}+p_{n-1}}{q_n\dfrac{p_{k-1}(u_{n+1})}{q_{k-1}(u_{n+1})}+q_{n-1}}.$$

令$k\to\infty$, 就得到

$$u=\frac{p_nu_{n+1}+p_{n-1}}{q_nu_{n+1}+q_{n-1}}. \tag{3.2.9}$$

根据$u_{n+1}=\frac{1}{T^n(u)}$ 就得到

$$u=\frac{p_n+p_{n-1}T^n(u)}{q_n+q_{n-1}T^n(u)}. \tag{3.2.10}$$

定理 3.2.4 连分数系统$([0,1],\mathcal{B}([0,1]),\mu,T)$ 为遍历系统.

证明. 设

$$\phi:\mathbb{N}^{\mathbb{N}}\to\mathbb{R},\quad (a_1,a_2,\ldots)\mapsto u=[a_1,a_2,\ldots]$$

σ 为$\mathbb{N}^{\mathbb{N}}$ 上的转移映射, 那么我们得到图表

$$\begin{array}{ccc}\mathbb{N}^{\mathbb{N}} & \xrightarrow{\sigma} & \mathbb{N}^{\mathbb{N}}\\ {\scriptstyle\phi}\big\downarrow & & \big\downarrow{\scriptstyle\phi}\\ [0,1] & \xrightarrow[T]{} & [0,1].\end{array}$$

我们将结合上面图表对定理进行证明. 注意μ 在$\mathbb{N}^{\mathbb{N}}$ 中的拉回测度比较复杂, 我们希望对其上的可测方体进行估计. 在证明中我们用符号$f\asymp g$表示存在常数$C_1,C_2>0$ 使得

$$C_1f\le g\le C_2f.$$

取定$\mathbf{a}=(a_1,a_2,\ldots,a_n)\in\mathbb{N}^n$, 其长度$|\mathbf{a}|=n$. 定义$\mathbb{N}^{\mathbb{N}}$ 中的方体的像集

$$I(\mathbf{a})=\{[x_1,x_2,\ldots]\in[0,1]:x_i=a_i,1\le i\le n\}.$$

注意$I(\mathbf{a})$ 为以$[a_1,a_2,\ldots,a_n],[a_1,a_2,\ldots,a_n+1]$ 为端点的区间与$Y=[0,1]\setminus\mathbb{Q}$ 的交集.

我们下面要证明

$$\mu(T^{-n}A\cap I(\mathbf{a}))\asymp\mu(A)\mu(I(\mathbf{a})),\quad \forall A\in\mathcal{B}([0,1]). \tag{3.2.11}$$

为此我们仅需要对形如$A=[d,e]\subseteq[0,1]$ 的A 证明即可.

令

$$\frac{p_n}{q_n}=[a_1,a_2,\ldots,a_n],\quad \frac{p_{n-1}}{q_{n-1}}=[a_1,a_2,\ldots,a_{n-1}].$$

则

$$u \in I(\mathbf{a}) \Leftrightarrow u = [a_1, \ldots, a_n, a_{n+1}(u), a_{n+1}(u), \ldots].$$

于是由(3.2.10)

$$u \in T^{-n}A \cap I(\mathbf{a}) \Leftrightarrow u = \frac{p_n + p_{n-1}T^n(u)}{q_n + q_{n-1}T^n(u)} \text{ 且 } T^n(u) \in A = [d, e].$$

因为T^n 限制在$I(\mathbf{a})$ 是连续且单调的（如n 为偶数时为单调增；n 为奇数时为单调减），所以$T^{-n}A \cap I(\mathbf{a})$ 为以下面两点为端点的区间：

$$\frac{p_n + p_{n-1}d}{q_n + q_{n-1}d}, \quad \frac{p_n + p_{n-1}e}{q_n + q_{n-1}e}.$$

设m 为$[0,1]$ 上的Lebesgue 测度, 那么就有

$$m(T^{-n}A \cap I(\mathbf{a})) = \left| \frac{p_n + p_{n-1}d}{q_n + q_{n-1}d} - \frac{p_n + p_{n-1}e}{q_n + q_{n-1}e} \right|.$$

化简上式,

$$\begin{aligned}
&\left| \frac{p_n + p_{n-1}d}{q_n + q_{n-1}d} - \frac{p_n + p_{n-1}e}{q_n + q_{n-1}e} \right| \\
&= \left| \frac{(p_n + p_{n-1}d)(q_n + q_{n-1}e) - (p_n + p_{n-1}e)(q_n + q_{n-1}d)}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \right| \\
&= \left| \frac{p_n q_{n-1}e + p_{n-1}q_n d - p_n q_{n-1}d - p_{n-1}q_n e}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \right| \\
&= (e-d)\frac{|p_n q_{n-1} - p_{n-1}q_n|}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \\
&= (e-d)\frac{1}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \quad (\text{根据 } (3.1.8)).
\end{aligned}$$

另外在上面取$d=0, e=1$就得到

$$\mu(I(\mathbf{a})) = \left| \frac{p_n}{q_n} - \frac{p_n + p_{n-1}}{q_n + q_{n-1}} \right| = \frac{1}{q_n(q_n + q_{n-1})}. \tag{3.2.12}$$

于是我们就得到

$$\begin{aligned}
m(T^{-n}A \cap I(\mathbf{a})) &= m(A)m(I(\mathbf{a}))\frac{q_n(q_n + q_{n-1})}{(q_n + q_{n-1}d)(q_n + q_{n-1}e)} \\
&\asymp m(A)m(I(\mathbf{a})).
\end{aligned} \tag{3.2.13}$$

因为

$$\frac{m(B)}{2\log 2} \leq \mu(B) \leq \frac{m(B)}{\log 2}, \quad \forall B \in \mathcal{B}([0,1]),$$

根据(3.2.13), 我们得到(3.2.11), 即

$$\mu(T^{-n}A \cap I(\mathbf{a})) \asymp \mu(A)\mu(I(\mathbf{a})), \quad \forall A \in \mathcal{B}([0,1]).$$

另外注意到对于$n \in \mathbb{N}$, $\{I(\mathbf{a}) : \mathbf{a} \in \mathbb{N}^n\}$ 为$[0,1]$ 的剖分, 并且由(3.2.12)和(3.1.6),

$$\operatorname{diam}(I(\mathbf{a})) = \frac{1}{q_n(q_n + q_{n-1})} \le \frac{1}{2^{n-2}}.$$

因此

$$\{I(\mathbf{a}) : \mathbf{a} \in \mathbb{N}^n, n \in \mathbb{N}\}$$

生成了$\mathcal{B}([0,1])$. 于是根据(3.2.11)就有

$$\mu(T^{-n}A \cap B) \asymp \mu(A)\mu(B), \quad \forall A, B \in \mathcal{B}([0,1]).$$

所以根据定理2.1.3-4, T 为遍历的. □

§3.2.3 遍历性的应用

作为上面定理的应用, 我们得到:

定理 3.2.5 对于m几乎处处的$x = [a_1, a_2, \ldots] \in [0,1]$, 自然数$j$ 出现在连分数展开中的频率为

$$\frac{2\log(1+j) - \log j - \log(2+j)}{\log 2}, \tag{3.2.14}$$

$$\lim_{n\to\infty} (a_1 a_2 \ldots a_n)^{\frac{1}{n}} = \prod_{a=1}^{\infty} \left(\frac{(a+1)^2}{a(a+2)} \right)^{\frac{\log a}{\log 2}}, \tag{3.2.15}$$

$$\lim_{n\to\infty} \frac{1}{n}(a_1 + a_2 + \ldots + a_n) = \infty, \tag{3.2.16}$$

$$\lim_{n\to\infty} \frac{1}{n} \log q_n(x) = \frac{\pi^2}{12\log 2}, \tag{3.2.17}$$

以及

$$\lim_{n\to\infty} \frac{1}{n} \log \left| x - \frac{p_n(x)}{q_n(x)} \right| = -\frac{\pi^2}{6\log 2}. \tag{3.2.18}$$

证明. (1) 设$j \in \mathbb{N}$, 那么对于几乎处处的$x = [a_1, a_2, \ldots]$, 数字j 出现的频率为

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N} |\{i : 1 \le i \le N, a_i = j\}| \\
&= \lim_{N\to\infty} \frac{1}{N} |\{i : 1 \le i \le N, T^{i-1}x \in (\frac{1}{j+1}, \frac{1}{j}]\}| \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{1}_{(\frac{1}{j+1}, \frac{1}{j}]}(T^i x) = \frac{1}{\log 2} \int_{\frac{1}{j+1}}^{\frac{1}{j}} \frac{1}{1+y} dy \\
&= \frac{2\log(1+j) - \log j - \log(2+j)}{\log 2}.
\end{aligned}$$

(2) 定义函数$f:(0,1)=\bigcup_{a=1}^{\infty}(\frac{1}{a+1},\frac{1}{a}]\to\mathbb{R}$ 如下, 如果$x\in(\frac{1}{a+1},\frac{1}{a}]$, 那么定义$f(x)=\log a$. 首先由

$$\begin{aligned}\int_0^1 f(x)dx &= \sum_{a=1}^{\infty}\left(\frac{1}{a}-\frac{1}{a+1}\right)\log a\\ &\le \sum_{a=1}^{\infty}\frac{1}{a^2}\log a<\infty,\end{aligned}$$

以及

$$\frac{d\mu}{dx}=\frac{1}{(1+x)\log 2}$$

知道$\int_0^1 fd\mu<\infty$, 即$f\in L^1([0,1],\mu)$. 所以由Birkhoff 逐点收敛定理,

$$\begin{aligned}&\frac{1}{n}\sum_{j=0}^{n-1}\log a_j=\frac{1}{n}\sum_{j=0}^{n-1}f(T^jx)\\ &\xrightarrow{a.e}\int_0^1 fd\mu=\sum_{a=1}^{\infty}\frac{\log a}{\log 2}\int_{\frac{1}{1+a}}^{\frac{1}{a}}\frac{1}{1+x}dx\\ &=\sum_{a=1}^{\infty}\frac{\log a}{\log 2}\Big(\log(1+\frac{1}{a})-\log(1+\frac{1}{a+1})\Big),\quad n\to\infty.\end{aligned}$$

即$\lim_{n\to\infty}(a_1a_2\dots a_n)^{\frac{1}{n}}=\prod_{a=1}^{\infty}\left(\frac{(a+1)^2}{a(a+2)}\right)^{\frac{\log a}{\log 2}}$.

(3) 设$g:(0,1)\to\mathbb{R},g(x)=e^{f(x)}$, 即$g(x)=a_1(x)$, 其中$a_1(x)$ 为$x=[a_1(x),a_2(x),\dots]$ 连分数展开的第1 项系数. 于是

$$\frac{a_1+a_2+\ldots+a_n}{n}=\frac{1}{n}\sum_{j=0}^{n-1}g(T^jx).$$

但是因为$\int_0^1 gd\mu=\infty$, 所以不能直接用Birkhoff 遍历定理.

对于$N>0$, 令

$$g_N(x)=\begin{cases} g(x), & \text{如果}g(x)\le N;\\ 0, & \text{否则}.\end{cases}$$

因为

$$\int_0^1 g_Nd\mu=\frac{1}{\log 2}\sum_{a=1}^{N}\int_{\frac{1}{a+1}}^{\frac{1}{a}}adx=\frac{1}{\log 2}\sum_{a=1}^{N}\frac{1}{a+1}<\infty.$$

注意$\lim_{N\to\infty}\int_0^1 g_Nd\mu=\infty$. 对$g_N$ 用遍历定理,

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}g(T^jx)&\ge\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}g_N(T^jx)\\ &=\int_0^1 g_Nd\mu\to\infty,N\to\infty.\end{aligned}$$

于是我们有(3.2.16).

(4) 首先注意到

$$\begin{aligned}\frac{p_n(x)}{q_n(x)} &= \frac{1}{a_1+[a_2,\ldots,a_n]}\\ &= \frac{1}{a_1+\dfrac{p_{n-1}(Tx)}{q_{n-1}(Tx)}}\\ &= \frac{q_{n-1}(Tx)}{p_{n-1}(Tx)+q_{n-1}(Tx)a_1}.\end{aligned}$$

于是根据p_n, q_n定义约定我们有

$$p_n(x) = q_{n-1}(Tx).$$

注意约定$p_1 = q_0 = 1$，我们得到

$$\frac{1}{q_n(x)} = \frac{p_n(x)}{q_n(x)}\cdot\frac{p_{n-1}(Tx)}{q_{n-1}(Tx)}\cdots\frac{p_1(T^{n-1}x)}{q_1(T^{n-1}x)}.$$

即有

$$-\frac{1}{n}\log q_n(x) = \frac{1}{n}\sum_{j=0}^{n-1}\log\left(\frac{p_{n-j}(T^jx)}{q_{n-j}(T^jx)}\right).$$

令$h(x) = \log x \in L^1([0,1],\mu)$. 则

$$-\frac{1}{n}\log q_n(x) = \frac{1}{n}\sum_{j=0}^{n-1}h(T^jx) - \frac{1}{n}\sum_{j=0}^{n-1}\left[\log(T^jx) - \log\left(\frac{p_{n-j}(T^jx)}{q_{n-j}(T^jx)}\right)\right].$$

令

$$S_n = \frac{1}{n}\sum_{j=0}^{n-1}h(T^jx),$$

$$R_n = \frac{1}{n}\sum_{j=0}^{n-1}\left[\log(T^jx) - \log\left(\frac{p_{n-j}(T^jx)}{q_{n-j}(T^jx)}\right)\right].$$

下面我们分别计算S_n, R_n. 首先由遍历定理, 我们有

$$\lim_{n\to\infty}\frac{1}{n}S_n = \frac{1}{\log 2}\int_0^1\frac{\log x}{1+x}dx = -\frac{\pi^2}{12\log 2}.$$

下面我们证明

$$\lim_{n\to\infty}\frac{1}{n}R_n = 0,$$

从而完成证明. 根据(3.1.6)(3.1.7)(3.1.13), 我们有

$$\left|\frac{x}{p_k/q_k} - 1\right| = \frac{q_k}{p_k}\left|x - \frac{p_k}{q_k}\right| \le \frac{1}{p_kq_{k+1}} \le \frac{1}{2^{k-1}}.$$

运用不等式

$$|\log u| \leq 2|u-1|,\ u \in [1/2, 3/2],$$

我们得到

$$\begin{aligned}|R_n| &\leq \sum_{j=0}^{n-1} \left|\log \frac{T^j x}{p_{n-j}(T^j x)/q_{n-j}(T^j x)}\right| \\ &\leq \underbrace{2\sum_{j=0}^{n-2} \left|\frac{T^j x}{p_{n-j}(T^j x)/q_{n-j}(T^j x)} - 1\right|}_{T_n} + \underbrace{\left|\log \frac{T^{n-1} x}{p_1(T^{n-1}x)/q_1(T^{n-1}x)}\right|}_{U_n}\end{aligned}$$

下面估计T_n, U_n.

$$T_n \leq \sum_{j=0}^{n-2} \frac{2}{2^{n-j-1}} \leq 2, \forall n.$$

注意到

$$U_n = \left|\log[(T^{n-1}x)]a_1(T^{n-1}x)\right|,$$

由(3.2.3) 以及$a_1(T^{n-1}x) \geq 1$,

$$1 \geq (T^{n-1}x)]a_1(T^{n-1}x) \geq \frac{a_1(T^{n-1}x)}{1 + a_1(T^{n-1}x)} \geq \frac{1}{2}.$$

所以

$$U_n = \left|\log[(T^{n-1}x)]a_1(T^{n-1}x)\right| \leq \log 2.$$

综合上面所得, 我们得到

$$\lim_{n\to\infty} \frac{1}{n} R_n = 0.$$

(5) 根据(3.1.13)(3.1.17), 我们有

$$\log q_n + \log q_{n+1} \leq -\log\left|x - \frac{p_n}{q_n}\right| \leq \log q_n + \log q_{n+2}.$$

于是根据(3.2.17)得到(3.2.18). $\square$

习　题

1. 证明$\phi : \mathbb{N}^{\mathbb{N}} \to [0,1] \setminus \mathbb{Q}, \quad (a_1, a_2, \ldots) \mapsto u = [a_1, a_2, \ldots]$ 为同胚, 其中$\mathbb{N}$ 取离散拓扑, $\mathbb{N}^{\mathbb{N}}$ 取乘积拓扑.

§3.3　坏逼近数与Lagrange定理

本节介绍著名的Lagrange定理.

§3.3.1 坏逼近数

一个数x为**代数数**是指它为某个整系数多项式的根, 其中满足条件多项式的最小阶数称为代数数的阶. 例如任何有理数为代数数, 其阶为1；$\sqrt{2}, \sqrt{3}$ 为阶2的代数数；$\sqrt{2}+\sqrt{3}$ 为阶4的代数数. 一个不是代数数的实数称为**超越数**(transcendental). 容易证明代数数全体是可数的, 所以超越数总是存在的. Liouville 给出了一个超越数存在的简单证明. 一个实数x 称为**Liouville 数**, 是指它为无理数, 并且对于每个$n \in \mathbb{N}$, 存在$p \in \mathbb{Z}, q \in \mathbb{N}$ 使得

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^n}.$$

例如$x = \sum_{k=1}^{\infty} \frac{1}{10^{k!}}$ 为Liouville 数（取$q = 10^{n!}$）. Liouville 证明了任何Liouville数为超越的. 从而证明了超越数的存在性. 在他的证明中, 他还指出了一个有趣的结论：对于n 阶代数数x, 存在$M \in \mathbb{N}$ 使得

$$\left| x - \frac{p}{q} \right| > \frac{1}{Mq^n}, \quad \forall p \in \mathbb{Z}, q \in \mathbb{N}.$$

在这里我们不讨论超越数的问题, 我们在本节介绍Lagrange的一个漂亮定理, 指出连分数可以刻画2阶无理数.

定义 3.3.1 一个数$u = [a_1, a_2, \ldots] \in [0,1]$ 称为**坏逼近的**(*badly approximable*) 是指存在$M > 0$ 使得$a_n \leq M, \ \forall n \in \mathbb{N}$.

例如, $\frac{\sqrt{5}-1}{2} = [1,1,1,\ldots]$ 为坏逼近的. 根据(3.2.16), 我们知道坏逼近数的全体是m零测集.

命题 3.3.2 一个数$u \in [0,1]$ 为坏逼近的当且仅当存在$\varepsilon > 0$ 使得对于任何有理数$\frac{p}{q}$ 有

$$\left| u - \frac{p}{q} \right| \geq \frac{\varepsilon}{q^2}.$$

证明. 设$u \in [0,1]$ 为坏逼近的, 那么存在$M > 0$ 使得$a_n \leq M, \ \forall n \in \mathbb{N}$. 根据(3.1.4),

$$q_{n+1} \leq (M+1)q_n, \ \forall n \geq 0.$$

对于任何q, 存在n 使得$q \in (q_{n-1}, q_n]$. 根据命题3.1.10 以及(3.1.17),

$$\left| \frac{p}{q} - u \right| > \left| \frac{p_n}{q_n} - u \right| > \frac{1}{q_n q_{n+2}} > \frac{1}{(M+1)^2 q^2}.$$

反之, 设存在$\varepsilon > 0$ 使得对于任何有理数$\frac{p}{q}$ 有

$$\left| u - \frac{p}{q} \right| \geq \frac{\varepsilon}{q^2}.$$

尤其由(3.1.13), 就有
$$\frac{\varepsilon}{q_n^2} \le \left|u - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}}.$$
于是
$$a_{n+1}q_n < a_{n+1}q_n + q_{n-1} = q_{n+1} < \frac{1}{\varepsilon}q_n,$$
即
$$a_n \le \frac{1}{\varepsilon},\ \forall n \ge 1$$
证毕! □

§3.3.2　二次无理数与Lagrange定理

定义 3.3.3 $u \in \mathbb{R}$ 称为**二次无理的**(*quadratic irrational*) 是指$u \notin \mathbb{Q}$, 并且存在$a, b, c \in \mathbb{Z}$ 使得$au^2 + bu + c = 0$. 即$u \in \mathbb{R}$ 为二次无理的当且仅当$\mathbb{Q}(u)$ 为相对于$\mathbb{Q}$ 的二次扩张域.

定义 3.3.4 连分数$[a_0; a_1, a_2, \ldots]$ 称为**最终周期的**(*eventually periodic*) 指存在数$N \ge 0$ 以及$k \ge 1$ 使得
$$a_{n+k} = a_n,\ \forall n \ge N.$$
此时把连分数记为
$$[a_0; a_1, \ldots, a_{N-1}, \overline{a_N, \ldots, a_{N+k-1}}].$$
例如
$$\sqrt{2} = [1; \overline{2}], \quad \sqrt{3} = [1; \overline{1, 2}],$$
$$\sqrt{5} = [2; \overline{4}], \quad \sqrt{7} = [2; \overline{1, 1, 1, 4}].$$

定理 3.3.5 (Lagrange定理) 设$u \notin \mathbb{Q}$. 那么u 为二次无理的当且仅当u 的连分数展开是最终周期的.

证明. 首先设$u = [\overline{a_0; a_1, a_2, \ldots, a_k}]$. 由定理3.1.3, $u \notin \mathbb{Q}$. 根据u 的定义
$$u_{k+1} = u_0 = u.$$
由(3.2.9),
$$u = \frac{up_k + p_{k-1}}{uq_k + q_{k-1}}.$$
于是
$$q_k u^2 + (q_{k-1} - p_k)u - p_{k-1} = 0,$$
即u 为二次无理的.

现在假设
$$u = [a_0; a_1, \ldots, a_{N-1}, \overline{a_N, \ldots, a_{N+k}}].$$

由(3.2.9),
$$u = \frac{[\overline{a_N; a_{N+1}, \dots, a_{N+k}}]p_{N-1} + p_{N-2}}{[\overline{a_N; a_{N+1}, \dots, a_{N+k}}]q_{N-1} + q_{N-2}}.$$
所以$\mathbb{Q}(u) = \mathbb{Q}([\overline{a_N; a_{N+1}, \dots, a_{N+k}}])$, u 为二次无理的.

反之, 设u 为二次无理的. 设
$$f_0(u) = \alpha_0 u^2 + \beta_0 u + \gamma_0 = 0,$$
其中$\alpha_0, \beta_0, \gamma_0 \in \mathbb{Z}$ 并且$\delta = \beta_0^2 - 4\alpha_0\gamma_0$ 不是完全平方数.

断言 *对于每个$n \geq 0$, 存在多项式*
$$f_n(x) = \alpha_n x^2 + \beta_n x + \gamma_n,$$
使得$f_n(u_n) = 0$ 且
$$\beta_n^2 - 4\alpha_n\gamma_n = \delta.$$

断言证明: 我们对n 归纳证明. $n = 0$ 已然. 设n 是断言成立.

因为$u_n = a_n + \frac{1}{u_{n+1}}$, 所以由归纳假设
$$f_n(a_n + \frac{1}{u_{n+1}}) = 0.$$
带入到$f_n(x) = \alpha_n x^2 + \beta_n x + \gamma_n$ 整理得到
$$f_{n+1}(x) = \alpha_{n+1}x^2 + \beta_{n+1}x + \gamma_{n+1},$$
其中
$$\alpha_{n+1} = a_n^2\alpha_n + a_n\beta_n + \gamma_n, \tag{3.3.1}$$
$$\beta_{n+1} = 2a_n\alpha_n + \beta_n \tag{3.3.2}$$
$$\gamma_{n+1} = \alpha_n. \tag{3.3.3}$$
明显的, $\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1} \in \mathbb{Z}$. 直接验证得到:
$$\beta_{n+1}^2 - 4\alpha_{n+1}\gamma_{n+1} = \beta_n^2 - 4\alpha_n\gamma_n = \delta.$$
断言证毕. □

根据断言, 每个多项式f_n 都有相同的非完全平方数的判别式δ, 由此
$$\alpha_n \neq 0, \forall n \geq 0.$$

如果存在某个$N > 0$ 使得$n \geq N$ 时$\alpha_n > 0$, 那么根据(3.3.2), $\beta_N, \beta_{N+1}, \dots$ 单调递增. 结合(3.3.3) 就会得到n充分大以后, $\alpha_n, \beta_n, \gamma_n > 0$. 但这与

$$f_n(u_n) = 0,\ u_n > 0$$

矛盾. 类似的不存在N 使得$n \geq N$ 时$\alpha_n < 0$. 所以α_n 在不停的变化正负号. 令

$$A = \{n \in \mathbb{N} : \alpha_n \alpha_{n-1} < 0\}.$$

则$|A| = \infty$.

由(3.3.3)

$$\alpha_n \gamma_n < 0, \forall n \in A.$$

因为

$$\beta_n^2 - 4\alpha_n \gamma_n = \delta, \forall n \geq 0,$$

所以对于任何$n \in A$ 有

$$|\alpha_n| \leq \frac{1}{4}\delta, \quad |\beta_n| < \sqrt{\delta}, \quad |\gamma_n| \leq \frac{1}{4}\delta.$$

于是A 为无穷集, 而$(\alpha_n, \beta_n, \gamma_n)$ 只有有限种可能性, 从而存在$n_1 < n_2 < n_3$ 使得

$$f_{n_0} = f_{n_1} = f_{n_2}.$$

因为二次多项式至多两个根, 所以$u_{n_1}, u_{n_2}, u_{n_3}$ 必有两个是相等的, 这就意味着u 为最终周期的. □

推论 3.3.6 任何二次无理数必为坏逼近的.

习　题

1. 对于$u \in \mathbb{R}$ 中任何两个连续的渐近分数中至少有一个满足
$$\left| u - \frac{p}{q} \right| < \frac{1}{2q^2}.$$
于是对于无理数u, 满足上面不等式的有理数有无穷多个.

2. (Hurwitz 定理) 对于$u \in \mathbb{R}$ 中任何三个连续的渐近分数中至少有一个满足
$$\left| u - \frac{p}{q} \right| < \frac{1}{\sqrt{5}q^2}.$$
于是对于无理数u, 满足上面不等式的有理数有无穷多个.

3. 在Hurwitz 定理中, $\sqrt{5}$ 为最佳数, 即如果$A > \sqrt{5}$, 纯存在$u \in \mathbb{R}$ 使得
$$\left| u - \frac{p}{q} \right| < \frac{1}{Aq^2}$$
不能有无穷组解.

4. 称两个实数ξ, η 为相似的, 指存在$a, b, c, d \in \mathbb{Z}, ad - bc = \pm 1$ 使得

$$\xi = \frac{a\eta + b}{c\eta + d}.$$

证明：

(a) 相似为等价关系.

(b) 有理数彼此相似.

(c) $\xi = \frac{a\eta+b}{c\eta+d}$ 可以表示为$\xi = [a_0; a_1, \ldots, a_{k-1}, \eta], k \geq 2$ 当且仅当$c > d > 0$.

(d) 两个无理数ξ, η 相似当且仅当它们连分数展开式中若干项后完全一样, 即

$$\xi = [a_0; a_1, \ldots, a_m, c_1, c_2, \ldots],\ \eta = [b_0; b_1, \ldots, b_m, c_1, c_2, \ldots].$$

§3.4 注记

关于连分数有许多优秀的专著, 本章的素材来自[54] 和[211]. 动力系统在连分数的进一步应用参见[45] 和[131] 等.

第四章　Lebesgue空间与同构

在本章中我们讨论一个基本的问题：什么时候两个保测系统是"一样的"？为此我们需要讨论如何定义保测系统的同构，这里有几个不同的定义，我们在这章的目的就是讨论这些定义的异同.

§4.1　Lebesgue空间

§4.1.1　Lebesgue空间

对于拓扑空间$(X, \mathcal{T})$，我们总是用$\mathcal{B}(X) = \sigma(\mathcal{T})$ 表示Borel σ 代数，即开集生成的子σ代数.

定理 4.1.1 *设$(X, \mathcal{X})$ 为可测空间，那么以下等价：*

1. *存在可分的度量空间Y 使得$(X, \mathcal{X})$ 同构于$(Y, \mathcal{B}(Y))$.*
2. *存在Cantor集的子集Y使得$(X, \mathcal{X})$ 同构于$(Y, \mathcal{B}(Y))$.*
3. *$\mathcal{X}$ 为可数生成的，并且分离点，即存在可数子集$\{A_n\}_{n\in\mathbb{N}}$ 使得$\mathcal{X} = \sigma(\{A_n\}_{n=1}^{\infty})$，以及对于任何$x \neq x'$，存在$A \in \mathcal{X}$ 使得$x \in A$ 且$x' \notin A$.*

*满足上面条件的可测空间称为**Borel 空间**，此时一般也记$\mathcal{X} = \mathcal{B}(X)$.*

注记 4.1.2 *1. 在数学中词语"同构"在各个分支里面使用比较频繁，我们需要根据上下文来理解碰到的同构是哪一个定义. 例如在定理4.1.1中，根据上下文，我们知道此处的同构是指可测空间的同构，即$\phi : (X, \mathcal{X}) \to (Y, \mathcal{B}(Y))$ 为双射且ϕ, ϕ^{-1} 可测.*

2. 定理4.1.1证明中唯一困难的一步是3 推2. 这个证明是标准的，我们把构造陈述如下. 设$\{A_n\}_{n\in\mathbb{N}}$ 为生成$\mathcal{X}$ 的可数子集族，定义

$$\phi : X \to \{0, 1\}^{\mathbb{N}},\ (\phi(x))_n = \mathbf{1}_{A_n}(x).$$

因为$\mathcal{X}$ 分离点，所以ϕ 为单射. 令$Y = \phi(X) \subseteq \{0, 1\}^{\mathbb{N}}$，那么

$$\phi(A_n) = Y \cap {}_n[1]_n = \{y \in Y : y_n = 1\}.$$

容易验证，ϕ, ϕ^{-1} 都是可测的.

设X是一个拓扑空间. 如果存在一个相容的度量d 使得(X, d) 是一个可分完备度量空间，则称X 是一个**Polish 空间**.

定义 4.1.3 *设$(X, \mathcal{X})$是一个可测空间，如果存在Polish 空间Y 使得$(X, \mathcal{X})$ 同构于$(Y, \mathcal{B}(Y))$，那么称$(X, \mathcal{X})$ 是一个**标准Borel 空间**.*

Polish 空间要么有限, 要么可数, 要么具有连续统的势. 下面的Kuratowski 同构定理告诉我们, 具有连续统势的标准Borel空间只有一个. 因为我们很少对有限或可数的系统感兴趣, 所以下面定理对我们而言十分重要.

定理 4.1.4 (Kuratowski 同构定理) 设$(X,\mathcal{X}),(Y,\mathcal{Y})$ 为标准*Borel* 空间, 那么$(X,\mathcal{X}),(Y,\mathcal{Y})$ 为同构的当且仅当

$$\mathrm{Card}(X)=\mathrm{Card}(Y).$$

尤其, 任何两个具有相同不可数势的标准*Borel* 空间为同构的.

下面的定理罗列了标准Borel空间的三个性质, 第一个是说标准Borel空间的子空间仍然是标准Borel的；第二指出Borel可测集的单射可测像仍为Borel可测集；第三个指出, 对于标准Borel空间而言, 可数生成且分离点的σ 代数是唯一的.

定理 4.1.5 *1.* 设$(X,\mathcal{X})$ 为标准*Borel* 空间, $Y\in\mathcal{X}$, 那么$(Y,\mathcal{X}\cap Y)$ 也是标准*Borel*空间.

2. 设$(X,\mathcal{X}),(Y,\mathcal{Y})$ 为标准*Borel* 空间, $f:X\to Y$ 为可测映射. 设$A\in\mathcal{X}$ 并且$f|_A$ 为单射, 那么$f(A)\in\mathcal{Y}$, 并且

$$f|_A:(A,\mathcal{X}\cap A)\to(f(A),f(A)\cap\mathcal{Y})$$

为同构.

3. 设$(X,\mathcal{X})$ 为标准*Borel*空间. 如果$\mathcal{A}\subseteq\mathcal{X}$ 为可数生成并且分离点的σ-代数, 那么$\mathcal{A}=\mathcal{X}$.

设$(X,\mathcal{B}(X)),(Y,\mathcal{B}(Y))$ 为标准Borel 空间, $f:Y\to X$ 为可测映射. 一般而言, 可测集$A\in\mathcal{B}(Y)$ 的像$B=f(A)$ 不一定为可测集. 于是有下面的定义.

定义 4.1.6 设$(X,\mathcal{B}(X))$ 为标准*Borel* 空间, 子集$B\subseteq X$ 称为**分析集***(analytic set)* , 如果存在标准*Borel* 空间$(Y,\mathcal{B}(Y))$, 可测映射$f:Y\to X$ 以及$A\in\mathcal{B}(Y)$ 使得$B=f(A)$.

用$\mathcal{A}(X)$ 表示X 上全体分析集生成的σ-代数.

注记 4.1.7 *1. Souslin* 指出, 任何不可数标准*Borel*空间一定存在不是*Borel*可测的分析集.

2. 分析集的补集一般不再是分析集, 一个重要的结果是: $B\in\mathcal{B}(X)$ 当且仅当$B\in\mathcal{A}(X)$ 且$B^c\in\mathcal{A}(X)$.

设$(X,\mathcal{X},\mu)$ 为测度空间, 一个子集$A\subseteq X$ 称为μ-**零集**(μ-null)的是指, 存在$B\in\mathcal{X}$ 使得$A\subseteq B$ 并且$\mu(B)=0$, 即它包含在某个零测集中. 令$\mathfrak{I}_\mu$ 为全体μ-零集的全体. 注意一般而言, $\mathfrak{I}_\mu\not\subseteq\mathcal{X}$. 令

$$\mathcal{X}_\mu=\{A\cup N:A\in\mathcal{X},N\in\mathfrak{I}_\mu\}.$$

那么容易验证, $\mathcal{X}_\mu$ 仍是一个σ 代数, 称$\mathcal{X}_\mu$ 中元素为**μ-可测集**. 将μ 自然扩充到$\mathcal{X}_\mu$ 上的测度$\overline{\mu}$: $\overline{\mu}(A\cup N)=\mu(A), \forall A\in\mathcal{X}, N\in\mathfrak{I}_\mu$. 那么$(X,\mathcal{X}_\mu,\overline{\mu})$ 成为一个完备的测度空间, 称为$(X,\mathcal{X},\mu)$ 的**完备化**. 如无歧义, 记$(X,\mathcal{X}_\mu,\overline{\mu})$ 为$(X,\mathcal{X}_\mu,\mu)$.

设$(X,\mathcal{B}(X))$ 为标准Borel 空间, μ 为其上σ有限的测度, 记全体μ 可测集为$\mathcal{B}_\mu(X)$, 即$\mathcal{B}_\mu(X)$ 为$\mathcal{B}(X)$ 在μ 下的完备化. 称一个子集$A\subseteq X$ 为**普遍可测的**(universally measurable), 是指对于任何σ 有限的测度, $A\in\mathcal{B}_\mu(X)$, 即

$$A\in\bigcap_\mu \mathcal{B}_\mu(X).$$

定理 4.1.8 (Lusin) *设$(X,\mathcal{B}(X))$ 为标准Borel空间, 那么任何分析集为普遍可测的, 即*

$$\mathcal{A}(X)\subseteq\bigcap_\mu \mathcal{B}_\mu(X).$$

定理 4.1.9 (Jankov-von Neumann) *设$(X,\mathcal{X}),(Y,\mathcal{Y})$ 为标准Borel 空间, $f:X\to Y$ 为可测映射. 则存在分析集可测映射$g:f(X)\to X$ (即X 可测集的逆像为$f(X)$ 上的分析集) 使得*

$$f\circ g=\mathrm{id}_{f(X)}.$$

定理 4.1.10 *设$(X,\mathcal{B}(X))$ 为标准Borel 空间, μ 为连续概率测度. 那么$(X,\mathcal{B}(X),\mu)$ 同构于$([0,1],\mathcal{B}([0,1]),m)$, 其中$m$ 为Lebesgue 测度.*

上面定理中两个测度空间的同构定义可参见下文定义4.1.20.

定义 4.1.11 (Lebesgue空间) *称概率空间$(X,\mathcal{X},\mu)$ 是一个***标准Lebesgue 空间** *(或者直接称为***Lebesgue 空间***), 如果$(X,\mathcal{X})$ 为标准的Borel 空间并且μ 为Borel 概率测度.*

注记 4.1.12 *1. 也有很多作者把Lebesgue 空间定义为标准Borel 空间的完备化. 此时如果是连续测度, 那么就同构于$([0,1],\mathcal{L},m)$, 其中$\mathcal{L}$ 为Lebesgue 可测集全体.*

2. 于是根据定理4.1.10, Lebesgue 空间同构于具有下面形式的概率空间

$$\Big([0,s]\cup A,\mathcal{B},m+\sum_{a\in A}p_a\delta_a\Big),$$

其中A 是一个至多可数集, $A\cap[0,s]=\emptyset$, $s\in[0,1]$, $p_a>0$ 且$\sum_{a\in A}p_a=1-s$, $\mathcal{B}$ 由$\mathcal{B}([0,s])$ (如要求是完备的测度, 那么此处换为Lebesgue 可测集) 和A 的所有子集生成的σ 代数, m 为Lebesgue 测度. 特别地, 当$s=1$时, $A=\emptyset$, 测度为连续的; 当$s=0$ 时, 这个测度是纯原子的.

布尔代数（Boolean algebra）是指一个集合R，其上有运算$\vee, \wedge$ 和 $'$ 满足

(1) $A \vee A = A$, $A \wedge A = A$, $\forall A \in R$;

(2) $A \vee B = B \vee A$, $A \wedge B = B \wedge A$, $\forall A, B \in R$;

(3) $(A \vee B) \vee C = A \vee (B \vee C)$, $(A \wedge B) \wedge C = A \wedge (B \wedge C)$, $\forall A, B, C \in R$;

(4) $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$, $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$, $\forall A, B, C \in R$;

(5) $(A \wedge B)' = A' \vee B'$;

(6) $(A')' = A$;

(7) 存在元素$0 \in R$ 满足$A \wedge 0 = 0$, $A \vee 0 = A$, $\forall A \in R$;

(8) $A' \wedge A = 0$, $\forall A \in R$.

给定布尔代数R, 其上可以定义半序$A \leq B \iff A \wedge B = A$. 易见0 为此半序下最小元, $R = 0'$ 为最大元.

布尔代数称为**布尔σ 代数**指，对于任何$\{A_n\}_{n=1}^{\infty} \subseteq R$, 存在最小的$B \in R$ 使得$A_n \leq B, \forall n \in \mathbb{N}$. 此时将$B$ 记为$\bigvee_{n=1}^{\infty} A_n$. 注意, $C = (\bigvee_{n=1}^{\infty} A_n')'$ 为满足$C \leq A_n, \forall n \in \mathbb{N}$ 的最大元, 记为$C = \bigwedge_{n=1}^{\infty} A_n$.

定义 4.1.13 *设X为集合, $\mathfrak{I} \subseteq \mathcal{P}(X)$ 称为***理想***(ideal)是指*

1. $\emptyset \in \mathfrak{I}$,

2. $A \subseteq B, B \in \mathfrak{I}$, *那么*$A \in \mathfrak{I}$,

3. $A, B \in \mathfrak{I}$, *那么*$A \cup B \in \mathfrak{I}$.

如果第三条换为对于可数并封闭, 那么称$\mathfrak{I}$*为*σ **理想**.

设$(X, \mathcal{X})$ 为可测空间, $\mathfrak{I}$ 为X上的σ 理想. 根据$\mathfrak{I}$ 可以定义$\mathcal{X}$ 上的等价关系:

$$A \sim B \iff A \Delta B \in \mathfrak{I}.$$

令$[A]$ 为$A \in \mathcal{X}$ 对应的等价类, 将得到的布尔σ 代数记为$\mathcal{X}/\mathfrak{I}$. 即

$$\mathcal{X}/\mathfrak{I} = \{[A] : A \in \mathcal{X}\}.$$

定理 4.1.14 *设$(X, \mathcal{X})$ 为可测空间, $\mathfrak{I}$ 为X 上的σ 理想, 设$(Y, \mathcal{B}(Y))$ 为非空的标准Borel空间. 如果*

$$\Phi : \mathcal{B}(Y) \to \mathcal{X}/\mathfrak{I}$$

为布尔σ 代数的同态, 那么存在可测映射$\phi : X \to Y$ 使得

$$\Phi(A) = [\phi^{-1}(A)],\ \forall A \in \mathcal{B}(Y).$$

并且在模$\mathfrak{I}$ 意义下ϕ 唯一, 即如果$\psi : X \to Y$ 也满足条件, 那么

$$\{x \in X : \phi(x) \neq \psi(x)\} \in \mathfrak{I}.$$

定理 4.1.15 设$(X, \mathcal{B}(X)), (Y, \mathcal{B}(Y))$ 为标准*Borel* 空间, $\mathfrak{I}_X, \mathfrak{I}_Y$ 分别为X, Y 上的σ 理想. 那么

$$\Phi : \mathcal{B}(Y)/\mathfrak{I}_Y \to \mathcal{B}(X)/\mathfrak{I}_X$$

为布尔σ 代数的同构当且仅当存在$X_0 \in \mathcal{B}(X), Y_0 \in \mathcal{B}(Y)$ 以及*Borel* 同构$\phi : X_0 \to Y_0$ 使得

$$\Phi([A]) = [\phi^{-1}(A \cap Y_0)],\ \forall A \in \mathcal{B}(Y).$$

并且在模$\mathfrak{I}_X$ 意义下ϕ 唯一.

如果$\mathfrak{I}_X, \mathfrak{I}_Y$ 都包含了一个不可数集合, 那么可以要求$X_0 = X, Y_0 = Y$.

§4.1.2　测度代数

定义 4.1.16 **测度代数***(measure algebra)* 是指二元组$(\mathcal{A}, \mu)$, 其中$\mathcal{A}$ 为布尔σ 代数, 而$\mu : \mathcal{A} \to \mathbb{R}_+$ 满足

1. $\mu(A) = 0$ 当且仅当$A = \emptyset$,
2. $\mu(X) < \infty$,
3. 对于互不相交序列$\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ 有$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

在定义中X 和$\emptyset$ 表示$\mathcal{A}$ 中单位元素与零元素, $\bigcap, \bigcup$ 表示布尔代数中的算子, 上面$\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ 互不相交是指$A_i \cap A_j = \emptyset, \forall i \neq j$.

设$(\mathcal{A}, \mu)$ 为测度代数, 那么可以定义度量

$$\rho(A, B) = \mu(A \Delta B).$$

容易验证, ρ为度量, 并且$(\mathcal{A}, \rho)$ 为完备度量空间, 以及$\mu : \mathcal{A} \to \mathbb{R}_+$ 为一致连续的函数. 还容易验证$A \mapsto A^c$ 为等距映射, 以及对于任何$A, B, C, D \in \mathcal{A}$

$$\rho(A \cup B, C \cup D) + \rho(A \cap B, C \cap D) \leq \rho(A, C) + \rho(B, D).$$

由此知道在此度量下, 布尔代数的运算是连续的.

称测度代数$(\mathcal{A}, \mu)$ 为**可分的**, 是指度量空间$(\mathcal{A}, \rho)$ 为可分的.

定义 4.1.17 设$(\mathcal{B},\nu),(\mathcal{A},\mu)$ 为测度代数, 那么$\Phi:(\mathcal{B},\nu)\to(\mathcal{A},\mu)$ 称为**同态**_(homomorphism)_是指它满足以下条件:

1. $\Phi(B_1\cup B_2)=\Phi(B_1)\cup\Phi(B_2),\forall B_1,B_2\in\mathcal{B}$,

2. $\Phi(B^c)=(\Phi(B))^c,\forall B\in\mathcal{B}$,

3. $\mu(\Phi(B))=\nu(B),\forall B\in\mathcal{B}$.

如果Φ 为一对一的, 那么称$(\mathcal{B},\nu),(\mathcal{A},\mu)$ 为同构的测度代数, Φ称为**同构**.

当$(\mathcal{B},\nu)=(\mathcal{A},\mu)$时, 同构$\Phi:(\mathcal{A},\mu)\to(\mathcal{A},\mu)$ 也称为**自同构**.

注记 4.1.18 1. 测度代数的同构自然为单射, 并且为σ-同态. 证明如下: 设$\Phi(B_1)=\Phi(B_2)$, 那么

$$\Phi(B_1\cup B_2)=\Phi(B_1)\cup\Phi(B_2)=\Phi(B_1)\cap\Phi(B_2)=\Phi(B_1\cap B_2).$$

于是$\nu(B_1\cap B_2)=\nu(B_1\cup B_2)$. 由此得到$\nu(B_1\Delta B_2)=0$, 即$B_1=B_2$. 下证为$\sigma$同态. 设$B=\cup_{n=1}^\infty B_n$, 则

$$\mu\left(\Phi(B)\setminus\bigcup_{n=1}^N\Phi(B_n)\right)=\nu\left(B\setminus\bigcup_{n=1}^N B_n\right)\to 0,\ N\to\infty.$$

于是$\mu(\Phi(B))=\mu(\bigcup_{n=1}^\infty\Phi(B_n))$, 所以$\Phi(B)=\bigcup_{n=1}^\infty\Phi(B_n)$.

2. 易验证: 测度代数的同构为等距的.

设$(X,\mathcal{X},\mu)$ 是一个有限测度空间. 令$\mathfrak{I}_\mu$ 为全体μ-零集的全体. 那么$\mathfrak{I}_\mu$ 是σ 理想. 在$\mathcal{X}_\mu$ 上定义一个等价关系: $A\sim B$ 当且仅当$A\Delta B\in\mathfrak{I}_\mu$. 对任意$A\in\mathcal{B}$, 令$[A]=\{B\in\mathcal{X}\colon A\sim B\}$. 记

$$\widetilde{\mathcal{X}_\mu}=\{[A]\colon A\in\mathcal{X}\}.$$

按照从$\mathcal{X}_\mu$ 上诱导下来的补, 并和交运算, $\widetilde{\mathcal{X}_\mu}$ 是一个Boolean σ-代数. 定义

$$\widetilde{\mu}\colon\widetilde{\mathcal{X}}\to\mathbb{R}_+,[A]\mapsto\mu(A).$$

于是$(\widetilde{\mathcal{X}_\mu},\widetilde{\mu})$ 成为一个测度代数.

注记 4.1.19 可以证明, 任何可分测度代数都同构于某个测度空间$(X,\mathcal{X},\mu)$ 上诱导的测度代数$(\widetilde{\mathcal{X}_\mu},\widetilde{\mu})$. 所以提及测度代数我们可以直接理解为测度空间上得到的测度代数.

§4.1.3 测度空间的同构

设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为两个概率空间, 我们研究什么时候它们是"一样的". 这里涉及三种观点.

定义 4.1.20 (同构) *设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为两个概率空间, 称它们是***同构***(isomorphic) 是指存在$X_0\in\mathcal{X},Y_0\in\mathcal{Y}$ 使得$\mu(X_0)=\nu(Y_0)=1$ 并且存在可逆保测变换*

$$\phi:(X_0,\mathcal{X}\cap X_0,\mu)\to(Y_0,\mathcal{Y}\cap Y_0,\nu).$$

定义 4.1.21 (共轭) *设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为两个概率空间, 如果它们的测度代数是同构的, 即存在$\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 为测度代数同构, 那么称$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为***共轭的***(conjugate).*

一般而言, 两个概率空间为同构的, 那么它们必为共轭的, 但是反之不然.

例 4.1.22 *设*

$$X=\{x\},\mathcal{X}=\{X,\emptyset\},\mu(X)=1,\mu(\emptyset)=0,$$

$$Y=\{y_1,y_2\},\mathcal{Y}=\{Y,\emptyset\},\nu(Y)=1,\nu(\emptyset)=0.$$

那么二者共轭, 但是不同构.

但是根据定理4.1.15, 对于Lebesgue 空间, 同构等价于共轭.

第三种观点是用函数空间来刻画. 设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为两个具有可数基的概率空间, 那么$L^2(X,\mu)$ 与$L^2(Y,\nu)$ 为可分的Hilbert 空间, 于是它们是酉等价的, 即存在双线性算子

$$W:L^2(Y,\nu)\to L^2(X,\mu),\ \langle Wf,Wg\rangle=\langle f,g\rangle,\forall f,g\in L^2(Y,\nu).$$

所以如果不附加条件, 不能有效地区分两个概率空间.

定理 4.1.23 *设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为两个概率空间. 那么二者共轭当且仅当存在双线性算子$W:L^2(Y,\nu)\to L^2(X,\mu)$ 满足以下条件*

1. $\langle Wf,Wg\rangle=\langle f,g\rangle,\forall f,g\in L^2(Y,\nu)$;
2. W,W^{-1} *将有界函数映为有界函数*;
3. $W(fg)=(Wf)(Wg),\forall f,g\in L^\infty(Y,\nu)$.

注记 4.1.24 *设$\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 为测度代数同构. 与之相联系的线性算子$W:L^2(Y,\nu)\to L^2(X,\mu)$满足*

$$W(\mathbf{1}_B)=\mathbf{1}_{\Phi([B])},\ \forall B\in\mathcal{Y}.$$

因为L^2空间中元素已经是模去零测集的等价类, 所以在这里我们不区分$\mathbf{1}_B$ 和$\mathbf{1}_{[B]}$.

证明. 首先设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 共轭, $\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 为测度代数同构. 首先令

$$W(\mathbf{1}_B)=\mathbf{1}_{\Phi([B])},\ \forall B\in\mathcal{Y}.$$

注意此时

$$\|W(\mathbf{1}_B)\|_2=\|\mathbf{1}_B\|_2,\forall B\in\mathcal{Y}.$$

接着直接将定义延拓到简单函数上. 对于任何$f\in L^2(Y,\nu)$, 取简单函数$\{f_n\}_{n\in\mathbb{N}}$ 使得$\|f_n-f\|_2\to 0,n\to\infty$. 于是$\{f_n\}_{n\in\mathbb{N}}$ 为$L^2(Y,\nu)$ 的Cauchy 列, 根据等距性, $\{W(f_n)\}_{n\in\mathbb{N}}$ 为$L^2(X,\mu)$ 的Cauchy 列, 所以存在极限$W(f)$. 下说明定义合理. 设$\{g_n\}_{n\in\mathbb{N}}$ 为另一列简单函数使得$\|g_n-f\|_2\to 0,n\to\infty$. 那么$\|f_n-g_m\|_2\to 0,n,m\to\infty$, 根据等距性, $\|Wf_n-Wg_m\|_2\to 0,n,m\to\infty$, 所以$W(f)$ 定义唯一. 接着对于Φ^{-1} 定义W^{-1}. 性质(1)(2)(3)是容易验证的. 例如(3):

$$W(\mathbf{1}_A\mathbf{1}_B)=W(\mathbf{1}_{A\cap B})=\mathbf{1}_{\Phi([A\cap B])}=\mathbf{1}_{\Phi([A])\cap\Phi([B])}=\mathbf{1}_{\Phi([A]}\mathbf{1}_{\Phi([B]}=W(\mathbf{1}_A)W(\mathbf{1}_B).$$

然后再拓广到$L^2(Y,\nu)$.

反过来, 我们设存在双线性算子$W:L^2(Y,\nu)\to L^2(X,\mu)$ 满足所列三个条件. 我们下面定义$\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 使得

$$W(\mathbf{1}_B)=\mathbf{1}_{\Phi([B])},\ \forall B\in\mathcal{Y}.$$

设$B\in\mathcal{Y}$, 则$\mathbf{1}_B^2=\mathbf{1}_B$. 于是根据性质(3), 有

$$W(\mathbf{1}_B)=W(\mathbf{1}_B^2)=W(\mathbf{1}_B)W(\mathbf{1}_B).$$

所以$W(\mathbf{1}_B)$ 取值1和0, 它为特征函数, 即存在$A\in\mathcal{X}$ 使得

$$W(\mathbf{1}_B)=\mathbf{1}_A.$$

我们令

$$\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu}),\ \Phi([B])=[A].$$

对于W^{-1}, 我们同样可以定义$\Psi:(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})\to(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})$, 然后验证$\Psi$为$\Phi$的逆. 另外我们有

$$\widetilde{\nu}([B])=\langle\mathbf{1}_B,\mathbf{1}_B\rangle=\langle W(\mathbf{1}_B),W(\mathbf{1}_B)\rangle=\langle\mathbf{1}_{\Phi([B])},\mathbf{1}_{\Phi([B])}\rangle=\widetilde{\mu}(\Phi([B])),\forall B\in\mathcal{Y}.$$

下面我们需要验证Φ保持补和有限并(由注记4.1.18, 保持可数运算). 根据W保距以及将特征函数映为特征函数, 易见$W(\mathbf{1})=\mathbf{1}$. 于是由$\mathbf{1}_B+\mathbf{1}_{Y\setminus B}=\mathbf{1}$, 我们两边作用$W$ 有

$$\mathbf{1}_{\Phi([B])}+\mathbf{1}_{\Phi([Y\setminus B])}=\mathbf{1}.$$

所以

$$[X]\setminus\Phi([B])=\Phi([Y]\setminus[B]),$$

即$\Phi([B]^c) = (\Phi([B]))^c,\ \forall B \in Y$.

设$B, C \in \mathcal{Y}$, 则

$$\mathbf{1}_{B\cup C} = \mathbf{1}_B + \mathbf{1}_C - \mathbf{1}_{B\cap C} = \mathbf{1}_B + \mathbf{1}_C - \mathbf{1}_B\mathbf{1}_C.$$

两边作用W得到

$$\mathbf{1}_{\Phi([B\cup C])} = \mathbf{1}_{\Phi([B])} + \mathbf{1}_{\Phi([C])} - \mathbf{1}_{\Phi([B])}\mathbf{1}_{\Phi([C])} = \mathbf{1}_{\Phi([B])\cup\Phi([C])}.$$

于是

$$\Phi([B]\cup[C]) = \Phi([B])\cup\Phi([C]).$$

所以Φ为测度代数同构. □

习　题

1. 设$(\mathcal{A}, \mu)$ 为测度代数, 证明：(1) $\rho(A, B) = \mu(A\Delta B)$ 为度量, 并且$(\mathcal{A}, \rho)$ 为完备度量空间, 以及$\mu : \mathcal{A} \to \mathbb{R}_+$ 为一致连续的函数; (2) $A \mapsto A^c$ 为等距映射, 以及对于任何$A, B, C, D \in \mathcal{A}$: $\rho(A\cup B, C\cup D) + \rho(A\cap B, C\cap D) \le \rho(A, C) + \rho(B, D)$.
2. 证明：测度代数的同构为等距的.

§4.2　保测系统的同构、共轭与谱同构

有几种方式来定义保测系统是否“一样”，我们在本节分别进行讨论.

§4.2.1　保测系统的同构

保测的系统的同构我们在第一章已经定义过，我们在此次再回顾一下. 设$(X, \mathcal{X}, \mu, T)$和$(Y, \mathcal{Y}, \nu, S)$ 为两个保测系统. 如果存在$X_0 \in \mathcal{X}, Y_0 \in \mathcal{Y}$ 使得

$$\mu(X_0) = \nu(Y_0) = 1,\ TX_0 \subseteq X_0, SY_0 \subseteq Y_0$$

以及存在可逆可测保测变换

$$\phi : (X_0, \mathcal{X}\cap X_0, \mu) \to (Y_0, \mathcal{Y}\cap Y_0, \nu)$$

满足$\phi\circ T(x) = S\circ\phi(x), \forall x \in X_0$, 即

$$\begin{array}{ccc} X_0 & \xrightarrow{T} & X_0 \\ \phi\downarrow & & \downarrow\phi \\ Y_0 & \xrightarrow[S]{} & Y_0 \end{array}$$

那么我们称$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 为**同构**(isomorphic), 或者直接称T **同构于** S.

如果ϕ 仅是保测变换, 那么我们称ϕ 为**因子映射** 或**同态**. 此时称$(Y, \mathcal{Y}, \nu, S)$ 为$(X, \mathcal{X}, \mu, T)$的**因子**, 称$(X, \mathcal{X}, \mu, T)$ 为$(Y, \mathcal{Y}, \nu, S)$ 的**扩充**.

注记 4.2.1 *1.* 同构是等价关系.

2. 如果T与S同构, 那么对于任何$n\in\mathbb{N}$, T^n 与S^n 同构.

3. 如果$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为可逆动力系统, 我们可以取X_0,Y_0 使得

$$TX_0=X_0, SY_0=Y_0.$$

设$X_0'\in\mathcal{X},Y_0'\in Y$ 满足$TX_0'\subseteq X_0,SY_0'\subseteq Y_0'$, 取$X_0=\bigcap_{n=-\infty}^{\infty}T^nX_0',Y_0=\bigcap_{n=-\infty}^{\infty}Y_0'$ 就满足要求.

4. 根据定义, 两个系统同构只需要在一个全测集上"一样", 这个观点在测度动力系统研究中非常重要. 例如, 我们定义一个测度动力系统, 可以忽略一个零测集.

§4.2.2 保测系统的共轭

设$(\mathfrak{X},\mu^*)$ 是一个可分的测度代数, 设$T^*:\mathfrak{X}\to\mathfrak{X}$ 为同构, 那么称$(\mathfrak{X},\mu^*,T^*)$ 为**测度代数动力系统**(measure algebra dynamical system). 设$(\mathfrak{Y},\nu^*,S^*)$ 为另一个测度代数系统, $\Phi:(\mathfrak{Y},\nu^*,S^*)\to(\mathfrak{X},\mu^*,T^*)$ 为**同构(同态)**是指Φ 为测度代数同构（同态）并且下图表交换

$$\begin{tikzcd}
(\mathfrak{X},\mu^*) & (\mathfrak{X},\mu^*) \arrow[l, "T^*"'] \\
(\mathfrak{Y},\nu^*) \arrow[u, "\Phi"] & (\mathfrak{Y},\nu^*), \arrow[l, "S^*"] \arrow[u, "\Phi"']
\end{tikzcd}$$

即$\Phi\circ S^*=T^*\circ\Phi$.

对于$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为概率空间, 我们有测度代数$(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 和$(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})$. 设$\pi:X\to Y$ 为保测映射, 我们可以定义测度代数的同态:

$$\widetilde{\pi}^{-1}:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu}),\quad \widetilde{\pi}^{-1}([B])=[\pi^{-1}(B)].$$

设$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 为保测系统, 那么根据上面的定义, 可以定义

$$\widetilde{T}^{-1}:(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu}),\ \text{和}\ \widetilde{S}^{-1}:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu}).$$

于是$(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu},\widetilde{T}^{-1}),(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu},\widetilde{S}^{-1})$ 为测度代数系统.

定义 4.2.2 (保测系统共轭) 设$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为两个保测系统. 如果存在测度代数系统同构

$$\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu},\widetilde{S}^{-1})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu},\widetilde{T}^{-1}),$$

那么我们称$(X,\mathcal{X},\mu,T)$ **共轭于***(conjugate)* $(Y,\mathcal{Y},\nu,S)$, 或称$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$为共轭的.

如果Φ 为测度代数系统同态, 那么称$(Y,\mathcal{Y},\nu,S)$ **半共轭于***(semi-conjugate)* $(X,\mathcal{X},\mu,T)$, 或称$(Y,\mathcal{Y},\nu,S)$ 为$(X,\mathcal{X},\mu,T)$ 的**半共轭像**.

注记 4.2.3 根据定义, 两个系统是同构的, 那么它们为共轭的, 反之不然. 下面的结论指出, 对于*Lebesgue* 空间二者等价.

设$\phi: X \to Y$ 为因子映射, 那么它诱导了测度代数系统的同态

$$\widetilde{\phi}^{-1}: (\widetilde{\mathcal{Y}}_\nu, \widetilde{\nu}, \widetilde{S}^{-1}) \to (\widetilde{\mathcal{X}}_\mu, \widetilde{\mu}, \widetilde{T}^{-1}),$$

即有图表:

$$\begin{tikzcd}
(\widetilde{\mathcal{X}}_\mu, \widetilde{\mu}) & (\widetilde{\mathcal{X}}_\mu, \widetilde{\mu}) \arrow[l, "\widetilde{T}^{-1}"'] \\
(\widetilde{\mathcal{Y}}_\nu, \widetilde{\nu}) \arrow[u, "\widetilde{\phi}^{-1}"] & (\widetilde{\mathcal{Y}}_\nu, \widetilde{\nu}). \arrow[l, "\widetilde{S}^{-1}"] \arrow[u, "\widetilde{\phi}^{-1}"']
\end{tikzcd}$$

定理 4.2.4 设$(\mathfrak{X}, \mu^*, T^*)$, $(\mathfrak{Y}, \nu^*, S^*)$ 为测度代数系统, $\Phi: (\mathfrak{Y}, \nu^*, S^*) \to (\mathfrak{X}, \mu^*, T^*)$ 为同构. 那么存在保测系统$(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 以及同构$\phi: (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ 使得如下图表交换:

$$\begin{tikzcd}
(\mathfrak{X}, \mu^*, T^*) & (\widetilde{\mathcal{X}}_\mu, \widetilde{\mu}, \widetilde{T}^{-1}) \arrow[l, "\alpha"'] \\
(\mathfrak{Y}, \nu^*, S^*) \arrow[u, "\Phi"] & (\widetilde{\mathcal{Y}}_\nu, \widetilde{\nu}, \widetilde{S}^{-1}), \arrow[l, "\beta"] \arrow[u, "\widetilde{\phi}^{-1}"']
\end{tikzcd}$$

其中

$$\alpha: (\widetilde{\mathcal{X}}_\mu, \widetilde{\mu}, \widetilde{T}^{-1}) \to (\mathfrak{X}, \mu^*, T^*), \quad \beta: (\widetilde{\mathcal{Y}}_\nu, \widetilde{\nu}, \widetilde{S}^{-1}) \to (\mathfrak{Y}, \nu^*, S^*)$$

为同构.

定理 4.2.5 设$(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 为保测系统, 其中$(X, \mathcal{X}, \mu), (Y, \mathcal{Y}, \nu)$ 为*Lebesgue*空间. 设$\Psi: (\widetilde{\mathcal{Y}}_\nu, \widetilde{\nu}, \widetilde{S}^{-1}) \to (\widetilde{\mathcal{X}}_\mu, \widetilde{\mu}, \widetilde{T}^{-1})$ 为测度代数系统的同构, 那么存在$X_0 \in \mathcal{X}, Y_0 \in \mathcal{Y}$ 使得

$$\mu(X_0) = \nu(Y_0) = 1, \ TX_0 \subseteq X_0, \ SY_0 \subseteq Y_0$$

以及*Borel* 同构

$$\psi: (X_0, \mu, T) \to (Y_0, \nu, S)$$

使得

$$\Psi = \widetilde{\psi}^{-1}.$$

注记 4.2.6 根据定理*4.2.5*, 定义保测系统, 空间是相对次要的, 重要的是测度代数上的动力系统.

定理 4.2.7 设$(X, \mathcal{X}, \mu, T)$ 为*Lebesgue* 空间上的保测系统. 对于任何T 不变子σ 代数$\mathcal{F} \subseteq \mathcal{X}$ *(*即$T^{-1}\mathcal{F} \subseteq \mathcal{F}$*)*, 存在因子系统$(Y, \mathcal{Y}, \nu, S)$ 以及因子映射$\pi: (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ 使得$\mathcal{F} = \pi^{-1}(\mathcal{Y})$.

因子系统$(Y, \mathcal{Y}, \nu, S)$ 是可逆的当且仅当子σ 代数$\mathcal{F} \subseteq \mathcal{X}$ 为严格T 不变的*(*即$T^{-1}\mathcal{F} = \mathcal{F}$*)*.

证明. 设$\mathcal{F}_0=\{F_n\}_{n\in\mathbb{N}}$ 为生成$\mathcal{F}$ 的可数T 不变代数. 令

$$\pi: X\to\{0,1\}^{\mathbb{N}},\quad (\pi x)_n=\mathbf{1}_{F_n}(x).$$

易见

$$\pi^{-1}({}_1[a_1,a_2,\ldots,a_n]_n)=\bigcap_{k=1}^{n}F_{a_k}\in X,$$

所以π 为可测的, 且$\pi^{-1}:\mathcal{B}(\{0,1\}^{\mathbb{N}})\cap\pi(X)\to\mathcal{F}$ 的布尔同构. 设$Y=\pi(X)$ 定义概率测度

$$\nu:\mathcal{B}(\{0,1\}^{\mathbb{N}})\cap Y\to[0,1],\quad \nu=\mu\circ\pi^{-1}.$$

易见, $\pi(x)=\pi(y)$ 当且仅当$\mathbf{1}_F(x)=\mathbf{1}_F(y)$, $\forall F\in\mathcal{F}$. 根据$\mathcal{F}$ 为T 不变的, $\pi(x)=\pi(y)$ 蕴含$\pi(Tx)=\pi(Ty)$. 于是我们可以定义变换$S:Y\to Y$ 为

$$S(\pi(x)):=\pi(Tx).$$

因为对于任何$A\in\mathcal{B}(\{0,1\}^{\mathbb{N}})$,

$$S^{-1}(A\cap Y)=S^{-1}\pi(\pi^{-1}A)=\pi(T^{-1}\pi^{-1}A)\in\mathcal{B}(\{0,1\}^{\mathbb{N}})\cap Y,$$

所以S 可测. 根据定理4.1.8, $Y=\pi(X)\in\mathcal{B}_\nu(\{0,1\}^{\mathbb{N}})$, 所以存在$Y_0\in\mathcal{B}(\{0,1\}^{\mathbb{N}})$ 使得$Y_0\subseteq Y$ 且$\nu\big(\{0,1\}^{\mathbb{N}}\setminus Y_0\big)=0$.

下证明第二部分. 我们证明: 设$(X,\mathcal{X},\mu,T)$ 为标准概率空间上的保测系统, 如果$T^{-1}\mathcal{X}=\mathcal{X}$, 那么$T$ 可逆. 设可数集$\mathcal{A}\subseteq\mathcal{X}$ 生成$\mathcal{X}$, 且分离点的子集:

$$\mathbf{1}_A(x)=\mathbf{1}_A(y),\ \forall A\in\mathcal{A}\ \Rightarrow\ x=y.$$

我们假设$T^{-1}\mathcal{A}=\mathcal{A}$, 否则我们可以扩大$\mathcal{A}$ 为一个满足条件的子集族:

$$\bigcup_{n=0}^{\infty}T^{-n}\mathcal{A}\cup\{A\in\mathcal{X}:\exists n\geq 1\ s.t.\ T^{-n}A\in\mathcal{A}\}.$$

设

$$S:\{0,1\}^{\mathcal{A}}\to\{0,1\}^{\mathcal{A}},\quad (Sx)_A=x_{T^{-1}A}.$$

明显地, S 为$\{0,1\}^{\mathcal{A}}$ 上的同胚. 定义

$$\pi: X\to\{0,1\}^{\mathcal{A}},\quad \pi(x)_A=\mathbf{1}_A(x).$$

易见π 为可测的, 一对一, 且$\pi\circ T=S\circ\pi$. 令

$$\nu=\mu\circ\pi^{-1}.$$

于是

$$\pi:(X,\mu,T)\to(\{0,1\}^{\mathcal{A}},\nu,S)$$

为同构. 证毕! □

根据定理4.2.7, 可以将$(X,\mathcal{X},\mu,T)$ 的因子与$\mathcal{X}$ 的T不变子σ 代数(即满足$T^{-1}\mathcal{A}\subseteq\mathcal{A}$ 的子σ 代数)可以等同视之.

推论 4.2.8 *设$(X,\mathcal{X},\mu,T)$为保测系统, 那么T是可逆的(忽略零测集意义下)当且仅当$\widetilde{T}^{-1}\widetilde{\mathcal{X}}_\mu=\widetilde{\mathcal{X}}_\mu$.*

§4.2.3 动力系统谱同构

最后我们介绍谱同构.

定义 4.2.9 *设$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 为两个保测系统. 如果存在线性算子$W:L^2(Y,\nu)\to L^2(X,\mu)$ 满足以下条件*

1. *W可逆,*
2. *$\langle Wf,Wg\rangle=\langle f,g\rangle,\forall f,g\in L^2(Y,\nu)$,*
3. *$U_TW=WU_S$,*

$$\begin{array}{ccc} L^2(X,\mu) & \xrightarrow{U_T} & L^2(X,\mu) \\ {\scriptstyle W}\uparrow & & \uparrow{\scriptstyle W} \\ L^2(Y,\nu) & \xrightarrow[U_S]{} & L^2(Y,\nu) \end{array}$$

*那么我们称$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 为***谱同构***(spectrally isomorphic).*

注记 4.2.10 *在定义中(1)(2)就是指W为Hilbert空间的同构, 而(3)是加入动力系统因素.*

下面定理指出了谱同构与共轭的关系.

定理 4.2.11 *设$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 为两个保测系统. 那么T 与S 为共轭的当且仅当如果存在线性算子$W:L^2(Y,\nu)\to L^2(X,\mu)$ 满足以下条件*

1. *W可逆,*
2. *$\langle Wf,Wg\rangle=\langle f,g\rangle,\forall f,g\in L^2(Y,\nu)$,*
3. *$U_TW=WU_S$,*

$$\begin{array}{ccc} L^2(X,\mu) & \xrightarrow{U_T} & L^2(X,\mu) \\ {\scriptstyle W}\uparrow & & \uparrow{\scriptstyle W} \\ L^2(Y,\nu) & \xrightarrow[U_S]{} & L^2(Y,\nu) \end{array}$$

4. *W,W^{-1} 将有界函数映为有界函数,*

5. $W(fg)=(Wf)(Wg), \forall f,g\in L^\infty(Y,\nu)$,

即T,S为谱同构并且满足*(4)(5)*.

证明. 设$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 共轭, $\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 为测度代数动力系统同构. 根据定理4.1.23证明的构造, 我们得到$W:L^2(Y,\nu)\to L^2(X,\mu)$ 满足:

$$W(\mathbf{1}_B)=\mathbf{1}_{\Phi([B])},\ \forall B\in\mathcal{Y}.$$

根据

$$U_TW(\mathbf{1}_B)=U_T(\mathbf{1}_{\Phi([B])})=\mathbf{1}_{\widetilde{T}^{-1}\Phi([B])}=\mathbf{1}_{\Phi\widetilde{S}^{-1}([B])}=W(\mathbf{1}_{\widetilde{S}^{-1}([B])})=WU_S(\mathbf{1}_B),\forall B\in\mathcal{Y}.$$

于是得到$U_TW=WU_S$.

反之, 我们根据定理4.1.23证明的构造, 得到$\Phi:(\widetilde{\mathcal{Y}}_\nu,\widetilde{\nu})\to(\widetilde{\mathcal{X}}_\mu,\widetilde{\mu})$ 为测度代数同构. 需要验证它满足

$$\widetilde{T}^{-1}\Phi=\Phi\widetilde{S}^{-1}.$$

根据$U_TW(\mathbf{1}_B)=WU_S(\mathbf{1}_B),\forall B\in\mathcal{Y}$, 我们得到

$$\mathbf{1}_{\widetilde{T}^{-1}\Phi([B])}=\mathbf{1}_{\Phi\widetilde{S}^{-1}([B])},\forall B\in\mathcal{Y}.$$

于是得到我们所需. □

根据定义和定理4.2.11, 我们知道共轭蕴含谱等价. 下一节中我们给出例子说明谱等价远远弱于共轭. 我们将给了一大类系统它们相互是谱同构的, 但是它们相互不是共轭的.

习 题

1. 如果$(X,\mathcal{X},\mu,T)$ 为保测系统, 那么$U_T:L^2(X,\mu)\to L^2(X,\mu)$ 为满射当且仅当$\widetilde{T}^{-1}:\widetilde{\mathcal{X}}_\mu\to\widetilde{\mathcal{X}}_\mu$ 为满射, 即$U_T:L^2(X,\mu)\to L^2(X,\mu)$ 为酉算子当且仅当$\widetilde{T}^{-1}:\widetilde{\mathcal{X}}_\mu\to\widetilde{\mathcal{X}}_\mu$ 为同构. (因为U_T等距, 为单射.)

§4.3 Kolmogorov 系统

为了说明谱同构弱于同构, 在本节我们介绍两类很重要的系统：Kolmogorov 系统和Bernouli 系统. 我们将说明所有Kolmogorov 系统为谱同构的, 因为Kolmogorov 系统中有许多相互不同构的例子, 由此说明谱同构弱于同构. Kolmogorov 系统和Bernouli 系统是非常重要的两类动力系统, 它们在熵理论中具有极其重要的地位.

1958 年Kolmogorov 借鉴Shannon 在信息论中不确定性的描述在遍历论中引入了熵的概念[133]. 熵是重要的同构不变量, 它反映了系统的混乱程度. 对于正熵系统, Kolmogorov 引入了一类正熵保测系统作为正规等分布随机过程的抽象化[133]. Rohlin 和Sinai 将这类保测系统称为Kolmogorov 系统或K系统, 他们考虑了具有平凡Pinsker σ-代数的保测系

统(现称之为完全正熵系统)并证明了完全正熵系统与Kolmogorov 系统是一致的[169]. 由于Bernoulli 系统是独立等分布随机过程, 所以Bernoulli 系统是Kolmogorov系统. 从1958年到1969年期间, 是否每个测度Kolmogorov 系统一定为Bernoulli 系统一直为公开的问题. 1970 年Ornstein 首先构造了一个Lebesgue 空间上的非Bernoulli 的Kolmogorov 系统, 解决了这一问题. Katok 构造了光滑的非Bernoulli 的Kolmogorov 系统[122]. 1982 年Kalikow [121] 给出了一个更容易验证的例子. Kolmogorov 系统具有非常好的回复属性和谱属性: Kolmogorov 系统是一致强混合系统; Kolmogorov 和Rohlin 证明了Lebesgue 空间上的Kolmogorov 系统具有可数Lebesgue 谱, 从而所有Lebesgue 空间上的Kolmogorov 系统是谱同构的. 1967 年Furstenberg 在遍历理论与拓扑动力系统中引入不交性的概念来研究系统之间的差异[62], Furstenberg 证明了Kolmogorov 系统不交于遍历的零熵系统, 由此可见Kolmogorov 系统是完全不同于零熵系统的. 关于Kolmogorov 系统的研究多年来一直是遍历理论中的热点.

§4.3.1 Kolmogorov系统和Bernouli系统

下面我们给出Kolmogorov 系统和Bernouli 系统的定义.

定义 4.3.1 (Kolmogorov 系统) 设$(X,\mathcal{X},\mu,T)$ 为非平凡可逆*Lebesgue* 系统. 如果存在子σ-代数$\mathcal{A}$ 使得满足以下条件:

1. $\mathcal{A}\subseteq T\mathcal{A}$,
2. $\bigvee_{n=0}^{+\infty}T^n\mathcal{A}=\mathcal{X}$,
3. $\bigcap_{n=0}^{+\infty}T^{-n}\mathcal{A}=\mathcal{N}=\{X,\emptyset\}$.

那么我们称$(X,\mathcal{X},\mu,T)$ 为***Kolmogorov* 系统**, 简称为***K*-系统**.

定义 4.3.2 (Bernouli系统) 设$(Y,\mathcal{Y},\nu)$为概率空间, 设

$$(X,\mathcal{X},\mu)=\prod_{n=-\infty}^{\infty}(Y,\mathcal{Y},\nu)=(Y,\mathcal{Y},\nu)^{\mathbb{Z}}$$

为乘积系统, 定义

$$T:X\to X,\ T(\mathbf{x})_n=x_{n+1},\ \forall\mathbf{x}=(x_j)_{j=-\infty}^{\infty},n\in\mathbb{Z}$$

为转移映射, 那么称$(X,\mathcal{X},\mu,T)$ 为***Bernouli* 系统** 或***Bernouli* 转移**, 简称为***B*-系统**.

注记 4.3.3 *1.* 符号系统为B-系统.

2. B-系统乘积仍为B-系统.

定理 4.3.4 任何*Bernouli* 系统为*Kolmogorov* 系统.

证明. 设$(Y,\mathcal{Y},\nu)$ 为B-系统$(X,\mathcal{X},\mu,T)$ 的底空间. 对于$F\in\mathcal{Y}$, 令

$$\widetilde{F}=\{(x_n)_{n\in\mathbb{Z}}:x_0\in F\}\in\mathcal{X}.$$

设$\mathcal{G}=\{\widetilde{F}:F\in\mathcal{Y}\}$, 令

$$\mathcal{A}=\bigvee_{i=-\infty}^{0}T^i\mathcal{G}.$$

则

(1) $\mathcal{A}=\bigvee\limits_{i=-\infty}^{0}T^i\mathcal{G}\subseteq\bigvee\limits_{i=-\infty}^{1}T^i\mathcal{G}=T\mathcal{A}$.

(2) $\bigvee\limits_{n=0}^{\infty}T^n\mathcal{A}=\bigvee\limits_{n=0}^{\infty}\bigvee\limits_{i=-\infty}^{n}T^i\mathcal{G}=\bigvee\limits_{i=-\infty}^{\infty}T^i\mathcal{G}=\mathcal{X}$.

(3) 下证明$\bigcap\limits_{n=0}^{+\infty}T^{-n}\mathcal{A}=\mathcal{N}=\{X,\emptyset\}$. 设

$$A\in\bigcap_{n=0}^{+\infty}T^{-n}\mathcal{A}=\bigcap_{n=0}^{\infty}\bigvee_{i=-\infty}^{-n}T^i\mathcal{G}.$$

设$B\in\bigvee\limits_{k=j}^{\infty}T^k\mathcal{G}$, 其中$j\in\mathbb{Z}$. 因为$A\in\bigvee\limits_{i<j}T^i\mathcal{G}$, 所以$A,B$ 为独立的, 即$\mu(A\cap B)=\mu(A)\mu(B)$.
令

$$\mathcal{M}=\{B\in\mathcal{X}:\mu(A\cap B)=\mu(A)\mu(B)\}.$$

易验证$\mathcal{M}$ 为单调类, 根据上面分析, $\mathcal{M}\supset\bigvee\limits_{k=j}^{\infty}T^k\mathcal{G}$, $\forall j\in\mathbb{Z}$. 即

$$\mathcal{M}\supset\bigcup_{j=-\infty}^{\infty}\bigvee_{k=j}^{\infty}T^k\mathcal{G}.$$

所以$\mathcal{M}=\mathcal{X}$. 于是

$$\mu(A\cap B)=\mu(A)\mu(B),\forall B\in\mathcal{X}.$$

令$B=A$, 则$\mu(A)=\mu(A)^2$. 由此$\mu(A)=0$ 或$\mu(A)=1$. 即$\bigcap\limits_{n=0}^{+\infty}T^{-n}\mathcal{A}=\mathcal{N}=\{X,\emptyset\}$. □

例 4.3.5 (Kalikow) 下面例子是K-系统但不是B-系统.

设$(\Sigma_2,\mathcal{B},\mu_{(\frac12,\frac12)},\sigma)$为双边$(\frac12,\frac12)$-转移, 我们令$X=\Sigma_2\times\Sigma_2$, 定义

$$T:X\to X,\ (\mathbf{x},\mathbf{y})\mapsto(\sigma\mathbf{x},\sigma^{\alpha(\mathbf{x})}\mathbf{y}),$$

其中$\mathbf{x},\mathbf{y}\in\Sigma_2$, $\alpha(\mathbf{x})=\begin{cases}-1, & x_0=0;\\ 1, & x_0=1.\end{cases}$ 则可证明$(X,\mathcal{X},\mu,T)$ 是K-系统但不是B-系统.

§4.3.2 可数Lebesgue谱

定义 4.3.6 设$(X,\mathcal{X},\mu,T)$ 为可逆*Lebesgue* 系统. 称T 具有**可数*Lebesgue*谱**(*countable Lebesgue spectrum*) 如果存在函数列$\{f_j\}_{j=0}^{\infty}\subseteq L^2(X,\mu)$ 使得$f_0\equiv 1$ 并且

$$\{f_0\}\cup\{U_T^n f_j : j\in\mathbb{N}, n\in\mathbb{Z}\}$$

为$L^2(X,\mu)$ 的一组正交基.

$$\begin{array}{ccccccc} & & & f_0=1 & & & \\ \dots & U_T^{-2}f_1 & U_T^{-1}f_1 & f_1 & U_Tf_1 & U_T^2f_1 & \dots \\ \dots & U_T^{-2}f_2 & U_T^{-1}f_2 & f_2 & U_Tf_2 & U_T^2f_2 & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{array}$$

定理 4.3.7 任何具有可数*Lebesgue* 谱的可逆*Lebesgue* 系统是强混合的.

证明. 设$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 具有可数Lebesgue 谱, 那么存在函数列$\{f_j\}_{j=0}^{\infty}\subseteq L^2(X,\mu)$ 使得$f_0\equiv 1$ 并且$\{f_0\}\cup\{U_T^n f_j : j\in\mathbb{N}, n\in\mathbb{Z}\}$ 为$L^2(X,\mu)$ 的一组正交基. 因为$\{f_0\}\cup\{U_T^n f_j : j\in\mathbb{N}, n\in\mathbb{Z}\}$ 是正交的, 于是对于任何$j,q\geq 0$ 下式是显然成立的:

$$\lim_{p\to\infty}\langle U_T^pU_T^nf_j,U_T^kf_q\rangle=\langle U_T^nf_j,1\rangle\langle 1,U_T^kf_q\rangle,\forall k,n\in\mathbb{Z}.$$

固定k,q, 令

$$\mathcal{H}_{k,q}=\left\{f\in L^2(X,\mu):\lim_{p\to\infty}\langle U_T^pf,U_T^kf_q\rangle=\langle f,1\rangle\langle 1,U_T^kf_q\rangle\right\}.$$

因为$\{f_0\}\cup\{U_T^n f_j : j\in\mathbb{N}, n\in\mathbb{Z}\}\subseteq\mathcal{H}_{k,q}$, 以及$\mathcal{H}_{k,q}$ 为闭子空间, 所以$\mathcal{H}_{k,q}=L^2(X,\mu)$. 对于取定的$f\in L^2(X,\mu)$,

$$\mathcal{H}_f=\left\{g\in L^2(X,\mu):\lim_{p\to\infty}\langle U_T^pf,g\rangle=\langle f,1\rangle\langle 1,g\rangle\right\}.$$

同样, $\{f_0\}\cup\{U_T^n f_j : j\in\mathbb{N}, n\in\mathbb{Z}\}\subseteq\mathcal{H}_f$, 以及$\mathcal{H}_f$ 为闭子空间, 所以$\mathcal{H}_f=L^2(X,\mu)$. 于是我们得到

$$\lim_{p\to\infty}\langle U_T^pf,g\rangle=\langle f,1\rangle\langle 1,g\rangle,\forall f,g\in L^2(X,\mu).$$

即$(X,\mathcal{X},\mu,T)$ 为强混合的. □

定理 4.3.8 任何两个具有可数*Lebesgue* 谱的可逆*Lebesgue* 系统是谱同构的.

证明. 设$(X,\mathcal{X},\mu,T),(Y,\mathcal{Y},\nu,S)$ 具有可数Lebesgue 谱, 那么存在函数列$\{f_j\}_{j=0}^{\infty}\subseteq L^2(X,\mu)$ 使得$f_0\equiv 1$ 并且$\{f_0\}\cup\{U_T^n f_j : j\in\mathbb{N}, n\in\mathbb{Z}\}$ 为$L^2(X,\mu)$ 的一组正交基, 以及存在函数列$\{g_j\}_{j=0}^{\infty}\subseteq L^2(Y,\nu)$ 使得$g_0\equiv 1$ 并且$\{g_0\}\cup\{U_S^n g_j : j\in\mathbb{N}, n\in\mathbb{Z}\}$ 为$L^2(Y,\nu)$ 的一组正交基.

定义$W:L^2(Y,\nu)\to L^2(X,\mu)$ 使得

$$g_0\mapsto f_0;\quad U_S^ng_j\mapsto U_T^nf_j,\ \forall n\in\mathbb{Z},j\in\mathbb{N}.$$

因为$\{g_0\}\cup\{U_S^n g_j : j\in\mathbb{N}, n\in\mathbb{Z}\}$ 是一组基, 所以上面W 可定义. 容易验证$WU_S=U_TW$, T 与S 为谱同构的. □

§4.3.3 Kolmogorov系统与可数Lebesgue谱

定理 4.3.9 (Rohlin) *Kolmogorov 系统具有可数Lebesgue谱.*

证明. 设$(X, \mathcal{X}, \mu, T)$ 为K-系统, 则它为非平凡可逆Lebesgue 系统（此时$\mathcal{X} \neq \mathcal{N}$）, 且存在子$\sigma$- 代数$\mathcal{A}$ 使得满足以下条件：

1. $\mathcal{A} \subseteq T\mathcal{A}$,
2. $\bigvee_{n=0}^{+\infty} T^n \mathcal{A} = \mathcal{X}$,
3. $\bigcap_{n=0}^{+\infty} T^{-n} \mathcal{A} = \mathcal{N} = \{X, \emptyset\}$.

下面我们证明它具有可数Lebesgue 谱.

(1) 首先证明$\mathcal{A}$ 中无原子, 即如果$C \in \mathcal{A}$, $\mu(C) > 0$, 那么存在$D \in \mathcal{A}$ 使得$D \subseteq C$ 并且$\mu(D) < \mu(C)$.

否则设存在原子$C \in \mathcal{A}$, $\mu(C) > 0$. 那么TC 为$T\mathcal{A}$ 的原子. 因为$\mathcal{A} \subseteq T\mathcal{A}$, 所以要么$TC \subseteq C$ 要么$\mu(C \cap TC) = 0$. 如果$TC \subseteq C$, 由于$\mu(TC) = \mu(C)$, 所以$TC = C$ *a.e.*, 于是

$$C \in \bigcap_{n=0}^{\infty} T^{-n} \mathcal{A} = \mathcal{N}.$$

因为$\mu(C) > 0$, 所以$\mu(C) = 1$. 于是$\mathcal{A} = \mathcal{N}$. 由此得到$\mathcal{X} = \mathcal{N}$, 矛盾. 如果$\mu(TC \cap C) = 0$, 那么要么存在某个$k \in \mathbb{N}$ 使得$T^k C \subseteq C$, 要么对于任何$k \in \mathbb{N}$, $\mu(T^k C \cap C) = 0$. 对于前者, 根据上面同样的讨论得出矛盾, 对于后者$\{T^k C\}_{k \geq 0}$ 互不相交并且每个元素测度一样, 这与$\mu(X) < \infty$ 矛盾. 所以不存在原子.

(2) 令

$$H = L^2(X, \mathcal{A}, \mu) = \{f \in L^2(X, \mathcal{X}, \mu) : f\text{为}\mathcal{A}\text{-可测的}\}.$$

则$U_T H \subseteq H$. 设$V \subseteq L^2(X, \mu)$ 使得

$$H = V \bigoplus U_T H.$$

易见$V = L^2(X, \mathcal{A}, \mu) \ominus L^2(X, T^{-1}\mathcal{A}, \mu)$. 因为

$$U_T^{-n} H = \bigoplus_{i=-n}^{m} U_T^i V \bigoplus U_T^{m+1} H, \ \forall n, m \geq 0,$$

我们有

$$L^2(X, \mu) = \bigoplus_{n=-\infty}^{\infty} U_T^n V \bigoplus \mathbb{C},$$

其中$\mathbb{C}$ 表示常值函数全体. 如果能说明V 为无限维的, 那么取它的正交基$\{f_1, f_2, \ldots\}$, 则$\{f_0\} \cup \{U_T^n f_j : j \in \mathbb{N}, n \in \mathbb{Z}\}$ 为$L^2(X, \mu)$ 的一组正交基.

(3) 下面证明V 为无穷维的. 因为$\mathcal{A} \neq T\mathcal{A}$, 所以$V \neq \{0\}$. 设$g \in V, g \neq 0$. 则$G = \{x : g(x) \neq 0\}$ 具有正测度. 因为g 为$\mathcal{A}$可测的, 所以$G \in \mathcal{A}$. 根据(1), $\mathbf{1}_G H = \{\mathbf{1}_G f : f \in H\}$ 为无穷维的. 于是

$$\mathbf{1}_G H = V' \bigoplus \mathbf{1}_G U_T H,$$

其中$V' \subseteq V$. 于是要么V' 为无穷维的, 要么$\mathbf{1}_G U_T H$ 为无穷维的. 如果是前者, $V \supset V'$ 也为无穷维的. 如果是后者, 存在线性无关组$\{\mathbf{1}_G U_T f_n\}_{n\in\mathbb{N}}$, 其中$f_n$ 为H 中有界函数. 于是$\{gU_T f_n\}_{n\in\mathbb{N}}$ 为H 的线性无关组. 对于$f \in H$,

$$\langle gU_T f_n, U_T f\rangle = \langle g, U_T(f\overline{f_n})\rangle = 0,$$

所以$gU_T f_n \in V$. 于是$\{gU_T f_n\}_{n\in\mathbb{N}}$ 为V 的线性无关组, 从而V 为无穷维的. □

根据定理4.3.8和定理4.3.9, 我们知道Kolmogorov 系统是谱同构的, 但是在Kolmogorov系统中有着非常多互相不共轭的系统, 例如$(\{0,1\}^{\mathbb{Z}}, \mu_{(\frac12,\frac12)}, \sigma)$, $(\{0,1,2\}^{\mathbb{Z}}, \mu_{(\frac13,\frac13,\frac13)}, \sigma)$.

最后我们陈述一条K系统的等价刻画, 在后面我们将给出其证明.

定理 4.3.10 *系统为K-系统当且仅当为一致混合的.*

习　题

1. 双边$(\vec{p}, P)$-Markov 转移为K-系统当且仅当P为不可约且非周期的.

§4.4　离散谱系统

我们已知同构蕴含谱同构, 但是反之不然. 但是对于一些系统, 谱同构也可以蕴含同构. 离散谱系统就是这样一类系统, 具有离散谱的系统谱同构是等价于同构的. 我们在本节中将证明这一结果, 并且说明实际上任何遍历离散谱系统同构于群旋转.

§4.4.1　离散谱系统及Halmos-von Neumann定理

首先回顾离散谱系统的定义.

定义 4.4.1 *设$(X, \mathcal{X}, \mu, T)$ 为一个保测系统, 称T 具有***离散谱***(discrete spectrum) 或者***纯点谱** *(pure-point spectrum)是指$L^2(X,\mu)$ 存在一组由T 的特征函数组成的正交基.*

注记 4.4.2 *如果$(X, \mathcal{X}, \mu, T)$ 为离散谱的, 那么显然$U_T : L^2(X,\mu) \to L^2(X,\mu)$ 为满射, 于是$\widetilde{T}^{-1}\widetilde{\mathcal{X}}_\mu = \widetilde{\mathcal{X}}_\mu$. 于是对于Lebesgue 系统, 如果它具有离散谱, 那么它是可逆的.*

引理 4.4.3 *设$(X, \mathcal{X}, \mu)$ 为概率空间, $h \in L^2(X,\mu)$. 那么$h \in L^\infty(X,\mu)$ 当且仅当*

$$h \cdot f \in L^2(X,\mu), \quad \forall f \in L^2(X,\mu).$$

证明. 设$h \in L^\infty(X,\mu)$, 那么存在$c \in \mathbb{R}$ 使得$\mu(\{x : |h(x)| > c\}) = 0$. 于是易见$h \cdot f \in L^2(X,\mu), \forall f \in L^2(X,\mu)$.

反之, 设$h \cdot f \in L^2(X,\mu), \forall f \in L^2(X,\mu)$. 下证$h \in L^\infty(X,\mu)$. 设

$$X_n = \{x \in X : n-1 \le |h(x)| < n\},\ n \in \mathbb{N}.$$

则$\{X_n\}_{n=1}^\infty$ 组成了X 的一个剖分. 令

$$f(x) = \sum_{k \in F} \frac{1}{k\sqrt{\mu(X_k)}} \mathbf{1}_{X_k}(x),$$

其中$F = \{k \in \mathbb{N} : \mu(X_k) \neq 0\}$.

于是

$$\int_X |f|^2 d\mu \le \sum_{k=1}^\infty \frac{1}{k^2} < \infty,$$

但是

$$\int_X |hf|^2 d\mu \ge \sum_{k \in F} \left(\frac{k-1}{k}\right)^2.$$

因为$hf \in L^2(X,\mu)$, 所以$|F| < \infty$, 于是h 有界. □

引理 4.4.4 *设H 为离散交换群, K 为H 的可除子群(divisible subgroup), 即对于任何$k \in K, n \in \mathbb{N}$, 存在$a \in K$ 使得$a^n = k$. 那么存在同态$\phi : H \to K$ 使得$\phi|_K = \mathrm{id}_K$.*

证明. 设

$$\mathcal{A} = \{(M,\psi) : K \le M \le H, \psi : M \to K\text{为同态, 并且}\psi|_K = \mathrm{id}\}.$$

因为$(K, \mathrm{id}_K) \in \mathcal{A}$, $\mathcal{A} \neq \emptyset$. 在$\mathcal{A}$ 中定义序关系如下:

$$(M_1,\psi_1) < (M_2,\psi_2) \ \Leftrightarrow\ M_1 \subseteq M_2, \psi_2|_{M_1} = \psi_1.$$

这是个半序, 并且任何全序集有上界. 根据Zorn 引理, 取此序下极大元(L,ϕ). 下面我们证明$L = H$ 即可.

如果$L \neq H$, 那么取$g \in H \setminus L$. 设$M = \langle g, L\rangle$.

情况1. $\langle g\rangle \cap L = \emptyset$. 则$M$ 中元素唯一的表示为$g^i a, a \in L, i \in \mathbb{Z}$. 此时令

$$\psi : M \to K,\ \psi(g^i a) = \phi(a).$$

易验证$(M,\psi) \in \mathcal{A}$ 并且$(L,\phi) < (M,\psi)$, 与(L,ϕ) 极大性矛盾.

情况2. $\langle g\rangle \cap L \neq \emptyset$. 取最小的自然数$n$ 使得$g^n \in L$. 则M 中元素唯一的表示为$g^i a, a \in L, i \in \{0,1,\dots,n-1\}$. 因为$K$ 为可除的, 取$g_0 \in K$ 使得$\phi(g^n) = g_0^n$. 令

$$\psi : M \to K,\ \psi(g^i a) = g_0^i \phi(a).$$

易验证$(M,\psi)\in\mathcal{A}$ 并且$(L,\phi)<(M,\psi)$, 与(L,ϕ) 极大性矛盾.

所以$L=H$. 证毕! □

注意如果两个系统是谱同构的, 那么它们具有相同的特征值集合. 对于离散谱系统反之也成立.

定理 4.4.5 (Halmos-von Neumann 定理, 1942) *设$(X,\mathcal{X},\mu,T)$, $(Y,\mathcal{Y},\nu,S)$ 为具有离散谱的遍历系统. 那么以下等价:*

1. *T 与S 为谱同构的.*
2. *T 的全体特征值的集合Λ_T 与S 的全体特征值集合Λ_S 相等, 即$\Lambda_T=\Lambda_S$.*
3. *T 与S 共轭.*

证明. (1) ⇒ (2), (3) ⇒ (1) 是显然的.

(2) ⇒ (1). 设$\Lambda=\Lambda_T=\Lambda_S$. 对于任何$\lambda\in\Lambda$, 取$f_\lambda\in L^2(X,\mu), g_\lambda\in L^2(Y,\nu)$ 使得

$$U_Tf_\lambda=\lambda f_\lambda, U_Sg_\lambda=\lambda g_\lambda,\ |f_\lambda|=|g_\lambda|=1.$$

定义

$$W:L^2(Y,\nu)\to L^2(X,\mu),\ g_\lambda\mapsto f_\lambda$$

再线性扩充. 因为把基映为基, W 为可逆线性算子. 容易验证$WU_S=U_TW$. 所以T 与S 为谱同构的.

(2) ⇒ (3). 同(2) ⇒ (1) 证明, 设$\Lambda=\Lambda_T=\Lambda_S$. 对于任何$\lambda\in\Lambda$, 取$f_\lambda\in L^2(X,\mu), g_\lambda\in L^2(Y,\nu)$ 使得

$$U_Tf_\lambda=\lambda f_\lambda, U_Sg_\lambda=\lambda g_\lambda,\ |f_\lambda|=|g_\lambda|=1.$$

定义

$$W:L^2(Y,\nu)\to L^2(X,\mu),\ g_\lambda\mapsto f_\lambda.$$

那么W 为谱同构. 根据定理4.2.11, 如果我们还能证明W,W^{-1} 将有界函数映为有界函数, 以及$W(fg)=(Wf)(Wg),\forall f,g\in L^\infty(Y,\nu)$, 那么我们就可以推出$T$ 与S 为共轭的.

断言: 我们可以选取$\{f_\lambda\}_{\lambda\in\Lambda}$ 和$\{g_\lambda\}_{\lambda\in\Lambda}$ 使得

$$f_\lambda f_\xi=f_{\lambda\xi},\ g_\lambda g_\xi=g_{\lambda\xi},\ \forall\lambda,\xi\in\Lambda.$$

断言证明. 对于$\lambda,\xi\in\Lambda$,

$$U_Tf_{\lambda\xi}=\lambda\xi f_{\lambda\xi},$$

又有

$$U_T(f_\lambda f_\xi)=U_Tf_\lambda U_Tf_\xi=\lambda\xi(f_\lambda f_\xi).$$

因为特征空间是一维的, 所以存在常数$r(\lambda,\xi)\in\mathbb{S}^1$ 使得

$$f_\lambda f_\xi = r(\lambda,\xi) f_{\lambda\xi} \quad a.e.$$

设H 为全体X 到$\mathbb{S}^1$ 的函数, 那么在逐点相乘下H 为交换群. 令

$$K=\{f_c\in H: c\in\mathbb{S}^1,\ f_c:X\to\mathbb{S}^1, f_c(x)\equiv c\}.$$

则K 为H 子群. 对$c\in\mathbb{S}^1$, 将f_c 与c 等同视之, 于是$K=\mathbb{S}^1$.

根据引理4.4.4, 存在同态$\phi:H\to K$ 使得$\phi|_K=\mathrm{id}_K$. 令

$$f_\lambda^*=\overline{\phi(f_\lambda)}f_\lambda.$$

那么

$$|f_\lambda^*|=1,\ U_T f_\lambda^*=\lambda f_\lambda^*,\ \forall \lambda\in\Lambda.$$

并且

$$\begin{aligned} f_\lambda^* f_\xi^* &= \overline{\phi(f_\lambda)}f_\lambda\overline{\phi(f_\xi)}f_\xi=\overline{\phi(f_\lambda f_\xi)}f_\lambda f_\xi\\ &=\overline{\phi(r(\lambda,\xi))\phi(f_{\lambda\xi})}r(\lambda,\xi)f_{\lambda\xi}\\ &=\overline{r(\lambda,\xi)\phi(f_{\lambda\xi})}r(\lambda,\xi)f_{\lambda\xi}\\ &=f_{\lambda\xi}^*. \end{aligned}$$

用$\{f_\lambda^*\}_{\lambda\in\Lambda}$ 替代$\{f_\lambda\}_{\lambda\in\Lambda}$ 即为所求. 对于$\{g_\lambda\}_{\lambda\in\Lambda}$ 同样处理. □

选用断言中的$\{f_\lambda\}_{\lambda\in\Lambda}$ 和$\{g_\lambda\}_{\lambda\in\Lambda}$, 根据上面做法构造$W: L^2(Y,\nu)\to L^2(X,\mu)$, $g_\lambda\mapsto f_\lambda$. 于是

$$W(g_\lambda g_\xi)=W(g_{\lambda\xi})=f_{\lambda\xi}=f_\lambda f_\xi=W(g_\lambda)W(g_\xi).$$

因为$\{f_\lambda\}_{\lambda\in\Lambda}$ 和$\{g_\lambda\}_{\lambda\in\Lambda}$ 是基, 我们容易由此证明

$$W(fg)=(Wf)(Wg),\forall f,g\in L^2(Y,\nu).$$

接着我们证明W,W^{-1} 将有界函数映为有界函数. 设$f\in L^\infty(Y,\nu)$. 根据引理4.4.3, $fh\in L^2(Y,\nu),\forall h\in L^2(Y,\nu)$. 于是$W(f)W(h)=W(fh)\in L^2(X,\mu)$, $\forall h\in L^2(Y,\nu)$. 即$W(f)k\in L^2(X,\mu),\forall k\in L^2(X,\mu)$. 再根据引理4.4.3, $W(f)\in L^\infty(X,\mu)$. 同理证明, W^{-1} 将有界函数映为有界函数.

根据定理4.2.11, T 与S 为共轭. □

注记 4.4.6 回顾上面定理证明, 其中最大的困难在于断言的证明, 即说明可以选取$\{f_\lambda\}_{\lambda\in\Lambda}$ 使得

$$f_\lambda f_\xi=f_{\lambda\xi},\ \forall\lambda,\xi\in\Lambda.$$

上面证明我们选取了*[195]*中的证明, 实际上我们可以有下面简单的证明:

先任意取$\{f_\lambda\}_{\lambda\in\Lambda}$. 因为遍历性, 存在常数$r(\lambda,\xi)\in\mathbb{S}^1$ 使得

$$f_\lambda f_\xi = r(\lambda,\xi) f_{\lambda\xi}\quad a.e.$$

因为特征值全体是可数的, 存在T-不变子集$N\in\mathcal{X}$, $\mu(N)=0$, 使得对于任何$\lambda,\xi\in\Lambda$,

$$f_\lambda(x) f_\xi(x) = r(\lambda,\xi) f_{\lambda\xi}(x)\quad \forall x\in X\setminus N.$$

取$x_0\in X\setminus N$, 则

$$r(\lambda,\xi)=\frac{f_\lambda(x_0)f_\xi(x_0)}{f_{\lambda\xi}(x_0)}.$$

令

$$f_\lambda^*=\frac{1}{f_\lambda(x_0)}f_\lambda.$$

则

$$f_\lambda^* f_\xi^*(x)=\frac{1}{f_\lambda(x_0)}f_\lambda\frac{1}{f_\xi(x_0)}f_\xi=\frac{1}{f_{\lambda\xi}(x_0)}f_{\lambda\xi}=f_{\lambda\xi}^*,\ \forall \lambda,\xi\in\Lambda.$$

注记 4.4.7 在第八章 中我们会运用交理论给出*Halmos-von Neumann* 定理的一个非常简单的证明.

推论 4.4.8 设$(X,\mathcal{X},\mu,T)$ 为具有离散谱的保测系统, 那么T 可逆, 并且T 与T^{-1} 共轭.

§4.4.2　离散谱与群旋转

设G 是一个紧致交换群和$a\in G$. 定义群旋转

$$R_a\colon G\to G, x\mapsto ax.$$

我们之前易证, $(G,\mathcal{B}(G),m,R_a)$ 是一个保测动力系统, 其中m 为Haar 测度. R_a 为遍历的当且仅当$\langle a\rangle=\{a^n: n\in\mathbb{Z}\}$ 在G 中稠密.

定理 4.4.9 设G 是一个紧致交换群和$a\in G$, $(G,\mathcal{B}(G),m,R_a)$ 为遍历的群旋转系统. 那么每个R_a 每个特征函数为特征的数乘, 并且全体特征值为

$$\Lambda_{R_a}=\{\gamma(a):\gamma\in\widehat{G}\}.$$

尤其, R_a 具有离散谱.

证明. 设$\gamma\in\widehat{G}$, 则

$$U_{R_a}\gamma(x)=\gamma(R_a x)=\gamma(ax)=\gamma(a)\gamma(x).$$

于是γ 为以$\gamma(a)$ 为特征值的特征函数. 因为$\widehat{G}$ 组成了$L^2(X,\mu)$ 的一组基, 所以R_a 具有离散谱.

因为每个特征值是简单的, 并且不同特征值的特征函数互相垂直, 所以

$$\Lambda_{R_a} = \{\gamma(a) : \gamma \in \widehat{G}\}$$

为全体特征值. 否则其它的特征函数将与所有$\widehat{G}$ 正交, 而$\widehat{G}$ 组成了$L^2(X,\mu)$ 的一组基, 这不可能. □

定理 4.4.10 *设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 那么T 具有离散谱当且仅当T 共轭于某个紧致交换群G 上的遍历旋转.*

如果$(X,\mathcal{X},\mu)$ 为Lebesgue 空间, 那么G 为可度量的.

证明. 根据定理4.4.9, 我们需证明, 如果T 具有离散谱, 那么T 共轭于某个紧致交换群G 上的遍历旋转. 设Λ_T 为T 全体特征值的集合, 赋予Λ_T 离散拓扑. 如果$(X,\mathcal{X},\mu)$ 为Lebesgue 空间, 那么$L^2(X,\mu)$ 可分, 此时Λ_T 为可数的.

令$G=\widehat{\Lambda_T}$, 那么G 为紧致交换群. 根据对偶定理, 定义

$$\phi : \widehat{G} = \widehat{\widehat{\Lambda}}_T \to \Lambda_T, \tilde{\lambda} \mapsto \lambda,\ \tilde{\lambda}(g) = g(\lambda),\ \forall g \in G = \widehat{\Lambda_T}$$

为相应同构. 令

$$a : \Lambda_T \to \mathbb{S}^1,\ \lambda \mapsto \lambda,$$

则$a \in G = \widehat{\Lambda_T}$, 并且$\tilde{\lambda}(a) = a(\lambda) = \lambda$. 令

$$R_a : G \to G, g \mapsto ag.$$

我们下面证明R_a 为遍历的, 并且与T 共轭.

设m 为G 的Haar测度, 设$f \in L^2(G,m)$ 满足$f \circ R_a = f$. 设f 的Fourier 展开为

$$f = \sum_{\lambda \in \Lambda_T} b_\lambda \tilde{\lambda}.$$

由$f \circ R_a = f$, 我们有

$$\sum_{\lambda \in \Lambda_T} b_\lambda \tilde{\lambda}(g) = f(g) = f \circ R_a(g) = \sum_{\lambda \in \Lambda_T} b_\lambda \tilde{\lambda}(ag) = \sum_{\lambda \in \Lambda_T} b_\lambda \tilde{\lambda}(a)\tilde{\lambda}(g).$$

于是

$$b_\lambda \tilde{\lambda}(a) = b_\lambda.$$

因为$\tilde{\lambda}(a) = a(\lambda) = \lambda$, 所以

$$b_\lambda \lambda = b_\lambda.$$

如果$b_\lambda \neq 0$, 那么$\lambda = 1$, 即$\tilde{\lambda} \equiv 1$. 于是$f = b_1$, 所以R_a 遍历.

根据定理4.4.9, R_a 为离散谱的. 根据定理4.4.5, 为证明T 与R_a 共轭, 仅需说明二者具有相同的特征值集合. 但是根据

$$\Lambda_{R_a} = \{\gamma(a) : \gamma \in \widehat{G}\} = \{a(\lambda) : \lambda \in \Lambda_T\} = \{\lambda : \lambda \in \Lambda_T\} = \Lambda_T.$$

我们得到所求.

群G 为可度量的当且仅当Λ_T 为可数的当且仅当$L^2(G,m)$ 可分. 所以如果$(X,\mathcal{X},\mu)$ 为Lebesgue 空间, 那么G 为可度量的. □

根据上面定理证明, 我们得到如下推论：

推论 4.4.11 对于$\mathbb{S}^1$ 任何子群Λ, 存在群旋转系统使得其特征值全体集合恰为Λ.

§4.5　Rohlin斜积定理

在本小节我们介绍一类重要的构造因子映射的方法：斜积. Rohlin斜积定理告诉我们遍历系统间的因子映射都可以表示为斜积.

§4.5.1　群扩充

设$(X,\mathcal{X},\mu,T)$ 为保测系统, 令$\mathrm{End}(X,\mu,T)$ 为全体与T 交换的保测变换$\phi : X \to X$ 的全体, 而令$\mathrm{Aut}(X,\mu,T)$ 为全体与T 交换的可逆保测变换的全体. 则$\mathrm{End}(X,\mu,T)$ 为半群, 而$\mathrm{Aut}(X,\mu,T) \subseteq \mathrm{End}(X,\mu,T)$ 为群. 赋予$(X,\mathcal{X},\mu,T)$ Cantor 模型, 设$\{A_n\}_n$ 为一组既开又闭集合组成的基, 那么可以定义$\mathrm{End}(X,\mu,T)$ 的度量如下：

$$d(\phi,\psi) = \sum_{n=1}^{\infty} \frac{\mu(\phi^{-1}(A_n)\Delta\psi^{-1}(A_n))}{2^n}.$$

对于$\mathrm{Aut}(X,\mu,T)$, 定义度量

$$\widehat{d}(\phi,\psi) = \sum_{n=1}^{\infty} \frac{\mu(\phi^{-1}(A_n)\Delta\psi^{-1}(A_n)) + \mu(\phi(A_n)\Delta\psi(A_n))}{2^{n+1}}.$$

此度量不依赖于$(X,\mathcal{X},\mu,T)$ 拓扑模型的选取, 并且在此度量下$\mathrm{Aut}(X,\mu,T)$ 为Polish 群.

定义 4.5.1 (群扩充) 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $K \subseteq \mathrm{Aut}(X,\mu,T)$ 为紧致子群. 设

$$\mathcal{A}(K) = \{A \in \mathcal{X} : \phi A = A,\ \forall \phi \in K\}.$$

那么$\mathcal{A}(K)$ 为$\mathcal{X}$ 的T 不变子σ-代数. 于是$\mathcal{A}(K)$ 定义了一个因子系统$(Y,\mathcal{Y},\nu,S)$ 以及因子映射$\pi : (X,\mathcal{X},\mu,T) \to (Y,\mathcal{Y},\nu,S)$, $\nu = \pi_*\mu$. 我们称$(X,\mathcal{X},\mu,T)$ 为$(Y,\mathcal{Y},\nu,S)$ 的**群扩充**, 记为$Y = X/K$.

§4.5.2 斜积系统

设$(Y,\mathcal{Y},\nu,\Gamma)$ 为保测系统, 其中Γ 为可数离散群. 设$(U,\mathcal{U},\rho)$ 为标准Borel 空间. 设$\mathrm{Aut}(U,\rho)$ 为(U,ρ) 全体可逆保测变换的全体. 称可测映射

$$\alpha:\Gamma\times Y\to \mathrm{Aut}(U,\rho)$$

为**可测cocycle** 是指它满足下面的**cocycle方程:**

$$\alpha(\gamma\gamma',y)=\alpha(\gamma,\gamma'y)\alpha(\gamma',y).$$

取定一个cocycle α, 我们就可以定义系统$(Y\times U,\mathcal{Y}\times\mathcal{U},\nu\times\rho,\Gamma)$ 为

$$\gamma(y,u)=(\gamma y,\alpha(\gamma,y)u).$$

根据cocycle 方程易验证这是一个动力系统, 称之为**斜积系统**(skew-product), 记为

$$(Y,\mathcal{Y},\nu,\Gamma)\underset{\alpha}{\times}(U,\rho).$$

一类重要的斜积系统是取$(U,\rho)=(K/L,\widetilde{m})$, 其中$K$ 为紧致群, $L\subseteq K$ 为满足$\bigcap_{k\in K}k^{-1}Lk=\{e\}$ 的闭子群, m 为K 上的Haar 测度而$\widetilde{m}$ 为m 在K/L 上的像. cocycle $\alpha:\Gamma\times Y\to K$ 使得$\alpha(\gamma,x)\cdot kL$ 为$\alpha(\gamma,x)\cdot k$ 相对L 的陪集. 我们得到的斜积系统$Y\underset{\alpha}{\times}K/L$ 称为**齐性斜积空间**(homogeneous skew-product). 当$L=\{e\}$ 为平凡群时, 得到的斜积系统$Y\underset{\alpha}{\times}K$ 称为**群斜积系统**.

在本书中, 我们一般讨论离散系统$\Gamma=\mathbb{Z}$. 设$(Y,\mathcal{Y},\nu,S)$ 为保测系统, $(U,\mathcal{U},\rho)$ 为标准Borel 空间. 那么此时cocycle 是指可测映射

$$\alpha:Y\to \mathrm{Aut}(U,\rho).$$

斜积系统为

$$(X,\mathcal{X},\mu,T_\alpha)=(Y,\mathcal{Y},\nu,S)\underset{\alpha}{\times}(U,\rho),$$

其中

$$T_a(y,u)=(Sy,\alpha(y)u)$$

设$T_\alpha^n(y,u)=(S^ny,\alpha(n,y)u),n\in\mathbb{Z}$, 易验证

$$\alpha(n,y)=\begin{cases}\prod\limits_{i=0}^{n-1}\alpha(S^iy)=\alpha(S^{n-1}y)\dots\alpha(Sy)\alpha(y), & n\ge 0;\\ \big(\prod\limits_{i=n}^{-1}\alpha(S^iy)\big)^{-1}=\alpha(S^ny)^{-1}\dots\alpha(S^{-1}y)^{-1}, & n<0.\end{cases}$$

§4.5.3 局部可逆定理

设$(X,\mathcal{X},\mu)$ 为一个概率测度空间. 我们称集族$\mathcal{H}\subseteq\mathcal{X}$ 为**遗传的**(hereditary) 是指它满足遗传向下性, 即如果$A\in\mathcal{H}$ 且$\mathcal{X}\ni B\subseteq A$, 则$B\in\mathcal{H}$. 一个遗传集族$\mathcal{H}$ 称为**浸透**$\mathcal{X}$**的**(saturate), 是指对于任何$A\in\mathcal{X}$, $\mu(A)>0$, 那么存在$B\in\mathcal{H}$ 使得$B\subseteq A$ 且$\mu(B)>0$. 称遗传集族$\mathcal{H}$ **浸透**$A\in\mathcal{X}$, 是指对于任何$B\in\mathcal{X}$, 如果$B\subseteq A$ 且$\mu(B)>0$, 那么存在$C\in\mathcal{H}$ 使得$C\subseteq B$ 且$\mu(C)>0$. 一个集合$U\in\mathcal{X}$ 称为遗传集族$\mathcal{H}$ 的**覆盖**(cover), 是指$A\underset{\mu}{\subseteq}U,\forall A\in\mathcal{H}$; 称为遗传集族$\mathcal{H}$ 的**可测并**, 是指它为$\mathcal{H}$ 覆盖并且$\mathcal{H}$ 浸透U. 易见如果$\mathcal{H}$ 浸透$\mathcal{X}$, 那么它的覆盖就是X.

对于一个取定的遗传集族$\mathcal{H}$, 它的可测并是唯一的. 否则设$U,U'\in\mathcal{X}$ 为两个可测并, 且$\mu(U\setminus U')>0$, 那么根据$\mathcal{H}$ 浸透U, 存在$C\in\mathcal{H},\mu(C)>0$ 且$C\subseteq U\setminus U'$. 但是U' 为覆盖, 所以$C\subseteq U'\pmod\mu$. 此二者矛盾!

引理 4.5.2 (耗尽引理) *(Exhaustion Lemma)[1, 1.0.7]* 设$(X,\mathcal{X},\mu)$ 为一个概率测度空间, 集族$\mathcal{H}\subseteq\mathcal{X}$ 为遗传的. 那么存在互不相交的子集列$A_1,A_2,\ldots\in\mathcal{H}$ 使得$U(\mathcal{H})=\bigcup_{n=1}^{\infty}A_n$ 为遗传集族$\mathcal{H}$ 的可测并.

证明. 设

$$\varepsilon_1=\sup\{\mu(A):A\in\mathcal{H}\},$$

取$A_1\in\mathcal{H}$ 使得$\mu(A_1)\geq\frac{\varepsilon_1}{2}$. 设

$$\varepsilon_2=\sup\{\mu(A):A\in\mathcal{H},\ A\cap A_1=\emptyset\},$$

取$A_2\in\mathcal{H}$ 使得$\mu(A_1)\geq\frac{\varepsilon_2}{2}$ 且$A_2\cap A_1=\emptyset$.

继续此过程, 我们得到互不相交的子集列$\{A_n\}_{n=1}^{\infty}\subseteq\mathcal{H}$ 以及递减数列ε_n 使得

$$\varepsilon_n=\sup\{\mu(A):A\in\mathcal{H},\ A\cap(A_1\cup\ldots\cup A_{n-1})=\emptyset\},\quad \mu(A_n)\geq\frac{\varepsilon_n}{2}.$$

于是

$$\sum_{n=1}^{\infty}\varepsilon_n\leq 2\sum_{n=1}^{\infty}\mu(A_n)\leq 2,$$

尤其$\varepsilon_n\to 0,n\to\infty$.

令$U=\bigcup_{n=1}^{\infty}A_n$, 下说明$U$ 为$\mathcal{H}$ 的可测并. 易见$\mathcal{H}$ 浸透U, 所以仅需说明它是$\mathcal{H}$ 的覆盖. 否则存在$A\in\mathcal{H}$ 使得$\mu(A)>0$ 且$A\cap A_n=\emptyset,\forall n\in\mathbb{N}$. 根据$\varepsilon_n$ 定义, $\mu(A)\leq\varepsilon_n\to 0,\forall n$, 于是与$\mu(A)>0$ 矛盾. □

推论 4.5.3 设$(X,\mathcal{X},\mu)$ 为一个概率测度空间, 集族$\mathcal{H}\subseteq\mathcal{X}$ 为遗传的并且浸透$\mathcal{X}$. 那么存在互不相交的子集列$A_1,A_2,\ldots\in\mathcal{H}$ 为X 的剖分.

定理 4.5.4 (局部可逆定理) 设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为标准*Borel*空间, 设$\pi : X \to Y$ 为保测映射使得对于任何$y \in Y$, $\pi^{-1}(y)$ 为可数的. 那么存在X 的可数剖分$\xi = \{A_n\}_{n\in\mathbb{N}}$ 使得$\pi|_{A_n} : A_n \to \pi(A_n)$ 为可逆保测的, $\forall n \in \mathbb{N}$.

证明. 设

$$\mu = \int_Y \mu_y d\nu(y)$$

为积分分解. 因为对于$\nu - a.e.\ y \in Y$, $\mu_y(\pi^{-1}(y)) = 1$ 以及$\pi^{-1}(y)$ 为可数集, 所以我们有对于$\nu - a.e.\ y \in Y$ μ_y 为纯原子的, 于是$\mu - a.e.\ x \in X$, $\mu_{\pi x}(\{x\}) > 0$.

称$A \in \mathcal{X}$ 为π-截面是指$\pi(A) = B \in \mathcal{Y}$ 且$\pi|_A : A \to B$ 为可测双射. 我们断言存在X 的可数可测剖分$\xi = \{A_n\}_{n\in\mathbb{N}}$ 使得每个A_n 为π-截面. 设$\mathcal{H}$ 为全体π-截面的集合, 易见它为遗传的. 下面我们说明它浸透$\mathcal{X}$. 设$B \in \mathcal{X}$, $\mu(B) > 0$. 则$\pi(B) = C$ 为Y 的分析集. 根据Jankov-von Neumann 定理, 存在普遍可测可测映射$f : C \to B$ 使得$\pi \circ f = \mathrm{id}_C$. 由$\nu(C) = \mu(B) > 0$, 取Borel 集$C' \in \mathcal{Y}$ 使得$C' \subseteq C$ 且$\nu(C') = \nu(C)$. 设$A = f(C')$, 由f 单射, $A \in \mathcal{X}$. 易见$A \in \mathcal{H}$ 且根据$\nu(C') > 0$, $\mu_y(f(y)) > 0, \nu - a.e.$ 有

$$\mu(A) = \int_Y \mu_y(A) d\nu(y) = \int_{C'} \mu_y(A) d\nu(y) = \int_{C'} \mu_y(f(y)) d\nu(y) > 0.$$

根据推论4.5.3, 存在可测剖分ξ 满足条件. □

§4.5.4 Rohlin斜积定理

下面我们给出著名的Rohlin斜积定理.

定理 4.5.5 (Rohlin斜积定理) 设$\pi : (X,\mathcal{X},\mu,T) \to (Y,\mathcal{Y},\nu,S)$ 为因子映射, 其中$(X,\mathcal{X},\mu,T)$ $(Y,\mathcal{Y},\nu,S)$为可逆的遍历系统. 那么$(X,\mathcal{X},\mu,T)$ 同构于$(Y,\mathcal{Y},\nu,S)$ 的一个斜积系统. 即存在标准概率空间$(U,\mathcal{U},\rho)$ 以及*cocycle* $\alpha : Y \to \mathrm{Aut}(U,\rho)$ 使得

$$(X,\mathcal{X},\mu,T_\alpha) \cong (Y,\mathcal{Y},\nu,S) \underset{\alpha}{\times} (U,\rho).$$

证明. 设

$$\mu = \int_Y \mu_y d\nu(y)$$

为测度分解. 我们假设$X = [0,1]$, 令

$$p : X \to [0,1],\ x \mapsto \mu_{\pi(x)}(\{x\}).$$

因为$p(x) = \lim_{\delta\to 0} \mu_{\pi(x)} B_\delta(x)$, 所以$p$ 是可测的.

$$\mu_{\pi(Tx)}(\{Tx\}) = T_*\mu_{\pi(x)}(\{Tx\}) = \mu_{\pi(x)}(\{x\}),$$

所以p 为T 不变的. 根据遍历性, $p \equiv c < 1$. 有两种情况: $c = 0$, 此时$\nu - a.e.\ y \in Y$ 测度μ_y 为连续的; $c > 0$, 此时$\nu - a.e.\ y \in Y$ 测度μ_y 为离散的.

情况1. $c=0$. 此时$\nu-a.e.\ y\in Y$ 测度μ_y 为连续的. 此时设$U=[0,1]$, ρ 为U 为Lebesgue 测度, 定义

$$\phi:X\to Y\times U;\quad \psi:Y\times U\to X$$

为

$$\phi(x)=(\pi(x),\mu_{\pi(x)}([0,x])),\quad \psi(y,u)=\min\{x\in X:\mu_y([0,x])\geq u\}.$$

则$\psi\circ\phi=\mathrm{id}$, 并且对于任何$A\in\mathcal{Y},t\in[0,1]$,

$$\begin{aligned}\mu\left(\phi^{-1}(A\times[0,1])\right)&=\int_A\mu_y\left(\psi(A\times[0,t])\right)d\nu(y)\\&=\int_A\mu_y\left(\psi(\{y\}\times[0,t])\right)d\nu(y)\\&=\int_A\mu_y\left([0,\psi(y,t)]\right)d\nu(y)\\&=t\nu(A).\end{aligned}$$

于是$\phi_*\mu=\nu\times\rho$.

情况2. $c>0$. 此时$\nu-a.e.\ y\in Y$, μ_y 为纯原子的, 所以$\pi^{-1}(y)$ 至多可数. 根据局部可逆定理4.5.4, 存在X 的可数可测剖分$\xi=\{A_n\}_{n\in\mathbb{N}}$ 使得每个A_n 为π- 截面.

函数$y\mapsto|\pi^{-1}(y)|=\sum_{n=1}^{\infty}\mathbf{1}_{\pi(A_n)}(y)$ 为可测且S 不变的. 根据遍历性, 存在$r\in\mathbb{N}$ 使得$c=\frac{1}{r}$, 并且$\nu-a.e.\ y\in Y$ 有$\mu_y=\frac{1}{r}\sum_{i=1}^{r}\delta_{x_i}$, 其中$x_1,\ldots,x_r\in X$ 互异. 下面我们构造r 个满的π- 截面$B_1,\ldots,B_r$ 使得它们为X 的剖分, 即$X=B_1\cup\ldots\cup B_r$, 且$\pi(B_1)=\ldots=\pi(B_r)=Y$.

设$\nu(Y\setminus\pi(A_1))>0$, 令

$$C_{n+1}=A_{n+1}\cap\pi^{-1}\left(Y\setminus\bigcup_{k=1}^{n}\pi(A_k)\right),\ n\geq1,$$

以及

$$B_1=A_1\cup\bigcup_{n=2}^{\infty}C_k.$$

则$\{B_1,A_2\setminus C_2,A_3\setminus C_3,\ldots\}$ 为X 的π-截面剖分, 且$\pi(B_1)=Y$.

接着设$\pi_1=\pi|_{X\setminus B_1}:X\setminus B_1\to Y$. 则

$$|\pi_1^{-1}(y)\cap(X\setminus B_1)|=r-1,\ \nu-a.e.\ y\in Y.$$

注意$\{A'_n=A_n\setminus C_n\}_{n=2}^{\infty}$ 为$X\setminus B_1$ 的π_1-截面组成的剖分. 如$\nu(Y\setminus\pi(A'_2))>0$, 令

$$C'_{n+1}=A'_{n+1}\cap\pi_1^{-1}\left(Y\setminus\bigcup_{k=2}^{n}\pi(A'_k)\right),\ n\geq2,$$

以及

$$B_2=A'_2\cup\bigcup_{n=3}^{\infty}C'_k.$$

则$\{B_2, A'_3 \setminus C'_3, A'_4 \setminus C'_4, \ldots\}$ 为X 的π-截面剖分, 且$\pi(B_2) = Y$.

经过有限次处理后我们就得到r 个满的π-截面$B_1, \ldots, B_r$ 使得它们为X 的剖分, 且$\pi(B_1) = \ldots = \pi(B_r) = Y$. 于是我们可以定义映射

$$\phi : X \to Y \times \{1, 2, \ldots, r\}$$

使得$\phi_* \mu = \nu \times \rho$, 其中$\rho$ 为$U = \{1, 2, \ldots, r\}$ 上的等分布测度.

在两种情况下我们都定义了测度空间同构

$$\phi : X \to Y \times U, \ \phi_* \mu = \nu \times \rho.$$

通过ϕ 来定义斜积定义中的cocycle 映射α 和T_α, 即α 和T_α 如下式子来定义:

$$T_\alpha : Y \times U \to Y \times U, \ T_\alpha(y, u) = (Sy, \alpha(y)u) = \phi(Tx).$$

这样我们就完成了整个定理的证明. □

最后我们陈述一个定理, 在本书不再给出证明.

定理 4.5.6 *设$(X, \mathcal{X}, \mu, T)$ 为遍历系统, $K \subseteq \mathrm{Aut}(X, \mu, T)$ 为紧致子群. 设$(Y, \mathcal{Y}, \nu, S) = X/K$ 为群因子, 那么存在cocycle $\alpha : Y \to \mathrm{Aut}(K)$ 使得*

$$(X, \mathcal{X}, \mu, T_\alpha) \cong (Y, \mathcal{Y}, \nu, S) \underset{\alpha}{\times} K.$$

即X 为Y 的群斜积系统.

§4.6 遍历分解定理

在本节我们给出遍历分解定理. 遍历分解定理告诉我们任何保测系统总是可以根据某种方式分解为若干个遍历系统. 在这里我们用Birkhoff遍历定理给出一个证明, 后面还会根据Choquet 定理给出另一个证明方法.

设$(X, \mathcal{X}, \mu, T)$ 为保测系统, 之前我们定义了

$$\mathcal{I} = \mathcal{I}(T) = \{A \in \mathcal{X} : \mu(A \Delta T^{-1} A) = 0\}.$$

于是遍历定理可以陈述为:

$$\mathbb{A}_n f = \frac{1}{n} \sum_{j=0}^{n-1} T^j f \xrightarrow{L^2/a.e.} \mathbb{E}(f|\mathcal{I}), n \to \infty.$$

定理 4.6.1 (遍历分解定理) 设$(X,\mathcal{X},\mu,T)$ 为保测系统, 其中$(X,\mathcal{X},\mu)$ 为$Borel$ 概率空间. 那么存在$Borel$ 概率空间$(Y,\mathcal{Y},\nu)$ 以及可测映射[1]

$$Y\to\mathcal{M}(X),\quad y\to\mu_y$$

使得

1. 对于$\nu-a.e.\ y\in Y$, μ_y 为$(X,\mathcal{X})$ 上遍历测度,
2. $\mu=\int_Y\mu_y d\nu(y)$.

等价的, 如果取$(Y,\mathcal{Y},\nu)=(X,\mathcal{X},\mu)$, 那么

$$\mu_x=\mu_x^{\mathcal{I}}.$$

证明. 因为$(X,\mathcal{X},\mu)$ 为Borel 概率空间, 所以取可数生成的σ-子代数$\widetilde{\mathcal{I}}\subseteq\mathcal{I}$ 且$\widetilde{\mathcal{I}}\underset{\mu}{=}\mathcal{I}$. 令$\widetilde{\mathcal{I}}=\sigma(\{E_n\}_n=1^\infty)$. 令

$$N'=\bigcup_{n=1}^\infty T^{-1}E_n\Delta E_n=\bigcup_{E\in\widetilde{\mathcal{I}}}T^{-1}E\Delta E,$$

设N'' 为使得定理0.3.9, 定理0.3.12, 定理0.3.13 中所排除的零测集. 令

$$N=\bigcup_{n=0}^\infty T^{-n}(N'\cup N''),$$

则N 为零测集, 且$T^{-1}N\subseteq N$.

对于$x\notin N$, $T^{-1}\mathcal{I}\underset{\mu}{=}\mathcal{I}$, 根据定理0.3.13,

$$T_*\mu_x^{\mathcal{I}}=\mu_{Tx}^{\mathcal{I}}.$$

因为$Tx\notin N$, 根据定理0.3.9, $[x]_{\widetilde{\mathcal{I}}}=[Tx]_{\widetilde{\mathcal{I}}}$, 所以

$$\mu_{Tx}^{\mathcal{I}}=\mu_x^{\mathcal{I}}.$$

于是对于任何$x\in N$, $\mu_x^{\mathcal{I}}$ 为T-不变测度. 下面我们证明$\mu_x^{\mathcal{I}}$ 为遍历的.

取$C(X)$ 的可数稠密子集$\{f_i\}_{i=1}^\infty$. 进一步放大N 使得对于任何$x\notin N$ 有

$$\frac{1}{M}\sum_{n=0}^{M-1}f_i(T^nx)\to\mathbb{E}(f_i|\mathcal{I})(x)=\int_X f_i d\mu_x^{\mathcal{I}},M\to\infty.$$

根据定理0.3.9,

$$N_1=N\cup\{x:\mu_x^{\mathcal{I}}(N)>0\}$$

[1]此处$\mathcal{M}(X)$ 为X 上的全体概率测度, 我们将在第第六章 章具体给出$\mathcal{M}(X)$ 的拓扑.

的测度为零. 如果$[x]_{\widetilde{\mathcal{I}}}=[y]_{\widetilde{\mathcal{I}}}$, $x\notin N_1, y\notin N_1$, 则$\mu_x^{\mathcal{I}}=\mu_y^{\mathcal{I}}$, 且对于任何$i\in\mathbb{N}$

$$\frac{1}{M}\sum_{n=0}^{M-1} f_i(T^n y)\to\int_X f_i d\mu_x^{\mathcal{I}}, M\to\infty.$$

因为$\mu_x^{\mathcal{I}}(N)=0$, 所以上式意味着对于$\mu_x^{\mathcal{I}}-a.e.\ y\in X$ 有

$$\frac{1}{M}\sum_{n=0}^{M-1} f(T^n y)\to\int_X f d\mu_x^{\mathcal{I}}, M\to\infty,\ \forall f\in L^2(X,\mu_x^{\mathcal{I}}).$$

即$\mu_x^{\mathcal{I}}$ 为遍历的.

接着根据定理0.3.12, 存在因子映射$\phi: X\to Y$ 满足条件. □

我们可以把遍历分解定理重新表述如下：

定理 4.6.2 (遍历分解定理) *设$(X,\mathcal{X},\mu,T)$ 为保测系统, 其中$(X,\mathcal{X},\mu)$ 为Borel 概率空间. 那么存在Borel 概率空间系统$(Y,\mathcal{Y},\nu,S)$ 以及因子映射*

$$\phi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$$

使得$\mathcal{I}=\phi^{-1}(\mathcal{Y})$. 设

$$\mu=\int_Y \mu_y d\nu(y)$$

为测度分解, 对于$y\in Y$ 设$X_y=\phi^{-1}(y)$ 以及$\mathcal{X}_y=\mathcal{X}\cap X_y$. 那么

1. 对于$\nu-a.e.\ y\in Y$, $(X_y,\mathcal{X}_y,\mu_y,T)$ 为遍历系统.

2. $S=\mathrm{id}:Y\to Y$.

§4.7 注记

关于Lebesgue 空间理论我们采取了和[76, 195] 等书中的处理方式, 介绍涉及的基本结论而不涉及定理的证明. 主要原因是这些定理是测度论中很深刻的结论, 具体给出证明将占较大篇幅且会偏离本书的主旨, 感兴趣的读者可以参见[126, 167] 等. Kolmogorov系统是非常重要的一类系统, 我们在熵理论中还会介绍. Rohlin 斜积定理的材料来自[1]. 遍历分解定理的不同证明将在第六章中给出.

Holmos-von Neumann定理有许多证明, 本章中定理4.4.5的证明选自[195], 注记4.4.6 中的简化证明来自[155]. 在第八章我们还有一个从不交性出发的简单证明.

第五章　拓扑动力系统基础

在这一章中我们介绍拓扑动力系统的一些基本概念和结论. 拓扑动力系统和遍历论是动力系统的两个密切关联的分支. 自学科产生之初, 遍历理论和拓扑动力系统二者就有着不可分割的联系. 一方面, 拓扑动力系统可以自然地视为一个保测系统；另一方面任何遍历系统有其拓扑实现. 两套理论中有许多相对应的概念：遍历—-极小性, 离散谱—-等度连续, 测度熵—-拓扑熵, 等等. 这些因素导致了这两套理论有着惊人的平行性, 二者许多结论有着极为相似的陈述, 但各自的证明方法却可能完全不同.

在前文中我们实际上已经介绍过拓扑动力系统的一些基本概念, 在本章中我们系统的再总结之前的概念并介绍一些新的内容. 在第一节我们回顾基本概念, 并罗列许多例子. 第二、三节介绍与遍历性相应的传递性和极小性, 二者都反映了系统的不可分割性. 在第四节我们详细介绍各种拓扑混合性. 第五节总结了其它常用的不变子集. 对应于遍历论中离散谱系统, 在拓扑动力系统中我们有等度连续系统, 也称为拓扑离散谱系统, 它是动力学性质相对简单的系统. 最后我们介绍可扩同胚, 这是一种非常重要的性质, 与微分动力系统、符号系统等有着密切关系.

在下一章我们将会强调拓扑动力系统与遍历论的关系. 到时读者自然会理解为什么介绍遍历论的书籍中一般都要介绍拓扑动力系统, 反之亦然.

§5.1　基本概念

§5.1.1　一般群作用的拓扑动力系统

设X 为紧致的Hausdorff 空间, G 为拓扑群或含单位元拓扑半群, 如果$\phi: G \times X \to X$ 连续且满足：

- 对任意$x \in X$ 有$\phi(e, x) = x$, 其中e 为G 的单位元,
- 对任意$x \in X$ 和$g_1, g_2 \in G$, $\phi(g_1, \phi(g_2, x)) = \phi(g_1 g_2, x)$ 成立,

那么我们就称(X, G, ϕ) 为一个**拓扑动力系统**. 一般地我们也直接用(X, G) 记一个拓扑动力系统. 如果G 为群，此时对于每个$g \in G$, $\phi(g, \cdot): X \to X$, $x \mapsto \phi(g, x)$ 为同胚；如果G 为半群，那么$\phi(g, \cdot)$ 为连续的. 为方便计, 我们将$\phi(g, x)$ 简记为gx. 当X为独点集时, 我们称系统(X, G)为**平凡系统**.

设(X, G) 为拓扑动力系统. 对$x \in X$, 称$\mathrm{orb}(x, G) = \{gx : g \in G\}$ 为x **的轨道**. 设A 为X 的子集, 如果

$$gA = \{gx : x \in A\} \subseteq A, \forall g \in G,$$

则称A 为**不变集**. 如果$A \subseteq X$ 为闭的不变集, 则将G 作用限制在A 上也成为一个动力系统,称之为(X, G) 的**子系统**, 记为(A, G). 对任意$x \in X$, 易见$\overline{\mathrm{orb}(x, G)}$ 为闭的不变集, 进而$(\overline{\mathrm{orb}(x, G)}, G)$ 是(X, G)的一个子系统. 这是一个常用的构造新动力系统的方法.

设(X,G) 和(Y,G) 为两个拓扑动力系统, 定义它们的**乘积系统**为$(X\times Y,G)$, 其中

$$g(x,y)=(gx,gy),\forall g\in G.$$

任意多个系统的乘积系统可以类似地定义.

就像别的数学分支一样,拓扑动力系统的一个中心问题便是系统的分类问题.于是一个自然的问题是：两个拓扑动力系统何时是"一样的"？在拓扑动力系统中, 我们有如下定义：

定义 5.1.1 设(X_1,G,ϕ_1) 和(X_2,G,ϕ_2) 为两个拓扑动力系统. 如果存在一个连续满射$\pi:X_1\to X_2$ 使得以下图表交换

$$\begin{array}{ccc} X_1 & \xrightarrow{g} & X_1 \\ \downarrow{\scriptstyle\pi} & & \downarrow{\scriptstyle\pi} \\ X_2 & \xrightarrow{g} & X_2 \end{array}$$

亦即, $\pi(gx)=g(\pi x)$, $\forall g\in G$, $\forall x\in X_1$, 那么我们就称(X_1,G,ϕ_1) 为(X_2,G,ϕ_2) 的一个**扩充**, 或者称(X_2,G,ϕ_2) 是(X_1,G,ϕ_1) 的一个**因子**. 此时称π 为一个**因子映射** 或称为**半共轭**.

如果π为同胚, 我们就称(X_1,G,ϕ_1) 和(X_2,G,ϕ_2) 为**同构的**或者**共轭**.

之前我们定义了保测系统的同构和共轭，在这里我们用了同样的术语. 在具体情况下，我们需要根据上下文区分是哪个定义. 有时为了表述更为精确，把保测系统的同构（共轭）称为"测度同构"（"测度共轭"），二把拓扑动力系统的同构（共轭）称为"拓扑同构"（"拓扑共轭"）.

在拓扑动力系统中, 两个系统如果为同构的, 我们就认为它们是"一样的". 于是,寻求同构不变量是拓扑动力系统的一个重要主题.

§5.1.2 离散拓扑动力系统

在本书中,如非特别指出,我们一般研究$\mathbb{Z}_+$ 作用下的系统. 亦即, 我们所指的拓扑动力系统是指偶对(X,T), 其中X 为紧度量空间而$T:X\to X$ 为连续映射.

对$x\in X$, 称

$$\mathrm{orb}(x,T)=\{x,Tx,T^2x,\cdots\}$$

为x**的轨道**. 设A为X的子集, 如果$T(A)\subseteq A$, 则称A为**正不变集**或**不变集**; 如果$T^{-1}A\subseteq A$, 则称A为**负不变集**; 如果$T(A)=A$, 则称A为**强不变集**. 如果$Y\subseteq X$为闭的不变集, 则$(Y,T|_Y)$也成为一个动力系统,称之为(X,T)的**子系统**, 有时我们就直接将它记为(Y,T). 对任意$x\in X$, 易见$\overline{\mathrm{orb}(x,T)}$为闭的不变集, 进而$(\overline{\mathrm{orb}(x,T)},T)$ 是(X,T)的一个子系统.

当我们考虑两个半离散拓扑动力系统(X,T)和(Y,S)时, 因子映射$\pi : X \to Y$就是满足$\pi \circ T = S \circ \pi$的连续满射.

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "\pi"'] & X \arrow[d, "\pi"] \\ Y \arrow[r, "S"'] & Y \end{tikzcd}$$

下面我们给出因子映射的等价描述.

设(X,T) 为拓扑动力系统. 我们可以将$X \times X$ 的子集R 视为X 上的一个关系. 如果R为$X\times X$ 的闭子集, 就称R为**闭关系**; 如果$(T\times T)(R) \subseteq R$, 就称关系$R$ 为**不变的**. 设$R \subseteq X \times X$ 为X 上闭的不变的等价关系. 对$x \in X$ 考虑x 所在的等价类$[x]_R = \{y \in X : (x, y) \in R\}$. 所有的这些等价类形成了一个新的空间$X/R = \{[x]_R : x \in X\}$, 如果$X/R$ 的拓扑取商拓扑, 则X/R 为紧度量空间. 映射T 自然地诱导了X/R 上的连续映射$T_R : X/R \to X/R, [x]_R \to [Tx]_R$, 从而$(X/R, T_R)$ 为拓扑动力系统. 设$\pi : X \to X/R$ 为商映射, 则$\pi : (X,T) \to (X/R, T_R)$为因子映射且

$$R_\pi = \{(x, y) \in X \times X : \pi(x) = \pi(y)\} = R.$$

反之, 设$\pi : (X,T) \to (Y,S)$ 为因子映射, 通过π 我们可以定义X 上的一个闭的不变的等价关系:

$$R_\pi = \{(x, y) \in X \times X : \pi(x) = \pi(y)\}.$$

易见$(X/R_\pi, T_{R_\pi})$ 拓扑同构于(Y,S). 因此, 在把拓扑同构的两个动力系统不加区分的意义下, 我们有

命题 5.1.2 设(X,T) 为拓扑动力系统. 则(X,T) 的因子系统一一对应于X 上闭的不变的等价关系.

另外注意到, 如果$T : X \to X$ 为连续的, 那么

$$T : \bigcap_{i=0}^{\infty} T^i(X) \to \bigcap_{i=0}^{\infty} T^i(X)$$

为满的. 因为一个系统几乎所有的动力学性状都集中体现在$\bigcap_{i=0}^{\infty} T^i(X)$ 上, 所以我们经常在系统(X,T) 定义中假设映射T 为满射.

§5.1.3　自然扩充

类似于定义1.1.9和定义1.1.10, 我们可以定义拓扑动力系统的逆极限和自然扩充. 因为定义是类似的, 我们在这里仅回顾下自然扩充.

将半离散系统与离散系统联系在一起的一个桥梁是所谓的**自然扩充**. 设$T : X \to X$为连续满的自映射. 设

$$\widetilde{X} = \Big\{(x_1, x_2, \ldots) \in \prod_{i=1}^{\infty} X : Tx_{i+1} = x_i,\ i \geq 1\Big\}.$$

作为乘积空间$\prod_{i=1}^{\infty} X$（取乘积拓扑）的子集, $\widetilde{X}$ 是非空闭的. 如果我们定义

$$\widetilde{T}:\widetilde{X}\to\widetilde{X},\ \widetilde{T}(x_1,x_2,\ldots)=(Tx_1,x_1,x_2,\ldots)$$

那么$\widetilde{T}$ 为同胚. 易见对每个$n\in\mathbb{N}$, 向第n 分量的投影映射$p_n:\widetilde{X}\to X$ 为连续满射. 尤其p_1 为$(\widetilde{X},\widetilde{T})$ 到(X,T) 的因子映射.

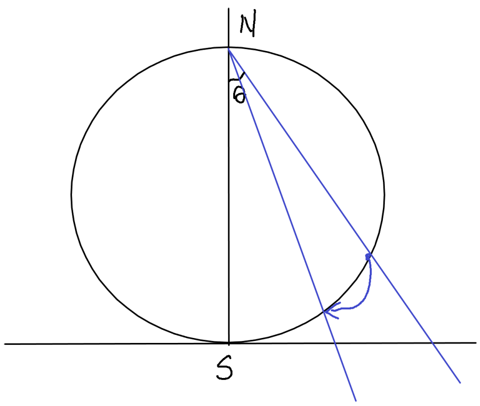
$$\begin{tikzcd}
\widetilde{X} \arrow[r, "\widetilde{T}"] \arrow[d, "p_1"'] & \widetilde{X} \arrow[d, "p_1"] \\
X \arrow[r, "T"'] & X
\end{tikzcd}$$

自然扩充$(\widetilde{X},\widetilde{T})$ 是可逆系统, 它保持了(X,T)几乎所有的动力学性质.在许多情形下, 我们的结论首先是对可逆系统证明的, 然后通过其自然扩充将一般情形转化为可逆的情况来获得相应的结论. 这是一种十分常用的技巧.

§5.1.4 一些例子

我们在第二章第(§1.2)节中介绍的大部分例子都是拓扑动力系统. 在这里我们在补充几个例子. 在后面的小节中我们还会继续介绍许多新的例子.

例 5.1.3 (南北极系统) 设$\mathbb{S}^1$ 为$\mathbb{R}^2$ 中单位圆周, 设其圆心在$(0,1)$, 南极S在$(0,0)$, 北极N在$(0,2)$. 以N 为起点, 设射线与y轴夹角为θ, 则$\theta\in(-\frac{\pi}{2},\frac{\pi}{2})$, 射线交$\mathbb{S}^1$ 于点z_θ.



南北极系统

定义

$$T:\mathbb{S}^1\to\mathbb{S}^1,\ z_\theta\mapsto z_{\arctan(\frac{\tan\theta}{2})}$$

并且$T(N)=N, T(S)=S$.

则对于$x\neq S,N$,

$$T^n x\to S, n\to\infty;\quad T^{-n}x\to N, n\to\infty.$$

例 5.1.4 (有限型子转移) 设(Σ_k,σ) 为双边全转移系统（$k\geq 2$）. 设$A=[a_{ij}]_{i,j=0}^{k-1}$ 为$k\times k$ 矩阵，其中$a_{ij}\in\{0,1\}$. 设

$$X_A=\{x=(x_n)_{n\in\mathbb{Z}}:a_{x_nx_{n+1}}=1,\forall n\in\mathbb{Z}\}.$$

易见X_A 的补集为开集, 所以X_A 为Σ_k 的闭子集. 另外易见

$$\sigma X_A=X_A.$$

于是(X_A,σ) 为拓扑动力系统, 称之为**有限型子转移系统**(*subshift of finite type*).

1. $A=[a_{ij}],a_{ij}=1,\forall i,j$. 则$X_A=\Sigma_k$.
2. $A=I$ 为恒同矩阵, 那么$X_A=\{j^{\mathbb{Z}}=(\dots,j,j,j,\dots):j\in\{0,1,\dots,k-1\}\}$.
3. 对于(Σ_2,σ), $A=\begin{pmatrix}0&0\\1&0\end{pmatrix}$, $X_A=\emptyset$.
4. 两个不同矩阵可以定义同一个有限型子转移. 例如对于(Σ_2,σ), $A_1=\begin{pmatrix}0&1\\0&1\end{pmatrix}$, $A_2=\begin{pmatrix}0&0\\0&1\end{pmatrix}$, $X_{A_1}=X_{A_2}$.

类似可以定义单边的系统.

例 5.1.5 (测地流、极限环流) 设$G=SL(2,\mathbb{R})$ 为2阶特殊线性群, m 为G 上的*Haar* 测度. 设Γ 为G 余紧的离散群, 即G/Γ 为紧致的. 对于任何$t,s\in\mathbb{R}$

$$h_t=\begin{pmatrix}1&t\\0&1\end{pmatrix},\quad g_s=\begin{pmatrix}e^{-s}&0\\0&e^s\end{pmatrix}.$$

设$X=G/\Gamma$, 其中$\mu=\mu_\Gamma$ 为X 上唯一的G 不变概率测度. 令

$$\mathbb{R}\times X\to X:(t,x\Gamma)\mapsto h_tx\Gamma.$$

称$\mathbb{R}$ 作用的系统$(X,\mathcal{B}(X),\mu,\{h_t\}_{t\in\mathbb{R}})$ 为**极限环流**(*horocycle flow*). 我们定义$h=h_1$, 则得到一个离散的**极限环系统** $(X,\mathcal{B}(X),\mu_0,h)$.

类似定义系统$(X,\mathcal{B}(X),\mu,\{g_s\}_{s\in\mathbb{R}})$, 称之为**测地流**(*geodesic flow*). 易见二者有如下关系:

$$g_sh_tg_s^{-1}=h_{e^{-2s}t},\ \forall t,s\in\mathbb{R}.\tag{5.1.1}$$

习　题

1. 验证式子5.1.1.
2. 一般地, 一个子转移$(X,\sigma)\subseteq\Sigma_l$ 称为**有限型子转移**（subshift of finite type, SFT）是指存在有限个词$\mathcal{W}=\{w_1,\dots,w_k\}$ 使得$x\in X$ 当且仅当$\mathcal{W}$ 中词从不在x 中出现. $m=\max\{|w_i|:1\le i\le k\}$ 称为X 的阶数. 证明：
   (a) 任何有限型子转移都同构于阶数为2的有限型子转移.
   (b) 对于每个阶数为2的有限型子转移$(X,\sigma)\subseteq\Sigma_k$, 可以定义取值为$0,1$ 的$k\times k$ 矩阵$A=[a_{ij}]$：
   $$a_{ij}=1\ \Leftrightarrow\ (ij)\notin\mathcal{W}.$$
   证明：$\mathrm{Card}(L_n(X))=\sum_{i,j=0}^{k}a_{ij}^{(n)}$, 其中$L_n(X)$ 为X 长为n 的词全体.

§5.2 传递性

回复性是十分重要的动力学性质. 这是因为在研究自然现象时, 那些可以重复观察的现象才是我们最关心的. 从这节开始, 我们将研究各种回复性质.

§5.2.1 传递的定义

定义 5.2.1 *设(X,T) 为拓扑动力系统, $x\in X$.*

1. *如果$Tx=x$, 那么点$x\in X$ 称为*不动点*;*
2. *如果存在某个$n\in\mathbb{N}$ 使得$T^nx=x$ 成立, 那么点$x\in X$ 称为*周期点*. 满足$T^nx=x$ 的自然数n 称为x 的*周期*; 而满足$T^nx=x$ 最小的自然数n 称为x 的*最小周期*.*

我们用$\mathrm{Fix}(X,T)$ 或$\mathrm{Fix}(T)$ 表示系统(X,T) 的不动点全体, 用$\mathrm{Per}(X,T)$ 或$\mathrm{Per}(T)$ 表示系统(X,T) 的周期点全体.

设x 为以n 为最小周期的周期点, 那么$Y=\mathrm{orb}(x,T)=\{x,Tx,\cdots,T^{n-1}x\}$ 为有限集合. (Y,T) 成为一个拓扑动力系统. 这是一类最简单的拓扑动力系统, 并且每个点都是周期回复的.

定义 5.2.2 *设(X,T) 为拓扑动力系统, $x\in X$. 定义x 的*ω 极限集*$\omega(x,T)$ 为$\mathrm{orb}(x,T)$ 的全体极限点集, 即*

$$\omega(x,T)=\{y\in X:\exists n_i\to+\infty\ s.t.\ T^{n_i}x\to y\}=\bigcap_{n\ge0}\overline{\bigcup_{k\ge n}\{T^kx\}}.$$

如果$U,\ V\subseteq X$, 我们定义**回复时间集**为：

$$N(U,V)=\{n\in\mathbb{Z}_+:U\cap T^{-n}V\neq\emptyset\}.$$

定义 5.2.3 称拓扑动力系统(X,T) 或T 为**传递的**是指对X 的任意两个非空开集U, V 我们有$N(U,V)\neq\emptyset$.

如果存在点$x\in X$ 满足$\overline{\mathrm{orb}(x,T)}=X$, 那么我们称$(X,T)$ 为**点传递的**, 而称x 为一个**传递点**. X 的全体传递点记为Trans_T .

注记 5.2.4 *(1)* 当X 是一个任意的拓扑空间, T 是其上的一个连续自映射, 我们也可类似地引入传递和点传递的概念. 如果对于拓扑空间X 不加任何限制, 传递与点传递是不同的概念. 反例如下:

设$X=\{0\}\cup\{\frac{1}{n}:n=1,2,\cdots\}$（取$\mathbb{R}$的遗传拓扑）, 而$T:X\to X$ 定义为$T(0)=0$,$T(\frac{1}{n})=\frac{1}{n+1}$. 于是$\mathrm{Tran}_T(X)=\{1\}$, 但$(X,T)$ 不是拓扑传递的.

设$g:I\to I$,其中$I=[0,1]$,$g(x)=1-|2x-1|$ 为帐篷映射. 则$\overline{\mathrm{Per}(g)}=I$ [209]. 令$X=\mathrm{Per}(g)$ 及$f=g|_{\mathrm{Per}(g)}$. 于是(X,f) 为传递但是不为点传递的.

(2) 从后面定理的证明我们可以看到: 如X 没有孤立点, 则点传递推出传递; 如果X 为可分的第二纲集, 则传递推出点传递.

由于在定义一个动力系统(X,T) 时我们已经假设X 为紧度量的, 于是对一个动力系统而言总有传递推出点传递,而如果再加上没有孤立点, 则两概念等价.

记$\mathcal{F}_{\mathrm{inf}}$ 为$\mathbb{Z}_+$ 的全体无限子集组成的集合.

定理 5.2.5 设(X,T) 为拓扑动力系统, 则以下*(1)*到*(17)*等价.

(1) T 为拓扑传递的;

(2) 对X 任二非空开集U,V, 存在$n\in\mathbb{Z}_+$ 使得$T^nU\bigcap V\neq\emptyset$;

(3) 对X 任二非空开集U,V, 存在$n\in\mathbb{N}$ 使得$T^nU\bigcap V\neq\emptyset$;

(4) 对X 任二非空开集U,V, 存在$n\in\mathbb{Z}_+$ 使得$U\bigcap T^{-n}V\neq\emptyset$;

(5) 对X 任意非空开集U, 有$\bigcup_{n=1}^{\infty}T^nU$ 在X 中稠密;

(6) 对X 任意非空开集U, 有$\bigcup_{n=0}^{\infty}T^nU$ 在X 中稠密;

(7) 对X 任意非空开集U, 有$\bigcup_{n=1}^{\infty}T^{-n}U$ 在X 中稠密;

(8) 对X 任意非空开集U, 有$\bigcup_{n=0}^{\infty}T^{-n}U$ 在X 中稠密;

(9) X 的任意闭正不变子集或为X 本身, 或为无处稠密集;

(10) X 的任意开负不变子集或在X 中稠密, 或为空集;

(11) 存在x 使得$\omega(x,T)=X$;

(12) $\{x\in X:\omega(x,T)=X\}$ 为X 的稠密G_δ 子集;

(13) $Trans_T$ 为X 的稠密G_δ 子集;

(14) T 为满射且$Trans_T\neq\emptyset$;

(15) $\Omega(X,T)=X$ 且$Trans_T\neq\emptyset$;（其中$\Omega(X,T)$定义参见定义*5.5.5*.）

(16) 存在x 使得$orb(Tx,T)$ 在X 中稠密;

(17) 对X 中任二非空开集U,V, 有$N(U,V)=\{n\in\mathbb{Z}_+:U\cap T^{-n}V\neq\emptyset\}\in\mathcal{F}_{\mathrm{inf}}$, 即它为无限集;

(1)-(17) 推出*(18)*, 如果X 没有孤立点则*(1)-(18)* 等价;

(18) $Trans_T \neq \emptyset$.

注记 5.2.6 取$\{U_i\}_{i=1}^{\infty}$ 为X 的可数基, 则

$$Trans_T = \{x \in X : \overline{orb(x,T)} = X\} = \bigcap_{n=1}^{\infty}\bigcup_{m=0}^{\infty} T^{-m}U_n$$

$$\{x \in X : \omega(x,T) = X\} = \bigcap_{i=1}^{\infty}\bigcap_{m=0}^{\infty}\bigcup_{n=m}^{\infty} T^{-n}U_i.$$

根据这两个式子容易给出上定理的证明.

设$\pi : X \to Y$ 为拓扑动力系统(X,T) 到(Y,S) 的因子映射, π 为**极小的**是指X 为唯一满足$\pi(A) = Y$ 的非空闭不变子集A. 我们有:

定理 5.2.7 设$\pi : X \to Y$ 为拓扑动力系统(X,T) 到(Y,S) 的因子映射,则

1. 如果(X,T) 为传递的, 那么Trans_T 为X 的稠密G_δ 子集.
2. 如果T 为传递的, 那么S 也是传递的, 且$\mathrm{Trans}_T \subseteq \pi^{-1}(\mathrm{Trans}_S)$. 如果$\pi$ 还为极小的, 则等号成立.

证明. (1) 根据定理5.2.5.

(2) 前半部分易证, 我们证后半部分. 设π 为极小的且$y \in \mathrm{Trans}_S$, 于是对任意$x \in \pi^{-1}(y)$ 有$\pi(\overline{\mathrm{orb}(x,T)}) = Y$. 根据$\pi$ 的极小性, 我们就有$\overline{\mathrm{orb}(x,T)} = X$, 即$x \in \mathrm{Trans}_T$. □

对应于定理2.1.7, 我们有:

定理 5.2.8 拓扑传递的系统没有非常值的连续不变函数.

证明. 设$x \in \mathrm{Trans}_T$, $f \in C(X)$ 使得$f(Tx) = f(x)$. 于是$f(T^n x) = f(x), \forall n \in \mathbb{Z}_+$. 因为$\overline{\mathrm{orb}(x,T)} = X$, 所以$f(y) = f(x), \forall y \in X$, 即$f$ 为常值函数. □

例 5.2.9 (没有非常值的连续不变函数的非传递系统) 没有非常值的连续不变函数的系统不一定为传递的. 设

$$X = (\mathbb{T}^2 \times \{0\}) \cup (\mathbb{T}^2 \times \{1\})/(e,0) \sim (e,1),$$

其中e 为$\mathbb{T}^2$ 单位元. 即X 为将两个二位环面在单位元处粘合在一起. 设$A : \mathbb{T}^2 \to \mathbb{T}^2$ 为遍历自同构, 令

$$T : X \to X, T(x,0) = (Ax,0); T(x,1) = (Ax,1).$$

则$\mathbb{T}^2 \times \{0\}$ 和$\mathbb{T}^2 \times \{1\}$ 为T 的不变集, 所以(X,T) 不是传递的. 但是对于任何T 不变连续函数在$\mathbb{T}^2 \times \{0\}$ 和$\mathbb{T}^2 \times \{1\}$ 分别是常数, 由于$\mathbb{T}^2 \times \{0\}$ 和$\mathbb{T}^2 \times \{1\}$ 在$(e,0),(e,1)$ 处粘合, 所以这个常数是相等的.

§5.2.2 回复点

与传递性紧密联系在一起的一个概念是回复点.

定义 5.2.10 设(X,T) 为拓扑动力系统. $x\in X$ 称为一个**回复点**是指存在序列$n_i\to+\infty$ 使得$T^{n_i}x\to x$,亦即$x\in\omega(x,T)$.

以$\mathrm{Rec}(T)$ 记全体回复点的集合.

一个重要的事实是, 如果x为回复点, 那么$(\overline{\mathrm{orb}(x,T)},T)$ 为传递系统; 对传递系统(X,T), 我们有$\mathrm{Trans}_T\subseteq\mathrm{Rec}(T)$. 下面著名的Birkhoff 定理说明对动力系统, 回复点总是存在的.

定理 5.2.11 (Birkhoff回复定理) 每个拓扑动力系统都存在回复点.

我们将这个定理的证明放在下一节给出，在第六章的定理6.3.6 我们可以运用Poincaré回复定理得到Birkhoff回复定理.

明显地, $\mathrm{Per}(T)\subseteq\mathrm{Rec}(T)$. 如果$x$ 是周期为n 周期点, 那么$(\{x,Tx,\dots,T^{n-1}x\},T)$为传递系统.

对于回复点集,我们有如下基本性质：

定理 5.2.12 设$\pi:X\to Y$为拓扑动力系统(X,T)到(Y,S)的因子映射,则

1. $T(\mathrm{Rec}(T))=\mathrm{Rec}(T)$.
2. $\mathrm{Rec}(T^n)=\mathrm{Rec}(T),\ \forall n\in\mathbb{N}$.
3. $\pi(\mathrm{Rec}(T))=\mathrm{Rec}(S)$.

证明. 1. 首先易验证$T(\mathrm{Rec}(T))\subseteq\mathrm{Rec}(T)$. 现在设$x\in\mathrm{Rec}(T)$, 则存在序列$n\to+\infty$ 使得$T^{n_i}x\to x$. 由于X 紧致, 不失一般性, 我们设$T^{n_i-1}x\to y$. 于是$T(y)=x$ 且$T^{n_i}y=T^{n_i-1}x\to y$, 亦即$y\in\mathrm{Rec}(T)$ 且$T(y)=x$. 这样就有$\mathrm{Rec}(T)\subseteq T(\mathrm{Rec}(T))$.

2. 根据定义容易验证$\mathrm{Rec}(T^n)\subseteq\mathrm{Rec}(T)$. 我们现在设$x\in\mathrm{Rec}(T)$, 那么存在序列$n_i\to+\infty$ 使得$T^{n_i}x\to x$. 不失一般性我们可以假设$n_i=k_in+j$,$0\le j\le n-1$. 于是$x\in\overline{\mathrm{orb}(T^jx,T^n)}$, 进而

$$\overline{\mathrm{orb}(x,T^n)}\subseteq\overline{\mathrm{orb}(T^jx,T^n)}.$$

用T^j 作用于此包含关系, 且注意到$x\in\overline{\mathrm{orb}(T^jx,T^n)}$, 我们有

$$\overline{\mathrm{orb}(x,T^n)}\subseteq\overline{\mathrm{orb}(T^jx,T^n)}\subseteq\ldots\subseteq\overline{\mathrm{orb}(T^{(n-1)j}x,T^n)}\subseteq\overline{\mathrm{orb}(T^{nj}x,T^n)}.$$

于是$x\in\overline{\mathrm{orb}(T^{jn}x,T^n)}$, 尤其有$x\in\mathrm{Rec}(T^n)$. 这样就有$\mathrm{Rec}(T)\subseteq\mathrm{Rec}(T^n)$.

3. $\pi(\mathrm{Rec}(T))\subseteq\mathrm{Rec}(S)$ 是明显的. 现在设$y\in\mathrm{Rec}(S)$ 且$B=\omega(y,S)$. 则$A'=\pi^{-1}B$ 为非空闭不变子集, 记$\mathcal{A}$ 为满足$\pi(A)=B$ 的A' 的闭不变子集的全体. 由Zorn 引理, 存在$A\in\mathcal{A}$

为包含关系下的极小元. 因为$\pi(A)=B$, 所以存在$x\in A$ 使得$\pi(x)=y$. 根据A 的选取, 我们有$\omega(x,T)=A$, 尤其有$x\in \mathrm{Rec}(T)$. 这样我们就完成了整个证明. □

一个拓扑动力系统称为**完全传递的**, 如果对任意$n\in\mathbb{N}$, 系统(X,T^n) 仍为传递的. 易见传递系统不必为完全传递的, 一个简单的例子为周期为2的周期轨. 一般而言我们有:

定理 5.2.13 设(X,T) 为传递的拓扑动力系统及$n\in\mathbb{N}$. 那么存在$k\in\mathbb{N}$ 使得$k|n$ 且有分解

$$X=X_0\cup X_1\cup\ldots\cup X_{k-1},$$

其中当$i\neq j$时$X_i\neq X_j$, (X_i,T^n) 为传递的且$T(X_i)=X_{i+1(\mathrm{mod}\ k)}$.

证明. 设$x\in \mathrm{Trans}_T$, 则由定理5.2.12, T^jx 为T^n 的回复点, 其中$0\le j\le n-1$. 令

$$Y_j=\overline{\mathrm{orb}(T^jx,T^n)},$$

则(Y_j,T^n) 为传递的且$T(Y_i)=Y_{i+1(\mathrm{mod}\ n)}$. 设$k$ 为满足$j\neq 0$ 和$Y_0=Y_j$ 的最小自然数. 则易验证$k|n$ 且当$0\le i<j\le k-1$ 时$Y_i\neq Y_j$. 于是$X_j=Y_j$, $0\le j\le k-1$,即为所求. □

§5.2.3 一些例子

为得到更多传递系统的例子, 我们先回顾**符号系统**的概念. 设$k\geq 2$为自然数且记$A=\{0,1\ldots,k-1\}$. 之前我们讨论的是$A^{\mathbb{Z}}$ 和$A^{\mathbb{Z}_+}$, 我们在这里取$A^{\mathbb{N}}$. $A^{\mathbb{N}}$ 与$A^{\mathbb{Z}_+}$ 没有本质区别. 赋予A以离散拓扑,而

$$\Sigma_k=A^{\mathbb{N}}=\{(x_1,x_2,\ldots):x_i\in A,i\in\mathbb{N}\}$$

取乘积拓扑. 则Σ_k为紧致可度量空间,一个相容的度量为

$$d(x,y)=\begin{cases}0, & \text{如果 } x=y,\\ \frac{1}{i}, & \text{如果 } x\neq y \text{ 且 } i=\min\{j:x_j\neq y_j\}.\end{cases}$$

定义

$$\sigma:\Sigma_k\to\Sigma_k,\ (x_1,x_2,x_3,\ldots)\mapsto(x_2,x_3,\ldots).$$

易见σ为连续的满射, 称(Σ_k,σ)或σ为**全转移**或直接称之为转移. 如果Y为Σ_k的非空闭正不变子集, 那么称(Y,σ)为**子转移**.

命题 5.2.14 (Σ_k,σ) 为传递的, 并且$\mathrm{Per}(T)$ 在Σ_k中稠密.

首先我们有一个简单但重要的结论:

命题 5.2.15 $x\in\Sigma_k$为回复点当且仅当每个x中的词在x中出现无限多次.

下面举一个具体的例子:

例 5.2.16 设$A_1 = (1), A_2 = (101), \ldots, A_{n+1} = A_n 0^{(n)} A_n$,则$x = \lim_{n\to\infty} A_n^\infty$ 为一个非周期点的回复点.

定理 5.2.17 设$A : \mathbb{T}^n \to \mathbb{T}^n$ 为环面遍历自同构. 那么它为传递的, 且

$$\mathrm{Per}(A) = \{(x_1, \ldots, x_n) \in \mathbb{T}^n : x_j \text{ 为有理数}\}.$$

即使A不是遍历的, 我们仍有

$$\mathrm{Per}(A) \supseteq \{(x_1, \ldots, x_n) \in \mathbb{T}^n : x_j \text{ 为有理数}\},$$

于是周期点稠密.

证明. 设A 为自同构. 我们用$(\mathbb{S}^1)^n$ 记环面. 设$\vec{w} = (w_1, \ldots, w_n) \in (\mathbb{S}^1)^n$, w_i 为单位根. 存在$k \in \mathbb{N}$ 使得$\vec{w}^k = e$, e 为$(\mathbb{S}^1)^n$单位元素. 对于每个$k \in \mathbb{N}$,

$$Y_k = \{\vec{z} \in (\mathbb{S}^1)^n : \vec{z}^k = e\}$$

为$(\mathbb{S}^1)^n$ 的有限子群, 并且$AY_k = Y_k$. 于是Y_k 每个元素都是A的周期点. 尤其$\vec{w}$ 为A 周期点.

下面设A 遍历, 此时我们用$\mathbb{T}^n$ 表示环面. 设$\vec{x} = (x_1, \ldots, x_n) \in \mathbb{T}^n$ 为A周期点, 即存在$k \in \mathbb{N}$ 使得

$$A^k \vec{x} = \vec{x} + \mathbb{Z}^n.$$

用矩阵表达, 即存在$\vec{p} = (p_1, \ldots, p_n) \in \mathbb{Z}^n$ 使得

$$\left([A]^k - I\right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}.$$

因为A 遍历, 矩阵$[A]^k - I$ 为可逆的整系数方阵, 其逆为有理系数方阵. 尤其x_i 为有理数. 于是每个周期点$\vec{x} = (x_1, \ldots, x_n)$ 中$x_i \in \mathbb{Q}$. □

§5.2.4 回复性与IP集

对于一个拓扑动力系统 (X, T), 设$x \in X$ 及$U \subseteq X$, 令x 进入U 的**回复时间集**为

$$N(x, U) = \{n \in \mathbb{Z}_+ : T^n x \in U\}.$$

从定理5.2.12我们看到如果x 为回复点, 那么它也是$T^n (n \in \mathbb{N}$）的回复点. 由此可见回复点的回复时间集并不是任意的. 为刻画回复点的回复时间, 首先回顾IP子集的概念：$A \subseteq \mathbb{N}$为**IP集**是指存在正整数序列$p_1, p_2, \ldots$使得

$$A = \{p_{i_1} + \ldots + p_{i_k} : i_1 < \ldots < i_k,\ k \in \mathbb{N}\}.$$

此时称A 为由$p_1, p_2, \ldots$ 生成的, 记为$\mathrm{FS}(\{p_i\})$. 记集合$\{A \subseteq \mathbb{N} : A \text{ 包含一个IP集}\}$ 为$\mathcal{F}_{\mathrm{ip}}$. $A \subseteq \mathbb{N}$ 称为IP***集**是指A 与任何IP 集相交非空.

需要注意的是, 在IP 集的定义中我们并没有要求p_i 为互异的.

定理 5.2.18 设(X,T) 为拓扑动力系统. 如果$x_0 \in \mathrm{Rec}(T)$, 那么对每个$\delta > 0$, $N(x_0, B_\delta(x_0))$包含了一个*IP*集.

反之,如果$R \subseteq \mathbb{N}$ 为一个*IP*集, 那么存在传递系统(X,T) 和$x_0 \in \mathrm{Trans}_T$ 使得

$$R \cup \{0\} \supseteq N(x_0, B_1(x_0)).$$

证明. 设$x_0 \in \mathrm{Rec}(T)$ 及$\delta > 0$. 取p_1 满足

$$d(T^{p_1}x_0, x_0) < \delta. \tag{5.2.1}$$

取$\delta_2 > 0$ 使得$\delta_2 \leq \delta$ 且满足

$$d(x, x_0) < \delta_2 \Rightarrow d(T^{p_1}x, x_0) < \delta. \tag{5.2.2}$$

对此δ_2 取p_2 使得

$$d(T^{p_2}x_0, x_0) < \delta_2. \tag{5.2.3}$$

由式子(5.2.1), (5.2.2) 和(5.2.3)我们就有

$$d(T^m x_0, x_0) < \delta \tag{5.2.4}$$

对$m = p_1, p_2$ 及$p_1 + p_2$ 均成立. 我们继续这个归纳过程. 假设$p_1, \ldots, p_n$ 已经取定并且(5.2.4)式对所有$m = p_{i_1} + \ldots + p_{i_k}$,$1 \leq i_1 < \ldots < i_k \leq n$, $k \in \mathbb{N}$ 均成立. 我们再取$\delta_{n+1} \leq \delta$使得一旦$d(x, x_0) < \delta_{n+1}$, 那么

$$d(T^m x, x_0) < \delta \tag{5.2.5}$$

对所有上述m 成立. 于是如果取p_{n+1} 使得

$$d(T^{p_{n+1}}x_0, x_0) < \delta_{n+1}, \tag{5.2.6}$$

那么(5.2.4)将还对形如$m + p_{n+1}$ 及p_{n+1} 的指数成立. 这样我们完成了归纳过程, 并且易见由$p_1, p_2, \ldots$ 生成的IP 集包含在$N(x_0, B_\delta(x_0))$中.

反之, 设$R \subseteq \mathbb{N}$ 为由$p_1, p_2, \ldots$ 生成的IP集. 如果$H_1, H_2, \ldots$ 为$\mathbb{N}$ 互不相交的子集, 令$p'_n = \sum_{i \in H_n} p_i$, 那么由$p'_1, p'_2, \ldots$ 生成的IP 集为R 的子集. 我们可以选取H_n 使得

$$p'_{n+1} > p'_1 + p'_2 + \ldots + p'_n.$$

这样, 不失一般性, 我们可以直接假设$p_{n+1} > p_1 + p_2 + \ldots + p_n$.

我们在$\{0,1\}^{\mathbb{Z}_+}$ 中定义点x_0 使得

$$(x_0)_n = \begin{cases} 1, & \text{如果 } n = 0 \text{ 或者 } n \in R, \\ 0, & \text{如果 } n > 0 \text{ 并且 } n \notin R. \end{cases}$$

注意我们取$\{0,1\}^{\mathbb{Z}_+}$ 度量为：

$$d(x,y)=\begin{cases}0, & \text{如果 } x=y,\\ \frac{1}{i+1}, & \text{如果 } x\neq y \text{ 且 } i=\min\{j: x_j\neq y_j\}.\end{cases}$$

易见x_0 为转移σ 的回复点, 并且对$n>0$ 有：

$$d(\sigma^n x_0, x_0)<1 \Leftrightarrow (\sigma^n x_0)_0=1 \Leftrightarrow (x_0)_n=1 \Leftrightarrow n\in R.$$

由此,我们完成了整个证明. □

定义 5.2.19 *设$\mathcal{F}$为$\mathbb{Z}_+$的子集族,我们称$\mathcal{F}$具有***Ramsey 性质***,如果$F\in\mathcal{F}$且$F=F_1\cup F_2$,则必有$F_1\in\mathcal{F}$或$F_2\in\mathcal{F}$成立.*

下面是著名的Hindman定理, 我们在本书中不给出证明, 感兴趣的读者可以参见[95].

定理 5.2.20 *(Hindman 定理) $\mathcal{F}_{\mathrm{ip}}$具有Ramsey 性质.*

习　题

1. 证明定理5.2.5.
2. 证明：(Σ_k,σ) 为传递的, 并且$\mathrm{Per}(T)$ 在Σ_k中稠密.
3. 证明：如果(X,T) 为传递系统,那么要么X 为有限集, 要么X 为不可数集.
4. 证明：传递系统的自然扩充仍为传递系统.
5. 证明：一个IP集与一个IP*集的交为无限子集.

§5.3　极小性

在本节中我们将讨论一类特殊的传递系统：极小系统.

§5.3.1　极小性

定义 5.3.1 *拓扑动力系统(X,T) 称为***极小的***是指它不真包含任何闭不变子集.如果子系统(Y,T) 为极小的, 那么我们称子集Y为X的***极小集***. 如果一个点包含在某个极小集中, 那么就称它为一个***极小点***.*

在后面的某些章节中我们需要考虑一般群作用下的极小集. 它的定义是完全类似的. 设(X,G) 为动力系统, 系统(X,G) 为极小的是指它没有真的非空不变子集. 同样可以定义极小集和极小点. 下面的许多定理(例如定理5.3.2, 5.3.4及5.3.7) 对一般群作用系统仍成立, 请读者验证之.

定理 5.3.2 *设(X,T) 为拓扑动力系统, 则以下命题等价:*

1. *(X,T) 为极小的;*
2. *对任何$x\in X$, $\text{orb}(x,T)$ 为X 的稠密子集;*
3. *对每个非空开集U, 存在有限子集$A\subseteq\mathbb{Z}_+$ 使得$\bigcup_{n\in A}T^{-n}U=X$.*

证明. (1) ⇒ (2) 显然.

(2) ⇒ (3) 设$E=X\setminus\bigcup_{i=1}^{\infty}T^{-i}U$,则$E$为闭不变的. 于是由假设$E=\emptyset$.根据$X$的紧性,就有(3)成立.

(3) ⇒ (1) 设E为非空闭不变子集.则$U=X\setminus E$为开集且$T^{-1}U\subseteq U$. 如果$E\neq X$, 那么$U\neq\emptyset$. 由假设, 存在有限集$A\subseteq\mathbb{Z}_+$ 使得$\bigcup_{n\in A}T^{-n}U=X$. 于是$U=X$, 矛盾! 从而$E=X$, 亦即$(X,T)$ 为极小的. □

注记 5.3.3 *根据定理5.2.8,设$T:X\to X$ 为极小同胚, $f\in C(X)$. 如果$f\circ T=f$, 那么$f=const$.*

易见, 每个极小点为回复点, 于是作为下面定理的推论，我们得到Birkhoff 回复定理.

定理 5.3.4 *每个拓扑动力系统都存在极小集.*

证明. 设(X,T) 为拓扑动力系统, $\{U_i\}_{i=1}^{\infty}$ 为X 的一组可数基. 设$X_0=X$, 归纳地, 对$i\in\mathbb{N}$, 如果$\bigcup_{n=0}^{\infty}T^{-n}U_i\supseteq X_{i-1}$, 则令$X_i=X_{i-1}$; 否则令

$$X_i=X_{i-1}\setminus\bigcup_{n=0}^{\infty}T^{-n}U_i.$$

易见$X_i\neq\emptyset$ 为闭不变的. 设

$$X_\infty=\bigcap_{i=1}^{\infty}X_i.$$

则X_∞ 也为非空闭不变的. 而且对每个满足$U_i\cap X_\infty\neq\emptyset$ 的U_i 有

$$\bigcup_{n=0}^{\infty}T^{-n}(U_i\cap X_\infty)\supseteq X_\infty.$$

根据定理5.3.2,X_∞为极小集. □

注记 5.3.5 *上面证明是Weiss给出的[199], 它依赖于空间的可度量性. 如果没有这个条件, 我们需要用Zorn 引理来证明. 具体证明如下. 设*

$$\mathcal{M}=\{Y\subseteq X:Y\neq\emptyset,\ TY\subseteq Y,\ \overline{Y}=Y\}.$$

即X 全体非空子系统的全体. 因为$X\in\mathcal{M}$, 所以$\mathcal{M}\neq\emptyset$. 在包含关系下, $\mathcal{M}$ 为偏序集. 容易验证任何线性序子列的交为其最小元, 所以根据Zorn 引理, $\mathcal{M}$ 有极小元M. 根据定义, (M,T) 为极小的.

§5.3.2 极小性与syndetic集

我们在前一小节已经看到, 如果x 为回复点, U 为其邻域, 那么$N(x,U)$ 包含了一个IP集. 对于极小点这一特殊的回复点, 我们自然期望它的回复时间集有些特殊的性质.下面我们将会看到, 事实上也的确如此. 对于以n 为周期的周期点x, 它的回复时间集包含了子集$n\mathbb{Z}_+$. 前面我们已对这一概念进行推广：集合$A=\{a_1<a_2<\dots\}\subseteq\mathbb{Z}_+$ 称为**syndetic的**是指它具有有界的间距, 即存在$N>0$ 使得

$$a_{i+1}-a_i\le N,\ \forall i\in\mathbb{N}.$$

记全体syndetic 集为$\mathcal{F}_{\mathrm{s}}$. 与syndetic 子集关联的概念是thick 子集. 集合$A\subseteq\mathbb{Z}_+$ 称为**thick的**是指它包含了任意长的整数段, 亦即存在序列$n_i\to\infty$ 使得

$$A\supseteq\bigcup_{i=1}^{\infty}\{n_i,n_i+1,\dots,n_i+i\}.$$

记全体thick 集为$\mathcal{F}_{\mathrm{t}}$. 容易证明：一个子集为syndetic 的当且仅当它与任何thick 子集相交非空；一个子集为thick 的当且仅当它与任何syndetic 子集相交非空.

定义 5.3.6 设(X,T)为拓扑动力系统, 其中点$x\in X$ 称为一个**几乎周期点**是指对x 的任意邻域U, $N(x,U)\in\mathcal{F}_{\mathrm{s}}$. 记全体几乎周期点的集合为$\mathrm{AP}(T)$.

我们有：

定理 5.3.7 设$\pi:X\to Y$ 为拓扑动力系统系统(X,T) 到(Y,S) 的因子映射.则

1. 如果$x\in\mathrm{AP}(T)$, 那么$\overline{\mathrm{orb}(x,T)}$ 为X 的极小集.
2. 如果$M\subseteq X$ 为极小集, 那么$M\subseteq\mathrm{AP}(T)$.
3. 如果(X,T) 为极小系统, 那么(Y,S) 也为极小的.
4. 如果(X,T)为极小的,那么π为半开的,即对X的任意非空开集U, $\pi(U)$有非空的内部.

证明. 1. 设$x\in\mathrm{AP}(T)$ 及$A=\overline{\mathrm{orb}(x,T)}$. 如果$A$ 不是极小集, 则由定理5.3.4存在极小集$A_1\subseteq A$ 且$A_1\neq A$. 易见$x\notin A_1$. 取x 和A_1 的不交邻域U, V. 则$N(x,U)$ 为syndetic的. 由于T 连续及A_1 不变, $N(x,V)$为thick 的.这样$N(x,U)\cap N(x,V)\neq\emptyset$, 矛盾.

2. 设M 为极小集, $x\in M$ 并且U 为x 的邻域. 如果$N(x,U)$ 不是syndetic的, 那么存在$\{n_i\}$ 使得对任意$i\in\mathbb{N}$ 有$T^{n_i}x,\dots,T^{n_i+i}x\notin U$.不失一般性, 可设$\lim T^{n_i}x=y$. 则对每个$j\in\mathbb{N}$, 我们有$\lim T^{n_i+j}x=T^jy$. 于是对任意$j\in\mathbb{Z}_+$ 有$T^jy\notin U$, 所以$M=\overline{\mathrm{orb}(y,T)}\subseteq M\setminus U$, 矛盾!

3. 假设(X,T) 为极小的. 设$Y_1\subseteq Y$ 为非空闭不变子集. 因为$\pi^{-1}(Y_1)$ 为非空闭不变的, 所以它为全空间. 于是$Y_1=\pi(X)=Y$. 亦即,(Y,S) 为极小的.

4. 设U,V 为X 的非空开集, 且满足$\overline{V}\subseteq U$. 根据定理5.3.2, 存在有限集$B\subseteq \mathbb{Z}_+$ 使得$\bigcup_{n\in B}T^{-n}\overline{V}=X$. 于是有$\bigcup_{n\in B}\pi(T^{-n}\overline{V})=Y$, 即$\bigcup_{n\in B}S^{-n}\pi(\overline{V})=Y$. 根据Baire 定理, 存在$n\in B$ 使得$S^{-n}\pi\overline{V}$ 内部非空. 易验证$\pi(U)(\supseteq \pi\overline{V})$的内部是非空的. □

注记 5.3.8 由上定理知道, 极小点与几乎周期点是同一回事, 后面我们经常会混用这两个概念.

需要注意的是, 对于非紧空间上的动力系统而言极小点与几乎周期点是不一样的.

下面为集合$\mathrm{AP}(T)$ 的若干性质:

定理 5.3.9 设$\pi:X\to Y$ 为拓扑动力系统系统(X,T) 到(Y,S) 的因子映射.则

1. $T(\mathrm{AP}(T))=\mathrm{AP}(T)$.
2. $\mathrm{AP}(T^n)=\mathrm{AP}(T)$, $\forall n\in\mathbb{N}$.
3. $\pi\mathrm{AP}(T)=\mathrm{AP}(S)$.

证明. 1. 首先易见$T(\mathrm{AP}(T))\subseteq \mathrm{AP}(T)$. 现设$x\in\mathrm{AP}(T)$, 则由定理5.3.7 $A=\overline{\mathrm{orb}(x,T)}$ 为极小的. 因为$T:A\to A$ 为满射,所以存在$y\in A$ 使得$T(y)=x$. 在由定理5.3.7 得到$y\in\mathrm{AP}(T)$.

2. 根据定理5.2.13和定理5.3.7即可得到结论.

3. 首先易见$\pi\mathrm{AP}(T)\subseteq\mathrm{AP}(S)$. 现设$y\in\mathrm{AP}(S)$ 且$A=\overline{\mathrm{orb}(y,T)}$. 则$A$ 为极小集. 由于$B=\pi^{-1}A$ 为非空闭不变的, 存在极小集$C\subseteq B$. 易见$\pi(C)=A$, 于是存在$x\in C$ 使得$\pi(y)=x$. 根据定理5.3.7, $x\in\mathrm{AP}(T)$. □

§5.3.3 极小系统的例子

例 5.3.10 下面几个例子都是关于群的:

- 设$\mathbb{S}^1$为复平面上的单位圆周, α为无理数. 定义$T:\mathbb{S}^1\to\mathbb{S}^1$ 为$z=e^{2\pi i\theta}\mapsto e^{2\pi i(\theta+\alpha)}$, $\forall z\in\mathbb{S}^1$. 则系统(X,T)为极小的.

- 加法机器为极小的.

  实际上, 只要注意到$\mathrm{orb}(\mathbf{0},T)=\{A0^{(\infty)}:A\in\bigcup_{i\geq 1}\{0,1\}^i\}$, 其中$\mathbf{0}=(0,0,\ldots)$, 就不难说明对任意$x\in\Sigma_2$, $\mathrm{orb}(x,T)$ 为稠密的.

- 设G为紧致度量交换群及$g_0\in G$. 令T为G在g_0下的转移映射, 即$T(g)=g_0g$, $\forall g\in G$. 我们一般称(G,T)为**Kroneker系统**.

  下面我们说明G的每个点都是几乎周期的. 设g_1为取定的一个几乎周期点, 而V为单位元的一个邻域. 如果$T^ng_1\in Vg_1$, 那么任何$g\in G$就有

  $$T^ng=g_0^ng=g_0^ng_1g_1^{-1}g\in Vg_1g_1^{-1}g=Vg.$$

  于是g为几乎周期的. 尤其如果(G,T)为传递的, 那么它为极小的.

一个常用且实用的事实为：

命题 5.3.11 设(Σ_k,σ) 为全转移，那么$x\in\Sigma_k$ 为几乎周期点当且仅当每个x 中出现的词出现在x 的位置组成的集合为*syndetic* 的.

例 5.3.12 (替换系统) 这里我们给出在符号系统中构造几乎周期点的一种方法. 设$A=\{0,1,\dots,k-1\}$, $w_0,\dots,w_{k-1}$ 为词且每个词包含了A 中所有字母.

定义

$$\phi:\bigcup_{i\geq 1}A^i\longrightarrow\bigcup_{i\geq 1}A^i$$

满足$\phi(i)=w_i$ 和$\phi(a_1,\dots,a_n)=\phi(a_1)\dots\phi(a_n)$. 自然地，这个映射$\phi$ 可以延拓到Σ_k 上. 容易说明，ϕ 的不动点w 实际上为几乎周期点，于是$\overline{\mathcal{O}(w)}$ 成为单边转移系统$(A^{\mathbb{Z}_+},\sigma)$ 中的极小集. 为得到一个双边几乎周期的序列，我们仅需将w 的右边序列按第一个坐标镜像对称到左边即可. 具体的讲，对$n\leq 0$, 令$w(n)=w(-n-1)$. 这样$w\in A^{\mathbb{Z}}$ 称为双边序列，它的轨道闭包成为$(A^{\mathbb{Z}},\sigma)$ 的一个极小系统. 一般我们称ϕ 为一个**替换**, 而由它产生的极小系统称为**替换系统**.

一个十分著名的例子就是所谓的***Morse*序列**. 设$w_0=01$ 以及$w_1=10$, 于是

$$0\longrightarrow 01\longrightarrow 0110\longrightarrow 01101001\longrightarrow 0110100110010110\longrightarrow\dots,$$

而单边*Morse* 序列为：$w=(0110100110010110\dots)$, 它是$(\{0,1\}^{\mathbb{Z}_+},\sigma)$ 的一个几乎周期点. 而$(X=\overline{\mathrm{orb}(w,\sigma)},\sigma)$ 称为***Morse*系统**.

例 5.3.13 (Morse系统) 上面我们已经给出了*Morse* 系统的一种构造方法. 下面我们给出别的方式来构造这个极小系统. 首先定义运算：$0'=1,1'=0$. 如果$a=a_1a_2\cdots a_n$ 为一个词，那么定义$a'=a_1'a_2'\cdots a_n'$. 下面我们归纳定义词如下：$a_1=0,a_2=a_1a_1'=01,a_3=a_2a_2'=0110,a_4=a_3a_3'=01101001$, 一般的，如果$a_n$ 已定义，那么令$a_{n+1}=a_na_n'$. 令$\omega=\lim_{n\to\infty}a_n$, 我们得到一个右边无穷的序列，它称为*(*单边的*)***Morse 序列**.

我们来证明这与上面用子替换来定义是等价的. 设b为一个词，那么令$b^*=\phi(b)$, 其中$0^*=01,1^*=10$. 设$b_1=a_1=0$, $b_{n+1}=b_n^*,n\in\mathbb{N}$. 下证明$b_n=a_n$, 这样就可以说明两种定义一致. 首先，观察到$(b^*)'=(b')^*$. 然后归纳假设$b_k=a_k,k\leq n$, 于是

$$b_{n+1}=b_n^*=a_n^*=(a_{n-1}a_{n-1}')^*=a_{n-1}^*(a_{n-1}^*)'=b_{n-1}^*(b_{n-1}^*)'=b_nb_n'=a_na_n'=a_{n+1}.$$

所以$b_n=a_n,\forall n\in\mathbb{N}$.

例 5.3.14 (Chacón 系统) 归纳定义

$$B_1=0010,\ B_2=0010001010010,\ \dots,B_{n+1}=B_nB_n1B_n.$$

设$w \in \Sigma_2$ 使得对于任何$n \in \mathbb{N}$

$$w_{[-l_n,l_n-1]} = w_{-l_n}w_{-l_n+1}\dots w_0w_1\dots w_{l_n-1} = B_nB_n,$$

其中$l_n = |B_n|$. 称w 为***Chacón* 序列**, 而$(X = \overline{\mathrm{orb}(w,\sigma)},\sigma)$ 称为***Chacón* 系统**. 可以证明*Chacón*系统为极小系统.

例 5.3.15 (Toeplitz 系统) $w \in \Sigma_k = \{0,1,\cdots,k-1\}^{\mathbb{Z}}$ 定义为: 对任意$n \in \mathbb{Z}$, 存在$p \geq 1$ 使得对任意$l \in \mathbb{Z}$ 成立$w(n+lp) = w(n)$. **Toeplitz 系统**定义为*Toeplitz* 序列在转移映射下的轨道闭包.

可以证明*Toeplitz* 系统为极小的, 并且一个子转移系统为*Toeplitz* 系统当且仅当它为加法机器的几乎一对一扩充.[1]

例 5.3.16 (Sturmian系统) 设$\mathbb{T} = \mathbb{R}/\mathbb{Z}$, α为无理数以及$\beta \in [0,1)$与α有理线性无关的. 定义$w \in \{0,1\}^{\mathbb{Z}}$为$w(n) = \mathbf{1}_{[0,\beta]}(n\alpha), \forall n \in \mathbb{Z}$. 令$X = \overline{\mathrm{orb}(w,\sigma)}$, 其中$\sigma$ 为转移映射. (X,σ) 为极小系统, 称之为**Sturmian 系统**.

§5.3.4 P系统与M系统

定义 5.3.17 拓扑动力系统(X,T) 称为
- 一个**P系统**是指它为传递的并且$\mathrm{Per}(T)$ 在X 中稠密;
- 一个**M系统**是指它为传递的并且$\mathrm{AP}(T)$ 在X 中稠密.

注记 5.3.18 *1.* 根据定义, 任何P系统是M系统. 全转移(Σ_k,T)为一个P系统; 任何极小但非周期的系统为M系统, 而非P系统.

2. 我们提及一个相反的例子: 设x 为例子*5.2.16* 中的回复点而$X = \overline{\mathrm{orb}(x,T)}$, 则系统$(X,T)$为传递的并且只有一个唯一的极小集$\{(0,0,\dots)\}$. 事实上, 如果$y$ 为X 的几乎周期点, 则存在序列$n_i \to +\infty$ 使得$T^{n_i}x \to y$. 于是y 中有任意长的*0* 词,这意味着$(0,0,\dots) \in \overline{\mathrm{orb}(y,T)}$. 于是由极小性就有$y = (0,0,\dots)$.

定义 5.3.19 集合$A \subseteq \mathbb{Z}_+$ 称为**piecewise syndetic的** 是指它为一个*syndetic* 集合与一个*thick* 集合的交. 记所有*piecewise syndetic* 集合为$\mathcal{F}_{\mathrm{ps}}$.

一个集合$A \subseteq \mathbb{Z}_+$ 称为**thickly syndetic的**是指对任意$n \in \mathbb{N}$ 存在一个*syndetic*集$\{s_1^n < s_2^n < \dots\}$ 使得

$$A \supseteq \bigcup_{j=1}^{\infty}\{s_j^n, s_j^n+1,\dots,s_j^n+n\}.$$

记全体*thickly syndetic* 集为$\mathcal{F}_{\mathrm{ts}}$.

---

[1]设X,Y 为拓扑空间, 映射$\pi: X \to Y$ 称为是几乎一对一的是指存在稠密G_δ 子集$X_0 \subseteq X$ 使得对于任何$x \in X_0$, $\pi^{-1}(\pi(x)) = \{x\}$.

易见,任何$\mathcal{F}_{\rm ps}$ 集与任何$\mathcal{F}_{\rm ts}$ 集相交非空. 下面的定理体现了它们与动力系统的联系.

定理 5.3.20 设(X,T) 为传递拓扑动力系统且$x \in {\rm Trans}_T$. 则

1. (X,T) 为M 系统当且仅当对x 的任意邻域U, $N(x,U) \in \mathcal{F}_{\rm ps}$.
2. 设K 为(X,T) 的极小集, 那么(X,T) 以K 为其唯一极小集当且仅当对K 的任意邻域U, $N(x,U) \in \mathcal{F}_{\rm ts}$.

证明. (1) 如果(X,T) 为M 系统, 则易见对x 的任何邻域U, $N(x,U) \in \mathcal{F}_{\rm ps}$. 这是因为对每个极小点$y \in U$, $N(y,U)$ 为syndetic 的, 并且存在序列$n_i \to +\infty$ 使得$T^{n_i}(x) \to y$.

下面我们假设对x 的任意邻域U, $N(x,U) \in \mathcal{F}_{\rm ps}$. 取$\varepsilon > 0$ 使得$\overline{B_\varepsilon(x)} \subseteq U$. 于是存在$p \in \mathbb{N}$ 及

$$\{m_j^i : i \in \mathbb{N}, 1 \le j \le i\} \subseteq N(x, B_{\varepsilon/2}(x))$$

使得$m_1^i < \ldots < m_i^i$ 且$m_{j+1}^i - m_j^i \le p$, $\forall 1 \le j \le i-1$. 令y 为$\{T^{m_1^i}(x)\}$ 的极限点.则易见$y \in B_\varepsilon(x)$ 且$N(y, B_\varepsilon(x))$为syndetic的. 设M 为y 在T 下轨道闭包中的极小集, 我们断言$M \cap \overline{B_\varepsilon(x)} \neq \emptyset$. 实际上, 如果$M \cap \overline{B_\varepsilon(x)} = \emptyset$, 那么存在开集$V \supseteq M$ 及开集$U_1 \supseteq \overline{B_\varepsilon(x)}$ 使得$U_1 \cap V = \emptyset$. 因为$N(y,V)$ 为thick 的, 所以$N(y,U_1)$ 不能为syndetic 的, 矛盾!

根据$M \cap U \neq \emptyset$ 以及(X,T) 的传递性,(X,T) 为M系统.

(2) 假设(X,T) 有唯一极小集K. 对任意K 的邻域U, 设$U_i \subseteq U$ 为K 的邻域,且满足如果$T^j(x) \in U_i$, 就有$T^{j+k}(x) \in U$, $\forall 1 \le k \le i$. 因为对每个i, $N(x,U_i)$ 为syndetic的, 所以$N(x,U) \in \mathcal{F}_{\rm ts}$.

反之, 如果T 有极小集K_1 使得$K_1 \cap K = \emptyset$. 那么对于K_1 的每个不交于U 的邻域V, $N(x,V)$ 为thick 的. 于是$N(x,U) \subseteq \mathbb{N} \setminus N(x,V)$ 不可能为syndetic 的, 矛盾! □

作为定理5.3.4的应用,我们有:

定理 5.3.21 设$\mathbb{N} = B_1 \cup \ldots \cup B_q$, 则存在$j \in \{1,2,\ldots,q\}$ 使得$B_j \in \mathcal{F}_{\rm ps}$.

证明. 设$A = \{1,\ldots,q\}$, 定义$w \in A^{\mathbb{N}}$ 为

$$w_n = i \text{ 当且仅当 } n \in B_i, \quad \forall n \in \mathbb{N}.$$

令$X = \overline{{\rm orb}(w,\sigma)}$, 设$\sigma$ 为转移映射. 则(X,σ) 为动力系统, 根据定理5.3.4在X 中存在极小点ξ. 假设j 在ξ 中出现, 那么j 出现的位置形成一个syndetic 集, 设其间距不大于l. 由于$\xi \in X$, 存在$\{m_i\}$ 使得$\sigma^{m_i} w$ 能任意接近ξ. 这意味着存在序列$n_i \to +\infty$ 使得

$$(\sigma^{m_i} w)_1 = \xi_1,\ (\sigma^{m_i} w)_2 = \xi_2, \ldots, (\sigma^{m_i} w)_{n_i} = \xi_{n_i}.$$

上式说明j 在$(w_{m_i+1}, \ldots, w_{m_i+n_i})$ 中以间距不大于l 出现. 于是$B_j \in \mathcal{F}_{\rm ps}$. □

类似于传递的情形, 我们可以定义**完全极小性**, 即对任意$n \in \mathbb{N}$, (X,T^n) 为极小的.圆周上的无理旋转是完全传递的, 而加法机器不是.

引理 5.3.22 设(X,T) 为极小动力系统, $m\in\mathbb{N}$. 那么X 可以分解为$l_m=l_m(X,T)\in\mathbb{N}$ 个不交的子集并

$$X=X^{m,1}\cup\ldots\cup X^{m,l_m},$$

其中l_m 整除m, $TX^{m,j}=X^{m,j+1\ (\mathrm{mod}\ l_m)}$, 并且每个系统$(X^{m,j},T^m), j=1,\ldots,l_m$, 为极小的.

§5.3.5 回复集

Birkhoff回复定理告诉我们每个拓扑动力系统都有回复点, 这个事实启发我们给出**回复集**的概念.

定义 5.3.23 子集$A\subseteq\mathbb{N}$称为**回复集**是指对每个动力系统(X,T),存在$\{n_i\}\subseteq A$使得$n_i\to+\infty$以及$x\in X$使得$T^{n_i}x\to x$.

于是Birkhoff定理说明$\mathbb{Z}_+$为回复集.由下面的定理5.3.24,我们可以看到每个thick集为回复集.令人惊奇的是,回复集与组合数学中的染色问题密切相关,而且也与syndetic集的差集联系在一起. 回复集的一个刻画为:

定理 5.3.24 集合A 为回复集当且仅当对每个*syndetic* 集S, $A\cap(S-S)\neq\emptyset$.

证明. $(\Rightarrow)$ 设A 为回复集且$S\in\mathcal{F}_{\mathrm{s}}$. 令$\Sigma_2=\{0,1\}^{\mathbb{Z}_+}$ 且σ 为转移映射. 定义$x\in\Sigma_2$ 为:

$$x_n=1 \text{ 当且仅当 } n\in S.$$

令$X=\overline{\mathrm{orb}(x,\sigma)}$. 设$Y$ 为X 的极小集且$U=\{y\in Y: y_0=1\}$. 由于S 为syndetic 的, $U\neq\emptyset$. 我们断言$N(U,U)\subseteq S-S$. 实际上, 令$n\in N(U,U)$, 则$U\cap\sigma^{-n}(U)\neq\emptyset$, 亦即, 存在$y\in U$ 使得$\sigma^n(y)\in U$. 这样就有某$m\in\mathbb{Z}_+$ 使得$x_m=1$ 且$x_{m+n}=1$, 即$n\in S-S$.

因为Y 为极小的, 根据定理5.3.2存在N 使得$Y=\bigcup_{i\le N}\sigma^{-i}U$. 这样对每个$y\in Y$, 存在$i\le N$ 使得$\sigma^i(y)\in U$. 因为A 为回复集, 所有存在$a\in A$ 及$y\in Y$ 使得$d(T^a y,y)$ 充分小且$\sigma^a(\sigma^i y),\sigma^i y\in U$, 其中$i\le N$ 使得$\sigma^i y\in U$. 这意味着$A\cap N(U,U)\neq\emptyset$, 继而$A\cap(S-S)\neq\emptyset$.

$(\Leftarrow)$ 设Y为极小集, U 为Y 满足$\mathrm{diam}(U)<\varepsilon$ 的非空开集, 其中$\varepsilon>0$. 对$y_0\in U$, 由于$N(U,U)=N(y_0,U)-N(y_0,U)$ 及$N(y_0,U)$ 为syndetic 的,我们有$A\cap N(U,U)\neq\emptyset$. 于是存在$a\in A$ 及$z_0\in U$ 使得$d(T^a z_0,z_0)<\varepsilon$. 令

$$B_\varepsilon=\{y\in Y: \text{存在} a\in A \text{ 使得} d(T^a y,y)<\varepsilon\}.$$

由于在每个非空开集U 中存在$y\in U\cap B_\varepsilon$, 所以B_ε 为稠密开集.

令

$$D=\bigcap_{n=1}^{\infty}B_{\frac{1}{n}}.$$

则D 为Y 的稠密子集且D 中每个点均按A 的子集元素回复. 因为每个系统均有极小子集, 所以A 为回复集. □

根据上面定理证明, 我们实际上得到:

定理 5.3.25 *集合A为回复集当且仅当对于任何极小系统(X,T) 以及非空开集$U \subseteq X$, 存在$n \in A$ 使得*

$$U \cap T^{-n}U \neq \emptyset.$$

§5.3.6　几个注记

最后我们以一个关于非传递点集的结构定理结束本节.

定理 5.3.26 *设(X,T) 为传递且非极小的拓扑动力系统, 则$X \setminus \mathrm{Trans}_T$ 为X 的稠密子集.*

证明. 设V 为X 的非空开集. 因为Trans_T 稠密, 存在$x \in V \cap \mathrm{Trans}_T$. 由于$x$ 为非极小点, 所以存在非空开集U 使得$\overline{U} \subseteq V$ 且$N(x,U)$ 不为syndetic的. 于是存在序列$n_i \to +\infty$ 使得

$$T^{n_i}(x) \in U \quad \text{且 } T^{n_i+j}(x) \notin U, \forall j = 1,2,\ldots,i.$$

根据紧性, 不妨设$T^{n_i}(x) \to y$. 于是$y \in \overline{U}$ 且$Ty, T^2y, \ldots \notin U$. 这样就有$y \in X \setminus \mathrm{Trans}_T$. 由$V$ 的任意性, $X \setminus \mathrm{Trans}_T$ 为X 的稠密子集. □

对于同胚, 作为半离散系统, $(X,T), (X,T^{-1})$ 许多时候表现的性质相差很大, 下面给出传递和极小性的一个注记.

注记 5.3.27 (双边传递与单边传递, 双边极小以及单边极小) *1. 双边点传递但是不是点单边传递例子. 设*

$$X = \{0,1\} \cup \{\frac{1}{n} : n \geq 2\} \cup \{1 - \frac{1}{n} : n \geq 2\},$$

$$T : X \to X,\ T(0) = 0, T(1) = (1)$$

其余点映到其左边的点. 那么(X,T) 为双边点传递但是不是点单边传递的. 此时$\Omega(T) = \{0,1\}$.

2. 等价的条件.

3. 双边极小等价于单边极小(留作习题).

习　题

1. 设(X,T) 为拓扑动力系统. 如果两个子集M_1 和M_2 均为极小的, 那么或者$M_1 = M_2$, 或者$M_1 \cap M_2 = \emptyset$.

2. 设(X,T) 为极小但是不是完全极小的. 取p_1 为最小的自然数使得(X,T^{p_1}) 非极小, 那么p_1 为素数, 并且(X,T^{p_1}) 所以极小子集为同构的. 设X_1 为(X,T^{p_1}) 的一个极小子集, 设$T_1=T^{p_1}|_{X_1}$. 如果(X_1,T_1) 不是完全极小的, 那么取最小的自然数p_2 使得$(X_1,T_1^{p_2})$ 不是极小的, 那么p_2 为素数, 且$p_2\geq p_1$. 继续此过程, 我们要么得到一个完全极小的系统(X_n,T_n), 要么得到素数序列$p_1\leq p_2\leq\ldots$.
3. 证明：双边极小等价于单边极小.
4. 证明: Chacón系统为极小系统.
5. 证明: Toeplitz 系统为极小的, 并且一个子转移系统为Toeplitz 系统当且仅当它为加法机器的几乎一对一扩充.
6. 证明：Sturmian系统为极小系统.

§5.4 拓扑混合性

在这一节,我们讨论另一类具有较强回复属性的传递系统——混合系统. 回顾两个动力系统的乘积系统$T\times S:X\times Y\to X\times Y$ 的作用定义为

$$T\times S(x,y)=(Tx,Sy),\ \forall (x,y)\in X\times Y.$$

§5.4.1 混合性与滤子

定义 5.4.1 *设(X,T)和(Y,S)为两个拓扑动力系统, 称它们为***弱不交的***是指乘积系统$(X\times Y,T\times S)$为传递的, 记为$(X,T)\curlywedge(Y,S)$.*

$\mathbb{Z}_+$的子集A称为**有限余的**是指$\mathbb{Z}_+\setminus A$为有限的. 我们用$\mathcal{F}_{\rm cf}$ 来表示全体有限余集构成的集合. 设(X,T) 为拓扑动力系统，$A,B\subseteq X$，令

$$N_T(A,B)=\{n\in\mathbb{Z}_+:A\cap T^{-n}B\neq\emptyset\}.\tag{5.4.1}$$

定义 5.4.2 *1. 一个拓扑动力系统(X,T)称为***(拓扑)弱混合的***是指它弱不交于自己,即$(X\times X,T\times T)$ 为传递的;*

*2. 一个拓扑动力系统(X,T)称为***(拓扑) 强混合的***是指对每个非空开集U 和V, $N_T(U,V)$ 为有限余的, 即存在$N>0$ 使得*

$$U\cap T^{-n}V\neq\emptyset,\ \forall n\geq N.$$

由定义知,强混合的系统必为弱混合的,而弱混合系统必为传递的（实际上也为完全传递的）. 圆周上的无理旋转为完全传递而非弱混合的,后面我们会看到许多弱混合而非强混合的例子. 为刻画混合性,我们引入如下概念:

定义 5.4.3 *$\mathbb{Z}_+$的一个子集族$\mathcal{F}$称为一个***滤子***是指它满足:*

1. $\emptyset \notin \mathcal{F}$;

2. 如果$F_1 \in \mathcal{F}$且$F_1 \subseteq F_2$,那么就有$F_2 \in \mathcal{F}$;

3. 对任意$F_1, F_2 \in \mathcal{F}$, $F_1 \cap F_2 \in \mathcal{F}$.

下面的定理被称为“Furstenberg相交引理”, 这是研究混合性的一个重要工具. 对于$\mathbb{Z}_+$ 的子集族$\mathcal{F}$, 我们令

$$[\mathcal{F}] = \{A \subseteq \mathbb{Z}_+ : \text{存在 } F \in \mathcal{F} \text{ 使得 } A \supseteq F\}.$$

定理 5.4.4 (Furstenberg相交引理) 一个拓扑动力系统(X,T) 为拓扑弱混合的当且仅当$[\mathcal{F}]$为滤子, 其中

$$\mathcal{F} = \{N_T(U,V) : U, V \text{ 为}X\text{的非空子集}\}.$$

证明. 如果$[\mathcal{F}]$ 为滤子, 那么对于X 的任意非空开集U_1, U_2, V_1, V_2 有

$$N_{T\times T}(U_1 \times U_2, V_1 \times V_2) = N_T(U_1, V_1) \cap N_T(U_2, V_2) \in [\mathcal{F}].$$

特别地, $N_{T\times T}(U_1 \times U_2, V_1 \times V_2) \neq \emptyset$. 所以$(X,T)$ 为弱混合的.

反之, 设(X,T) 为弱混合的, 且$N_T(U_1, V_1)$, $N_T(U_2, V_2) \in [\mathcal{F}]$. 由弱混合的定义, 存在$m \in N_T(U_1, U_2) \cap N_T(V_1, V_2)$. 设

$$A = U_1 \cap T^{-m}U_2,\ B = V_1 \cap T^{-m}V_2.$$

对任意$k \in N_T(A,B)$, 有

$$A \cap T^{-k}B = (U_1 \cap T^{-m}U_2) \cap T^{-k}(V_1 \cap T^{-m}V_2) = (U_1 \cap T^{-k}V_1) \cap T^{-m}(U_2 \cap T^{-k}V_2).$$

这意味着$U_1 \cap T^{-k}V_1 \neq \emptyset$ 且$U_2 \cap T^{-k}V_2 \neq \emptyset$. 于是

$$N_T(A,B) \subseteq N_T(U_1, V_1) \cap N_T(U_2, V_2).$$

证毕! □

下面的定理表明弱混合系统与集族$\mathcal{F}_t$有着密切关联.

定理 5.4.5 设(X,T)为拓扑动力系统,则以下各命题等价:

1. (X,T)为拓扑弱混合的;

2. 对任意非空开集U和V, $N_T(U,U) \cap N_T(U,V) \neq \emptyset$;

3. 对任意非空开集U和V, $N_T(U,U) \cap N_T(V,U) \neq \emptyset$;

4. 对任意非空开集U和V, $N_T(U,V)$为*thick*的.

证明. (1)⇒(2). 由定义即得.

(2)⇒(3). 设U, V为X的非空开集, 那么存在$n\in\mathbb{Z}_+$使得$V_1=V\cap T^{-n}U\neq\emptyset$.于是

$$N_T(U,U)\cap N_T(V,U)\supseteq N_T(T^{-n}U,T^{-n}U)\cap N_T(V_1,U)\supseteq N_T(V_1,V_1)\cap N_T(V_1,U)\neq\emptyset.$$

(3)⇒(1). 设U_1,U_2,U_3,U_4为X的非空开集, 那么存在$n_1,n_2\in\mathbb{Z}_+$使得

$$E=U_1\cap T^{-n_1}U_2\neq\emptyset$$

及

$$F=T^{-n_1}U_3\cap T^{-n_2}E\neq\emptyset.$$

同样地,存在$n_3\in\mathbb{Z}_+$使得$F\cap T^{-n_3}F\neq\emptyset$及$U_4\cap T^{-n_3}F\neq\emptyset$.令$n=n_2+n_3$,则有

$$T^{-n_1}(T^{-n}U_2\cap U_3)\supseteq T^{-(n_1+n)}U_2\cap T^{-n}U_1\cap T^{-n_1}U_3\supseteq T^{-n}E\cap F\supseteq T^{-n_3}F\cap F\neq\emptyset.$$

于是$T^{-n}U_2\cap U_3\neq\emptyset$.另外

$$T^{-n}U_1\cap U_4\supseteq T^{-(n_1+n)}U_2\cap T^{-n}U_1\cap U_4=T^{-n}E\cap U_4\supseteq T^{-n_3}F\cap U_4\neq\emptyset.$$

即$N_T(U_3,U_2)\cap N_T(U_4,U_1)\neq\emptyset$.

(1)⇒(4). 设U, V为X的非空开集. 由定理5.4.4,对任意$N\in\mathbb{N}$有

$$N_T(U,V)\cap N_T(U,T^{-1}V)\cap\ldots\cap N_T(U,T^{-N}V)\neq\emptyset.$$

所以$N_T(U,V)$为thick的.

(4)⇒(2). 设U, V为X的非空开集及$m\in N_T(U,V)$, 那么$W=U\cap T^{-m}V\neq\emptyset$. 由于$N_T(W,W)$为thick的, 存在$n\in\mathbb{Z}_+$使得$W\cap T^{-n}W\neq\emptyset$及$W\cap T^{-(n-m)}W\neq\emptyset$. 于是

$$U\cap T^{-n}U\supseteq W\cap T^{-n}W\neq\emptyset,\ U\cap T^{-n}V=U\cap T^{-(n-m)}T^{-m}V\supseteq W\cap T^{-(n-m)}W\neq\emptyset.$$

即, $N_T(U,U)\cap N_T(U,V)\neq\emptyset$. □

§5.4.2 拓扑mild 混合

介于弱混合性与强混合性之间有一类非常重要的混合性：mild混合.

定义 5.4.6 一个拓扑动力系统称为**拓扑mild混合的**是指它弱不交与任何传递系统.

由定义,两个mild混合系统的乘积系统仍为mild混合的. 实际上mild混合是严格介于强、弱混合之间的性质. 根据定理 5.4.5, 一个动力系统为拓扑弱混合的当且仅当对任意非空开集U,V, $N(U,V)$为thick的. 设$\mathcal{F}$为$\mathbb{Z}_+$的子集族, 定义$\mathcal{F}-\mathcal{F}=\{F-F:F\in\mathcal{F}\}$, 而子集$A\subseteq\mathbb{Z}_+$属于集合$(\mathcal{F}_{\rm ip}-\mathcal{F}_{\rm ip})^*$是指它与任何$\mathcal{F}_{\rm ip}-\mathcal{F}_{\rm ip}$中元相交. 对mild混合我们有：

定理 5.4.7 设(X,T) 为拓扑动力系统. 则(X,T) 为拓扑$mild$混合的当且仅当对任意非空开集U,V,

$$N_T(U,V)\in(\mathcal{F}_{\rm ip}-\mathcal{F}_{\rm ip})^*.$$

证明. 设(X,T) 为拓扑mild混合的, 即对任意传递系统(Y,S), $(X\times Y,T\times S)$仍为传递的. 因此(X,T)为弱混合的. 下证对任意IP子集F以及X的任意非空开集U_1,U_2, $N_T(U_1,U_2)\cap(F-F)\neq\emptyset$成立. 由定理5.2.18存在传递系统$(Y,S)$, 传递点$y\in Y$, 以及$y$的邻域$V$使得$N(y,V)\subseteq F$. 由于$(X,T)$为mild混合的, $(X\times Y,T\times S)$为传递的. 于是$N_T(U_1,U_2)\cap N_S(V,V)=N_{T\times S}(U_1\times V,U_2\times V)\neq\emptyset$. 由于

$$N_S(V,V)=N(y,V)-N(y,V)\subseteq F-F,$$

我们得到$N_T(U_1,U_2)\cap(F-F)\neq\emptyset$.

反之, 假设对X的任意两个非空开集U_1,U_2均有$N_T(U_1,U_2)\in(\mathcal{F}_{\rm ip}-\mathcal{F}_{\rm ip})^*$. 设$(Y,S)$为任一传递系统, 且设$U_1,U_2$为$X$的任意非空开集而$V_1,V_2$ 为Y任意的非空开集. 则由定义有

$$N_{T\times S}(U_1\times V_1,U_2\times V_2)=\{n\in\mathbb{Z}_+:(T\times S)^{-n}(U_2\times V_2)\cap(U_1\times V_1)\neq\emptyset\}.$$

由于(Y,S)为传递的, 存在$k\in\mathbb{Z}_+$使得$V=S^{-k}V_2\cap V_1$ 为X非空开集. 这样就有：

$$\begin{aligned}
& N_{T\times S}(U_1\times V_1,U_2\times V_2)\\
= & \{n\in\mathbb{Z}_+:(T\times S)^{-n}(U_2\times V_2)\cap(U_1\times V_1)\neq\emptyset\}\\
\supseteq & k+\{m\in\mathbb{Z}_+:(T^{-(m+k)}U_2\cap U_1)\times(S^{-(m+k)}V_2\cap V_1)\neq\emptyset\}\\
\supseteq & k+\{m\in\mathbb{Z}_+:(T^{-m}(T^{-k}U_2\cap U_1)\times(S^{-m}V\cap V)\neq\emptyset\}\\
= & k+N_T(U_1,T^{-k}U_2)\cap N_S(V,V).
\end{aligned}$$

由定理5.2.18, $N(V,V)$包含子集$(F-F)$, 其中F 为IP子集. 由于(X,T)为$(\mathcal{F}_{\rm ip}-\mathcal{F}_{\rm ip})^*$ 传递的,我们有$N_T(U_1,T^{-k}U_2)\cap N_S(V,V)\neq\emptyset$. 于是$N_{T\times S}(U_1\times V_1,U_2\times V_2)\neq\emptyset$. 因为$U_1,U_2$与$V_1,V_2$为任意的, 所以$(X\times Y,T\times S)$为传递的, 亦即$(X,T)$为mild混合的. □

习 题

1. 证明：全转移(Σ_k,T)为拓扑强混合的.
2. 一个拓扑动力系统称为**拓扑遍历的**是指对任意非空开集U,V, $N(U,V)\in\mathcal{F}_{\rm s}$; 一个拓扑动力系统称为**扩散的**是指它弱不交于任意极小系统; 一个拓扑动力系统称为**极端扩散的**是指它弱不交于任意拓扑遍历系统. 证明:
   (a) 任何弱混合系统弱不交与拓扑遍历系统.
   (b) 任何弱混合系统为极端扩散的.
   (c) 任何极端扩散系统为扩散的.
3. 证明：Chacón系统为弱混合的.

§5.5 其它不变集

在前面几节中,我们研究了传递性、极小性和混合性,介绍了周期点、几乎周期点和回复点的概念. 它们有如下包含关系：

$$\mathrm{Fix}(T) \subseteq \mathrm{Per}(T) \subseteq \mathrm{AP}(T) \subseteq \mathrm{Rec}(T).$$

在这一节,我们再介绍一些诸如ω极限集、非游荡集、链回复点等其它不变集.

§5.5.1 ω极限集

定义 5.5.1 设(X,T)为拓扑动力系统. 令$\omega(T) = \bigcup_{x\in X}\omega(x,T)$, 我们称$\omega(T)$为$(X,T)$的$\omega$**极限集**.

由定义易见, $x \in \mathrm{Rec}(T)$当且仅当$x \in \omega(x,T)$. 需要注意的是一个ω极限点不必为回复点. 例如, 令$x = (10100100010000\dots) \in \Sigma_2$, T为转移映射. 则$y = (100000\dots) \in \omega(x,T)$但它不是回复点. ω极限点集有如下性质：

定理 5.5.2 设(X,T) 为拓扑动力系统, $x \in X$.则

1. $\omega(x,T)$为非空闭集;
2. $T\omega(x,T) = \omega(x,T)$. 于是$T\omega(T) = \omega(T)$. 另外, 对每个$i \geq 0$ 和$n \in \mathbb{N}$, $T\omega(T^ix,T^n) = \omega(T^{i+1}x,T^n)$;
3. 对每个$n \in \mathbb{N}$ 均成立$\omega(x,T) = \bigcup_{i=0}^{n-1}\omega(T^ix,T^n)$, 于是对每个$n \in \mathbb{N}$ 有$\omega(T^n) = \omega(T)$;
4. $\mathrm{Rec}(T) \subseteq \omega(T)$.

证明. 1. 由X的紧性, $\omega(x,T)$为非空的.对$y \in X\setminus\omega(x,T)$, 存在$y$的邻域$U$使得$U\cap\mathrm{orb}(x,T)$有限. 于是存在$y$的邻域$U'$使得$U' \subseteq X\setminus\omega(x,T)$.此意味着$\omega(x,T)$为闭集.

2. 易验证, 留作习题.

3. 易见$\bigcup_{i=0}^{n-1}\omega(T^ix,T^n) \subseteq \omega(x,T)$. 下设$y \in \omega(x,T)$. 那么存在$n_i \to +\infty$使得$T^{n_i}x \to y$. 不失一般性, 设$n_i = k_in + r$,其中$0 \leq r \leq n-1$. 于是$y \in \omega(T^rx,T^n) \subseteq \bigcup_{i=0}^{n-1}\omega(T^ix,T^n)$.

4. 易证. □

注记 5.5.3 如果$T: X \to X$ 为同胚, 那么T^{-1} 的ω极限集称为α- **极限集**.

需要特别注意的是, 一般而言$\omega(T)$ 并非为闭的. 下面的定理告诉我们并非X的每一个闭不变子集都可以作为某个$x \in X$的ω极限集.

定理 5.5.4 设(X,T)为拓扑动力系统, $x \in X$. 如果存在一个周期点$p \in \omega(x,T)$使得p 为$\omega(x,T)$的孤立点, 那么$\omega(x,T)$为周期轨. 尤其, 如果$\omega(x,T)$为有限的, 那么它必为周期轨.

证明. 设p的周期为n.根据定理5.5.2 (3), 存在i使得$p\in\omega(T^ix,T^n)$.令$y=T^i(x)$及$g=T^n$.于是$p\in\omega(y,g)$为g的不动点,且为$\omega(y,g)$的孤立点.下证$\omega(y,g)=\{p\}$.

假设$\omega(y,g)\neq\{p\}$.令U为p的开邻域满足$U\cap(\omega(y,g)\setminus\{p\})=\emptyset$. 则存在$n_i\to+\infty$使得$g^{n_i}y\in U$且$g^{n_i+1}y\notin U$.因为$p$为$\omega(y,g)$的孤立点, $g^{n_i}y\to p$.由g的连续性, $g^{n_i+1}y\to p$.因为$g^{n_i+1}y\notin U$, 矛盾!

于是$\omega(y,g)=\{p\}$.这说明$\{p\}=\omega(T^ix,T^n)$, 从而$\omega(x,T)=\mathrm{orb}(p,T)$为周期轨. □

§5.5.2 非游荡集

定义 5.5.5 *设(X,T)为拓扑动力系统. 一个点$x\in X$称为***非游荡点***是指对x的任意邻域U, 存在$n\in\mathbb{N}$使得$U\cap T^{-n}U\neq\emptyset$. 如果$x$不是非游荡点, 那么称它为***游荡点***.*

记全体X非游荡点的集合为$\Omega(X,T)$或者$\Omega(T)$.

非游荡点集有如下性质：

定理 5.5.6 *设(X,T)为拓扑动力系统, 则*

1. *$\Omega(T)$为闭的;*
2. *$\overline{\omega(T)}\subseteq\Omega(T)$, 尤其$\Omega(T)$为非空的;*
3. *$T(\Omega(T))\subseteq\Omega(T)$.*

证明. 1. 因为$X\setminus\Omega(T)$为开集,所以$\Omega(T)$为闭集.

2. 设$x\in X$及$y\in\omega(x,T)$.则存在序列$n_i\to+\infty$使得$T^{n_i}x\to y$.对y的任意邻域U,存在$n_j>n_i$使得$T^{n_j}x,T^{n_i}x\in U$.于是$U\cap T^{-(n_j-n_i)}U\neq\emptyset$,继而$y\in\Omega(T)$.所以$\omega(T)\subseteq\Omega(T)$.因为$\Omega(T)$为闭集,所以$\overline{\omega(T)}\subseteq\Omega(T)$.

3. 令$x\in\Omega(T)$且U为Tx的邻域.由T的连续性,存在x的邻域V使得$TV\subseteq U$.因为$x\in\Omega(T)$, 存在$n\in\mathbb{N}$使得$V\cap T^{-n}V\neq\emptyset$.于是$U\cap T^{-n}U\supseteq TV\cap T^{-n}TV\supseteq T(V\cap T^{-n}V)\neq\emptyset$. □

需要指出的是$T(\Omega(T))=\Omega(T)$一般不成立. 另外迭代不变性也一般不成立,即$\Omega(T^n)=\Omega(T)$一般不再成立.

当$\Omega(T)=X$时, 我们有：

定理 5.5.7 *设(X,T)为一个拓扑动力系统. 如果$\Omega(T)=X$, 那么$\mathrm{Rec}(T)$为X的一个稠密G_δ集.*

证明. 对任意$\varepsilon>0$, 令

$$A_\varepsilon=\{x\in X:\text{存在 }n\geq 1\text{ 使得 }d(T^nx,x)<\varepsilon\}.$$

易见A_ε 为稠密开集. 设

$$A=\bigcap_{i=1}^{\infty}A_{\frac{1}{i}},$$

则$\mathrm{Rec}(T)=A$为稠密的G_δ集. □

设(X,T) 为拓扑动力系统, 则$\Omega(X,T)$ 为非空闭正不变子集, 于是$(\Omega(X,T),T|_{\Omega(X,T)})$ 为子系统. 同样的对$(\Omega(X,T),T|_{\Omega(X,T)})$ 也有非游荡集, 定义$\Omega_2(X,T)=\Omega(\Omega(X,T),T|_{\Omega(X,T)})$. 而$(\Omega_2(X,T),T|_{\Omega_2(X,T)})$ 仍为动力系统, 从而可定义其非游荡点集为$\Omega_3(X,T)$, $\cdots$.

一般的, 令γ 为大于$|X|$ 的最小序数, 如下定义$\{\Omega_\lambda(X,T):\lambda<\gamma\}$:

(1) 如$\lambda<\gamma$ 为后继序数, 则设

$$\Omega_\lambda(X,T)=\Omega(\Omega_{\lambda-1}(X,T),T|_{\Omega_{\lambda-1}(X,T)});$$

(2) 如λ 为极限序数, 则设

$$\Omega_\lambda(X,T)=\bigcap_{\mu<\lambda}\Omega_\mu(X,T).$$

由超限归纳法, 存在$\theta<\gamma$ 使得$\Omega_{\theta+1}(X,T)=\Omega_\theta(X,T)$. 于是对任何$\lambda>\theta$, 我们有$\Omega_\lambda(X,T)=\Omega_\theta(X,T)$, 即

$$\Omega(X,T)\supseteq\Omega_2(X,T)\supseteq\cdots\supseteq\Omega_\theta(X,T)=\Omega_{\theta+1}(X,T)=\cdots.$$

我们称$\Omega_\infty=\Omega_\theta$ 为(X,T) 的**中心**, θ 称为(X,T) 的**中心深度**(Depth of the centre). 由定理5.5.7, 系统的中心实际上就是$\overline{\mathrm{Rec}(T)}$. 另外, 对任意可数序数$\tau$, 我们都可以找到一个系统$(X,T)$, 其中心深度为$\tau$. 我们收集中心的性质如下:

定理 5.5.8 *设(X,T) 为动力系统, 则*

(1) $\Omega_\infty(X,T)\neq\emptyset$;

(2) 如X 为度量空间, 则(X,T) 中心深度可数;

(3) $(\Omega_\infty(X,T),T)$ 任一点为非游荡点;

(4) $\overline{\mathrm{Rec}(T)}=\Omega_\infty(X,T)$.

对系统(X,T), 一般我们把它的回复点全体的闭包$\overline{\mathrm{Rec}(T)}$ 称为系统的**Birkhoff 中心**. 如果系统(X,T) 的每个点都是非游荡点, 则称(X,T) 为**中心的**(Central). 此时系统的Birkhoff中心即为全空间. 对于传递系统, 如果它没有孤立点, 那它就是中心的.

例 5.5.9 ($\Omega_2(T)\neq\Omega(T)$) 设X 为极坐标下单位圆盘, 即

$$X=\{(r,2\pi\theta):0\le r\le 1,\theta\in[0,1)\}.$$

定义

$$T:X\to X,\ (r,2\pi\theta)\mapsto(r^{\frac{1}{2}},2\pi(\theta^2+1-r)\mod 2\pi).$$

则T 为同胚.

在单位圆周上,

$$(1,2\pi\theta)\mapsto(1,2\pi\theta^2).$$

它保持点$(1,0)$ 不动, 其余点逆时针方向向点$(1,0)$ 移动. 于是

$$\Omega(T|_{\partial X})=\{(1,0)\}.$$

圆盘上其余点在T 作用下都螺旋形向边界逼近. 我们断言

$$\Omega(T)=\{(0,0)\}\cup\partial X.$$

如果$(r,2\pi\theta)\not\in\{(0,0)\}\cup\partial X$, 则它为游荡点; 如果$(1,2\pi\theta)\in\partial X$, 那么对于它的任何连通邻域$U$, 其像$T^nU$ 也为连通的, 并且螺旋形逆时针向∂X 逼近, n 充分大后会与U 相交, 所以为非游荡点.

综上,

$$\Omega(T)=\{(0,0)\}\cup\partial X;\ \Omega_2(T)=\{(0,0),(1,0)\}.$$

§5.5.3　链回复点

设$T:X\to X$ 为同胚, $\alpha>0$, $a,b\in\mathbb{Z}$ (允许$a=-\infty$ 和$b=\infty$). 点列$\{x_i\}_{i=a}^b\subseteq X$ 称为T 的一个α**伪轨**, 如果

$$d(Tx_i,x_{i+1})<\alpha,\quad \forall i=a,\ldots,b.$$

我们称α伪轨$\{x_i\}_{i=a}^b$ 被从点y 出发的轨道β **跟踪**, 如果

$$d(T^iy,x_i)<\beta,\quad \forall i=a,\ldots,b.$$

α 伪轨$\{x_i\}_{i=-\infty}^{\infty}$称为**周期的**, 如果存在$n\in\mathbb{N}$ 使得

$$x_{i+n}=x_i,\quad \forall i\in\mathbb{Z}.$$

定义 5.5.10 (链回复点) 设$T:X\to X$ 为同胚. 点$x\in X$ 称为T 的**链回复点**, 如果对于任何$\alpha>0$都存在通过点x 的周期α 伪轨. 全体链回复点记为$\mathrm{CR}(T)$.

§5.5.4　总结

综上我们有:

$$\mathrm{Fix}(T)\subseteq\mathrm{Per}(T)\subseteq\mathrm{AP}(T)\subseteq\mathrm{Rec}(T)\subseteq\Omega(T)\subseteq\mathrm{CR}(T).$$

最后我们从二元关系的观点重新审视上面定义的概念.

设$W\subseteq X\times X$为二元关系, 记$W(x)=\{y\in X:(x,y)\in W\}$. 设$T:X\to X$为映射, 可以将它等同于二元关系$T=\{(x,Tx):x\in X\}\subseteq X\times X$. 我们称$\mathcal{O}T=\bigcup\limits_{n=0}^{\infty}T^n$为轨道关系.

此时$\mathcal{O}T(x)$即为x的轨道$Orb(x)$. 定义ω极限点关系为$\omega T=\{(x,y):y\in\omega(x,T)\}$. 虽然对任意$x\in X$, $\omega(x,T)$为闭的, 但是一般而言ωT并不是$X\times X$中闭集. 我们定义

$$\Omega T=\bigcap_{N\geq 0}\overline{\bigcup_{n\geq N}T^n}\subseteq X\times X.$$

则$y\in\Omega T(x)$当且仅当存在序列$\{n_k\}$及$\{x'_k\}\subseteq X$使得$n_k\to\infty$, $x'_k\to x$且$T^{n_k}(x'_k)\to y$. 显然, ΩT为包含ωT的闭集.

对关系W, 定义

$$|W|=\{x\in X:x\in W(x)\}.$$

则有:

$$\mathrm{Fix}(X,T)=|\mathcal{O}T|;\quad \mathrm{Rec}(X,T)=|\omega T|;\quad \Omega(X,T)=|\Omega T|.$$

习 题

1. 设$T:X\to X$ 为同胚. 那么T 为单边拓扑传递当且仅当T 为双边拓扑传递并且$\Omega(T)=X$.
2. 证明定理5.5.8.

§5.6 等度连续性

在这一节我们研究等度连续系统. 在本节中, 我们总是假设对于拓扑动力系统(X,T), 映射T 为满射.

§5.6.1 等度连续与一致几乎周期

定义 5.6.1 *拓扑动力系统(X,T) 称为***等度连续的***是指函数族$\{T^n:n\in\mathbb{Z}_+\}$ 为等度连续的,亦即对任意$\varepsilon>0$, 存在$\delta>0$ 使得当$d(x_1,x_2)<\delta$ 时$d(T^nx_1,T^nx_2)<\varepsilon,\forall n\in\mathbb{Z}_+$ 成立.*

如果令

$$d_T(x,y)=\sup_{n\in\mathbb{Z}_+}d(T^nx,T^ny),$$

后面我们会说明当(X,T) 为等度连续的时候, d_T 与d 是等价的, 并且它为T 不变的, 即

$$d_T(Tx,Ty)=d_T(x,y),\forall x,y\in X.$$

直观上讲, 等度连续系统中的任何两个不同点随着时间的推移将保持同样的差距,亦即它们的轨道是"平行"的. 所以从某种意义上讲, 等度连续性是最简单的一种动力学性状.

例 5.6.2 *1. 设$\mathbb{S}^1$为圆周, T为圆周上的旋转映射.则$(\mathbb{S}^1,T)$为等度连续的.*

2. 加法机器为等度连续的.

下面的定理说明等度连续系统中的每个点按一致的步调回复.

定义 5.6.3 一个拓扑动力系统(X,T) 称为**一致几乎周期的** 是指对任意$\varepsilon>0$, 存在$syndetic$集A 使得对任意$x\in X, n\in A$, $d(x,T^nx)<\varepsilon$ 成立.

我们用$C(X,X)$ 表示从X 到X 的全体连续映射.

定理 5.6.4 一个拓扑动力系统(X,T) 为等度连续的当且仅当它为一致几乎周期的.

证明. 设(X,T) 为等度连续的. 对任意$\varepsilon>0$, 根据Ascoli 定理$\{T^n:n\in\mathbb{Z}_+\}$ 在$C(X,X)$（取一致拓扑）中为相对紧的, 于是存在有限ε 网$\{T^n\}_{n=1}^k$. 亦即, 对任意$n\in\mathbb{Z}_+$, 存在$1\le j\le k$ 使得

$$\sup_{x\in X} d(T^nx,T^jx)<\varepsilon.$$

因T 是满射, 易见

$$\{n\in\mathbb{Z}_+:\sup_{x\in X} d(T^nx,x)<\varepsilon\}$$

为以k 为间距的syndetic 集, 从而(X,T) 为一致几乎周期的.

反之, 设(X,T) 为一致几乎周期的. 对任意$\varepsilon>0$, 存在syndetic 集A 使得对任意$x\in X, n\in A$, $d(x,T^nx)<\frac{\varepsilon}{3}$ 成立. 设A 间距为k. 取$\delta>0$ 使得对任意满足$d(x,y)<\delta$ 的x,y, $d(T^ix,T^iy)<\frac{\varepsilon}{3},\forall i=1,2,\cdots,k$.

下证对满足$d(x,y)<\delta$ 的x,y, $d(T^nx,T^ny)<\varepsilon,\forall n\in\mathbb{Z}_+$成立. 对任意$n\in\mathbb{Z}_+$, 存在$a\in A$ 及$i\in\{0,1,\dots,k-1\}$ 使得$n=a+i$. 于是

$$d(T^nx,T^ny)\le d(T^a(T^ix),T^ix)+d(T^ix,T^iy)+d(T^iy,T^a(T^iy))<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.$$

所以(X,T) 为等度连续的. □

§5.6.2 拓扑离散谱

从拓扑同构的角度看, 等度连续系统相当于一个紧致交换群上的旋转. 为了说明这点, 我们首先引入一些概念.

定义 5.6.5 设(X,T) 为拓扑动力系统,记$C(X)$为全体复值连续函数的集合. 称非零函数$f\in C(X)$ 为T 的**特征函数**是指存在$\lambda\in\mathbb{C}$ 使得$f(Tx)=\lambda f(x),\forall x\in X$. 此时称$\lambda$ 为相应于f 的**特征值**.

命题 5.6.6 设(X,T)为传递拓扑动力系统,则

1. 如$f\in C(X)$ 为以λ 为特征值的非零特征函数, 则$|\lambda|=1$ 且$|f|=c$, c为常数;
2. 如f,g 为同一特征值的特征函数, 则$f=cg$, 其中c 为常数;
3. 在$C(X)$ 中, 相应于不同特征值的特征函数为线性无关的;

4. T 的全体特征值形成S^1 的一个可数子群.

证明. 留作习题. □

定义 5.6.7 设(X,T) 为拓扑动力系统, 称之有**拓扑离散谱** 是指系统的特征函数张成了$C(X)$.

由前命题可见, 如(X,T) 为有拓扑离散谱的传递系统,则存在可数个特征值$\{\lambda_n\}_{n=1}^{\infty}$ 及线性无关组$\{f_n\}_{n=1}^{\infty} \subseteq C(X)$ 使得

$$\overline{\mathrm{span}\{f_n\}} = C(X) \ \ 且 f_n \circ T = \lambda_n f_n.$$

定理 5.6.8 (Halmos-von Neumann) 设(X,T)为可逆拓扑动力系统,则以下等价:

1. T 为极小等度连续的;

2. T 传递且存在X 上的度量使得T 在此度量下为等距的;

3. T 拓扑共轭于紧致交换群上的一个极小旋转;

4. T 极小且有拓扑离散谱;

5. T 传递且有拓扑离散谱.

证明. $(1) \Leftrightarrow (2)$: 易证.

$(2) \Rightarrow (3)$: 设ρ 为等距度量. 设x_0 为传递点, 所以$X = \overline{\mathrm{orb}(x_0,T)}$. 定义运算:

$$* : \mathrm{orb}(x_0,T) \times \mathrm{orb}(x_0,T) \to \mathrm{orb}(x_0,T),\ T^n x_0 * T^m x_0 \mapsto T^{n+m} x_0.$$

因为

$$\begin{aligned}
&\rho(T^n x_0 * T^m x_0, T^p x_0 * T^q x_0)\\
=&\rho(T^{n+m} x_0, T^{p+q} x_0)\\
\le&\rho(T^{n+m} x_0, T^{p+m} x_0) + \rho(T^{p+m} x_0, T^{p+q} x_0)\\
=&\rho(T^n x_0, T^p x_0) + \rho(T^m x_0, T^q x_0).
\end{aligned}$$

所以$*$ 为一致连续的, 于是可以唯一扩充为 $* : X \times X \to X$.

另外, 因为

$$\rho(T^{-n} x_0, T^{-m} x_0) = \rho(T^{m+n} T^{-n} x_0, T^{m+n} T^{-m} x_0) = \rho(T^m x_0, T^n x_0),$$

所以映射

$$T^n x_0 \mapsto T^{-n} x_0, \quad \mathrm{orb}(x_0,T) \to \mathrm{orb}(x_0,T)$$

为一致连续的, 于是可以唯一扩充为X 到X 的映射.

这样, 我们说明了$(X,*)$ 为拓扑群. 因为$\{T^n x_0 : n\in\mathbb{Z}\}$ 在X 中稠密, 所以$(X,*)$为交换的. 因为$T(T^n x_0) = Tx_0 * T^n x_0, \forall n\in\mathbb{Z}$, 所以$Tx = Tx_0 * x, \forall x\in X$. 这说明$(X,T)$ 为T 在群$(X,*)$ 上的Tx_0 旋转.

(3) $\Rightarrow$ (4) : 设X为紧交换群, 且$Tx = ax, \forall x\in X$. 则X的每个特征为X的特征函数(因为如果$f\in\hat{X}$, 那么$f(Tx) = f(ax) = f(a)f(x)$). 设A为由所有特征的线性组合生成的代数, 则$A\subseteq C(X)$. 由于A包含常值函数、对共轭运算封闭并且分离点, 所以根据Stone-Weierstrass 定理就有$\overline{A} = C(X)$.

(4) $\Rightarrow$ (5) : 显然.

(5) $\Rightarrow$ (2) : 由前面分析, 存在可数个特征值$\{\lambda_n\}_{n=1}^{\infty}$ 及线性无关组$\{f_n\}_{n=1}^{\infty}\subseteq C(X)$ 使得$|f_n| = 1$, $\text{span}\{f_n\} = C(X)$ 且$f_n\circ T = \lambda_n f_n$. 定义

$$\rho(x,y) = \sum_{n=1}^{\infty}\frac{|f_n(x) - f_n(y)|}{2^n},$$

它为X 上的伪度量. 由于$\text{span}\{f_n\} = C(X)$, 所以$\{f_n\}_{n=1}^{\infty}$ 分离X 的点. 于是$\rho(x,y) = 0$ 蕴含$x = y$, 亦即ρ 为度量.

因为$|\lambda_n| = 1$, 所以我们有：

$$\rho(Tx,Ty) = \sum_{n=1}^{\infty}\frac{|\lambda_n f_n(x) - \lambda_n f_n(y)|}{2^n} = \rho(x,y).$$

于是(X,ρ) 为等距的.

设d 为X 的原始度量, 下说明(X,d) 与(X,ρ) 等价. 我们仅需证明恒同映射$\text{id} : (X,d)\to(X,\rho)$ 连续即可(紧致空间到Hausdorff空间的连续双射是同胚).

对任意$\varepsilon > 0$, 取$N\in\mathbb{N}$ 使得$\displaystyle\sum_{n=N+1}^{\infty}\frac{2}{2^n} < \frac{\varepsilon}{2}$. 由于$f_n$ 连续, 所以存在$\delta > 0$ 使得

$$d(x,y) < \delta \Rightarrow |f_n(x) - f_n(y)| < \frac{\varepsilon}{2}, \quad \forall 1\le n\le N.$$

于是我们就有

$$d(x,y) < \delta \Rightarrow \rho(x,y) < \sum_{n=1}^{N}\frac{\varepsilon}{2^{n+1}} + \sum_{n=N+1}^{\infty}\frac{2}{2^n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

所以id 连续. □

§5.6.3 distal性质与等度连续性

对于一个系统, 其中的任何两个点随着时间的转移它们要么会在许多时候变得越来越接近, 要么它们将永远保持着一定的距离. 下面我们在数学上准确的给出上面的描述.

定义 5.6.9 设(X,T) 为动力系统, d 为X 上的度量. 两点$x,y \in X$ 称为**proximal的**是指$\liminf\limits_{n\to+\infty} d(T^n x, T^n y)=0$; 又如点对满足$\lim\limits_{n\to+\infty} d(T^n x, T^n y)=0$, 则称之为**渐近(asymptotic)的**.

两点$x,y\in X$ 如果不是*proximal* 的, 那么就称为**distal的**. 一个系统如果其任何两个不同的点都是*distal* 的, 就称之为**distal系统**.

distal 这个名称是由Gottschalk 引入的, 但是这个概念最早是由Hilbert 在试图给刚性群一个拓扑刻画时提出的. 从定义我们容易推导出等度连续系统一定是distal 的. 一个简单的distal 非等度连续的例子是圆盘上的不等速旋转. 设D 为平面上单位圆盘, 我们在平面上取极坐标(r,θ). 定义$T:D\to D$, $T(r,\theta)=(r,\theta+r)$. 则(D,T) 为distal 的但不是等度连续的. 但是要举一个极小的distal 但不是等度连续的例子要困难一些, Furstenberg[61] 给出了这样的例子, 并且他分析了二者的区别, 给出了极小distal 系统的结构定理.

给定拓扑动力系统$(X,T),(Y,S)$, 设$\pi: X\to Y$ 为扩充. 扩充称为**等度连续的**是指对任意$\varepsilon>0$, 存在$\delta>0$ 使得对满足$\pi(x_1)=\pi(x_2)$ 及$d(x_1,x_2)<\delta$ 的x_1,x_2 成立$d(T^n(x_1),T^n(x_2))<\varepsilon, \forall n\in\mathbb{Z}$. 等度连续扩充也称为**等距扩充**.

下面是著名的Furstenberg 的极小distal 系统结构定理.

定理 5.6.10 *[61]* 一个极小拓扑动力系统为*distal* 的当且仅当它为等度连续扩充的逆极限,即存在如下图表:

$$\{pt\}\leftarrow X_0 \xleftarrow{\pi_0} X_1 \xleftarrow{\pi_1} X_2 \xleftarrow{\pi_2} \cdots \xleftarrow{\pi_{n-1}} X_n \xleftarrow{\pi_n} X_{n+1} \xleftarrow{\pi_{n+1}} \cdots \longleftarrow X_\theta \xleftarrow{\pi_\theta} X.$$

其中$\{pt\}$ 指平凡系统,而每个π_i 为等度连续扩充.

习 题

1. 证明命题5.6.6.
2. 设
$$T:\mathbb{T}^2\to\mathbb{T}^2,\ (x,y)\mapsto(x+\alpha, y+2x+\alpha),$$
其中$\alpha\notin\mathbb{Q}$. 证明:
   (a) $(\mathbb{T}^2,T)$ 为distal 但是不是等度连续的极小拓扑动力系统.
   (b) 验证
   $$T^n(0,0)=(n\alpha, n^2\alpha).$$
   于是$\{n^2\alpha\}_{n\in\mathbb{N}}$ 在$\mathbb{T}$ 中稠密.

§5.7 可扩同胚

在本节我们介绍可扩同胚性.

§5.7.1 可扩同胚

定义 5.7.1 (可扩同胚) 设(X,d) 为紧致度量空间, $T:X\to X$ 为同胚. 称T 为**可扩的**(*expansive*), 是指存在$\delta>0$ 使得对于任何$x\neq y\in X$ 存在$n\in\mathbb{Z}$ 满足$d(T^nx,T^ny)>\delta$.

称满足上面条件的δ 为T 的**可扩常数**.

注记 5.7.2 我们将这个概念与初值敏感对比. 给定$\varepsilon>0$, T 称为在点x 处 ***Lyapunov*** ε**不稳定的** 是指对x 的任意邻域U, 存在点$y\in U$ 及$n\geq 0$ 使得$d(T^nx,T^ny)>\varepsilon$; T 称为在x **处不稳定**是指存在$\varepsilon>0$ 使得T 在x 处为*Lyapunov* ε 不稳定的.如果系统为处处不稳定的, 一般而言并不能保证存在一个$\varepsilon>0$ 使得所有点都是ε 不稳定的. 但对于传递系统, 这是能得到保证的, 即, 此时系统如为处处不稳定的, 则存在一个$\varepsilon>0$ 使得所有点都是ε 不稳定的. 这就是初值敏感的定义.

具体地讲, 称动力系统(X,T) 具有**初值敏感性**是指存在$\epsilon>0$ 使得对任意$\delta>0$ 和$x\in X$ 我们都能找到$y\in B(x,\delta)$ 和$n\in\mathbb{N}$ 满足$d(T^nx,T^ny)>\epsilon$.

1986 年*Devaney* 以初值敏感性为核心定义了一类重要的混沌. *Devaney* 称系统(X,T) 为混沌的, 如果它满足以下三条:

1. (X,T) 是传递的,

2. T 的周期点集在X 中稠密,

3. (X,T) 具有初值敏感性.

人们习惯把满足上述三条属性的系统称为**Devaney混沌的**. 可以证明传递的周期点稠密的非周期系统具有初值敏感性, 亦即, 我们有$(1)+(2)\Rightarrow(3)$. 鉴于这些原因, 我们干脆将上面定义中第二条舍弃, 而直接把满足第一, 第三个条件的系统成为混沌的. 准确的说,我们称系统为**Auslander-Yorke 混沌的** 是指它为初值敏感的传递系统. 例如,任何非极小的E系统（进而M系统）为*Auslander-Yorke* 混沌的.

§5.7.2 生成子

设X 为拓扑空间,$\mathcal{U}$、$\mathcal{V}$ 为X 的两个覆盖,$\mathcal{U}$ 和$\mathcal{V}$的**交** $\mathcal{U}\vee\mathcal{V}$ 定义为

$$\mathcal{U}\vee\mathcal{V}=\{U\cap V:U\in\mathcal{U},V\in\mathcal{V}\},$$

它仍为X 的覆盖. 类似地我们可以定义至多可数个覆盖$\{\mathcal{U}_n\}_n$ 的**交** $\bigvee_n\mathcal{U}_n$. 设$T:X\to X$ 为同胚,$\mathcal{U}$ 为X 的覆盖, 对于$n\in\mathbb{Z}$ 我们定义

$$T^{-n}\mathcal{U}=\{T^{-n}U:U\in\mathcal{U}\}.$$

显然$T^{-n}\mathcal{U}$ 仍为X 的覆盖.

定义 5.7.3 (生成子) 设X 为紧致度量空间, $T:X\to X$ 为同胚. X 的一个有限开覆盖α 称为T 的**生成子**(*generator*), 是指对于任何α 元素组成的序列$\{A_n\}_{n=-\infty}^{\infty}$ (即$A_n\in\alpha,\forall n\in\mathbb{Z}$

), 我们有

$$\mathrm{Card}\left(\bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_n}\right) \le 1.$$

如果上面条件换为

$$\mathrm{Card}\left(\bigcap_{n=-\infty}^{\infty} T^{-n}A_n\right) \le 1,$$

则称α 为T 的**弱生成子**(*weak generator*).

注记 5.7.4 设$\alpha = \{A_1, A_2, \ldots, A_k\}$ 为X 的有限开覆盖. 那么我们可以如下重新表述生成子的定义: 对于任何$(a_n)_{n\in\mathbb{Z}} \in \{1, 2, \ldots, k\}^{\mathbb{Z}}$, 我们有

$$\mathrm{Card}\left(\bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_{a_n}}\right) \le 1.$$

类似定义弱生成子.

设$\overline{\alpha} = \{\overline{A_1}, \ldots, \overline{A_k}\}$. 注意$\bigvee_{n\in\mathbb{Z}} T^{-n}\overline{\alpha}$ 中元素具有形式$\bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_{a_n}}$. 所以$\alpha$ 为生成子当且仅当$\bigvee_{n\in\mathbb{Z}} T^{-n}\overline{\alpha}$ 中元素为独点集或空集.

定理 5.7.5 设X 为紧致度量空间, $T : X \to X$ 为同胚. 那么T 具有生成子当且仅当它具有弱生成子.

证明. 由定义, 生成子自然为弱生成子. 下证反之. 设$\beta = \{B_1, \ldots, B_s\}$ 为T 的弱生成子, 设δ 为β 的Lebesgue 数. 设$\alpha = \{A_1, \ldots, A_t\}$ 为X 的开覆盖, 且$\max\limits_{1\le i\le t} \mathrm{diam}\overline{A_i} \le \delta$. 设$\{A_{a_n}\}_{n=-\infty}^{\infty}$ 为α 元素序列, 其中$(a_n)_n \in \{1, 2, \ldots, t\}^{\mathbb{Z}}$. 那么对于任何$n$, 存在$b_n \in \{1, \ldots, s\}$ 使得$\overline{A_{a_n}} \subseteq B_{b_n}$. 于是

$$\bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_{a_n}} \subseteq \bigcap_{n=-\infty}^{\infty} T^{-n}B_{b_n}.$$

尤其

$$\mathrm{Card}\left(\bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_{a_n}}\right) \le \mathrm{Card}\left(\bigcap_{n=-\infty}^{\infty} T^{-n}B_{b_n}\right) \le 1.$$

亦即α 为生成子. □

定理 5.7.6 设(X, d) 为紧致度量空间, $T : X \to X$ 为同胚. 设α 为生成子. 那么对于任何$\varepsilon > 0$, 存在N 使得

$$\mathrm{diam} \bigvee_{n=-N}^{N} T^{-n}\alpha \le \varepsilon.$$

反之, 对于任何$N > 0$, 存在$\varepsilon > 0$ 使得$d(x, y) < \varepsilon$ 蕴含

$$x, y \in \bigcap_{n=-N}^{N} T^{-n}A_n,$$

其中$A_{-N}, A_{-N+1} \ldots, A_N \in \alpha$.

证明. 设α 为生成子. 下证对于任何$\varepsilon > 0$, 存在N 使得

$$\operatorname{diam} \bigvee_{n=-N}^{N} T^{-n}\alpha \le \varepsilon.$$

否则, 存在$\varepsilon > 0$ 使得对于任何$j > 0$, 存在$x_j, y_j \in X$, $d(x_j, y_j) > \varepsilon$, 以及存在$A_{j,i} \in \alpha, -j \le i \le j$ 使得

$$x_j, y_j \in \bigcap_{n=-j}^{j} T^{-i}A_{j,i}.$$

根据紧致性, 不妨设$x_j \to x, y_j \to y, j \to \infty$ (否则取子列). 因为$d(x_j, y_j) > \varepsilon$, 所以$x \neq y$. 因为$\alpha$ 为有限的, 所以$A_{j,0}$ 有无穷个是相同的, 设为$A_0 \in \alpha$. 因为对于无穷多个j, $x_j, y_j \in A_0$, 所以$x, y \in \overline{A_0}$. 类似的, 对于任何$n \in \mathbb{Z}$, 存在无穷多个$A_{j,n}$ 相同的, 设为A_n, 同理$x, y \in T^{-n}\overline{A_n}$. 于是

$$x, y \in \bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_n},$$

这与α 为生成子矛盾.

下面证明对于任何$N > 0$, 存在$\varepsilon > 0$ 使得$d(x, y) < \varepsilon$ 蕴含$x, y \in \bigcap_{n=-N}^{N} T^{-n}A_n$, 其中$A_{-N}, \ldots, A_N \in \alpha$. 取$\delta > 0$ 为α 的Lebesgue 数. 取$\varepsilon > 0$ 使得$d(x, y) < \varepsilon$ 蕴含$d(T^i x, T^i y) < \delta, -N \le i \le N$. 于是如果$d(x, y) < \varepsilon$, 那么对于任何$-N \le i \le N$, 存在$A_i \in \alpha$ 使得$T^i x, T^i y \in A_i$. 由此

$$x, y \in \bigcap_{n=-N}^{N} T^{-n}A_n.$$

定理证毕! □

§5.7.3　可扩性与生成子

定理 5.7.7 (Reddy-Keynes-Robertson) 设(X, d) 为紧致度量空间, $T : X \to X$ 为同胚. 那么T 为可扩的当且仅当T 具有生成子, 当且仅当T 具有弱生成子.

证明. 根据定理5.7.5, 我们已经知道T 具有生成子当且仅当T 具有弱生成子. 下面证明它们等价于可扩性.

设T 为可扩同胚, δ 为其可扩系数. 设α 为任何直径小于δ 的有限开覆盖. 设

$$x, y \in \bigcap_{n=-\infty}^{\infty} T^{-n}\overline{A_n}, \quad A_n \in \alpha.$$

则

$$d(T^n x, T^n y) \le \delta, \quad \forall n \in \mathbb{Z}.$$

根据可扩性定义, $x = y$. 于是α 为生成子.

反之, 设α为弱生成子, 设δ为它的Lebesgue数. 如果对于任何$n \in \mathbb{Z}$有$d(T^n x, T^n y) \le \delta$, 那么我们需要证明$x = y$. 对于任何$n \in \mathbb{Z}$, 因为$d(T^n x, T^n y) \le \delta$, 所以存在$A_n \in \alpha$ 使得$T^n x, T^n y \in A_n$. 于是

$$x, y \in \bigcap_{n=-\infty}^{\infty} T^{-n} A_n.$$

因为α 为弱生成子, 所以$x = y$. 即T 可扩. □

注记 5.7.8 *1. 根据定理5.7.7, 可扩性不依赖与空间的度量选取, 但是可扩常数会依赖于度量的选取.*

2. 对于任何$k \neq 0$, T 可扩当且仅当T^k 可扩. 设α 为T 生成子, 那么$\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-(k-1)}\alpha$ 为T^k 的生成子; 反之显然.

3. 可扩性是同构不变性质.

4. 可扩系统的子系统也是可扩的.

5. 有限个可扩系统的乘积仍然为可扩的, 但是无穷个可扩系统的乘积不一定为可扩的.

注记 5.7.9 *对于连续映射$T: X \to X$ 我们也可以定义可扩性: 称T 为***正向可扩的***, 是指存在$\delta > 0$ 使得对于任何$x \neq y \in X$ 存在$n \in \mathbb{Z}_+$ 满足$d(T^n x, T^n y) > \delta$.*

§5.7.4 扩充同胚与符号扩充

例 5.7.10 (可扩系统的因子不一定可扩的例子1) *设$\mathbb{T}^2$ 为二维环面, 在$\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ 上定义等价关系:*

$$(x, y) + \mathbb{Z}^2 \sim (-x, -y) + \mathbb{Z}^2.$$

于是得到商空间为$\mathbb{S}^2 = \mathbb{T}^2/\sim$. 设

$$\phi : \mathbb{T}^2 \to \mathbb{S}^2$$

为商映射. 注意ϕ 除了在四个点$(1, 1), (1, -1), (-1, 1), (-1, -1)$ 外为二对一的.

设$A : \mathbb{T}^2 \to \mathbb{T}^2$ 为连续自同构, 易见通过ϕ它诱导了$\mathbb{S}^2$ 上同胚, 记为$T : \mathbb{S}^2 \to \mathbb{S}^2$. 根据定义, $(\mathbb{S}^2, T)$ 为$(\mathbb{T}^2, A)$ 的因子. 下说明如果矩阵$[A]$ 没有绝对值等于1 的特征值, 那么A 为可扩的, 但是T 不是. 与之前一样, 我们设$\widetilde{A} : \mathbb{R}^2 \to \mathbb{R}^2$ 为$[A]$ 对应的线性映射.

设矩阵$[A]$ 没有绝对值等于1 的特征值, 那么它有两个特征值λ_s, λ_u. 因为$|\det[A]| = 1$, 所以设

$$|\lambda_s| < 1, |\lambda_u| > 1.$$

此处s 指"stable", 而u 指"unstable". 设V_s 和V_u 为λ_s, λ_u 对应的特征空间. 对于$\varepsilon > 0$, 取点$(x, y) \in B_\varepsilon(0, 0)$. 考虑过$(x, y)$ 以$(0, 0)$ 为中心, 边平行于V_s, V_u 的平行四边形. 设它的四个顶点为

$$P_1(x, y),\ P_2(u, v),\ P_3(-x, -y),\ P_4(-u, -v).$$

其中$\overline{P_1P_2}, \overline{P_3P_4}$ 平行于V_u, $\overline{P_1P_4}, \overline{P_2P_4}$ 平行于V_s 则四个顶点, 以及整个平行四边形包含在球$B_{c\varepsilon}(0,0)$ 中, 其中c 为仅依赖于V_s, V_u 斜率的常数.

注意到$(u,v)-(x,y)\in V_u$, 所以

$$\|\widetilde{A}^n(x,y)-\widetilde{A}^n(u,v)\| = \lambda_u^n\|(x,y)-(u,v)\|, n\le 0.$$

因为$(-u,-v)-(x,y)\in V_s$, 所以

$$\|\widetilde{A}^n(x,y)-\widetilde{A}^n(-u,-v)\| = \lambda_s^n\|(x,y)-(-u,-v)\|, n\ge 0.$$

上面$\|\cdot\|$ 为$\mathbb{R}^2$ 的模, 设d 为环面上由$\|\cdot\|$ 诱导的度量, 那么我们得到

$$d(A^n(x,y),A^n(u,v)) < 2c\varepsilon, \forall n\le 0,$$

$$d(A^n(x,y),A^n(-u,-v)) < 2c\varepsilon, \forall n\ge 0.$$

如果$(\mathbb{S}^2,T)$ 有生成子$\gamma=\{C_1,\ldots,C_k\}$. 那么$\phi^{-1}\gamma$ 具有性质: 任何形如

$$\bigcap_{n=-\infty}^{\infty} A^{-n}\phi^{-1}\overline{C_{a_n}}$$

的子集至多会包含一个等价类. 取ε 充分小, 使得$c\varepsilon$ 为覆盖$\phi^{-1}\gamma$ 的*Lebesgue* 数. 根据上面分析, 我们知道存在某个形如

$$\bigcap_{n=-\infty}^{\infty} A^{-n}\phi^{-1}\overline{C_{a_n}}$$

的子集同时包含了(x,y) 和(u,v) 的等价类. 但是上面两个事实是矛盾的.

所以$(\mathbb{S}^2,T)$ 不可扩.

例 5.7.11 (可扩系统的因子不一定可扩的例子2) 设$T:\mathbb{S}^1\to\mathbb{S}^1, z\mapsto az$ 为无理旋转, 即a 非单位根. 对于$x,y\in\mathbb{S}^1$, 我们用$\overrightarrow{[x,y]}$ 表示从x 出发逆时针到y 的弧. 设

$$A_0=\{z=e^{2\pi ix}:\frac{1}{2}\le x\le 1\};\ A_0=\{z=e^{2\pi ix}:0\le x\le \frac{1}{2}\},$$

即$A_0=\overrightarrow{[-1,1]}, A_1=\overrightarrow{[1,-1]}$ 分别为下、上半闭圆周. 令

$$Y=\mathbb{S}^1\setminus\{a^n,-a^n:n\in\mathbb{Z}\}.$$

定义

$$\psi:Y\to\Sigma_2=\{0,1\}^{\mathbb{Z}}, z\mapsto(a_n)_{n\in\mathbb{Z}},$$

其中$(a_n)_{n\in\mathbb{Z}}$ 由下式定义:

$$T^nz\in A_{a_n}.$$

设$\Lambda=\psi(Y)$, 我们要证明ψ 为单射, 并且其逆可以扩充为连续映射

$$\phi:\overline{\Lambda}\to\mathbb{S}^1.$$

为此我们证明: 对于任何$\varepsilon>0$, 存在$N\in\mathbb{N}$ 使得如果$x,y\in Y$, 并且$(\psi(x))_n=(\psi(y))_n,\forall|n|\le N$, 那么$d(x,y)<\varepsilon$, 即$\psi^{-1}$ 为一致连续的.

设$\varepsilon>0$. 取$N\in\mathbb{N}$ 使得$\{a^n:|n|\le N\}$ 为$\mathbb{S}^1$ 的$\varepsilon/2$-网. 设$x,y\in Y$, 并且$(\psi(x))_n=(\psi(y))_n,\forall|n|\le N$. 根据条件, 对于$|n|\le N$, $T^nx=a^nx$ 与$T^ny=a^ny$ 都在A_0 中, 或者都在A_1 中. 这就排除了$y=-x$ 的可能性. 假设y 到x 逆时针方向弧长小于顺时针方向弧长. 根据$\{a^n:|n|\le N\}$ 为$\mathbb{S}^1$ 的$\varepsilon/2$-网, 存在某个$|n|\le N$, 使得a^nx 位于从*1* 出发逆时针方向长为ε 的开弧内. 此时$a^nx\in A_1$. 于是根据我们的假设, $a^ny\in A_1$, 并且y 在1 和a^nx 之间, 即$y\in\overrightarrow{[1,a^nx]}$. 于是$d(a^nx,a^ny)<\varepsilon$, 从而$d(x,y)<\varepsilon$.

综上, 我们将$\psi^{-1}:\Lambda\to\mathbb{S}^1$ 可以扩充为连续映射

$$\phi:\overline{\Lambda}\to\mathbb{S}^1.$$

容易验证ϕ 为因子映射. 根据例子*5.7.14*, 符号系统一定为可扩的, 而等度连续系统一定不是可扩的. 于是$(\overline{\Lambda},\sigma)$ 为可扩的, 但是其因子$(\mathbb{S}^1,T)$ 不是可扩的.

定理 5.7.12 设(X,d) 为紧致度量空间, $T:X\to X$ 为可扩同胚. 设δ 为可扩常数. 设$\gamma=\{C_1,\dots,C_r\}$ 为X 的任何直径小于δ 的有限覆盖（即$\mathrm{diam}\gamma<\delta$）. 那么

$$\mathrm{diam}\left(\bigvee_{j=-n}^{n}T^{-j}\gamma\right)\to 0,\ n\to\infty.$$

证明. 如果命题不成立, 那么存在$\varepsilon_0>0$, 序列$\{n_i\}_{i\in\mathbb{N}}\subseteq\mathbb{N}$, 以及$x_i,y_i\in X$ 使得$d(x_i,y_i)\ge\varepsilon_0$, 且

$$x_i,y_i\in\bigcap_{j=-n_i}^{n_i}T^{-j}C_{i,j},$$

其中$C_{i,j}\in\gamma$.

不妨设$x_i\to x,y_i\to y,i\to\infty$ (否则取子列). 由于$d(x_i,y_i)\ge\varepsilon_0$, 所以$d(x,y)\ge\varepsilon_0$. 因为$\gamma$ 有限, 所以无穷多个$C_{i,0}$ 取到γ 同一个元素, 记为C_{i_0}. 因为对于无穷多个i, $x_i,y_i\in C_{i_0}$, 所以$x,y\in\overline{C_{i_0}}$. 类似的, 无穷多个$C_{i,j}$ 取到γ 同一个元素, 记为C_{i_j}. 因为对于无穷多个i, $x_i,y_i\in T^{-j}C_{i_j}$, 所以$x,y\in T^{-j}\overline{C_{i_j}}$. 由于$\mathrm{diam}\gamma<\delta$, $d(T^jx,T^jy)\le\delta,\forall j\in\mathbb{Z}$, 根据可扩性$x=y$. 此与$d(x,y)\ge\varepsilon_0$ 矛盾. □

类似于例子5.7.11的证明, 可以证明任何可扩同胚都有符号扩充.

定理 5.7.13 设(X,d) 为紧致度量空间, $T:X\to X$ 为可扩同胚. 那么存在$k\in\mathbb{N}$ 以及子转移系统$\Omega\subseteq\Sigma_k$ 使得(Ω,σ) 为(X,T) 的扩充.

证明. 设δ 为T 的可扩常数. 我们首先构造一个覆盖$\gamma = \{C_1, \ldots, C_{k-1}\}$ 满足：

1. $\mathrm{diam}(\gamma) < \delta$,
2. $C_i \cap C_j = \partial C_i \cap \partial C_j, i \neq j \in \{1, 2, \ldots, k-1\}$,
3. $\bigcup\limits_{i=0}^{k-1} \partial C_i$ 内部为空.

为此, 我们先取一个半径为$\delta/3$ 的开球覆盖$\{B_0, B_1, \ldots, B_{k-1}\}$, 然后再改造成所需. 设$C_0 = \overline{B_0}$. 对于$n > 0$, 令

$$C_n = \overline{B_n} \setminus (B_0 \cup \ldots \cup B_{n-1}).$$

于是对于$i < j$, 我们有

$$\begin{aligned} C_i \cap C_j &= \partial C_j \cap C_i \quad (\because \mathrm{int}(C_j) = B_j \setminus \overline{(B_0 \cup \ldots \cup B_{j-1})}) \\ &= \partial C_j \cap \partial C_i \quad (\because \partial C_j \cap \mathrm{int}(C_i) \subseteq B_i \setminus (B_0 \cup \ldots \cup B_{j-1}) = \emptyset). \end{aligned}$$

根据$\bigcup\limits_{i=0}^{k-1} \partial C_i \subseteq \bigcup\limits_{i=0}^{k-1} \partial B_i$, 知道其内部为空.

设$D = \bigcup\limits_{i=0}^{k-1} \partial C_i$ 以及$D_\infty = \bigcup\limits_{n \in \mathbb{Z}} T^n D$.

$$Y = X \setminus D_\infty.$$

定义

$$\psi : Y \to \Sigma_k = \{0, 1, \ldots, k-1\}^{\mathbb{Z}}, z \mapsto (a_n)_{n \in \mathbb{Z}},$$

其中$(a_n)_{n \in \mathbb{Z}}$ 由下式定义：

$$T^n z \in A_{a_n}.$$

设$\Lambda = \psi(Y)$, 我们要证明ψ 为单射, 并且其逆可以扩充为连续映射

$$\phi : \overline{\Lambda} \to X.$$

为此我们证明：对于任何$\varepsilon > 0$, 存在$N \in \mathbb{N}$ 使得如果$x, y \in Y$, 并且$(\psi(x))_n = (\psi(y))_n, \forall |n| \leq N$, 那么$d(x, y) < \varepsilon$, 即$\psi^{-1}$ 为一致连续的.

设$\varepsilon > 0$. 由定理5.7.12, 取$N \in \mathbb{N}$ 使得$\mathrm{diam}\left(\bigvee\limits_{j=-n}^{n} T^{-j}\gamma\right) < \varepsilon$. 如果$x, y \in Y$, 并且$(\psi(x))_n = (\psi(y))_n, \forall |n| \leq N$, 那么$x, y$ 在$\bigvee\limits_{j=-n}^{n} T^{-j}\gamma$ 同一元素中, 尤其$d(x, y) < \varepsilon$.

于是ψ^{-1} 为一致连续的, 我们可以将$\psi^{-1} : \Lambda \to X$ 扩充为连续映射

$$\phi : \overline{\Lambda} \to X.$$

因为$X \setminus D_\infty$ 稠密, 所以ϕ 为满射. 容易验证ϕ 为因子映射. 证毕! □

§5.7.5 可扩同胚的例子

例 5.7.14 *1.* 除非为有限个点, 等度连续系统不是可扩的.

2. 设$A:\mathbb{T}^n\to\mathbb{T}^n$ 为自同构, $[A]$ 为对应的矩阵. A 为可扩的当且仅当$[A]$ 没有模为*1* 的特征值.

3. 任何符号子转移系统(X,σ)为可扩的. 设$X\subseteq\Sigma_k, k\geq 2$. 因为

$$\alpha=\{{}_0[i]_0\cap X: i\in\{0,1,\ldots,k-1\}\}$$

为生成子, 所以(X,σ) 可扩.

注记 *5.7.15* 一个重要的问题是: 对于什么样的拓扑空间一定存在可扩同胚?

容易证明在紧致一维流形上不存在可扩同胚, 于是任何带边紧致曲面上没有可扩同胚. *O'Brien* 与*Reddy*（*1970*）证明了任何具有正亏格可定向曲面上存在可扩同胚*[158]*. 一个重要的结果是*Hiraide (1990) [97]*和*Lewowicz (1989) [146]*独立证明了在一个闭曲面上, 任何可扩同胚共轭于*pseudo-Anosov* 映射, 于是在二维球面、射影平面以及*Klein* 瓶上不存在可扩同胚. 对于高维流形, 情况更为复杂, 目前只有部分成果.

定理 5.7.16 设(X,d) 为紧致度量空间, $T:X\to X$ 为可扩同胚. 那么对于任何$p\in\mathbb{N}$,

$$|\mathrm{Fix}(X,T^p)|<\infty.$$

证明. 设$\delta>0$ 为T^p 的可扩常数. 设$T^px=x, T^py=y$. 那么$x=y$ 或$d(x,y)>\delta$. 于是

$$|\mathrm{Fix}(X,T^p)|<\infty.$$

证毕! □

定理 5.7.17 闭区间上不存在可扩同胚.

证明. 设$I=[0,1]$, $T:I\to I$ 为同胚, 于是

$$T(0)=0,\ T(1)=1;\ \text{或者}\ T(0)=1, T(1)=0.$$

无论何种情况, 我们总是有

$$T^2(0)=0;\quad T^2(1)=1.$$

所以T^2 为保向同胚.

(1) 设T^2 除了$0,1$ 外没有不动点, 那么于是对于任何$x\in I$,

$$T^{2n}x\to 1, n\to\infty;\ \ T^{2n}x\to 0, n\to-\infty.$$

于是任何子区间都收缩地趋向端点, T^2 不可能为可扩的.

(2) 根据定理5.7.16, 如果T^2 为可扩的, 那么它只有有限个不动点. 设x_1, x_2 为其中任何两个相邻的不动点. 那么T^2 将区间$[x_1, x_2]$ 映为$[x_1, x_2]$, 并且在其中除了端点外没有不动点, 于是归结到情况(1), T^2 不为可扩的.

综上, T 不为可扩的. □

定理 5.7.18 *圆周$\mathbb{S}^1$ 上不存在可扩同胚.*

证明. 设$T:\mathbb{S}^1\to\mathbb{S}^1$ 为同胚. 设它为保向同胚, 否则研究T^2 即可.

情况*(1)*: $\mathrm{Per}(T)\neq\emptyset$.

于是存在$p\in\mathbb{N}$ 使得$\mathrm{Fix}(T^p)\neq\emptyset$. 取$w_1\in\mathrm{Fix}(T^p)$. 如果$T$ 为可扩的, 那么根据定理5.7.16, $|\mathrm{Fix}(T^p)|<\infty$. 设$w_2$ 为从w_1 出发逆时针方向上出现的第一个T^p 不动点（当然可能出现$w_2=w_1$ 的情况）. 因为T^p 保向, 所以

$$T^p:\overrightarrow{[w_1,w_2]}\to\overrightarrow{[w_1,w_2]}$$

为区间自映射, 其中$\overrightarrow{[w_1,w_2]}$ 为从w_1出发逆时针方向到w_2 的弧. 根据定理5.7.17, T^p 不是可扩的, 进而T 不是可扩的.

情况*(2)*: $\mathrm{Per}(T)=\emptyset$.

那么根据定理6.4.8, 存在因子映射$\phi:(\mathbb{S}^1,T)\to(\mathbb{S}^1,S)$, 其中$S$ 为无理旋转. 对于任何点$w\in\mathbb{S}^1$, $\phi^{-1}(w)$ 要么为独点集, 要么为闭区间.

$$\begin{tikzcd}
\mathbb{S}^1 \arrow[r, "T"] \arrow[d, "\phi"'] & \mathbb{S}^1 \arrow[d, "\phi"] \\
\mathbb{S}^1 \arrow[r, "S"'] & \mathbb{S}^1
\end{tikzcd}$$

因为$(\mathbb{S}^1,S)$ 为无理旋转, 所以S 不是可扩的. 如果对于任何$w\in\mathbb{S}^1$, $\phi^{-1}(w)$ 要么为独点集, 那么ϕ 为同胚. 于是T,S 为同构, T 也不可扩. 如果存在点w_0 使得$\phi^{-1}(w_0)$ 为非平凡闭区间, 那么易见

$$\{T^{-n}(\phi^{-1}(w_0)):n\in\mathbb{Z}\}$$

为$\mathbb{S}^1$ 互不相交闭区间列. 任取$\delta>0$, 那么存在N 使得对于任何$|n|\geq N$, $T^{-n}(\phi^{-1}(w_0))$ 的长度小于δ. 在区间$\phi^{-1}(w_0)$ 中取点z_1,z_2 使得$z_1\neq z_2$ 且满足

$$d(T^nz_1,T^nz_2)<\delta,\ \forall|n|\leq N.$$

由此得到

$$d(T^nz_1,T^nz_2)<\delta,\ \forall n\in\mathbb{Z}.$$

于是δ 不可能为T 的可扩常数. 因为δ 是任何正实数, 所以T 不可扩. □

习　题

1. 设(X,T) 为拓扑动力系统. 证明：如果(X,T) 是传递的, 并且T 的周期点集在X 中稠密, 那么(X,T) 具有初值敏感性.

§5.8 注记

拓扑动力系统有许多非常优秀的专著, 例如[3, 4, 9, 56, 64, 73, 84, 190, 199, 209] 等. 本章内容主要参考了笔者们的另一本专著[207], 可扩同胚的内容来自[195]. 关于可扩同胚还可参见[6, 153, 208] 等. 替换系统是非常重要的一类符号系统，系统研究可参见[57, 165].

第六章　拓扑动力系统的不变测度

本章主要介绍拓扑动力系统上的不变测度. 我们将看到对于任何拓扑动力系统, 总是可以找到不变测度使得它成为一个保测系统. 反之, 对于任何保测系统, 也可以建立适当的拓扑模型而不影响其测度结构. 本章的目的之一就是希望介绍拓扑动系统与保测系统的相互关系.

在第一节我们回顾测度论中的一些基本知识, 在第二节我们将说明任何拓扑动力系统, 总是可以找到相对于其Borel σ 代数的不变测度, 即总是可以将拓扑动力系统视为某个不变测度下的保测系统. 在第三、四节我们介绍通用点与支撑、唯一遍历等, 并且给出一些数论中的应用.　最后我们简单介绍拓扑模型理论, 说明保测系统也可以视为拓扑动力系统.

§6.1　测度空间的一些基本性质

设(X,d) 是一个度量空间, 我们总是以$\mathcal{B}(X)$ 记其上的Borel σ 代数. 记$(X,\mathcal{B}(X))$ 上全体Borel 概率测度构成的集合为$\mathcal{M}(X)$. 注意对于$\mu,\nu\in\mathcal{M}(X)$, μ 与ν 相等（记为$\mu=\nu$）, 是指对任意的$B\in\mathcal{B}(X)$, $\mu(B)=\nu(B)$. 在本节中, 我们将赋予$\mathcal{M}(X)$ 适当的拓扑使得它成为紧致度量空间.

§6.1.1　测度的正则性

命题 6.1.1 *设X 为度量空间, 则$\mu\in\mathcal{M}(X)$ 为正则的, 即对任意$B\in\mathcal{B}(X)$ 和$\varepsilon>0$, 存在开集U和闭集C 使得*

$$C\subseteq B\subseteq U, \text{且 } \mu(U\setminus C)<\varepsilon.$$

证明. 令

$$\mathcal{A}=\{B\in\mathcal{B}(X):\forall\varepsilon>0,\exists\text{开集}U,\ \text{闭集}C\ s.t.\ C\subseteq B\subseteq U,\ \mu(U\setminus C)<\varepsilon\}.$$

我们先验证$\mathcal{A}$ 是一个σ 代数.

(1) 首先显然有$\emptyset,X\in\mathcal{A}$.

(2) 设$B\in\mathcal{A}$. 对任意的$\varepsilon>0$, 存在开集U 和闭集C 使得$C\subseteq B\subseteq U$ 且$\mu(U\setminus C)<\varepsilon$. 则$X\setminus U\subseteq X\setminus B\subseteq X\setminus C$ 和$\mu((X\setminus C)\setminus(X\setminus U))=\mu(U\setminus C)$. 由于$X\setminus U$ 是闭的和$X\setminus C$ 是开的, $X\setminus B\in\mathcal{A}$.

(3) 设$B_n\in\mathcal{A}$, $n\geq 1$. 对任意的$\varepsilon>0$, 存在存在开集U_n 和闭集C_n 使得$C_n\subseteq B_n\subseteq U_n$ 且$\mu(U_n\setminus C_n)<\frac{\varepsilon}{2^{n+1}}$. 令$U=\bigcup_{n=1}^\infty U_n$ 和$C_0=\bigcup_{n=1}^\infty C_n$. 则$C_0\subseteq\bigcup_{n=1}^\infty B_n\subseteq U$ 和

$$\mu(U\setminus C_0)\leq\sum_{n=1}^\infty\mu(U_n\setminus C_n)<\frac{\varepsilon}{2}.$$

$\{U\setminus\bigcup_{n=1}^m C_n\}_{m=1}^\infty$ 单调递减收敛到$U\setminus C_0$. 故存在$m\in\mathbb{N}$ 使得

$$\mu\Big(U\setminus\bigcup_{n=1}^m C_n\Big)-\mu(U\setminus C_0)<\frac{\varepsilon}{2}.$$

令$C=\bigcup_{n=1}^m C_n$. 则U 是开的和C 是闭的, 且满足$C\subseteq\bigcup_{n=1}^\infty B_n\subseteq U$ 和$\mu(U\setminus C)<\varepsilon$. 因此$\bigcup_{n=1}^\infty B_n\in\mathcal{A}$. 这证明了$\mathcal{A}$ 是一个σ代数.

要证$\mathcal{A}=\mathcal{B}(X)$, 只需要证明$\mathcal{A}$ 包含所有非空闭集. 设A 是一个非空闭集. 对任意的$n\in\mathbb{N}$, 令

$$V_n=B_{\frac{1}{n}}(A)=\{x\in X: d(x,A)<\tfrac{1}{n}\}.$$

易见$\{V_n\}_{n=1}^\infty$是一列单调下降的开集列, 并收敛到A. 则$\lim_{n\to\infty}\mu(V_n\setminus A)=0$. 故$A\in\mathcal{A}$. 这就完成了整个证明. □

根据上定理, 我们马上有如下推论:

推论 6.1.2 设X 为度量空间, 则对于任何$\mu\in\mathcal{M}(X)$和$B\in\mathcal{B}(X)$, 有

$$\begin{aligned}\mu(B)&=\sup\{\mu(C): C\subseteq B,\ C\text{是}X\text{的闭子集}\}\\&=\inf\{\mu(U): B\subseteq U,\ U\text{是}X\text{的开子集}\}.\end{aligned}$$

下面简单的命题提供了一个判定两个测度是否一样的方法. 下面命题根据Riesz 表示定理即得, 我们给出一个直接的证明.

命题 6.1.3 设X 为度量空间, $\mu,\nu\in\mathcal{M}(X)$. 则$\mu=\nu$当且仅当对任意的$f\in C(X)$,

$$\int_X f d\mu=\int_X f d\nu.$$

证明. ($\Rightarrow$)是显然的. 下证($\Leftarrow$). 由推论6.1.2, 我们只需证对X中任意一个非空闭集A有$\mu(A)=\nu(A)$. 设A 是一个非空闭集. 对任意的$n\in\mathbb{N}$, 令

$$V_n=\{x\in X: d(x,A)<\tfrac{1}{n}\}$$

和

$$f_n(x)=\frac{d(x,X\setminus V_n)}{d(x,A)+d(x,X\setminus V_n)}.$$

则f_n 连续, $0\le f_n(x)\le 1$, 对任意$x\in A$ 有$f_n(x)=1$ 和对任意$x\in X\setminus V_n$ 有$f_n(x)=0$. 于是

$$\mu(A)\le\int_X f_n d\mu\le\mu(V_n)\ \text{和}\ \nu(A)\le\int_X f_n d\nu\le\nu(V_n).$$

由于$\{V_n\}_{n=1}^\infty$ 单调下降收敛到A,

$$\begin{aligned}\mu(A)&=\lim_{n\to\infty}\mu(V_n)\\&=\lim_{n\to\infty}\int_X f_n d\mu=\lim_{n\to\infty}\int_X f_n d\nu\\&=\lim_{n\to\infty}\nu(V_n)=\nu(A).\end{aligned}$$

证毕! □

§6.1.2 $\mathcal{M}(X)$的拓扑

X 上全体连续函数集合$C(X)$ 按函数的加法和数乘构成一个线性空间. 对$f \in C(X)$, 定义

$$\|f\|_{\sup} = \sup_{x \in X} |f(x)|.$$

则$(C(X), \|\cdot\|_{\sup})$ 构成一个Banach 空间. 由于X 是一个紧致度量空间, $C(X)$ 是可分的.

$C(X)$ 的对偶空间$C(X)^*$ 由$C(X)$ 的所有连续线性泛函$L: C(X) \to \mathbb{C}$ 组成. $C(X)^*$ 上的弱* 拓扑是使得所有由$C(X)$ 的中元素f 诱导的线性算子, $C(X)^* \to \mathbb{C}$, $L \mapsto L(f)$, 都连续的最小拓扑.

如果$f \geq 0$ 蕴含$L(f) \geq 0$, 则称线性泛函$L: C(X) \to \mathbb{C}$ 是正的. 记

$$C(X)_1^* = \{L \in C(X)^* : L\text{是正的且}L(1) = 1.\}$$

易见, 对于$L \in C(X)_1^*$,

$$\|L\| = \sup\{|L(f)| : \|f\|_{\sup} = 1\} = 1.$$

所以$C(X)_1^*$ 是$C(X)^*$ 单位球面上的一个凸子空间.

定理 6.1.4 (Riesz表示定理) 设X 为紧致度量空间.

$$\Phi: \mathcal{M}(X) \to C(X)_1^*,\ \mu \mapsto L_\mu,$$

是一个既单又满的仿射, 其中$L_\mu(f) = \int_X f d\mu,\ \forall f \in C(X)$.

于是可以把$C(X)^*$ 的拓扑诱导到$\mathcal{M}(X)$ 上, 称此拓扑为$\mathcal{M}(X)$ 上的弱* 拓扑, 即$\mathcal{M}(X)$ 上的弱* 拓扑是使得对任意$f \in C(X)$, 映射

$$\mathcal{M}(X) \to \mathbb{C},\ \mu \mapsto \int_X f d\mu,$$

都连续的最小拓扑. 设$\mu_n, \mu \in \mathcal{M}(X)$, 那么在弱* 拓扑下

$$\lim_{i \to \infty} \mu_i = \mu \Longleftrightarrow \lim_{i \to \infty} \int_X f d\mu_i = \int_X f d\mu,\ \forall f \in C(X).$$

这个拓扑的一个拓扑基可由如下形式的集合给出:

$$V_\mu(f_1, \ldots, f_k, \varepsilon) = \Big\{\nu \in \mathcal{M}(X) : \Big|\int_X f_i d\mu - \int_X f_i d\nu\Big| < \varepsilon, i = 1, \ldots, k\Big\},$$

其中$\mu \in \mathcal{M}(X)$, $\varepsilon > 0$, $f_i \in C(X)$ $(i = 1, \ldots, k)$ 和$k \in \mathbb{N}$.

定理 6.1.5 设X 为紧致度量空间, 那么$\mathcal{M}(X)$ 在弱* 拓扑下是一个紧致可度量化空间.

它的一个相容度量定义为

$$P(\mu, \nu) = \sum_{n=1}^{\infty} \frac{1}{2^n \|f_n\|_{\sup}} \Big|\int_X f_n d\mu - \int_X f_n d\nu\Big|,$$

其中$\{f_n\}_{n=1}^{\infty}$ 是$C(X)$ 中一列稠密的非零函数列.

证明. 首先容易验证$P(\cdot,\cdot)$ 是$\mathcal{M}(X)$ 上的一个度量. 下证明拓扑相容性. 设$\mu\in\mathcal{M}(X)$ 和$\varepsilon>0$. 存在$N\in\mathbb{N}$ 使得$\sum\limits_{n=N+1}^{\infty}\frac{1}{2^n}<\frac{\varepsilon}{4}$. 若$\nu\in V_\mu(f_1,\dots,f_N,\frac{\varepsilon}{4}\min\limits_{1\le n\le N}\|f_n\|_{\sup})$, 则

$$\begin{aligned}P(\mu,\nu)&=\sum_{n=1}^{\infty}\frac{1}{2^n\|f_n\|_{\sup}}\Big|\int_X f_nd\mu-\int_X f_nd\nu\Big|\\&\le\sum_{n=1}^{N}\frac{1}{2^n\|f_n\|_{\sup}}\cdot\frac{\varepsilon}{4}\cdot\min_{1\le n\le N}\|f_n\|_{\sup}+\sum_{n=N+1}^{\infty}\frac{1}{2^n}\cdot 2\le\varepsilon.\end{aligned}$$

设$g_1,\dots,g_k\in C(X)$ 以及$\nu\in V_\mu(g_1,\dots,g_k,\varepsilon)$. 由于$\{f_n\}_{n=1}^\infty$ 在$C(X)$ 中稠密, 对任意的$i=1,\dots,k$, 存在$n_i\in\mathbb{N}$ 使得$\|g_i-f_{n_i}\|_{\sup}<\frac{\varepsilon}{4}$. 若

$$P(\nu,\eta)<\frac{1}{\max_{1\le i\le k}2^{n_i}\|f_{n_i}\|_{\sup}}\cdot\frac{\varepsilon}{2},$$

则

$$\begin{aligned}\Big|\int_X g_id\nu-\int_X g_id\eta\Big|&\le\Big|\int_X f_{n_i}d\nu-\int_X f_{n_i}d\eta\Big|+\frac{\varepsilon}{2}\\&\le 2^{n_i}\|f_{n_i}\|_{\sup}P(\nu,\eta)+\frac{\varepsilon}{2}<\varepsilon.\end{aligned}$$

这说明$\mathcal{M}(X)$ 的弱*拓扑与度量$P(\cdot,\cdot)$ 诱导的拓扑一致.

为证明紧致性, 只需要证明每个序列具有收敛子列. 设$\{\mu_k\}_{k=1}^\infty$ 是$\mathcal{M}(X)$ 中一个序列. 由于$\Big\{\int_X f_1d\mu_k\Big\}_{k=1}^\infty$是一个有界序列, 存在一个收敛子列$\Big\{\int_X f_1d\mu_{k_{1,i}}\Big\}_{i=1}^\infty$. 对任意的$n\ge 2$, 我们递归选取$k_{n-1,i}$ 的子序列$k_{n,i}$ 使得$\Big\{\int_X f_nd\mu_{k_{n,i}}\Big\}_{i=1}^\infty$ 收敛. 根据对角线原则, 令$k_i=k_{i,i}$, 则对任意的$n\in\mathbb{N}$, $\Big\{\int_X f_nd\mu_{k_i}\Big\}_{i=1}^\infty$ 收敛, 从而令

$$L(f_n)=\lim_{i\to\infty}\int_X f_nd\mu_{k_i}.$$

对任意的$f\in C(X)$, 存在序列$\{n_s\}$使得

$$\lim_{s\to\infty}\|f_{n_s}-f\|_{\sup}=0.$$

则对任意的$s,i,j\in\mathbb{N}$,

$$\begin{aligned}\Big|\int_X fd\mu_{k_i}-\int_X fd\mu_{k_j}\Big|&\le\Big|\int_X f_{n_s}d\mu_{k_i}-\int_X f_{n_s}d\mu_{k_j}\Big|+\int_X|f-f_{n_s}|d\mu_{k_i}+\int_X|f-f_{n_s}|d\mu_{k_j}\\&\le\Big|\int_X f_{n_s}d\mu_{k_i}-\int_X f_{n_s}d\mu_{k_j}\Big|+2\|f-f_{n_s}\|_{\sup}.\end{aligned}$$

故$\Big\{\int_X fd\mu_{k_i}\Big\}_{i=1}^\infty$ 是一个Cauchy 列, 从而收敛. 令

$$L(f)=\lim_{i\to\infty}\int_X fd\mu_{k_i}.$$

于是我们得到算子

$$L : C(X) \to \mathbb{C}, \quad f \mapsto L(f).$$

易见它为线性的. 若$f \geq 0$, 则对任意的$i \in \mathbb{N}$, $\int_X f d\mu_{k_i} \geq 0$, 从而$L(f) \geq 0$. 若$\|f\|_{\sup} \leq 1$, 则对任意的$i \in \mathbb{N}$, $\left|\int_X f d\mu_{k_i}\right| \leq \|f\|_{\sup}$, 从而$|L(f)| \leq \|f\|_{\sup}$. 若$f \equiv 1$, 则对任意的$i \in \mathbb{N}$, $\int_X f d\mu_{k_i} = 1$, 从而$L(f) = 1$. 这说明L 是$C(X)$ 上的一个连续正线性泛函且$L(1) = 1$. 由Riesz 表示定理, 存在$\mu \in \mathcal{M}(X)$ 使得对任意的$f \in C(X)$, $\int_X f d\mu = L(f)$. 故μ_{k_i} 按照度量$P(\cdot,\cdot)$ 收敛到μ. 证毕! □

容易证明：

命题 6.1.6 *$\varphi : X \to \mathcal{M}(X)$, $x \mapsto \delta_x$, 是一个拓扑嵌入.*

定理 6.1.7 (Portmanteau定理) *设X是一个度量空间, 设$\mu_i, \mu \in \mathcal{M}(X)$, $i = 1, 2, \ldots$. 则下列论断等价*

1. *弱*拓扑下$\mu_i \to \mu$, $i \to \infty$.*
2. *对X中任意一个闭子集E, $\limsup\limits_{i\to\infty} \mu_i(E) \leq \mu(E)$.*
3. *对X中任意一个开子集U, $\liminf\limits_{i\to\infty} \mu_i(U) \geq \mu(U)$.*
4. *对任意满足$\mu(\partial A) = 0$的$A \in \mathcal{B}(X)$, $\lim\limits_{i\to\infty} \mu_i(A) = \mu(A)$, 其中$\partial A$表示集合$A$的边界.*

证明. (1)⇒(2) 设E是X一个非空闭子集. 对任意的$n \in \mathbb{N}$, 令

$$U_n = B_{\frac{1}{n}}(E) = \{x \in X : d(x, E) < \tfrac{1}{n}\}$$

和

$$f_n(x) = \frac{d(x, X \setminus U_n)}{d(x, E) + d(x, X \setminus U_n)}.$$

则f_n 连续, $0 \leq f(x) \leq 1$, 对任意$x \in E$ 有$f(x) = 1$ 和对任意$x \in X \setminus U_n$ 有$f(x) = 0$. 于是

$$\limsup_{i\to\infty} \mu_i(E) \leq \limsup_{i\to\infty} \int_X f_n d\mu_i = \int_X f_n d\mu \leq \mu(U_n).$$

由于$\{U_n\}_{n=1}^{\infty}$ 单调下降收敛到E, $\lim_{n\to\infty} \mu(U_n) = \mu(E)$. 所以

$$\limsup_{i\to\infty} \mu_i(E) \leq \mu(E).$$

(2)⇔(3) 设$U \subseteq X$. 则U 是X 中一个开集当且仅当$X \setminus U$ 是X 中的一个闭集. 注意到$\mu(U) = 1 - \mu(X \setminus U)$ 和$\mu_i(U) = 1 - \mu_i(X \setminus U)$. 所以

$$\limsup_{i\to\infty} \mu_i(X \setminus U) \leq \mu(X \setminus U) \text{ 当且仅当 } \liminf_{i\to\infty} \mu_i(U) \geq \mu(U).$$

(2)⇒(4) 设$A \in \mathcal{B}(X)$ 且满足$\mu(\partial A) = 0$. 则$\mu(\mathrm{int}(A)) = \mu(A) = \mu(\mathrm{cl}(A))$, 其中$\mathrm{int}(A)$和$\mathrm{cl}(A)$ 分别为A 的内部和闭包. 由(2) 和(3),

$$\liminf_{i\to\infty} \mu_i(A) \geq \liminf_{i\to\infty} \mu_i(\mathrm{int}(A)) \geq \mu(\mathrm{int}(A)) = \mu(A)$$

和

$$\limsup_{i\to\infty} \mu_i(A) \leq \limsup_{i\to\infty} \mu_i(\mathrm{cl}(A)) \leq \mu(\mathrm{cl}(A)) = \mu(A).$$

故$\lim\limits_{i\to\infty} \mu_i(A) = \mu(A)$.

(4)⇒(1) 设$f \in C(X)$. 不妨设f 为实值函数. 则存在$a, b \in \mathbb{R}$使得$f(X) \subseteq (a, b)$. 对任意的$z \in (a, b)$, 令$F_z = f^{-1}(z)$. 则$\{F_z : z \in (a, b)\}$ 是X 中一族两两不交的闭集组成的覆盖, 从而只有至多可数个这样的集相对于μ 的测度是正的. 对任意的$\varepsilon > 0$, 取$a = t_0 < t_1 < \cdots < t_m = b$ 使得$\mu(F_{t_i}) = 0$ 和$t_{j+1} - t_j < \varepsilon$, $j = 1, 2, \ldots, m-1$. 令

$$A_j = f^{-1}([t_{j-1}, t_j)), j = 1, 2, \ldots, m.$$

易见$X = \bigcup_{j=1}^m A_j$ 和$\partial A_j \subseteq F_{t_{j-1}} \cup F_{t_j}$, $j = 1, 2, \ldots, m-1$ 和$\partial A_m \subseteq F_{t_{m-1}}$. 则由$A_j$ 取法知$\mu(\partial A_j) = 0$, $j = 1, 2, \ldots, m$. 由(4), $\lim_{i\to\infty} \mu_i(A_j) = \mu(A_j)$, $j = 1, 2, \ldots, m$. 令

$$g = \sum_{j=1}^m t_{j-1} \mathbf{1}_{A_j}.$$

则$\|f - g\|_{\sup} < \varepsilon$. 从而

$$\begin{aligned}\left|\int_X f d\mu_i - \int_X f d\mu\right| &\leq \left|\int_X f d\mu - \int_X g d\mu\right| + \left|\int_X f d\mu_i - \int_X g d\mu_i\right| + \left|\int_X g d\mu_i - \int_X g d\mu\right| \\ &\leq 2\|f - g\|_{\sup} + \sum_{j=1}^m |t_{j-1}| \cdot |\mu_i(A_j) - \mu(A_j)|.\end{aligned}$$

故

$$\limsup_{i\to\infty} \left|\int_X f d\mu_i - \int_X f d\mu\right| \leq 2\varepsilon.$$

由ε的任意性得

$$\lim_{i\to\infty} \int_X f d\mu_i = \int_X f d\mu.$$

证毕! □

设X和Y是两个紧致度量空间. 若$\phi : X \to Y$是一个可测映射, 则ϕ诱导一个映射

$$\phi_* : \mathcal{M}(X) \to \mathcal{M}(Y), \quad \mu \mapsto \phi_*\mu = \mu \circ \phi^{-1},$$

即对任意$B \in \mathcal{B}(Y)$, $(\phi_*\mu)(B) = \mu(\phi^{-1}B)$.

引理 6.1.8 若$\phi : X \to Y$ 是一个连续映射, 则$\phi_* : \mathcal{M}(X) \to \mathcal{M}(Y)$ 是一个连续的仿射. 并且ϕ_* 是满射当且仅当ϕ 是满射.

证明. 易见ϕ_* 是一个仿射, 即对任意的$\mu, \nu \in \mathcal{M}(X)$ 和$\alpha \in [0, 1]$,

$$\phi_*(\alpha\mu + (1-\alpha)\nu) = \alpha\phi_*(\mu) + (1-\alpha)\phi_*(\nu).$$

故只需证ϕ_* 连续. 注意到对任意的$f \in C(Y)$,

$$\int_Y f d\phi_*\mu = \int_X f \circ \phi d\mu.$$

设$\mu_n, \mu \in \mathcal{M}(X)$, 且$\mu_n \to \mu, n \to \infty$, 则对任意的$f \in C(Y)$,

$$\int_Y f d\phi_*\mu_n = \int_X f \circ \phi d\mu_n \to \int_X f \circ \phi d\mu = \int_Y f d\phi_*\mu.$$

故$\phi_*\mu_n \to \phi_*\mu, n \to \infty$. 这就证明了$\phi_*$ 的连续性.

命题后一部分留作习题. □

习 题

1. 证明命题6.1.6.
2. 设X 是一个度量空间和$A \subseteq X$. 对任意的$\varepsilon > 0$, 令
$$B_\varepsilon(A) = \{x \in X : \exists y \in A \text{ s.t. } d(x, y) < \varepsilon\}.$$
对于$\mu, \nu \in \mathcal{M}(X)$, 定义
$$\rho(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(A) < \nu(B_\varepsilon(A)), \nu(A) < \mu(B_\varepsilon(A)), \quad \forall A \in \mathcal{B}(X)\}.$$
证明ρ 是与$\mathcal{M}(X)$ 弱* 拓扑相容的一个度量. 我们称之为**Prohorov 度量**(Prohorov metric).
3. 点测度的凸包
$$co(X) = \left\{\sum_{i=1}^n \alpha_i \delta_{x_i} : n \in \mathbb{N},\ x_i \in X,\ \alpha_i \geq 0,\ i = 1, \ldots, n,\ \sum_{i=1}^n \alpha_i = 1\right\}$$
在$\mathcal{M}(X)$中稠密.
4. 证明引理6.1.8后面部分. 若$\phi : X \to Y$ 是一个连续映射, 则$\phi_* : \mathcal{M}(X) \to \mathcal{M}(Y)$ 是满射当且仅当ϕ 是满射.
5. 设$X = [0, 1]$. 设$\mu_i = \delta_{\frac{1}{i}}$和$\mu = \delta_0$. 则$\mu_i \to \mu$, $i \to \infty$. 取$E = \{0\}$ 和$U = (0, 1]$. 考察$\lim_{i\to\infty} \mu_i(E)$ 和$\lim_{i\to\infty} \mu_i(U)$.
6. 设$f \in C(X)$. 如果存在$L > 0$使得对任意的$x, y \in X$,
$$d(f(x), f(y)) \leq L \cdot d(x, y),$$
则称f是**Lipschitz连续**(Lipschitz continuous) 和L 为f 的一个**Lipschitz常数**(Lipschitz constant). 记$C(X)$ 中所有Lipschitz连续函数的组成的集合为$BL(X)$. 证明$BL(X)$在$C(X)$中稠密. 从而定理6.1.5中可取$C(X)$中满足下列要求的子列$\{f_n\}_{n=1}^\infty$:

(a) $0 < \|f_n\|_{\sup} \le 1$, $n \in \mathbb{N}$;

(b) f_n具有Lipschitz常数1, $n \in \mathbb{N}$;

(c) $\{\sum_{i=1}^n \alpha_i f_i : \alpha_i \in \mathbb{C}, n \in \mathbb{N}\}$ 在$C(X)$中稠密.

§6.2 拓扑动力系统的不变测度空间

对于一个紧致度量空间X, 它有一个自然的σ 代数与之对应, 即由全体开集生成的σ 代数$\mathcal{B}(X)$. 对于一个动力系统(X,T), 一个自然而且重要的问题是：在$\mathcal{B}(X)$ 上是否存在一个不变测度? 如果是, 那么我们就可以将遍历理论的方法运用到拓扑动力系统中. 幸运的是, 答案是肯定的. 我们在本节讨论这个问题.

§6.2.1 不变测度

设(X,T) 是一个拓扑动力系统和$\mu \in \mathcal{M}(X)$. 如果对任意的$B \in \mathcal{B}(X)$, $\mu(B) = \mu(T^{-1}B)$, 则称μ 是一个**不变测度**(invariant measure). 有时为了强调T, 也称之为一个T 不变测度. 记$\mathcal{M}(X,T) \subseteq \mathcal{M}(X)$ 为全体T 不变Borel 概率测度组成的集合, 而$\mathcal{M}^e(X,T) \subseteq \mathcal{M}(X,T)$ 为全体遍历测度组成的集合.

令

$$T_* : \mathcal{M}(X) \to \mathcal{M}(X), \quad \mu \mapsto \mu \circ T^{-1}.$$

则由引理6.1.8, T_* 为连续的，从而$(\mathcal{M}(X), T_*)$ 为拓扑动力系统, 称之为**测度空间系统**. 令

$$\delta : (X,T) \to (\mathcal{M}(X), T_*), \quad x \mapsto \delta_x.$$

那么动力系统(X,T) 嵌入$(\mathcal{M}(X), T_*)$ 中(命题6.1.6).

注意$\mathcal{M}(X,T)$ 为拓扑动力系统$(\mathcal{M}(X), T_*)$ 的全体不动点构成的集合. 即

$$\mathcal{M}(X,T) = \mathrm{Fix}(\mathcal{M}(X), T_*) = \{\mu \in \mathcal{M}(X) : T_*\mu = \mu\}.$$

一个常用的事实是（引理1.1.6）：

$$\int_X f d(T_*\mu) = \int_X f \circ T d\mu, \quad \forall f \in C(X).$$

由此我们有：

引理 6.2.1 设(X,T)是一个拓扑动力系统和$\mu \in \mathcal{M}(X)$. 那么$\mu \in \mathcal{M}(X,T)$ 当且仅当

$$\int_X f \circ T d\mu = \int_X f d\mu, \quad \forall f \in C(X).$$

由Markov-Kakutani定理(例如参见[173, 定理5.11]), 每个紧致凸集上的连续仿射必有不动点. 所以$\mathcal{M}(X,T) \neq \emptyset$. [1] 下面命题6.2.2 给出这个结论的一个直接证明.

[1]很多书籍中将这个结论叫做Krylov-Bogolyubov 定理.

命题 6.2.2 设(X,T) 是一个拓扑动力系统和$\{\sigma_n\}_{n=1}^{\infty}$是$\mathcal{M}(X)$ 中的一个序列. 对任意的$n \in \mathbb{N}$, 令$\mu_n = \dfrac{1}{n}\sum_{i=0}^{n-1} T_*^i \sigma_n$. 则序列$\{\mu_n\}_{n=1}^{\infty}$ 在弱* 拓扑下的极限测度是T 不变的. 尤其

$$\mathcal{M}(X,T) \neq \emptyset.$$

证明. 设$\{n_j\}_{j\in N} \subseteq \mathbb{N}$ 使得$\mu_{n_j} \to \mu, j \to \infty$. 则对任意的$f \in C(X)$,

$$\begin{aligned}
\left|\int_X f d\mu - \int_X f\circ T d\mu\right| &= \lim_{j\to\infty}\left|\int_X f d\mu_{n_j} - \int_X f\circ T d\mu_{n_j}\right| \\
&= \lim_{j\to\infty}\left|\int_X \frac{1}{n_j}\sum_{i=0}^{n_j-1}\big(f\circ T^i - f\circ T^{i+1}\big)d\sigma_{n_j}\right| \\
&= \lim_{j\to\infty}\left|\int_X \frac{1}{n_j}\big(f - f\circ T^{n_j}\big)d\sigma_{n_j}\right| \\
&\leq \lim_{j\to\infty}\frac{2\|f\|_{\sup}}{n_j} = 0.
\end{aligned}$$

所以$T_*\mu = \mu$, 即$\mu \in \mathcal{M}(X,T)$. □

一个常用的推论如下, 它是构造不变测度常用的方式.

推论 6.2.3 设(X,T)是一个拓扑动力系统和$x \in X$. 对任意的$n \in \mathbb{N}$, 令

$$\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} T_*^i \delta_x = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i x}.$$

则序列$\{\mu_n\}_{n=1}^{\infty}$ 在弱* 拓扑下的极限测度是T 不变的.

定理 6.2.4 设(X,T)为动力系统, 则$\mathcal{M}(X,T)$为$\mathcal{M}(X)$的紧致凸子集.

证明. 仅需证明$\mathcal{M}(X,T)$ 为闭的. 设$\{\mu_n\}_{n=1}^{\infty}$ 为$\mathcal{M}(X,T)$ 中的序列且$\mu_n \to \mu, n \to \infty$. 那么

$$\int_X f\circ T d\mu = \lim_{n\to\infty}\int_X f\circ T d\mu_n = \lim_{n\to\infty}\int_X f d\mu_n = \int_X f d\mu, \quad \forall f \in C(X).$$

于是$\mu \in \mathcal{M}(X,T)$. □

§6.2.2 $\mathcal{M}$的凸集结构

定义 6.2.5 (凸集端点) 设M是某向量空间的一个凸子集和$x \in M$. 如果x 满足下面的性质: 对任意$y, z \in M$ 若存在$\alpha \in (0,1)$ 使得$x = \alpha y + (1-\alpha)z$ 则必有$y = z = x$, 那么称x 是M 的一个**端点** *(extreme point)*. 记M 的所有端点的集合为$ex(M)$.

我们有下面关于凸集性质的Krein - Milman定理（证明可参见[173, 定理3.23]）.

定理 6.2.6 (Krein - Milman定理) 设X 是一个局部凸*Hausdorff*的拓扑向量空间, M 是X 的一个紧凸子集. 则$ex(M)$ 是M 的一个非空G_δ 子集, 并且$ex(M)$ 的凸包在M 中稠密.

命题 6.2.7 设(X,T) 为拓扑动力系统. $\mu\in\mathcal{M}(X,T)$是一个端点当且仅当$\mu\in\mathcal{M}^e(X,T)$.

证明. $(\Leftarrow)$ 设存在$\alpha\in[0,1]$ 和$\mu_1,\mu_2\in\mathcal{M}(X,T)$ 使得$\mu=\alpha\mu_1+(1-\alpha)\mu_2$. 则$\mu_1\ll\mu$. 由Radon-Nikodym 定理(见定理0.3.3), 存在$f\in L^1(X,\mu)$ 使得$f\ge 0$ 和对任意的$A\in\mathcal{B}(X)$,

$$\mu_1(A)=\int_A f d\mu.$$

设$B=\{x\in X:0\le f(x)<1\}$. 则$B\in\mathcal{B}(X)$ 和

$$\begin{aligned}\int_{B\cap T^{-1}B} fd\mu+\int_{B\setminus T^{-1}B} fd\mu&=\int_B fd\mu=\mu_1(B)=\mu_1(T^{-1}B)\\&=\int_{T^{-1}B} fd\mu=\int_{B\cap T^{-1}B} fd\mu+\int_{(T^{-1}B)\setminus B} fd\mu.\end{aligned}$$

从而

$$\int_{B\setminus T^{-1}B} fd\mu=\int_{(T^{-1}B)\setminus B} fd\mu.$$

由B 的定义, 当$x\in B\setminus T^{-1}B$ 时, $0\le f(x)<1$. 又有$f\ge 0$, 当$x\in T^{-1}B\setminus B$ 时, $f(x)\ge 1$. 注意到

$$\begin{aligned}\mu((T^{-1}B)\setminus B)&=\mu(T^{-1}B)-\mu(B\cap T^{-1}B)\\&=\mu(B)-\mu(B\cap T^{-1}B)=\mu(B\setminus T^{-1}B).\end{aligned}$$

于是根据

$$\mu(B\setminus T^{-1}B)>\int_{B\setminus T^{-1}B} fd\mu=\int_{(T^{-1}B)\setminus B} fd\mu\ge\mu(T^{-1}B\setminus B),$$

我们得到$\mu((T^{-1}B)\setminus B)=\mu(B\setminus T^{-1}B)=0$, 即$\mu(B\Delta T^{-1}B)=0$. 由于$\mu$ 是遍历的, $\mu(B)=0$或$\mu(B)=1$. 如果$\mu(B)=1$, 则

$$\mu_1(X)=\int_X fd\mu<\mu(B)=1,$$

矛盾. 故$\mu(B)=0$.

类似地可以证明$\{x\in X:f(x)>1\}$ 的测度也是0. 所以$f\equiv 1$, 从而$\mu_1=\mu$. 这就证明了μ 是一个端点.

$(\Rightarrow)$ 如果μ 不是遍历的, 则存在则存在$B\in\mathcal{B}(X)$ 使得$T^{-1}B=B$ 和$\mu(B)\in(0,1)$. 对任意的$A\in\mathcal{B}(X)$, 令

$$\nu_1(A)=\frac{\mu(A\cap B)}{\mu(B)}\ \text{ 和 }\ \nu_2(A)=\frac{\mu(A\cap X\setminus B)}{\mu(X\setminus B)}.$$

则$\nu_1,\nu_2\in\mathcal{M}(X,T)$ 和$\mu=\mu(B)\nu_1+(1-\mu(B))\nu_2$. 这与$\mu$ 是端点矛盾. 所以μ 是遍历的. □

根据上面证明, 我们得到:

命题 6.2.8 设(X,T) 为拓扑动力系统. $\mu,\nu\in\mathcal{M}^e(X,T)$.

1. 若$\mu \neq \nu$, 则$\mu \perp \nu$.

2. 若$\mu \ll \nu$, 则$\mu = \nu$.

根据Krein - Milman 定理, 我们有

定理 6.2.9 设(X,T) 是一个拓扑动力系统. 则$\mathcal{M}^e(X,T)$ 是$\mathcal{M}(X,T)$ 的一个非空G_δ 子集, 并且$\mathcal{M}^e(X,T)$ 的凸包在$\mathcal{M}(X,T)$ 中稠密.

§6.2.3 Choquet定理与遍历分解定理

前面我们已经给出了遍历分解定理, 根据著名的Choquet 理论我们可以给出遍历分解定理的另一个证明.

定理 6.2.10 (Choquet定理) 设X 是一个局部凸的*Hausdorff* 的拓扑向量空间和M是X的一个紧凸子集. 则对任意的$m \in M$, 存在$ex(M)$ 上唯一一个*Borel* 概率测度τ 使得

$$m = \int_{ex(M)} x d\tau(x).$$

这里的积分定义为: 对X 上的任意连续线性泛函f,

$$f(m) = \int_{ex(M)} f(x) d\tau(x).$$

设(X,T) 是一个拓扑动力系统. 将Choquet 定理应用到$C(X)^*$ 和$\mathcal{M}(X,T)$, 我们得到下面的遍历分解定理.

定理 6.2.11 (遍历分解定理) 设(X,T) 是一个拓扑动力系统和$\mu \in \mathcal{M}(X,T)$. 则存在$\mathcal{M}^e(X,T)$ 上唯一一个*Borel* 概率测度τ 使得

$$\mu = \int_{\mathcal{M}^e(X,T)} \nu d\tau(\nu).$$

特别地, 对任意的$f \in C(X)$,

$$\int_X f d\mu = \int_{\mathcal{M}^e(X,T)} \left(\int_X f(x) d\nu(x) \right) d\tau(\nu).$$

注意这种证明方法可以推广到一般群作用的动力系统, 而前面方法（第§4.6节）却不太容易做到这点.

§6.2.4 因子映射诱导的映射

定理 6.2.12 设$\pi:(X,T)\to(Y,S)$ 是一个因子映射. 则

$$\pi_*(\mathcal{M}(X,T))=\mathcal{M}(Y,S),\ \text{且}\ \pi_*(\mathcal{M}^e(X,T))=\mathcal{M}^e(Y,S).$$

证明. (1) 设$\mu\in\mathcal{M}(X,T)$ 和令$\nu=\pi_*\mu$. 对任意的$A\in\mathcal{B}(Y)$,

$$\nu(S^{-1}(A))=\mu(\pi^{-1}(S^{-1}(A)))=\mu(T^{-1}(\pi^{-1}(A)))=\mu(\pi^{-1}(A))=\nu(A).$$

故$\nu\in\mathcal{M}(Y,S)$. 设$A\in\mathcal{B}(Y)$满足$S^{-1}A=A$. 则

$$T^{-1}(\pi^{-1}(A))=\pi^{-1}(S^{-1}(A))=\pi^{-1}(A).$$

如果μ 是遍历的, 则$\mu(\pi^{-1}(A))=0$ 或1, 从而$\nu(A)=0$ 或1. 故此时ν 也是遍历的. 由上, 我们得到了

$$\pi_*(\mathcal{M}(X,T))\subseteq\mathcal{M}(Y,S),\ \text{且}\ \pi_*(\mathcal{M}^e(X,T))\subseteq\mathcal{M}^e(Y,S).$$

(2) 设$\nu\in\mathcal{M}(Y,S)$. 易见$\{f\circ\pi:f\in C(Y)\}$ 为$C(X)$ 的一个闭子空间，以及

$$L:\{f\circ\pi:f\in C(Y)\}\to\mathbb{C}f\circ\pi\mapsto\int_Y fd\nu$$

是$\{f\circ\pi:f\in C(Y)\}$ 上的一个范数为1 正线性泛函且$L_\nu(\mathbf{1})=1$. 根据Hahn-Banach 定理, L_ν 可延拓为$C(X)$ 上一个范数为1 正线性泛函L.

由Riesz 表示定理, 存在$\hat{\mu}\in\mathcal{M}(X)$使得对任意的$g\in C(X)$, $L(f)=\int gd\hat{\mu}$. 特别地, 对任意的$f\in C(Y)$,

$$\int_X f\circ\pi d\hat{\mu}=L(f\circ\pi)=L_\nu(f\circ\pi)=\int_Y fd\nu.$$

所以$\pi_*\hat{\mu}=\nu$.

设$\mu_k=\dfrac{1}{k}\sum_{i=0}^{k-1}T_*^i\hat{\mu}$. 由紧致性，存在$\{\mu_k\}$ 的一个子列$\{\mu_{k_j}\}$ 和$\mu\in\mathcal{M}(X)$ 使得

$$\lim_{j\to\infty}\mu_{k_j}=\mu.$$

由命题6.2.2, $\mu\in\mathcal{M}(X,T)$. 因为$\pi_*(\mu_k)=\nu,\forall k\in\mathbb{N}$, 所以$\pi_*\mu=\nu$, 从而$\pi_*^{-1}(\nu)\cap\mathcal{M}(X,T)$ 是$\mathcal{M}(X,T)$ 的一个非空闭凸子集.

设$\nu\in\mathcal{M}^e(Y,S)$. 设μ 是$\pi_*^{-1}(\nu)\cap\mathcal{M}(X,T)$ 的一个端点. 下证μ 是$\mathcal{M}(X,T)$ 的一个端点. 如果μ 不是$\mathcal{M}(X,T)$的一个端点, 则存在$\alpha\in(0,1)$ 和$\mu_1,\mu_2\in\mathcal{M}(X,T)$ 使得

$$\mu=\alpha\mu_1+(1-\alpha)\mu_2$$

且μ_1 或μ_2 不在$\pi_*^{-1}(\nu)$中. 注意$\nu=\pi_*\mu=\alpha\pi_*\mu_1+(1-\alpha)\pi_*\mu_2$, $\pi_*\mu_1,\pi_*\mu_2\in\mathcal{M}(Y,S)$. 由于$\nu$ 是遍历的, 则ν 是$\mathcal{M}(Y,S)$ 的一个端点, 从而$\pi_*\mu_1=\pi_*\mu_2=\nu$. 这与$\mu_1$ 或μ_2 不在$\pi_*^{-1}(\nu)$ 中矛盾. 所以μ 是$\mathcal{M}(X,T)$ 的一个端点, 从而根据命题6.2.7, $\mu\in\mathcal{M}^e(X,T)$. □

习 题

1. 设(X,T)和(Y,S)是两个动力系统. 若$\mu \in \mathcal{M}(X,T)$和$\nu \in \mathcal{M}(Y,S)$, 证明$\mu \times \nu \in \mathcal{M}(X \times Y, T \times S)$.

2. μ为$\mathcal{M}(X)$的一个端点当且仅当μ是一个Dirac测度.

3. 设$\{n_k\}$和$\{m_k\}$是两个递增的自然数序列且满足$\lim_{k\to\infty}(n_k - m_k) = \infty$. 设$(X,T)$是一个动力系统和$x \in X$. 对任意的$k \in \mathbb{N}$, 令
$$\mu_k = \frac{1}{n_k - m_k} \sum_{i=m_k}^{n_k-1} T_*^i \delta_x = \frac{1}{n_k - m_k} \sum_{i=m_k}^{n_k-1} \delta_{T^i x}.$$
则序列$\{\mu_k\}_{k=1}^{\infty}$ 在弱*-拓扑下的极限测度是T-不变的.

4. 设(X,T)是一个可逆动力系统. 则$\mathcal{M}(X,T) = \mathcal{M}(X,T^{-1})$.

5. 设(X,T) 是一个拓扑动力系统和$\mu \in \mathcal{M}^e(X,T)$. 证明$\mu$ 要么是一个周期轨上的等分布测度要么是无原子的.

6. 设(X,T)是一个可逆拓扑动力系统. 则$\mathcal{M}^e(X,T) = \mathcal{M}^e(X,T^{-1})$

7. 设(X,T)是一个拓扑动力系统和$\mu \in \mathcal{M}(X,T)$. 则μ是遍历的当且仅当对任意的$\nu \in \mathcal{M}(X)$, $\nu \ll \mu$蕴含
$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} T_*^i \nu = \mu.$$

§6.3　不变测度的通用点与支撑

通用点是由Furstenberg引入的[61], 现在它已经是动力系统中的重要概念.

§6.3.1　通用点

定义 6.3.1 (Furstenberg) 设(X,T) 是一个拓扑动力系统, $x \in X$ 和$\mu \in \mathcal{M}(X,T)$. 如果
$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x} = \mu,$$
即
$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f d\mu,\ \forall f \in C(X),$$
则称x 为μ 的一个**通用点**(*generic point*). 记μ 的全体通用点构成的集合为$G_\mu(T)$ 或G_μ.

一般而言, G_μ 可能是空集, 但是对于遍历测度它为非空的.

引理 6.3.2 设(X,T) 是一个拓扑动力系统, $\mu \in \mathcal{M}(X,T)$. 则G_μ 是X 的一个*Borel* 子集. μ 是遍历的当且仅当$\mu(G_\mu) = 1$.

证明. 设$\{f_k\}_{k=1}^{\infty}$ 是$C(X)$ 的一个由非零函数构成的稠密子列. 对任意的$k \in \mathbb{N}$, 令

$$G_\mu(f_k) = \left\{ x \in X : \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f_k(T^i x) = \int_X f_k d\mu \right\}.$$

则$G_\mu(f_k)$ 是一个Borel 子集. 下说明

$$G_\mu = \bigcap_{k=1}^{\infty} G_\mu(f_k).$$

从而它为Borel 子集. 由于$\{f_k\}_{k=1}^{\infty}$ 在$C(X)$ 中稠密, 对于$f \in C(X)$, $\varepsilon > 0$, 取$i \in \mathbb{N}$ 使得

$$\|f - f_i\|_{\sup} < \varepsilon.$$

于是

$$\begin{aligned} &\left| \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) - \frac{1}{n} \sum_{j=0}^{n-1} f_i(T^j x) - \int_X f d\mu + \int_X f_i d\mu \right| \\ \le &\left| \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) - \frac{1}{n} \sum_{j=0}^{n-1} f_i(T^j x) \right| + \left| \int_X f d\mu - \int_X f_i d\mu \right| < 2\varepsilon \end{aligned}$$

于是我们有

$$\int_X f d\mu - 2\varepsilon \le \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) \le \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) \le \int_X f d\mu + 2\varepsilon.$$

由此我们得到, 对于任何$x \in \bigcap\limits_{k=1}^{\infty} G_\mu(f_k)$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f d\mu, \ \forall f \in C(X).$$

所以就有$G_\mu = \bigcap_{k=1}^{\infty} G_\mu(f_k)$.

如果μ 是遍历的, 则由逐点遍历定理2.3.1, $\mu(G_\mu(f_k)) = 1$ 对任意的$k \ge 1$ 成立, 从而$\mu(G_\mu) = 1$.

如果$\mu(G_\mu) = 1$, 则对任意的$f \in C(X)$和$g \in L^1(X, \mu)$,

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i \cdot g \to \int_X f d\mu \cdot g, \ n \to \infty,$$

几乎处处成立. 由Lebesgue控制收敛定理,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_X f \circ T^i \cdot g d\mu = \int_X f d\mu \int_X g d\mu.$$

所以μ 是遍历的. □

§6.3.2　支撑

定义 6.3.3 设X 为紧致度量空间，$\mu \in \mathcal{M}(X)$. 定义测度μ 的**支撑**(*support*)为

$$\mathrm{Supp}(\mu) = \{x \in X : \text{对 } x \text{ 的任意一个开邻域 } U,\ \mu(U) > 0.\}$$

定理 6.3.4 设(X,T) 为拓扑动力系统, $\mu \in \mathcal{M}(X,T)$.

1. $\mathrm{Supp}(\mu)$ 为非空闭的不变集, 且$\mu(\mathrm{Supp}(\mu)) = 1$.
2. $\Omega(\mathrm{Supp}(\mu), T) = \mathrm{Supp}(\mu)$.
3. 如果μ 为遍历的, 那么$(\mathrm{Supp}(\mu), T)$ 为传递的，并且$\mu(\mathrm{Trans}(\mathrm{Supp}(\mu), T)) = 1$.

证明. 1. 如果$\mathrm{Supp}(\mu)$ 为空集, 那么由X 的紧性知道存在开覆盖, 其每个元素都是μ 零测集. 于是$\mu(X) = 0$, 这不可能! 于是$\mathrm{Supp}(\mu)$ 非空. 因为$X \setminus \mathrm{Supp}(\mu)$ 为开集, 所以$\mathrm{Supp}(\mu)$ 为闭集. 下面说明$\mathrm{Supp}(\mu)$ 为T 不变的. 设$x \in \mathrm{Supp}(\mu)$ 及U 为Tx 的邻域. 因为T 为连续的, 所以存在x 的邻域V 使得$V \subseteq T^{-1}U$. 这样就有$\mu(U) = \mu(T^{-1}U) \geq \mu(V) > 0$.

2. 由于$\mu(\mathrm{Supp}(\mu)) = 1$, μ 也可视为$(\mathrm{Supp}(\mu), T)$ 上的一个不变测度. 设V 是$\mathrm{Supp}(\mu)$ 的一个非空开集, 则$\mu(V) > 0$. 由命题1.3.5, $N^{\mu}(V,V) = \{n : \mu(V \cap T^{-n}V) > 0\}$ 是一个Δ^* 集. 特别地, $N(V,V) = \{n : V \cap T^{-n}V \neq \emptyset\}$ 是一个无限集. 故$\Omega(\mathrm{Supp}(\mu), T) = \mathrm{Supp}(\mu)$.

3. 假设μ 为遍历的. 则$(\mathrm{Supp}(\mu), \mathcal{B}(\mathrm{Supp}(\mu)), \mu, T)$ 也为遍历的. 如果U, V 为$\mathrm{Supp}(\mu)$ 的非空开集, 则$\mu(U)\mu(V) > 0$. 由遍历性,存在$n > 0$ 使得$\mu(U \cap T^{-n}V) > 0$, 尤其$N(U,V) \neq \emptyset$.

由于$X' = \mathrm{Supp}(\mu)$ 紧致, 取一组由可数个非空开集构成的拓扑基$\{U_n\}_{n=1}^{\infty}$. 因为

$$\mathrm{Trans}(X', T) = \bigcap_{n=1}^{\infty} \bigcup_{k=0}^{\infty} T^{-k}U_n.$$

对任意的$n \in \mathbb{N}$, $\mu(U_n) > 0$. 所以$\mu(\bigcup_{k=0}^{\infty} T^{-k}U_n) = 1$, 进而$\mu(\mathrm{Trans}(X', T)) = 1$. □

推论 6.3.5 设(X,T)是一个极小系统. 则对任意的$\mu \in \mathcal{M}(X,T)$, $\mathrm{Supp}(\mu) = X$.

由于$\Omega(\mathrm{Supp}(\mu), T) = \mathrm{Supp}(\mu)$, 所以$\mathrm{Supp}(\mu)$ 中回复点稠密. 这给出了Birkhoff 回复定理的另一个证明. 事实上我们有下面更强的结果.

定理 6.3.6 设(X,T) 是一个拓扑动力系统. 则对任意的$\mu \in \mathcal{M}(X,T)$, $\mu(\mathrm{Rec}(X,T)) = 1$.

证明. 由于X 是一个紧致度量空间, 取X 的一组可数拓扑基$\{U_n\}_{n=1}^{\infty}$ 使得对任意的$m \in \mathbb{Z}_+$, $\bigcup_{n=m}^{\infty} U_n = X$; 对任意的$m \in \mathbb{Z}_+$, $x \in X$ 和x 的一个邻域U, 存在$n \geq m$ 使得$x \in U_n \subseteq U$.

对任意的$n \in \mathbb{N}$, 令$V_n = \{x \in U_n : N(x, U_n) \text{ 是一个无限集}\}$ 以及

$$V = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} V_n.$$

不难验证$V = \mathrm{Rec}(X,T)$. 设$\mu \in \mathcal{M}(X,T)$. 由Poincaré 回复定理(定理1.3.6)证明, $\mu(V_n) = \mu(U_n)$. 则

$$\mu(\mathrm{Rec}(X,T)) = \lim_{m\to\infty} \mu\left(\bigcup_{n=m}^{\infty} V_n\right) = \lim_{m\to\infty} \mu\left(\bigcup_{n=m}^{\infty} U_n\right) = \mu(X) = 1.$$

证毕! □

注记 6.3.7 根据下面方式容易给出命题*6.3.6*另一个证明. 取X 的一组可数拓扑基$\{U_n\}_{n=1}^{\infty}$. 对任意的$n \in \mathbb{N}$, 令

$$W_n = U_n \setminus \bigcup_{i=1}^{\infty} T^{-i} U_n.$$

易见$\mu(W_n) = 0$ 和$X \setminus \mathrm{Rec}(X,T) = \bigcup_{n=1}^{\infty} W_n$. 于是$\mu(\mathrm{Rec}(X,T)) = 1$.

习　题

1. 设(X,T) 是一个动力系统. 定义(X,T) 的**支撑**(support)为

$$\mathrm{Supp}(X,T) = \{x \in X : \text{对任意 } x \text{ 的开邻域 } U,\ \text{存在} \mu \in \mathcal{M}(X,T) \text{使得} \mu(U) > 0\}.$$

证明$\mathrm{Supp}(X,T)$ 具有下列性质:

(a) $\mathrm{Supp}(X,T)$ 是一个非空不变闭子集.

(b) 对任意的$k \in \mathbb{N}$, $\mathrm{Supp}(X,T) = \mathrm{Supp}(X,T^k)$.

(c) 存在$\mu \in \mathcal{M}(X,T)$ 使得$\mathrm{Supp}(X,T) = \mathrm{Supp}(\mu)$. 但不一定存在$\mu \in \mathcal{M}^e(X,T)$ 使得$\mathrm{Supp}(X,T) = \mathrm{Supp}(\mu)$.

(d) $\mathrm{Supp}(X,T) = \{x \in X : \text{对 } x \text{ 的任意一个开邻域 } U,\ \text{存在} \mu \in \mathcal{M}^e(X,T) \text{使得} \mu(U) > 0.\}$

(e) $\mathrm{Supp}(X,T) = \overline{\bigcup_{\mu \in \mathcal{M}^e(X,T)} \mathrm{Supp}(\mu)}$.

(f) 设Y 是X 的一个闭子集. 如果对任意的$\mu \in \mathcal{M}(X,T)$, $\mu(Y) = 1$, 则$\mathrm{Supp}(X,T) \subseteq Y$.

(g) 设Y 是X 的一个闭子集. 如果对任意的$\mu \in \mathcal{M}^e(X,T)$, $\mu(Y) = 1$, 则$\mathrm{Supp}(X,T) \subseteq Y$.

§6.4 唯一遍历性

唯一遍历性是一个重要的动力系统性质，它指拓扑动力系统只有唯一一个不变测度的性质. 它由Furstenberg引入[?], 具有很多重要的刻画和应用. 我们首先给出定义和刻画, 再给出它在等分布等中的应用.

§6.4.1　唯一遍历

从定理6.2.4我们看到对于一个拓扑动力系统(X,T), $\mathcal{M}(X,T)\neq\emptyset$. 下面我们考虑一种特殊的情况.

定义 6.4.1 *设(X,T) 为一个拓扑动力系统, 我们称(X,T) 为***唯一遍历***的是指$\mathcal{M}(X,T)$ 只有一个元素.*

因为$\mathcal{M}(X,T)$ 只有一个元素, 所以$\mathcal{M}(X,T)=\mathcal{M}^e(X,T)$, 即此测度为遍历的.

下面是唯一遍历性的一些刻画：

定理 6.4.2 *设(X,T) 为拓扑动力系统, 则以下命题等价：*

1. *对任意$f\in C(X)$, $\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x))$ 一致收敛到一个常值函数;*
2. *对任意$f\in C(X)$, $\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x))$ 逐点收敛到一个常值函数;*
3. *存在$\mu\in\mathcal{M}(X,T)$ 使得对任意的$x\in X$,*
$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^ix}=\mu,$$
即对任意的$f\in C(X)$ 和$x\in X$,
$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)=\int_X f d\mu.$$
4. *T为唯一遍历的.*

证明. (1) ⇒ (2). 显然.

(2) ⇒ (3). 定义$k:C(X)\to\mathbb{C}$ 使得
$$k(f)=\lim\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x)).$$
由于$|\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(x))|\le||f||$, 易见$k$ 为连续线性算子. 并且$k(1)=1$ 以及对$f\ge 0$ 有$k(f)\ge 0$. 这样就由Riesz 表示定理存在Borel 概率测度μ 使得$k(f)=\int_X f d\mu$. 由于$k(f\circ T)=k(f)$, 于是$\int_X f\circ T\ d\mu=\int_X f d\mu$. 所以有$\mu\in\mathcal{M}(X,T)$.

(3) $\Rightarrow$ (4). 设$\nu \in \mathcal{M}(X,T)$. 由(3) 我们有: 对于任何$x \in X$ 和任何$f \in C(X)$

$$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \to f^* = \int_X f d\mu, n \to \infty.$$

对ν 积分之, 根据Lebesgue 控制收敛定理就有

$$\int_X f d\nu = \int_X f^* d\nu = f^* = \int_X f d\mu,\ \forall f \in C(X).$$

于是就有$\nu = \mu$.亦即, T 为唯一遍历的.

(4) $\Rightarrow$ (1). 设$\{\mu\} = \mathcal{M}(X,T)$. 如果$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ 一致收敛于一个常值, 那么此常值必为$\int_X f d\mu$. 假设(1)不成立, 那么存在$g \in C(X)$, $\varepsilon > 0$ 使得对任意$N \in \mathbb{N}$ 存在$n > N$ 及$x_n \in X$ 使得

$$\left|\frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x_n)) - \int_X g d\mu\right| \geq \varepsilon.$$

如果设$\mu_n = \frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i x_n}$, 那么上式改写为

$$\left|\int_X g d\mu_n - \int_X g d\mu\right| \geq \varepsilon.$$

取$\{\mu_n\}_{n\in\mathbb{N}}$ 的收敛子列$\{\mu_{n_i}\}_{i\in\mathbb{N}}$. 如果$\mu_{n_i} \to \mu_\infty$, 那么由命题6.2.2, $\mu_\infty \in \mathcal{M}(X,T)$. 因为$\left|\int_X g d\mu_\infty - \int_X g d\mu\right| \geq \varepsilon$, 所以$\mu_\infty \neq \mu$. 这与$T$ 的唯一遍历性矛盾. □

定理 6.4.3 设(X,μ,T) 为唯一遍历的拓扑动力系统, 那么$(\mathrm{Supp}(\mu),T)$ 为极小的.

证明. 设$Y \subseteq \mathrm{Supp}(\mu)$ 为非空闭的不变子集. 存在$\nu \in \mathcal{M}(Y,T)$ 使得$\mathrm{Supp}(\nu) \subseteq Y$. 因为$(X,T)$ 为唯一遍历的, $\nu = \mu$ 及$Y = \mathrm{Supp}(\mu)$. 所以$(\mathrm{Supp}(\mu),T)$ 为极小的. □

命题 6.4.4 设(X,T) 是一个动力系统. 则(X,T) 是唯一遍历的当且仅当对任意的$f \in C(X)$, 函数列$\{\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\}_{n=1}^{\infty}$ 有一个子列逐点收敛到一个常值函数.

证明. "$\Rightarrow$"根据定理6.4.2即可.

"$\Leftarrow$" 设$\mu,\nu \in \mathcal{M}^e(X,T)$ 和$f \in C(X)$. 则存在$c \in \mathbb{R}$ 和递增的正整数序列$\{n_k\}$ 使得

$$\lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1} f(T^i x) = c, \quad \forall x \in X.$$

由Lebesgue 控制收敛定理, 函数列$\{\frac{1}{n_k}\sum_{i=0}^{n_k-1} f(T^i x)\}_{k=1}^{\infty}$ 在$L^1(X,\mu)$ 中依范数收敛到c 和在$L^1(X,\nu)$ 中依范数收敛到c. 由L^1-平均遍历定理(定理2.2.6), 函数列$\{\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\}_{n=1}^{\infty}$

在$L^1(X,\mu)$中依范数收敛到$\int_X fd\mu$ 和在$L^1(X,\nu)$ 中依范数收敛到$\int_X fd\nu$. 故

$$\int_X fd\mu = c = \int_X fd\nu.$$

由f 的任意性, $\mu=\nu$. 故$\mathcal{M}^e(X,T)$ 是一个单点集, 从而(X,T) 是唯一遍历的. □

命题 6.4.5 设(X,T) 是一个传递系统. 如果对任意的$f\in C(X)$, 函数列$\{\frac{1}{n}\sum_{i=0}^{n-1} f(T^ix)\}_{n=1}^\infty$ 是等度连续的, 则(X,T) 是唯一遍历的.

证明. 设$f\in C(X)$. 不难验证函数列$\{\frac{1}{n}\sum_{i=0}^{n-1} f(T^ix)\}_{n=1}^\infty$ 是一致有界. 根据题设, 这个函数列还是等度连续的. 由Arezlá-Ascoli 定理(见定理0.2.44), 存在$f^*\in C(X)$ 和子列$\{n_k\}$ 使得$\{\frac{1}{n_k}\sum_{i=0}^{n_k-1} f(T^ix)\}_{k=1}^\infty$ 一致收敛于f^*. 对任意的$x\in X$, 不难验证$f^*(Tx)=f^*(x)$. 由于(X,T) 是传递的和f^* 连续, f^* 是一个常值函数. 由命题6.4.4, (X,T) 是唯一遍历的. □

§6.4.2 一些例子

定义 6.4.6 设(X,T)是一个动力系统. 如果它既是极小的又是唯一遍历的, 则称之为**严格遍历的**(*strictly ergodic*).

例 6.4.7 *1.* 设$n\in\mathbb{Z}$, $(\mathbb{Z}_n,R_1)$是严格遍历的.

2. 当$\alpha\in\mathbb{T}$是无理数时, $(\mathbb{T},R_\alpha)$是严格遍历的.

3. 加法机器系统$(\Sigma_2,R_{\mathbf{1}})$是严格遍历的.

4. 设$k\geq 2$和$\alpha_1,\dots,\alpha_k\in(0,1)\setminus\mathbb{Q}$. 若这些$\alpha_i$是有理无关的, 则$(\mathbb{T}^k,\prod_{i=1}^k R_{\alpha_i})$是严格遍历的.

定理 6.4.8 设$T:\mathbb{S}^1\to\mathbb{S}^1$ 为没有周期点的同胚. 那么T 为唯一遍历的, 并且它半共轭于圆周无理旋转, 即存在因子映射$\phi:(\mathbb{S}^1,T)\to(\mathbb{S}^1,S)$, 其中$S$ 为无理旋转.

$$\begin{tikzcd}
\mathbb{S}^1 \arrow[r, "T"] \arrow[d, "\phi"'] & \mathbb{S}^1 \arrow[d, "\phi"] \\
\mathbb{S}^1 \arrow[r, "S"'] & \mathbb{S}^1
\end{tikzcd}$$

对于任何点$w\in\mathbb{S}^1$, $\phi^{-1}(w)$ 要么为独点集, 要么为闭区间.

如果T 为极小的, 那么ϕ 为同构.

证明. 因为T 没有周期点, 所以任何不变测度是无原子的. 设$\mu_1, \mu_2 \in \mathcal{M}(\mathbb{S}^1, T)$, 我们需要证明$\mu_1 = \mu_2$. 令$\nu = \frac{1}{2}(\mu_1 + \mu_2) \in \mathcal{M}(\mathbb{S}^1, T)$. 定义

$$\phi : \mathbb{S}^1 \to \mathbb{S}^1,\ \phi(z) = e^{2\pi i \nu(\overrightarrow{[1,z]})},$$

其中$\overrightarrow{[a,b]}$ 表示从a 逆时针到b 的闭弧. 因为ν 无原子, 所以ϕ 为连续满射. 一个重要的事实是:

$$\nu(\overrightarrow{[z_1,z_2]}) + \nu(\overrightarrow{[z_2,z_3]}) = \nu(\overrightarrow{[z_1,z_3]}) \pmod 1,\ \forall z_1, z_2, z_3 \in \mathbb{S}^1.$$

令

$$S : \mathbb{S}^1 \to \mathbb{S}^1,\ z \mapsto az,$$

其中$a = e^{2\pi i\alpha}, \alpha = \nu(\overrightarrow{[1,T(1)]})$. 于是

$$\begin{aligned}\phi(Tz) &= e^{2\pi i \nu(\overrightarrow{[1,T(z)]})} \\ &= e^{2\pi i\left(\nu(\overrightarrow{[1,T(1)]}) + \nu(\overrightarrow{[T(1),T(z)]})\right)} \\ &= e^{2\pi i \nu(\overrightarrow{[1,T(1)]})} e^{2\pi i \nu(\overrightarrow{[1,z]})} = a\phi(z) \\ &= S(\phi(z)).\end{aligned}$$

即ϕ 为因子映射.

下证α 为无理数. 否则存在$p \in \mathbb{N}$ 使得$a^p = 1$, 于是$\phi(T^p z) = \phi(z), \forall z \in \mathbb{S}^1$. 所以

$$\nu(\overrightarrow{[1,T^p z]}) = \nu(\overrightarrow{[1,z]}) \pmod 1, \quad \forall z \in \mathbb{S}^1.$$

因为T^p 没有不动点, 所以存在$\delta > 0$ 使得

$$d(z, T^p z) > \delta,\ \forall z \in \mathbb{S}^1,$$

其中d 为$\mathbb{S}^1$ 上的度量. 于是$\mathbb{S}^1$ 每个点都是某个长为δ 区间的端点, 并且区间的ν 测度为0. 这将推出$\nu(\mathbb{S}^1) = 0$, 此不可能. 于是$\alpha \notin \mathbb{Q}$.

$(\mathbb{S}^1, S)$ 为无理旋转, 所以它是唯一遍历的, 遍历测度为Haar测度m. 于是

$$\phi_*\nu = \phi_*\mu_1 = \phi_*\mu_2 = m.$$

设$\overrightarrow{[a,b]}$ 为$\mathbb{S}^1$ 区间, 那么

$$\phi(\overrightarrow{[a,b]}) = [e^{2\pi i \nu(\overrightarrow{[1,a]})}, e^{2\pi i \nu(\overrightarrow{[1,b]})}].$$

设

$$\phi^{-1}\Big(\phi(\overrightarrow{[a,b]})\Big) = \overrightarrow{[c,d]},$$

其中

$$c = \inf\{z : \nu(\overrightarrow{[a,z]}) = 0\};\ d = \sup\{w : \nu(\overrightarrow{[b,w]}) = 0\}.$$

因为$\nu\Big(\phi^{-1}(\phi(\overrightarrow{[a,b]}))\Delta\overrightarrow{[a,b]}\Big)=0$, 所以

$$\mu_i\Big(\phi^{-1}(\phi(\overrightarrow{[a,b]}))\Delta\overrightarrow{[a,b]}\Big)=0, i=1,2.$$

于是

$$\mu_i(\overrightarrow{[a,b]})=\mu_i\Big(\phi^{-1}(\phi(\overrightarrow{[a,b]}))\Big)=m(\phi(\overrightarrow{[a,b]})),\ i=1,2.$$

尤其

$$\mu_1(\overrightarrow{[a,b]})=\mu_2(\overrightarrow{[a,b]}).$$

因此$\mu_1=\mu_2$, 即T 唯一遍历.

设$w\in\mathbb{S}^1$, 下面考虑$\phi^{-1}(w)$. 设$z_1\in\phi^{-1}(w)$, 那么

$$z_2\in\phi^{-1}(w)\Rightarrow\nu(\overrightarrow{[z_1,z_2]})=0\text{ 或 }1.$$

于是$\phi^{-1}(w)$ 为包含z_1 的最大ν 零测的闭区间.

如果T 为极小的, 那么对于任何非平凡开区间I 有$\nu(I)>0$. 根据上面证明, $\phi^{-1}(w)$ 为包含z_1 的最大ν 零测的闭区间. 所以此时只能是独点集, 即ϕ 为同胚. □

例 6.4.9 (更多的例子) *下面系统为严格遍历的, 但是大部分例子证明都不容易, 所以我们不再给出证明, 感兴趣的读者可以参见[76, Theorem 4.11].*

1. *设$T_\phi:\mathbb{T}^2\to\mathbb{T}^2,(x,y)\mapsto(x+\alpha,y+\phi(x))$, 其中$\alpha\notin\mathbb{Q}$, $\phi:\mathbb{T}\to\mathbb{T}$ 满足Lipschitz 条件且具有非退化的度.*
2. *极小幂零系统.*
3. *horocycle 系统.*
4. *Morse 系统.*
5. *Chacón 系统.*
6. *规则的几乎自守系统. 即$\pi:(X,T)\to(X_{eq},T)$ 使得$m(\{y\in X_{eq}:\mathrm{Card}\ \pi^{-1}(y)=1\})=1$.*

§6.4.3　等分布

定义 6.4.10 设$\{x_n\}_{n=1}^\infty$ 是$[0,1]$中的一个序列. 如果

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n f(x_k)=\int_0^1 f dx,\ \forall f\in C([0,1]),$$

则称$\{x_n\}_{n=1}^\infty$ 在$[0,1]$ 是**等分布的** *(equidistributed)* 或**一致分布的** *(uniformly distributed)*.

引理 6.4.11 设$\{x_n\}_{n=1}^{\infty}$ 是$[0,1]$ 中的一个序列.那么以下等价:

1. 序列$\{x_n\}_{n=1}^{\infty}$ 为等分布的.
2. 对于任何$k \neq 0$,
$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} e^{2\pi i k x_j} = 0.$$
3. 对任意的$0 \le a < b \le 1$,
$$\lim_{n\to\infty} \frac{|\{x_1, x_2, \dots, x_n\} \cap [a,b]|}{n} = b - a.$$

证明. (1) 蕴含(2) 是显然的, (2) 蕴含(1) 是因为三角多项式在$C([0,1])$ 中稠密.

(3) 蕴含(1) 是因为连续函数被区间上的简单函数一致逼近. 所以仅需证明(1) 蕴含(3). 设$0 \le a < b \le 1$. 对于任何$\varepsilon > 0$, 我们定义逼近$\mathbf{1}_{[a,b]}$ 的函数如下:

$$f^+(x) = \begin{cases} 1, & a \le x \le b; \\ \frac{x-(a-\varepsilon)}{\varepsilon}, & \max\{0, a-\varepsilon\} \le x < a; \\ \frac{(b+\varepsilon)-x}{\varepsilon}, & b < x \le \min\{b+\varepsilon, 1\}; \\ 0, & \text{其它}x \end{cases}$$

以及

$$f^-(x) = \begin{cases} 1, & a+\varepsilon \le x \le b-\varepsilon; \\ \frac{x-a}{\varepsilon}, & a \le x < a+\varepsilon; \\ \frac{b-x}{\varepsilon}, & b-\varepsilon < x \le b; \\ 0, & \text{其它}x. \end{cases}$$

注意到

$$f^-(x) \le \mathbf{1}_{[a,b]} \le f^+(x), \forall x \in [0,1]$$

且

$$\int_0^1 \Big(f^+(x) - f^-(x)\Big) dx \le 2\varepsilon.$$

于是

$$\frac{1}{n}\sum_{j=1}^{n} f^-(x_j) \le \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{[a,b]}(x_j) \le \frac{1}{n}\sum_{j=1}^{n} f^+(x_j).$$

于是根据等分布定义,

$$\begin{aligned} b - a - 2\varepsilon \le \int_0^1 f^-(x)dx &\le \liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{[a,b]}(x_j) \\ &\le \limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{[a,b]}(x_j) \le \int_0^1 f^+(x)dx \le b - a + 2\varepsilon. \end{aligned}$$

于是

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\mathbf{1}_{[a,b]}(x_j)=b-a.$$

证毕! □

定理 6.4.12 (Weyl等分布定理) 设α是一个无理数．则$\{n\alpha \pmod 1\}_{n=1}^{\infty}$ 在$[0,1]$是等分布的．

证明. 由于$\alpha\in\mathbb{T}$ 是无理数时, $(\mathbb{T},R_\alpha)$ 是唯一遍历的．设m 为其上Haar 测度. 对任意的$x\in\mathbb{T}$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\delta_{R_\alpha^i x}=m.$$

设$0\le a<b\le 1$. 则$m(\partial([a,b]))=m(\{a,b\})=0$. 由命题6.1.7,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\delta_{R_\alpha^i 0}([a,b])&=\lim_{n\to\infty}\frac{|\{0,\alpha,\dots,(n-1)\alpha \ (\mathrm{mod} 1)\}\cap[a,b]|}{n}\\&=m([a,b])=b-a.\end{aligned}$$

证毕! □

作为应用，我们有下面例子.

例 6.4.13 用

$$S=(1,2,4,8,1,3,6,1,\dots)$$

表示序列

$$\{2^n:n\in\mathbb{Z}_+\}=\{1,2,4,8,16,32,64,128,\dots\}$$

的十进制表示中首数字组成的序列．设$k\in\{1,2,\dots,9\}$．用$p_k(n)$ 表示k 在S前n个位置中出现的次数. 那么我们有

$$\lim_{n\to\infty}\frac{p_k(n)}{n}=\log_{10}(\frac{k+1}{k}).$$

事实上, 2^i 的十进制表示的首位数字为k 但且仅当存在$r\ge 0$ 使得

$$k\cdot 10^r\le 2^i<(k+1)\cdot 10^r.$$

此时

$$r+\log_{10}k\le i\log_{10}2<r+\log_{10}(k+1).$$

即$r=[i\log_{10}2]$ 以及

$$\log_{10}k\le i\log_{10}2-[i\log_{10}2]<\log_{10}(k+1).$$

设$J_k = [\log_{10} k, \log_{10}(k+1))$ 以及$\alpha = \log_{10} 2$. 考虑无理旋转$(\mathbb{T}, R_\alpha)$. 根据上面分析: 2^i 的十进制表示的首位数字为k 但且仅当$R_\alpha^i 0 \in J_k$. 于是

$$\lim_{n\to\infty} \frac{p_k(n)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{J_k}(R_\alpha^i 0) = \log_{10}(k+1) - \log_{10} k = \log_{10}(\frac{k+1}{k}).$$

定义 6.4.14 设X 为紧致度量空间, $\mu \in \mathcal{M}(X)$, $\{x_n\}_{n\in\mathbb{N}} \subseteq X$. 称序列$\{x_n\}_{n\in\mathbb{N}}$ 关于μ 为**等分布的**(*equidistributed*), 是指

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(x_j) = \int_X f(x) d\mu(x),\ \forall f \in C(X).$$

等价的, 序列$\{x_n\}_{n\in\mathbb{N}}$ 关于μ 为等分布的当且仅当

$$\frac{1}{n} \sum_{j=1}^{n} \delta_{x_j} \to \mu, n \to \infty,$$

收敛取弱*拓扑.

注记 6.4.15 根据定义, 设(X,T) 为拓扑动力系统, $\mu \in \mathcal{M}(X,T)$. 那么点$x \in X$ 为通用的当且仅当$\{T^n x\}_{n\in\mathbb{N}}$ 为关于μ 等分布的.

定理 6.4.16 (Furstenberg) 设(Y,S) 是一个唯一遍历系统和ν 为其唯一的不变测度, G 是紧致可度量化的交换群和m 为其上的$Haar$ 测度. 设$\xi : X \to G$ 为连续映射, 定义$X = Y \times G$ 以及斜积映射

$$T : X \to X,\ (y,g) \mapsto (Sy, \xi(y)g).$$

如果$\nu \times m$ 是(X,T) 上的一个遍历测度, 那么(X,T) 是唯一遍历的.

证明. 因为$\nu \times m$ 为遍历的, $\nu \times m$ 通用点集合非空. 对于任何$\nu \times m$ 的通用点$(y,g) \in X$ 以及任意的$h \in G$. 设

$$R_h : X \to X,\ (y,g) \mapsto (y, gh).$$

由于m 是R_h-不变的, 所以(y,hg) 也是$\nu \times m$ 的一个通用点, 即$\{y\} \times G$ 均为通用点. 根据$\nu \times m$ 是遍历的, 则$\nu \times m$ 几乎处处的点为$\nu \times m$ 通用的. 则存在一个ν 满测集Y_0 使得每个$Y_0 \times G$ 中的点是$\nu \times m$ 通用的.

下设μ 是(X,T) 上的一个遍历测度和$p_1 : X \to Y$, $(y,g) \mapsto y$. 则$p_{1*}(\mu) = \nu$, 从而$\mu(Y_0 \times G) = 1$. 所以存在$(y,g) \in Y_0 \times G$ 使得(y,g) 是μ 通用的. 此时必有$\mu = \nu \times m$, 即(X,T) 是唯一遍历的. □

命题 6.4.17 设$\alpha \notin \mathbb{Q}$ 和$k \geq 2$,

$$T : \mathbb{T}^k \to \mathbb{T}^k, \quad (x_1, x_2, \ldots, x_k) \mapsto (x_1 + \alpha, x_2 + x_1, \ldots, x_k + x_{k-1}).$$

则$(\mathbb{T}^k, T)$ 是严格遍历的.

证明. 根据定理6.4.16, 我们仅需要证明$(\mathbb{T}^k, T)$ 为遍历的. 设$f \in L^2(\mathbb{T}^k, m)$ 为T 不变的, 其中m 为Haar 测度(此时即Lebesgue 测度). 设其Fourier 级数为

$$f(\mathbf{x}) = \sum_{\mathbf{n} \in \mathbb{Z}^k} c_{\mathbf{n}} e^{2\pi i \mathbf{n} \cdot \mathbf{x}}.$$

设

$$T' : \mathbb{Z}^k \to \mathbb{Z}^k, \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_{k-1} \\ n_k \end{pmatrix} \mapsto \begin{pmatrix} n_1 + n_2 \\ n_2 + n_3 \\ \vdots \\ n_{k-1} + n_k \\ n_k \end{pmatrix}.$$

根据$f(\mathbf{x}) = f(T\mathbf{x})$, 我们得到

$$\sum_{\mathbf{n} \in \mathbb{Z}^k} c_{\mathbf{n}} e^{2\pi i \mathbf{n} \cdot \mathbf{x}} = \sum_{\mathbf{n} \in \mathbb{Z}^k} c_{\mathbf{n}} e^{2\pi i n_1 \alpha} e^{2\pi i T' \mathbf{n} \cdot \mathbf{x}}.$$

根据Fourier 级数的唯一性, 我们得到

$$c_{T'\mathbf{n}} = e^{2\pi i \alpha n_1} c_{\mathbf{n}}.$$

于是

$$|c_{T'\mathbf{n}}| = |c_{\mathbf{n}}|, \ \forall \mathbf{n} \in \mathbb{Z}^k.$$

于是对于任何$\mathbf{n} \in \mathbb{Z}^k$, 要么$\mathbf{n}, T'\mathbf{n}, (T')^2\mathbf{n}, \ldots$ 互异（此时因为$\sum_{\mathbf{n} \in \mathbb{Z}^k} |\mathbf{n}|^2 < \infty$, 我们有$c_{\mathbf{n}} = 0$）; 要么存在$p > q$ 使得$(T')^p \mathbf{n} = (T')^q(\mathbf{n})$（此时可以归纳证明$n_2 = n_3 = \ldots = n_k = 0$）. 对于$\mathbf{n} = (n_1, 0, 0, \ldots, 0)$ 带入前面关系式$c_{T'\mathbf{n}} = e^{2\pi i \alpha n_1} c_{\mathbf{n}}$ 我们得到$n_1 = 0$ 或$c_{\mathbf{n}} = 0$. 所以f 为常值函数, 即T 为遍历的. □

定理 6.4.18 (Weyl) 设$P(n) = a_k n^k + \cdots + a_1 n + a_0$ 是一个实多项式. 如果$a_1, a_2, \ldots, a_k$ 中至少有一个是无理数, 那么$\{P(n) \pmod 1\}_{n=1}^{\infty}$ 在$[0,1]$ 是等分布的.

证明. 假设命题对于任何阶数少于k 的多项式成立. 如果$a_k \in \mathbb{Q}$, 那么存在$q \in \mathbb{Z}$ 使得$q a_k \in \mathbb{Z}$. 于是对于任何$j = 0, 1, \ldots, q-1$, $\{P(qn+j) \pmod 1\}_n$ 取值与为阶数严格小于k 的多项式吻合. 根据假设, 每个$\{P(qn+j) \pmod 1\}_n$ 为等分布的, 从而$\{P(n) \pmod 1\}_{n=1}^{\infty}$ 是等分布的. 于是根据上面分析, 我们不妨可以假设$a_k \notin \mathbb{Q}$. 令$T : \{\alpha\} \times \mathbb{T}^k \to \{\alpha\} \times \mathbb{T}^k$ 为

$$\begin{pmatrix} \alpha \\ x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} \mapsto \begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & & \ddots & \\ & & & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} \alpha \\ x_1 + \alpha \\ x_2 + x_1 \\ \vdots \\ x_{k_1} + x_k \end{pmatrix}.$$

迭代之,

$$\begin{pmatrix}1 & & & & & \\ 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & 1\end{pmatrix}^n \begin{pmatrix}\alpha \\ x_1 \\ x_2 \\ \vdots \\ x_k\end{pmatrix} = \begin{pmatrix}1 & & & & \\ n & 1 & & & \\ \binom{n}{2} & n & 1 & & \\ \vdots & & & \ddots & \\ \binom{n}{k} & & & n & 1\end{pmatrix} \begin{pmatrix}\alpha \\ x_1 \\ x_2 \\ \vdots \\ x_k\end{pmatrix}$$

$$= \begin{pmatrix}\alpha \\ n\alpha + x_1 \\ \binom{n}{2}\alpha + nx_1 + x_2 \\ \vdots \\ \binom{n}{k}\alpha + \binom{n}{k-1}x_1 + \ldots + nx_{k-1} + x_k\end{pmatrix}$$

我们取$\alpha = k!a_k$, 取$x_1, \ldots, x_k$ 使得

$$P(n) = \binom{n}{k}\alpha + \binom{n}{k-1}x_1 + \ldots + nx_{k-1} + x_k.$$

根据命题6.4.17, 这个系统的任何点轨道是等分布的, 尤其是最后个分量取值也是等分布的. □

习 题

1. 设$\{a_i\}_{i=0}^{\infty}$是一个有界序列. 定义

$$\limsup_{n-m\to\infty} \frac{1}{n-m}\sum_{i=m}^{n-1} a_i = \lim_{k\to\infty}\left(\sup_{n-m\geq k}\frac{1}{n-m}\sum_{i=m}^{n-1} a_i\right)$$

和

$$\liminf_{n-m\to\infty} \frac{1}{n-m}\sum_{i=m}^{n-1} a_i = \lim_{k\to\infty}\left(\inf_{n-m\geq k}\frac{1}{n-m}\sum_{i=m}^{n-1} a_i\right).$$

如果

$$\limsup_{n-m\to\infty} \frac{1}{n-m}\sum_{i=m}^{n-1} a_i = A = \liminf_{n-m\to\infty} \frac{1}{n-m}\sum_{i=m}^{n-1} a_i,$$

则称极限$\lim\limits_{n-m\to\infty} \frac{1}{n-m}\sum\limits_{i=m}^{n-1} a_i$ 存在并定义它的极限是A.

设$\{a_i\}_{i=0}^{\infty}$是一个有界序列. 证明

$$\limsup_{n-m\to\infty} \frac{1}{n-m}\sum_{i=m}^{n-1} a_i = \sup\Big\{\limsup_{k\to\infty}\frac{1}{n_k-m_k}\sum_{i=m_k}^{n_k-1} a_i :$$
$$\{n_k\}\text{和}\{m_k\}\text{递增且}\lim_{k\to\infty}(n_k-m_k)=\infty\Big\},$$

并且存在递增的序列$\{n_k\}$和$\{m_k\}$使得$\lim_{k\to\infty}(n_k-m_k)=\infty$ 和

$$\lim_{k\to\infty}\frac{1}{n_k-m_k}\sum_{i=m_k}^{n_k-1}a_i=\limsup_{n-m\to\infty}\frac{1}{n-m}\sum_{i=m}^{n-1}a_i.$$

2. 设$\{a_i\}_{i=0}^{\infty}$是一个有界序列. 如果极限$\lim\limits_{n-m\to\infty}\frac{1}{n-m}\sum\limits_{i=m}^{n-1}a_i$ 存在且等于A, 则对任意满足$\lim_{k\to\infty}(n_k-m_k)=\infty$的递增的自然数序列$\{n_k\}$ 和$\{m_k\}$, 有

$$\lim_{k\to\infty}\frac{1}{n_k-m_k}\sum_{i=m_k}^{n_k-1}a_i=A.$$

3. 设(X,T)是一个动力系统. 则则下列论断等价:
   (a) (X,T)是唯一遍历的;
   (b) 对任意的$f\in C(X)$, $\lim\limits_{n-m\to\infty}\frac{1}{n-m}\sum\limits_{i=m}^{n-1}f(T^ix)$ 逐点收敛到一个常值函数;
   (c) 存在$\mu\in\mathcal{M}(X,T)$使得对任意的$x\in X$,

$$\lim_{n-m\to\infty}\frac{1}{n-m}\sum_{i=m}^{n-1}\delta_{T^ix}=\mu,$$

   即对任意的$f\in C(X)$ 和$x\in X$,

$$\lim_{n-m\to\infty}\frac{1}{n-m}\sum_{i=m}^{n-1}f(T^ix)=\int fd\mu.$$

4. 设(X,T)是一个动力系统. 若(X,T)是唯一遍历的, 则对任意的$f\in C(X)$和$x\in X$,

$$\lim_{n-m\to\infty}\frac{1}{n-m}\sum_{i=m}^{n-1}f(T^ix)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix).$$

5. 设$F\subseteq\mathbb{Z}_+$. F的**上Banach密度**(upper Banach density)定义为

$$\overline{BD}(F)=\limsup_{n-m\to\infty}\frac{|F\cap\{m,m+1,\dots,m+n-1\}|}{n-m}.$$

   类似地, F的**下Banach密度**(lower Banach density)定义为

$$\underline{BD}(F)=\liminf_{n-m\to\infty}\frac{|F\cap\{m,m+1,\dots,m+n-1\}|}{n-m}.$$

   如果$\overline{BD}(F)=\underline{BD}(F)$, 则称$F$的**Banach密度**(Banach density)存在, 它的Banach密度$BD(F)=\overline{BD}(F)$.

   设(X,T)是一个唯一遍历系统和μ为其唯一的不变测度. 则对任意的$x\in X$和开集U, $N(x,U)$的下Banach密度至少为$\mu(U)$.

6. 设$k\in\mathbb{N}$. 如果(X,T^k)是唯一遍历的, 则(X,T)也是唯一遍历的. 逆命题是否成立?
7. 设$\pi:(X,T)\to(Y,S)$是一个因子映射. 如果(X,T)是唯一遍历的, 则(Y,S)也是唯一遍历的.
8. 设(X,T)是一个可逆系统. 则(X,T)是唯一遍历的当且仅当(X,T^{-1})是唯一遍历的.

§6.5 测度动力系统的拓扑模型

动力系统的模型理论是动力系统的一个重要组成部分, 也是遍历理论与拓扑动力系统的一个交汇点. 给定一个拓扑动力系统, 那么自然由全体开集生成一个σ 代数（即Borel σ 代数）, 并且可以证明上面必定存在一个不变测度. 于是拓扑动力系统自然可以看成一个保测系统, 从而可以将遍历论的方法运用到拓扑动力系统的研究中去. 反之, 给定一个保测系统, 是否也可以赋予它拓扑结构, 使上面拓扑结构与测度结构相容? 这个问题要困难很多, 也是动力系统模型理论涉及的主要内容. 它的一个直接应用就是使得拓扑动力系统的方法也可以用到遍历论的研究中. 如果对拓扑结构不加苛刻的限制条件, 那么这个问题相对容易回答, 例如经典的结果告诉我们任何具有一定可数性要求的保测系统一定测度同构于一个Lebesgue 系统; 任何Lebesgue 系统测度同构于一个Cantor 空间上的拓扑动力系统等. 但是如果要对拓扑结构加上限制条件使得系统的测度结构与拓扑结构更加融洽, 那么问题就不是那么容易了.

在模型理论中最为令人吃惊的结论是著名的Jewett-Krieger 定理, 它指出任何遍历系统都测度同构于一个唯一遍历的极小拓扑动力系统. "唯一遍历"是指系统仅有唯一一个不变概论测度. 这个定理指出了遍历理论和拓扑动力系统的深刻内在联系. 在Jewett-Krieger 定理出现之后, 包括著名数学家美国艺术与科学院院士B. Weiss 在内的很多数学工作者对此方面进行了研究. 例如, Weiss 证明了存在一个极小系统使得任何遍历系统都测度同构于这个极小系统上的某个测度系统[197]. 另外他还证明了Jewett-Krieger 定理对于相对化和更一般的群作用也是成立的[196]. 此外我们还能够在拓扑模型上要求更多的性质, 例如：E. Lehrer 证明了任何遍历系统都测度同构与一个唯一遍历的极小拓扑强混合系统[142]；Glasner 和Weiss 证明了系统具有正熵当且仅当它有u.p.e. 的唯一遍历模型, [2] 而具有零熵当且仅当它有prime 的唯一遍历模型[78, 81]. [3]

定义 6.5.1 (拓扑模型) *设$(X,\mathcal{B}(X),\mu,T)$ 是一个保测系统, 其中$(X,\mathcal{B}(X),\mu)$ 是一个Borel 概率空间. 设$(\widehat{X},\widehat{T})$ 是一个拓扑动力系统. 如果存在$\widehat{\mu}\in\mathcal{M}(\widehat{X},\widehat{T})$ 使得$(X,\mathcal{B}(X),\mu,T)$ 和$(\widehat{X},\mathcal{B}(\widehat{X}),\widehat{\mu},\widehat{T})$ 是测度同构的, 则称$(\widehat{X},\mathcal{B}(\widehat{X}),\widehat{\mu},\widehat{T})$ 是$(X,\mathcal{B}(X),\mu,T)$ 的一个***拓扑模型** *(topological model)或***拓扑实现** *(topological realization).*

之前我们已经知道每个保测映射都有一个拓扑模型.

定理 6.5.2 (Furstenberg) *[64] 任何可逆保测系统都存在一个拓扑模型.*

Weiss 说明存在万有的拓扑模型：证明了存在一个极小系统使得任何遍历系统都测度同构于这个极小系统上的某个测度系统. 具体讲：

[2] u.p.e. 的定义参见后面定义7.12.1

[3] prime 系统指没有非平凡因子的系统. 例如Chacón 系统为prime 的.

定理 6.5.3 (Weiss) *[81, 197] 存在一个可逆极小系统$(\widehat{X},\widehat{T})$, 使得对任意的可逆非原子遍历系统$(X,\mathcal{B}(X),\mu,T)$, 存在$\widehat{\mu}\in\mathcal{M}(\widehat{X},\widehat{T})$ 使得$(\widehat{X},\mathcal{B}(\widehat{X}),\widehat{\mu},\widehat{T})$ 是$(X,\mathcal{B}(X),\mu,T)$ 的一个拓扑模型.*

在模型理论中最为著名的就是Jewett-Krieger定理. 它的证明比较复杂, 我们在此略去, 请参见[199]中Weiss给的简化证明.

定理 6.5.4 (Jewett-Krieger定理) *[120, 138] 任何可逆遍历系统都存在一个唯一遍历的极小的拓扑模型.*

我们还可以在Jewett-Krieger定理基础上添加一些性质, 例如有下面的Lehrer 定理. 这个定理在构造很多问题的反例时十分有用.

定理 6.5.5 (Lehrer) *[142] 任何可逆非原子遍历系统都存在一个唯一遍历的极小强混合的拓扑模型.*

Weiss 还给出了Jewett-Krieger定理的相对化版本. 我们称$\hat{\pi}:\hat{X}\to\hat{Y}$ 为因子映射$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 的拓扑实现或拓扑模型, 是指$\hat{\pi}$ 为拓扑因子映射, 并且存在测度同构ϕ 和ψ 使得下图表交换

$$\begin{tikzcd} X \arrow[r, "\phi"] \arrow[d, "\pi"'] & \hat{X} \arrow[d, "\hat{\pi}"] \\ Y \arrow[r, "\psi"] & \hat{Y} \end{tikzcd}$$

即$\hat{\pi}\phi=\psi\pi$. Weiss 于1985 年将Jewett-Krieger 定理推广到相对化的情况. 他证明了：

定理 6.5.6 (Weiss) *[199] 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为因子映射, 其中$(X,\mathcal{X},\mu,T)$ 遍历. 设$(\hat{Y},\mathcal{B}(\hat{Y}),\hat{\nu},\hat{S})$ 为$(Y,\mathcal{Y},\nu,T)$ 唯一遍历模型, 那么存在$(X,\mathcal{X},\mu,T)$ 唯一遍历模型$(\hat{X},\mathcal{B}(\hat{X}),\hat{\mu},\hat{T})$ 以及拓扑因子映射$\hat{\pi}:\hat{X}\to\hat{Y}$ 使得它为$\pi:X\to Y$ 的模型.*

§6.6 注记

测度空间的基本性质可以参见测度论的专著, 例如[32]. 涉及关于凸集的结构的Krein-Milman定理等可以参见[173]以及[163]等. 更多关于唯一遍历及其应用的内容可以参见[54, 64, 76, 199] 等. 关于拓扑模型理论方面的证明参见[81, 76, 199].

拓扑动力系统与遍历理论密切相关，关于介绍二者关联的综述性文章参见[81, 114].

第七章　测度熵与拓扑熵

保测系统上的熵的概念是由Kolmogorov 在1958 年给出的[133], 之后Adler 等人在拓扑动力系统中引入了拓扑熵的定义[2]. 熵是目前为止发现的最重要的共扼不变量之一, 并得到广泛、深入地研究.

本章主要涉及熵的经典理论, 具体来说我们先介绍测度熵, 给出其基本性质, 而后给出拓扑熵的定义并研究它们的基本属性. 在此处我们会重点介绍测度熵与拓扑熵的计算, 熵的变分原理等. 之后我们将介绍测度Pinsker σ 代数和熵的Pinsker 公式; 重新介绍测度Kolmogorov 系统, 然后研究其基本属性并证明Rohlin-Sinai定理等.

§7.1　测度熵

1958 年Kolmogorov [133] 借鉴Shannon [180] 信息论中不确定性的描述在遍历论中引入了测度熵的概念, 测度熵是重要的同构不变量, 它反映了保测系统的混乱程度.

§7.1.1　一些关于剖分和σ代数的概念和符号

设$(X, \mathcal{X}, \mu)$ 为概率空间. 由X 的有限个互不相交的可测集构成的X 的覆盖, 我们称其为X 的**有限可测剖分**. 确切地说, $\alpha = \{A_1, A_2, \cdots, A_n\}$ 是X 的有限可测剖分, 如果$A_i \in \mathcal{X}$, $A_i \cap A_j = \emptyset$ 对任意$1 \le i \neq j \le n$ 成立且$\bigcup_{i=1}^n A_i = X$.

对X 的有限可测剖分α, 我们用$\hat{\alpha}$ 或者$\sigma(\alpha)$ 表示由α 生成的σ代数. 对X的剖分α, β, 如果$\sigma(\alpha) \subset \sigma(\beta)$, 我们就说$\beta$ 为α 的**加细**, 记为$\beta \succeq \alpha$ 或$\alpha \preceq \beta$.

对于有限σ 代数$\mathcal{C}$, 我们记由$\mathcal{C}$ 确定的剖分为$\xi(\mathcal{C})$. 设$\mathcal{C} = \{C_1, C_2, \ldots, C_n\}$, 那么易见$\xi(\mathcal{C})$ 中元形如$B_1 \cap \ldots \cap B_n$, 其中$B_i = C_i$ 或$X \setminus C_i$. 易见对于有限σ 代数$\mathcal{C}$ 和有限剖分η, 我们有

$$\sigma(\xi(\mathcal{C})) = \mathcal{C}, \quad \xi(\sigma(\eta)) = \eta.$$

于是$\mathcal{X}$ 的有限子σ 代数与有限剖分η 一一对应.

设α, β 为X 的剖分, 那么剖分的**交** (join) 定义为

$$\alpha \vee \beta = \{A \cap C : A \in \alpha, C \in \beta\}.$$

对$\mathcal{X}$ 的两个子σ 代数$\mathcal{F}_1, \mathcal{F}_2$, 我们用$\mathcal{F}_1 \vee \mathcal{F}_2$ 表示$\mathcal{X}$ 的同时包含$\mathcal{F}_1, \mathcal{F}_2$的最小子$\sigma$代数. 易见对于$\sigma$ 代数$\mathcal{C}, \mathcal{D}$ 和剖分α, β 我们有

$$\xi(\mathcal{C} \vee \mathcal{D}) = \xi(\mathcal{C}) \vee \xi(\mathcal{D}); \quad \sigma(\alpha \vee \beta) = \sigma(\alpha) \vee \sigma(\beta).$$

为方便起见, 在本书中如果上下文没有歧义的情况下, 我们有时会混用剖分和其对应的σ 代数. 例如, $\alpha \vee \mathcal{F}$ 表示$\sigma(\alpha) \vee \mathcal{F}$.

设$T : X \to X$ 为保测映射, $\alpha = \{A_1, \ldots, A_k\}$ 为可测剖分, $n \in \mathbb{N}$, 那么

$$T^{-n}\alpha = \{T^{-n}A_1, \ldots, T^{-n}A_k\}$$

仍为剖分. 设$\mathcal{A}$ 为$\mathcal{X}$ 的子σ 代数, 那么

$$T^{-n}\mathcal{A} = \{T^{-n}A : A \in \mathcal{A}\}$$

仍为子σ 代数.

§7.1.2 剖分的熵

根据信息论和概率论, 我们有如下关于剖分熵的定义.

定义 7.1.1 设$(X, \mathcal{X}, \mu)$ 为概率空间, $\alpha = \{A_1, A_2, \cdots, A_n\}$ 是X 的有限可测剖分, 令

$$H_\mu(\alpha) = \sum_{j=1}^{n} -\mu(A_j)\log\mu(A_j).$$

我们将数$H_\mu(\alpha)$称为可测剖分α的**熵**.

显然$H_\mu(\alpha) \in [0, +\infty)$. 进一步, 我们考虑函数$\phi : [0,1] \to [0, +\infty)$,

$$\phi(t) = \begin{cases} -t\log t, & t > 0 \\ 0, & t = 0 \end{cases}.$$

我们不难看出函数ϕ 是严格凸函数: 即对满足$\sum_{i=1}^{k} p_i = 1$ 的$p_i \geq 0$ 和$t_i \in [0,1]$, 我们有

$$\phi(\sum_{i=1}^{k} p_i t_i) \geq \sum_{i=1}^{k} p_i\phi(t_i),$$

其中等号成立当且仅当对所有满足$p_i \neq 0$ 的t_i 是彼此相等的.

注记 7.1.2 设$(X, \mathcal{X}, \mu)$ 为概率空间.

1. 设$\mathcal{N} = \{X, \emptyset\}$, 那么$H_\mu(\mathcal{N}) = 0$.

2. 设$\alpha = \{A_1, \ldots, A_k\}$ 为剖分, 并且$\mu(A_1) = \ldots = \mu(A_k) = \frac{1}{k}$, 那么我们有

$$H_\mu(\hat{\alpha}) = -\sum_{i=1}^{k} \frac{1}{k}\log\frac{1}{k} = \log k.$$

   用ϕ 的凸性容易证明, 在元素个数为k 的所有剖分中, $\log k$ 为熵的最大值.

3. 设$\mathcal{A}, \mathcal{C}$ 为有限子σ 代数, 并且$\mathcal{A} \underset{\mu}{=} \mathcal{C}$, 那么$H_\mu(\mathcal{A}) = H_\mu(\mathcal{C})$.

4. $T : X \to X$ 保测, $\mathcal{A}$ 为有限子σ 代数, 那么$H_\mu(T^{-1}\mathcal{A}) = H_\mu(\mathcal{A})$.

§7.1.3 条件期望与信息函数

设$(X,\mathcal{X},\mu)$ 为概率空间. 我们把$\mathcal{X}$ 中的元素称为X 的**可测集**. 对X 的一个有限可测剖分$\alpha=\{A_1,A_2,\cdots,A_n\}$, 我们将$\alpha$ 的**信息函数**定义为

$$I(\alpha)(x)=\sum_{j=1}^{n}-\mathbf{1}_{A_j}(x)\log\mu(A_j),$$

其中$\mathbf{1}_{A_j}(x)$ 为A_j 的特征函数.

更一般地, 对$\mathcal{X}$ 的给定的子σ 代数$\mathcal{F}$, 可以定义相应的信息函数和条件熵.

定义 7.1.3 *设$(X,\mathcal{X},\mu)$ 为概率空间, $\alpha=\{A_1,A_2,\cdots,A_n\}$ 是X 的有限可测剖分. 对$\mathcal{X}$ 的给定的子σ 代数$\mathcal{F}$, 可测剖分α 相对于$\mathcal{F}$ 的***条件信息函数***定义为*

$$I(\alpha|\mathcal{F})(x)=\sum_{j=1}^{n}-\mathbf{1}_{A_j}(x)\log\mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F})(x),$$

在这里$\mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F})$ 为函数$\mathbf{1}_{A_j}$ 相对于$\mathcal{F}$ 的数学期望.

注记 7.1.4 *1. 设$\mathcal{N}=\{\emptyset,X\}$为$\mathcal{X}$的平凡子$\sigma$代数, 显然$I(\alpha|\mathcal{N})=I(\alpha)$.*

2. 当$\mathcal{F}=\mathcal{X}$时, 对于任意的$A\in\mathcal{X}$我们有$\mathbb{E}(\mathbf{1}_A|\mathcal{F})=\mathbf{1}_A$, 进而$I(\alpha|\mathcal{F})=0$. .

对于X 的有限可测剖分$\alpha=\{A_1,A_2,\cdots,A_n\}$ 和$\beta=\{B_1,B_2,\cdots,B_m\}$ 容易验证

$$I(\alpha|\hat{\beta})=\sum_{i=1}^{n}\sum_{j=1}^{m}-\mathbf{1}_{A_i\cap B_j}(x)\log\frac{\mu(A_i\cap B_j)}{\mu(B_j)}. \tag{7.1.1}$$

由公式(7.1.1), 对X 的有限可测剖分α, β 和γ 容易验证

$$I(\alpha\vee\beta|\hat{\gamma})=I(\alpha|\hat{\gamma})+I(\beta|\hat{\alpha}\vee\hat{\gamma}) \tag{7.1.2}$$

以下我们将上述公式(7.1.2)一般化.

命题 7.1.5 *设$(X,\mathcal{X},\mu)$ 为概率空间, α,β 为X 的有限可测剖分以及$\mathcal{F}$ 为$\mathcal{X}$ 的子σ 代数. 则对μ-a.e. x 有*

$$I(\alpha\vee\beta|\mathcal{F})(x)=I(\alpha|\mathcal{F})(x)+I(\beta|\hat{\alpha}\vee\mathcal{F})(x).$$

证明. 首先我们说明对$B\in\beta$, 作为$L^1(X,\hat{\alpha}\vee\mathcal{F},\mu)$ 函数而言成立

$$\mathbb{E}(\mathbf{1}_B|\alpha\vee\mathcal{F})(x)=\sum_{A\in\alpha}\mathbf{1}_A(x)\frac{\mathbb{E}(\mathbf{1}_{B\cap A}|\mathcal{F})(x)}{\mathbb{E}(\mathbf{1}_A|\mathcal{F})(x)}. \tag{7.1.3}$$

为此目的, 设$A' \in \alpha$, $F \in \mathcal{F}$, 将特征函数$\mathbf{1}_{A'\cap F}$ 与(7.1.3) 的右边项相乘, 再积分我们有

$$\begin{aligned}
&\int_X \mathbf{1}_{A'\cap F}(x)\cdot(\sum_{A\in\alpha}\mathbf{1}_A(x)\frac{\mathbb{E}(\mathbf{1}_{B\cap A}|\mathcal{F})(x)}{\mathbb{E}(\mathbf{1}_A|\mathcal{F})(x)})d\mu(x) = \int_F \mathbf{1}_{A'}(x)\frac{\mathbb{E}(\mathbf{1}_{B\cap A'}|\mathcal{F})(x)}{\mathbb{E}(\mathbf{1}_{A'}|\mathcal{F})(x)}d\mu(x)\\
&= \int_F \mathbb{E}(\mathbf{1}_{A'}\frac{\mathbb{E}(\mathbf{1}_{B\cap A'}|\mathcal{F})}{\mathbb{E}(\mathbf{1}_{A'}|\mathcal{F})}|\mathcal{F})(x)d\mu(x) = \int_F \mathbb{E}(\mathbf{1}_{A'}|\mathcal{F})(x)\cdot\mathbb{E}(\frac{\mathbb{E}(\mathbf{1}_{B\cap A'}|\mathcal{F})}{\mathbb{E}(\mathbf{1}_{A'}|\mathcal{F})}|\mathcal{F})(x)d\mu(x)\\
&= \int_X \mathbf{1}_F(x)\mathbb{E}(\mathbf{1}_{B\cap A'}|\mathcal{F})(x)d\mu(x) = \int_X \mathbb{E}(\mathbf{1}_F\cdot\mathbf{1}_{B\cap A'}|\mathcal{F})(x)d\mu(x)\\
&= \int_X \mathbf{1}_F(x)\cdot\mathbf{1}_{B\cap A'}(x)d\mu(x) = \int_{A'\cap F}\mathbf{1}_B(x)d\mu(x) = \int_{A'\cap F}\mathbb{E}(\mathbf{1}_B|\hat{\alpha}\vee\mathcal{F})(x)d\mu(x)\\
&= \int_X \mathbf{1}_{A'\cap F}(x)\cdot\mathbb{E}(\mathbf{1}_B|\hat{\alpha}\vee\mathcal{F})(x)d\mu(x).
\end{aligned}$$

因A', F是任意的, 从期望的定义我们获得(7.1.3). 现在在等式(7.1.3)两边取对数即得

$$\log(\mathbb{E}(\mathbf{1}_B|\alpha\vee\mathcal{F})(x)) = \sum_{A\in\alpha}\mathbf{1}_A(x)\left(\log(\mathbb{E}(\mathbf{1}_{B\cap A}|\mathcal{F})(x)) - \log(\mathbb{E}(\mathbf{1}_A|\mathcal{F})(x))\right),$$

因此

$$\begin{aligned}
I(\alpha\vee\beta|\mathcal{F})(x) &= -\sum_{A\in\alpha}\sum_{B\in\beta}\mathbf{1}_{A\cap B}(x)\log\mathbb{E}(\mathbf{1}_{A\cap B}|\mathcal{F})(x)\\
&= -\sum_{A\in\alpha}\sum_{B\in\beta}\mathbf{1}_A(x)\mathbf{1}_B(x)\log(\mathbb{E}(\mathbf{1}_A|\mathcal{F})(x)) - \sum_{B\in\beta}\mathbf{1}_B(x)\log(\mathbb{E}(\mathbf{1}_B|\alpha\vee\mathcal{F})(x))\\
&= -\sum_{A\in\alpha}\mathbf{1}_A(x)\log(\mathbb{E}(\mathbf{1}_A|\mathcal{F})(x)) - \sum_{B\in\beta}\mathbf{1}_B(x)\log(\mathbb{E}(\mathbf{1}_B|\alpha\vee\mathcal{F})(x))\\
&= I(\alpha|\mathcal{F})(x) + I(\beta|\hat{\alpha}\vee\mathcal{F})(x),
\end{aligned}$$

这就完成了命题的证明. □

推论 7.1.6 设$(X,\mathcal{X},\mu)$ 为概率空间, $\beta_1,\beta_2,\ldots,\beta_n$ 为有限剖分, 那么

$$I(\bigvee_{j=1}^n \beta_j) = I(\beta_1) + \sum_{j=1}^{n-1} I(\beta_{j+1}|\hat{\beta}_1\vee\ldots\vee\hat{\beta}_j).$$

§7.1.4　熵的定义

定义 7.1.7 设$(X,\mathcal{X},\mu)$ 为概率空间, $\alpha = \{A_1, A_2, \cdots, A_n\}$ 是X 的有限可测剖分. 对$\mathcal{X}$ 的给定的子σ 代数$\mathcal{F}$, 可测剖分α 相对于$\mathcal{F}$ 的**条件熵**定义为

$$\begin{aligned}
H_\mu(\alpha|\mathcal{F}) &= \int_X I(\alpha|\mathcal{F})(x)d\mu(x)\\
&= \sum_{j=1}^n \int_X -\mathbf{1}_{A_j}(x)\log\mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F})d\mu(x).
\end{aligned}$$

由于

$$\mathbb{E}(I(\alpha|\mathcal{F})|\mathcal{F}) = \sum_{j=1}^{n} -\mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F}) \log \mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F}) = \sum_{j=1}^{n} \phi(\mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F})),$$

因此

$$\begin{aligned} H_\mu(\alpha|\mathcal{F}) &= \int_X I(\alpha|\mathcal{F})(x) d\mu(x) = \int_X \mathbb{E}(I(\alpha|\mathcal{F})|\mathcal{F}) d\mu(x) \\ &= \sum_{j=1}^{n} \int_X \phi(\mathbb{E}(\mathbf{1}_{A_j}|\mathcal{F})) d\mu(x). \end{aligned} \tag{7.1.4}$$

注记 7.1.8 *1.* 设$\mathcal{N} = \{\emptyset, X\}$为$\mathcal{X}$的平凡子$\sigma$代数, 显然$H_\mu(\alpha) = H_\mu(\alpha|\mathcal{N})$.

2. 当$\mathcal{F} = \mathcal{X}$时, 对于任意的$A \in \mathcal{X}$我们有$\mathbb{E}(\mathbf{1}_A|\mathcal{F}) = \mathbf{1}_A$, 进而$H_\mu(\alpha|\mathcal{F}) = 0$.

可测剖分α的条件熵具有以下性质:

命题 7.1.9 设$(X, \mathcal{X}, \mu)$ 为概率空间. 对X 的有限测度α, β 以及$\mathcal{X}$ 的子σ 代数$\mathcal{F}, \mathcal{F}_1, \mathcal{F}_2$, 我们有

1. 如果$\alpha = \{A_1, A_2, \cdots, A_k\}$, 则

$$H_\mu(\alpha|\mathcal{F}) \le \log k.$$

等号成立当且仅当$\mathbb{E}(1_{A_i}|\mathcal{F}) \equiv \frac{1}{k}$对每个$i \in \{1, 2, \cdots, k\}$. 特别地$H_\mu(\alpha) \le \log k$且等号成立当且仅当$\mu(A_i) = \frac{1}{k}$对每个$i \in \{1, 2, \cdots, k\}$;

2. $H_\mu(\alpha \vee \beta|\mathcal{F}) = H_\mu(\alpha|\mathcal{F}) + H_\mu(\beta|\hat{\alpha} \vee \mathcal{F})$. 特别地$H_\mu(\alpha \vee \beta) = H_\mu(\alpha) + H_\mu(\beta|\hat{\alpha})$;

3. $H_\mu(\alpha|\mathcal{F}) = 0$成立当且仅当$\alpha$ 为$\mathcal{F}$可测的.特别地$H_\mu(\alpha|\beta) = 0$ 成立当且仅当$\alpha \preceq \beta$;

4. $\mathcal{F}_1 \subset \mathcal{F}_2 \Rightarrow H_\mu(\alpha|\mathcal{F}_1) \ge H_\mu(\alpha|\mathcal{F}_2)$;

5. $\alpha \preceq \beta \Rightarrow H_\mu(\alpha|\mathcal{F}) \le H_\mu(\beta|\mathcal{F})$. 特别地$H_\mu(\alpha) \le H_\mu(\beta)$;

6. $H_\mu(\alpha \vee \beta|\mathcal{F}) \le H_\mu(\alpha|\mathcal{F}) + H_\mu(\beta|\mathcal{F})$. 特别地$H_\mu(\alpha \vee \beta) \le H_\mu(\alpha) + H_\mu(\beta)$;

7. 设$T : (X, \mathcal{X}) \to (X, \mathcal{X})$为可测映射, 则

$$H_\mu(T^{-1}\alpha|T^{-1}\mathcal{F}) = H_{T\mu}(\alpha|\mathcal{F}).$$

如果$(X, \mathcal{X}, \mu, T)$为保测系统, 则$H_\mu(T^{-1}\alpha|T^{-1}\mathcal{F}) = H_\mu(\alpha|\mathcal{F})$.

证明. (1) 首先由于ϕ 是区间$[0,\infty)$ 上严格凸的函数, 因此

$$\begin{aligned} H_\mu(\alpha|\mathcal{F}) &= \int_X \sum_{j=1}^{k} \phi(\mathbb{E}(1_{A_j}|\mathcal{F}))d\mu(x) \\ &\le \int_X k\phi(\sum_{j=1}^{k}\frac{1}{k}\mathbb{E}(1_{A_j}|\mathcal{F}))d\mu(x) \\ &= \int_X \log k d\mu(x) = \log k. \end{aligned}$$

且等号成立当且仅当$\mathbb{E}(1_{A_i}|\mathcal{F}) \equiv \frac{1}{k}$ 每个$i \in \{1,2,\cdots,k\}$. 这就证明了(1).

(2,5,6) 由命题7.1.5 知对μ-a.e. x 有$I^{\mathcal{F}}(\alpha\vee\beta)(x) = I^{\mathcal{F}}(\alpha)(x) + I^{\alpha\vee\mathcal{F}}(\beta)(x)$ 成立. 对上式两边积分即得

$$H_\mu(\alpha\vee\beta|\mathcal{F}) = H_\mu(\alpha|\mathcal{F}) + H_\mu(\beta|\alpha\vee\mathcal{F}).$$

这就证明了(2). 现在由(2) 我们得到(5) 和(6).

(3) 注意到α 为$\mathcal{F}$ 可测的当且仅当$I^{\mathcal{F}}(\alpha)=0$, 由公式(7.1.4) 即得.

(4) 设$\alpha=\{A_1,A_2,\cdots,A_k\}$. 则

$$\begin{aligned} H_\mu(\alpha|\mathcal{F}_1) &= \sum_{i=1}^{k}\int_X \phi(\mathbb{E}(1_{A_i}|\mathcal{F}_1)(x))d\mu(x) = \sum_{i=1}^{k}\int_X \phi(\mathbb{E}(\mathbb{E}(1_{A_i}|\mathcal{F}_2))|\mathcal{F}_1)(x))d\mu(x) \\ &\ge \sum_{i=1}^{k}\int_X \mathbb{E}(\phi(\mathbb{E}(1_{A_i}|\mathcal{F}_2))|\mathcal{F}_1)(x))d\mu(x) \quad (\text{Jensen不等式}) \\ &= \sum_{i=1}^{k}\int_X \phi(\mathbb{E}(1_{A_i}|\mathcal{F}_2)(x))d\mu(x) = H_\mu(\alpha|\mathcal{F}_2). \end{aligned}$$

这就证明了(4).

(7) 对可测函数f, 我们有

$$\mathbb{E}(f\circ T|T^{-1}\mathcal{F})(x) = \mathbb{E}(f|\mathcal{F})(Tx)$$

对μ-a.e. $x\in X$ 成立. □

注记 7.1.10 *在上面我们用到了如下形式的Jensen 不等式. 设$f\in L^1(X,\mathcal{X},\mu)$, $0\le f(x)\le 1$, a.e., 以及$\mathcal{A}\subset\mathcal{X}$ 为子σ 代数. 设$\phi:[0,1]\to\mathbb{R}$ 为凸函数(即$\phi(px+(1-p)y)\ge p\phi(x)+(1-p)\phi(y), p\in[0,1]$), 那么$\phi(\mathbb{E}(f|\mathcal{A}))\ge\mathbb{E}(\phi(f)|\mathcal{A})$.*

定义 7.1.11 *设α,β 为$(X,\mathcal{X},\mu)$ 的两个有限可测剖分, 称α* **独立于** *β, 是指*

$$\mu(A\cap B)=\mu(A)\mu(B),\ \forall A\in\alpha, B\in\beta,$$

此时记$\alpha\perp_\mu\beta$.

命题 7.1.12 设$(X,\mathcal{X},\mu,T)$ 为保测系统和α,β 为X 的两个有限可测剖分. 则以下性质彼此等价:

1. α 独立于β;
2. $H_\mu(\alpha\vee\beta)=H_\mu(\alpha)+H_\mu(\beta)$;
3. $H_\mu(\alpha|\hat{\beta})=H_\mu(\alpha)$.

证明. 由命题7.1.9 (2), 我们知(2)等价于(3).以下证明(1)等价于(3). 首先假设(1)成立, 即$\alpha\perp_\mu\beta$. 因此对任意$A\in\alpha$, $B\in\beta$, $\mu(A\cap B)=\mu(A)\mu(B)$, 进而

$$I(\alpha|\hat{\beta})(x)=\sum_{A\in\alpha,B\in\beta}-1_{A\cap B}(x)\log\frac{\mu(A\cap B)}{\mu(B)}=\sum_{A\in\alpha,B\in\beta}-1_{A\cap B}(x)\log\mu(A)=I(\alpha)(x).$$

这就说明$H_\mu(\alpha|\hat{\beta})=H_\mu(\alpha)$.

反之, 假设$H_\mu(\alpha|\hat{\beta})=H_\mu(\alpha)$.使用(7.1.1), 对凸函数$\phi(t)=-t\log t$我们有

$$0=\sum_{A\in\alpha}\phi(\mu(A))-\sum_{A\in\alpha}\sum_{B\in\beta}\mu(B)\phi(\frac{\mu(A\cap B)}{\mu(B)}). \tag{7.1.5}$$

因ϕ是严格凸的, 对每个$A\in\alpha$我们有

$$\phi(\mu(A))\geq\sum_{B\in\beta}\mu(B)\phi(\frac{\mu(A\cap B)}{\mu(B)}), \tag{7.1.6}$$

且等号成立当且仅当$\frac{\mu(A\cap B)}{\mu(B)}=\mu(A)$对$B\in\beta,\mu(B)>0$成立.

现在结合(7.1.5)和(7.1.6), 对$A\in\beta$我们得到

$$\phi(\mu(A))=\sum_{B\in\beta}\mu(B)\phi(\frac{\mu(A\cap B)}{\mu(B)}).$$

再注意到(7.1.6)不等式成立条件, 对$B\in\beta,\mu(B)>0$ 我们有

$$\frac{\mu(A\cap B)}{\mu(B)}=\mu(A).$$

这就证明$\alpha\perp_\mu\beta$. □

§7.1.5 动力系统测度熵的定义

设$(X,\mathcal{X},\mu,T)$ 为保测系统, α 为X 的有限或可数可测剖分. 对整数$m\leq n\in\mathbb{Z}$ 我们记

$$\alpha_m^n=\bigvee_{i=m}^{n}T^{-i}\alpha=T^{-m}\alpha\vee T^{-m+1}\alpha\vee\ldots\vee T^{-n}\alpha.$$

尤其

$$\alpha_0^{n-1}=\alpha\vee T^{-1}\alpha\vee\ldots\vee T^{-n+1}\alpha.$$

我们还常用下面符号：

$$\alpha^- = \alpha_1^\infty = \bigvee_{n=1}^{+\infty} T^{-n}\alpha \text{ 以及 } \alpha^T = \alpha_{-\infty}^\infty = \bigvee_{n=-\infty}^{\infty} T^{-n}\alpha,$$

其中$\bigvee_{n=1}^{+\infty} T^{-n}\alpha$ 表示$\mathcal{X}$ 的包含$\bigcup_{n=1}^{\infty} T^{-n}\alpha$ 的最小子σ 代数.

为定义动力系统的熵, 我们需要下面引理：

引理 7.1.13 设$\{a_n\} \subset \mathbb{R}$ 为次可加非负序列, 即

$$a_{m+n} \le a_n + a_m.$$

那么

$$\lim_{n\to\infty} \frac{1}{n} a_n = \inf_n \frac{1}{n} a_n.$$

证明. 设$a = \inf\limits_{n\ge 1} \dfrac{a_n}{n}$. 固定$\ell \ge 1$, 对每个$m \in \mathbb{N}$ 存在$k_m \in \mathbb{Z}_+$ 和$r_m \in \{0, 1, \cdots, \ell - 1\}$ 使得$m = k_m\ell + r_m$. 利用$\{a_n\}$ 的次可加性可得

$$\frac{a_m}{m} \le \frac{k_m a_\ell + a_{r_m}}{k_m \ell + r_m},$$

再让$m \to +\infty$ 便得$\limsup_{m\to\infty} \frac{a_m}{m} \le \frac{a_\ell}{\ell}$. 现在从$\ell$ 的任意性, 我们有

$$\limsup_{m\to\infty} \frac{a_m}{m} \le \inf_{\ell \ge 1} \frac{a_\ell}{\ell}.$$

而相反方向的不等式是明显成立的. □

命题 7.1.14 设$(X, \mathcal{X}, \mu, T)$ 为保测系统. 对X 的有限可测剖分α 序列$H_\mu(\alpha_0^{n-1})$ 为次可加非负序列, 进而极限$\lim\limits_{n\to+\infty} \dfrac{1}{n} H_\mu(\alpha_0^{n-1})$ 存在且

$$\lim_{n\to+\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\alpha_0^{n-1}).$$

证明. 序列a_n的次可加性来自于以下不等式

$$\begin{aligned}
& a_{n+m} \\
&= H_\mu(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(n+m-1)}\alpha) \\
&\le H_\mu(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(n-1)}\alpha) + H_\mu(T^{-n}(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(m-1)}\alpha)) \\
&= a_n + a_m.
\end{aligned} \tag{7.1.7}$$

于是

$$\lim_{n\to+\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\alpha_0^{n-1}).$$

□

根据上面命题我们可以定义剖分的熵：

定义 7.1.15 设$(X,\mathcal{X},\mu,T)$为保测系统.对X的有限可测剖分α, 我们用$h_\mu(T,\alpha)$ 表示极限

$$\lim_{n\to+\infty}\frac{1}{n}H_\mu(\alpha_0^{n-1})$$

且称$h_\mu(T,\alpha)$为**剖分α相对于T的熵**.

定义 7.1.16 设$(X,\mathcal{X},\mu,T)$为保测系统. 我们称

$$h_\mu(T)=\sup_\alpha h_\mu(T,\alpha)$$

为**系统$(X,\mathcal{X},\mu,T)$ 的熵**, 其中α 取遍X 的有限可测剖分.

注记 7.1.17 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为因子映射. 对Y的有限可测剖分α, 容易验证$h_\nu(S,\alpha)=h_\mu(T,\pi^{-1}\alpha)$. 进而从定义*7.1.16* 我们知$h_\mu(T)\geq h_\nu(S)$. 特别地, 当$\pi$为测度同构时, $h_\mu(T)=h_\nu(S)$. 这说明**测度熵是测度同构不变量!**

命题 7.1.18 设$(X,\mathcal{X},\mu,T)$为保测系统, α,β 为有限剖分, 那么

$$h_\nu(T,\alpha)\leq h_\mu(T,\beta)+H(\alpha|\beta).$$

证明. 首先我们有:

$$\begin{aligned}
H_\mu(\alpha_0^n)&\leq H_\mu(\alpha_0^n\vee\beta_0^n)=H_\mu(\beta_0^n)+H_\mu(\alpha_0^n|\beta_0^n)\\
&\leq H_\mu(\beta_0^n)+\sum_{j=0}^n H_\mu(T^{-j}\alpha|\beta_0^n)\\
&\leq H_\mu(\beta_0^n)+\sum_{j=0}^n H_\mu(T^{-j}\alpha|T^{-j}\beta)\\
&=H(\beta_0^n)+(n+1)H_\mu(\alpha|\beta).
\end{aligned}$$

于是两边除以$n+1$ 趋向无穷即有命题. □

由此我们得到:

推论 7.1.19 设$(X,\mathcal{X},\mu,T)$为保测系统, α,β 为有限剖分, 那么

$$|h_\mu(\alpha)-h_\mu(\beta)|\leq H(\beta|\alpha)+H(\alpha|\beta).$$

命题 7.1.20 对保测系统$(X,\mathcal{X},\mu,T)$ 和每个$m\in\mathbb{N}$, 我们有

$$h_\mu(T^m)=mh_\mu(T).$$

证明. 首先对X每个可测剖分α, 我们不难算得

$$h_\mu(T^m,\alpha_0^{m-1})=\lim_{k\to\infty}\frac{1}{k}H_\mu(\bigvee_{i=0}^{k-1}T^{-im}\alpha_0^{m-1})=\lim_{k\to\infty}\frac{1}{k}H_\mu(\alpha_0^{km-1})=mh_\mu(T,\alpha).$$

因此由定义我们可有

$$h_\mu(T^m)=\sup_\alpha h_\mu(T^m,\alpha_0^{m-1})=m\sup_\alpha h_\mu(T,\alpha)=mh_\mu(T).$$

其中α取遍X的所有有限可测剖分. □

注记 7.1.21 当保测系统$(X,\mathcal{X},\mu,T)$ 可逆时, 我们可以说明对X 的每个可测剖分α有$h_\mu(T,\alpha)=h_\mu(T^{-1},\alpha)$ 成立, 进而

$$h_\mu(T^m)=|m|h_\mu(T),\quad m\in\mathbb{Z}.$$

习　题

1. 设保测系统$(X,\mathcal{X},\mu,T)$ 为可逆的, 证明：对X 的每个可测剖分α 有$h_\mu(T,\alpha)=h_\mu(T^{-1},\alpha)$ 成立.

§7.2　鞅定理、Kolmogorov-Sinai定理以及一些例子的计算

鞅理论是概率论的重要理论之一, 它在动力系统研究中也有着重要应用.

§7.2.1　鞅定理

在此我们介绍鞅定理. 首先介绍一个很有用的引理, 此引理类似于著名的Chebyshev不等式[1]

引理 7.2.1 (Maximal 引理) 设$(X,\mathcal{X},\mu)$ 是概率空间, $\mathcal{X}_1\subset\mathcal{X}_2\subset\ldots\subset\mathcal{X}_N\subset\mathcal{X}$ 为子σ 代数列, $\lambda>0$, $f\in L^1(X,\mathcal{X},\mu)$. 设$E=\{x\in X:\max\limits_{1\le n\le N}\mathbb{E}(f|\mathcal{X}_n)(x)>\lambda\}$, 那么

$$\mu(E)\le\frac{1}{\lambda}\int_X|f|d\mu=\frac{1}{\lambda}\|f\|_1.$$

证明. 不失一般性, 我们假设$f\ge0$ (否则用$f^+=\max\{f(x),0\}$ 代替f). 设

$$E_n=\{x\in X:\mathbb{E}(f|\mathcal{X}_n)(x)>\lambda;\ \mathbb{E}(f|\mathcal{X}_j)\le\lambda,\ 1\le i\le n-1\}.$$

那么我们就有$E_n\in\mathcal{X}_n$, $E_i\cap E_j=\emptyset,i\ne j$, 且

$$E=E_1\cup\ldots\cup E_N.$$

[1]对于$f\in L^1(X,\mathcal{X},\mu)$, $\lambda>0$, Chebyshev 不等式指$\mu(\{x\in X:f(x)>\lambda\})\le\frac{1}{\lambda}\int_X|f|d\mu=\frac{\|f\|_1}{\lambda}$.

于是
$$\int_E fd\mu = \sum_{n=1}^N \int_{E_n} fd\mu = \sum_{n=1}^N \int_{E_n} \mathbb{E}(f|\mathcal{X}_n)d\mu \geq \sum_{n=1}^N \lambda\mu(E_n) = \lambda\mu(E).$$
进而,
$$\mu(E) \leq \frac{1}{\lambda}\int_X fd\mu = \frac{1}{\lambda}\int_X |f|d\mu.$$
证毕! □

定理 7.2.2 *(递增鞅定理; Increasing Martingale Theorem)* 设$(X,\mathcal{X},\mu)$为概率空间, $f \in L^1(X,\mathcal{X},\mu)$. 如果$\{\mathcal{X}_n\}_{n=1}^\infty$ 为$\mathcal{X}$ 递增的子σ 代数且满足$\mathcal{X}_n \nearrow \mathcal{X}_\infty$ ($\mathcal{X}_\infty$ 为由$\bigcup_{n=1}^\infty \mathcal{X}_n$ 生成的σ 代数), 则
$$\lim_{n\to+\infty} \mathbb{E}(f|\mathcal{X}_n) = \mathbb{E}(f|\mathcal{X}_\infty)$$
在$L^1(\mu)$ 和μ-*a.e.* 意义下同时成立.

证明. 不失一般性, 我们设$\mathcal{X}_\infty = \mathcal{X}$. 首先定理对于$L^1(X,\mathcal{X},\mu)$ 的稠密子集$\bigcup_{k=1}^\infty L^1(X,\mathcal{X}_k,\mu)$ 成立：如果$f \in L^1(X,\mathcal{X}_k,\mu)$, 那么$\mathbb{E}(f|\mathcal{X}_n) = g, \forall n \geq k$. 证明的主体在于用稠密子集逼近一般的函数.

设$f \in L^1(X,\mathcal{X},\mu)$. 对于任何取定$\varepsilon > 0$, 存在$k \in \mathbb{N}$ 以及$g \in L^1(X,\mathcal{X}_k,\mu)$ 使得$\|f-g\|_1 < \varepsilon$. 于是对于任何$n \geq k$, $\mathbb{E}(g|\mathcal{X}_n) = g$, 进而
$$\begin{aligned}
&\|\mathbb{E}(f|\mathcal{X}_n) - f\|_1 \\
\leq &\|\mathbb{E}(f|\mathcal{X}_n) - \mathbb{E}(g|\mathcal{X}_n)\|_1 + \|\mathbb{E}(g|\mathcal{X}_n) - g\|_1 + \|g-f\|_1 \\
\leq &2\|f-g\|_1 < 2\varepsilon.
\end{aligned}$$
于是
$$\limsup_{n\to\infty} \|\mathbb{E}(f|\mathcal{X}_n) - f\|_1 \leq 2\varepsilon.$$
因为ε 任意, 我们得到$\lim_{n\to\infty} \|\mathbb{E}(f|\mathcal{X}_n) - f\|_1 = 0$.

下面证明逐点收敛. 首先我们有：
$$\begin{aligned}
&\mu(\{x \in X : \limsup_{n\to\infty} |\mathbb{E}(f|\mathcal{X}_n)(x) - f(x)| > \varepsilon^{1/2}\}) \\
\leq &\mu\left(\{x \in X : \limsup_{n\to\infty} \left(|\mathbb{E}((f-g)|\mathcal{X}_n)(x) - (f-g)(x)| + |\mathbb{E}(g|\mathcal{X}_n)(x) - g(x)|\right) > \varepsilon^{1/2}\}\right) \\
\leq &\mu\left(\{x \in X : \limsup_{n\to\infty} \left(|\mathbb{E}((f-g)|\mathcal{X}_n)(x)| + |(f-g)(x)|\right) > \varepsilon^{1/2}\}\right) \\
\leq &\mu\left(\{x \in X : \limsup_{n\to\infty} |\mathbb{E}((f-g)|\mathcal{X}_n)(x)| > \frac{1}{2}\varepsilon^{1/2}\}\right) + \mu\left(\{x \in X : \limsup_{n\to\infty} |(f-g)(x)| > \frac{1}{2}\varepsilon^{1/2}\}\right) \\
\leq &2\frac{1}{\frac{1}{2}\varepsilon^{1/2}}\|f-g\|_1 \leq 4\varepsilon^{1/2}.
\end{aligned}$$

其中最后不等式根据引理7.2.1 以及Chebyshev不等式. 因为ε 任意, 所以

$$\mathbb{E}(f|\mathcal{X}_n)(x) \xrightarrow{a.e.} f(x),\ n \to \infty.$$

证毕! □

定理 7.2.3 *(递减鞅定理; Decreasing Martingale Theorem)* 设$(X, \mathcal{X}, \mu)$ 为概率空间, $f \in L^1(X, \mathcal{X}, \mu)$. 如果$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为$\mathcal{X}$ 递减的子σ 代数且满足$\mathcal{X}_n \searrow \mathcal{X}_\infty = \bigcap_{n=1}^{\infty} \mathcal{X}_n$, 则

$$\lim_{n\to+\infty} \mathbb{E}(f|\mathcal{X}_n) = \mathbb{E}(f|\mathcal{X}_\infty)$$

在$L^1(\mu)$ 和μ-*a.e.* 意义下同时成立.

证明. 因为对于取定$N \in \mathbb{N}$, $\mathcal{X}_N \subset \mathcal{X}_{N-1} \subset \ldots \subset \mathcal{X}_1$ 为递增的, 根据引理7.2.1, 对$f \in L^1(X, \mathcal{X}, \mu), \lambda > 0$

$$\mu(\{x \in X : \sup_{1\le n\le N} \mathbb{E}(f|\mathcal{X}_n) > \lambda\}) \le \frac{1}{\lambda}\|f\|_1.$$

令$N \to \infty$, 于是

$$\mu(\{x \in X : \sup_{1\le n\le \infty} \mathbb{E}(f|\mathcal{X}_n) > \lambda\}) \le \frac{1}{\lambda}\|f\|_1,$$

以及

$$\mu(\{x \in X : \limsup_{n\to\infty} \mathbb{E}(f|\mathcal{X}_n) > \lambda\}) \le \frac{1}{\lambda}\|f\|_1.$$

我们后面将用到这个不等式.

设$V_n = \mathrm{Ker}\mathbb{E}(\cdot|\mathcal{X}_n) = \{f \in L^1(X, \mathcal{X}, \mu) : \mathbb{E}(f|\mathcal{X}_n) = 0\}$. 易见$V_1 \subset V_2 \subset \cdots$. 令$V_\infty = \mathrm{Ker}\mathbb{E}(\cdot|\mathcal{X}_\infty)$, 下证明$\bigcup_{n=1}^{\infty} V_n$ 在V_∞ 中稠密, 进而

$$V = L^1(X, \mathcal{X}_\infty, \mu) + \bigcup_{n=1}^{\infty} V_n$$

在$L^1(X, \mathcal{X}, \mu)$ 中稠密. 为了证明$\bigcup_{n=1}^{\infty} V_n$ 在V_∞ 中稠密, 我们证明在$\bigcup_{n=1}^{\infty} V_n$ 上消失的泛函J 也在V_∞ 上消失即可.

设$J : L^1(X, \mu) \to \mathbb{C}$ 为$\bigcup_{n=1}^{\infty} V_n$ 上消失的泛函. 因为$(L^1)^* = L^\infty$, 存在$h \in L^\infty(X, \mathcal{X}, \mu)$ 使得

$$J : L^1(X, \mu) \to \mathbb{C},\ f \mapsto \int_X fh d\mu.$$

因为J 在$\bigcup_{n=1}^{\infty} V_n$ 上消失, 所以对于任何$n \in \mathbb{N}$

$$\int_X (f - \mathbb{E}(f|\mathcal{X}_n))hd\mu = 0, \quad \forall f \in L^1(X, \mathcal{X}, \mu.)$$

尤其

$$\int_X (h - \mathbb{E}(h|\mathcal{X}_n))hd\mu = 0.$$

由此易得

$$\int_X (h - \mathbb{E}(h|\mathcal{X}_n))\mathbb{E}(h|\mathcal{X}_n)d\mu = 0.$$

从而

$$\int_X (h - \mathbb{E}(h|\mathcal{X}_n))^2 d\mu = 0.$$

即$h = \mathbb{E}(h|\mathcal{X}_n) \in L^\infty(X, \mathcal{X}_n, \mu)$. 因为$n$ 任意, 所以$h \in L^\infty(X, \mathcal{X}_\infty, \mu)$. 于是$\mathbb{E}(f|\mathcal{X}_\infty) = 0$ 蕴含

$$J(f) = \int_X fh d\mu = \int_X \mathbb{E}(fh|\mathcal{X}_\infty) d\mu = \int_X h\mathbb{E}(f|\mathcal{X}_\infty) d\mu = 0.$$

亦即我们证明了在$\bigcup_{n=1}^\infty V_n$ 上消失的泛函J 也在V_∞ 上消失, 即$\bigcup_{n=1}^\infty V_n$ 在V_∞ 中稠密.

注意需要证明的命题对于V 上函数成立. 下面用V 上函数逼近一般的函数. 设$f \in L^1(X, \mathcal{X}, \mu)$. 对于任何取定$\varepsilon > 0$, 存在$k \in \mathbb{N}$ 以及$g \in V$ 使得$\|f - g\|_1 < \varepsilon$. 于是对于任何$n \geq k$, $\mathbb{E}(g|\mathcal{X}_n) = g$, 进而

$$\begin{aligned}
&\|\mathbb{E}(f|\mathcal{X}_n) - \mathbb{E}(f|\mathcal{X}_\infty)\|_1 \\
\leq &\|\mathbb{E}(f|\mathcal{X}_n) - \mathbb{E}(g|\mathcal{X}_n)\|_1 + \|\mathbb{E}(g|\mathcal{X}_n) - \mathbb{E}(g|\mathcal{X}_\infty)\|_1 + \|\mathbb{E}(g|\mathcal{X}_\infty) - \mathbb{E}(f|\mathcal{X}_\infty)\|_1 \\
\leq &2\|f - g\|_1 + \|\mathbb{E}(g|\mathcal{X}_n) - \mathbb{E}(g|\mathcal{X}_\infty)\|_1.
\end{aligned}$$

于是

$$\limsup_{n\to\infty} \|\mathbb{E}(f|\mathcal{X}_n) - \mathbb{E}(f|\mathcal{X}_\infty)\|_1 \leq 2\|f - g\|_1 \leq 2\varepsilon.$$

因为ε 任意, 我们得到$\lim_{n\to\infty} \|\mathbb{E}(f|\mathcal{X}_n) - \mathbb{E}(f|\mathcal{X}_\infty)\|_1 = 0$.

下面证明逐点收敛. 首先我们有:

$$\begin{aligned}
&\mu(\{x \in X : \limsup_{n\to\infty} |\mathbb{E}(f|\mathcal{X}_n)(x) - \mathbb{E}(f|\mathcal{X}_\infty)(x)| > \varepsilon^{1/2}\}) \\
\leq &\mu\left(\{x \in X : \limsup_{n\to\infty} \left(|\mathbb{E}((f-g)|\mathcal{X}_n)(x) - \mathbb{E}((f-g)|\mathcal{X}_\infty)(x)| + |\mathbb{E}(g|\mathcal{X}_n)(x) - \mathbb{E}(g|\mathcal{X}_\infty)(x)|\right) > \varepsilon^{1/2}\}\right) \\
\leq &\mu\left(\{x \in X : \limsup_{n\to\infty} \left(|\mathbb{E}((f-g)|\mathcal{X}_n)(x) - \mathbb{E}((f-g)|\mathcal{X}_\infty)(x)|\right) > \varepsilon^{1/2}\}\right) \\
\leq &\mu\left(\{x \in X : \limsup_{n\to\infty} |\mathbb{E}((f-g)|\mathcal{X}_n)(x)| > \frac{1}{2}\varepsilon^{1/2}\}\right) + \mu\left(\{x \in X : |\mathbb{E}((f-g)|\mathcal{X}_\infty)(x)| > \frac{1}{2}\varepsilon^{1/2}\}\right) \\
\leq &2\frac{1}{\frac{1}{2}\varepsilon^{1/2}}\|f - g\|_1 \leq 4\varepsilon^{1/2},
\end{aligned}$$

其中最后一个不等式用到了证明开始时建立的不等式和Chebyshev 不等式. 因为ε 任意, 所以

$$\mathbb{E}(f|\mathcal{X}_n)(x) \xrightarrow{a.e.} \mathbb{E}(f|\mathcal{X}_\infty)(x),\ n \to \infty.$$

证毕! □

定理 7.2.4 (钟开莱定理) 设$(X,\mathcal{X},\mu)$ 为概率空间, 设α 为可数剖分且$H(\alpha)<\infty$. 设$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为递增子σ 代数列, 那么

$$\int_X \sup_{n\geq 1} I(\alpha|\mathcal{X}_n)d\mu \leq H(\alpha)+1. \tag{7.2.1}$$

尤其, $f(x)=\sup_{n\geq 1} I(\alpha|\mathcal{X}_n)(x)\in L^1(X,\mathcal{X},\mu)$.

如$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为递减子σ 代数列, 结果仍成立.

证明. 设α 为可数剖分且$H(\alpha)<\infty$. 设$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为递增子σ 代数列, 令

$$f(x)=\sup_{n\geq 1} I(\alpha|\mathcal{X}_n)(x)\in L^1(X,\mathcal{X},\mu).$$

我们下面估计$C(t)=\{x\in X: f(x)>t\}, t\in\mathbb{R}$, 的测度.

取$A\in\alpha$, 设$g_n=\mu(A|\mathcal{X}_n)$. 于是对于$x\in A$, 我们有$I(\alpha|\mathcal{X}_n)(x)=-\log g_n(x)$. 根据定义, 由$f(x)>t$ 可以推出存在$n\in\mathbb{N}$ 使得$g_n(x)<e^{-t}$. 令

$$C_n=\{x\in X: g_k(x)\geq e^{-t},\ k<n;\ g_n(x)<e^{-t}\}.$$

则$C_n\in\mathcal{X}_n$, $\{C_n\}_{n=1}^{\infty}$ 互不相交, 并且$C\cap A\subset\bigcup_{n=1}^{\infty}C_n$. 因为$C_n$ 为$\mathcal{X}_n$ 可测的,

$$\mu(A\cap C_n)=\int_{C_n}1_A d\mu=\int_{C_n}g_n d\mu<e^{-t}\mu(C_n).$$

对n 求和, 就得到

$$\mu(\{x\in A: f(x)>t\})\leq e^{-t}.$$

于是

$$\mu(C(t))=\mu(\{x\in X: f(x)>t\})\leq\sum_{A\in\alpha}\min\{\mu(A),e^{-t}\}.$$

设$F(t)=\mu(C(t))$, 则

$$\begin{aligned}\int_X f d\mu &= -\int_0^{\infty} t dF(t)=[-tF(t)]_0^{\infty}+\int_0^{\infty}F(t)dt\\ &=\int_0^{\infty}\mu(\{x\in X: f(x)>t\})dt\\ &\leq\int_0^{\infty}\sum_{A\in\alpha}\min\{\mu(A),e^{-t}\}dt\\ &=\sum_{A\in\alpha}\left(\int_0^{-\log\mu(A)}+\int_{-\log\mu(A)}^{\infty}e^{-t}\right)\\ &=H_\mu(\alpha)+\sum_{A\in\alpha}\mu(A)=H_\mu(\alpha)+1.\end{aligned}$$

下设$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为递减子σ 代数列. 对于每个n, 那么$\mathcal{X}_n\subset\mathcal{X}_{n-1}\subset\ldots\subset\mathcal{X}_1$ 为递增的. 设$f_n(x)=\sup_{1\leq j\leq n} I(\alpha|\mathcal{X}_j)$, 那么根据上面结论我们有

$$\int_X f_n d\mu\leq H_\mu(\alpha)+1.$$

于是$f_n \in L^1(\mu)$, 并且$f_n \nearrow f = \sup_{n\geq 1} I(\alpha|\mathcal{X}_n)$. 根据单调收敛定理,

$$\int_X f_n d\mu \to \int_X f d\mu \leq H_\mu(\alpha) + 1.$$

证毕! □

根据鞅定理和钟开莱定理我们马上就得到下面结论:

定理 7.2.5 设$(X, \mathcal{X}, \mu)$ 为概率空间, α 为X 的可数可测剖分, $H(\alpha) < \infty$.

1. 如果$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为$\mathcal{X}$ 递增的子σ 代数且满足$\mathcal{X}_n \nearrow \mathcal{X}_\infty$, 则

$$I(\alpha|\mathcal{X}_n) \to I(\alpha|\mathcal{X}_\infty),\ n \to \infty$$

在$L^1(\mu)$ 和μ-$a.e.$ 意义下同时成立. 进而$H_\mu(\alpha|\mathcal{X}_n) \searrow H_\mu(\alpha|\mathcal{X}_\infty)$;

2. 如果$\{\mathcal{X}_n\}_{n=1}^{\infty}$ 为$\mathcal{X}$ 递减的子σ 代数且满足$\mathcal{X}_n \searrow \mathcal{X}_\infty$, 则

$$I(\alpha|\mathcal{X}_n) \to I(\alpha|\mathcal{F}_\infty), n \to \infty$$

在$L^1(\mu)$ 和μ-$a.e.$ 意义下同时成立. 进而$H_\mu(\alpha|\mathcal{X}_n) \nearrow H_\mu(\alpha|\mathcal{X}_\infty)$.

推论 7.2.6 设$(X, \mathcal{X}, \mu)$ 为概率空间, ξ, η 为可数剖分, $H(\xi) < \infty, H(\eta) < \infty$, $\mathcal{A}$ 为子σ 代数, 那么

$$I(\xi \vee \eta|\mathcal{A}) = I(\xi|\mathcal{A}) + I(\eta|\widehat{\xi} \vee \mathcal{A}),$$

$$H_\mu(\xi \vee \eta|\mathcal{A}) = H_\mu(\xi|\mathcal{A}) + H_\mu(\eta|\widehat{\xi} \vee \mathcal{A}).$$

证明. 用有限递增子σ 代数列取逼近$\mathcal{A}$, 然后根据鞅定理证明结论. 请读者自己补充细节. □

命题 7.2.7 设$(X, \mathcal{X}, \mu, T)$为保测系统. 对X的有限可测剖分α,

$$h_\mu(T, \alpha) = H_\mu(\alpha|\alpha^-).$$

特别地$h_\mu(T, \alpha) \leq H_\mu(\alpha)$.

证明. 由于$\lim\limits_{n\to\infty} H_\mu(\alpha|\bigvee_{i=1}^{n-1} T^{-i}\alpha) = H_\mu(\alpha|\alpha^-)$ (参见定理7.2.5的(1)), 我们有

$$\begin{aligned}
h_\mu(T, \alpha) \quad &= \lim_{n\to\infty} \tfrac{1}{n} H_\mu(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-(n-1)}\alpha) \\
&= \lim_{n\to\infty} \tfrac{1}{n}(H_\mu(\alpha) + H_\mu(\alpha|T^{-1}\alpha) + \cdots + H_\mu(\alpha|\bigvee_{i=1}^{n-1} T^{-i}\alpha)) \\
&= \lim_{n\to\infty} H_\mu(\alpha|\bigvee_{i=1}^{n-1} T^{-i}\alpha) \\
&= H_\mu(\alpha|\alpha^-).
\end{aligned}$$

证毕! □

§7.2.2 Kolmogorov-Sinai 定理

我们需要一些更易于操作的方法去计算测度熵. 下面的定理7.2.8以及Kolmogorov-Sinai 定理给我们一种易于操作的计算测度熵的办法.

定理 7.2.8 (Abramov) 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 如果X 的有限可测剖分序列$\{\alpha_n\}_{n=1}^{\infty}$ 满足$\alpha_1\preceq\alpha_2\preceq\cdots$ 且$\hat{\alpha}_n\nearrow\mathcal{X}$, 则

$$h_\mu(T)=\lim_{n\to\infty}h_\mu(T,\alpha_n).$$

证明. 对X的每个有限可测剖分β和$n\in\mathbb{N}$,

$$\begin{aligned}
h_\mu(T,\beta)\le h_\mu(T,\beta\vee\alpha_n)&=\lim_{m\to\infty}\frac{1}{m}H_\mu(\bigvee_{i=0}^{m-1}T^{-i}(\beta\vee\alpha_n))\\
&=\lim_{m\to\infty}\frac{1}{m}(H_\mu(\bigvee_{i=0}^{m-1}T^{-i}\alpha_n)+H_\mu(\bigvee_{i=0}^{m-1}T^{-i}\beta|\bigvee_{i=0}^{m-1}T^{-i}\alpha_n))\\
&\le\lim_{m\to\infty}\frac{1}{m}(H_\mu(\bigvee_{i=0}^{m-1}T^{-i}\alpha_n)+\sum_{i=0}^{m-1}H_\mu(T^{-i}\beta|T^{-i}\alpha_n))\\
&=\lim_{m\to\infty}\frac{1}{m}(H_\mu(\bigvee_{i=0}^{m-1}T^{-i}\alpha_n)+mH_\mu(\beta|\alpha_n))\quad\text{(利用命题7.1.5 (7))}\\
&=h_\mu(T,\alpha_n)+H_\mu(\beta|\alpha_n),
\end{aligned}\tag{7.2.2}$$

在上面不等式中令$n\to\infty$ 并利用鞅定理我们得到

$$h_\mu(T,\beta)\le\lim_{n\to\infty}h_\mu(T,\alpha_n)+H_\mu(\beta|\mathcal{X})=\lim_{n\to\infty}h_\mu(T,\alpha_n).$$

最后由β的任意性, $h_\mu(T)=\lim_{n\to\infty}h_\mu(T,\alpha_n)$. □

定理 7.2.9 设$(X,\mathcal{X},\mu,T)$和$(Y,\mathcal{Y},\nu,S)$为两个保测系统.则

$$h_{\mu\times\nu}(T\times S)=h_\mu(T)+h_\nu(S).$$

证明. 分别取X 和Y 的递增有限可测剖分序列$\{\alpha_n\}_{n=1}^{\infty}$ 和$\{\beta_n\}_{n=1}^{\infty}$ 使得$\hat{\alpha}_n\nearrow\mathcal{X}$ 且$\hat{\beta}_n\nearrow\mathcal{Y}$. 则$\{\alpha_n\times\beta_n\}_{n=1}^{\infty}$ 为$X\times Y$ 递增的有限可测剖分序列且满足$\hat{\alpha}_n\times\hat{\beta}_n\nearrow\mathcal{X}\times\mathcal{Y}$. 利用定理7.2.8, 我们有

$$\begin{aligned}
h_{\mu\times\nu}(T\times S)&=\lim_{n\to\infty}h_{\mu\times\nu}(T\times S,\alpha_n\times\beta_n)\\
&=\lim_{n\to\infty}\lim_{m\to\infty}\frac{1}{m}H_{\mu\times\nu}(\bigvee_{i=0}^{m-1}(T\times S)^{-i}\alpha_n\times\beta_n)\\
&=\lim_{n\to\infty}\lim_{m\to\infty}\frac{1}{m}(H_\mu(\bigvee_{i=0}^{m-1}T^{-i}\alpha_n)+H_\nu(\bigvee_{i=0}^{m-1}S^{-i}\beta_n))\\
&=\lim_{n\to\infty}(h_\mu(T,\alpha_n)+h_\nu(S,\beta_n))\\
&=h_\mu(T)+h_\nu(S).
\end{aligned}$$

证毕! □

定理 7.2.10 (Kolmogorov-Sinai定理) 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 如果有限可测剖分α 满足$\bigvee_{i=0}^{\infty} T^{-i}\alpha=\mathcal{X}$ (如此的α 称为**生成子**), [2] 则$h_\mu(T)=h_\mu(T,\alpha)$.

证明. 首先对$k\in\mathbb{N}$ 我们有

$$\begin{aligned} h_\mu(T,\alpha_0^{k-1}) &= \lim_{n\to+\infty}\frac{1}{n}H_\mu((\alpha_0^{k-1})_0^{n-1}) = \lim_{n\to+\infty}\frac{1}{n}H_\mu(\alpha_0^{k+n-1}) \\ &= \lim_{n\to+\infty}\frac{k+n}{n}\frac{1}{k+n}H_\mu(\alpha_0^{k+n-1}) = h_\mu(T,\alpha). \end{aligned}$$

类似于不等式(7.2.2), 对X 的每个有限可测剖分β 和$n\in\mathbb{N}$ 我们有

$$h_\mu(T,\beta)\le h_\mu(T,\alpha_0^{n-1})+H_\mu(\beta|\alpha_0^{n-1})=h_\mu(T,\alpha)+H_\mu(\beta|\alpha_0^{n-1}).$$

在上式中令$n\to+\infty$, 利用鞅定理我们有

$$h_\mu(T,\beta)\le h_\mu(T,\alpha)+H_\mu(\beta|\mathcal{X})=h_\mu(T,\alpha).$$

最后, 因β 是任意的, $h_\mu(T)=h_\mu(T,\alpha)$. □

注记 7.2.11 当保测系统$(X,\mathcal{X},\mu,T)$ 可逆时, 我们把满足$\bigvee_{i=-\infty}^{+\infty}T^{-i}\alpha=\mathcal{X}$ 的有限可测剖分α 称为**生成子**. 此时, 我们同样可以证明: 如果有限可测剖分α 为可逆保测系统$(X,\mathcal{X},\mu,T)$ 的生成子,则$h_\mu(T)=h_\mu(T,\alpha)$. 我们容易看出具有生成子的保测系统其测度熵一定有限, 与之对应的是, *1970* 年*W. Krieger [137]* 证明了每个具有有限熵的遍历保测系统都存在生成子. 因为这个定理的证明比较困难, 我们在本书中不再介绍, 例如可参见*[76]*.

定理 7.2.12 (Krieger生成子定理) 设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 如果$h_\mu(T)<\infty$, 那么T 有有限的生成子$\xi=\{A_1,A_2,\dots,A_n\}$, 其中$e^{h_\mu(T)}\le n\le e^{h_\mu(T)+1}$.

§7.2.3 一些例子

例 7.2.13 恒同映射$\mathrm{Id}:(X,\mathcal{X},\mu)\to(X,\mathcal{X},\mu)$ 的熵为零.

例 7.2.14 设$T:\mathbb{T}\to\mathbb{T}, x\mapsto 2x \pmod 1$, $m_{\mathbb{T}}$ 为$\mathbb{T}$ 上的*Haar* 测度, 那么$h_{m_{\mathbb{T}}}(T)=\log 2$.

证明. 设$\alpha=\{[0,\frac{1}{2}),[\frac{1}{2},1)\}$. 则

$$\alpha_0^{n-1}=\left\{\left[\frac{i}{2^n},\frac{i+1}{2^n}\right):i=0,1,\dots,2^n-1\right\}.$$

[2]在可扩同胚的研究中我们也有生成子的概念, 请读者自己对比二者.

于是α 为生成子, 并且

$$\begin{aligned} H_{m_{\mathbb{T}}}(\vee_{i=0}^{n-1} T^{-i}\alpha) &= -\sum_{i=0}^{2^n-1} \mu\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) \log\left(\mu\Big(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\Big)\right) \\ &= -\sum_{i=0}^{2^n-1} \frac{1}{2^n} \log(\frac{1}{2^n}) \\ &= n\log 2. \end{aligned}$$

于是

$$h_{m_{\mathbb{T}}}(T) = h_{m_{\mathbb{T}}}(T,\alpha) = \lim_{n\to\infty} \frac{1}{n} H_{m_{\mathbb{T}}}(\alpha_0^{n-1}) = \log 2.$$

□

例 7.2.15 设$T : \mathbb{T} \to \mathbb{T}, x \mapsto x+\alpha \pmod 1$ 上的旋转, $m_{\mathbb{T}}$ 为$\mathbb{T}$ 上的*Haar* 测度. 则$h_{m_{\mathbb{T}}}(T) = 0$.

证明. 情况1：$\alpha = \frac{p}{q}$ 为有理数. 此时$T^q = \mathrm{Id}$. 所以$h_{m_{\mathbb{T}}}(T) = \frac{1}{q} h_{m_{\mathbb{T}}}(\mathrm{Id}) = 0$.

情况2：α 为无理数. 设$\xi = \{[0,\frac{1}{2}), [\frac{1}{2}, 1)\}$. 因为$\{\frac{1}{2} + n\alpha \pmod 1\}_{n=0}^{\infty}$ 在$[0,1)$ 中稠密, 所以ξ 为生成子, 并且$\mathcal{X} = \xi_1^{\infty}$. 所以

$$h_{m_{\mathbb{T}}}(T) = h_{m_{\mathbb{T}}}(T,\xi) = \lim_{n\to\infty} H_{m_{\mathbb{T}}}(\xi|\xi_1^{\infty}) = H_{m_{\mathbb{T}}}(\xi|\mathcal{X}) = 0.$$

无论哪种情况，我们均得到$h_{m_{\mathbb{T}}}(T) = 0$. □

例 7.2.16 设$k \geq 2$ 和$\vec{p} = (p_0, p_1, \ldots, p_{k-1})$ 是一个概率向量. 设$(\Sigma_k, \mathcal{B}, \mu_{\vec{p}}, \sigma)$为双边边$\vec{p}$转移系统. 则

$$h_\mu(\sigma) = -\sum_{j=0}^{k-1} p_j \log p_j.$$

证明. 设$\alpha = \{{}_0[j]_0 : 0 \leq j \leq k-1\}$. 则$\alpha$ 为(Σ_k, μ, σ) 的有限可测剖分且

$$\bigvee_{i=-\infty}^{\infty} T^{-i}\alpha = \mathcal{X}(\Sigma_k).$$

由于$\alpha_0^{(n-1)} \perp_\mu T^{-n}\alpha$ 和对任意$n \in \mathbb{N}$ 成立, 利用命题7.1.12我们有

$$h_\mu(T,\alpha) = \lim_{n\to+\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = H_\mu(\alpha) = -\sum_{j=0}^{k-1} p_j \log p_j.$$

最后, 由注记7.2.11,

$$h_\mu(\sigma) = h_\mu(\sigma,\alpha) = -\sum_{j=0}^{k-1} p_j \log p_j.$$

□

例 7.2.17 设$I=[0,1]$ 为区间, $X=I^{\mathbb{Z}}$, 其上μ 为乘积测度, $T:X\to X$ 为转移映射. 那么$h_\mu(T)=\infty$.

证明. 对$n>0, 1\le i\le n$, 设

$$A_{n,i}=\{x=(x_j)_{j\in\mathbb{Z}}\in X:\frac{i-1}{n}<x_0<\frac{i}{n}\}.$$

则$\xi_n=\{A_{n,1},\dots,A_{n,n}\}$ 为X 剖分, 且$\mu(A_{n,j})=\frac{1}{n}, 1\le i\le n$. 类似于上面例子的计算, $h_\mu(T,\xi_n)=\log n$. 于是

$$h_\mu(T)\ge h_\mu(T,\xi_n)\ge\log n,\ \forall n\in\mathbb{N}.$$

即$h_\mu(T)=\infty$. □

例 7.2.18 (Markov转移) 双(单)边$(\overrightarrow{p},P)$ *Markov* 转移的熵为

$$h_\mu(\sigma)=-\sum_{i,j=0}^{k-1}p_ip_{ij}\log p_{ij}.$$

证明. 设$\alpha=\{A_j:0\le j\le k-1\}$, 其中$A_j={}_0[j]_0$. 则$\alpha$ 为(Σ_k,μ,σ) 的有限可测剖分且为生成子. 于是α_0^{n-1} 中元形如

$$A_{i_0}\cap T^{-1}A_{i_1}\cap\ldots\cap T^{-(n-1)}A_{i_{n-1}}=[i_0i_1\ldots i_{n-1}].$$

其测度为$p_{i_0}p_{i_0i_1}\ldots p_{i_{n-2}i_{n-1}}$. 于是

$$\begin{aligned}H_m(\vee_{i=0}^{n-1}T^{-i}\alpha)&=-\sum_{i_0,\dots,i_{n-1}=0}^{k-1}p_{i_0}p_{i_0i_1}\ldots p_{i_{n-2}i_{n-1}}\log(p_{i_0}p_{i_0i_1}\ldots p_{i_{n-2}i_{n-1}})\\&=-\sum_{i_0,\dots,i_{n-1}=0}^{k-1}p_{i_0}p_{i_0i_1}\ldots p_{i_{n-2}i_{n-1}}(\log p_{i_0}+\log p_{i_0i_1}+\ldots+\log p_{i_{n-2}i_{n-1}})\\&=-\sum_{i_0=0}^{k-1}p_{i_0}\log p_{i_0}-(n-1)\sum_{i,j=0}^{k-1}p_ip_{ij}\log p_{ij}.\end{aligned}$$

此处我们用了$\sum_{i=0}^{k-1}p_ip_{ij}=p_j$ 以及$\sum_{j=0}^{k-1}p_{ij}=1$.

于是$h_\mu(\sigma)=-\sum_{i,j=0}^{k-1}p_ip_{ij}\log p_{ij}$. □

习 题

1. 证明公式(7.1.2).
2. 设$(X,\mathcal{X},\mu,T)$为保测系统α, α为X的有限可测剖分.证明:$\frac{1}{n}H_\mu(\alpha_0^n)$关于$n$单调递减. 因此$h_\mu(T,\alpha)=\lim\limits_{n\to\infty}\frac{1}{n}H_\mu(\alpha_0^n)$.
3. 设$(X,\mathcal{X},\mu,T)$为保测系统.如果α为X的有限可测剖分, 则$h_\mu(T,\alpha)\le|\alpha|_\mu$, 其中$|\alpha|_\mu=\mathrm{Card}\{A\in\alpha:\mu(A)>0\}$.
4. 设$(X,\mathcal{X},\mu,T)$为保测系统.如果存在有限可测剖分α 满足$\alpha^-=\mathcal{X}$ (如此的α称为**强生成子**), 则$h_\mu(T)=0$.

§7.3　Shannon-McMillan-Breiman定理

Shannon-McMillan-Breiman定理是熵理论的一个重要定理，例如根据它也可定义熵.

引理 7.3.1 (Breiman引理) 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $g_n,g\in L^1(\mu),\forall n\in\mathbb{N}$ 且$g_n\to g,n\to\infty$ *a.e.* 如果

$$\int_X \sup_{n\in\mathbb{N}}|g_n|d\mu<\infty,$$

那么

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}g_n(T^nx)=\mathbb{E}(g|\mathcal{I}),$$

其中收敛为*a.e.*, 以及$L^1(\mu)$ 意义下.

证明. 由遍历定理, 我们有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}g(T^jx)=\mathbb{E}(g|\mathcal{I}),$$

其中收敛为a.e., 以及$L^1(\mu)$ 意义下. 因为

$$\frac{1}{n}\sum_{j=0}^{n-1}g_j(T^jx)=\frac{1}{n}\sum_{j=0}^{n-1}g(T^jx)+\frac{1}{n}\sum_{j=0}^{n-1}(g_j(T^jx)-g(T^jx)),$$

所以我们仅需证明

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}|g_j(T^jx)-g(T^jx)|=0.$$

设

$$G_n=\sup_{j\natural\geq n}|g_j(x)-g(x)|\leq\sup_{j\geq n}(|g_j(x)|+|g(x)|).$$

根据Lebesgue 控制收敛定理, 易有$\lim_{n\to\infty}\int_X G_nd\mu=0$, 尤其$\lim_{n\to\infty}\mathbb{E}(G_n|\mathcal{I})=0$. 取$N,n\in\mathbb{N}$ 使得$n>N+1$. 注意我们有不等式

$$\begin{aligned}&\frac{1}{n}\sum_{j=0}^{n-1}|g_j(T^jx)-g(T^jx)|\\ \leq&\frac{1}{n}\sum_{j=0}^{N-1}|g_j(T^jx)-g(T^jx)|+\frac{n-N-1}{n}\left(\frac{1}{n-N-1}\sum_{j=0}^{n-N-1}G_N(T^j(T^Nx))\right).\end{aligned}$$

于是

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}|g_j(T^jx)-g(T^jx)|\leq 0+\mathbb{E}(G_N|\mathcal{I}).$$

其中收敛为a.e., 以及$L^1(\mu)$ 意义下. 再由$\lim_{N\to\infty}\mathbb{E}(G_N|\mathcal{I})=0$ 我们就得到需要的结论. □

注记 7.3.2 根据同样的证明, 我们有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}g_{N-n-1}(T^nx)=\mathbb{E}(g|\mathcal{I}).$$

定理 7.3.3 (Shannon-McMillan-Breiman定理) *设$(X,\mathcal{X},\mu,T)$ 为保测系统, α 为可数剖分且$H(\alpha)<\infty$. 设$g(x)=I(\alpha|\vee_{n=1}^{\infty}T^{-n}\hat{\alpha})=I(\alpha|\hat{\alpha}^-)$ 那么*

$$\lim_{N\to\infty}\frac{1}{N}I(\alpha_0^{N-1})=\mathbb{E}(g|\mathcal{I}),$$

其中收敛为$a.e.$, 以及$L^1(\mu)$ 意义下, 并且

$$\lim_{N\to\infty}\frac{1}{N}H(\alpha_0^{N-1})=H(\alpha|\hat{\alpha}^-)=h_\mu(\alpha,T).$$

当T 为遍历的时候, 我们有

$$\lim_{N\to\infty}\frac{1}{N}I(\alpha_0^{N-1})=H(\alpha|\hat{\alpha}^-)=h_\mu(\alpha,T)$$

其中收敛为$a.e.$, 以及$L^1(\mu)$ 意义下.

证明. 对$k\in\mathbb{N}$, 设$g_k(x)=I(\alpha|\hat{\alpha}_1^k)(x)$, $g_0(x)=I(\alpha)(x)$. 则

$$\begin{aligned}I(\alpha_0^{N-1})&=I(\alpha\vee T^{-1}\alpha\vee\ldots\vee T^{N-1}\alpha)\\&=I(\alpha|\vee_{n=1}^{N-1}T^{-n}\hat{\alpha})+I(\vee_{n=0}^{N-2}T^{-n}\alpha)\circ T\\&=g_{N-1}+g_{N-2}\circ T+I(\vee_{n=0}^{N-3}T^{-n}\alpha)\circ T^2\\&=g_{N-1}+g_{N-2}\circ T+\ldots+g_0\circ T^{N-1}=\sum_{n=0}^{N-1}g_{N-n-1}(T^nx).\end{aligned}$$

由鞅定理和钟开莱定理, $g_k(x)=I(\alpha|\hat{\alpha}_1^k)(x)\to I(\alpha|\hat{\alpha}^-)(x)=g(x),k\to\infty$, 且

$$\int_X\sup g_k=\int_X g<H_\mu(\alpha)+1<\infty.$$

根据Breiman引理以及注记, 我们有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}g_n(T^nx)=\mathbb{E}(g|\mathcal{I}),$$

其中收敛为a.e., 以及$L^1(\mu)$ 意义下. 证明剩余部分是简单的. □

推论 7.3.4 *设$(X,\mathcal{X},\mu,T)$ 为遍历系统, α 为可数剖分且$H(\alpha)<\infty$. 那么对于任何$\varepsilon>0,\delta>0$, 存在$N=N(\varepsilon.\delta)\in\mathbb{N}$ 使得当$n\geq N$ 时, 我们有*

1. α_0^{n-1} 中元素分成两类$\mathcal{F},\mathcal{G}$ 使得

(a) $\mu(\bigcup_{B\in\mathcal{F}})B<\delta$,

(b) $e^{-n(h_\mu(\alpha)+\varepsilon)}\le\mu(A)\le e^{-n(h_\mu(\alpha)-\varepsilon)},\forall A\in\mathcal{G}$

2. 存在α_0^{n-1}-可测集$E=E(\varepsilon)$, $\mu(E)<\delta$, 并且包含在$X\setminus E$ 中α_0^{n-1} 的元素总数M满足

$$e^{n(h_\mu(\alpha)-\varepsilon)}\le M\le e^{n(h_\mu(\alpha)+\varepsilon)}.$$

习　题

1. 完成上面推论7.3.4的证明.
2. 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, α 为可数剖分且$H(\alpha)<\infty$. 对于$x\in X$, 设$\alpha_n(x)$ 为剖分$\bigvee_{i=0}^{n-1}T^{-i}\alpha$ 中包含x 的元素. 设ν 为$(X,\mathcal{X})$ 上另外一个概率测度，且存在$0<C_1<C_2$ 使得$C_1\nu(A)<\mu(A)<C_2\nu(A),\forall A\in\mathcal{X}$. 证明：对于$\nu$ 几乎处处的$x\in\mathcal{X}$, 我们有
$$\lim_{n\to\infty}-\frac{\log\nu(\alpha_n(x))}{n}=h_\mu(\alpha,T).$$
3. 设$([0,1],\mathcal{B}([0,1]),\mu,T)$ 为连分数系统，其中μ 为Gauss 测度. 证明：
$$h_\mu(T)=\frac{\pi^2}{6\log 2}.$$

§7.4　拓扑熵

在本小节, 我们将介绍拓扑动力系统中一个重要的不变量—-拓扑熵. 它是由Adler, Konheim 和McAndrew [2] 在1965年首先引入的, 后来Dinabury [?] 和Bowen [36] 使用分离集和张成集给出了一个新的等价定义. 拓扑熵在拓扑动力系统共轭分类中扮演了相当重要的角色, 反映了拓扑动力系统的复杂性程度.

§7.4.1　开覆盖的定义方式

设X为一非空集合, 我们通常用花写字母$\mathcal{U}$、$\mathcal{V}$ 等来表示X 的覆盖. 设$\mathcal{U}$、$\mathcal{V}$ 为X 的两个覆盖, $\mathcal{U}$和$\mathcal{V}$的**交** $\mathcal{U}\vee\mathcal{V}$定义为
$$\mathcal{U}\vee\mathcal{V}=\{U\cap V:U\in\mathcal{U},V\in\mathcal{V}\},$$
它仍为X 的覆盖. 类似地我们可以定义有限个覆盖$\mathcal{U}_1,\mathcal{U}_2,\cdots,\mathcal{U}_n$ 的**交** $\mathcal{U}_1\vee\cdots\vee\mathcal{U}_n$. 设$T:X\to X$为映射, $\mathcal{U}$ 为X 的覆盖, 我们定义
$$T^{-1}\mathcal{U}=\{T^{-1}U:U\in\mathcal{U}\}.$$
显然$T^{-1}\mathcal{U}$ 仍为X 的覆盖. 一般来说, 对两个非负整数m, n $(n\ge m)$ 我们可以定义
$$\mathcal{U}_m^n=\bigvee_{j=m}^n T^{-j}\mathcal{U},$$

特别地对$n\geq 1$
$$\mathcal{U}_0^{n-1}=\mathcal{U}\vee T^{-1}\mathcal{U}\vee\cdots\vee T^{-(n-1)}\mathcal{U}.$$

设$\mathcal{U},\mathcal{V}$为X的两个覆盖, 如果对每个$V\in\mathcal{V}$, 我们能找到某个$U\in\mathcal{U}$使得$V\subset U$, 则称$\mathcal{V}$为$\mathcal{U}$的**加细**, 记为$\mathcal{V}\succeq\mathcal{U}$或$\mathcal{U}\preceq\mathcal{V}$. 特别地, 如果$\mathcal{V}$的每个元素均为$\mathcal{U}$中的元素, 则称$\mathcal{V}$为$\mathcal{U}$的**子覆盖**, 显然此时$\mathcal{V}\succeq\mathcal{U}$.

定义 7.4.1 *如果$\mathcal{U}$为非空集合X的覆盖, 我们用$N(\mathcal{U})$表示$\mathcal{U}$的所有子覆盖中具有最少元素个数的子覆盖的元素个数. 当$\mathcal{U}$没有有限子覆盖时, 我们约定$N(\mathcal{U})=+\infty$.*

易见当X为紧度量空间以及$\mathcal{U}$为X的开覆盖时, $N(\mathcal{U})$为有限数. 为方便起见, 对覆盖$\mathcal{U}$我们记
$$H(\mathcal{U})=\log N(\mathcal{U}).$$

性质 7.4.2 *设X为一非空集合, $\mathcal{U},\mathcal{V}$为X的两个覆盖. 则*

1. $H(\mathcal{U})\geq 0$.
2. *如果$\mathcal{V}\succeq\mathcal{U}$, 则$H(\mathcal{V})\geq H(\mathcal{U})$.*
3. $H(\mathcal{U}\vee\mathcal{V})\leq H(\mathcal{U})+H(\mathcal{V})$.
4. *对任意映射$T:X\to X$, 我们有$H(\mathcal{U})\geq H(T^{-1}\mathcal{U})$; 进而当$T$为满射时, $H(\mathcal{U})=H(T^{-1}\mathcal{U})$.*

证明. 由定义(1)、(2)和(4)是明显成立的. 对于(3), 当$H(\mathcal{U})=+\infty$或$H(\mathcal{V})=+\infty$时, 它是明显成立地. 现假设$\{U_1,\cdots,U_{N(\mathcal{U})}\}$和$\{V_1,\cdots,V_{N(\mathcal{V})}\}$分别为$\mathcal{U}$和$\mathcal{V}$具有最少元素个数的子覆盖. 则
$$\{U_i\cap V_j:1\leq i\leq N(\mathcal{U}),1\leq j\leq N(\mathcal{V})\}$$
为$\mathcal{U}\cap\mathcal{V}$的子覆盖, 从而
$$N(\mathcal{U}\vee\mathcal{V})\leq N(\mathcal{U})N(\mathcal{V}).$$

□

命题 7.4.3 *设X为一非空集合, $T:X\to X$为映射, $\mathcal{U}$为X的覆盖. 则非负序列$a_n=H(\mathcal{U}_0^{n-1})$具有次可加性, 即$a_{m+n}\leq a_m+a_n,\forall m,n\geq 1$. 进而极限*
$$\lim_{n\to\infty}\frac{1}{n}H(\mathcal{U}_0^{n-1})=\inf_{n\geq 1}\frac{1}{n}H(\mathcal{U}_0^{n-1}),$$
*我们将该极限称为$\mathcal{U}$相对于T的**组合熵**, 记为$h_c(T,\mathcal{U})$.*

*当(X,T)为动力系统以及$\mathcal{U}$为X的开覆盖时, 以上方式定义的$h_c(T,\mathcal{U})$称为$\mathcal{U}$相对于T的**拓扑熵**, 记为$h_{\rm top}(T,\mathcal{U})$.*

证明. 对$m,n\geq 1$, 由性质7.4.2的(3)和(4)我们有

$$\begin{aligned} a_{m+n} &= H(\bigvee_{i=0}^{m+n-1} T^{-i}\mathcal{U}) \\ &\leq H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}) + H(T^{-n}(\bigvee_{j=0}^{m-1} T^{-j}\mathcal{U})) \\ &\leq a_n + a_m \end{aligned}$$

这说明$\{a_n\}$ 为次可加非负序列. 这就完成了引理的证明. □

从命题7.4.2和7.4.3, 我们不难得到

性质 7.4.4 *设X 为一非空集合和$T:X\to X$为映射. 则对X 的任意两个覆盖$\mathcal{U},\mathcal{V}$, 我们有*

1. $H(\mathcal{U})\geq h_c(T,\mathcal{U})\geq 0$.
2. 如果$\mathcal{V}\succeq\mathcal{U}$, 则$h_c(T,\mathcal{V})\geq h_c(T,\mathcal{U})$.
3. $h_c(T,\mathcal{U}\vee\mathcal{V})\leq h_c(T,\mathcal{U})+h_c(T,\mathcal{V})$.
4. $h_c(T,\mathcal{U})\geq h_c(T,T^{-1}\mathcal{U})$; 进而当$T$为满射时, $h_c(T,\mathcal{U})=h_c(T,T^{-1}\mathcal{U})$.

定义 7.4.5 *对动力系统(X,T), 我们用$\mathcal{C}_X^o$ 表示空间X 的全体开覆盖. 令*

$$h_{\rm top}(T)=\sup_{\mathcal{U}\in\mathcal{C}_X^o} h_{\rm top}(T,\mathcal{U}),$$

*我们把$h_{\rm top}(T)$ 称为系统(X,T) 的***拓扑熵***. 在必要时, 为强调空间X 我们也可将其记为$h_{\rm top}(T,X)$.*

需要提及的是在这里可能出现$h_{\rm top}(T)$为$+\infty$的情形.

命题 7.4.6 *设(X,T) 为拓扑动力系统.*

1. 如果(Y,T) 为(X,T) 的子系统, 则$h_{\rm top}(T,X)\geq h_{\rm top}(T,Y)$.
2. 如果$\pi:(X,T)\to(Y,S)$ 为因子映射, 则$h_{\rm top}(T)\geq h_{\rm top}(S)$.
3. 如果(X,T) 为可逆系统, 则$h_{\rm top}(T)=h_{\rm top}(T^{-1})$.

证明. 由定义7.4.5和性质7.4.2, (1)是明显成立地. (2)来自于以下事实: 对任意$\mathcal{U}\in\mathcal{C}_Y^o$, $h_{\rm top}(T,\pi^{-1}(\mathcal{U}))=h_{\rm top}(S,\mathcal{U})$. 最后, (2)蕴含着(3). □

注记 7.4.7 *由命题7.4.6的性质(2), 如果$\pi:(X,T)\to(Y,S)$ 为拓扑共轭,则$h_{\rm top}(T)=h_{\rm top}(S)$. 这说明***熵是系统的拓扑共轭不变量***.*

一般来说, 跟测度熵一样, 系统的拓扑熵是相当不容易计算的, 但对于一些特殊系统我们仍然有章可寻. 我们会在后面介绍一些基本的方法和例子.

例 7.4.8 设$T = \mathrm{Id} : X \to X$ 为恒同映射, 则对X 任意的开覆盖$\mathcal{U} = \{U_1, \ldots, U_t\}$ 我们有$T^{-i}\mathcal{U} = \mathcal{U}$ 对所有$i \in \mathbb{Z}_+$ 成立. 因此存在k 使得$\mathcal{U}_0^{n-1} = \mathcal{U}_0^{k-1}$, $n \geq k$. 这说明

$$H(\mathcal{U}_0^{n-1}) = H(\mathcal{U}_0^{k-1}), \quad n \geq k.$$

因此

$$h_{\mathrm{top}}(T, \mathcal{U}) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{U}_0^{n-1}) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{U}_0^{k-1}) = 0.$$

进而由$\mathcal{U}$ 的任意性, 我们得到$h_{\mathrm{top}}(T) = 0$.

§7.4.2 Bowen 的定义

以下我们介绍Bowen 关于拓扑熵的定义, 对相当多的系统而言Bowen 的定义使我们更容易计算其拓扑熵. 注意Bowen 熵可以对于非紧致空间类似定义, 我们先对紧致空间给出定义, 然后对于非紧空间上熵的定义给出简单的描述.

当(X, T) 为动力系统, d 为X 上与拓扑相容的度量以及$n \in \mathbb{N}$时, 对$x, x' \in X$ 我们定义

$$d_n(x, x') = \max_{0 \leq k \leq n-1} d(T^k x, T^k x').$$

容易证明d_n 也为X 上与拓扑相容的度量. 注意在d_n 下以x 为中心, r 为半径的球为

$$\bigcap_{i=0}^{n-1} T^{-i} B(T^i x, r).$$

定义 7.4.9 设(X, T)为具有度量d的动力系统.

1. 对$n \geq 1$ 和$\epsilon > 0$, 我们称有限集$A \subset X$为一个(n, ϵ)**分离集**(*(n, ε)-separated set*), 如果对A中任意两个不同的点x, y 均有$d_n(x, y) \geq \epsilon$成立; 我们用$\mathrm{sr}(n, \epsilon, T)$表示$(X, T)$具有最多元素个数的$(n, \epsilon)$分离集的元素个数;

2. 对$n \geq 1$和$\epsilon > 0$, 我们称有限集$A \subset X$为一个(n, ϵ)**张成集**(*(n, ε)-spanning set*), 如果对X的任意点x 存在$y \in A$ 使得$d_n(x, y) < \epsilon$; 我们用$\mathrm{sp}(n, \epsilon, T)$表示$(X, T)$具有最少元素个数的$(n, \epsilon)$张成集的元素个数.

由于空间X是紧的, $\mathrm{sr}(n, \epsilon, T)$和$\mathrm{sp}(n, \epsilon, T)$ 均为有限数. 现在我们定义

$$\mathrm{sr}(\epsilon, T) = \limsup_{n\to+\infty} \frac{1}{n} \log \mathrm{sr}(n, \epsilon, T),$$

$$\mathrm{sp}(\epsilon, T) = \limsup_{n\to+\infty} \frac{1}{n} \log \mathrm{sp}(n, \epsilon, T).$$

易见当$\epsilon \to 0+$时, $\mathrm{sr}(\epsilon, T)$和$\mathrm{sp}(\epsilon, T)$ 关于ϵ单调上升. 因此极限

$$\mathrm{sr}(d, T) = \lim_{\epsilon\to 0+} \mathrm{sr}(\epsilon, T) \ \text{ 和 } \ \mathrm{sp}(d, T) = \lim_{\epsilon\to 0+} \mathrm{sp}(\epsilon, T)\text{存在.}$$

引理 7.4.10 对每个$n \in \mathbb{N}$.

1. $\mathrm{sp}(n,\epsilon,T) \le \mathrm{sr}(n,\epsilon,T) \le \mathrm{sp}(n,\frac{\epsilon}{2},T)$;
2. $\mathrm{sp}(\epsilon,T) \le \mathrm{sr}(\epsilon,T) \le \mathrm{sp}(\frac{\epsilon}{2},T)$;
3. $\mathrm{sp}(d,T) = \mathrm{sr}(d,T)$.

证明. 由于(1) $\Rightarrow$ (2) $\Rightarrow$ (3)是明显成立的, 我们只需证明(1). 显然X的一个具有最多元素个数的(n,ϵ)分离集也为(n,ϵ) 张成集, 所以$\mathrm{sp}(n,\epsilon,T) \le \mathrm{sr}(n,\epsilon,T)$.相反地, 如果$A$为具有最少元素个数的$(n,\frac{\epsilon}{2})$ 张成集以及E为任意(n,ϵ)分离集, 则对每个$x \in E$存在点$\phi(x) \in A$使得$d_n(x,\phi(x)) < \frac{\epsilon}{2}$. 因$E$为任意$(n,\epsilon)$分离集, ϕ为一一对应.这说明$|A| \ge |E|$, 进而$\mathrm{sr}(n,\epsilon,T) \le \mathrm{sp}(n,\frac{\epsilon}{2},T)$. □

以下命题说明通过分离集和张成集可给出拓扑熵的等价定义.

命题 7.4.11 设(X,T) 为动力系统. 则

$$h_{\mathrm{top}}(T) = \mathrm{sp}(d,T) = \mathrm{sr}(d,T).$$

特别地, $\mathrm{sp}(d,T) = \mathrm{sr}(d,T)$ 不依赖于相容度量d 的选取.

证明. 首先我们固定一个$\epsilon > 0$ 和$n \in \mathbb{N}$. 设$\mathcal{U}$ 为具有Lebesque 数2ϵ 的开覆盖以及E 为具有最少元素个数的(n,ϵ) 张成集; 即$|E| = \mathrm{sp}(n,\epsilon,T)$. 从$\mathrm{sp}(n,\epsilon,T)$的定义知

$$\bigcup_{x\in E}\bigcap_{i=0}^{n-1} T^{-i}B_\epsilon(T^i x) = X.$$

现在对每个$x \in E$和$1 \le i \le n$, $B_\epsilon(T^i x)$ 包含在覆盖$\mathcal{U}$的某个元素中, 由此可见

$$N(\bigvee_{i=1}^{n} T^{-i}\mathcal{U}) \le \mathrm{sp}(n,\epsilon,T).$$

取$\mathcal{V}$为X的满足$\mathrm{diam}(\mathcal{V}) < \epsilon$的开覆盖.设$A$ 为具有最多元素个数的(n,ϵ)分离集, 即$|A| = \mathrm{sr}(n,\epsilon,T)$.注意到$\bigvee_{i=0}^{n-1} T^{-i}\mathcal{V}$每个元素中至多含有$A$中一个点, 由此可见

$$\mathrm{sr}(n,\epsilon,T) \le N(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{V}).$$

由以上分析

$$N(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}) \le \mathrm{sp}(n,\epsilon,T) \le \mathrm{sr}(n,\epsilon,T) \le N(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{V}),$$

因此
$$\frac{1}{n}N(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U}) \le \frac{1}{n}\mathrm{sp}(n,\epsilon,T) \le \frac{1}{n}\mathrm{sr}(n,\epsilon,T) \le \frac{1}{n}N(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{V}). \tag{7.4.1}$$
令$n\to\infty$，得到
$$h_{\mathrm{top}}(T,\mathcal{U}) \le \mathrm{sp}(\epsilon,T) \le \mathrm{sr}(\epsilon,T) \le h_{\mathrm{top}}(T,\mathcal{V}).$$
然后让$\epsilon\searrow 0$, 同时要求$\mathrm{diam}(\mathcal{U})\searrow 0$ (这样$\mathcal{V}$ 也满足$\mathrm{diam}(\mathcal{V})\searrow 0$), 利用引理7.5.1 可得
$$h_{\mathrm{top}}(T)=\mathrm{sp}(d,T)=\mathrm{sr}(d,T).$$
证毕! □

从公式(7.4.1)我们同时获得了如下的推论:

推论 7.4.12 设
$$\underline{\mathrm{sr}}(\epsilon,T)=\liminf_{n\to\infty}\frac{1}{n}\log \mathrm{sr}(n,\epsilon,T),$$
$$\underline{\mathrm{sp}}(\epsilon,T)=\liminf_{n\to\infty}\frac{1}{n}\log \mathrm{sp}(n,\epsilon,T).$$
则: 极限$\underline{\mathrm{sr}}(d,T)=\lim\limits_{\epsilon\to 0+}\underline{\mathrm{sr}}(\epsilon,T)$ 和$\underline{\mathrm{sp}}(d,T)=\lim\limits_{\epsilon\to 0+}\underline{\mathrm{sp}}(\epsilon,T)$ 存在, 且
$$h_{\mathrm{top}}(T)=\underline{\mathrm{sp}}(d,T)=\underline{\mathrm{sr}}(d,T).$$

例 7.4.13 设T 在(X,d) 上等距, 那么$d_n=d,\forall n\in\mathbb{N}$, 由此$h(T)=0$. 所以等度连续系统的熵为零.

下面我们讨论下非紧空间上的熵的定义. 下面设(X,d) 是一个度量空间, 用$UC(X,d)$ 表示(X,d) 上全体一致连续自映射的全体. 下面我们取定$T\in UC(X,d)$. 首先需要定义子集上的分离集与张成集.

定义 7.4.14 设K 为X 的紧致子集.

1. 对$n\ge 1$ 和$\epsilon>0$, 我们称子集$A\subset K$ 为一个K 上的(n,ϵ) **分离集**((n,ε)-separated set), 如果对A 中任意两个不同的点x,y 均有$d_n(x,y)\ge\epsilon$成立; 我们用$\mathrm{sr}(n,\epsilon,K,T)$ 表示K 上具有最多元素个数的(n,ϵ)分离集的元素个数;

2. 对$n\ge 1$ 和$\epsilon>0$, 我们称子集$A\subset X$ 为一个K 的(n,ϵ)**张成集**((n,ε)-spanning set), 如果对K 的任意点x 存在$y\in A$ 使得$d_n(x,y)<\epsilon$; 我们用$\mathrm{sp}(n,\epsilon,K,T)$ 表示(X,T) 具有最少元素个数的(n,ϵ) 张成集的元素个数.

设

$$\mathrm{sr}(\varepsilon,K,T)=\limsup_{n\to\infty}\frac{1}{n}\log \mathrm{sr}(n,\varepsilon,K,T),$$

$$\mathrm{sp}(\varepsilon,K,T)=\limsup_{n\to\infty}\frac{1}{n}\log \mathrm{sp}(n,\varepsilon,K,T).$$

如需要强调度量d, 就记上式为$\mathrm{sr}(\varepsilon,K,T,d)$ 和$\mathrm{sp}(\varepsilon,K,T,d)$. 接着我们可以定义拓扑熵为

$$h_{\mathrm{top}}(T)=\sup_K\lim_{\varepsilon\to 0}\mathrm{sp}(\varepsilon,K,T),$$

其中K 取遍X 的所有紧致子集. 容易验证下面性质：$\mathrm{sp}(n,\epsilon,K,T)\le \mathrm{sr}(n,\epsilon,K,T)\le \mathrm{sp}(n,\frac{\epsilon}{2},K,T)$, 从而我们有

$$h_{\mathrm{top}}(T)=\sup_K\lim_{\varepsilon\to 0}\mathrm{sr}(\varepsilon,K,T).$$

即用生成集和分离集定义熵的方式是等价的. 注意对于非紧的情况, 上面的定义会依赖于度量d 的选取. 如果X 上度量d 和d' 为一致等价的(即$\mathrm{Id}:(X,d)\to(X,d')$ 为一致同胚), 那么$T\in UC(X,d)$ 当且仅当$T\in UC(X,d')$.

命题 7.4.15 *设X 上度量d 和d' 为一致等价的, 那么按照d 和d' 定义的$Bowen$ 熵是相等的.*

注记 7.4.16 *如果X 上度量d 和d' 只是等价而非一致等价的, 那么熵算出来可能不一样. 例如设$X=(0,\infty),T:X\to X,T(x)=2x$. 如果$d$ 为欧式度量, 那么$T\in UC(X,d)$, 并且容易估计$h_d(T)\ge\log 2$. 设d' 定义如下：在区间$[1,2]$ 是欧式度量并且T 为d' 等距的(具体如下构造：$X=\bigcup_{n\in\mathbb{Z}}(2^{n-1},2^n]$ 为剖分, 因为$T((2^{n-1},2^n])=(2^n,2^{n+1}]$, 将$[1,2]$ 上度量诱导到每个区间$(2^{n-1},2^n]$ 上). 此时$h_{d'}(T)=0$, 但是d,d' 等价.*

习 题

1. 证明命题7.4.15.
2. 设(X,d) 为度量空间, $T\in UC(X,d)$, $K,K_1,\ldots,K_m$ 为紧致子集满足$K\subset K_1\cup\ldots\cup K_m$. 那么

$$h_{\mathrm{top}}(T;K)\le\max_{1\le i\le m}h_{\mathrm{top}}(T;K_i).$$

于是对于任何$\delta>0$, $h_{\mathrm{top}}(T)=\sup_K h_{\mathrm{top}}(T;K)$, 其中$K$ 取遍直径小于δ 的全体紧致子集.

3. 存在非紧度量空间$(X_1,d_1),(X_2,d_2)$, $T_1\in UC(X_1,d_1),T_2\in UC(X_2,d_2)$ 使得

$$h_{\mathrm{top}}(T_1\times T_2)<h_{\mathrm{top}}(T_1)+h_{\mathrm{top}}(T_2).$$

4. 存在非紧度量空间(X,d), $T\in UC(X,d)$ 使得

$$h_{\mathrm{top}}(T)\ne h_{\mathrm{top}}(T^{-1}).$$

§7.5 拓扑熵的计算

在本节我们计算一些常见系统的拓扑熵, 首先介绍几个常用的定理.

§7.5.1 一些有用的性质

下面的引理提供给我们一个相对来说更易于操作的计算熵的方法.

引理 7.5.1 设(X,T) 为具有度量d 的动力系统. 如果$\{\mathcal{U}_i\}_{i=1}^{\infty}$ 为X 满足

$$\mathrm{diam}(\mathcal{U}_n)=\sup\{\mathrm{diam}(U):U\in\mathcal{U}_n\}\to 0$$

的开覆盖序列,则

$$\lim_{n\to\infty}h_{\mathrm{top}}(T,\mathcal{U}_n)=h_{\mathrm{top}}(T).$$

证明. 设$\mathcal{V}$为X的任一开覆盖, δ 为$\mathcal{V}$ 的Lebesgue 数. 则对满足$\mathrm{diam}(\mathcal{U}_n)<\delta$ 的n 而言, 我们有$\mathcal{U}_n\succeq\mathcal{V}$, 进而

$$\liminf_{n\to+\infty}h_{\mathrm{top}}(T,\mathcal{U}_n)\geq h_{\mathrm{top}}(T,\mathcal{V}).$$

再由$\mathcal{V}$的任意性, 我们得到

$$\liminf_{n\to+\infty}h_{\mathrm{top}}(T,\mathcal{U}_n)\geq h_{\mathrm{top}}(T).$$

同时, 由于

$$\limsup_{n\to+\infty}h_{\mathrm{top}}(T,\mathcal{U}_n)\leq h_{\mathrm{top}}(T)$$

是明显成立地, 结论成立. □

引理 7.5.2 设(X,T)为动力系统, 对X的每个开覆盖$\mathcal{U}$和$n\in\mathbb{N}$, 我们有

$$h_{\mathrm{top}}(T,\mathcal{U}_0^n)=h_{\mathrm{top}}(T,\mathcal{U}).$$

证明. 注意到

$$\begin{aligned}h_{\mathrm{top}}(T,\mathcal{U})\leq h_{\mathrm{top}}(T,\mathcal{U}_0^n)&=\lim_{k\to+\infty}\frac{1}{k}H((\mathcal{U}_0^n)_0^{k-1})=\lim_{k\to+\infty}\frac{1}{k}H(\mathcal{U}_0^{n+k-1})\\&=\lim_{k\to+\infty}\frac{n+k}{k}\frac{1}{n+k}H(\mathcal{U}_0^{n+k-1})=h_{\mathrm{top}}(T,\mathcal{U}),\end{aligned}$$

引理得证. □

注记 7.5.3 当(X,T)为可逆动力系统时, 对X的每个开覆盖$\mathcal{U}$和$n\in\mathbb{N}$我们有

$$h_{\mathrm{top}}(T,\mathcal{U}_{-n}^n)=h_{\mathrm{top}}(T,\mathcal{U}).$$

命题 7.5.4 *(Abramov定理)* 设(X,T)为动力系统且$m\in\mathbb{N}$, 则

$$h_{\mathrm{top}}(T^m)=m\ h_{\mathrm{top}}(T).$$

证明. 首先对每个$\mathcal{U}\in\mathcal{C}_X^o$, 我们不难得到

$$h_{\rm top}(T^m,\mathcal{U}_0^{m-1})=\lim_{k\to\infty}\frac{1}{k}H(\bigvee_{i=0}^{k-1}T^{-im}\mathcal{U}_0^{m-1})=\lim_{k\to\infty}\frac{1}{k}H(\mathcal{U}_0^{km-1})=mh_{\rm top}(T,\mathcal{U}).$$

因此

$$h_{\rm top}(T^m)=\sup_{\mathcal{U}\in\mathcal{C}_X^o}h_{\rm top}(T^m,\mathcal{U}_0^{m-1})=m\sup_{\mathcal{U}\in\mathcal{C}_X^o}h_{\rm top}(T,\mathcal{U})=m\ h_{\rm top}(T).$$

上面第一个和最后一个等号的成立由定义所保证. □

注记 7.5.5 当(X,T)为可逆动力系统且$m\in\mathbb{Z}$时, 我们有$h_{\rm top}(T^m)=|m|h_{\rm top}(T)$.

定理 7.5.6 设$(X,T),(Y,S)$ 为拓扑动力系统, 那么

$$h_{\rm top}(X\times Y,T\times S)=h_{\rm top}(X,T)+h_{\rm top}(Y,S).$$

证明. 设X,Y 的度量为d_X,d_Y, 对于乘积空间$X\times Y$, 我们取度量

$$d((x,y),(x',y'))=\max\{d_X(x,x'),d_Y(y,y')\}.$$

在此度量下, 如果E,F 为X,Y 的(n,ε) 张成集, 那么$E\times F$ 为$X\times Y$ 的张成集. 由此我们有

$$\mathrm{sp}(n,\varepsilon,X\times Y)\le\mathrm{sp}(n,\varepsilon,X)\cdot\mathrm{sp}(n,\varepsilon,Y),$$

从而

$$h_{\rm top}(X\times Y,T\times S)\le h_{\rm top}(X,T)+h_{\rm top}(Y,S).$$

另一方面, 如果E,F 为X,Y 的(n,ε) 分离集, 那么$E\times F$ 为$X\times Y$ 的分离集. 于是

$$\mathrm{sr}(n,\varepsilon,X\times Y)\ge\mathrm{sr}(n,\varepsilon,X)\cdot\mathrm{sr}(n,\varepsilon,Y).$$

由此

$$\begin{aligned}\mathrm{sr}(\varepsilon,X\times Y)&\ge\limsup_{n\to\infty}\frac{1}{n}\left(\log\mathrm{sr}(n,\varepsilon,X)+\log\mathrm{sr}(n,\varepsilon,Y)\right)\\&\ge\liminf_{n\to\infty}\frac{1}{n}\log\mathrm{sr}(n,\varepsilon,X)+\liminf_{n\to\infty}\frac{1}{n}\log\mathrm{sr}(n,\varepsilon,Y)\\&=\underline{\mathrm{sr}}(\varepsilon,X)+\underline{\mathrm{sr}}(\varepsilon,Y).\end{aligned}$$

所以我们得到

$$h_{\rm top}(X\times Y,T\times S)\ge h_{\rm top}(X,T)+h_{\rm top}(Y,S).$$

由此我们完成了整个证明. □

定理 7.5.7 设$T:X\to X$ 为可扩同胚, 那么

1. 如果α 为生成子, 则$h_{\rm top}(T)=h_{\rm top}(T,\alpha)$.

2. 如果δ 为T 可扩常数, 那么

$$h_{\rm top}(T)={\rm sp}(\delta_0,T)={\rm sr}(\delta_0,T),$$

其中$\delta_0<\delta/4$.

证明. 1. 设β 为开覆盖, 设δ 为β 的Lebesgue 数. 根据定理5.7.12, 存在$N\in\mathbb{N}$ 使得${\rm diam}\bigvee_{n=-N}^N T^{-n}\alpha<\delta$. 于是$\beta\prec\bigvee_{n=-N}^N T^{-n}\alpha$. 于是

$$\begin{aligned}
h_{\rm top}(T,\beta)&\le h_{\rm top}(T,\bigvee_{n=-N}^{N}T^{-n}\alpha)\\
&=\lim_{k\to\infty}\frac{1}{k}H\Big(\bigvee_{i=0}^{k-1}T^{-i}(\bigvee_{n=-N}^{N}T^{-n}\alpha)\Big)=\lim_{k\to\infty}\frac{1}{k}H\Big(\bigvee_{n=-N}^{N+k-1}T^{-n}\alpha\Big)\\
&=\lim_{k\to\infty}\frac{1}{k}H\Big(\bigvee_{n=0}^{2N+k-1}T^{-n}\alpha\Big)\\
&=\lim_{k\to\infty}\frac{2N+k-1}{k}\frac{1}{2N+k-1}H\Big(\bigvee_{n=0}^{2N+k-1}T^{-n}\alpha\Big)\\
&=h_{\rm top}(T,\alpha).
\end{aligned}$$

于是对于任何开覆盖β, 我们得到$h_{\rm top}(T,\beta)\le h_{\rm top}(T,\alpha)$, 尤其

$$h_{\rm top}(T)=h_{\rm top}(T,\alpha).$$

2. 设$\delta_0<\delta/4$. 取$x_1,\dots,x_k$ 使得$X=\bigcup_{i=1}^k B(x_i,\frac{\delta}{2}-2\delta_0)$. 易见$2\delta_0$ 为覆盖$\alpha=\{B(x_i,\frac{\delta}{2}):1\le i\le k\}$ 的Lebesgue 数. 于是根据(7.4.1)

$$h_{\rm top}(T,\alpha)\le{\rm sp}(\delta_0,T)\le{\rm sr}(\delta_0,T)\le h_{\rm top}(T).$$

于是根据(1) 我们得到所求. □

由上命题我们直接得到如下推论.

推论 7.5.8 可扩同胚具有有限的拓扑熵.

定理 7.5.9 设$T:X\to X$ 为紧致度量空间上的可扩同胚, δ 为可扩常数. 设ξ 为满足${\rm diam}\,\xi<\delta$ 的有限可测剖分. 那么

$$\hat{\xi}^T=\bigvee_{n=-\infty}^{\infty}T^{-n}\hat{\xi}=\mathcal{X}(X).$$

尤其对于任何$\mu\in\mathcal{M}(X,T)$, 我们有

$$h_\mu(T)=h_\mu(T,\xi).$$

证明. 根据定理5.7.6, 对于任何$n \in \mathbb{N}$, 存在N_n 使得

$$\mathrm{diam}(\bigvee_{i=-N_n}^{N_n} T^{-i}\xi) < \frac{1}{n}.$$

设E_n 为$\bigvee_{i=-N_n}^{N_n} T^{-i}\xi$ 中与$B(x, \varepsilon - \frac{1}{n})$ 相交非空元素的并(如果$\varepsilon - \frac{1}{n} \le 0$, 那么约定$E_n = \emptyset$). 则$B(x, \varepsilon - \frac{1}{n}) \subset E_n \subset B(x, \varepsilon)$. 于是

$$B(x, \varepsilon) = \bigcup_{n=1}^{\infty} E_n \in \bigvee_{i=-\infty}^{\infty} T^{-i}\hat{\xi}.$$

由此任何开集都包含在$\bigvee_{i=-\infty}^{\infty} T^{-i}\hat{\xi}$ 中, 于是$\bigvee_{n=-\infty}^{\infty} T^{-n}\hat{\xi} = \mathcal{X}(X)$. □

计算熵时我们经常会证明下面命题, 根据这个结论我们计算熵仅需要限制在非游荡点集上计算即可. 定理的证明后面会给出.

定理 7.5.10 设(X, T)为动力系统且$\Omega(T)$为其非游荡点集. 那么

$$h_{\mathrm{top}}(T, X) = h_{\mathrm{top}}(T|_{\Omega(T)}, \Omega(T)).$$

§7.5.2　一些例子

例 7.5.11 设$\Sigma_k = \{0, 1, \ldots, k-1\}^{\mathbb{Z}}$, σ 为转移映射. 设$Y \subset \Sigma_k$ 为闭不变子集, 设$\theta_n(Y)$ 为Y 中长为n 的词的个数, 即

$$\theta_n(Y) = \mathrm{Card}\ \{(i_0 i_1 \ldots i_{n-1}) : \exists w \in Y\ s.t.\ w_0 = i_0, w_1 = i_1, \ldots, w_{n-1} i_{n-1}\}.$$

则

$$h_{\mathrm{top}}(\sigma|_Y) = \lim_{n\to\infty} \frac{1}{n} \log \theta_n.$$

特别地当$Y = \Sigma_k$ 时, 我们有

$$h_{\mathrm{top}}(\sigma) = \log k.$$

证明. 我们定义它的基本柱形集$[j] = \{x \in X : x_0 = j\}$, $0 \le j \le k-1$ 和Y 的基本剖分$\mathcal{U} = \{[j] \cap Y : 0 \le j \le k-1\}$. 易见$\theta_n = N(\mathcal{U}_0^{n-1}) = \mathrm{Card}(\mathcal{U}_0^{n-1})$.

明显地序列$\{\mathcal{U}_0^{n-1}\}_{n=1}^{\infty}$满足引理7.5.1 的条件, 因此利用引理7.5.1 我们有

$$\begin{aligned} h_{\mathrm{top}}(\sigma, X) &= \lim_{n\to+\infty} h_{\mathrm{top}}(\sigma, \mathcal{U}_0^{n-1}) = h_{\mathrm{top}}(\sigma, \mathcal{U}) \\ &= \lim_{n\to+\infty} \frac{1}{n} \log N(\mathcal{U}_0^{n-1}) = \lim_{n\to+\infty} \frac{1}{n} \log \theta_n. \end{aligned}$$

当$Y = \Sigma_k$ 时, $\theta_n = k^n$. 因此我们得到

$$h_{\mathrm{top}}(\sigma, \Omega) = \lim_{n\to+\infty} \frac{1}{n} \log k^n = \log k.$$

□

注意我们也可用定理7.5.7 处理上面例子.

例 7.5.12 设A 为不可约$k\times k$ 的$\{0,1\}$ 矩阵, λ 为A 最大正特征值, $T_A: X_A\to X_A$ 为$Markov$ 转移, 那么

$$h_{\text{top}}(T_A)=\log\lambda.$$

证明. 注意$\{(x_n)_{n=-\infty}^{\infty}\in X_A: x_0=i_0,\dots,x_{n-1}=i_{n-1}\}\neq\emptyset$ 当且仅当$a_{i_0i_1}a_{i_1i_2}\dots a_{i_{n-2}i_{n-1}}=1$. 沿用例子7.5.11 符号,

$$\theta_n(X_A)=\sum_{i_0,\dots,i_{n-1}=0}^{k-1}a_{i_0i_1}a_{i_2i_3}\dots a_{i_{n-2}i_{n-1}}=\sum_{i_0,\dots,i_{n-1}=0}^{k-1}(A^{n-1})_{i_0i_{n-1}},$$

其中$(A^{n-1})_{ij}$ 表示A^{n-1} 位置(ij) 的值. 取$k\times k$ 矩阵的模为$\|B\|=\sum_{i,j=0}^{k-1}|b_{ij}|$, 那么我们有

$$\theta_n(X_A)=\|A^{n-1}\|.$$

根据谱半径公式,

$$h_{\text{top}}(T_A)=\lim_{n\to\infty}\frac{1}{n}\log\theta_n(X_A)=\lim_{n\to\infty}\|A^{n-1}\|^{\frac{1}{n}}=\log\lambda.$$

□

例 7.5.13 设$T: I\to I$ 为区间自同胚, 那么$h_{\text{top}}(T)=0$

证明. 因为$\Omega(T^2)=Fix(T^2)$, 所以

$$h_{\text{top}}(T)=\frac{1}{2}h_{\text{top}}(T^2)=\frac{1}{2}h_{\text{top}}(T^2|_{Fix(T^2)})=0.$$

□

例 7.5.14 设$T:\mathbb{T}\to\mathbb{T}$ 为圆周自同胚, 则$h_{\text{top}}(T)=0$.

证明. 因为T 为同胚, 它将周周上子区间因为子区间. 设$\mathbb{T}$ 的长度为1. 取$\varepsilon>0$ 使得$d(x,y)\le\varepsilon$ 蕴含$d(T^{-1}x,T^{-1}y)\le 1/4$. 下面我们考虑张成集. 首先明显有$\text{sp}(1,\varepsilon,\mathbb{T})\le[\frac{1}{\varepsilon}]+1$. 下面我们证明$\text{sp}(n,\varepsilon,\mathbb{T})\le n([\frac{1}{\varepsilon}]+1)$.

设F 为基数$\text{sp}(n-1,\varepsilon,\mathbb{T})$ 的$(n-1,\varepsilon)$ 的张成集. 考虑点集$T^{n-1}F$ 以及它所决定的$\mathbb{T}$ 的所有子区间. 至多加上$[\frac{1}{\varepsilon}]+1$ 个点可以使得$T^{n-1}F$ 加上这些点得到的子区间长度都小于ε. 把新的点集记为E. 再设

$$F'=F\cup T^{-(n-1)}E.$$

我们证明F' 为$\mathbb{T}$ 的(n,ε) 张成集. 设$x\in\mathbb{T}$, 则存在$y\in F$ 使得

$$d_{n-1}(x,y)=\max_{0\le i\le n-2}d(T^ix,T^iy)\le\varepsilon$$

设I' 为端点为$T^{-(n-2)}x,T^{-(n-2)}y$ 长度小于等于ε 的子区间. 设I 为端点为$T^{-(n-1)}x,T^{-(n-1)}y$ 且在T^{-1} 下映射到I' 的子区间. 取点$z\in F'$ 使得$T^{n-1}z\in I$

且$d(T^{n-1}x,T^{n-1}z)\le\varepsilon$. 则$T^{n-2}z\in I'$ 且$d(T^{n-2}x,T^{n-2}z)\le\varepsilon$. 注意$I'$ 在T^{-1} 下像为端点为$T^{n-3}x,T^{n-3}y$ 长小于1/4 的区间, 设之为I''. 根据$d_{n-1}(x,y)\le\varepsilon$, 我们推出$I''$ 的长度实际上小于等于ε. 因为$T^{n-3}z\in I''$, 我们得到$d(T^{n-3}x,T^{n-3}z)\le\varepsilon$. 归纳地, 我们可以证明

$$d(T^ix,T^iz)\le\varepsilon,\quad 0\le i\le n-1.$$

亦即我们证明了F' 为一个(n,ε) 张成集, 尤其

$$\mathrm{sp}(n,\varepsilon,\mathbb{T})\le\mathrm{sp}(n-1,\varepsilon,\mathbb{T})+\left[\frac{1}{\varepsilon}\right]+1,$$

从而

$$\mathrm{sp}(n,\varepsilon,\mathbb{T})\le n(\left[\frac{1}{\varepsilon}\right]+1).$$

由此我们得到$h_{\mathrm{top}}(T)=0$. □

例 7.5.15 *对于任何给定的正数$a>0$, 存在系统(X,T) 使得$h_{\mathrm{top}}(T)=a$.*

证明. 设$a=\log\beta$, 则$\beta\ge1$. 因为全符号转移(Σ_k,σ) 的拓扑熵为$\log k$, 我们仅需要考虑β 为非整数的情况.

设$1=\sum_{n=1}^{\infty}a_n\beta^{-n}$ 为1 的$\frac{1}{\beta}$ 展开. 易见

$$a_1=[\beta],\quad a_n=\left[\beta^n-\sum_{i=1}^{n-1}a_i\beta^{n-i}\right].$$

设$k=[\beta]+1$, 那么$0\le a_n\le k-1$, $\forall n\in\mathbb{N}$. 则$a=(a_n)_{n=1}^{\infty}$ 为$\Sigma_k=\{0,1,\dots,k-1\}^{\mathbb{N}}$ 的点. 设$<$ 为Σ_k 的字典序, $\sigma:\Sigma_k\to\Sigma_k$ 为单边转移映射. 注意$\sigma^na\le a,\forall n\ge0$.

设

$$X_\beta=\{x=(x_n)_{n=1}^{\infty}:x\in\Sigma_k,\ \sigma^nx\le a,\forall n\ge0\}.$$

那么X_β 为Σ_k 闭不变子集$(\sigma X_\beta=X_\beta)$. 下面我们运用例子7.5.11 证明$h_{\mathrm{top}}(\sigma|_{X_\beta})=\log\beta$.

设θ_n 为X_β 中长为n 词的个数. 注意一个词$b_1b_2\dots b_n$ 出现在X_β 中当且仅当对于任何$k\in\{1,2,\dots,n\}$ 我们有$(b_kb_{k+1}\dots b_n)\le(a_1a_2\dots a_{n-k+1})$. 设$\theta_0=1$, $a_0=0$. 我们断言

$$\theta_n=1+a_0\theta_n+a_1\theta_{n-1}+\ldots+a_n\theta_0,\quad\forall n\ge0.$$

设词$b_1b_2\dots b_n$ 出现在X_β 中, 那么有以下情况:

(1). $b_1<a_1$, 且$(b_2b_3\dots b_n)$ 出现在X_β 中, 此时一共有$a_1\theta_{n-1}$ 种可能.

(2). $b_1=a_1$, $b_2<a_2$ 且$(b_3b_4\dots b_n)$ 出现在X_β 中, 此时一共有$a_2\theta_{n-1}$ 种可能.

......

(n). $(b_1b_2\dots b_{n-2})=(a_1a_2\dots a_{n-2})$, $b_{n-1}\le a_{n-1}$, 且b_n 出现在X_β 中, 此时一共有$a_{n-1}\theta_1$ 种可能.

(n+1). $(b_1b_2\dots b_{n-1})=(a_1a_2\dots a_{n-1})$, $b_n\le a_n$, 且b_n 出现在X_β 中, 此时一共有a_n+1种可能.

于是

$$\beta^{-n}\theta_n=\beta^{-n}+\beta^{-1}a_1\beta^{-n-1}\theta_{n-1}+\dots+\beta^{-n}a_n\theta_0.$$

根据更新定理,

$$\lim_{n\to\infty}\beta^{-n}\theta_n=(\sum_{n=0}^{\infty}\beta^{-n})(\sum_{n=0}^{\infty}na_n\beta^{-n})>0.$$

于是

$$h_{\mathrm{top}}(\sigma|_{X_\beta})=\lim_{n\to\infty}\frac{1}{n}\log\theta_n=\log\beta.$$

□

注记 7.5.16 在上面例子中, 如果我们需要同胚映射, 那么就令

$$X'_\beta=\{x=(x_n)_{n=-\infty}^{\infty}\in\{0,1,\dots,k-1\}^{\mathbb{Z}}:(x_ix_{i+1}\dots)\in X_\beta,\ \forall i\in\mathbb{Z}\}.$$

那么$\theta_n(X'_b)=\theta_n(X_\beta)$, 从而熵仍为$\log\beta$.

习 题

1. 设X为非空集合, $T:X\to X$为满射, $\mathcal{U}$ 为X的有限覆盖.证明：对$l\in\mathbb{N}$有$h_c(T,\mathcal{U})\ge\frac{1}{l}h_c(T^l,\mathcal{U})$.
2. 设(X,σ)为单边的k符号全转移(Ω,σ) 的子系统且$h_{\mathrm{top}}(\sigma,X)=\log k$, 证明$X=\Omega$.
3. 设(X,T)为动力系统.假设$X=X_1\cup X_2\cup\dots\cup X_n$, 其中$X_1,X_2,\cdots,X_n$ 为X的一些互不相交闭的不变集, 证明：$h_{\mathrm{top}}(T,X)=\max_{1\le i\le n}h_{\mathrm{top}}(T|_{X_i},X_i)$.

§7.6 熵映射

§7.6.1 熵映射的定义

定义 7.6.1 设(X,T) 为拓扑动力系统, $\mathcal{B}(X)$ 为$Borel\ \sigma$代数, $\mathcal{M}(X,T)$ 为全体不变概率测度的集合. 映射

$$h:\mathcal{M}(X,T)\to[0,\infty],\quad \mu\mapsto h_\mu(T)$$

称为**熵映射**.

定理 7.6.2 熵映射为仿射, 即对于任何$\mu,m\in\mathcal{M}(X,T)$ 以及$p\in[0,1]$, 有

$$h_{p\mu+(1-p)m}(T)=ph_\mu(T)+(1-p)h_m(T).$$

证明. 因为$\phi(x)=-x\log x$ 为凸的, 对于任何$B\in\mathcal{B}(X)$, 我们有

$$\begin{aligned}
0 &\le \phi(p\mu(B)+(1-p)m(B))-p\phi(\mu(B))-(1-p)\phi(m(B)) \\
&= (p\mu(B)+(1-p)m(B))\log(p\mu(B)+(1-p)m(B))-p\mu(B)\log\mu(B)-(1-p)m(B)\log m(B) \\
&= p\mu(B)\left[\log(p\mu(B)+(1-p)m(B))-\log(p\mu(B))\right] \\
&\quad +(1-p)m(B)\left[\log(p\mu(B)+(1-p)m(B))-\log((1-p)m(B))\right] \\
&\quad +p\mu(B)[\log(p\mu(B))-\log(\mu(B))]+(1-p)m(B)[\log((1-p)m(B))-\log(m(B))] \\
&\le 0+0+p\mu(B)\log p+(1-p)m(B)\log(1-p).
\end{aligned}$$

于是对于任何有限可测剖分ξ,

$$\begin{aligned}
0 &\le H_{p\mu+(1-p)m}(\xi)-pH_\mu(\xi)-(1-p)H_m(\xi) \\
&\le -\left(p\log p+(1-p)\log(1-p)\right) \\
&\le \log 2.
\end{aligned}$$

设η 为有限剖分, 在上面令$\xi=\bigvee_{i=0}^{i-1}T^{-i}\eta$, 我们有

$$h_{p\mu+(1-p)m}(T,\eta)=ph_\mu(T,\eta)+(1-p)h_m(T,\eta).$$

根据上式, 我们自然有

$$h_{p\mu+(1-p)m}(T)\le ph_\mu(T)+(1-p)h_m(T).$$

还需要证明反向不等式. 设$\varepsilon>0$, 取有限剖分η_1,η_2 使得

$$h_\mu(T,\eta_1)>\begin{cases} h_\mu(T)-\varepsilon, & \text{如果}h_\mu(T)<\infty; \\ \frac{1}{\varepsilon}, & \text{如果}h_\mu(T)=\infty. \end{cases}$$

$$h_m(T,\eta_1)>\begin{cases} h_m(T)-\varepsilon, & \text{如果}h_m(T)<\infty; \\ \frac{1}{\varepsilon}, & \text{如果}h_m(T)=\infty. \end{cases}$$

令$\eta=\eta_1\bigvee\eta_2$, 则

$$h_{p\mu+(1-p)m}(T,\eta)>\begin{cases} ph_\mu(T)+(1-p)h_m(T)-\varepsilon, & \text{如果}h_\mu(T),h_m(T)<\infty; \\ \frac{1}{\varepsilon}, & \text{如果}h_\mu(T)=\infty\ \text{或}h_m(T)=\infty. \end{cases}$$

于是

$$h_{p\mu+(1-p)m}(T)\ge ph_\mu(T)+(1-p)h_m(T).$$

证毕! □

下面例子说明$h:M(X,T)\to[0,\infty]$ 不必为连续的, 甚至不必为上半连续的.

例 7.6.3 设$X=\Sigma_2=\{0,1\}^{\mathbb{Z}}$. $\mathrm{Fix}(X,\sigma^n)$ 为X 的n 周期点全体, 共2^n 个元素. 设$\mu_n\in\mathcal{M}(X,T)$ 为原子测度, 它在$\mathrm{Fix}(X,\sigma^n)$ 每个点取测度$\frac{1}{2^n}$. 明显地, 我们有$h_{\mu_n}(\sigma)=0$.

设μ 为X 上$(\frac{1}{2},\frac{1}{2})$乘积测度, 那么之前我们计算过, 它的熵为$h_\mu(\sigma)=\log 2$. 下证$\mu_n\to\mu,n\to\infty$. 设$A\subset C(X)$ 为全体全体由有限个坐标确定的连续函数全体, 根据*Stone-Weierstrass* 定理, A 在$C(X)$ 中稠密. 对于任何$f\in A$, 存在N 使得当$n\geq N$ 时$\int_X fd\mu_n=\int_X fd\mu$. 即$\mu_n\to\mu,n\to\infty$.

综上, 我们有$\mu_n\to\mu,n\to\infty$, 但是$h_{\mu_n}\not\to h_\mu,n\to\infty$.

例 7.6.4 设$Y=\{0\}\cup\{\frac{1}{n}:n\in\mathbb{N}\}$, 拓扑取$\mathbb{R}$ 的诱导拓扑, 它为紧致集合. 令$X=Y^{\mathbb{Z}}$,$T:X\to X$ 为转移映射. 对$n\in\mathbb{N}$, 设ν_n 为Y 上使得$\nu_n(\{\frac{1}{n}\})=\nu(\{\frac{1}{n+1}\})=\frac{1}{2}$ 的测度, 而μ_n 为相应的乘积测度. 因为(X,μ_n,T) 测度同构于(Σ_2,σ) 上$(\frac{1}{2},\frac{1}{2})$ 乘积测度, 所以$h_{\mu_n}(T)=\log 2,\forall n\in\mathbb{N}$. 类似于上面例子, 容易验证$\mu_n\to\mu,n\to\infty$, 其中$\mu=\delta_{\mathbf{0}}$ 为点$\mathbf{0}=(\ldots 000\ldots)$ 上*Dirac* 测度. $h_\mu(T)=0$.

综上, 我们有$\mu_n\to\mu,n\to\infty$, 但是$h_{\mu_n}\not\to h_\mu,n\to\infty$. 此时熵映射不是上半连续的.

§7.6.2 熵映射及上半连续性

定理 7.6.5 对于可扩同胚$T:X\to X$, 熵映射是上半连续的. 即对于任何$\mu\in\mathcal{M}(X,T)$, $\varepsilon>0$, 存在μ 的邻域U 使得对于任何$m\in U$, 我们有

$$h_m(T)<h_\mu(T)+\varepsilon.$$

证明. 设δ 为T 的可扩常数. 设$\mu\in\mathcal{M}(X,T)$, $\varepsilon>0$. 设$\gamma=\{C_1,\ldots,C_k\}$ 为满足$\mathrm{diam}\gamma<\delta$ 的可测剖分. 由定理7.5.9, $h_\mu(T)=h_\mu(T,\gamma)$. 取N 使得

$$\frac{1}{N}H_\mu(\bigvee_{j=0}^{N-1}T^{-j}\gamma)<h_\mu(T)+\frac{\varepsilon}{2}.$$

因为μ 为正则测度, 取紧致子集

$$K(i_0,i_1,\ldots,i_{N-1})\subset\bigcap_{j=0}^{N-1}T^{-j}C_{i_j}$$

使得$\mu\left(\bigcap_{j=0}^{N-1}T^{-j}C_{i_j}\setminus K(i_0,i_1,\ldots,i_{N-1})\right)<\varepsilon_1$, 其中$\varepsilon_1$ 为后面再取的常数. 则

$$C_i\supset L_i=\bigcup_{j=0}^{N-1}\{T^jK(i_0,\ldots,i_{N-1}):i_j=i\}.$$

根据定义, $L_1,\ldots,L_k$ 为互不相交的紧致子集. 取剖分$\gamma'=\{C'_1,\ldots,C'_k\}$ 使得$\mathrm{diam}\gamma'<\delta$ 且$L_i\subset\mathrm{int}(C'_i),1\leq i\leq k$. 根据定义,

$$K(i_0,i_1,\ldots,i_{N-1})\subset\mathrm{int}\Big(\bigcap_{j=0}^{N-1}T^{-j}C'_{i_j}\Big).$$

根据Urysohn 引理, 取连续函数$f_{i_0,\dots,i_{N-1}} \in C(X)$ 使得$0 \le f_{i_0,\dots,i_{N-1}} \le 1$, 并且$f_{i_0,\dots,i_{N-1}}$在$K(i_0,\dots,i_{N-1})$ 上取1, 在$X \setminus \operatorname{int}\left(\bigcap_{j=0}^{N-1} T^{-j}C'_{i_j}\right)$ 取0. 令

$$U(i_0,\dots,i_{N-1}) = \left\{ m \in \mathcal{M}(X,T) : \left| \int_X f_{i_0,\dots,i_{N-1}} dm - \int_X f_{i_0,\dots,i_{N-1}} d\mu \right| < \varepsilon_1 \right\}.$$

那么$U(i_0,\dots,i_{N-1})$ 为$\mathcal{M}(X,T)$ 的开集, 并且对于$m \in U(i_0,\dots,i_{N-1})$,

$$m(\bigcap_{j=0}^{N-1} T^{-j}C'_{i_j}) \ge \int_X f_{i_0,\dots,i_{N-1}} dm > \int_X f_{i_0,\dots,i_{N-1}} d\mu - \varepsilon_1 \ge \mu(K(i_0,\dots,i_{N-1})) - \varepsilon_1.$$

于是对于$m \in U(i_0,\dots,i_{N-1})$,

$$\mu\Big(\bigcap_{j=0}^{N-1} T^{-j}C_{i_j}\Big) - m\Big(\bigcap_{j=0}^{N-1} T^{-j}C'_{i_j}\Big) < 2\varepsilon_1.$$

令

$$U = \bigcap_{i_0,\dots,i_{N-1}=1}^{k} U(i_0,\dots,i_{N-1}).$$

那么对于任何$m \in U$, 我们有[3]

$$\left|\mu\Big(\bigcap_{j=0}^{N-1} T^{-j}C_{i_j}\Big) - m\Big(\bigcap_{j=0}^{N-1} T^{-j}C'_{i_j}\Big)\right| < 2\varepsilon_1 k^N.$$

于是如果$m \in U$, 取ε_1 充分小使得根据$x \log x$ 连续性, 我们有

$$\frac{1}{N} H_m\left(\bigvee_{j=0}^{N-1} T^{-j}\gamma'\right) < \frac{1}{N} H_\mu\left(\bigvee_{j=0}^{N-1} T^{-j}\gamma\right) + \frac{\varepsilon}{2}.$$

由此我们有

$$h_m(T) = h_m(T,\gamma') \le \frac{1}{N} H_m\left(\bigvee_{j=0}^{N-1} T^{-j}\gamma'\right) < \frac{1}{N} H_\mu\left(\bigvee_{j=0}^{N-1} T^{-j}\gamma\right) + \frac{\varepsilon}{2} < h_\mu(T) + \varepsilon.$$

证毕! □

注意紧致空间上上半连续实值函数能取到最大值, 所以对于可扩同胚, 存在$m \in \mathcal{M}(X,T)$ 使得$h_m(T) = \sup\{h_\mu(T) : \mu \in \mathcal{M}(X,T)\}$.

定理 7.6.6 设(X,T) 为拓扑动力系统, $\{\xi_n\}_{n=1}^\infty$ 为X 可测剖分列, 且满足$\operatorname{diam} \xi_n \to 0, n \to \infty$. 那么对于任何$\mu \in \mathcal{M}(X,T)$, 我们有

$$h_\mu(T) = \lim_{n\to\infty} h_\mu(T,\xi_n).$$

[3]这里我们用了下面简单的事实：如果$\sum_{i=1}^m a_i = \sum_{i=1}^m b_i = 1$, 存在$c > 0$ 使得$a_i - b_i < c, 1 \le i \le m$, 那么$|a_i - b_i| = |\sum_{j\ne i}(a_j - b_j)| < mc, 1 \le i \le m$.

证明. 设$\mu \in \mathcal{M}(X,T)$. 设$\varepsilon > 0$. 取可测剖分$\xi = \{A_1, \dots, A_k\}$ 使得

$$h_\mu(T,\xi) > \begin{cases} h_\mu(T) - \varepsilon, & \text{如果} h_\mu(T) < \infty; \\ \frac{1}{\varepsilon}, & \text{如果} h_\mu(T) = \infty. \end{cases}$$

取$\delta > 0$ 使得对于任何k 元剖分$\xi = \{A_1, \dots, A_k\}$, $\eta = \{C_1, \dots, C_k\}$, $\sum_{i=1}^k \mu(A_i \mathcal{D} C_i) < \delta$ 蕴含$H_\mu(\xi|\eta) + H_\mu(\eta|\xi) < \varepsilon$. 取紧致子集$K_i \subset A_i, 1 \le i \le k$ 使得$\mu(A_i \setminus K_i) < \frac{\delta}{k+1}$. 令$\delta' = \inf_{i \neq j} d(K_i, K_j)$. 取$n \in \mathbb{N}$ 使得$\mathrm{diam}\xi_n < \delta'/2$.

对$1 \le i \le k-1$, 令$E_{n,i}$ 为与K_i 相交非空的ξ_n 元的并, 令$E_{n,k} = X \setminus \bigcup_{i=1}^{k-1} E_{n,i}$. 设$\xi_n' = \{E_{n,1}, \dots, E_{n,k}\}$. 则$\xi_n' \preccurlyeq \xi_n$. 因为$\mathrm{diam}\xi_n < \delta'/2$, 所以每个$\xi_n$ 中元只与至多一个K_i 相交非空, 于是

$$\sum_{i=1}^k \mu(E_{n,i} \mathcal{D} A_i) = \sum_{i=1}^k \mu(E_{n,i} \setminus A_i) + \mu(A_i \setminus E_{n,i}) \le \mu(X \setminus \bigcup_{i=1}^k K_i) + \sum_{i=1}^k \mu(A_i \setminus K_i) < \delta.$$

根据δ 选取, $H_\mu(\xi_n|\xi_n') < \varepsilon$. 所以

$$h_\mu(T,\xi) = h_\mu(T,\xi_n') + H_\mu(\xi|\xi_n') < h_\mu(T,\xi_n') + \varepsilon \le h_\mu(T,\xi_n) + \varepsilon.$$

综上, 当n 使得$\mathrm{diam}(\xi_n) < \delta'/2$ 时, 我们有

$$h_\mu(T,\xi_n) > \begin{cases} h_\mu(T) - 2\varepsilon, & \text{如果} h_\mu(T) < \infty; \\ \frac{1}{\varepsilon} - \varepsilon, & \text{如果} h_\mu(T) = \infty. \end{cases}$$

所以$h_\mu(T) = \lim_{n\to\infty} h_\mu(T,\xi_n)$. □

§7.6.3 上半连续与遍历分解

设$\mu = \int_\Omega \mu_\omega dm(\omega)$ 为μ 的遍历分解. 设$F : \mathcal{M}(X,T) \to \mathbb{R}$ 为上半连续仿射, 那么它为单调递增连续仿射的极限, 从而我们有

$$F(\mu) = \int_\Omega F(\mu_\omega) dm(\omega).$$

根据这个事实我们有如下命题:

定理 7.6.7 设(X,T) 为拓扑动力系统. 对$\mu \in \mathcal{M}(X,T)$ 以及有限可测剖分ξ. 如果$\mu = \int_\Omega \mu_\omega dm(\omega)$ 为μ 的遍历分解, 那么

$$h_\mu(T,\xi) = \int_\Omega h_{\mu_\omega}(T,\xi) dm(\omega), \quad h_\mu(T) = \int_\Omega h_{\mu_\omega}(T) dm(\omega).$$

证明. 设$\xi = \{A_1, \dots, A_k\}$, $\Sigma_k = \{1, 2, \dots, k\}^{\mathbb{Z}}$. 定义

$$\phi : X \to \Sigma_k \text{ 使得 } \phi(x) = (i_n)_{n\in\mathbb{Z}} \text{当且仅当 } T^n x \in A_{i_n}.$$

于是诱导了映射

$$\phi_*:\mathcal{M}(X,T)\to\mathcal{M}(\Sigma_k,\sigma),\ m\mapsto m\circ\phi^{-1}.$$

易验证, 如$\mu=\int_{\mathcal{M}^e(X,T)}\nu dm(\nu)$ 为遍历分解, 那么

$$\phi_*\mu=\int_{\mathcal{M}^e(\Sigma_k,\sigma)}\nu dm(\phi_*^{-1}(\nu))$$

为$\phi_*\mu$ 遍历分解. 因为(Σ_k,σ) 为可扩系统, 所以其熵映射为上半连续的, 于是

$$h_{\phi_*\mu}(\sigma)=\int_{\mathcal{M}^e(\Sigma_k,\sigma)}h_\nu(\sigma)dm(\phi_*^{-1}(\nu))=\int_{\mathcal{M}^e(X,T)}h_{\phi_*\nu}(\sigma)dm(\nu).$$

因为$\xi=\phi^{-1}\eta$, 其中$\eta=\{{}_0[1]_0,\dots,{}_0[k]_0\}$. 于是对于任意$\nu\in\mathcal{M}(X,T)$, $h_{\phi_*\nu}(\sigma)=h_{\phi_*\nu}(\sigma,\eta)=h_\nu(T,\xi)$. 由此

$$h_\mu(T,\xi)=\int_{\mathcal{M}^e(X,T)}h_\nu(T,\xi)dm(\nu).$$

取$\{\xi_n\}_{n=1}^\infty$ 为X 可测剖分列, 且满足$\mathrm{diam}\ \xi_n\to 0, n\to\infty$. 那么根据定理7.6.6, 对于任何$\mu\in\mathcal{M}(X,T)$, 我们有$h_\mu(T)=\lim_{n\to\infty}h_\mu(T,\xi_n)$. 于是

$$h_\mu(T)=\lim_{n\to\infty}h_\mu(T,\xi_n)=\lim_{n\to\infty}\int_{\mathcal{M}^e(X,T)}h_\nu(T,\xi_n)dm(\nu)=\int_{M^e(X,T)}h_\nu(T)dm(\nu).$$

证毕! □

§7.7　熵的变分原理

熵的变分原理指出拓扑熵与测度熵的关系. 我们在本节介绍这个经典结果，后面还会提及局部的变分原理.

定理 7.7.1 *设(X,T) 为拓扑动力系统, 那么*

1. $h_\mu(T)\le h_{\mathrm{top}}(T),\forall\mu\in\mathcal{M}(X,T)$,
2. $h_{\mathrm{top}}(T)=\sup_{\mu\in\mathcal{M}(X,T)}h_\mu(T)$.

证明. (1) 取定可测剖分$\alpha=\{A_1,\dots,A_k\}$. 对于任何$\varepsilon>0$, 取紧致子集$\widetilde{A}_i\subset A_i, 1\le i\le k$ 使得$\mu(A_i\setminus\widetilde{A}_i)<\varepsilon$. 定义新的剖分

$$\widetilde{\alpha}=\{\widetilde{A}_1,\dots,\widetilde{A}_k,V\},$$

其中$V=X\setminus\bigcup_{i=1}^k\widetilde{A}_i$. 定义开覆盖

$$\mathcal{U}=\{\widetilde{A}_1\cup V,\dots,\widetilde{A}_k\cup V\}.$$

对比$\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U}$ 和$\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha}$, 我们有

$$N(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha})\le 2^nN(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U}). \tag{7.7.1}$$

注意上面$N(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha})$ 为$\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha}$ 非平凡元素个数, 而$N(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U})$ 为覆盖$\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U}$ 子覆盖最小个数.

根据命题7.1.9-(1), $H_\mu(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha})\le\log N(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha})$. 结合(7.7.1) 得到

$$H_\mu(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha})\le\log N(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{\alpha})\le n\log 2+\log N(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U}).$$

注意到

$$h_{\rm top}(T)\ge h_{\rm top}(T,\mathcal{U})=\lim_{n\to\infty}\frac{1}{n}H(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{U}),$$

以及

$$h_\mu(T,\alpha)=\lim_{n\to\infty}H_\mu(\bigvee_{i=0}^{n-1}T^{-i}\alpha),$$

我们得到

$$h_\mu(T,\widetilde{\alpha})\le\log 2+h_{\rm top}(T).$$

取ε 充分小使得

$$|h_\mu(T,\widetilde{\alpha})-\mathfrak{h}_\mu(T,\alpha)|\le H_\mu(\alpha|\widetilde{\alpha})+H(\widetilde{\alpha}|\alpha)<1.$$

于是

$$h_\mu(T,\alpha)\le\log 2+h_{\rm top}(T)+1.$$

综上我们得到

$$h_\mu(T)=\sup\{h_\mu(T,\alpha):\alpha\text{ 有限剖分}\}\le h_{\rm top}(T)+\log 2+1.$$

在上式中用T^k 替代T,

$$h_\mu(T^k)\le h_{\rm top}(T^k)+\log 2+1$$

从而

$$h_\mu(T)=\lim_{k\to\infty}\frac{h_\mu(T^k)}{k}\le\lim_{k\to\infty}\frac{h_{\rm top}(T^k)}{k}+\lim_{k\to\infty}\frac{\log 2+1}{k}=h_{\rm top}(T).$$

(2) 我们证明对于任何$\delta>0$, 存在$\mu\in\mathcal{M}(X,T)$ 使得$h_\mu(T)\ge h_{\rm top}(T)-\delta$. 取充分小$\varepsilon>0$ 使得$\limsup_{n\to\infty}\frac{1}{n}\log{\rm sr}(n,\varepsilon)\ge h_{\rm top}(T)-\delta$. 取子列$\{n_i\}_{i=1}^\infty$ 使得

$$\limsup_{i\to\infty}\frac{1}{n_i}\log{\rm sr}(n_i,\varepsilon)\ge h_{\rm top}(T)-\delta.$$

设S_{n_i} 为个数为$\mathrm{sr}(n_i,\varepsilon)$ 的(n_i,ε) 分离集. 对每个n_i, 定义

$$\nu_{n_i}=\frac{1}{\mathrm{sr}(n_i,\varepsilon)}\sum_{x\in S_{n_i}}\delta_x.$$

再定义

$$\mu_{n_i}=\frac{1}{n_i}\sum_{r=0}^{n_i-1}(T^r)_*\nu_{n_i}.$$

不妨设$\lim_{i\to\infty}\mu_{n_i}=\mu$ (否则取子列), 则$\mu\in\mathcal{M}(X,T)$. 下面验证μ 为所求: $h_\mu(T)\geq h_{\mathrm{top}}(T)-\delta$.

取可测剖分$\alpha=\{A_1,\dots,A_k\}$ 满足$\mathrm{diam}(\alpha)<\varepsilon$ 并且$\mu(\partial A_i)=0, 1\leq i\leq k$. 因为$S_{n_i}$ 为(n_i,ε) 分离集, 对于每个$C\in\alpha_0^{n_i-1}=\bigvee_{j=0}^{n_i-1}T^{-j}\alpha$ 至多包含一个点$x=x_C\in S_{n_i}$. 所以$\alpha_0^{n_i-1}$ 的元素中, 有$\mathrm{sr}(n_i,\varepsilon)$ 个元素具有ν_{n_i} 测度$\frac{1}{\mathrm{sr}(n_i,\varepsilon)}$, 其余$\nu_{n_i}$测度为0. 尤其

$$\log\mathrm{sr}(n_i,\varepsilon)=-\sum_{C\in\alpha_0^{n_i-1}}\nu_{n_i}(C)\log\nu_{n_i}(C). \tag{7.7.2}$$

取定$1<N<n_i$, 再取定$j\in\{0,1,\dots,N-1\}$. 则

$$\alpha_0^{n_i-1}=\bigvee_{i=0}^{n_i-1}T^{-i}\alpha=\left(\bigvee_{l\equiv j\pmod N),0\leq l\leq n_i-N}T^{-l}\Big(\bigvee_{i=0}^{N-1}T^{-i}\alpha\Big)\right)\bigvee\left(\bigvee_{i\in E}T^{-i}\alpha\right),$$

其中$E=\{0,1,\dots,j-1\}\bigcup\{M_j,M_j+1,\dots,n_i-1\}$, $M_j=N[\frac{n_i-j}{N}]$. 注意$\mathrm{Card}(E)\leq 2N$.

于是

$$\begin{aligned}&-\sum_{C\in\alpha_0^{n_i-1}}\nu_{n_i}(C)\log\nu_{n_i}(C)\\ \leq&\sum_{l\equiv j\pmod N),0\leq l\leq n_i-N}\left(-\sum_{C\in T^{-l}\alpha_0^{N-1}}\nu_{n_i}(C)\log\nu_{n_i}(C)\right)+\sum_{i\in E}\left(-\sum_{C\in T^{-i}\alpha_0^{N-1}}\nu_{n_i}(C)\log\nu_{n_i}(C)\right)\\ \leq&\sum_{r=0}^{M_j}\left(-\sum_{D\in\alpha_0^{N-1}}(T^{rN+j})_*\nu_{n_i}(D)\log(T^{rN+j})_*\nu_{n_i}(D)\right)+2N\log k\end{aligned} \tag{7.7.3}$$

[[注意对于$l=rN+j$, $D\in\alpha_0^{N-1}$ 对应于$C\in T^{-l}\alpha_0^{N-1}$ 使得$(T^l)_*\nu_{n_i}(D)=\nu_{n_i}(T^{-l}D)=\nu_{n_i}(C)$ 且$C=T^{-l}D$.]]

将(7.7.3) 对$j=0,1,\dots,N-1$ 做和, 结合(7.7.2) 我们得到

$$N\log\mathrm{sr}(n_i,\varepsilon)\leq\sum_{l=0}^{n_i-1}\left(-\sum_{D\in\alpha_0^{N-1}}(T^l)_*\nu_{n_i}(D)\log(T^l)_*\nu_{n_i}(D)\right)+2N^2\log k. \tag{7.7.4}$$

根据$-x\log x$ 的凸性以及(7.7.4),

$$\begin{aligned}
&\frac{1}{n_i}\log \mathrm{sr}(n_i,\varepsilon)\\
&\le \frac{1}{n_i}\sum_{l=0}^{n_i-1}\left(-\frac{1}{N}\sum_{C\in\alpha_0^{N-1}}(T^l)_*\nu_{n_i}(C)\log(T^l)_*\nu_{n_i}(C)\right)+\frac{2N\log k}{n_i}\\
&\le -\frac{1}{N}\sum_{C\in\alpha_0^{N-1}}\mu_{n_i}(C)\log\mu_{n_i}(C)+\frac{2N\log k}{n_i}.
\end{aligned}$$

因为$\mu(\partial A_i)=0, 1\le i\le k$, 所以在上式中取$n_i\to\infty$ 我们得到

$$\begin{aligned}
h_{\mathrm{top}}(T)-\delta &\le \lim_{i\to\infty}\frac{1}{n_i}\log \mathrm{sr}(n_i,\varepsilon)\\
&\le -\frac{1}{N}\lim_{i\to\infty}\sum_{C\in\alpha_0^{N-1}}\mu_{n_i}(C)\log\mu_{n_i}(C)+\lim_{i\to\infty}\frac{2N\log k}{n_i}\\
&=\frac{1}{N}H_\mu(\alpha_0^{N-1}).
\end{aligned}$$

再令$N\to\infty$,

$$h_{\mathrm{top}}(T)-\delta\le \lim_{N\to\infty}\frac{1}{N}H_\mu(\alpha_0^{N-1})=h_\mu(T,\alpha)\le h_\mu(T).$$

证毕! □

推论 7.7.2 设(X,T) 为拓扑动力系统, 那么

1. $h_{\mathrm{top}}(T)=\sup_{\mu\in\mathcal{M}^e(X,T)}h_\mu(T)$.
2. $h_{\mathrm{top}}(T)=h_{\mathrm{top}}(T|_{\Omega(T)})$.
3. $h_{\mathrm{top}}(T)=h_{\mathrm{top}}(T|_{\cap_{n=0}^\infty T^nX})$

习 题

1. 证明推论7.7.2.
2. 能否不运用变分原理给出推论7.7.2-2证明?

§7.8 最大熵测度

§7.8.1 最大熵测度的定义和性质

定义 7.8.1 设(X,T) 为动力系统, 如果$\mu\in\mathcal{M}(X,T)$ 满足$h_\mu(T)=h_{\mathrm{top}}(T)$, 那么称$\mu$ 为**最大熵测度** . 记全体最大熵测度的集合为$\mathcal{M}_{max}(X,T)$.

注记 7.8.2 最大熵测度集合$\mathcal{M}_{max}(X,T)$ 有可能为空集. 对于零熵系统$h_{\text{top}}(T) = 0$, 易见$\mathcal{M}_{max}(X,T) = \mathcal{M}(X,T)$.

命题 7.8.3 设(X,T) 为动力系统, 则

1. $\mathcal{M}_{max}(X,T)$ 为凸集;
2. 如果$h_{\text{top}}(T) < \infty$, 那么$\mathcal{M}_{max}(X,T)$ 的端点恰为$\mathcal{M}_{max}(X,T)$ 的遍历元;
3. 如果$h_{\text{top}}(T) < \infty$ 且$\mathcal{M}_{max}(X,T) \neq \emptyset$, 那么$\mathcal{M}_{max}(X,T)$ 有遍历元;
4. 如果$h_{\text{top}}(T) = \infty$, 那么$\mathcal{M}_{max}(X,T) \neq \emptyset$;
5. 如果T 的熵映射上半连续, 那么$\mathcal{M}_{max}(X,T) \neq \emptyset$ 且紧致.

证明. (1) 因为熵映射是仿射, 所以$\mathcal{M}_{max}(X,T)$ 为凸集.

(2) 如果$\mu \in \mathcal{M}_{max}(X,T)$ 是遍历的, 那么它为$\mathcal{M}(X,T)$ 的端点, 自然也为$\mathcal{M}_{max}(X,T)$ 的端点. 反之, 设$\mu \in \mathcal{M}_{max}(X,T)$ 为$\mathcal{M}_{max}(X,T)$ 的端点. 设$\mu = p\mu_1 + (1-p)\mu_2, p \in [0,1], \mu_1, \mu_2 \in \mathcal{M}(X,T)$. 则根据仿射性,

$$h_{\text{top}}(T) = h_\mu(T) = ph_{\mu_1}(T) + (1-p)h_{\mu_2}(T).$$

由变分原理, $h_{\mu_1}(T), h_{\mu_2}(T) \leq h_{\text{top}}(T)$. 于是我们有$h_{\mu_1}(T) = h_{\mu_2}(T) = h_{\text{top}}(T) = h_\mu(T)$, 即$\mu_1, \mu_2 \in \mathcal{M}_{max}(X,T)$. 于是$\mu = \mu_1 = \mu_2$, 所以$\mu$ 为$\mathcal{M}(X,T)$ 的端点, 它为遍历的.

(3) 设$\mu \in \mathcal{M}_{max}(X,T)$, 设其遍历分解为$\mu = \int_{\mathcal{M}^e(X,T)} m d\tau(m)$. 根据定理7.6.7,

$$h_{\text{top}}(T) = h_\mu(T) = \int_{\mathcal{M}^e(X,T)} h_m(T) d\tau(m).$$

由变分原理（定理7.7.1）, $h_m(T) \leq h_{\text{top}}(T)$. 于是对于$\tau$-a.e. m, 有$h_m(T) = h_{\text{top}}(T)$, 尤其$\mathcal{M}_{max}(X,T)$ 有遍历元.

(4) 由变分原理（定理7.7.1）, 对于任何$n \in \mathbb{N}$, 取$h_{\mu_n}(T) > 2^n$. 令

$$\mu = \sum_{n=1}^{\infty} \frac{1}{2^n} \mu_n \in \mathcal{M}(X,T)$$

对于$N \in \mathbb{N}$, 记$\mu = \sum_{n=1}^{N} \mu_n + \frac{1}{2^N}\nu$, 其中$\nu \in \mathcal{M}(X,T)$. 根据熵映射的仿射性,

$$h_\mu(T) \geq \sum_{n=1}^{N} \frac{1}{2^n} h_{\mu_n}(T) > N.$$

由N 的任意性, $h_\mu(T) = \infty$, 即$\mu \in \mathcal{M}_{max}(X,T)$.

(5) 如果熵映射是上半连续的, 那么上半连续函数在紧致子集$\mathcal{M}(X,T)$ 上取到最大值$h_{\rm top}(T)$, 所以$\mathcal{M}_{max}(X,T)\neq\emptyset$. 下说明它是紧致的. 设$\mu_n\in\mathcal{M}_{max}(X,T)$, 且$\lim_{n\to\infty}\mu_n=\mu\in\mathcal{M}(X,T)$. 于是

$$h_\mu(T)\geq\limsup_{n\to\infty}h_{\mu_n}(T)=h_{\rm top}(T).$$

即$\mu\in\mathcal{M}_{max}(X,T)$, 由此$\mathcal{M}_{max}(X,T)$ 为紧致的. □

注记 7.8.4 *存在极小系统(X,T), $h_{\rm top}(T)=\infty$ 但是对于任何$\mu\in\mathcal{M}^e(X,T)$ 有$h_\mu(T)<\infty$. 此说明当$h_{\rm top}(T)=\infty$ 时, 不一定存在遍历的最大测度.*

例 7.8.5 (Gurevič) *存在系统(X,T) 使得$\mathcal{M}_{max}(X,T)=\emptyset$.*

证明. 取单调递增数论$\beta_n\in(1,2)$, 且$\beta_n\nearrow 2, n\to\infty$. 设$(X_n,T_n)$ 为例子7.5.15 中系统使得$h_{\rm top}(T_n)=\log\beta_n$. 取$X_n$ 上度量d_n 使得${\rm diam}_{d_n}(X)\leq 1$.

设X 为所有X_n 的无交并附加上紧化点x_∞, 使得X_n 收缩为x_∞. 具体如下操作：定义$X=\bigsqcup_{n=1}^\infty X_n\sqcup\{x_\infty\}$, 定义$X$ 上度量ρ:

$$\rho(x,y)=\begin{cases}\dfrac{1}{n^2}d_n(x,y), & \text{如果}x,y\in X_n;\\ \displaystyle\sum_{i=n}^m\frac{1}{i^2}, & \text{如果}x\in X_n,y\in X_m,n<m;\\ \displaystyle\sum_{i=n}^\infty\frac{1}{i^2}, & \text{如果}x_n\in X,y=x_\infty.\end{cases}$$

可以验证(X,ρ) 为紧致度量空间. 定义

$$T:X\to X,\quad T|_{X_n}=T_n,\ T(x_\infty)=x_\infty.$$

如果$\mu\in\mathcal{M}(X,T)$, 那么

$$\mu=\sum_{n=1}^\infty p_n\mu_n+(1-\sum_{n=1}^\infty p_n)\delta_{x_\infty},$$

其中$\mu_n\in\mathcal{M}(X_n,T_n),p_n\geq 0,\sum_{n=1}^\infty p_n\leq 1$. 于是如果$\mu\in\mathcal{M}^e(X,T)$ 当且仅当存在n 使得$\mu\in\mathcal{M}^e(X_n,T_n)$, 或$\mu=\delta_{x_\infty}$. 根据变分原理

$$h_{\rm top}(T)=\sup_{\mu\in\mathcal{M}^e(X,T)}h_\mu(T)=\sup_{n\in\mathbb{N}}\sup_{\mu\in\mathcal{M}^e(X_n,T_n)}h_\mu(T_n)=\sup_{n\in\mathbb{N}}h_{\rm top}(T_n)=\log 2.$$

如果$\mathcal{M}_{max}(X,T)\neq\emptyset$, 根据命题7.8.3, $\mathcal{M}_{max}(X,T)$ 中存在遍历元μ. 根据构造存在$n\in\mathbb{N}$ 使得$\mu\in\mathcal{M}^e(X_n,T_n)$. 但是$h_\mu(T)\leq\log\beta_n\leq\log 2$, 矛盾！所以$\mathcal{M}_{max}(X,T)=\emptyset$. □

注记 7.8.6 *1. 存在极小系统, 也存在紧流形上微分同胚使得其上不存在最大熵测度.*

2. Denker-Grillenberger-Sigmund 定理指出：$h_{\rm top}(T)<\infty$ 且$\mathcal{M}_{max}(X,T)\neq\emptyset$ 当且仅当存在X 的有限覆盖列$\{\alpha_n\}_{n=1}^\infty$ 使得$\sum_{n=1}^\infty h_{\rm top}(T,\alpha_n)<\infty$ 且$\lim_{k\to\infty}h_{\rm top}(T,\vee_{n=1}^k\alpha_n)=h_{\rm top}(T)$.

§7.8.2 唯一最大熵测度

定义 7.8.7 设(X,T) 为拓扑动力系统. 如果它的最大熵测度存在且唯一, 那么就称系统具有**唯一的最大熵测度**.

注记 7.8.8 *1.* 如果系统(X,T) 为唯一遍历的, 那么它自然具有唯一最大熵测度.

2. 如果$h_{\rm top}(X,T)=\infty$ 并且$\mathcal{M}_{max}(X,T)=\{\mu\}$, 那么$\mathcal{M}(X,T)=\{\mu\}$. 否则设$m\in\mathcal{M}(X,T)$ 异于μ, 根据$h_{\frac{1}{2}\mu+\frac{1}{2}m}(T)=\infty$,与具有唯一最大熵测度矛盾!

3. 如果$\mathcal{M}_{max}(X,T)=\{\mu\}$, 那么$\mu$ 必为遍历的. 如果$h_{\rm top}(T)=\infty$, 那么根据(*2*)即有结论, 如果$h(T)<\infty$, 那么根据命题*7.8.3-(3)*得到结论.

例 7.8.9 设(Σ_k,σ) 为转移系统, 那么σ 具有唯一最大熵测度, 且此测度为$(\frac{1}{k},\frac{1}{k},\dots,\frac{1}{k})$乘积测度.

证明. 我们已知拓扑熵$h_{\rm top}(\sigma)=\log k$. 下设$\mu\in\mathcal{M}(\sigma_k,\sigma)$ 使得$h_\mu(T)=\log k$. 设$\alpha=\{{}_0[0]_0,{}_0[1]_0,\dots,{}_0[k-1]_0\}$ 为生成子. 则

$$\log k=h_\mu(T)\le\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^{-i}\alpha)\le\frac{1}{n}\log k^n=\log k,\forall n\in\mathbb{N}.$$

于是$H_\mu(\bigvee_{i=0}^{n-1}T^{-i}\alpha)=\log k^n$, 所以$\bigvee_{i=0}^{n-1}T^{-i}\alpha$ 每个元素测度需为$\frac{1}{k^n}$, 即μ 为$(\frac{1}{k},\dots,\frac{1}{k})$乘积测度. □

例 7.8.10 设$T_A:X_A\to X_A$ 为*Markov* 转移系统, 其中$P,\overrightarrow{p}$ 是由不可约矩阵A 确定的随机矩阵和概率向量. 那么σ 具有唯一最大熵测度μ, 其中μ 由$P,\overrightarrow{p}$ 确定.

证明. 首先回顾一下符号. 设A为不可约非负矩阵, $\lambda,\vec{u},\vec{v}$ 为最大特征值以及对应的严格正值左右特征向量. 定义$k\times k$ 矩阵$P=[p_{ij}]$ 如下:

$$p_{ij}=\frac{a_{ij}v_j}{\lambda v_i},\forall i,j\in\{1,\dots,k\}$$

因为$A\vec{v}=\lambda\vec{v}$, 容易验证

$$0\le p_{ij}\le 1,\ 且\ \sum_{j=1}^k p_{ij}=\sum_{j=1}^k\frac{a_{ij}v_j}{\lambda v_i}=\frac{\sum_{j=1}^k a_{ij}v_j}{\lambda v_i}=1.$$

满足这种条件的P称为随机矩阵. 令$\vec{p}=(p_1,\dots,p_k)$为

$$p_i=u_iv_i,1\le i\le k.$$

正规化$\vec{u},\vec{v}$, 我们可以假设

$$\sum_{i=1}^k p_i=1.$$

即$\vec{p}$为概率向量. 对于$P,\vec{p}$我们有:

$$\sum_{i=1}^{k} p_i p_{ij} = \sum_{i=1}^{k} u_i v_i \frac{a_{ij} v_j}{\lambda v_i} = \sum_{i=1}^{k} \frac{a_{ij} u_i}{\lambda} v_j = u_j v_j = p_j, \forall j.$$

即

$$\vec{p}P = \vec{p}.$$

对应的测度μ 我们称之为**Parry 测度**. 我们要证明μ 为最大熵测度, 即$h_\mu(T) = \log\lambda$.

根据例子7.2.18 的公式

$$\begin{aligned}
h_\mu(T_A) &= -\sum_{i,j=0}^{k-1} u_i v_i \frac{a_{ij} v_j}{\lambda v_i} \log\left(\frac{a_{ij} v_j}{\lambda v_i}\right) \\
&= -\sum_{i,j=0}^{k-1} \frac{u_i a_{ij} v_j}{\lambda} \Big[\log a_{ij} + \log v_j - \log\lambda - \log v_i\Big] \\
&= 0 - \sum_{j=0}^{k-1} u_j v_j \log v_j + \log\lambda + \sum_{i=0}^{k-1} u_i v_i \log v_i \quad (\text{因为} a_{ij} \in \{0,1\}) \\
&= \log\lambda.
\end{aligned}$$

下证唯一性, 即μ 为唯一的具有最大熵的测度. 首先因为A 不可约性, μ 为遍历测度的. 于是如果$\mathcal{M}_{max}(X,T) \neq \{\mu\}$, 那么还会存在另一遍历的最大熵测度$m$. 因为$\mu \perp m$, 所以存在$E \in \mathcal{X}(X_A)$ 使得$\mu(E) = 0, m(E) = 1$.

令$\alpha = \{A_0, \ldots, A_{k-1}\}$ 为标准生成子, 即$A_j = {}_0[j]_0 \cap X_A, 0 \le j \le k-1$. 因为$\hat{\alpha}_{-(n-1)}^{n-1} \nearrow \mathcal{X}(X_A)$, 取$E_n \in \hat{\alpha}_{-(n-1)}^{n-1}$ 使得$(m+\mu)(E_n \Delta E) \to 0, n \to \infty$. 于是$\mu(E_n) \to 0, m(E_n) \to 1, n \to \infty$.

设$\eta_n = \{E_n, X \setminus E_n\}$, 则

$$\begin{aligned}
\log\lambda = h_m(T_A) &\le \frac{1}{2n-1} H_m(\alpha_0^{2n-2}) = \frac{1}{2n-1} H_m(\alpha_{-(n-1)}^{n-1}) \\
&\le \frac{1}{2n-1}\Big[H_m(\eta_n) + H_m(\alpha_{-(n-1)}^{n-1} | \eta_n)\Big] \\
&\le \frac{1}{2n-1}[-m(E_n)\log m(E_n) - (1-m(E_n))\log(1-m(E_n))] \\
&\quad + \frac{1}{2n-1}[m(E_n)\log\theta_n(E_n) + (1-m(E_n))\log\theta_n(X \setminus E_n)],
\end{aligned}$$

其中$\theta_n(B)$ 表示$\alpha_{-(n-1)}^{n-1} = \bigvee_{i=-(n-1)}^{n-1} T^{-i}\alpha$ 与$B \in \hat{\alpha}_{-(n-1)}^{n-1}$ 相交非空元素的个数. 于是

$$\log\lambda \le \frac{1}{2n-1}\left[m(E_n)\log\frac{\theta_n(E_n)}{m(E_n)} + (1-m(E_n))\log\frac{\theta_n(X \setminus E_n)}{1-m(E_n)}\right]. \tag{7.8.1}$$

如果$C \in \alpha_{-(n-1)}^{n-1} = \bigvee_{i=-(n-1)}^{n-1} T^{-i}\alpha$, 设

$$C = {}_{-(n-1)}[j_{-(n-1)} \cdots j_{n-1}]_{n-1} = \{x : (x_{-(n-1)} \cdots x_{n-1}) = (j_{-(n-1)} \cdots j_{n-1})\},$$

则根据$a_{j_pj_{p+1}}=1$,

$$\mu(C)=u_{j-(n-1)}v_{j-(n-1)}\prod_{p=-(n-1)}^{n-2}\frac{a_{j_pj_{p+1}}}{\lambda v_{j_p}}v_{j_{p+1}}=\frac{u_{j-(n-1)}v_{j_{n-1}}}{\lambda^{2n-1}}.$$

于是如果设

$$a=\min_{0\le i,j\le k-1}u_iu_j,\quad b=\max_{0\le i,j\le k-1}u_iu_j,$$

则

$$\frac{a}{\lambda^{2n-1}}\le\mu(C)\le\frac{b}{\lambda^{2n-1}},\ \forall C\in\bigvee_{i=-(n-1)}^{n-1}T^{-i}\alpha.$$

于是

$$\frac{a}{\lambda^{2n-1}}\theta_n(B)\le\mu(B)\le\frac{b}{\lambda^{2n-1}}\theta_n(B),\ \forall B\in\bigvee_{i=-(n-1)}^{n-1}T^{-i}\hat{\alpha}.$$

结合上式将$B=E_n,B=X\setminus E_n$ 用到(7.8.1) 我们就得到

$$\log\lambda\le\frac{1}{2n-1}\left[m(E_n)\log\frac{\mu(E_n)\lambda^{2n-1}}{am(E_n)}+(1-m(E_n))\log\frac{(1-\mu(E_n))\lambda^{2n-1}}{a(1-m(E_n))}\right].$$

整理得到

$$0\le m(E_n)\log\frac{\mu(E_n)}{am(E_n)}+(1-m(E_n))\log(1-\mu(E_n))-(1-m(E_n))\log\big(a(1-m(E_n))\big).$$

令$n\to\infty$, 得到$0\le-\infty+0+0$, 矛盾! 所以$\mathcal{M}_{max}(X_A,T_A)=\{\mu\}$. □

习　题

1. 证明：紧致度量空间上的任何可扩同胚存在最大熵测度.

§7.9　仿射映射的熵公式

§7.9.1　紧致交换群上的仿射

定理 7.9.1 设G 为紧致交换度量群, 设$T=a\cdot A:G\to G$ 为仿射变换, m 为G 的$Haar$ 测度. 那么

$$h_m(T)=h_m(A)=h_{\text{top}}(A)=h_{\text{top}}(T).$$

如果d 为G 上左不变度量, 那么

$$h_{\text{top}}(T)=\lim_{\varepsilon\to0}\limsup_{n\to\infty}-\frac{1}{n}\log m\left(\bigcap_{i=0}^{n-1}A^{-i}B(e,\varepsilon)\right),$$

其中e 为G 单位元, $B(e,\varepsilon)=\{x\in G:d(e,x)<\varepsilon\}$.

证明. 根据变分原理$h_m(T) \le h_{\text{top}}(T)$. 令

$$D_n(x,\varepsilon,T) = \bigcap_{k=0}^{n-1} T^{-k}B(T^k,\varepsilon),$$

即$D_n(x,\varepsilon,T) = \{y \in G : d_n(x,y) < \varepsilon\}$. 首先我们有

$$T^{-k}B(T^k,\varepsilon) = x \cdot (A^k B(e,\varepsilon)).$$

我们归纳证明上式. $k=0$ 根据d 的左不变性就有$B(x,\varepsilon) = xB(e,\varepsilon)$. 假设上式对$k$ 成立, 那么对于$k+1$ 有:

$$T^{-(k+1)}B(T^{k+1}x,\varepsilon) = T^{-1}(T^{-k}B(T^k(Tx),\varepsilon)) = T^{-1}(Tx \cdot A^{-k}B(e,\varepsilon)) = x \cdot (A^{-(k+1)}B(e,\varepsilon)).$$

于是我们就有

$$D_n(x,\varepsilon,T) = x \bigcap_{k=0}^{n-1} A^{-k}B(e,\varepsilon) = x \cdot D_n(e,\varepsilon,A).$$

尤其

$$m(D_n(x,\varepsilon,T)) = m(D_n(e,\varepsilon,A)).$$

设$\varepsilon > 0$, 设$\alpha = \{A_1,\ldots,A_k\}$ 为G 满足$\mathrm{diam}\alpha < \varepsilon$ 的剖分. 设$x \in \bigcap_{j=0}^{n-1} T^{-j}A_{i_j}$, 则

$$\bigcap_{j=0}^{n-1} T^{-j}A_{i_j} \subset x \cdot D_n(e,\varepsilon,A).$$

上式成立的原因: 如果$y \in \bigcap_{j=0}^{n-1} T^{-j}A_{i_j}$, 则$T^jx, T^jy \in A_{i_j}, 0 \le j \le n-1$, 尤其$d_n(x,y) < \varepsilon$. 所以$y \in D_n(x,\varepsilon,T) = xD_n(e,\varepsilon,A)$. 于是$m(\bigcap_{j=0}^{n-1} T^{-j}A_{i_j}) \le m(D_n(e,\varepsilon,A))$. 由此

$$\begin{aligned}
&\sum_{i_0,\ldots,i_{n-1}=1}^{k} m\left(\bigcap_{j=0}^{n-1} T^{-j}A_{i_j}\right) \log m\left(\bigcap_{j=0}^{n-1} T^{-j}A_{i_j}\right) \\
\le &\sum_{i_0,\ldots,i_{n-1}=1}^{k} m\left(\bigcap_{j=0}^{n-1} T^{-j}A_{i_j}\right) \log m\left(D_n(e,\varepsilon,A)\right) \\
= &\log m(D_n(e,\varepsilon,A)).
\end{aligned}$$

于是

$$\begin{aligned}
h_m \ge h_m(T,\alpha) &= \lim_{n\to\infty} H_m\left(\bigvee_{j=0}^{n-1} T^{-j}\alpha\right) \\
&\ge \limsup_{n\to\infty}\left[-\frac{1}{n}\log m(D_n(e,\varepsilon,A))\right].
\end{aligned}$$

由ε 任意性,

$$h_m \geq \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left[-\frac{1}{n}\log m(D_n(e,\varepsilon,A))\right].$$

设E 为具有最大个数的相对于T 的(n,ε) 分离集. 则

$$\bigcup_{x\in E} D_n(x,\varepsilon/2,T)=\bigcup_{x\in E} x\cdot D_n(e,\varepsilon/2,A)$$

为无交并. 所以

$$\mathrm{sr}(n,\varepsilon,X,T)\cdot m(D_n(e,\varepsilon/2,A))\leq 1.$$

由此

$$\mathrm{sr}(\varepsilon,X,T)\leq \limsup_{n\to\infty}\left[-\frac{1}{n}\log m(D_n(e,\varepsilon/2,A))\right].$$

令$\varepsilon\to 0$, 则

$$h_{\mathrm{top}}(T)\leq \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left[-\frac{1}{n}\log m(D_n(e,\varepsilon/2,A))\right]\leq h_m(T).$$

综上,

$$h_m(T)=h_{\mathrm{top}}(T)=\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left[-\frac{1}{n}\log m(D_n(e,\varepsilon,A))\right].$$

因为上面证明中没有依赖于a, 所以

$$h_m(A)=h_{\mathrm{top}}(A)=h_m(T)=h_{\mathrm{top}}(T)=\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left[-\frac{1}{n}\log m(D_n(e,\varepsilon,A))\right].$$

证毕! □

§7.9.2 熵的提升

我们后面将计算环面上面仿射的熵. 首先回顾一些记号. 每个自同态$A:(\mathbb{S}^1)^n\to(\mathbb{S}^1)^n$ 形如

$$A(z_1,\ldots,z_n)=(z_1^{a_{11}}z_2^{a_{12}}\ldots z_n^{a_{1n}},\ldots,z_1^{a_{n1}}z_2^{a_{n2}}\ldots z_n^{a_{nn}}),$$

其中$a_{ij}\in\mathbb{Z}$. 等价的, $A:\mathbb{T}^n\to\mathbb{T}^n$表达式为

$$A\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+\mathbb{Z}^n\right)=[a_{ij}]\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+\mathbb{Z}^n,$$

其中$[a_{ij}]\in\mathbb{Z}^{n\times n}$为方阵, 第$(i,j)$元素为$a_{ij}$. 以后如果$A:\mathbb{T}^n\to\mathbb{T}^n$为自同态, 那么用$[A]$表示对应的方阵, 而用$\widetilde{A}:\mathbb{R}^n\to\mathbb{R}^n$表示由矩阵$[A]$定义的线性映射. 设$\pi:\mathbb{R}^n\to\mathbb{T}^n=\mathbb{R}^n/\mathbb{Z}^n,\mathbf{x}\mapsto\mathbf{x}+\mathbb{Z}^n$ 为自然投射, 于是我们有交换图标:

$$\begin{tikzcd}\mathbb{R}^n \arrow[r, "\widetilde{A}"] \arrow[d, "\pi"'] & \mathbb{R}^n \arrow[d, "\pi"] \\ \mathbb{T}^n \arrow[r, "A"'] & \mathbb{T}^n.\end{tikzcd}$$

我们用$\|\cdot\|$ 表示$\mathbb{R}^n$ 上标准模, $\widetilde{d}$ 为相应的度量, 而用d 表示它诱导的环面度量:

$$d(\mathbf{x}+\mathbb{Z}^n,\mathbf{y}+\mathbb{Z}^n)=\inf_{\mathbf{v}\in\mathbb{Z}^d}\|\mathbf{x}-\mathbf{y}+\mathbf{v}\|.$$

d 为左不变且右不变的度量, 并且在π 映射下每个以$\mathbf{x}\in\mathbb{R}^n$ 为球心半径小于$1/4$ 的球等距映为$\mathbb{T}^n$ 上球心$\pi(\mathbf{x})$ 相同半径的球.

定理 7.9.2 设$(X,d),(\widetilde{X},\widetilde{d})$ 为度量空间, 设$\pi:\widetilde{X}\to X$ 为连续满射满足下面条件: 存在$\delta>0$ 使得对于任何$\widetilde{x}\in\widetilde{X}$

$$\pi|_{B(\widetilde{x},\delta)}:B(\widetilde{x},\delta)\to B(\pi(\widetilde{x}),\delta)$$

为等距满射. 那么对于满足$\pi\widetilde{T}=T\pi$ 的$T\in UC(X,d),\widetilde{T}\in UC(\widetilde{X},\widetilde{d})$, 我们有

$$h_d(T,d)=h_{\widetilde{d}}(\widetilde{T}).$$

此处h_d 表示度量d 下定义的$Bowen$ 熵.

证明. 设$\widetilde{K}$ 为$\widetilde{X}$ 的直径小于δ 的紧致子集, 那么根据δ 的取法$\pi(\widetilde{K})$ 也为X 直径小于δ 的紧致子集. 设$\varepsilon>0$ 使得$\varepsilon<\delta$, 且$\widetilde{d}(\widetilde{x},\widetilde{y})<\varepsilon$ 蕴含$\widetilde{d}(\widetilde{T}\widetilde{x},\widetilde{T}\widetilde{y})<\delta$.

设$\widetilde{E}\subset\widetilde{K}$ 为相对于$\widetilde{T}$ 的(n,ε) 分离集. 我们先说明$\pi(\widetilde{E})\subset\pi(\widetilde{K})$ 为相对于T 的(n,ε) 分离集. 设$\widetilde{x}\neq\widetilde{y}\in\widetilde{E}$. 则$\pi(\widetilde{x})\neq\pi(\widetilde{y})$. 设$i_0\in\{0,1,\dots,n-2\}$ 使得$\widetilde{d}(\widetilde{T}^i\widetilde{x},\widetilde{T}^i\widetilde{y})\le\varepsilon,i\le i$ 且$\widetilde{d}(\widetilde{T}^{i_0+1}\widetilde{x},\widetilde{T}^{i_0+1}\widetilde{y})>\varepsilon$. 根据$\varepsilon$ 的选取, $\widetilde{d}(\widetilde{T}^{i_0+1}\widetilde{x},\widetilde{T}^{i_0+1}\widetilde{y})<\delta$, 所以

$$d(T^{i_0+1}\pi(\widetilde{x}),T^{i_0+1}\pi(\widetilde{y}))=\widetilde{d}(\widetilde{T}^{i_0+1}\widetilde{x},\widetilde{T}^{i_0+1}\widetilde{y})>\varepsilon.$$

所以$\pi(\widetilde{E})$ 为T 的(n,ε) 分离集. 由此我们得到

$$\mathrm{sr}(n,\varepsilon,\widetilde{K},\widetilde{T})\le\mathrm{sr}(n,\varepsilon,\pi(\widetilde{K}),T).$$

下证反向不等式. 设E 为$\pi(\widetilde{K})$ 的相对于T 的(n,ε) 分离集. 令$\widetilde{E}=\pi^{-1}(E)\cap\widetilde{K}$. 因为对于$\widetilde{x},\widetilde{y}\in\widetilde{E}$ 以及i, 不等式$\widetilde{d}(\widetilde{T}^i\widetilde{x},\widetilde{T}^i\widetilde{y})\le\varepsilon$ 蕴含不等式$d(T^i\pi\widetilde{x},T^i\pi\widetilde{y})\le\varepsilon$, 我们有$\widetilde{E}$ 为$\widetilde{T}$ 的(n,ε) 分离集. 于是

$$\mathrm{sr}(n,\varepsilon,\widetilde{K},\widetilde{T})\ge\mathrm{sr}(n,\varepsilon,\pi(\widetilde{K}),T).$$

所以

$$\mathrm{sr}(n,\varepsilon,\widetilde{K},\widetilde{T})=\mathrm{sr}(n,\varepsilon,\pi(\widetilde{K}),T).$$

即$h_{\widetilde{d}}(\widetilde{T},\widetilde{K})=h_d(T,\pi(\widetilde{K}))$. 由此我们就有

$$h_d(T,d)=h_{\widetilde{d}}(\widetilde{T}).$$

证毕! □

推论 7.9.3 设$A:\mathbb{T}^n\to\mathbb{T}^n$ 为环面自同态, 那么$h_d(A)=h_{\widetilde{d}}(\widetilde{A})$, 其中$\widetilde{A}:\mathbb{R}^n\to\mathbb{R}^n$ 为诱导A 的线性映射, $\widetilde{d}$ 为$\mathbb{R}^n$ 上诱导$\mathbb{T}^n$ 上度量d 的度量.

§7.9.3 环面上仿射熵公式

根据上面命题, 我们接着需要处理$\mathbb{R}^n$ 上的线性映射.

引理 7.9.4 设$A:\mathbb{R}^n\to\mathbb{R}^n$ 为线性映射, m 为*Lebesgue* 测度, ρ 为$\mathbb{R}^n$ 上由模确定的度量. 那么

$$h_\rho(A)=\lim_{\varepsilon\to 0}\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,\varepsilon,A)),$$

其中$D_k(0,\varepsilon,A)=\bigcap_{i=0}^{k-1}A^{-i}B_\rho(e,\varepsilon)$.

$h_\rho(A)$ 不依赖于$\mathbb{R}^n$ 上模的选取.

证明. 因为有限维线性空间的模都是等价的, 它们诱导了一致等价的度量, 所以$h_\rho(A)$ 不依赖于$\mathbb{R}^n$ 上模的选取. 我们不妨假设ρ 为欧式度量.

对于线性映射A, 我们有$m(AB)=|\det A|m(B),\forall B\in\mathcal{X}(\mathbb{R}^n)$. 设$K\subset\mathbb{R}^n$ 为紧致子集, $m(K)>0$. 设F 为K 的(k,ε) 张成集, 则

$$K\subset\bigcup_{x\in F}D_k(x,2\varepsilon,A)=\bigcup_{x\in F}\left(x+D_k(0,2\varepsilon,A)\right).$$

于是$m(K)\le \mathrm{sp}(k,\varepsilon,K,A)m(D_k(0,2\varepsilon,A))$. 所以

$$h_\rho(T)\ge \mathrm{sp}(\varepsilon,K,A)\ge\limsup_{k\to\infty}\frac{1}{k}\log\frac{m(K)}{m(D_k(0,2\varepsilon,A))}=\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,2\varepsilon,A)).$$

由此

$$h_\rho(T)\ge \mathrm{sp}(\varepsilon,K,A)\ge\lim_{\varepsilon\to 0}\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,\varepsilon,A)).$$

设K_r 为$\mathbb{R}^n$ 以0 为中心边长为$2m$ 的方体, 设E 为K_r 的(k,ε) 分离集. 则$\bigcup_{x\in E}D_k(x,\varepsilon/2,A)$ 为K_r 的无交并, 并且

$$\bigcup_{x\in E}D_k(x,\varepsilon/2,A)=\bigcup_{x\in E}(x+D_k(0,\varepsilon/2,A))\subset K_{r+\varepsilon}.$$

于是

$$\mathrm{sr}(k,\varepsilon,K_r)\cdot m(D_k(0,\varepsilon/2,A))\le 2^n(r+\varepsilon)^n.$$

由此

$$\mathrm{sr}(\varepsilon,K_r,A)\le\limsup_{k\to\infty}\frac{1}{k}\log\frac{2^n(r+\varepsilon)^n}{m(D_n(0,\varepsilon/2,A))}=\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,\varepsilon/2,A)).$$

如果K 为任何紧致子集, 取r 使得$K\subset K_r$, 则

$$\mathrm{sr}(\varepsilon,K,A)\le \mathrm{sr}(\varepsilon,K_r,A)\le\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,\varepsilon/2,A)).$$

尤其

$$h_\rho(T)=\sup_K \lim_{\varepsilon\to 0}\mathrm{sr}(\varepsilon,K,A)\le \lim_{\varepsilon\to 0}\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,\varepsilon,A)).$$

综上,

$$h_\rho(T)=\lim_{\varepsilon\to 0}\limsup_{k\to\infty}-\frac{1}{k}\log m(D_k(0,\varepsilon,A)).$$

证毕! □

定理 7.9.5 设V 为n为向量空间, $A:V\to V$ 为线性映射, ρ 为V 上由模确定的度量. 那么

$$h_\rho(A)=\sum_{\{i:|\lambda_i|>1\}}\log|\lambda_i|,$$

其中$\lambda_1,\dots,\lambda_n$ 为A 全体特征值(其中允许λ_i 相等).

证明. 首先注意上面公式与ρ 选取无关. 根据特征值将V 分解为

$$V=E_1\oplus E_2,$$

其中$AE_1\subset E_1, AE_2\subset E_2$, $A_1=A|_{E_1}$ 的特征值绝对值大于1, $A_2=A|_{E_2}$ 的特征值绝对值小于等于1. 选取合适的基, 我们不妨设$V=\mathbb{R}^p, E_1=\mathbb{R}^{p_1}, E_2=\mathbb{R}^{p_2}$, 其中$p_1+p_2=p$. 设$\mathbb{R}^{p_1},\mathbb{R}^{p_2}$ Lebesgue 测度为m_1,m_2, 则$m=m_1\times m_2$ 为$\mathbb{R}^p$ 的Lebesgue 测度. 设$\|\cdot\|_i$ 为$\mathbb{R}^{p_i}$ 的欧式模, ρ_i 为相应的度量$(i=1,2)$. 则

$$\|(\mathbf{x}_1,\mathbf{x}_2)\|=\max\{\|\mathbf{x}_1\|_1,\|\mathbf{x}_2\|_2\},\ \rho((\mathbf{x}_1,\mathbf{x}_2),(\mathbf{y}_1,\mathbf{y}_2))=\max\{\rho_1(\mathbf{x}_1,\mathbf{y}_1),\rho_2(\mathbf{x}_2,\mathbf{y}_2)\}$$

为$\mathbb{R}^p$ 上的模与度量. 根据引理7.9.4,

$$h_\rho(A)=\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left[-\frac{1}{n}\log m_1(D_n(0,\varepsilon,A_1))-\frac{1}{n}\log m_2(D_n(0,\varepsilon,A_2))\right],$$

其中$D_n(0,\varepsilon,A_i)=\bigcap_{j=0}^{n-1}A_i^{-j}B_{\rho_i}(0,\varepsilon), i=1,2$. 根据$A_1,A_2$ 的定义,

$$m_1(D_n(0,\varepsilon,A_1))\le m_1(A_1^{-(n-1)}B_{\rho_1}(0,\varepsilon))=|\mathrm{det}A_1|^{-(n-1)}m_1(B_{\rho_1}(0,\varepsilon)),$$

$$m_2(D_n(0,\varepsilon,A_2))\le m_2(B_{\rho_2}(0,\varepsilon)).$$

所以

$$\begin{aligned}&-\frac{1}{n}\log m_1(D_n(0,\varepsilon,A_1))-\frac{1}{n}\log m_2(D_n(0,\varepsilon,A_2))\\ \ge&\left(\frac{n-1}{n}\right)\log|\det A_1|-\frac{1}{n}\log m_1(B_{\rho_1}(0,\varepsilon))-\frac{1}{n}\log m_2(B_{\rho_2}(0,\varepsilon))\end{aligned}$$

于是

$$h_\rho(A)\ge\log|\det A_1|=\sum_{\{i:|\lambda_i|>1\}}\log|\lambda_i|.$$

下证反向不等式. 根据Jordan 分解, 我们可以将V 分解为

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_k,$$

其中$AV_i \subset V_i$, $A_i = A|_{V_i}$ 上的特征值绝对值为$\tau_i, 1 \le i \le k$. 选取合适的基, 我们不妨设$V = \mathbb{R}^p$, $V_i = \mathbb{R}^{p_i}$, m_i 为$\mathbb{R}^{p_i}$ 上Lebsgue 测度, ρ_i 为$\mathbb{R}^{p_i}$ 上欧式度量$(1 \le i \le k)$. 于是

$$V = \mathbb{R}^p = \mathbb{R}^{p_1} \oplus \ldots \oplus \mathbb{R}^{p_k}$$

并且$m = m_1 \times \ldots \times m_k$ 为$\mathbb{R}^p$ 的Lebesgue 测度, $\rho((\mathbf{x_1}, \ldots, \mathbf{x_k}), (\mathbf{y}_1, \ldots, \mathbf{y}_k)) = \max\{\rho_i(\mathbf{x}_i, \mathbf{y}_i) : 1 \le i \le k\}, \mathbf{x}_i, \mathbf{y}_i \in \mathbb{R}^{p_i}, 1 \le i \le k$, 为$\mathbb{R}^p$ 的度量.

根据引理7.9.4,

$$h_\rho(A) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \sum_{i=1}^{k} -\frac{1}{n} \log m_i(D_n(0, \varepsilon, A_i)),$$

其中$D_n(0, \varepsilon, A_i) = \bigcap_{j=0}^{n-1} A_i^{-j} B_{\rho_i}(0, \varepsilon), i = 1, \ldots, k$. 下证对于$1 \le i \le k$,

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} -\frac{1}{n} \log m_i(D_n(0, \varepsilon, A_i)) \le \max\{0, p_i \log \tau_i\},$$

从而完成证明.

取定$i \in \{1, 2, \ldots, k\}$, $\delta > 0$. 设$\| \cdot \|_i$ 为$\mathbb{R}^{p_i}$ 的欧式模, 定义新的模为

$$|||x||| = \sum_{n=0}^{\infty} \frac{\|A^n x\|_i}{(\tau_i + \delta)^n}.$$

因为

$$\sqrt[n]{\frac{\|A^n x\|_i}{(\tau_i + \delta)^n}} \le \frac{\|A^n\|_i^{\frac{1}{n}} \|x\|_i^{\frac{1}{n}}}{\tau_i + \delta} \to \frac{\tau_i}{\tau_i + \delta}, n \to \infty,$$

所以$|||x|||$ 可定义. 又

$$\|x\|_i \le |||x||| \le \|x\|_i \sum_{n=0}^{\infty} \frac{\|A^n\|_i}{(\tau_i + \delta)^n} = c^{-1} \|x\|_i,$$

其中$c = (\sum_{n=0}^{\infty} \frac{\|A^n\|_i}{(\tau_i + \delta)^n})^{-1}$, 所以$\| \cdot \|_i$ 与$||| \cdot |||$ 等价. 另外根据定义易见

$$|||Ax||| \le (\tau_i + \delta) \, ||| \, x \, ||| \, .$$

用$\hat{B}(x, \varepsilon)$ 表示$||| \cdot |||$ 诱导度量的球, 根据上式我们就有

$$A^{-j} \hat{B}(0, \varepsilon) \supset \hat{B}(0, \frac{\varepsilon}{(\tau_i + \delta)^j}), \forall j \in \mathbb{N}.$$

用$\hat{D}_n(0, \varepsilon, A_i)$ 表示$\mathbb{R}^{p_i}$ 上由$||| \cdot |||$ 诱导度量下的Bowen 球, 我们有

$$\hat{D}_n(0, \varepsilon, A_i) \supset \begin{cases} \hat{B}(0, \frac{\varepsilon}{(\tau_i + \delta)^{n-1}}) \supset B(0, \frac{c\varepsilon}{(\tau_i + \delta)^{n-1}}), & \text{如果} \tau_i + \delta \ge 1; \\ \hat{B}(0, \varepsilon) \supset B(0, c\varepsilon), & \text{如果} \tau_i + \delta \le 1. \end{cases}$$

注意上面$B(x,\varepsilon)$ 为标准欧式度量的球. 于是

$$m_i\Big(\hat{D}_n(0,\varepsilon,A_i)\Big) \geq \begin{cases} \frac{1}{(\tau_i+\delta)^{p_i(n-1)}} m_i(B(0,c\varepsilon)), & \text{如果}\tau_i+\delta \geq 1; \\ m_i(B(0,c\varepsilon)), & \text{如果}\tau_i+\delta \leq 1. \end{cases}$$

于是

$$\limsup_{n\to\infty} -\frac{1}{n}\log m_i(\hat{D}_n(0,\varepsilon,A_i)) \leq \max\{0, p_i\log(\tau_i+\delta)\}.$$

因为δ 维任意取的, 所以

$$\limsup_{n\to\infty} -\frac{1}{n}\log m_i(\hat{D}_n(0,\varepsilon,A_i)) \leq \max\{0, p_i\log(\tau_i)\}.$$

这样我们完成了整个证明. □

综合上面得到的结论, 我们容易得到下面关于环面的结果:

定理 7.9.6 *设$T=a\cdot A:\mathbb{T}^n\to\mathbb{T}^n$ 为仿射变换($a\in\mathbb{T}^n$, A 为满自同态), m 为$\mathbb{T}^n$ 的Haar 测度. 那么. 那么*

$$h_{\text{top}}(T)=h_m(T)=h_m(A)=\sum_{\{i:|\lambda_i|>1\}}\log|\lambda_i|,$$

其中$\lambda_1,\dots,\lambda_n$ 为A 对应矩阵$[A]$ 的全体特征值(其中允许λ_i 相等).

§7.10 零熵与Pinsker σ 代数

以下我们介绍一个重要的概念: Pinsker σ 代数. 一个保测系统的Pinsker σ 代数可以认为是该系统具有零熵的最大的子σ代数. Pinsker σ 代数在熵的分析中是一个不可缺少的工具.

§7.10.1 Pinsker σ代数

定义 7.10.1 *设$(X,\mathcal{X},\mu,T)$ 为保测系统, 包含$\bigcup\{\hat{\xi}:h_\mu(T,\xi)=0, H_\mu(\xi)<\infty\}$ 的最小σ代数称为***Pinsker σ 代数***, 记为$P_\mu(T)$ 或P_μ.*

注记 7.10.2 *容易验证$\bigcup\{\hat{\xi}:h_\mu(T,\xi)=0, H_\mu(\xi)<\infty\}$ 为代数.*

定理 7.10.3 *设$(X,\mathcal{X},\mu,T)$为保测系统. 则*

1. $P_\mu(T)=\{A\in\mathcal{X}: h_\mu(T,\{A,A^c\})=0\}$
2. $P_\mu(T)$ *为*$\mathcal{X}$ *的子*σ *代数;*
3. *有限可测剖分*$\alpha\subset P_\mu(T)$ *当且仅当*$h_\mu(T,\alpha)=H_\mu(\alpha|\alpha^-)=0$ *当且仅当*$\alpha\subseteq\alpha^-$*;*
4. $T^{-1}P_\mu(T)=P_\mu(T)(\mod \mu)$*;*

5. 对$k \geq 1$, $P_\mu(T) = P_\mu(T^k)$. 如果T 又是可逆的, 则$P_\mu(T) = P_\mu(T^{-1})$.

证明. 1. 显然$\emptyset, X \in P_\mu(T)$. 设$A, B \in P_\mu(T)$. 因$(A^c)^c = A$, $A^c \in P_\mu(T)$. 注意到$\{A, A^c\} \vee \{B, B^c\} \succeq \{A \cup B, (A \cup B)^c\}$, 我们有

$$\begin{aligned} h_\mu(T, \{A \cup B, (A \cup B)^c\}) &\leq h_\mu(T, \{A, A^c\} \vee \{B, B^c\}) \\ &\leq h_\mu(T, \{A, A^c\}) + h_\mu(T, \{B, B^c\}) = 0. \end{aligned}$$

因此$A \cup B \in P_\mu(T)$.

设$A_i \in P_\mu(T)$, $i \in \mathbb{N}$, 现在, 对$k \in \mathbb{N}$我们有

$$\begin{aligned} h_\mu(T, \{\bigcup_{i=1}^\infty A_i, (\bigcup_{i=1}^\infty A_i)^c\}) &\leq h_\mu(T, \bigvee_{i=1}^k \{A_i, A_i^c\}) + H_\mu(\{\bigcup_{i=1}^\infty A_i, (\bigcup_{i=1}^\infty A_i)^c\} | \bigvee_{i=1}^k \{A_i, A_i^c\}) \\ &\leq \sum_{i=1}^k h_\mu(T, \{A_i, A_i^c\}) + H_\mu(\{\bigcup_{i=1}^\infty A_i, (\bigcup_{i=1}^\infty A_i)^c\} | \bigvee_{i=1}^k \{A_i, A_i^c\}) \\ &= H_\mu(\{\bigcup_{i=1}^\infty A_i, (\bigcup_{i=1}^\infty A_i)^c\} | \bigvee_{i=1}^k \{A_i, A_i^c\}). \end{aligned}$$

在上述不等式中令$k \to +\infty$ 并利用Martingale 定理我们有

$$h_\mu(T, \{\bigcup_{i=1}^\infty A_i, (\bigcup_{i=1}^\infty A_i)^c\}) \leq H_\mu(\{\bigcup_{i=1}^\infty A_i, (\bigcup_{i=1}^\infty A_i)^c\} | \bigvee_{i=1}^\infty \{A_i, A_i^c\}) = 0.$$

因此$\cup_{i=1}^\infty A_i \in P_\mu(T)$. 综上所述我们知$P_\mu(T)$为$\mathcal{X}$ 的子σ代数.

2. 设$\alpha = \{A_1, A_2, \cdots, A_k\}$. 注意到$\vee_{j=1}^k \{A_j, A_j^c\} \succeq \alpha \succeq \{A_i, A_i^c\}$, 我们有$\alpha \subseteq P_\mu(T)$当且仅当$h_\mu(T, \alpha) = H_\mu(\alpha | \alpha^-) = 0$当且仅当$\alpha \subseteq \alpha^-$.

3. 由于对$A \in \mathcal{X}$,

$$h_\mu(T, \{A, A^c\}) = h_\mu(T, T^{-1}\{A, A^c\}) = h_\mu(T, \{T^{-1}A, (T^{-1}(A))^c\}),$$

我们得到$T^{-1}P_\mu(T) \subseteq P_\mu(T)$. 反之, 设$A \in P_\mu(T)$和$\alpha = \{A, A^c\}$. 则$\alpha \subset \alpha^-$. 因$\alpha \subset P_\mu(T)$, 我们有

$$T^{-i}\alpha \subset T^{-i}P_\mu(T) \subseteq T^{-(i-1)}P_\mu(T) \subset \cdots \subset T^{-1}P_\mu(T)$$

对每个$i \in \mathbb{N}$成立. 因$P_\mu(T)$为$\mathcal{X}$的子σ代数, $T^{-1}P_\mu(T)$也为$\mathcal{X}$的子σ代数. 再注意到α^-为$\mathcal{X}$的包含所有$T^{-i}\alpha, i \in \mathbb{N}$的最小的子$\sigma$代数, $\alpha^- \subset T^{-1}P_\mu(T)$. 这已经说明$A \in \alpha \subset \alpha^- \subset T^{-1}P_\mu(T)$. 因$A$是任意的, $P_\mu(T) \subseteq T^{-1}P_\mu(T)$.

4. 对X的任意有限可测剖分α和$k \in \mathbb{N}$, 首先我们有$h_\mu(T^k, \alpha_0^{k-1}) = kh_\mu(T, \alpha)$(参见命题7.1.20 的证明). 进而我们不难算得$\frac{1}{k}h_\mu(T, \alpha) \leq h_\mu(T^k, \alpha) \leq kh_\mu(T, \alpha)$. 这说明$\alpha \subset P_\mu(T)$当且仅当$\alpha \subset P_\mu(T^k)$. 因此$P_\mu(T) = P_\mu(T^k)$. 最后由注记7.1.21, 我们有$P_\mu(T) = P_\mu(T^{-1})$. □

注记 7.10.4 设$(X,\mathcal{X},\mu,T)$为保测系统. 根据上面命题的3, 我们知道

1. 设α 为有限可测剖分. $h_\mu(T,\alpha)=0$ 当且仅当$\hat{\alpha}\subset\hat{\alpha}^-$, 即过去决定现在.
2. $h_\mu(T)=0$ 当且仅当对于任何可数可测剖分α, $H(\alpha)<\infty$, 我们有$\hat{\alpha}\subset\hat{\alpha}^-$.
3. 如果$h_\mu(T)=0$, 那么$T^{-1}\mathcal{X}\underset{\mu}{=}\mathcal{X}$, 即$T$ 几乎处处可逆. 这个事实的证明如下: 设$B\in\mathcal{X}$, 令$\mathcal{A}=\{\emptyset,B,B^c,X\}$, 那么由$h_\mu(T)=0$, 我们有$\mathcal{A}\subset\vee_{i=1}^{\infty}T^{-i}\mathcal{A}\subset T^{-1}\mathcal{X}$. 由$B$ 任意, 我们有$\mathcal{X}\subset T^{-1}\mathcal{X}$.
4. 设$h_\mu(T)=0$. 如果$\mathcal{F}\subset\mathcal{X}$ 为子σ 代数且$T^{-1}\mathcal{F}\underset{\mu}{\subset}\mathcal{F}$, 那么$T^{-1}\mathcal{F}\underset{\mu}{=}\mathcal{F}$.

命题 7.10.5 设$(X,\mathcal{X},\mu,T)$为保测系统, η 为有限可测剖分且满足$H_\mu(\eta)<\infty$, $\hat{\eta}\subset\mathcal{P}_\mu(T)$. 那么$h_\mu(T,\eta)=0$.

证明. 由于$\mathcal{G}=\bigcup\{\hat{\xi}:h_\mu(T,\xi)=0,H_\mu(\xi)<\infty\}$ 为生成$\mathcal{P}_\mu(T)$ 的代数, 存在有限递增代数列$\{\xi_n\}_{n=1}^{\infty}$ 使得$\xi_n\subset\mathcal{G}$, 并且$\xi_n\nearrow\mathcal{P}_\mu(T)$.

但是

$$h_\mu(T,\eta)\le h_\mu(T,\xi_n)+H_\mu(\eta|\xi_n)=H(\eta|\xi_n).$$

令$n\to\infty$, 我们就有

$$h_\mu(T,\eta)\le H_\mu(\eta|\mathcal{P}_\mu(T))=0.$$

证毕! □

于是, Pinsker σ 代数可以视为最大有零熵的σ 代数.

§7.10.2 完全正熵

定义 7.10.6 我们将σ 代数P_μ 称为$(X,\mathcal{X},\mu,T)$ 的**Pinsker σ代数**. 当$(X,\mathcal{X},\mu,T)$ 为可逆*Lebesgue*系统时, 由P_μ 可以确定$(X,\mathcal{X},\mu,T)$ 的因子系统$\pi:(X,\mathcal{X},\mu,T)\to(Z,\mathcal{Z},\nu,S)$ (即$P_\mu=\pi^{-1}(\mathcal{Z})$), 我们称*Lebesgue* 系统$(Z,\mathcal{Z},\nu,S)$ 为$(X,\mathcal{X},\mu,T)$的**Pinsker 因子**.

定义 7.10.7 一个保测系统$(X,\mathcal{X},\mu,T)$ 称为**完全正熵系统**, 如果$P_\mu(T)=\{\emptyset,X\}$. 这等价于$(X,\mathcal{X},\mu,T)$ 的每个非平凡因子系统有正熵.

设$\alpha=\{A_1,A_2,\cdots,A_n\}$为$(X,\mathcal{X},\mu,T)$ 的可测剖分. 如果对每个$i\in\mathbb{N}$有$\mu(A_i)<1$, 则称α为**测度非平凡的剖分**.我们不难看出

命题 7.10.8 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 则以下性质彼此等价:

1. $(X,\mathcal{X},\mu,T)$ 为完全正熵系统;
2. X 的每个由两个元素构成的测度非平凡的剖分有正熵;
3. X 的每个测度非平凡的有限剖分有正熵.

§7.10.3 Pinsker公式

引理 7.10.9 设$(X,\mathcal{X},\mu,T)$为可逆保测系统, α,β,γ为X的有限可测剖分.则

1. 如果$\beta\preceq\alpha$ 或$\alpha\preceq\beta$, 那么

$$\lim_{n\to\infty}\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\beta^-)=H_\mu(\alpha|\alpha^-);$$

2. 如果$\alpha\preceq\beta$, 那么

$$\lim_{n\to\infty}H_\mu(\alpha|\beta^-\vee T^{-n}\gamma^-)=H_\mu(\alpha|\beta^-).$$

证明. (i). 首先假设$\beta\preceq\alpha$. 则$T^{-n}(\beta^-\vee\bigvee_{i=0}^{n-1}T^i\alpha)\nearrow\alpha^-$, 因此

$$H_\mu(\alpha|T^{-n}(\beta^-\vee\bigvee_{i=0}^{n-1}T^i\alpha))\to H_\mu(\alpha|\alpha^-).$$

注意到

$$\begin{aligned}H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\beta^-)&=H_\mu(\alpha|\beta^-)+H_\mu(T\alpha|\alpha\vee\beta^-)+\cdots+H_\mu(T^{n-1}\alpha|\bigvee_{i=0}^{n-2}T^i\alpha\vee\beta^-)\\&=H_\mu(\alpha|\beta^-)+H_\mu(\alpha|T^{-1}(\alpha\vee\beta^-))+\cdots\\&\quad+H_\mu(\alpha|T^{-(n-1)}(\bigvee_{i=0}^{n-2}T^i\alpha\vee\beta^-))\ (\text{由命题7.1.9 (6)}),\end{aligned}$$

因此

$$\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\beta^-)\to H_\mu(\alpha|\alpha^-).$$

(ii) 其次, 假设$\alpha\preceq\beta$.则一方面

$$\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\beta^-)\le\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\alpha^-)\to H_\mu(\alpha|\alpha^-)\ \ (\text{由(i)}).$$

另一方面,

$$\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\beta^-)=\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\beta|\beta^-)-\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i\beta|\bigvee_{i=0}^{n-1}T^i\alpha\vee\beta^-).$$

在上式中令$n \to \infty$我们得到

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^i\alpha|\beta^-) \\
&\geq H_\mu(\beta|\beta^-) - \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^i\beta|\bigvee_{i=0}^{n-1} T^i\alpha \vee \alpha^-) \\
&= \lim_{n\to\infty} \left(\frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^i\beta|\alpha^-) - \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^i\beta|\bigvee_{i=0}^{n-1} T^i\alpha \vee \alpha^-) \right) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^i\alpha|\alpha^-) = H_\mu(\alpha|\alpha^-).
\end{aligned}$$

(iii) 以下我们证明(2), 假设$\alpha \preceq \beta$. 首先

$$H_\mu(\alpha|\beta^- \vee T^{-n}\gamma^-) = H_\mu(\beta|\beta^- \vee T^{-n}\gamma^-) - H_\mu(\beta|\alpha \vee \beta^- \vee T^{-n}\gamma^-). \tag{7.10.1}$$

再考虑等式

$$\begin{aligned}
H_\mu(\bigvee_{i=0}^{n-1} T^i\beta|\beta^- \vee \gamma^-) &= H_\mu(\beta|\beta^- \vee \gamma^-) + H_\mu(T\beta|T\beta^- \vee \gamma^-) \\
&\quad + \cdots + H_\mu(T^n\beta|T^n\beta^- \vee \gamma^-) \\
&= H_\mu(\beta|\beta^- \vee \gamma^-) + H_\mu(\beta|\beta^- \vee T^{-1}\gamma^-) \\
&\quad + \cdots + H_\mu(\beta|\beta^- \vee T^{-n}\gamma^-) \text{ (由命题7.1.9 (6))}.
\end{aligned}$$

在上式中让$n \to \infty$并利用前面的讨论, 我们得到

$$H_\mu(\beta|\beta^-) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^i\beta|\beta^- \vee \gamma^-) = \lim_{n\to\infty} H_\mu(\beta|\beta^- \vee T^{-n}\gamma^-).$$

将此应用到等式(7.10.1)可得

$$\begin{aligned}
&\lim_{n\to\infty} H_\mu(\alpha|\beta^- \vee T^{-n}\gamma^-) \\
&= H_\mu(\beta|\beta^-) - \lim_{n\to\infty} H_\mu(\beta|\alpha \vee \beta^- \vee T^{-n}\gamma^-) \\
&\geq H_\mu(\beta|\beta^-) - H_\mu(\beta|\alpha \vee \beta^-) = H_\mu(\alpha|\beta^-).
\end{aligned}$$

反向的不等式是明显地. □

定理 7.10.10 (Pinsker公式) 设$(X, \mathcal{X}, \mu, T)$ 为可逆的保测系统, α, β 为X 的有限可测剖分. 则

$$h_\mu(T, \alpha \vee \beta) = h_\mu(T, \beta) + H_\mu(\alpha|\beta^T \vee \alpha^-),$$

其中$\beta^T = \bigvee_{n=-\infty}^{\infty} T^{-n}\beta$.

证明. 首先我们有

$$\begin{aligned}&\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^i(\alpha\vee\beta)|\alpha^-\vee\beta^-)\\&=\frac{1}{n}\{H_\mu(\bigvee_{i=0}^{n-1}T^i\beta|\alpha^-\vee\beta^-)+H_\mu(\bigvee_{i=0}^{n-1}T^i\alpha|\alpha^-\vee\beta^-\vee\bigvee_{i=0}^{n-1}T^i\beta)\}\\&=\frac{1}{n}\{H_\mu(\bigvee_{i=0}^{n-1}T^i\beta|\alpha^-\vee\beta^-)+\sum_{k=0}^{n-1}H_\mu(\alpha|\alpha^-\vee\beta^-\vee\bigvee_{i=0}^{k-1}T^i\beta)\}.\end{aligned}$$

在上式中让$n\to\infty$, 利用引理7.10.9的性质(1) 和定理7.2.5的(1) 我们可得

$$h_\mu(T,\alpha\vee\beta)=h_\mu(T,\beta)+H_\mu(\alpha|\beta^T\vee\alpha^-).$$

证毕! □

我们不难得到引理7.10.9和定理7.10.10 相对于一个严格T不变的子σ代数的相对化版本.例如:

定理 7.10.11 (相对的Pinsker公式) 设$(X,\mathcal{X},\mu,T)$为可逆的保测系统, α,β为X的有限可测剖分和$\mathcal{A}$为$\mathcal{X}$严格T不变的子σ代数(即$T^{-1}\mathcal{A}=\mathcal{A}$), 则

$$H_\mu(\alpha\vee\beta|\alpha^-\vee\beta^-\vee\mathcal{A})=H_\mu(\beta|\beta^-\vee\mathcal{A})+H_\mu(\alpha|\alpha^-\vee\beta^T\vee\mathcal{A}).$$

设$(X,\mathcal{X},\mu,T)$为保测系统, 我们用$\mathcal{P}_X$表示X的全体有限可测剖分.

定理 7.10.12 设$(X,\mathcal{X},\mu,T)$为可逆保测系统, 则

$$P_\mu(T)=\bigvee_{\beta\in\mathcal{P}_X}\bigcap_{n=0}^{\infty}T^{-n}\beta^-.$$

证明. 选取X递增的有限剖分序列$\xi_k\nearrow P_\mu(T)$.因$h_\mu(T,\xi_k)=H_\mu(\xi_k|\xi_k^-)=0$, 所以$\xi_k\subset\xi_k^-$. 进而

$$\xi_k\subset\xi_k^-=\bigcap_{n=0}^{\infty}T^{-n}\xi_k^-\subset\bigvee_{\beta\in\mathcal{P}_X}\bigcap_{n=0}^{\infty}T^{-n}\beta^-.$$

即

$$P_\mu(T)=\bigvee_{k=1}^{+\infty}\xi_k\subset\bigvee_{k=1}^{+\infty}\bigcap_{n=0}^{+\infty}T^{-n}\xi_k^-\subset\bigvee_{\beta\in\mathcal{P}_X}\bigcap_{n=1}^{\infty}T^{-n}\beta^-.$$

另一方面, 固定$\xi\in\mathcal{P}_X$, 设$\eta\in\mathcal{P}_X$满足$\eta\subset\bigcap_{n=0}^{+\infty}T^{-n}\xi^-$, 则$\eta^T\subset\xi^-$ 且

$$H_\mu(\xi\vee\eta|\xi^-\vee\eta^-)\le H_\mu(\xi\vee\eta|\xi^-)=H_\mu(\xi|\xi^-),$$

$$\begin{aligned}H_\mu(\xi\vee\eta|\xi^-\vee\eta^-)&=h_\mu(T,\xi\vee\eta)=h_\mu(T,\xi)+H_\mu(\eta|\xi^T\vee\eta^-)\\&\ge h_\mu(T,\xi)=H_\mu(\xi|\xi^-).\end{aligned}$$

结合上面两个不等式, 我们有

$$\begin{aligned}H_\mu(\xi \vee \eta|\xi^- \vee \eta^-) = H_\mu(\xi|\xi^-) &= H_\mu(\xi|\xi^- \vee \eta^T) + H_\mu(\eta|\eta^-)\\ &= H_\mu(\xi|\xi^-) + H_\mu(\eta|\eta^-).\end{aligned}$$

这说明$H_\mu(\eta|\eta^-) = 0$, 即$\eta \subset P_\mu(T)$.从而对每个$\xi \in \mathcal{P}_X$我们有

$$P_\mu(T) \supset \bigcap_{n=0}^{+\infty} T^{-n}\xi^-,$$

这就推得

$$P_\mu(T) = \bigvee_{\xi\in\mathcal{P}_X} \bigcap_{n=0}^{+\infty} T^{-n}\xi^-.$$

□

定理 7.10.13 设$(X, \mathcal{X}, \mu, T)$为可逆保测系统, $\xi \in \mathcal{P}_X$和$\mathcal{A}$为$\mathcal{X}$ 严格T不变的子σ代数. 则

$$H_\mu(\xi|\xi^- \vee \mathcal{A}) = H_\mu(\xi|\xi^- \vee P_\mu(T) \vee \mathcal{A}).$$

特别地,

$$h_\mu(T, \xi) = H_\mu(\xi|\xi^-) = H_\mu(\xi|\xi^- \vee P_\mu(T)).$$

证明. 任取有限可测剖分$\eta \subset P_\mu(T)$, 则由定理7.10.11可得

$$\begin{aligned}H_\mu(\xi \vee \eta|\xi^- \vee \eta^- \vee \mathcal{A}) &= H_\mu(\eta|\eta^- \vee \xi^T \vee \mathcal{A}) + H_\mu(\xi|\xi^- \vee \mathcal{A})\\ &= H_\mu(\xi|\xi^- \vee \eta^- \vee \mathcal{A}) + H_\mu(\eta|\xi \vee \xi^- \vee \eta^- \vee \mathcal{A}).\end{aligned}$$

即

$$H_\mu(\xi|\xi^- \vee \eta^- \vee \mathcal{A}) = H_\mu(\xi|\xi^- \vee \mathcal{A}).$$

现在选取满足$\eta_n \nearrow P_\mu(T)$的有限可测剖分序列η_n, 然后利用定理7.10.3 (3)便可得到

$$H_\mu(\xi|\xi^- \vee P_\mu(T) \vee \mathcal{A}) = H_\mu(\xi|\xi^- \vee \mathcal{A}).$$

证毕! □

定理 7.10.14 设$(X, \mathcal{X}, \mu, T)$ 为可逆保测系统, $\xi \in \mathcal{P}_X$. 则

$$\lim_{k\to\infty} h_\mu(T^k, \xi) = H_\mu(\xi|P_\mu(T)).$$

证明. 使用定理7.10.3 (4)和定理7.10.13, 一方面

$$\limsup_{k\to\infty} h_\mu(T^k, \xi) = \limsup_{k\to\infty} H_\mu(\xi|\bigvee_{j=1}^{+\infty} T^{-kj}\xi \vee P_\mu(T)) \le H_\mu(\xi|P_\mu(T)).$$

另一方面, 我们有

$$\liminf_{k\to\infty} h_\mu(T^k,\xi) = \liminf_{k\to\infty} H_\mu(\xi|\bigvee_{j=1}^{+\infty} T^{-kj}\xi) \geq \liminf_{k\to\infty} H_\mu(\xi|T^{-k}\xi^-)$$
$$\geq H_\mu(\xi|\bigcap_{k=1}^{+\infty} T^{-k}\xi^-) \geq H_\mu(\xi|P_\mu(T)) \quad (\text{由定理7.10.12}).$$

证毕! □

习 题

1. 设$(X,\mathcal{X},\mu,T)$ 为保测系统, 证明：如果$T^{-1}\mathcal{X} \neq \mathcal{X} \pmod{\mu}$, 则$h_\mu(T) > 0$.
2. 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 如果α, β 为X 的两个有限可测剖分, 则
$$h_\mu(T,\alpha) \leq h_\mu(T,\beta) + H_\mu(\alpha|\beta \vee P_\mu(T)).$$
3. 设$\mathbb{S}^1$ 为复平面的单位圆周, $p>1$, $T_p : \mathbb{S}^1 \to \mathbb{S}^1$ 满足$T_p(z) = z^p$ 以及$m_{\mathbb{S}^1}$ 为$\mathbb{S}^1$ 上的Haar 测度, 证明：$h_{m_{\mathbb{S}^1}}(T_p) > 0$ (实际上$h_{\mathbb{S}^1}(T_p) = \log p$).
4. 设$(X,\mathcal{X},\mu,T)$ 为完全正熵系统, 则它是遍历系统且当$(X,\mathcal{X},\mu,T)$ 不为平凡系统时, $\mu(\{x\}) = 0$ 对$x \in X$.
5. 证明相对的Pinsker 公式–定理7.10.11.
6. 设$(X,\mathcal{X},\mu,T)$ 为可逆的保测系统, α 为X 的有限可测剖分以及$\mathcal{A}$ 为$\mathcal{X}$ 的子σ 代数. 证明：如果$T^{-1}\mathcal{A} \subset \mathcal{A}$, 则$H_\mu(\alpha|\alpha^- \vee P_\mu(T) \vee \mathcal{A}) = H_\mu(\alpha|\alpha^- \vee \mathcal{A})$.

§7.11 Kolmogorov系统（续）

本小节, 我们将介绍一类重要的正熵保测系统, 它是1958 年Kolmogorov[133] 首先引入的. Rohlin 和Sinai [169]将这类保测系统称为Kolmogorov 系统. 在第四章中我们已经研究过这类系统，在这里我们从熵理论研究它.

§7.11.1 Kolmogorov系统

首先回顾下Kolmogorov系统的定义. 设$(X,\mathcal{X},\mu,T)$为可逆保测系统. 如果存在$\mathcal{X}$的子σ代数$\mathcal{K}$使得:

$$\mathcal{K} \subset T\mathcal{K}, \ \bigvee_{n=0}^{+\infty} T^n\mathcal{K} = \mathcal{X} \text{ 且 } \bigcap_{n=0}^{\infty} T^{-n}\mathcal{K} = \{\emptyset, X\}.$$

则称$(X,\mathcal{X},\mu,T)$为**Kolmogorov系统**或简称为**K系统** .

引理 7.11.1 *每个测度K系统$(X,\mathcal{X},\mu,T)$是遍历系统且当$(X,\mathcal{X},\mu,T)$不为平凡系统时, $\mu(\{x\}) = 0$对$x \in X$.*

证明. 设$A \in \mathcal{X}$为T-不变, $\epsilon > 0$和$m \in \mathbb{N}$. 则存在$n_0 \geq 0$和子集$B \in T^{n_0}\mathcal{K}$使得$\mu(A\Delta B) < \epsilon$. 从而$T^{-(n_0+m)}B \in T^{-m}\mathcal{K}$并且$\mu(A\Delta T^{-(m+n_0)}B) < \epsilon$. 因这对任意$\epsilon > 0$成立, 我们有$A \in T^{-m}\mathcal{K}$. 进而再由$m$ 的任意性知$A \in \cap_{n=0}^{\infty} T^{-n}\mathcal{K} = \{\emptyset, X\}$. 这说明$(X,\mathcal{X},\mu,T)$为遍历系统. 显然当$(X,\mathcal{X},\mu,T)$不为平凡系统时, $(X,\mathcal{X},\mu,T)$与有限系统不测度同构. 因此$\mu(\{x\}) = 0$对$x \in X$. □

§7.11.2 Rohlin-Sinai定理

我们将以下引理留作习题:

引理 7.11.2 设$(X,\mathcal{X},\mu,T)$为保测系统, $k \in \mathbb{N}$. 则对任意$\epsilon > 0$ 存在$\delta = \delta(\epsilon,k) > 0$ 使得如果可测剖分$\alpha = \{A_1,A_2,\cdots,A_k\}$和$\beta = \{B_1,B_2,\cdots,B_k\}$满足$\sum_{j=1}^k \mu(A_j\Delta B_j) < \delta$, 那么$H_\mu(\alpha|\beta) + H_\mu(\beta|\alpha) < \epsilon$. 进而

$$|h_\mu(T,\alpha) - h_\mu(T,\beta)| \leq H_\mu(\alpha|\beta) + H_\mu(\beta|\alpha) < \epsilon.$$

定理 7.11.3 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统, $\mathcal{A}$ 为$\mathcal{X}$的子σ 代数. 如果$T^{-1}\mathcal{A} \subseteq \mathcal{A}$ 且$\lim_{n\to+\infty} T^n\mathcal{A} \nearrow \mathcal{X}$, 则

$$P_\mu(T) \subseteq \mathcal{A}_{-\infty},$$

其中$\mathcal{A}_{-\infty} = \bigcap_{n=1}^{\infty} T^{-n}\mathcal{A}$.

证明. 首先, 我们有如下断言.

断言: 对每个有限可测剖分ξ我们有$H_\mu(\xi|P_\mu(T) \vee \mathcal{A}_{-\infty}) = H_\mu(\xi|\mathcal{A}_{-\infty})$. 断言的证明: 首先假设$\xi \subset \mathcal{A}$, 则

$$\begin{aligned} H_\mu(\xi|P_\mu(T) \vee \mathcal{A}_{-\infty}) &\geq \lim_{p\to\infty} H_\mu(\xi|\bigvee_{i=1}^{\infty} T^{-pi}\xi \vee P_\mu(T) \vee \mathcal{A}_{-\infty}) \\ &= \lim_{p\to\infty} H_\mu(\xi|\bigvee_{i=1}^{\infty} T^{-pi}\xi \vee \mathcal{A}_{-\infty}) \quad (\text{由定理7.10.13}) \\ &\geq \lim_{p\to\infty} H_\mu(\xi|T^{-p}\mathcal{A}) = H_\mu(\xi|\mathcal{A}_{-\infty}). \end{aligned}$$

其次对任意的$n \in \mathbb{N}$和有限可测剖分$\xi \subset T^n\mathcal{A}$, 我们可以进行同样的讨论. 最后对$X$的任何一个有限可测剖分$\xi = \{A_1,A_2,\cdots,A_k\}$和$\epsilon > 0$. 选$\delta = \delta(\epsilon,k) > 0$满足引理7.11.2的条件. 注意到$\lim_{n\to+\infty} T^n\mathcal{A} \nearrow \mathcal{X}$, 我们能找到$n \in \mathbb{N}$ 和有限可测剖分$\beta = \{B_1,B_2,\cdots,B_k\} \subset T^n\mathcal{A}$使得$\sum_{i=1}^k \mu(A_i\Delta B_i) < \delta$, 进而$H_\mu(\xi|\beta) + H_\mu(\beta|\xi) < \epsilon$. 现在

$$\begin{aligned} H_\mu(\xi|\mathcal{A}_{-\infty}) &\leq H_\mu(\xi \vee \beta|\mathcal{A}_{-\infty}) \leq H_\mu(\beta|\mathcal{A}_{-\infty}) + H_\mu(\xi|\beta) \\ &\leq H_\mu(\beta|\mathcal{A}_{-\infty} \vee P_\mu(T)) + \epsilon \leq H_\mu(\xi \vee \beta|\mathcal{A}_{-\infty} \vee P_\mu(T)) + \epsilon \\ &\leq H_\mu(\xi|\mathcal{A}_{-\infty} \vee P_\mu(T)) + H_\mu(\beta|\xi) + \epsilon \leq H_\mu(\xi|\mathcal{A}_{-\infty} \vee P_\mu(T)) + 2\epsilon. \end{aligned}$$

再因ϵ是任意的, 我们有$H_\mu(\xi|\mathcal{A}_{-\infty}) \le H_\mu(\xi|\mathcal{A}_{-\infty} \vee P_\mu(T))$, 而相反方向的不等式是明显成立地, 这就完成断言的证明.

现在从断言我们知道, 如果有限可测剖分ξ满足$\xi \subset P_\mu(T) \vee \mathcal{A}_{-\infty}$, 则$\xi \subset \mathcal{A}_{-\infty}$.因此$P_\mu(T) \subseteq \mathcal{A}_{-\infty}$. □

定理 7.11.4 (Rohlin-Sinai定理) 设$(X, \mathcal{X}, \mu, T)$ 为可逆保测系统, 则存在$\mathcal{X}$的子σ代数$\mathcal{K}$使得

1. $\mathcal{K} \subset T\mathcal{K}$,
2. $\bigvee_{n=0}^{+\infty} T^n\mathcal{K} = \mathcal{X}$,
3. $\bigcap_{n=0}^{\infty} T^{-n}\mathcal{K} = P_\mu$.

特别地, $(X, \mathcal{X}, \mu, T)$为K系统当且仅当$(X, \mathcal{X}, \mu, T)$为完全正熵系统.

证明. 取有限可测剖分$\xi_n \nearrow \mathcal{X}$, 归纳地定义$\eta_p = \eta_{p-1} \vee T^{-n_p}\xi_p$, 其中序列$n_p \in \mathbb{N}$(利用引理7.10.9的(2))满足对每个$q \ge 2$ 有以下一组不等式成立

$$H_\mu(\eta_p|\eta_{q-1}^-) - H_\mu(\eta_p|\eta_q^-) < (\frac{1}{p})\frac{1}{2^{q-p}},\ p = 1, 2, \cdots, q-1.$$

现在固定p, 再对$q = p+1, p+2, \cdots, n$求和即可得到

$$H_\mu(\eta_p|\eta_p^-) - H_\mu(\eta_p|\eta_n^-) < \frac{1}{p}\text{对所有}n > p\text{ 成立}.$$

设$\mathcal{E}$为有限代数$\{\eta_n\}$生成的σ代数, 则

$$\eta_n^- \to \bigvee_{n=1}^{\infty} T^{-i}\mathcal{E}.$$

再设

$$\mathcal{K} = \bigvee_{n=1}^{\infty} T^{-i}\mathcal{E}.$$

我们有

$$H_\mu(\eta_p|\eta_p^-) - H_\mu(\eta_p|\mathcal{K}) \le \frac{1}{p}.$$

显然$T^{-1}\mathcal{K} \subset \mathcal{K}$, $T^n\mathcal{K} \nearrow \mathcal{X}$且

$$\lim_{p\to\infty}(H_\mu(\eta_p|\eta_p^-) - H_\mu(\eta_p|\mathcal{K})) = 0.$$

设ξ 为有限可测剖分且$\xi \subset \bigcap_{n=0}^{\infty} T^{-n}\mathcal{K}$, 那么

$$\xi^T \subset \bigcap_{n=0}^{\infty} T^{-n}\mathcal{K}.$$

因此

$$\begin{aligned} H_\mu(\xi|\xi^-) &= H_\mu(\eta_p \vee \xi|\eta_p^- \vee \xi^-) - H_\mu(\eta_p|\eta_p^- \vee \xi^T) \\ &\le H_\mu(\eta_p|\eta_p^-) + H_\mu(\xi|\eta_p^-) - H_\mu(\eta_p|\mathcal{K}) \\ &= H_\mu(\xi|\eta_p^-) + (H_\mu(\eta_p|\eta_p^-) - H_\mu(\eta_p|\mathcal{K})). \end{aligned}$$

令$p \to \infty$, 注意到

$$\lim_{p\to\infty} H_\mu(\xi|\eta_p^-) = H_\mu(\xi|\mathcal{K}) = 0$$

以及

$$\lim_{p\to\infty} \big(H_\mu(\eta_p|\eta_p^-) - H_\mu(\eta_p|\mathcal{K})\big) = 0$$

我们就得到

$$h_\mu(T,\xi) = H_\mu(\xi|\xi^-) = 0.$$

因此如果$\xi \subset \bigcap_{n=0}^\infty T^{-n}\mathcal{K}$, 则$\xi \subset P_\mu(T)$, 这说明

$$\bigcap_{n=0}^\infty T^{-n}\mathcal{K} \subseteq P_\mu(T).$$

注意到$\mathcal{K}$满足定理 7.11.3的条件, 我们有

$$\bigcap_{n=0}^\infty T^{-n}\mathcal{K} \supseteq P_\mu(T).$$

定理得证. □

定理 7.11.5 设$(X,\mathcal{X},\mu,T)$和$(Y,\mathcal{Y},\nu,S)$为两个可逆保测系统, 则

$$P_{\mu\times\nu}(T\times S) = P_\mu(T)\times P_\nu(S).$$

特别地, 两个测度K系统的乘积仍为测度K系统.

证明. 利用Rohlin-Sinai 定理存在$\mathcal{X}$ 的子σ 代数$\mathcal{K}(T)$ 和$\mathcal{Y}$ 的子σ 代数$\mathcal{K}(S)$ 满足$T^n\mathcal{K}(T) \nearrow \mathcal{X}$, $T^{-n}\mathcal{K}(T) \searrow P_\mu(T)$, $S^n\mathcal{K}(S) \nearrow \mathcal{Y}$ 和$S^{-n}\mathcal{K}(S) \searrow P_\nu(S)$. 因此

$$(T\times S)^n\mathcal{K}(T)\times\mathcal{K}(S) \nearrow \mathcal{X}\times\mathcal{Y}.$$

由定理7.11.3我们得到

$$\begin{aligned} P_{\mu\times\nu}(T\times S) &\subset \bigcap_{n=1}^\infty [(T\times S)^{-n}\mathcal{K}(T)\times\mathcal{K}(S)] \\ &\subset \bigcap_{n=1}^\infty [T^{-n}\mathcal{K}(T)\times\mathcal{Y}] = P_\mu(T)\times\mathcal{Y}. \end{aligned}$$

同理

$$P_{\mu\times\nu}(T\times S)\subset\mathcal{X}\times P_\nu(S).$$

这就得到

$$P_{\mu\times\nu}(T\times S)\subset P_\mu(T)\times P_\nu(S).$$

而相反的包含关系是显然成立地, 从而定理得证. □

§7.11.3　K系统与混合性

一个保测系统$(X,\mathcal{X},\mu,T)$称为**序k强混合的**, 如果对任意$A_0,A_1,\cdots,A_k\in\mathcal{X}$,

$$\lim_{n_1,n_2,\cdots,n_k\to+\infty}\mu(A_0\cap T^{-n_1}A_1\cap\cdots\cap T^{-(n_1+n_2+\cdots+n_k)}A_k)=\mu(A_0)\mu(A_1)\cdots\mu(A_k).\quad(7.11.1)$$

如果对每个$k\in\mathbb{N}$, $(X,\mathcal{X},\mu,T)$均为序k强混合的, 则称$(X,\mathcal{X},\mu,T)$为**完全强混合** 系统. 显然序1强混合性等价于强混合性.

以下我们将证明测度K系统为完全强混合系统, 为此我们还需要另外几个引理. 首先我们将以下Pinsker不等式留给读者做为练习.

引理 7.11.6 (Pinsker不等式) 设$p_i,q_i>0,i=1,2,\cdots,l$ 满足$\sum_{i=1}^l p_i=1$和$\sum_{i=1}^l q_i\le 1$. 则

$$\sum_{i=1}^l p_i\log(\frac{p_i}{q_i})\ge\frac{(\sum_{i=1}^l|p_i-q_i|)^2}{2}\log 2.$$

引理 7.11.7 设$(X,\mathcal{X},\mu,T)$为保测系统, α,β为X的两个有限可测剖分以及$\epsilon>0$. 如果

$$H_\mu(\alpha)-H_\mu(\alpha|\beta)<\frac{\epsilon^2}{2}\log 2,$$

则

$$\sum_{B\in\beta}\sum_{A\in\alpha}|\mu(A\cap B)-\mu(A)\mu(B)|<\epsilon.$$

证明. 由Pinsker不等式,

$$\begin{aligned}H_\mu(\alpha)-H_\mu(\alpha|\beta)&=-\sum_{A\in\alpha}\mu(A)\log\mu(A)+\sum_{B\in\beta}\sum_{A\in\alpha}\mu(A\cap B)\log\frac{\mu(A\cap B)}{\mu(B)}\\&=\sum_{B\in\beta}\sum_{A\in\alpha}\mu(A\cap B)\log\frac{\mu(A\cap B)}{\mu(A)\mu(B)}\\&\ge(\frac{1}{2}\log 2)\cdot(\sum_{B\in\beta}\sum_{A\in\alpha}|\mu(A\cap B)-\mu(A)\mu(B)|)^2\end{aligned}$$

这说明

$$\sum_{B\in\beta}\sum_{A\in\alpha}|\mu(A\cap B)-\mu(A)\mu(B)|<\epsilon.$$

□

定理 7.11.8 测度K系统为完全强混合.

证明. 设$(X,\mathcal{X},\mu,T)$为测度K系统和$\mathcal{X}$的子σ代数$\mathcal{K}$满足

$$\mathcal{K}\subset T\mathcal{K},\ \bigvee_{n=0}^{+\infty}T^n\mathcal{K}=\mathcal{X}\ \text{且}\ \bigcap_{n=0}^{\infty}T^{-n}\mathcal{K}=\{\emptyset,X\}.$$

因$\bigvee_{n=0}^{+\infty}T^n\mathcal{K}=\mathcal{X}$, 为说明$(X,\mathcal{X},\mu,T)$为完全强混合的, 我们只需证明对任意$k\in\mathbb{N}$和$A_0,A_1,\cdots,A_k\in\bigcup_{n=0}^{+\infty}T^n\mathcal{K}$, (7.11.1)成立.

现设$A_0,A_1,\cdots,A_k\in\bigcup_{n=0}^{+\infty}T^n\mathcal{K}$, 则存在$m_0\in\mathbb{N}$ 使得$A_0,A_1,\cdots,A_k\in T^{m_0}\mathcal{K}$. 固定$\epsilon>0$.因$\cap_{n=0}^{\infty}T^{-n}\mathcal{K}=\{\emptyset,X\}$ 和$T^{-(n+1)}\mathcal{K}\subset T^{-n}\mathcal{K}$对$n\in\mathbb{Z}_+$,

$$\lim_{n\to+\infty}H_\mu(\{A_i,A_i^c\}|T^{-n}\mathcal{K})=H_\mu(\{A_i,A_i^c\})\text{对每个}i\in\{0,1,\cdots,k\}\text{成立}.$$

因此存在$N\in\mathbb{N}$使得当$n\ge N$时

$$H_\mu(\{A_i,A_i^c\})-H_\mu(\{A_i,A_i^c\}|T^{-n}\mathcal{K})<\frac{\epsilon^2}{2}\log 2\text{对每个}i\in\{0,1,\cdots,k\}\text{成立}.$$

特别地,对任意$n\ge N$和$B\in T^{-n}\mathcal{K}$, $H_\mu(\{A_i,A_i^c\})-H_\mu(\{A_i,A_i^c\}|\{B,B^c\})<\frac{\epsilon^2}{2}\log 2$ 对每个$i\in\{0,1,\cdots,k\}$成立. 由引理7.11.7, 对任意$n\ge N$和$B\in T^{-n}\mathcal{K}$我们有

$$|\mu(A_i\cap B)-\mu(A_i)\mu(B)|<\epsilon\text{对}i\in\{0,1,\cdots,k\}\text{成立}. \tag{7.11.2}$$

当$n_1,n_2,n_3,\cdots,n_k\ge N+m_0$时, 使用(7.11.2)和事实: $A_0,A_1\cdots,A_k\in T^{m_0}\mathcal{K}$, 我们有

$$|\mu(A_{k-1}\cap T^{-n_k}A_k)-\mu(A_{k-1})\mu(A_k)|<\epsilon$$
$$|\mu(A_{k-2}\cap T^{-n_{k-1}}(A_{k-1}\cap T^{-n_k}A_k))-\mu(A_{k_2})\mu(A_{k-1}\cap T^{-n_k}A_k)|<\epsilon$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$|\mu(A_0\cap T^{-n_1}(A_1\cap T^{-n_2}(A_2\cdots\cap T^{-n_k}A_k\cdots)))-\mu(A_0)\mu(A_1\cap T^{-n_2}(A_2\cdots\cap T^{-n_k}A_k\cdots))|<\epsilon$$

从上面的一组不等式我们容易推得: 当$n_1,n_2,n_3,\cdots,n_k\ge N+m_0$时,

$$|\mu(A_0\cap T^{-n_1}(A_1)\cap T^{-(n_1+n_2)}A_2\cap\cdots\cap T^{-(n_1+\cdots+n_k)}A_k)-\mu(A_0)\mu(A_1)\cdots\mu(A_k)|<k\epsilon.$$

因$\epsilon>0$是任意的,

$$\lim_{n_1,n_2,\cdots,n_k\to+\infty}\mu(A_0\cap T^{-n_1}A_1\cap T^{-(n_1+n_2)}A_2\cap\cdots\cap T^{-(n_1+n_2+\cdots+n_k)}A_k)=\mu(A_0)\mu(A_1)\cdots\mu(A_k).$$

这就完成了证明. □

§7.11.4 Ornstein 定理等

熵理论中最为深刻的结论之一就是Ornstein 定理:

定理 7.11.9 (Ornstein定理) 设$(X_1, \mathcal{X}_1, \mu_1, T_1)$ 和$(X_2, \mathcal{X}_2, \mu_2, T_2)$ 为两个底空间为*Lebesgue*系统的*Bernoulli* 系统. 如果它们的熵相同，即$h_{\mu_1}(T_1) = h_{\mu_2}(T_2)$, 那么它们为共轭的（因为此时两个*Bernoulli* 系统仍为*Lebesgue* 空间，故它们也为同构的.）.

Ornstein 定理的证明非常困难, 可以参见[174] 和[76]. 根据Ornstein 定理容易说明很多关于B 系统的性质：任何B 系统$(X, \mathcal{X}, \mu, T)$ 具有n 次根, 即存在$(Y, \mathcal{Y}, \nu, S)$ 使得$T = S^n$; 任何B 系统共轭与某两个B 系统的乘积系统；任何B 系统$(X, \mathcal{X}, \mu, T)$ 共轭于它的逆系统$(X, \mathcal{X}, \mu, T^{-1})$. Ornstein 还证明了B 系统的根、因子、逆极限系统仍为B 系统.

Ornstein 定理告诉我们对于B 系统，熵是全系不变量. 这基本上是我们能得到结果中最好的结论了. 因为对于K 系统, 熵就不是全系不变量了. 这个很容易解释清楚. 任取一个不是B 系统的K 系统$(X, \mathcal{X}, \mu, T)$，那么根据Rohlin-Sinai 定理，$h_\mu(T) > 0$. 再取一个B 系统$(Y, \mathcal{Y}, \nu, S)$ 使得$h_\mu(T) = h_\nu(S)$. 这样我们就得到两个不共轭的K 系统. 事实上Ornstein和Shields 证明了存在不可数个具有相同熵但是不共轭的K 系统.

上面提到任何B 系统$(X, \mathcal{X}, \mu, T)$ 具有n 次根；任何B 系统$(X, \mathcal{X}, \mu, T)$ 共轭于它的逆系统$(X, \mathcal{X}, \mu, T^{-1})$. 对于K 系统这两个结果都不成立.

我们以Sinai 关于B 系统的一个著名结果结束这一小节：

定理 7.11.10 (Sinai) 设$(X, \mathcal{X}, \mu, T)$ 为遍历的可逆保测系统. 如果$h_\mu(T) > 0$, 那么任何满足$h_\nu(S) \le h_\mu(T)$ 的B 系统$(Y, \mathcal{Y}, \nu, S)$ 都是$(X, \mathcal{X}, \mu, T)$ 的因子，即存在保测映射$\phi : X \to Y$ 使得$\phi T = S\phi$.

习 题

1. 设$(X, \mathcal{X}, \mu, T)$ 为K系统.证明: $(X, \mathcal{X}, \mu, T^k)$ $(k \neq 0)$ 也为K 系统.
2. 证明Kolmogorov 属性对因子遗传.
3. 证明引理7.11.2和引理7.11.6.
4. 设$\mathbb{S}^1$为复平面的单位圆周, $p > 1$, $T_p : \mathbb{S}^1 \to \mathbb{S}^1$ 满足$T_p(z) = z^p$ 以及$m_{\mathbb{S}^1}$ 为$\mathbb{S}^1$ 上的Haar 测度, 证明：$(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), m_{\mathbb{S}^1}, T_p)$ 为Bernoulli 系统.
5. 设$(X, \mathcal{X}, \mu, T)$ 为可逆保测系统, α 为X 的一个有限可测剖分. 则对任意的$\epsilon > 0$, 存在$K \in \mathbb{N}$ 使得对任意K 分离的非空有限集$E \subset \mathbb{Z}$, 我们有
$$|\frac{1}{|E|} H_\mu(\bigvee_{i \in E} T^{-i}\alpha | P_\mu(T)) - H_\mu(\alpha | P_\mu(T))| < \epsilon,$$
在这里$\mathbb{Z}$ 的子集称为K 分离, 如果它的任意两个不同元素之差的绝对值大于K.

6. 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统, 那么它要么为零熵的, 要么U_T 限制在$L^2(X,\mathcal{P}(T),\mu)$ 正交补上具有可数Lebsgue谱.
7. 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统, 如果T 具有纯粹奇异谱或具有有限乘数谱, 那么其熵为0.

§7.12 局部熵理论简介

§7.12.1 测度K 系统与拓扑K 系统

在本书中我们已经详细介绍过测度K 系统. 在这里我们先回顾一些概念, 然后介绍拓扑K 系统.

回顾一个保测系统$(X,\mathcal{X},\mu,T)$ 称为**Kolmogorov 系统** 或简称为**K 系统** 是指存在子σ 代数$\mathcal{K}\subset\mathcal{X}$ 满足以下条件: (1)$\mathcal{K}\subset T\mathcal{K}$, (2) $\bigvee_{n=0}^{\infty}T^n\mathcal{K}=\mathcal{X}$, (3) $\bigcap_{n=0}^{\infty}T^{-n}\mathcal{K}=\{X,\emptyset\}$.

K 系统完全异于零熵系统, 它可以视为在熵语言下最复杂的系统. 之前我们证明过以下等价:

- 保测系统为K 系统;
- 系统为完全正熵的(即每个非平凡因子具有正熵);
- 系统任何非平凡二元素剖分具有正熵;
- 系统任何非平凡有限剖分具有正熵.

在拓扑动力系统中, 拓扑K 系统由Blanchard 在上世纪九十年代最先研究. 他引入了c.p.e. 和u.p.e. 的概念, 这两个概念类似于上面测度K 系统刻画的前两个.

定义 7.12.1 称一个拓扑系统(X,T) 具有**(拓扑)完全正熵**_(completely positive entropy)_, 简称_c.p.e._ , 是指它的任何非平凡因子具有正拓扑熵; 称它具有**一致正熵**, 简称_u.p.e._, 是指任何有两个非平凡开集组成的开覆盖具有正拓扑熵.

Blanchard [24] 证明了任何一致正熵系统为拓扑弱混合的, 任何完全正熵系统蕴含具有全支撑的不变测度.

定义 7.12.2 _[115]_ 称系统(X,T) 具有n **阶的一致正熵** _(u.p.e. of order n)_ 是指X 的任何n 个非稠密开集组成的覆盖具有正拓扑熵. 我们称(X,T) 具有**任何阶一致正熵**或者**拓扑K 系统**, 是指对于任何$n\geq 2$ 它为n 阶一致正熵的.

明显地, 2阶一致正熵就是前面定义的一致正熵. 可以证明一致正熵系统为拓扑mildl 混合的[115]; 极小拓扑K 系统为拓扑强混合的[109].

黄文和叶向东在文章[115] 中回答了几个关于一致正熵和完全正熵的猜测: 对于每个$n\geq 2$, n 阶一致正熵不蕴含$n+1$ 阶一致正熵(回答了Host 在[80] 的问题); 存在传递对角

系统(diagonal system) 不为一致正熵的(回答了[25, 问题1]); 存在一致正熵系统没有具有全支撑的遍历测度(回答了[24, 问题2]).

在[78] 中Glasner 和Weiss 证明了：任何具有全支撑测度K 系统测度的拓扑动力系统为一致正熵的; 存在极小一致正熵系统使得对于其上任何遍历测度具有正的测度熵.这些结果对于任何阶一致正熵系统也是成立的[115]：任何具有全支撑测度K 系统测度的拓扑动力系统为任意阶一致正熵的, 即拓扑K 系统；一个拓扑系统(X,T)为拓扑K 系统当且仅当存在不变测度$\mu \in \mathcal{M}(X,T)$ 使得对于X 任何非平凡剖分α, $h_\mu(T,\alpha) > 0$.

§7.12.2 测度与拓扑熵串

熵对最早是由Blanchard 开始研究的[25], 之后发展为熵串、熵集等.

§7.12.2.1 n熵串

设(X,T) 为拓扑动力系统, $n \geq 2$, n阶乘积系统记为$(X^n,T^{(n)})$, 其中$T^{(n)} = T \times \ldots \times T$ (n 次). 设$\mathcal{B}_X$ 为X 上的Borel σ代数, $\mathcal{B}_X^{(n)}$ 为乘积空间X^n 上的Borel σ 代数. 我们将X^n 的对角线记为$\Delta_n(X) = \{(x_i)_1^n \in X^n : x_1 = \cdots = x_n\}$. 设$(x_i)_1^n \in X^n$. X 的一个有限覆盖$\mathcal{U}$ 称为相对于$(x_i)_1^n$ 的**可允许覆盖**是指不存在$\mathcal{U}$ 的某一个元素, 它的闭包包含了全体的x_i. 类似定义相对于$(x_i)_1^n$ 的可允许剖分.

定义 7.12.3 ([78, 115]) *设(X,T) 为拓扑动力系统. 一个n 串$(x_i)_1^n \in X^n$, $n \geq 2$, 称为n* **熵串***是指存在$1 \leq i \neq j \leq n$ 使得$x_i \neq x_j$, 并且对于$(x_i)_1^n$ 的任何可允许开覆盖$\mathcal{U}$ 我们有$h_{\rm top}(T,\mathcal{U}) > 0$.*

*2熵串通常称为***熵对***.*

我们用符号$E_n(X,T)$ 表示全体n 熵串的集合. 它的基本性质如下[25]:

1. 如果$\mathcal{U} = \{U_1, \cdots, U_n\}$ 为X 开覆盖且$h_{\rm top}(T,\mathcal{U}) > 0$, 那么对于每个$1 \leq i \leq n$ 存在$x_i \in U_i^c$ 使得$(x_i)_1^n$ 为n 熵串.

2. $E_n(X,T) \cup \Delta_n(X)$ 为X^n 的$T^{(n)}$ 不变闭子集.

3. 设$\pi : (X,T) \to (Y,S)$ 为因子映射. 那么$\pi^{(n)}\big(E_n(X,T) \cup \Delta_n(X)\big) = E_n(Y,S) \cup \Delta_n(Y)$.

4. 设W 为(X,T) 的T不变闭子集. 如果$(x_1,\ldots,x_n)$ 为$(W,T|_W)$ 的n熵串, 那么它也为(X,T) 的n 熵串.

根据(1), (X,T) 具有正拓扑熵当且仅当$E_2(X,T) \neq \emptyset$; (X,T) 为n 阶一致正熵的当且仅当每个不在对角线$\Delta_n(X)$ 上的串$(x_i)_1^n \in X^n$ 为n 熵串. 类似于测度K系统不交于零熵系统, Blanchard 证明了一致正熵系统不交于任何极小零拓扑熵系统[25].

之前我们证明过保测系统$(X,\mathcal{X},\mu,T)$ 的Pinsker 因子是具有最大的具有零测度熵的因子. 根据Rohlin-Sinai 定理, 保测系统$(X,\mathcal{X},\mu,T)$ 为K 系统当且仅当它的Pinsker 因子为平凡的, 即$P_\mu = \{X,\emptyset\} \pmod{\mu}$. 我们定义拓扑Pinsker 因子.

定义 7.12.4 设(X,T) 为拓扑动力系统. 它的**拓扑Pinsker 因子** 是指具有零拓扑熵的最大因子.

Blanchard 和Lacroix [31] 证明对于任何拓扑动力系统(X,T), 包含$E_2(X,T)$ 的最小的闭不变等价关系诱导的因子为拓扑Pinsker 因子.

在[29] 中作者对测度引入熵对的概念, 他们引入熵对的方式不能直接取定义测度n 熵串($n > 2$). 黄文和叶向东在文[115] 中对测度定义了n熵串, 并且他们的定义在$n = 2$ 时与[29] 中定义的熵对相容.

定义 7.12.5 设(X,T) 为拓扑动力系统, $\mu \in \mathcal{M}(X,T)$. 一个n 点串$(x_i)_1^n \in X^{(n)}$, $n \geq 2$, 称为**相对μ 的n 熵串**,是指存在$1 \leq i \neq j \leq n$ 使得$x_i \neq x_j$, 并且对于$(x_i)_1^n$ 的任何可允许*Borel* 剖分α 我们有$h_\mu(T,\alpha) > 0$. 用$E_n^\mu(X,T)$ 记全体相对μ 的n 熵串集合.

对于拓扑动力系统(X,T), 设$\mu \in \mathcal{M}(X,T)$. 设$\mathcal{P}_\mu$ 为保测系统$(X,\mathcal{B}_X,\mu,T)$ 的Pinsker σ 代数. 定义$(X^{(n)},\mathcal{B}_X^n,T^{(n)})$ 上的交$\lambda_n(\mu)$ 为

$$\lambda_n(\mu)\left(\prod_{i=1}^n A_i\right) = \int_X \prod_{i=1}^n \mathbb{E}(1_{A_i}|\mathcal{P}_\mu)d\mu,$$

其中$A_i \in \mathcal{B}_X, i = 1,\cdots,n$. 下面结果表明相对于$\mu$ 的n 熵串为$\lambda_n(\mu)$ 的支撑. 这个结果$n = 2$ 情况在[75] 给出, 一般情况在[115] 给出.

定理 7.12.6 设(X,T) 为拓扑动力系统, $\mu \in \mathcal{M}(X,T)$ *and* $n \geq 2$. 则

$$E_n^\mu(X,T) = \mathrm{supp}(\lambda_n(\mu)) \setminus \Delta_n(X).$$

根据定理7.12.6, $h_\mu(T) = 0$ 当且仅当$E_2^\mu(X,T) = \emptyset$; $E_n^\mu(X,T) \cup \Delta_n(X)$ 为X^n 的$T^{(n)}$ 不变闭子集.

§7.12.2.2 熵的局部变分原理

联系两种熵串的关键是熵的局部变分原理, 这个结果是经典熵的变分原理的推广. Blanchard, Glasner 和Host 证明了如下定理

定理 7.12.7 *[26]* 设(X,T) 为拓扑动力系统. 对于X 的有限开覆盖$\mathcal{U}$, 存在不变测度$\mu \in \mathcal{M}(X,T)$ 使得

$$\inf_\alpha h_\mu(T,\alpha) \geq h_{\mathrm{top}}(T,\mathcal{U}),$$

其中下确界取遍所有比$\mathcal{U}$ 细的有限剖分.

在[115] 中黄文和叶向东证明了：对于$\mu \in \mathcal{M}(X,T)$, 如果X 任何比$\mathcal{U}$ 细的Borel 剖分α 我们有$h_\mu(T,\alpha) > 0$, 那么

$$\inf_\alpha h_\mu(T,\alpha) > 0, \text{且 } h_{\text{top}}(T,\mathcal{U}) > 0.$$

这提供了定理7.12.7 某种意义下的逆定理.

对于开覆盖$\mathcal{U}$, 自然的问题是上面的结论中

$$\inf_\alpha h_\mu(T,\alpha) = h_{\text{top}}(T,\mathcal{U})$$

是否成立？为了研究这个问题Romagnoli [170] 引入下面的概念

$$h_\mu^+(T,\mathcal{U}) = \inf_{\alpha \succeq \mathcal{U}} h_\mu(T,\alpha) \text{ 以及}$$
$$h_\mu(T,\mathcal{U}) = \lim_{n\to+\infty} \frac{1}{n} \inf_{\alpha \succeq \bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}} H_\mu(\alpha),$$

其中α 为X 有限Borel 剖分. 事实上它们是相等的：$h_\mu^+(T,\mathcal{U}) = h_\mu(T,\mathcal{U})$.

定理 7.12.8 (熵的局部变分原理) *[170, 81] 设(X,T) 为拓扑动力系统, $\mathcal{U}$ 为X 的有限开覆盖. 那么*

$$\max_{\mu \in \mathcal{M}(X,T)} h_\mu^+(T,\mathcal{U}) = \max_{\mu \in \mathcal{M}(X,T)} h_\mu(T,\mathcal{U}) = h_{\text{top}}(T,\mathcal{U}).$$

注意到$h_\mu(T) = \sup_\mathcal{U} h_\mu(T,\mathcal{U})$, 其中上确界取遍$X$ 的所有有限开覆盖, 那么根据局部变分原理就可以得到经典的熵变分原理.

拓扑熵串与测度熵串密切相关. 对于每个不变测度测度熵对包含在熵对的集合中[29]; 反之亦然[26] . n 熵串的变分原理如下:

定理 7.12.9 *[115] 设(X,T) 为拓扑动力系统. 如果$\mu \in \mathcal{M}(X,T)$, 那么对于每个$n \geq 2$*

$$E_n(X,T) \supseteq E_n^\mu(X,T),$$

并且存在$\mu \in \mathcal{M}(X,T)$ 使得

$$E_n(X,T) = E_n^\mu(X,T).$$

Blanchard, Glasner 和Host [26] 构造了一个拓扑动力系统, 它的任何熵对都不是遍历测度的熵对.

若干个n 阶一致正熵系统的乘积系统仍为n 阶一致正熵的；若干个拓扑K系统的乘积系统仍为拓扑K的[115].

§7.12.2.3 遍历分解

设(X,T) 为拓扑动力系统, $\mu\in\mathcal{M}(X,T)$. 设$\mu=\int_X\mu_x d\mu(x)$ 为遍历分解. 我们之前证明过对于剖分α 我们有$h_\mu(T,\alpha)=\int_X h_{\mu_x}(T,\alpha)d\mu(x)$. 这个性质对于有限Borel 覆盖也成立:

定理 7.12.10 *[115] 设(X,T) 为拓扑动力系统, $\mu\in\mathcal{M}(X,T)$, $\mathcal{U}$ 为X 有限Borel 覆盖. 设$\mu=\int_X\mu_x d\mu(x)$ 为μ 的遍历分解, 那么*

$$h_\mu(T,\mathcal{U})=\int_X h_{\mu_x}(T,\mathcal{U})d\mu(x)$$

取定(X,T) 的有限开覆盖$\mathcal{U}$, 根据定理7.12.10 和局部变分原理可以证明：存在$\nu\in\mathcal{M}^e(X,T)$ 使得$h_\nu(T,\mathcal{U})=h_{\mathrm{top}}(T,\mathcal{U})$. 进一步地, 还可以证明映射$\mu\in\mathcal{M}(X,T)\mapsto h_\mu(T,\mathcal{U})$ 为上半连续的[117].

下面结论给出遍历测度熵串和不变测度熵串的关系（$n=2$ 见[26]；一般情况见[115]）.

定理 7.12.11 *设(X,T) 为拓扑动力系统, 设$\mu\in\mathcal{M}(X,T)$的遍历分解为$\mu=\int_X\mu_x d\mu(x)$. 则*

1. *对于μ-a.e. $x\in X$, $E_n^{\mu_x}(X,T)\subseteq E_n^\mu(X,T)$, $\forall n\geq 2$.*
2. *如果$(x_i)_1^n\in E_n^\mu(X,T)$, 那么对于$(x_i)_1^n$ 的任何邻域V ,*

$$\mu(\{x\in X: V\cap E_n^{\mu_x}(X,T)\neq\emptyset\})>0.$$

于是可以取$X_0\in\mathcal{B}_X$ 使得$\mu(X_0)=1$ 且

$$\overline{\bigcup\{E_n^{\mu_x}(X,T):x\in X_0\}}\setminus\Delta_n(X)=E_n^\mu(X,T).$$

§7.12.2.4 弱马蹄

马蹄是著名数学家Smale 引入的, 用于刻画系统的复杂性. 在[115] 黄文和叶向东得到如下关于拓扑n 熵串的刻画. 设(X,T) 为拓扑动力系统, $n\geq 2$. 那么$(x_1,\cdots,x_n)\in E_n(X,T)$ 当且仅当对于$(x_1,\cdots,x_n)$ 的任何邻域$U_1\times\cdots\times U_n$, 存在$\mathbb{Z}_+$ 的正密度子集$S=\{s_1<s_2<\cdots\}$ 使得

$$\bigcap_{i=1}^\infty T^{-s_i}U_{t(i)}\neq\emptyset,\quad \forall t\in\{1,2,\cdots,n\}^S.$$

上面结论也可参见[128, 129, 130].

基于上结论, 我们可以定义弱马蹄如下: 设J 为$\mathbb{Z}_+$ 的子集, 称系统(X,T) 具有**具有插集J 的弱马蹄** (weak horseshoe with an interpolating set J) 是指存在X 的两个不交闭子集U_0,U_1 使得对于任何$t\in\{0,1\}^J$,

$$\bigcap_{j\in J}T^{-j}U_{t(j)}\neq\emptyset,$$

即存在$x_t \in X$ 使得$T^j(x_t) \in U_{t(j)}, \forall j \in J$ [106] . 如果J 具有正密度, 那么就称(X,T) 具有**弱马蹄**.

定理 7.12.12 *[115] 一个拓扑动力系统具有正拓扑熵当且仅当它具有弱马蹄.*

这个结果最早是由Glasner 和Weiss 对于符号系统给出[79]. 上面结论被Kerr 和Li 推广到可数amenable 群[129]. 这个结论有许多应用, 例如最近, 黄文和吕克宁[106] 研究了无穷维随机动力系统复杂性, 证明了此时正熵蕴含弱马蹄.

§7.12.3 序列熵

Kushnirenko 于1967 年引入序列作为新的同构不变量[140]. 随后Goodman 于1974 年引入拓扑序列熵[83].

下面给出序列熵的定义. 记$\mathcal{F}_{inf}$ 为$\mathbb{Z}_+$ 全体无穷子集的集合. 设$S = \{0 \le t_1 < t_2 < \cdots\} \in \mathcal{F}_{inf}$, $\mathcal{U}$ 为X 有限开覆盖. 系统(X,T) $\mathcal{U}$ **沿着S 的序列熵** 定义为

$$h_{\rm top}^S(T,\mathcal{U}) = \limsup_{n\to+\infty} \frac{1}{n} \log N(\bigvee_{i=1}^n T^{-t_i}\mathcal{U}).$$

(X,T) **沿着S 的序列熵** 为

$$h_{\rm top}^S(T) = \sup\{h_{\rm top}^S(T,\mathcal{U}) : \mathcal{U} \text{ 有限开覆盖}\}.$$

如果$S = \mathbb{Z}_+$, 那么就得到经典熵的定义, 此时我们省去上标$\mathbb{Z}_+$.

设$(X,\mathcal{X},\mu,T)$ 为保测系统, 记全体有限可测剖分的集合为$\mathcal{P}_X$. 设$\xi \in \mathcal{P}_X$, $(X,\mathcal{X},\mu,T)$ ξ **沿着S 的序列熵** 定义为

$$h_\mu^S(T,\xi) = \limsup_{n\to+\infty} \frac{1}{n} H_\mu(\bigvee_{i=1}^n T^{-t_i}\xi).$$

$(X,\mathcal{X},T,\mu)$ **沿S 的序列熵** 为

$$h_\mu^S(T) = \sup_{\alpha\in\mathcal{P}_X} h_\mu^S(T,\alpha).$$

同样地, 如果$S = \mathbb{Z}_+$ 就得到经典熵的定义, 此时我们省去上标$\mathbb{Z}_+$.

$h_{\rm top}^S(T)$ 和$h_\mu^S(T)$ 为动力系统同构不变量.

设$(X,\mathcal{X},\mu,T)$ 为保测系统, 如果$h_\mu(T) > 0$, 则$h_\mu^S(T) = K(S)h_\mu(T)$, 其中$K(S)$ 为不依赖于T 的数[139]. 这个结论表明对于正熵系统, 序列熵作为新的同构不变量没有起太大作用, 然而对于零测度熵系统序列熵十分重要.

对于序列熵一个重要的事实是变分原理一般不成立[83]. 设(X,T) 为拓扑动力系统, Goodman [83] 证明了对于任何$S \in \mathcal{F}_{inf}$

$$h_{\rm top}^S(T) \ge \sup_{\mu\in\mathcal{M}(X,T)} h_\mu^S(T),$$

其中$h_{\mathrm{top}}(T) > 0$ 时等号成立. 在[83] 中上面结论有个限制条件, 此条件是多余的[51]（也可参见[116]）. 如果$h_{\mathrm{top}}(T) = 0$, 那么序列熵变分原理不一定成立. 有例子满足$h^S_{\mathrm{top}}(T) = \log 2$ 但$\sup_{\mu\in\mathcal{M}(X,T)} h^S_\mu(T) = 0$ [83].

对于保测系统$(X, \mathcal{X}, \mu, T)$, 如果$h_\mu(T) > 0$, 那么对于任何$S \in \mathcal{F}_{\mathrm{inf}}$, $h^S_\mu(T) > 0$ [181]. 类似地, 对于拓扑动力系统(X, T), 如果$h_{\mathrm{top}}(T) > 0$, 那么对于任何$S \in \mathcal{F}_{\mathrm{inf}}$, $h^S_{\mathrm{top}}(T) > 0$ [109].

类似于熵串, 我们也可以定义序列熵串, 感兴趣的读者可以参见[104, 108, 116, 129, 152] 等.

序列熵一个重要的应用在于它可以刻画各种混合性. 我们仅在这里举几个例子, 更多相关内容参见[39, 118, 119, 109, 181, 200, 201].

定理 7.12.13 *[109] 设$(X, \mathcal{X}, \mu, T)$ 为保测系统, 那么以下命题等价:*

1. *$(X, \mathcal{X}, \mu, T)$ 为mild 混合;*
2. *对于任何二元非平凡剖分$\alpha \in \mathcal{P}_X$, 任何IP 集F, 存在无限子集$A \subseteq F$ 使得$h^A_\mu(T, \alpha) > 0$;*
3. *对于任何有限非平凡剖分$\alpha \in \mathcal{P}_X$, 任何IP 集F, 存在无限子集$A \subseteq F$ 使得$h^A_\mu(T, \alpha) > 0$.*

上面非平凡指里面元素不是零测集或全测集.

定理 7.12.14 *[109] 设(X, T) 为拓扑动力系统, 那么以下等价:*

1. *(X, T) 为拓扑mild 混合.*
2. *X 的任何二元非平凡开覆盖$\mathcal{U}$, 任何IP子集F, 存在无穷子集$A \subseteq F$ 使得$h^A_{\mathrm{top}}(T, \mathcal{U}) > 0$.*
3. *X 的任何有限非平凡开覆盖$\mathcal{U}$, 任何IP子集F, 存在无穷子集$A \subseteq F$ 使得$h^A_{\mathrm{top}}(T, \mathcal{U}) > 0$.*

上面非平凡开覆盖指每个元素不是空集或稠密子集.

对于弱混合和强混合也有类似的刻画.

§7.12.4 Null系统

保测系统$(X, \mathcal{X}, \mu, T)$ 称为**null 的** 是指对于任何$S \in \mathcal{F}_{inf}$ 有$h^S_\mu(T) = 0$. null 等价于离散谱.

定理 7.12.15 (Kushnirenko) *[140] 保测系统$(X, \mathcal{X}, \mu, T)$ 具离散谱当且仅当它为null 的.*

事实上, 对于保测系统$(X, \mathcal{X}, \mu, T)$ 和$\alpha \in \mathcal{P}_X$, 我们有

$$\max_{S\in\mathcal{F}_{inf}} h^S_\mu(T, \alpha) = H_\mu(\alpha|\mathcal{K}_\mu),$$

其中$\mathcal{K}_\mu$ 为Kronecker σ 代数[108] . 可以证明[108, 109], 对于$B \in \mathcal{X}$ 以及$R \in \mathcal{F}_{inf}$, $\mathrm{cl}(\{U^n 1_B : n \in R\})$ 为$L^2(\mu)$ 的紧致子集当且仅当对于每个无穷子集$S \subseteq R$ 有$h_\mu^S(T, \{B, X \setminus B\}) = 0$. 尤其$B \in \mathcal{K}_\mu$ 当且仅当$h_\mu^S(T, \{B, X \setminus B\}) = 0, \forall S \in \mathcal{F}_{inf}$.

类似地, 拓扑系统(X, T) 称为**null 的** 是指$h_{\mathrm{top}}^S(T) = 0$, $\forall S \in \mathcal{F}_{inf}$. 容易证明等度连续系统为null的, 自然的问题是反之是否成立：如极小系统为null 的, 那么它是否为等度连续的？一般而言, 存在非等度连续的null系统[83]. 但是如果忽略几乎一对一扩充情况下, 结论是成立的. 回顾几乎一多一的定义：扩充$\pi : (X, T) \to (Y, S)$ 为几乎一对一的是指$\{x \in X : |\pi^{-1}(\pi(x))| = 1\}$ 为稠密G_δ 子集.

定理 7.12.16 *[104] 如果极小系统(X, T) 为null 的, 那么它为到其最大等度连续因子(X_{eq}, T_{eq}) 的几乎一对一扩充. 并且此时系统为唯一遍历的, 相对于唯一的不变测度为离散谱的.*

一个未解决的问题是[104]: 是否存在传递非极小的null 系统?

§7.12.5 最大型熵

最大型熵的概念是文[116] 中引入的. 对于拓扑系统(X, T), $n \in \mathbb{N}$ 以及有限开覆盖$\mathcal{U}$ 设

$$p^*_{X,\mathcal{U}}(n) = \max_{(t_1 < t_2 < \cdots < t_n) \in \mathbb{Z}_+^n} N\Big(\bigvee_{i=1}^n T^{-t_i}\mathcal{U}\Big).$$

T 相对于$\mathcal{U}$ 的**最大型熵**(maximal pattern entropy) 定义为

$$h^*_{\mathrm{top}}(T, \mathcal{U}) = \lim_{n \to +\infty} \frac{1}{n} \log p^*_{X,\mathcal{U}}(n).$$

容易验证$\{\log p^*_{X,\mathcal{U}}(n)\}_{n=1}^\infty$ 为次可加的, 于是可以定义(X, T)的**最大型熵**为

$$h^*_{\mathrm{top}}(T) = \sup\{h^*_{\mathrm{top}}(T, \mathcal{U}) : \mathcal{U} \text{ 有限开覆盖}\}.$$

类似地, 对于保测系统$(X, \mathcal{X}, \mu, T)$ 我们也可以定义最大型熵$h^*_\mu(T)$. 一个主要结论是：最大型熵可以用序列熵刻画.

定理 7.12.17 *[116] 设(X, T) 为拓扑动力系统, 那么*

$$h^*_{\mathrm{top}}(T) = \sup_{S \in \mathcal{F}_{inf}} h^S_{\mathrm{top}}(T).$$

设$(X, \mathcal{X}, \mu, T)$ 为保测系统, 那么

$$h^*_\mu(T) = \sup_{S \in \mathcal{F}_{inf}} h^S_\mu(T).$$

于是拓扑动力系统(X, T) 为null 的当且仅当$h^*_{\mathrm{top}}(T) = 0$; 保测系统$(X, \mathcal{X}, \mu, T)$ 为离散谱的当且仅当$h^*_\mu(T) = 0$. 最大型熵有很多有趣的性质, 例如它们为同构不变量；对于$k \in \mathbb{N} \setminus \{0\}$, $h^*_{\mathrm{top}}(T^k) = h^*_{\mathrm{top}}(T)$, $h^*_\mu(T^k) = h^*_\mu(T)$ 等.

对于熵而言, 它取值为连续的, 但是最大型熵却是离散化取值的. 即

定理 7.12.18 *1. 对于拓扑动力系统*(X,T), $h^*_{\text{top}}(T)\in\{\log k: k\in\mathbb{N}\}\cup\{\infty\}$ *[116].*

2. 对于遍历系统$(X,\mathcal{X},\mu,T)$, $h^*_{\mu}(T)\in\{\log k: k\in\mathbb{N}\}\cup\{\infty\}$ *[181].*

事实上, 如果引入序列熵串的概念, 那么可以证明对于拓扑动力系统(X,T), $k\in\mathbb{N}\cup\{\infty\}$, $h^*_{\text{top}}(T)=\log k$ 当且仅当系统存在各个坐标互异的k序列熵串但不存在相应的$(k+1)$序列熵串. 对于测度最大型熵也有类似结论[116].

由定理7.12.16, 极小系统(X,T) 为null 的, 即$h^*_{\text{top}}(T)=0$, 则(X,T) 为其最大等度连续因子的几乎一对一扩充且为唯一遍历的. 自然的问题是：如果(X,T) 为有界的, 即$h^*_{\text{top}}(T)<\infty$, 那么系统具有什么性质？可以证明, 此类系统也具有很好的结构：它为其最大等度连续因子的几乎有限对一扩充, 并且只有有限个遍历测度[105].

对于紧致度量空间X, 设$S(X)=\{h^*_{\text{top}}(T): T\in C(X,X)\}$. 由定理7.12.18, $\{0\}\subseteq S(X)\subseteq\{0,\log 2,\log 3,\cdots\}\cup\{\infty\}$. Snoha, 叶向东和张瑞丰[185] 证明了：

定理 7.12.19 *对于每个*$\{0\}\subseteq A\subseteq\{0,\log 2,\log 3,\cdots\}\cup\{\infty\}$, *存在空间*$X_A\subseteq\mathbb{R}^3$ *使得*$S(X_A)=A$.

§7.13 注记

本章的内容是熵的经典理论, 我们主要参考和借鉴了以下著作：[47, 76, 161, 162, 195]等. 我们在本章仅介绍熵理论最基本的内容, 著名的Ornstein理论可以参见[76, 175]; 局部熵理论参见[76, 207]; 微分动力系统方面的熵理论参见[123, 153] 等.

在混沌理论中，人们常把正熵作为混沌的定义之一. 在本书中，我们没有系统的对混沌理论进行讨论. 关于正熵有两个很重要的结论，一个是正熵蕴含Li-Yorke混沌[27]，另一个是正熵蕴含渐进对[30]. 关于混沌的进一步讨论可以参见[5, 207].

第八章　交理论简介

交的理论是由Furstenberg 于1963年引入到动力系统的, 它现在是动力系统研究的核心理论之一. 在这一章我们介绍关于交理论的一些基本结论和它的一些应用.

§8.1　系统交的基本概念与性质

§8.1.1　交的概念

定义 8.1.1 (交, 自交)　*1.* 设$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为保测系统. 一个乘积空间$(X\times Y,\mathcal{X}\times\mathcal{Y})$ 上的概率λ 称为$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 的**交** *(joining)* 是指λ 为$T\times S$ 不变的, 并且它到X,Y 的投射分别为μ 和ν, 即如设$p_X : X\times Y\to X, p_Y : X\times Y\to Y$ 为投射, 那么$(p_X)_*\lambda=\mu,(p_Y)_*\lambda=\nu$.

我们将所有X,Y 的交的集合记为$J(X,Y)$ 或$J(\mu,\nu)$. 将$J(X,Y)$ 中所有遍历元素记为$J_e(X,Y)$.

2. 如果$(X,\mathcal{X},\mu,T)=(Y,\mathcal{Y},\nu,S)$, 那么称它们的交为**自交***(self-joining)*. 此时$J(X,Y)$ 也记为$J(X)=J(\mu)$.

注记 8.1.2　*1.* 因为$\mu\times\nu\in J(X,Y)$, 所以$J(X,Y)\neq\emptyset$. 如果$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 之一不为遍历的, 那么$J_e(X,Y)=\emptyset$.

2. 对于测度空间$(X,\mathcal{X},\mu)$ 和$(Y,\mathcal{Y},\nu)$, 我们也可以定义$(X,\mathcal{X},\mu)$ 和$(Y,\mathcal{Y},\nu)$ 的交为乘积空间$X\times Y$ 上测度λ, 它到X,Y 的投射分别为μ 和ν.

交的概念可以推广到多个的情况.

定义 8.1.3 (多个系统的交)　*1.* 设$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i\in I}$ 为一族保测系统, 其中I 为指标集. 一个乘积空间$\prod_{i\in I}X_i$ 上的概率λ 称为$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i\in I}$ 的**交** *(joining)* 是指λ 为$\prod_{i\in I}T_i$ 不变的, 并且它到每个X_i 的投射为μ_i.

我们将所有X,Y 的交的集合记为$J(\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i\in I})$ 将$J(X,Y)$ 中所有遍历元素记为$J_e(\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i\in I})$.

2. 当$I=\{1,2,\dots,n\}$, 我们称$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i=1}^n$ 的交为n**重交**, 如果每个$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i\in I}$ 都相同, 那么称n重交为n**重自交***(n-fold self-joining)*. n重自交的全体记为$J_n(X)$.

设$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为保测系统. 那么$J(X,Y)$ 上可以赋予下面拓扑：

$$\lambda_n\to\lambda,n\to\infty\ \Leftrightarrow\ \lambda_n(A\times B)\to\lambda(A\times B),n\to\infty\ \forall A\in\mathcal{X},B\in\mathcal{Y}.$$

这是一个可度量化的拓扑, 例如设$\{A_n\}_{n\in\mathbb{Z}_+}, \{B_m\}_{m\in\mathbb{Z}_+}$ 分别为$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 的可数生成集, 那么$J(X,Y)$ 上的度量可以定义为

$$d(\lambda,\lambda')=\sum_{n,m=0}^{\infty}\frac{1}{2^{m+n}}\big|\lambda(A_n\times B_m)-\lambda'(A_n\times B_m)\big|.$$

在此拓扑下, $J(X,Y)$ 成为一个紧致度量空间. 如果赋予$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 拓扑模型, 那么$J(X,Y)$ 成为$\mathcal{M}(X\times Y,T\times S)$ 的一个紧凸子集. $J(X,Y)$ 的拓扑与$\mathcal{M}(X\times Y,T\times S)$ 上诱导的弱拓扑是吻合的. 这个拓扑与$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 拓扑模型的选取无关(习题).

定理 8.1.4 设$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为遍历系统, 那么$J_e(X,Y)$ 为$J(X,Y)$ 的端点, 它是非空的.

证明. 作为习题. □

§8.1.2 一些重要的交

例 8.1.5 (图交) 设$\phi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为因子映射. 令

$$\mathrm{id}_X\times\phi: X\to X\times Y,\ x\mapsto(x,\phi x).$$

我们称

$$\mathrm{gr}(\mu,\phi)=(\mathrm{id}_X\times\phi)_*\mu$$

为**图交***(graph joining)*. 等价地, $\mathrm{gr}(\mu,\phi)$ 可以定义为:

$$\mathrm{gr}(\mu,\phi)(A\times B)=\mu(A\cap\phi^{-1}B),\ \forall A\in\mathcal{X},B\in\mathcal{Y}.$$

如果$\phi: X\to Y$ 为同构, 那么$\mathrm{gr}(\mu,\phi)$ 也称为**同构图交**.

例 8.1.6 (非对角交) 对于自同构图交我们也称之为非对角交, 即$\phi:(X,\mathcal{X},\mu,T)\to(X,\mathcal{X},\mu,T)$ 为自同构, 那么也称$\mathrm{gr}(\mu,\phi)$ 为**非对角交**. 此时我们更多的是记为

$$\Delta_\mu^\phi=\mathrm{gr}(\mu,\phi).$$

尤其记$\mathrm{gr}(\mu,\mathrm{id})=\Delta_\mu$.

设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $\phi_1,\phi_2,\ldots,\phi_k\in\mathrm{Aut}(X)$, $k\ge 1$. 令

$$\mathrm{id}_X\times\phi_1\times\ldots\times\phi_k: X\to X^k,\quad x\mapsto(x,\phi_1x,\ldots,\phi_kx).$$

我们称$k+1$重交

$$\mathrm{gr}(\mu,\phi_1,\phi_2,\ldots,\phi_k)=(\mathrm{id}_X\times\phi_1\times\ldots\times\phi_k)_*\mu$$

为**非对角交***(off-diagonal)*.

一个k 重交λ 称为***POOD*** *(product of off-diagonal)* 是指存在$\{1,2,\ldots,k\}$ 的剖分$\{1,2,\ldots,k\}=J_1\cup J_2\cup\ldots\cup J_m$ 使得:

1. 对于每个$l \in \{1, 2, \ldots, m\}$, λ 到$\prod_{j\in J_l} X$ 的投射为非对角交;

2. $\prod_{j\in J_l} X, l = 1, 2, \ldots, m$ 为独立的.

定义 8.1.7 设$(X, \mathcal{X}, \mu, T)$ 为遍历系统, $k \geq 2$. 如果X 的每个k重遍历交为$POOD$, 那么我们称$(X, \mathcal{X}, \mu, T)$ 为k**阶简单的***(simple of order k, or k-simple)*. 如果对于任何$k \geq 2$, $(X, \mathcal{X}, \mu, T)$ 为k 阶简单的, 那么称为**简单的** *(simple)*.

称$(X, \mathcal{X}, \mu, T)$ 为k**阶极小自交***(minimal self-joining of order k, or k-MSJ)*, 是指它为k 阶简单的且$\mathrm{Aut}(X) = \{T^n : n \in \mathbb{Z}\}$.

易见遍历系统$(X, \mathcal{X}, \mu, T)$ 为2阶简单的当且仅当

$$J_e(X) = \{\mu \times \mu\} \cup \{\mathrm{gr}(\mu, \phi) : \phi \in \mathrm{Aut}(X)\}.$$

遍历系统$(X, \mathcal{X}, \mu, T)$ 为2阶极小自交的当且仅当

$$J_e(X) = \{\mu \times \mu\} \cup \{\mathrm{gr}(\mu, T^n) : n \in \mathbb{Z}\}.$$

下面的相对独立交的概念是由Furstenberg 引入的.

定义 8.1.8 (相对独立交) *(relatively independent joining)* 设$\pi : (X, \mathcal{X}, \mu, T) \to (Z, \mathcal{Z}, \eta, H)$, $\phi : (Y, \mathcal{Y}, \nu, S) \to (Z, \mathcal{Z}, \eta, H)$ 为遍历系统间的因子映射. 设积分分解为

$$\mu = \int_Z \mu_z d\eta(z), \quad \nu = \int_Z \nu_z d\eta(z).$$

$$\begin{tikzcd} X \arrow[dr, "\pi"'] & & Y \arrow[dl, "\phi"] \\ & Z & \end{tikzcd}$$

定义系统X 和Y 相对于公因子Z 的**相对独立交**为

$$\lambda_{\pi,\phi} = \int_Z (\mu_z \times \nu_z) d\eta(z).$$

一般把上面交记为

$$\mu \underset{\eta}{\times} \nu, \text{ 或 } \mu \underset{Z}{\times} \nu.$$

当Z 为平凡的时候, $\mu \underset{\eta}{\times} \nu = \mu \times \nu$.

我们可以推广上面的概念. 设$(X_i, \mathcal{X}_i, \mu_i, T_i), i = 1, \ldots, k$ 保测系统，设$(Y_i, \mathcal{Y}_i, \nu_i, S_i)$ 为相应的因子系统，$\pi_i : X_i \to Y_i$ 为对应的因子映射. 设ξ 为$Y_1, \ldots, Y_k$ 上的交，即它为乘积空间$\prod_i Y_i$ 上$S_1 \times \ldots \times S_k$ 不变并且到各个分量Y_j 上投射为ν_j 的测度. 对于每个$i = 1, \ldots, k$,

设$\mu_i = \int_{Y_i} \mu_{X_i,y_i}\, d\nu_i(y_i)$ 为测度μ_i 相对ν_i 的测度分解. 定义λ 为乘积空间$\prod_i X_i$ 上的测度如下：

$$\lambda = \int_{\prod_i Y_i} \mu_{X_1,y_1} \times \mu_{X_2,y_2} \times \ldots \times \mu_{X_k,y_k}\, d\xi(y_1, y_2, \ldots, y_k). \tag{8.1.1}$$

易验证λ 为$(X_i, \mathcal{X}_i, \mu_i, T_i), i = 1, \ldots, k$ 上的交，称之为**相对于ξ 的条件乘积测度**（conditional product measure relative to ν）. 等价地，λ 为相对于ξ 的条件乘积测度当且仅当对于任何$f_i \in L^\infty(X_i, \mu_i), i = 1, \ldots, k$

$$\begin{aligned}&\int_{\prod_i X_i} f_1(x_1)f_2(x_2)\ldots f_k(x_k)\, d\lambda(x_1, x_2, \ldots, x_k)\\ &\quad= \int_{\prod_i Y_i} \mathbb{E}(f_1|\mathcal{Y}_1)(y_1)\mathbb{E}(f_2|\mathcal{Y}_2)(y_2)\ldots\mathbb{E}(f_k|\mathcal{Y}_k)(y_k)\, d\xi(y_1, y_2, \ldots, y_k).\end{aligned} \tag{8.1.2}$$

如果$(Y_i, \mathcal{Y}_i, \nu_i, S_i) = (Y, \mathcal{Y}, \nu, S), 1 \leq i \leq k$ 并且ξ 取为Y^k 上的对角测度，则称此时得到的件乘积测度λ 为**相对于独立积**. 注意λ 表达式为

$$\lambda = \int_Y \mu_{X_1,y} \times \mu_{X_2,y} \times \ldots \times \mu_{X_k,y}\, d\nu(y).$$

记之为$\lambda = \mu \underset{\nu}{\times} \mu \underset{\nu}{\times} \ldots \underset{\nu}{\times} \mu$, 或者$\lambda = \mu \underset{Y}{\times} \mu \underset{Y}{\times} \ldots \underset{Y}{\times} \mu$.

更一般地，我们由如下概念. 设$(X_i, \mathcal{X}_i, \mu_i, T_i), i = 1, \ldots, k$ 保测系统，对于$i \in \{1, \ldots, k\}$ 设$(Y_i, \mathcal{Y}_i, \nu_i, S_i)$ 为$(X_i, \mathcal{X}_i, \mu_i, T_i)$ 的因子系统, 而$(Z_i, \mathcal{Z}_i, \eta_i, H_i)$ 为$(Y_i, \mathcal{Y}_i, \nu_i, S_i)$ 的因子. 设λ 为$(X_i, \mathcal{X}_i, \mu_i, T_i), i = 1, \ldots, k$ 上的交. 称$(Y_1, Y_2, \ldots, Y_k)$ **在交λ 下关于**$(Z_1, Z_2, \ldots, Z_k)$ **相对独立** 指对于任何$f_i \in L^\infty(Y_i, \nu_i), i = 1, \ldots, k$

$$\begin{aligned}&\int_{\prod_i X_i} f_1(x_1)f_2(x_2)\ldots f_k(x_k)\, d\lambda(x_1, x_2, \ldots, x_k)\\ &\quad= \int_{\prod_i Z_i} \mathbb{E}(f_1|\mathcal{Z}_1)(y_1)\mathbb{E}(f_2|\mathcal{Z}_2)(y_2)\ldots\mathbb{E}(f_k|\mathcal{Z}_k)(y_k)\, d\rho(z_1, z_2, \ldots, z_k),\end{aligned} \tag{8.1.3}$$

其中ρ 为μ 在$\prod_i Z_i$ 上的像.

注意如果上面空间只是测度空间，而非保测系统，我们可以定义类似的概念.

习 题

1. 证明：如果赋予$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 拓扑模型, 那么$J(X, Y)$ 成为$\mathcal{M}(X \times Y, T \times S)$ 的一个紧凸子集, 并且$J(X, Y)$ 的拓扑与$\mathcal{M}(X \times Y, T \times S)$ 上诱导的弱拓扑是吻合的, 这个拓扑与$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 拓扑模型的选取无关.
2. 证明定理8.1.4.

§8.2 从交到同构

在本节中我们给出交与同构的关系，作为应用给出Halmos-von Neumann 定理的简单证明.

§8.2.1　交与同构的关系

设$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 为保测系统, 定义

$$\mathcal{X} \times Y = \{A \times Y : A \in X\},$$

$$X \times \mathcal{Y} = \{X \times B : B \in \mathcal{Y}\}.$$

定理 8.2.1 设$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 为保测系统, $\lambda \in J(X, Y)$. 那么λ 为图交当且仅当

$$\mathcal{X} \times Y \supset X \times \mathcal{Y} \pmod{\lambda}.$$

λ 为同构图交当且仅当

$$\mathcal{X} \times Y = X \times \mathcal{Y} \pmod{\lambda}.$$

证明. 我们仅证明第二个命题, 第一个的证明是类似的.

设$\phi : X \to Y$ 为同构使得$\lambda = \mathrm{gr}(\mu, \phi)$. 于是对于任何$A \in \mathcal{X}$, 我们有

$$\begin{aligned}\lambda((A \times Y)\Delta(X \times \phi(A))) &= \lambda(A \times \phi(A)^c \cup A^c \times \phi(A)) \\ &= \mu(A \cap \phi^{-1}\phi(A^c)) + \mu(A^c \cap \phi^{-1}\phi(A)) \\ &= 0.\end{aligned}$$

类似地, 我们得到对于任何$B \in \mathcal{Y}$,

$$\lambda((\phi^{-1}B \times Y)\Delta(X \times B)) = 0.$$

反之, 假设$\mathcal{X} \times Y = X \times \mathcal{Y} \pmod{\lambda}$. 定义映射$\Phi : \mathcal{Y} \to \mathcal{X}$ 以及映射$\Psi : \mathcal{X} \to \mathcal{Y}$ 使得

$$\lambda((A \times Y)\Delta(X \times \Psi(A))) = 0$$

以及

$$\lambda((\Phi(B) \times Y)\Delta(X \times B)) = 0.$$

可验证由此定义的映射满足

$$\Phi \circ \Psi = \mathrm{id}_{\mathcal{X}}, \quad \Psi \circ \Phi = \mathrm{id}_{\mathcal{Y}}.$$

接着应用定理4.2.5 得到所需结论. □

由上结论, 我们容易得到下面定理.

定理 8.2.2 设$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 为保测系统, $\lambda \in J(X, Y)$. 那么

$$\mathcal{A} = \{A \in \mathcal{X} : \exists B \in \mathcal{Y}\ s.t.\ \lambda((A \times Y)\Delta(X \times B)) = 0\}$$

以及

$$\mathcal{B} = \{B \in \mathcal{Y} : \exists A \in \mathcal{X} \ s.t. \ \lambda((A \times Y)\Delta(X \times B)) = 0\}$$

分别为T-不变和S-不变子σ-代数. 并且我们有

$$\mathcal{A} \times Y = \mathcal{X} \times Y \cap X \times \mathcal{Y} = X \times \mathcal{B} \pmod{\lambda},$$

$\mathcal{A}, \mathcal{B}$ 它们对应的系统是同构的.

在上面定理中, 我们设$\mathcal{A}, \mathcal{B}$ 它们对应的系统为$(Z, \mathcal{Z}, \eta, H)$. 我们得到下面的图表

$$\begin{tikzcd}
 & (X \times Y, \lambda) \arrow[dl] \arrow[dr] & \\
(X, \mu) \arrow[dr] & & (Y, \nu) \arrow[dl] \\
 & (Z, \eta) &
\end{tikzcd}$$

称$(Z, \mathcal{Z}, \eta, H)$ 为$(X, \mathcal{X}, \mu, T)$ 和$(Y, \mathcal{Y}, \nu, S)$ 由λ **确定的公因子**. 当$X = Y$ 时, 我们直接称$(Z, \mathcal{Z}, \eta, H)$ 为$(X, \mathcal{X}, \mu, T)$ 由λ 确定的因子.

§8.2.2 Halmos-von Neumann定理的简单证明

下面作为应用, 我们给出Lemańczyk 关于Halmos-von Neumann 定理的简单证明.

定理 8.2.3 (Halmos-von Neumann) 设$(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 为遍历具有离散谱的*Lebesgue*系统. 那么$(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 谱同构当且仅当它们为同构的.

证明. 仅需证明, 如果$(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 谱同构, 那么它们为同构的. 设$\lambda \in J_e(X, Y)$ 为$(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 的一个遍历交. 对于T, S 的特征值α, 设$f \in L^2(X, \mu)$, $g \in L^2(Y, \nu)$ 为相应的特征函数. 那么在遍历系统$(X \times Y, \mathcal{X} \times \mathcal{Y}, \lambda, T \times S)$ 中,

$$\overline{f}(x, y) = f \otimes 1(x, y) = f(x), \quad \overline{g}(x, y) = 1 \otimes g(x, y) = g(y)$$

为特征值为α 的特征函数. 因为$(X \times Y, \mathcal{X} \times \mathcal{Y}, \lambda, T \times S)$ 遍历, 所以存在$c \in \mathbb{C}$ 使得

$$\overline{f} = c\overline{g} \pmod{\lambda}.$$

因为特征函数生成了$L^2(X, \mu), L^2(Y, \nu)$, 所以上式意味着

$$L^2(X, \mu) \otimes \mathbb{C} \underset{\lambda}{=} \mathbb{C} \times L^2(Y, \nu) \subseteq L^2(X \times Y, \lambda)$$

这等价于

$$\mathcal{X} \times Y = X \times \mathcal{Y} \pmod{\lambda}.$$

于是根据定理8.2.1, $(X, \mathcal{X}, \mu, T), (Y, \mathcal{Y}, \nu, S)$ 为同构的. $\square$

习 题

1. 设$\pi:(X,\mathcal{X},\mu,T)\to(Z,\mathcal{Z},\eta,H)$ 以及$\phi:(Y,\mathcal{Y},\nu,S)\to(Z,\mathcal{Z},\eta,H)$ 为因子映射. 证明: $(Z,\mathcal{Z},\eta,H)$ 为$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 由$\mu\underset{\eta}{\times}\nu$ 确定的公因子.

§8.3　交与混合

本节给出Ryzhikov 关于Furtenberberg 关于弱混合一些经典结论运用交的证明.

引理 8.3.1 设$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为保测系统. 设$\lambda\in J(X,Y)$ 为交, 设$\lambda\times_Y\lambda$ 为$(X\times Y,\lambda,T\times S)$ 相对于于因子$(Y,\mathcal{Y},\nu,S)$ 的相对独立交. 如果$\lambda\underset{\nu}{\times}\lambda$ 到$X\times X$ 的投射是$\mu\times\mu$, 那么

$$\lambda=\mu\times\nu.$$

证明. 设$A\in\mathcal{X}$, 根据$\lambda\times_Y\lambda$ 的定义,

$$\lambda\underset{\nu}{\times}\lambda(A\times Y\times A)=\int_Y\mathbb{E}_\lambda(\mathbf{1}_A|Y)^2d\nu(y).$$

于是, 如果$\lambda\times_Y\lambda$ 到$X\times X$ 的投射是$\mu\times\mu$, 那么我们就有

$$\mu(A)^2=\int_Y\mathbb{E}_\lambda(\mathbf{1}_A|Y)^2d\nu(y).$$

于是

$$\left(\int_Y\mathbb{E}_\lambda(\mathbf{1}_A|Y)d\nu(y)\right)^2=\mu(A)^2=\int_Y\mathbb{E}_\lambda(\mathbf{1}_A|Y)^2d\nu(y),\ \forall A\in\mathcal{X}.$$

根据Cauchy不等式等号成立条件, $\mathbb{E}_\lambda(\mathbf{1}_A|Y)$ 为常值, 由此$\lambda=\mu\times\nu$. 证毕! □

定理 8.3.2 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 如果$T\times T$ 为遍历的（即T 弱混合）, 那么对于任何遍历系统$(Y,\mathcal{Y},\nu,S)$, $(X\times Y,\mu\times\nu,T\times S)$ 为遍历的.

证明. 由条件, $\mu\times\mu\in J(X)$ 为遍历自交, 我们要证对于任何遍历系统$(Y,\mathcal{Y},\nu,S)$, $\mu\times\nu\in J_e(X,Y)$. 为此根据定理8.1.4我们需要证明$\mu\times\nu$ 为凸集$J(X,Y)$ 的端点.

我们取$p\in(0,1)$ 以及$\lambda_1,\lambda_2\in J(X,Y)$ 使得

$$\mu\times\nu=p\lambda_1+(1-p)\lambda_2.$$

注意$\mu\times\nu$ 关于Y 的相对独立交为$\mu\times\nu\times\mu$, 即$(\mu\times\nu)\underset{\nu}{\times}(\mu\times\nu)=\mu\times\nu\times\mu$. 注意

$$\begin{aligned}\mu\times\nu&=(p\lambda_1+(1-p)\lambda_2)\underset{\nu}{\times}(p\lambda_1+(1-p)\lambda_2)\\&=p^2\lambda_1\underset{\nu}{\times}\lambda_1+(1-p)^2\lambda_2\underset{\nu}{\times}\lambda_2+p(1-p)\lambda_1\underset{\nu}{\times}\lambda_2+p(1-p)\lambda_2\underset{\nu}{\times}\lambda_1.\end{aligned}$$

因为$\mu\times\mu$ 为遍历的, 所以上面每一项中的交到$X\times X$ 的投射为$\mu\times\mu$. 于是λ_1,λ_2 都满足引理8.3.1 条件, 从而

$$\lambda_1=\lambda_2=\mu\times\nu.$$

由此$\mu\times\nu$ 为$J(X,Y)$ 端点, 即它为遍历交. □

根据交的语言, 一个保测系统$(X,\mathcal{X},\mu,T)$ 为强混合的当且仅当

$$\Delta_\mu^{T^n} \xrightarrow{n\to\infty} \mu\times\mu.$$

定理 8.3.3 (Ornstein) 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 那么T 为强混合的当且仅当T 为弱混合的且存在$\theta>0$ 使得

$$\limsup_{n\to\infty}\mu(A\cap T^{-n}B)\le\theta\mu(A)\mu(B),\quad \forall A,B\in\mathcal{X}. \tag{8.3.1}$$

证明. 设T 为强混合的, 那么它为弱混合的. 我们也可以如下直接证明. 因为强混合, 所以$\Delta_\mu^{T^n} \xrightarrow{n\to\infty} \mu\times\mu$. 于是

$$\Delta_{\mu\times\mu}^{(T\times T)^n} \xrightarrow{n\to\infty} \mu^4=\mu\times\mu\times\mu\times\mu.$$

所以$T\times T$ 不变子集的测度必为1或0, 从而T 弱混合. (8.3.1) 是显然的.

反之, 设T 为弱混合且(8.3.1) 成立. 设λ 为$\{\Delta_\mu^{T^n}\}_n$ 在$J(X)$ 中得到任何聚点, 由(8.3.1),

$$\lambda\le\theta(\mu\times\mu).$$

尤其

$$\lambda\ll\mu\times\mu.$$

因为T 弱混合, $\mu\times\mu$ 遍历, 所以上式表明

$$\lambda=\mu\times\mu.$$

由此我们得到

$$\Delta_\mu^{T^n} \xrightarrow{n\to\infty} \mu\times\mu.$$

即T 为强混合的. □

在交理论中一个著名的问题是:

问题 8.3.4 *是否存在零熵系统$(X,\mathcal{X},\mu,T)$ 以及$\lambda\in J_3(X)$ 使得λ 两两独立, 但是$\lambda\ne\mu\times\mu\times\mu$?*

§8.4 不交与相对独立交

在本节中我们研究不交性，类似于两个自然数为互素，两个系统为不交的体现了两系统有较大的差异.

§8.4.1　不交的定义

定义 8.4.1 设$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为保测系统, 称它们是**不交的** *(disjoint)* 是指$J(X,Y)=\{\mu\times\nu\}$. 记为$X\perp Y$.

定义 8.4.2 设$\pi:(X,\mathcal{X},\mu,T)\to(Z,\mathcal{Z},\eta,H)$, $\phi:(Y,\mathcal{Y},\nu,S)\to(Z,\mathcal{Z},\eta,H)$ 为遍历系统间的因子映射.

$$\begin{tikzcd} X \arrow[dr, "\pi"'] & & Y \arrow[dl, "\phi"] \\ & Z & \end{tikzcd}$$

称X 和Y **相对于Z 是不交的**, 指$\mu\underset{\eta}{\times}\nu$ 为唯一映到Δ_η 的交. 记为$\pi\perp\phi$ 或者$X\perp_Z Y$.

命题 8.4.3 *1.* 设$(X,\mathcal{X},\mu,T)$ 为非平凡的保测系统, 那么$X\not\perp X$.

2. 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为保测系统间的非平凡因子映射, 并且$(Y,\mathcal{Y},\nu,S)$ 为非平凡的保测系统. 那么$X\not\perp Y$.

证明. 作为习题. □

§8.4.2　遍历与恒同

遍历性体现的是系统的不可分割性，而恒同系统每个点都是不动点，这两种性质是相对立的。下面定理体现了这种对立性.

定理 8.4.4 $(Y,\mathcal{Y},\nu,S)$ 为遍历的当且仅当它与所有恒同系统$(X,\mathcal{X},\mu,\mathrm{id}_X)$ 不交.

证明. 设$(Y,\mathcal{Y},\nu,S)$ 为遍历的，并且$(X,\mathcal{X},\mu,\mathrm{id}_X)$ 为恒同系统. 下证$X\perp Y$. 设$\lambda\in J(X,Y)$. 这对于任何$A\in\mathcal{X},B\in\mathcal{Y}$, 根据$\lambda$ 为$\mathrm{id}\times S$ 不变的, 我们得到

$$\begin{aligned}\lambda(A\times B)&=\lambda(A\times S^{-n}B)\\&=\frac{1}{N}\sum_{n=0}^{N-1}\lambda(A\times S^{-k}B)\\&=\int_{A\times Y}\frac{1}{N}\sum_{n=0}^{N-1}\mathbf{1}_B(S^ny)d\lambda(x,y).\end{aligned}$$

因为S 为遍历的, 所以

$$\frac{1}{N}\sum_{n=0}^{N-1}\mathbf{1}_B(S^ny)\xrightarrow{a.e}\nu(B),\quad N\to\infty.$$

于是

$$\lambda(A\times B)=\mu(A)\nu(B),\ \forall A\in\mathcal{X},B\in\mathcal{Y}.$$

即$\lambda=\mu\times\nu$.

反之， 如果$(Y,\mathcal{Y},\nu,S)$ 与所有恒同系统$(X,\mathcal{X},\mu,\mathrm{id}_X)$ 不交， 那么根据命题8.4.3, $(Y,\mathcal{Y},\nu,S)$ 没有非平凡的恒同因子，尤其它为遍历的. □

作为应用, 我们给出Ryzhikov 关于Furtenberberg关于多重弱混合定理的一个证明.

定理 8.4.5 (Furstenberg) 设$(X,\mathcal{X},\mu,T)$ 为弱混合系统. 对于任何$k\in\mathbb{N}$ 以及$A_0,A_1,\ldots,A_k$, 我们有

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(A_0\cap T^{-n}A_1\cap T^{-2n}A_2\cap\ldots\cap T^{-kn}A_k)\xrightarrow{N\to\infty}\mu(A_0)\mu(A_1)\ldots\mu(A_k).$$

证明. 对于任何$N\in\mathbb{N}$, 考虑$(X,\mathcal{X},\mu,T)$ 的$k+1$ 阶自交：

$$\lambda_N=\frac{1}{N}\sum_{n=0}^{N-1}\mu(A_0\cap T^{-n}A_1\cap T^{-2n}A_2\cap\ldots\cap T^{-kn}A_k).$$

于是命题所要证明的就是

$$\lambda_N\xrightarrow{N\to\infty}\mu\times\mu\times\ldots\times\mu.$$

我们设λ 为$\{\lambda_N\}_N$ 的任何聚点, 我们仅需证明

$$\lambda=\mu\times\mu\times\ldots\times\mu.$$

注意λ 为$\mathrm{id}\times T\times T^2\times\ldots\times T^k$-不变的.

我们对k 归纳证明. $k=1$, 那么根据定理8.4.4 以及T 为遍历的直接得到. 设$k-1$ 时已经成立, 那么根据归纳假设, λ 到后面k 个坐标的投射为μ^k, 于是λ 为(X,μ,id) 与$(X^k,\mu^k,T\times T^2\times\ldots\times T^k)$ 的交. 因为$T\times T^2\times\ldots\times T^k$ 为遍历的, 根据定理8.4.4 我们就有

$$\lambda=\mu^{k+1}=\mu\times\mu\times\ldots\times\mu.$$

证毕! □

下面给出定理8.4.4 的相对化版本.

定义 8.4.6 称因子映射$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为**遍历扩充** 是指X 的任何T 不变函数是$\mathcal{Y}$ 可测的, 即$\mathcal{I}(T)\subseteq\mathcal{Y}$.

定理 8.4.7 设$\phi:(Y,\mathcal{Y},\nu,S)\to(Z,\mathcal{Z},\eta,\mathrm{id}_Z)$ 为因子映射. 那么ϕ 为遍历扩充当且仅当$\phi\perp\pi$, 其中$\pi:(X,\mathcal{X},\mu,\mathrm{id}_X)\to(Z,\mathcal{Z},\eta,\mathrm{id}_Z)$ 为任何扩充.

尤其, $(Y,\mathcal{Y},\nu,S)$ 为遍历的当且仅当它与所有恒同系统$(X,\mathcal{X},\mu,\mathrm{id})$ 不交.

证明. 设ϕ 为遍历扩充. 设λ 为系统$(Y,\mathcal{Y},\nu,S)$ 和$(X,\mathcal{X},\mu,\mathrm{id}_X)$ 相对于系统$(Z,\mathcal{Z},\eta,\mathrm{id}_Z)$ 的交.

$$\begin{tikzcd}
(X,\mu,\mathrm{id}_X) \arrow[dr, "\pi"'] & & (Y,\nu,S) \arrow[dl, "\phi"] \\
& (Z,\eta,\mathrm{id}_Z) &
\end{tikzcd}$$

设

$$\mu=\int_Z \mu_z d\eta(z),\quad \nu=\int_Z \nu_z d\eta(z)$$

为μ,ν 相对η 的测度分解. 设λ 相对Y 的测度分解为

$$\lambda=\int_Y \lambda_y\times\delta_y d\nu(y).$$

于是根据λ 为$\mathrm{id}_X\times S$ 不变的, 我们有

$$\lambda=(\mathrm{id}_X\times S)_*\lambda=\int_Y \lambda_y\times\delta_{Sy}d\nu(y)=\int_Y \lambda_{S^{-1}y}\times\delta_y d\nu(y).$$

因为测度分解是唯一的, 所以

$$\lambda_y=\lambda_{S^{-1}y},\ \nu-a.e.\ y\in Y.$$

因为ϕ 是遍历扩充, 于是(作为习题验证)

$$\lambda_y=\lambda_{\phi(y)},\quad \nu-a.e.$$

设$p_X: X\times Y\to X$ 为投射, 根据$(p_X)_*\lambda=\mu$, 我们有

$$\mu=\int_Y \lambda_y d\nu(y)=\int_Y \lambda_{\phi(y)}d\nu(y).$$

于是我们就有(作为习题验证)

$$\lambda_y=\lambda_{\phi(y)}=\mu_{\phi(y)},\quad \nu-a.e.$$

由此

$$\begin{aligned}
\lambda&=\int_Y \mu_{\phi(y)}\times\delta_y d\nu(y)=\int_Z\int_Y \mu_z\times\delta_y d\nu_z(y)d\eta(z)\\
&=\int_Z \mu_z\times\left(\int_Y \delta_y d\nu_z(y)\right)d\eta(z)=\int_Z \mu_z\times\nu_z d\eta(z)\\
&=\mu\underset{\eta}{\times}\nu.
\end{aligned}$$

反之, 设ϕ 不是遍历的, 那么存在一个S-不变的$\mathcal{Y}$函数但不是$\mathcal{Z}$ 可测的函数f. 设f 满足$0\le f\le 1$, $\int_Y f d\nu=\frac{1}{2}$. 令$X=Z\times\{0,1\}$, $\mu=\eta\times\frac{1}{2}(\delta_0+\delta_1)$. 定义测度

$$\lambda_0=2\int_Z(\delta_z\times\delta_0\times f\nu_z)d\eta(z),\quad \lambda_1=2\int_Z(\delta_z\times\delta_1\times(1-f)\nu_z)d\eta(z).$$

令$\lambda=\frac{1}{2}(\lambda_0+\lambda_1)$. 易验证$\lambda$ 为映到Δ_η 但是不等于$\mu\underset{\eta}{\times}\nu$. 证毕! □

§8.4.3 不交性基本定理

类似于定理8.2.2, 对于测度空间$(X,\mathcal{X},\mu)$ 和$(Y,\mathcal{Y},\nu)$ 以及它上面的交λ (即λ 为$X\times Y$ 上测度, 它到X,Y 的投射分别为μ 和ν), 我们也可以定义由λ 确定的公因子. 设$\mathcal{A}=\{A\in\mathcal{X}: \exists B\in\mathcal{Y}$ 使得 $\lambda((A\times Y)\Delta(X\times B))=0\}$ 以及$\mathcal{B}=\{B\in\mathcal{Y}:\exists A\in\mathcal{X}$ 使得 $\lambda((A\times Y)\Delta(X\times B))=0\}$. 那么$\mathcal{A}\times Y=\mathcal{X}\times Y\cap X\times\mathcal{Y}=X\times\mathcal{B} \pmod\lambda$, $\mathcal{A}$, $\mathcal{B}$ 它们对应的测度空间是同构的. 设$\mathcal{A}$, $\mathcal{B}$ 它们对应的测度空间为$(Z,\mathcal{Z},\eta)$. 我们得到下面的图表

$$\begin{tikzcd}
 & (X\times Y,\lambda) \arrow[dl] \arrow[dr] & \\
(X,\mu) \arrow[dr] & & (Y,\nu) \arrow[dl] \\
 & (Z,\eta) &
\end{tikzcd}$$

称$(Z,\mathcal{Z},\eta)$ 为$(X,\mathcal{X},\mu)$ 和$(Y,\mathcal{Y},\nu)$ 由 λ 确定的公因子.

定理 8.4.8 设$(X,\mathcal{X},\mu),(Y,\mathcal{Y},\nu)$ 为*Lebesgue* 空间, λ 为$X\times Y$ 上到X,Y 投射分别为μ,ν 的测度. 设λ 相对Y 的测度分解为

$$\lambda=\int_Y \lambda_y\times\delta_y d\nu(y).$$

定义$X^{\mathbb{Z}}\times Y$ 上的测度λ_∞

$$\lambda_\infty=\int_Y(\ldots\times\lambda_y\times\lambda_y\times\ldots)\times\delta_y d\nu(y)$$

以及$X^{\mathbb{Z}}$ 上的测度μ_∞

$$\mu_\infty=\int_Y(\ldots\times\lambda_y\times\lambda_y\times\ldots)d\nu(y).$$

设$\mathcal{Z}$ 为σ-代数$\mathcal{X}^{\mathbb{Z}}$ 和$\mathcal{Y}$ 相对于测度λ_∞ 的最大子σ-代数, 相应的*Lebesgue*空间记为$(Z,\mathcal{Z},\eta)$. 那么在测度λ_∞ 下, σ-代数$\mathcal{X}^{\mathbb{Z}}$ 和$\mathcal{Y}$ 相对于$\mathcal{Z}$ 是相对独立的.

尤其, 如果$(X,\mathcal{X},\mu,T)$ 和$(Y,\mathcal{Y},\nu,S)$ 为保测系统, $\lambda\in J(X,Y)$, 那么所对应的λ_∞ 为$(X^{\mathbb{Z}},\mu_\infty)$ 与(Y,ν) 相对于公因子(Z,η) 的相对独立交. 我们有如下图标:

$$\begin{tikzcd}
 & & (X^{\mathbb{Z}}\underset{\eta}{\times}Y,\lambda_\infty) \arrow[dll, "\pi"'] \arrow[dl] \arrow[dr] & \\
(X\times Y,\lambda) & (X^{\mathbb{Z}},\mu_\infty) \arrow[dr] & & (Y,\nu) \arrow[dl] \\
 & & (Z,\eta) &
\end{tikzcd}$$

在图表中, $\pi=p_0\times\mathrm{id}_Y: X^{\mathbb{Z}}\times Y\to X\times Y$, 其中$p_0: X^{\mathbb{Z}}\to X$ 为向第0 坐标的投射.

证明. 定义变换

$$S : X^{\mathbb{Z}} \times Y \to X^{\mathbb{Z}} \times Y,\ S(\mathbf{x}, y) = (\sigma \mathbf{x}, y),$$

其中$\mathbf{x} = (x_n)_{n\in\mathbb{Z}}$, σ 为$X^{\mathbb{Z}}$ 上的转移. 如果$f(\mathbf{x}, y)$ 为$X^{\mathbb{Z}} \times Y$ 上的S 不变函数, 那么对于每个$y \in Y$, 函数$f_y(\mathbf{x}) = f(\mathbf{x}, y)$ 为$(X^{\mathbb{Z}}, \lambda_y^{\mathbb{Z}})$ 上σ 不变的函数. 因为$(X^{\mathbb{Z}}, \lambda_y^{\mathbb{Z}}, \sigma)$ 为Bernoulli 系统, 函数$f_y(\mathbf{x})$ 为常值函数, 即

$$f(\mathbf{x}, y) = f(y),\ \lambda_\infty - a.e.$$

于是每个S不变函数都是$\mathcal{Y}$ 可测的, 尤其扩充

$$(X^{\mathbb{Z}}, \mu_\infty, \sigma) \to (Z, \eta, \mathrm{id}_Z)$$

为遍历扩充. 对于下图表运用定理8.4.7.

$$\begin{tikzcd}
 & (X^{\mathbb{Z}} \times Y, \lambda_\infty, S \times \mathrm{id}_Y) \arrow[dl] \arrow[dr] & \\
(X^{\mathbb{Z}}, \mu_\infty, S) \arrow[dr] & & (Y, \nu, \mathrm{id}_Y) \arrow[dl] \\
 & (Z, \eta, \mathrm{id}_Z) &
\end{tikzcd}$$

我们得到$\mathcal{X}^{\mathbb{Z}}$ 与$\mathcal{Y}$ 相对于$\mathbb{Z}$ 独立. □

根据定理8.4.8, 我们可以得到下面定理. 其中第一条我们前面已经给出过，剩下几条的证明作为习题请读者完成.

定理 8.4.9 • *恒同$\perp$ 遍历. 事实上, 保测系统为遍历的当且仅当它与所有恒同系统不交.*

- *distal $\perp$ 弱混合. 事实上, 保测系统为弱混合的当且仅当它与所有distal 系统不交.*
- *rigid $\perp$ mild 混合. 事实上, 保测系统为mild混合的当且仅当它与所有rigid系统不交.*
- *零熵$\perp$ K系统. 事实上, 保测系统为K系统当且仅当它与所有零熵系统不交.*

习 题

1. 证明命题8.4.3.
2. 给出定理8.4.7证明中两处需要补充说明的细节.
3. 证明定理8.4.9.

§8.5 交在Sarnak 猜测研究中的应用

在本小节我们介绍著名的Sarnak 猜测. 我们首先给出Sarnak 猜测的陈述, 然后证明测度版本的Sarnak 定理, 最后运用交的方法证明著名的Chowla 猜测蕴含Sarnak 猜测.

§8.5.1 什么是Sarnak不交性猜测?

熵理论是动力系统的核心内容之一. 从熵的角度我们可以把系统分为正熵系统和零熵系统两大类. 从概率论的角度看, 零熵系统可以视为“确定性”的系统, 而正熵被认为是“不确定”的. 幂零系统是一类特殊零熵系统, 受 Green 和 Tao 关于 Hardy-Littlewood 猜测系列工作的影响[86, 87], Wolf 奖得主 Sarnak 在 2009 年提出一个与熵理论和数论有关的关于遍历平均的著名猜测：Möbius 函数 $\boldsymbol{\mu}$ 与零熵序列 ξ 渐近正交.

如在熵理论章节中所述, 1958 年Kolmogorov [133]借鉴Shannon 在信息论中不确定性的描述在遍历论中引入了熵的概念. 熵是重要的同构不变量, 它反映了系统的混乱程度. 随后, Adler 等人[2]在拓扑动力系统中引入了拓扑熵的概念, Bowen [36]使用分离集和张成集给出拓扑熵一种新的定义方式. Goodwyn、Dianburg 和Goodman 证明了熵的”变分原理”: 动力系统(X,T) 的拓扑熵$h_{\text{top}}(T)$ 是测度熵$h_\mu(T)$ 的上确界, 其中μ 取遍T 不变测度. 从熵的角度我们可以把系统分为正熵系统和零熵系统两大类. 从概率论的角度看, 零熵系统可以视为“确定性”的系统, 而正熵被认为是“不确定”的.

自从熵被引入遍历理论和拓扑动力系统中以来, 它一直是动力系统研究中的重要内容. 熵反映了系统的复杂程度, 熵越大则系统越复杂, 广义上的熵可以看作是某种集合特征量（比如维数, 信息量等）在动力系统作用下的平均. 由于正熵系统存在Shannon-McMillan-Breiman定理等性质, 在附加正熵的条件下Furstenberg $\times_2, \times_3$ 猜测[174]、Littlewood 猜测[53] 等测度刚性问题已经得到解决；也有一些问题根据Pinsker 因子可以化归到零熵系统的情况, 例如逐点收敛的多重遍历平均问题[48]、Rohlin问题[82] 等. 还有一些问题就是直接针对零熵系统提出的, 例如由2009 年Wolf 奖得主Sarnak 提出的一个与动力系统和数论有关的著名猜想－－Sarnak 猜测[179]. 对于零熵系统, 因为缺乏有效的工具, 所以相关问题的研究进展比较缓慢. 在本节我们主要介绍Sarnak 猜测的一些基本信息, 最新的进展请参见最新的相关文献.

定义 8.5.1 称$\boldsymbol{\mu}(n)$ 为 ***Möbius* 函数**, 指

$$\boldsymbol{\mu}(n)=\begin{cases}1, & n=1;\\ 0, & n \text{ 有某个素数的平方作为因子};\\ (-1)^t, & n \text{ 分解为} t \text{ 个互不一样素数的乘积}.\end{cases}$$

***Liouville* 函数** $\boldsymbol{\lambda}:\mathbb{N}\to\{-1,1\}$ 定义为

$$\boldsymbol{\lambda}(n)=(-1)^{\Omega(n)},$$

其中$\Omega(n)$ 为n 素因子的个数（同一个素因子记重数在内）.

Möbius 函数和Liouville 函数是数论中最重要函数之一. 一个基本的数论结果是素数定理等价于 Möbius 函数(Liouville 函数)与常值序列1 渐近正交, 即

$$\sum_{n=1}^{N}\boldsymbol{\mu}(n)=\sum_{n=1}^{N}\boldsymbol{\lambda}(n)=o(N).$$

而黎曼zeta 函数与Möbius 函数关系如下：对于任何满足Re $s>1$ 的$s\in\mathbb{C}$

$$\frac{1}{\zeta(s)}=\sum_{n=1}^{\infty}\frac{\boldsymbol{\mu}(n)}{n^s}.$$

黎曼猜测等价于

$$\sum_{n=1}^{N}\boldsymbol{\mu}(n)=O_\varepsilon(N^{\frac{1}{2}+\epsilon}),N\to\infty,\forall\varepsilon>0.$$

关于更多Möbius 函数和Liouville 函数的性质参见[150]等.

在本节我们主要讨论Chowla 猜测和Sarnak 猜测. 数论中的著名的Chowla 猜测指：

问题 8.5.2 (Chowla 猜测) 设$1\le a_1<a_2<\ldots<a_r$, 设$k_1,k_2,\ldots,k_r\in\{1,2\}$ 不全都是偶数, 那么

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\boldsymbol{\mu}^{k_0}(n)\boldsymbol{\mu}^{k_1}(n+a_1)\boldsymbol{\mu}^{k_2}(n+a_2)\ldots\boldsymbol{\mu}^{k_t}(n+a_t)=0.$$

定义 8.5.3 设(X,T) 是一个拓扑动力系统, 即X 为紧致度量空间, $T:X\to X$ 为连续的. 一个序列$\{\xi_n\}_n$ 称为与**系统**(X,T) **关联的序列**, 是指存在点$x\in X$ 和连续函数$f\in C(X)$ 使得$\xi_n=f(T^nx)$. 此时, 我们也称序列$\{\xi_n\}_n$ **可由系统**(X,T) **实现**.

问题 8.5.4 (Sarnak 猜测) *Möbius* 函数 μ 与零熵序列 ξ 渐近正交, 即

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\boldsymbol{\mu}(n)\xi(n)=0,$$

这里零熵序列是指由零熵系统实现的序列.

注记 8.5.5 *1.* $\xi(n)\equiv 1$ 的情况就是素数定理, 常值序列 *1* 是最简单零熵序列, 它相对应于不动点;

2. 等差数列形式的素数定理等价于 *Möbius* 函数与周期序列渐近正交, 而周期序列可以由周期系统实现.

3. *1937* 年 *Davenport* 证明了 *Möbius* 函数与拟周期序列渐近正交, 而拟周期序列可以视为由拟周期系统（即环面上的无理旋转）实现的序列.

在文献[179]中, Sarnak 详细解释了提出上面猜测的原因, 例如文章提到的下面结论有助于大家理解这个猜测. 设$X_{\boldsymbol{\mu}}$ 为$\boldsymbol{\mu} \in \{-1,0,1\}^{\mathbb{N}}$ 在$(\{-1,0,1\}^{\mathbb{N}}, S)$ 中的轨道闭包系统, Sarnak 证明了下面定理:

定理 8.5.6 *[179]*

1. $\frac{6}{\pi^2}\log 2 \le h_{\text{top}}(X_{\boldsymbol{\mu}}, S) \le \frac{6}{\pi^2}\log 3$;
2. $(X_{\boldsymbol{\mu}}, S)$ 为*proximal* 系统, $\mathbf{0} = (000\ldots)$ 为其唯一的极小子集;
3. 取$(X_{\boldsymbol{\mu}}, S)$ 上不变测度, 作为保测系统它与*Kronecker* 系统有非平凡的交, 尤其它不为测度弱混合的.

我们将证明Sarnak 猜测是比Chowla 猜测要弱一些的问题. 目前这个问题进展不大, 关于Sarnak 猜测与Chowla 猜测的关系以及Sarnak 猜测的一些进展可以参见[179].

§8.5.2 测度版本Sarnak 定理

在这部分我们运用Devenport 估计证明测度版本的Sarnak 定理.

引理 8.5.7 (Davenport估计(Davenport's estimation)) *[46]* 对于任何$A > 0$, 存在$C_A > 0$ 使得对于任何$N \ge 2$, 我们有

$$\max_{z\in\mathbb{T}} \left| \sum_{n=1}^{N} z^n \boldsymbol{\mu}(n) \right| \le C_A \frac{N}{\log^A N}.$$

定理 8.5.8 (Sarnak) 设$(X, \mathcal{X}, \mu, T)$ 为*Lebesgue*系统, $f \in L^1(X, \mathcal{X}, \mu)$. 那么对于$\mu - a.e\ x \in X$, 我们有

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\boldsymbol{\mu}(n) \xrightarrow{N\to\infty} 0. \tag{8.5.1}$$

证明. 根据遍历分解定理, 我们不妨假设系统为遍历的. 取定$f \in L^1(X, \mathcal{X}, \mu)$. 根据谱定理我们有

$$\left\| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)\boldsymbol{\mu}(n) \right\|_2 = \left\| \frac{1}{N}\sum_{n=1}^{N} z^n \boldsymbol{\mu}(n) \right\|_{L^2(\mathbb{S}^1, \sigma_f)},$$

其中σ_f 为f 的谱测度. 根据Davenport估计, 对于任何$A > 0$, 存在$C_A > 0$ 使得

$$\left\| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)\boldsymbol{\mu}(n) \right\|_2 \le C_A \frac{N}{\log^A N}. \tag{8.5.2}$$

取$\rho > 1$, 在上式中设$N = [\rho^m], m \in \mathbb{N}$, 我们得到

$$\left\| \frac{1}{N}\sum_{n=1}^{N} f(T^n x)\boldsymbol{\mu}(n) \right\|_2 \le C_A \frac{N}{(m\log\rho)^A}, \forall A > 0.$$

令$A=2$, 且对m 求和得到:

$$\sum_{m=1}^{\infty}\left\|\frac{1}{[\rho^m]}\sum_{n=1}^{[\rho^m]}f(T^nx)\boldsymbol{\mu}(n)\right\|_2\leq\infty.$$

由此得到

$$\sum_{m=1}^{\infty}\left|\frac{1}{[\rho^m]}\sum_{n=1}^{[\rho^m]}f(T^nx)\boldsymbol{\mu}(n)\right|\in L^2(X,\mathcal{X},\mu),$$

并且上面求和是几乎处处有限的. 于是对于几乎处处的点$x\in X$, 我们有

$$\frac{1}{[\rho^m]}\sum_{n=1}^{[\rho^m]}f(T^nx)\boldsymbol{\mu}(n)\longrightarrow 0, m\to\infty. \tag{8.5.3}$$

假设$f\in L^\infty(X,\mathcal{X},\mu)$. 如果$[\rho^m]\leq N<[\rho^{m+1}]+1$, 那么有

$$\begin{aligned}\left|\frac{1}{N}\sum_{n=1}^{N}f(T^nx)\boldsymbol{\mu}(n)\right|&=\left|\frac{1}{N}\sum_{n=1}^{[\rho^m]}f(T^nx)\boldsymbol{\mu}(n)+\frac{1}{N}\sum_{n=[\rho^m]+1}^{N}f(T^nx)\boldsymbol{\mu}(n)\right|\\&\leq\left|\frac{1}{[\rho^m]}\sum_{n=1}^{[\rho^m]}f(T^nx)\boldsymbol{\mu}(n)\right|+\frac{\|f\|_\infty}{[\rho^m]}(N-[\rho^m])\\&\leq\left|\frac{1}{[\rho^m]}\sum_{n=1}^{[\rho^m]}f(T^nx)\boldsymbol{\mu}(n)\right|+\frac{\|f\|_\infty}{[\rho^m]}([\rho^{m+1}]-[\rho^m]).\end{aligned}$$

因为$\lim_{m\to\infty}\frac{\|f\|_\infty}{[\rho^m]}([\rho^{m+1}]-[\rho^m])=\|f\|_\infty(\rho-1)$, 根据(8.5.3), 我么得到

$$\frac{1}{N}\sum_{n=1}^{N}f(T^nx)\boldsymbol{\mu}(n)\xrightarrow{a.e.}\|f\|_\infty(\rho-1), N\to\infty.$$

令$\rho\to 1$, 我们就得到

$$\frac{1}{N}\sum_{n=1}^{N}f(T^nx)\boldsymbol{\mu}(n)\xrightarrow{a.e.}0, N\to\infty.$$

下设$f\in L^1(X,\mathcal{X},\mu)$. 对于任何$\varepsilon>0$, 取$g\in L^\infty(X,\mathcal{X},\mu)$ 使得$\|f-g\|_1<\varepsilon$. 根据Birkhoff逐点遍历定理, 对于几乎处处$x\in X$,

$$\lim_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}(f-g)(T^nx)\right|<\varepsilon.$$

于是

$$\begin{aligned}&\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}f(T^nx)\boldsymbol{\mu}(n)\right|\\\leq&\lim_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}(f-g)(T^nx)\right|+\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}g(T^nx)\boldsymbol{\mu}(n)\right|\\\leq&\varepsilon.\end{aligned}$$

因为ε 任意, 我们完成证明. □

§8.5.3 Sarnak 猜测与Chowla 猜测

在这部分我们运用交的方法证明下述定理:

定理 8.5.9 (Sarnak) *Chowla 猜测蕴含Sarnak 猜测.*

我们实际上要证明更为一般的结论, 亦即下面满足定义的Chowla 条件的序列也必满足Sarnak 条件.

定义 8.5.10 我们称$z \in \{-1,0,1\}^{\mathbb{N}}$ 满足***Chowla* 条件** 是指对于任何$1 \le a_1 < a_2 < \ldots < a_r$, 以及$k_0, k_1, k_2, \ldots, k_r \in \{1,2\}$ 不全都是偶数, 那么

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} z^{k_0}(n) z^{k_1}(n+a_1) z^{k_2}(n+a_2) \ldots z^{k_r}(n+a_r) = 0.$$

定义 8.5.11 称序列$z \in \{-1,0,1\}^{\mathbb{N}}$ 满足***Sarnak* 条件** 是指

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} z(n)\xi(n) = 0,$$

其中ξ 为由零熵系统实现的序列.

设S 为符号系统σ_k 上的转移映射. 对于$w \in \Sigma_k$, 我们用$w(n)$ 和w_n 来记w 得第n 个分量, $n \in \mathbb{Z}$. 在本节中π 指下面因子映射:

$$\pi : \{-1,0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}},\ (\pi(w))_n = w_n^2,\ \forall w \in \{-1,0,1\}^{\mathbb{N}}, n \in \mathbb{N}.$$

易见π 为$(\{-1,0,1\}^{\mathbb{N}}, S)$ 到$(\{0,1\}^{\mathbb{N}}, S)$ 的因子映射. 对于$\nu \in \mathcal{M}(\{0,1\}^{\mathbb{N}}, S)$, 令$\widehat{\nu} \in \mathcal{M}(\{-1,0,1\}^{\mathbb{N}}, S)$ 为ν 相对独立交：对于任何$\{-1,0,1\}$ 上词B,

$$\widehat{\nu}(B) = 2^{-|\mathrm{supp}(B)|}\nu(\pi(B)) = 2^{-|\mathrm{supp}(B)|}\nu(B^2),$$

其中$\mathrm{supp}(B) = \{i : B(i) \neq 0\}$, $B^2(i) = B(i)^2$.

取定$z = (z_n)_{n\in\mathbb{Z}} \in \{-1,0,1\}^{\mathbb{N}}$, 设$z^2 = (z_n^2)_{n\in\mathbb{Z}} \in \{0,1\}^{\mathbb{N}}$ 关于测度ν 沿着$\{N_k\}_{k\in\mathbb{N}}$ 为quasi-generic 的, 即

$$\delta_{N_k,z^2} \triangleq \frac{1}{N_k} \sum_{n=1}^{N_k} \delta_{S^n z^2} \to \nu \in \mathcal{M}(X_{z^2}, S),\ k \to \infty,$$

其中X_{z^2} 为由z^2 确定的系统. 定义

$$F : \{-1,0,1\}^{\mathbb{N}} \to \{-1,0,1\},\ w \mapsto w(1).$$

引理 8.5.12 设$1 \leq a_1 < a_2 < \ldots < a_r, r \geq 0$ 以及$k_s \in \{1,2\}, 0 \leq s \leq r$ 不全为2. 那么

$$\int_{\{-1,0,1\}^{\mathbb{N}}} F^{k_0} \cdot F^{k_1} \circ S^{a_1} \cdot \ldots \cdot F^{k_r} \circ S^{a_r} d\widehat{\nu} = 0,$$

以及

$$\int_{\{-1,0,1\}^{\mathbb{N}}} F^2 \cdot F^2 \circ S^{a_1} \cdot \ldots \cdot F^2 \circ S^{a_r} d\widehat{\nu} = \int_{\{0,1\}^{\mathbb{N}}} F \cdot F \circ S^{a_1} \cdot \ldots \cdot F \circ S^{a_r} d\widehat{\nu}.$$

证明. 直接计算有

$$\begin{aligned}
&\int_{\{-1,0,1\}^{\mathbb{N}}} F^{k_0} \cdot F^{k_1} \circ S^{a_1} \cdot \ldots \cdot F^{k_r} \circ S^{a_r} d\widehat{\nu} \\
=& \sum_{j_0,j_1,\ldots,j_r=\pm 1} j_0^{k_0} j_1^{k_1} \ldots j_r^{k_r} \widehat{\nu}\left(\left\{y \in \{-1,0,1\}^{\mathbb{N}} : y_1 y_{1+a_1} \cdots y_{1+a_r} = j_0 j_1 \cdots j_r\right\}\right) \\
=& \left(\sum_{j_0,j_1,\ldots,j_r=\pm 1} j_0^{k_0} j_1^{k_1} \ldots j_r^{k_r}\right) \frac{1}{2^{r+1}} \nu\left(\left\{u \in \{0,1\}^{\mathbb{N}} : u_1 = u_{1+a_1} = \cdots = u_{1+a_r} = 1\right\}\right).
\end{aligned}$$

由此容易完成整个证明. □

引理 8.5.13 设$z^2 = (z_n^2)_{n\in\mathbb{Z}} \in \{0,1\}^2$ 关于测度ν 沿着$\{N_k\}_{k\in\mathbb{N}}$ 为*quasi-generic* 的, 即

$$\delta_{N_k,z^2} = \frac{1}{N_k}\sum_{n=1}^{N_k} \delta_{S^n z^2} \to \nu \in \mathcal{M}(X_{z^2}, S),\ k \to \infty.$$

那么以下等价

1. $\delta_{N_k,z} \to \widehat{\nu},\ k \to \infty$;
2. $\displaystyle\frac{1}{N_k}\sum_{n=1}^{N_k} z^{k_0}(n) z^{k_1}(n+a_1) z^{k_2}(n+a_2) \ldots z^{k_r}(n+a_r) \to 0,\ k \to \infty$, 对于任何$1 \leq a_1 < a_2 < \ldots < a_r$, 以及$k_0, k_1, k_2, \ldots, k_r \in \{1,2\}$ 不全都是偶数成立.

证明. (1) ⇒ (2): 如果$\delta_{N_k,z} \to \widehat{\nu},\ k \to \infty$ 成立, 那么对于任何$1 \leq a_1 < a_2 < \ldots < a_r$, 以及$k_0, k_1, k_2, \ldots, k_r \in \{1,2\}$ 不全都是偶数, 我们有

$$\begin{aligned}
&\frac{1}{N_k}\sum_{n=1}^{N_k} z^{k_0}(n) z^{k_1}(n+a_1) z^{k_2}(n+a_2) \ldots z^{k_r}(n+a_r) \\
=& \frac{1}{N_k}\sum_{n=1}^{N_k} \left(F^{k_0} \cdot F^{k_1} \circ S^{a_1} \cdot \ldots \cdot F^{k_r} \circ S^{a_r}\right)(S^{n-1} z) \\
\xrightarrow[k\to\infty]{}& \int_{\{-1,0,1\}^{\mathbb{N}}} F^{k_0} \cdot F^{k_1} \circ S^{a_1} \cdot \ldots \cdot F^{k_r} \circ S^{a_r} d\widehat{\nu} \\
=& 0 \quad (\text{引理8.5.12}).
\end{aligned}$$

(2) $\Rightarrow$ (1): 不妨设

$$\delta_{N_k,z} \to \rho,\ k \to \infty.$$

于是对于任何$1 \le a_1 < a_2 < \ldots < a_r$, 以及$k_0, k_1, k_2, \ldots, k_r \in \{1,2\}$ 不全都是偶数, 我们有

$$\begin{aligned}
&\frac{1}{N_k}\sum_{n=1}^{N_k} z^{k_0}(n) z^{k_1}(n+a_1) z^{k_2}(n+a_2)\ldots z^{k_r}(n+a_r) \\
&= \frac{1}{N_k}\sum_{n=1}^{N_k} \Big(F^{k_0}\cdot F^{k_1}\circ S^{a_1}\cdot \ldots \cdot F^{k_r}\circ S^{a_r}\Big)(S^{n-1}z) \\
&\underset{k\to\infty}{\longrightarrow} \int_{\{-1,0,1\}^{\mathbb{N}}} F^{k_0}\cdot F^{k_1}\circ S^{a_1}\cdot \ldots \cdot F^{k_r}\circ S^{a_r} d\rho \\
&= 0 \quad (\text{引理8.5.12}).
\end{aligned}$$

因为$F^2(u) = F(u^2), \forall u \in \{-1,0,1\}^{\mathbb{N}}$, 于是由$\delta_{N_k,z^2} = \frac{1}{N_k}\sum_{n=1}^{N_k}\delta_{S^n z^2} \to \nu \in \mathcal{M}(X_{z^2}, S),\ k \to \infty$, 我们得到

$$\int_{\{-1,0,1\}^{\mathbb{N}}} F^2\cdot F^2\circ S^{a_1}\cdot \ldots \cdot F^2\circ S^{a_r} d\rho = \int_{\{0,1\}^{\mathbb{N}}} F\cdot F\circ S^{a_1}\cdot \ldots \cdot F\circ S^{a_r} d\nu.$$

于是综合引理8.5.12, 我们得到: 对于任何$G \in \mathcal{A} \triangleq \{F^{k_0}\cdot F^{k_1}\circ S^{a_1}\cdot \ldots \cdot F^{k_r}\circ S^{a_r} : 1 \le a_1 < a_2 < \ldots < a_r, r \ge 0, i_s \in \mathbb{N}\}$,

$$\int_{\{-1,0,1\}^{\mathbb{N}}} G d\widehat{\nu} = \int_{\{-1,0,1\}^{\mathbb{N}}} G d\rho.$$

因为$\mathcal{A} \subseteq C(\{-1,0,1\}^{\mathbb{N}})$ 中对乘积封闭且分离点, 根据Stone-Weierstrass 定理, 我们得到$\rho = \widehat{\nu}$. 即

$$\delta_{N_k,z} \to \widehat{\nu},\ k \to \infty.$$

□

记

$$\mathrm{Q}-\mathrm{gen}(x) = \{\nu : \exists \{N_k\}_k\ s.t.\ \delta_{N_k,x} \underset{k\to\infty}{\longrightarrow} \nu\}.$$

注记 8.5.14 *根据引理8.5.13, 我们得到以下等价:*

1. *$z \in \{-1,0,1\}^{\mathbb{N}}$ 满足Chowla 条件;*
2. *$\mathrm{Q}-\mathrm{gen}(x) = \{\widehat{\nu} : \nu \in \mathrm{Q}-\mathrm{gen}(z^2)\}$;*
3. *$\delta_{N_k,z^2} \underset{k\to\infty}{\longrightarrow} \nu$ 当且仅当$\delta_{N_k,z} \underset{k\to\infty}{\longrightarrow} \widehat{\nu}$.*

根据上面注记, 我们容易推出唯一满足Chowla 条件的$\{-1,1\}^{\mathbb{N}}$ 中序列u 为Bernoulli 测度$\mu_{(\frac{1}{2},\frac{1}{2})}$ 的通用点. 这是因为此时u^2 为在$(111\ldots)$ 处Dirac 测度的通用点, 根据引理8.5.13, u 为此Dirac测度上的相对独立测度, 即为Bernoulli 测度$\mu_{(\frac{1}{2},\frac{1}{2})}$, 的通用点.

回顾

$$\pi : \{-1,0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}},\ (\pi(w))_n = w_n^2,\ \forall w \in \{-1,0,1\}^{\mathbb{N}}, n \in \mathbb{N}.$$

引理 8.5.15 *设$\nu \in \mathcal{M}(\{0,1\}^{\mathbb{N}}, S)$, $\widehat{\nu} \in \mathcal{M}(\{-1,0,1\}^{\mathbb{N}}, S)$ 为ν 相对独立交. 那么系统$(\{-1,0,1\}^{\mathbb{N}}, \widehat{\nu}, S)$ 为乘积系统*

$$(\{0,1\}^{\mathbb{N}}, \nu, S) \times (\{-1,1\}^{\mathbb{N}}, \mu_{(\frac{1}{2},\frac{1}{2})}, S)$$

的因子.

证明. 设$\xi : \{0,1\}^{\mathbb{N}} \times \{-1,1\}^{\mathbb{N}} \to \{-1,0,1\}$ 为

$$\xi(w,u)(n) = w(n)u(n).$$

那么根据$\widehat{\nu}$ 定义容易验证

$$\xi_*(\nu \times \mu_{(\frac{1}{2},\frac{1}{2})}) = \widehat{\nu}.$$

□

引理 8.5.16 *设$\nu \in \mathcal{M}(\{0,1\}^{\mathbb{N}}, S)$, $\widehat{\nu} \in \mathcal{M}(\{-1,0,1\}^{\mathbb{N}}, S)$ 为ν 相对独立交. 扩充*

$$\pi : (\{-1,0,1\}^{\mathbb{N}}, \widehat{\nu}, S) \to (\{0,1\}^{\mathbb{N}}, \nu, S),\ (\pi(w))_n = w_n^2,\ \forall w \in \{-1,0,1\}^{\mathbb{N}}, n \in \mathbb{N}$$

要么为平凡的(即几乎处处为1-1), 要么为相对K 的.

证明. 因为扩充

$$p : (\{0,1\}^{\mathbb{N}}, \nu, S) \times (\{-1,1\}^{\mathbb{N}}, \mu_{(\frac{1}{2},\frac{1}{2})}, S) \to (\{0,1\}^{\mathbb{N}}, \nu, S),\ (w,u) \mapsto w$$

为相对K 的, 于是所有此扩充的因子也为相对于$(\{0,1\}^{\mathbb{N}}, \nu, S)$ 的K 扩充. 根据等式$w = (w \cdot u)^2$, $\forall w \in \{0,1\}^{\mathbb{N}}, u \in \{-1,1\}^{\mathbb{N}}$, 容易验证$\pi \circ \xi$ 为到第一个坐标的投射. 于是根据引理8.5.15, π 为p 的因子, 从而完成了引理证明. □

引理 8.5.17 *设$\nu, \pi, \widehat{\nu}$ 同上, 那么*

$$\mathbb{E}_{\widehat{\nu}}(F|\pi(w) = u) = 0,\ \nu - a.e.\ u \in \{0,1\}^{\mathbb{N}}.$$

证明. 设测度分解为$\widehat{\nu} = \int_{\{0,1\}^{\mathbb{N}}} \widehat{\nu}_u d\nu(u)$. 则

$$\mathbb{E}_{\widehat{\nu}}(F|\pi(w) = u) = \mathbb{E}_{\widehat{\nu}}(F|\{0,1\}^{\mathbb{N}})(u) = \int_{\pi^{-1}(u)} F d\widehat{\nu}_u.$$

注意$\widehat{\nu}_u$ 为$(\frac{1}{2},\frac{1}{2})$ 与u 支撑位置的乘积. 如果$u(1) = 0$, 那么结果已然; 如果$u(1) = 1$, 那么F 在$\pi^{-1}(u)$ 上以相同概率取值于± 1, 积分依然为0. 证毕! □

引理 8.5.18 设(X,T) 为零熵系统, $x\in X$, 设$\nu,\pi,\widehat{\nu}$ 同上. 设$z\in\{-1,0,1\}^{\mathbb{N}}$ 为沿着$\{N_k\}_k$ 相对于$\widehat{\nu}$ 的*quasi-generic*点. 设

$$\delta_{T\times S,N_k,(x,z)}\underset{k\to\infty}{\longrightarrow}\rho\in\mathcal{M}(X\times\{-1,0,1\}^{\mathbb{N}},T\times S).$$

则

1. ρ 为(X,κ,T) 与$(\{-1,0,1\}^{\mathbb{N}},\widehat{\nu},S)$ 的交, 其中$\kappa\in\mathrm{Q-gen}(x)$;
2. 作为$(X\times\{-1,0,1\}^{\mathbb{N}},\rho,T\times S)$ 的因子, $(X,\kappa,T)\vee(\{0,1\}^{\mathbb{N}},\nu,S)$[1] 和$(\{-1,0,1\}^{\mathbb{N}},\widehat{\nu},S)$ 为相对于$(\{0,1\}^{\mathbb{N}},\nu,S)$ 的独立交.

证明. 由$\delta_{T\times S,N_k,(x,z)}\underset{k\to\infty}{\longrightarrow}\rho\in\mathcal{M}(X\times\{-1,0,1\}^{\mathbb{N}},T\times S)$, 令

$$\kappa=\lim_{k\to\infty}\delta_{T,N_k,x},$$

则ρ 为(X,κ,T) 与$(\{-1,0,1\}^{\mathbb{N}},\widehat{\nu},S)$ 的交. 因为(X,T) 零熵, 所以扩充

$$(X\times\{-1,0,1\}^{\mathbb{N}},\rho,T\times S)\to(\{0,1\}^{\mathbb{N}},\nu,S)$$

为相对零熵扩充. 根据引理8.5.16, $\pi:(\{-1,0,1\}^{\mathbb{N}},\widehat{\nu},S)\to(\{0,1\}^{\mathbb{N}},\nu,S)$为相对K 的, 所以二者为相对于$(\{0,1\}^{\mathbb{N}},\nu,S)$ 的独立交. 证毕! □

定理 8.5.19 如果$z\in\{-1,0,1\}^{\mathbb{N}}$ 满足*Chowla* 条件, 那么它也满足*Sarnak* 条件.

证明. 设$z\in\{-1,0,1\}^{\mathbb{N}}$ 满足Chowla 条件. 设(X,T) 为零熵系统, $x\in X$. 设

$$\delta_{T\times S,N_k,(x,z)}\underset{k\to\infty}{\longrightarrow}\rho\in\mathcal{M}(X\times\{-1,0,1\}^{\mathbb{N}},T\times S).$$

根据注记8.5.14, ρ 在第二个空间的投射为$\widehat{\nu}$, 其中$\nu\in\mathrm{Q-gen}(z^2)$. 设$f\in C(X)$, 于是

$$\frac{1}{N_k}\sum_{n=1}^{N_k}f(T^nx)z(n)=\frac{1}{N_k}\sum_{n=1}^{N_k}f(T^nx)F(S^nz)\underset{k\to\infty}{\longrightarrow}\int_{X\times\{-1,0,1\}^{\mathbb{N}}}f\otimes Fd\rho.$$

根据引理8.5.17,

$$\mathbb{E}_{\rho}(F|\{0,1\}^{\mathbb{N}})=\mathbb{E}_{\widehat{\nu}}(F|\{0,1\}^{\mathbb{N}})=0.$$

结合引理8.5.18, 我们就有

$$\mathbb{E}_{\rho}(f\otimes F|\{0,1\}^{\mathbb{N}})=\mathbb{E}_{\rho}(f|\{0,1\}^{\mathbb{N}})\mathbb{E}_{\rho}(F|\{0,1\}^{\mathbb{N}})=0.$$

于是$\int_{X\times\{-1,0,1\}^{\mathbb{N}}}f\otimes Fd\rho=0$, 即

$$\frac{1}{N_k}\sum_{n=1}^{N_k}f(T^nx)z(n)\underset{k\to\infty}{\longrightarrow}0.$$

证毕! □

显然根据定理8.5.19, 我们得到定理8.5.9.

[1]$(X,\kappa,T)\vee(\{0,1\}^{\mathbb{N}},\nu,S)$ 表示包含(X,κ,T) 和$(\{0,1\}^{\mathbb{N}},\nu,S)$ 最小的$(X\times\{-1,0,1\}^{\mathbb{N}},\rho,T\times S)$ 的因子.

§8.6 注记

交理论是由Furstenberg在1967年引入到动力系统中的[62], 时至今日它已经成为遍历论中最核心的工具之一. 关于交理论最详细的著作是[175]和[76], 也可参见综述性论文[176, 189]. 请感兴趣的读者参见这两本著作以及其中的参考文献.

测度版本Sarnak不交性定理以及Chowla猜测蕴含Sarnak猜测由Sarnak给出[179], 此处证明参见[55].

第九章　多重遍历定理与多重回复定理

在本章中我们介绍多重遍历定理和多重回复定理. 首先我们从van der Waerden 定理讲起, 给出它的动力系统证明. 之后我们介绍Furstenberg 对应原则, 说明如何将Ramsey 型的组合问题与动力系统关联在一起. 本章的主要目的之一是给出Furstenberg 关于Szemerédi 定理的动力系统证明, 为此目的我们先介绍几种特殊情况再给出Furstenberg-Zimmer 结构定理来完成整个证明. 事实上, 我们对于这些特殊情况给出的不止是回复性质, 还证明了它们的遍历平均收敛性质, 例如Furstenberg弱混合多重遍历定理、Furstenberg-Sárközy定理、Roth 定理等. 虽然我们不能介绍Host-Kra定理的证明, 但是我们将证明陶哲轩关于有限个交换变换的模多重遍历平均定理. 最后我们介绍这方面最新的进展.

§9.1　Birkhoff 多重回复定理与van der Waerden定理

Birkhoff 回复定理告诉我们, 如果T为紧致度量空间到自身的连续映射, 则$\mathrm{Rec}(T) \neq \emptyset$. 于是自然的问题是：如果$X$为紧致度量空间, $T_1, \ldots, T_l$ 为X 上l 个可交换的连续自映射, 那么是否存在一个序列$n_i \to +\infty$ 以及点$x \in X$ 使得$T_j^{n_i} x \to x, \forall 1 \leq j \leq l$ 成立? 这个问题的答案是肯定的,Furstenberg 等人在上世纪七十年代证明了这个结论, 并且运用它给出了著名的van der Waerden 定理的一个动力系统证明. van der Waerden 定理断言自然数任何有限染色必然有单染色包含任意长的等差数列，这个定理被Khinchin誉为“数论中的三颗明珠”之一[132].

目前关于Birkhoff 多重回复定理的证明有很多, 最简短的证明可能是Ellis半群的证明[76]. 　但在本节中, 我们介绍Furstenberg 等的原始证明[64], 希望通过这个证明读者更能体会证明的思想. 下面证明个别细节的处理引用了[162] 中的证明.

定理 9.1.1 *设X为紧致度量空间, $T_1, \ldots, T_l$为X 到自身的可交换的连续映射. 那么存在序列$n_i \to +\infty$ 以及$x \in X$ 使得$T_j^{n_i} x \to x, \forall 1 \leq j \leq l$.*

为证明定理9.1.1, 我们需要如下准备. 首先我们引入如下定义.

定义 9.1.2 *设(X, T) 是一个动力系统, 其中T 是可逆的. (X, T) 称为***齐性的***是指存在X 上与T 交换的同胚群G 使得(X, G) 为极小的. 一个闭子集$A \subseteq X$ 在(X, T) 中为***齐性***的是指存在X 上与T 交换的同胚群G 使得$GA = A$ 且(A, G) 为极小的.*

引理 9.1.3 (Bowen) *设(X, T)为动力系统,T可逆且$A \subseteq X$为闭的齐性子集.假设对任意$\varepsilon > 0$,存在$x, y \in A$及$n \in \mathbb{N}$使得$d(T^n x, y) < \varepsilon$.那么对任意$\varepsilon > 0$,存在$z \in A$及$n \in \mathbb{N}$使得$d(T^n z, z) < \varepsilon$.(注意$A$不必为$T$不变的.)*

证明. 首先我们证明定理的假设可以转化为：对任意$\varepsilon > 0$和$y \in A$,存在$x \in A$ 及$n \in \mathbb{N}$使得$d(T^n x, y) < \varepsilon$.

由A的齐性,存在群G使得(A,G)极小.对$\varepsilon>0$,我们断言,存在$g_1,\ldots,g_n\in G$使得

$$\min_i\ d(g_ix,y)<\frac{\varepsilon}{2}\ \forall x,y\in A: \tag{9.1.1}$$

事实上,用直径小于$\frac{\varepsilon}{2}$的有限多个开集V_j覆盖A,对每个j, $\{g^{-1}V_j:g\in G\}$为A的开覆盖, 于是有有限子覆盖

$$\{g_{1,j}^{-1}V_j,g_{2,j}^{-1}V_j,\ldots,g_{n,j}^{-1}V_j\}.$$

从而对任意$x,y\in A$, 存在j使得$y\in V_j$且对此j存在i使得$x\in g_{i,j}^{-1}V_j$. 于是$d(g_{i,j}x,y)<\frac{\varepsilon}{2}$, 这就证明了(9.1.1).

取充分小的$\delta>0$使得一旦$d(x,x')<\delta$,就有$d(g_ix,g_ix')<\frac{\varepsilon}{2}$, $\forall i$ (注意g_i为满足(9.1.1)中取定的有限个元素).根据假设,存在$x_0,y_0\in A$及$n_0\in\mathbb{N}$使得$d(T^{n_0}x_0,y_0)<\delta$.

于是对任意i有

$$d(T^{n_0}g_ix_0,g_iy_0)=d(g_iT^{n_0}x_0,g_iy_0)<\frac{\varepsilon}{2},$$

(9.1.1)允许我们取i使得$d(g_iy_0,y)<\frac{\varepsilon}{2}$, 于是

$$\min_i d(T^{n_0}g_ix_0,y)<\varepsilon\ ,\forall\ y\in A.$$

这就证明了,对任意$y\in A$,存在$x\in A$及$n\in\mathbb{N}$使得$d(T^nx,y)<\varepsilon$.

任取定点$z_0\in A$,由上结论取$z_1\in A$及$n_1\in\mathbb{N}$使得

$$d(T^{n_1}z_1,z_0)<\frac{\varepsilon}{2}. \tag{9.1.2}$$

同样取$z_2\in A$, $n_2\in\mathbb{N}$ 及$\varepsilon_2<\frac{\varepsilon}{2}$ 使得$d(T^{n_2}z_2,z_1)<\varepsilon_2$, 其中$\varepsilon_2$ 充分小使得(9.1.2) 当z_1被$T^{n_2}z_2$ 替代时仍成立.即, $d(T^{n_1+n_2}z_2,z_0)<\frac{\varepsilon}{2}$.

继续上面的归纳.如果$z_0,z_1,\ldots,z_r\in A$, $n_1,n_2,\ldots,n_r\in\mathbb{N}$, 及$\varepsilon_2,\ldots,\varepsilon_r\in(0,\frac{\varepsilon}{2})$已经取定使得

$$d(T^{n_j}z_j,z_{j-1})<\varepsilon_j,\quad j=1,2,\ldots,r, \tag{9.1.3}$$

取$\varepsilon_{r+1}<\frac{\varepsilon}{2}$充分小使得(9.1.3)当$z_r$被它附近距离小于$\varepsilon_{r+1}$的点替代时仍成立.取$z_{r+1}\in A$及$n_{r+1}\in\mathbb{N}$使得$d(T^{n_{r+1}}z_{r+1},z_r)<\varepsilon_{r+1}$. 这样我们有,当$i<j$时

$$d(T^{n_j+n_{j-1}+\ldots+n_i}z_j,z_i)<\frac{\varepsilon}{2}.$$

由A的紧性, 存在i,j使得$i<j$且$d(z_i,z_j)<\frac{\varepsilon}{2}$. 取$n=n_j+n_{j-1}+\ldots+n_i$, 就有$d(T^nz_j,z_j)<\varepsilon$.证毕. □

引理 9.1.4 *假设同上, 则存在$x\in A$在T作用下回复.*

证明. 对任意$n=1,2,\ldots,$令

$$E_n=\{x\in A:\inf_k\ d(T^kx,x)\geq\frac{1}{n}\}.$$

如果A中没有T的回复点, 那么
$$A = \bigcup_{n=1}^{\infty} E_n.$$
下面我们说明每个闭集E_n的内部E_n^0 (相对于A)为空集,这样就与Baire 定理矛盾.

如果存在n使得$E_n^0 \neq \emptyset$,那么由于(A,G)极小就有$A = GE_n^0$.由紧性, 存在$g_1,\ldots,g_m \in G$使得
$$A = g_1^{-1}E_n^0 \cup \ldots \cup g_m^{-1}E_n^0.$$

取$\delta > 0$使得$d(x,x') < \delta$蕴含$d(g_ix, g_ix') < \frac{1}{n}$, $i = 1,2,\cdots,m$.我们断言如果$x \in g_j^{-1}E_n^0$, 那么$\inf_k d(T^kx,x) \geq \delta$.这是因为如果存在$k$使得$d(T^kx,x) < \delta$, 那么对任意$j \in \{1,2,\cdots,m\}$有$d(T^kg_jx, g_jx) < \frac{1}{n}$,或者对某个$y \in E_n^0$ 有$d(T^ky,y) < \frac{1}{n}$.

因为任意$x \in A$必在某个$g_j^{-1}E_n^0$中, 我们得到对任意$x \in A$有$\inf_k d(T^kx,x) \geq \delta$.这与引理9.1.3矛盾. □

定理9.1.1的证明: 设$T_1,\ldots,T_l$为紧度量空间X上的交换同胚, 我们要寻找点$x \in X$满足对任意$\varepsilon > 0$存在$n \in \mathbb{N}$使得$d(T_i^nx,x) < \varepsilon$, $\forall i = 1,\ldots,l$.

我们用归纳法.当$l = 1$时,即为Birkhoff回复定理.假设上结论对任意$l-1$个交换同胚已经成立.

取G为由$T_1,\ldots,T_l$生成的群,不妨设(X,G)为极小的（否则限制于某个极小子集上）.令$\Delta \subseteq X^l$为对角线及$T = T_1\times\ldots\times T_l$. G中元素g在X^l上的作用为: $g(x_1,\ldots,x_l) = (gx_1,\ldots,gx_l)$,即$g$对应于$g\times\ldots\times g$.易见这些映射与$T$交换, 且$(\Delta,G)$ (它与(X,G)同构)为极小的,进而Δ为系统(X^l,T) 的齐性集.

下面我们验证前面引理中的假设.亦即验证：对任意$\varepsilon > 0$,存在$x^*,y^* \in \Delta$ 及$n \in \mathbb{N}$使得$d(T^nx^*,y^*) < \varepsilon$.令$R_i = T_iT_l^{-1}$,$\forall i = 1,\ldots,l-1$. 由归纳假设,存在$x \in X$及$n_m \to +\infty$使得$R_i^{n_m}x \to x$, $\forall i = 1,\ldots,l-1$. 设$\varepsilon > 0$, 令
$$y^* = (x,x,\ldots,x) \;\; 和 \;\; x^* = (T_l^{-n_m}x, T_l^{-n_m}x,\ldots,T_l^{-n_m}x).$$
则
$$\begin{aligned} d(T^{n_m}x^*,y^*) &= d(T_1^{n_m}\times T_2^{n_m}\ldots\times T_l^{n_m}x^*,y^*) \\ &= d((T_1^{n_m}T_l^{-n_m}x,\ldots,T_{l-1}^{n_m}T_l^{-n_m}x,x),(x,x,\ldots x)) \\ &= d((R_1^{n_m}x,\ldots,R_{l-1}^{n_m}x,x),(x,x,\ldots x)) \end{aligned}$$
取m充分大使得上式小于ε.

这样我们就可以应用引理9.1.4得到：存在$(x,x,\ldots x) \in \Delta$为在$T = T_1\times\ldots\times T_l$下回复的,此即为所求.

下面我们运用标准的方法将上结论推广到一般情况：

定理 9.1.5 设X为紧致度量空间,$T_1,\ldots,T_l$为X上交换的连续自映射.那么存在$x\in X$ 及序列$n_i\to+\infty$使得

$$T_j^{n_i}x\to x,\forall 1\le j\le l.$$

证明. 令$\Omega=X^{\mathbb{Z}^l}$及

$$(S_i\omega)_{(n_1,\ldots,n_i,\ldots,n_l)}=(\omega)_{(n_1,\ldots,n_i+1,\ldots,n_l)},i=1,\ldots,l.$$

设$\widetilde{X}\subseteq\Omega$满足对每个$i=1,\ldots,l$和每个格点$(n_1,\ldots,n_l)\in\mathbb{Z}^l$有

$$(S_i\omega)_{(n_1,\ldots,n_i,\ldots,n_l)}=T_i\omega_{(n_1,\ldots,n_i,\ldots,n_l)}. \tag{9.1.4}$$

集合$\widetilde{X}$ 为非空的. 为证明此点, 任取$x\in X$, 对$n\in\mathbb{N}$ 设

$$(\omega^n)_{(n_1,\ldots,n_l)}=T_1^{n_1+n}T_2^{n_2+n}\ldots T_l^{n_l+n}x,n_i\ge -n.$$

对于ω^n 在其余格点的值随意定义. 这样得到点列ω^n, 对$n_i\ge -n$ 的$(n_1,\ldots,n_l)$ 满足(9.1.4). 取ω^n 在Ω 中的极限点, 它在$\widetilde{X}$ 中. 于是$\widetilde{X}$ 为非空的. 并且易证在S_i 与S_i^{-1} 作用下为不变的.

于是应用定理9.1.1 于$\widetilde{X}$ 及交换同胚$S_1,\ldots,S_l$, 我们可找到相对于$S_1,\ldots,S_l$ 的多重回复点$\widetilde{x}$. 根据(9.1.4), $\widetilde{x}$ 的每个分量为X 相对于$T_1,\ldots,T_l$ 的多重回复点. 证毕. □

现在我们运用定理9.1.5证明:

定理 9.1.6 (van der Waerden定理) *[193]* 如果$\mathbb{N}=B_1\cup\ldots\cup B_l$, 那么存在$j$使得$B_j$包含了任意长的等差数列.

证明. 不失一般性,设$B_i\cap B_j=\emptyset$, $i\ne j$.设$A=\{1,\ldots,l\}$, $\Sigma_l=A^{\mathbb{N}}$且$T:\Sigma_l\to\Sigma_l$为转移映射.定义点$w\in\Sigma_l$使得

$$w_n=i \text{ 当且仅当 } n\in B_i.$$

设$X=\overline{\mathrm{orb}(w,T)}$.给定$k\ge 1$, 令$T_i=T^i$, $i=1,\ldots,k$.运用定理9.1.5,我们得到点$x\in X$及$n\ge 1$使得

$$d(T_i^n x,x)<1,\forall i=1,\ldots,k.$$

尤其$x,T^nx,T^{2n}x,\ldots,T^{kn}x$在坐标1处一致,于是

$$x_1=x_{n+1}=x_{2n+1}=\ldots=x_{kn+1}.$$

因为$x\in X$,所以存在m使得

$$w_{m+1}=w_{m+n+1}=\ldots=w_{m+kn+1}.$$

于是存在j_k使得B_{j_k}包含了长为$k+1$的等差数列.

因为$B_1,\ldots,B_l$是$\mathbb{N}$的有限剖分,必定存在j使得B_j包含任意长的等差数列. □

习　题

1. 设$T, S : X \to X$ 为同胚, 举例说明如果不限制T, S 的条件, 有可能不存在点$x \in X$ 以及$\{n_i\}_i$ 使得$T^{n_i}x \to x, S^{n_i}x \to x, i \to \infty$.

§9.2 从Poincaré多重回复定理到Szemerédi定理

§9.2.1 Szemerédi定理

定理 9.2.1 (Szemerédi定理) *[186] 设S为$\mathbb{Z}_+$具有正上Banach密度的子集,则S包含了任意长的算术级数.*

易见前面提到的V. D. Waerden定理是此定理的一个直接推论.在第一章我们运用拓扑的多重回复定理给出了V. D. Waerden定理的证明,此处我们运用遍历论的方法证明Szemerédi定理.

根据Poincaré 回复定理我们定义Poincaré序列:

定义 9.2.2 *$\mathbb{Z}_+$的一个子集S称为***Poincaré 序列***是指对任意保测系统$(X, \mathcal{B}, \mu, T)$ 以及任意满足$\mu(A) > 0$的$A \in \mathcal{B}$, 存在$0 < n \in S$使得$\mu(A \cap T^{-n}A) > 0$.*

为刻画Poincaré序列,我们需要下面的定义.

设S为$\mathbb{Z}_+$的子集. S的**上半Banach 密度**定义为

$$BD^*(S) = \limsup_{|I| \to +\infty} \frac{|S \cap I|}{|I|},$$

其中I取遍$\mathbb{Z}_+$所有子区间.其**下半Banach 密度**类似定义之. 记$\mathcal{F}_{\rm pubd}$ 为$\mathbb{Z}_+$中全体具有正上半Banach 密度的序列的集合.

§9.2.2 Furstenberg对应原则

定理 9.2.3 (Furstenberg对应原则) *设$\mathcal{F}(\mathbb{Z}_+)$为$\mathbb{Z}_+$的全体有限子集的集合,则*

(1) 如果$E \subseteq \mathbb{Z}_+$满足$BD^*(E) > 0$, 那么存在保测系统$(X, \mathcal{X}, \mu, T)$以及$A \in \mathcal{X}$ 使得$\mu(A) = BD^*(E)$. 进一步, 对任意$\alpha \in \mathcal{F}(\mathbb{Z}_+)$ 我们有

$$BD^*(\bigcap_{n \in \alpha}(E - n)) \geq \mu(\bigcap_{n \in \alpha}(T^{-n}A)).$$

(2) 设$(X, \mathcal{X}, \mu, T)$ 为保测系统, $A \in \mathcal{X}$ 满足$\mu(A) > 0$. 那么存在$E \subseteq \mathbb{Z}_+$ 使得$\overline{d}(E) \geq \mu(A)$ 且

$$\{\alpha \in \mathcal{F}(\mathbb{Z}_+) : \bigcap_{n \in \alpha}(E - n)) \neq \emptyset\} \subseteq \{\alpha \in \mathcal{F}(\mathbb{Z}_+) : \mu(\bigcap_{n \in \alpha} T^{-n}A) > 0\}.$$

证明. (1) 设$X=\{0,1\}^{\mathbb{Z}_+}$,$T:X\to X$为转移映射：$Tx(n)=x(n+1)$. 取$\mathbb{Z}_+$的区间列I_n满足$\lim_{n\to\infty}|I_n|=\infty$且

$$\lim_{n\to\infty}\frac{|E\cap I_n|}{I_n}=BD^*(E).$$

令$\xi=1_E\in X$及$A=\{x\in X:x(0)=1\}$.则

$$\lim_{n\to\infty}\frac{1}{|I_n|}\sum_{i\in I_n}1_A(T^i\xi)=\lim_{n\to\infty}\frac{1}{|I_n|}\sum_{i\in I_n}1_E(i)=BD^*(E).$$

设$\mu_n=\frac{1}{|I_n|}\sum\limits_{i\in I_n}\delta_{T^i\xi}$. 不失一般性,我们设$\mu_n$弱收敛于$\mu$（否则取子列）.易证$\mu$为不变测度,并且

$$\mu(A)=\lim_{n\to\infty}\frac{1}{|I_n|}\sum_{i\in I_n}1_A(T^i\xi)=BD^*(E).$$

于是对任意$\alpha=\{n_1,n_2,\cdots,n_k\}\in\mathcal{F}(\mathbb{Z}_+)$,我们就有

$$\begin{aligned}
&\mu(T^{-n_1}A\cap T^{-n_2}A\cap\cdots\cap T^{-n_k}A)\\
=\;&\lim_{n\to\infty}\frac{1}{|I_n|}\sum_{i\in I_n}1_{T^{-n_1}A\cap T^{-n_2}A\cap\cdots\cap T^{-n_k}A}(T^i\xi)\\
=\;&\lim_{n\to\infty}\frac{1}{|I_n|}\sum_{i\in I_n}1_{(E-n_1)\cap(E-n_2)\cap\cdots(E-n_k)}(i)\\
\le\;&BD^*((E-n_1)\cap(E-n_2)\cap\cdots(E-n_k)).
\end{aligned}$$

(2) 对$\alpha\in\mathcal{F}(\mathbb{Z}_+)$,令$E_\alpha=\bigcap_{n\in\alpha}T^{-n}A$.取$N$为所有具有零测度的$E_\alpha$的并集.由于它为可数个零测集的并,$\mu(N)=0$. 令$B=A\setminus N$.根据定义,我们易有断言：对$\alpha\in\mathcal{F}$, $\mu(\bigcap_{n\in\alpha}T^{-n}B)=0$当且仅当$\bigcap_{n\in\alpha}T^{-n}B=\emptyset$.

令$f_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}1_{T^{-i}B}$,则对任意$x\in X$和$n$有$f_n\le 1$.于是由Fatou引理有：

$$\int_X\limsup_{n\to\infty}f_nd\mu\ge\limsup_{n\to\infty}\int_X f_nd\mu.$$

于是存在$x\in X$使得

$$\overline{d}(\{n:x\in T^{-n}B\})=\limsup_{n\to\infty}f_n(x)\ge\limsup_{n\to\infty}\int_X\frac{1}{n}\sum_{i=1}^{n}1_{T^{-i}B}d\mu=\mu(B)=\mu(A).$$

令$E=\{n:x\in T^{-n}B\}$.则对$\alpha\in\mathcal{F}(\mathbb{Z}_+)$,如果$k\in\bigcap\limits_{n\in\alpha}(E-n)$, 那么$T^kx\in\bigcap\limits_{n\in\alpha}T^{-n}B$. 由前面的断言, 我们有$\mu(\bigcap\limits_{n\in\alpha}(T^{-n}A))>0$. □

下面我们可以证明：

定理 9.2.4 $R\subseteq\mathbb{Z}_+$为*Poincaré*序列当且仅当对任意$E\in\mathcal{F}_{\text{pubd}}$, $R\cap(E-E)\ne\emptyset$.

证明. 设R为Poincaré 序列. 对$E \in \mathcal{F}_{\text{pubd}}$, 由定理9.2.3 存在保测系统$(X, \mathcal{X}, \mu, T)$ 以及$A \in \mathcal{X}$ 使得$\mu(A) = BD^*(E) > 0$ 及$BD^*(E \cap (E - n)) \geq \mu(A \cap T^{-n}A), \forall n \in \mathbb{N}$. 因为存在$n \in R$ 使得$\mu(A \cap T^{-n}A) > 0$, 我们有$E \cap (E - n) \neq \emptyset$, 即$n \in E - E$. 于是就有$R \cap (E - E) \neq \emptyset$.

反之假设对任意$E \in \mathcal{F}_{\text{pubd}}$, 我们有$R \cap (E - E) \neq \emptyset$. 设$(X, \mathcal{X}, \mu, T)$ 为保测系统, $A \in \mathcal{A}$ 满足$\mu(A) > 0$. 根据定理9.2.3, 存在$E \subseteq \mathbb{Z}_+$ 使得$\overline{d}(E) \geq \mu(A)$, 且$E \cap (E - n) \neq \emptyset$ 蕴含$\mu(A \cap T^{-n}A) > 0$. 由假设存在$n \in R \cap (E - E)$, 于是$\mu(A \cap T^{-n}A) > 0$. 所以R 为Poincaré 序列. □

定理9.2.4 的一个直接推论为:

推论 9.2.5 *任何Poincaré 序列为回复集. 尤其对任意$S \in \mathcal{F}_{\text{pubd}}$, $S - S$ 为syndetic 的.*

证明. 因为syndetic 集合有正上半Banach 密度, 所以任何Poincaré 为回复集. 设$S \in \mathcal{F}_{\text{pubd}}$. 由定理9.2.4, $S - S$ 与任意Poincaré 序列相交非空. 因为thick 集为Poincaré 序列, 所以$S - S$ 与所有thick 集相交非空. 从而$S - S$ 为syndetic. □

§9.2.3 Poincaré 多重回复定理

根据Furstenberg 对应原则, 我们得到Szemerédi 定理等价于下面的Poincaré 多重回复定理:

定理 9.2.6 (Poincaré 多重回复定理) *设$(X, \mathcal{X}, \mu, T)$ 为保测系统, $k \in \mathbb{N}$, $A \in \mathcal{X}$ 满足$\mu(A) > 0$. 那么存在$n \in \mathbb{N}$ 使得*

$$\mu(A \cap T^{-n}A \cap T^{-2n}A \cap \ldots \cap T^{-kn}A) > 0.$$

事实上, 为了证明Poincaré 多重回复定理, Furstenberg 证明了更强的一个结论:[1]

定理 9.2.7 *设$(X, \mathcal{X}, \mu, T)$ 为保测系统, $k \in \mathbb{N}$, $A \in \mathcal{X}$ 满足$\mu(A) > 0$. 那么*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap T^{-2n}A \cap \ldots \cap T^{-kn}A) > 0. \tag{9.2.1}$$

我们在后文中将会证明上面结论.

习 题

1. 证明：如果对任意$\alpha \in (0, 2\pi)$ 有$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} e^{i\alpha s_k} = 0$ 成立, 那么$\{s_k\}$ 为Poincaré 序列. [提示: 参见[199].]
2. 证明：IP 集为Poincaré 序列.

[1]在文献[63] 中, Furstenberg 实际上证明的结论是: 设$(X, \mathcal{X}, \mu, T)$ 为保测系统, $f \in L^\infty(X, \mathcal{X}, \mu)$, $f \geq 0$, 且$f \neq 0$, 那么对于任何$k \in \mathbb{N}$,

$$\liminf_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N} \int_X f(x) f(T^n x) \cdots f(T^{(k-1)n} x) d\mu(x) > 0.$$

§9.3 van der Corput引理与Furstenberg弱混合多重遍历定理

van der Corput引理在等分布理论等中有着重要应用，将它运用来研究多重遍历性质最先开始于Bergelson的工作[14]. 在本节中我们先给出van der Corput引理，然后用它来给出Furstenberg弱混合多重遍历定理的证明.

§9.3.1 van der Corput引理

定理 9.3.1 (van der Corput 引理) 设$\{u_n\}_{n\in\mathbb{Z}_+}$ 为*Hilbert* 空间$\mathcal{H}$ 中的有界序列. 令

$$s_h=\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h},u_n\rangle\right|.$$

如果

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}s_h=0,$$

那么

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n\right\|=0.$$

证明. 取定$\varepsilon>0$, 存在$H_0\in\mathbb{N}$ 使得当$H>H_0$,

$$\frac{1}{H}\sum_{h=0}^{H-1}s_h<\varepsilon. \tag{9.3.1}$$

取比H 充分大的N 使得$\frac{1}{N}\sum_{n=0}^{N-1}u_n$ 与$\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}$ 仅在少数几项不相同, 即

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n-\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}\right\|\le\varepsilon. \tag{9.3.2}$$

由此我们转向估计第二个平均. 由Cauchy不等式,

$$\begin{aligned}\limsup_{N\to\infty}\left\|\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}\right\|^2&=\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}\Big(\frac{1}{H}\sum_{h=0}^{H-1}u_{n+h}\Big)\right\|^2\\&\le\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left\|\frac{1}{H}\sum_{h=0}^{H-1}u_{n+h}\right\|^2\\&=\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{H^2}\sum_{h,h'=0}^{H-1}\langle u_{n+h},u_{n+h'}\rangle\\&\le\limsup_{N\to\infty}\frac{1}{H^2}\sum_{h,h'=0}^{H-1}\left|\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h},u_{n+h'}\rangle\right|\end{aligned} \tag{9.3.3}$$

对于取定的h, h', 根据条件, 我们有

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h}, u_{n+h'}\rangle\right| = s_{|h-h'|}.$$

于是(9.3.3) 变为,

$$\limsup_{N\to\infty}\left\|\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}\right\|^2 \le \frac{1}{H^2}\sum_{h,h'=0}^{H-1}s_{|h-h'|}.$$

接下去再用(9.3.1) 去估计右边式子. 将$[0, H-1]^2$ 分解为三部分:

$$\begin{aligned} &(I) \quad 0\le h\le H-H_0-2, h\le h'\le H-1\\ &(II) \quad 1\le h'\le H-H_0-2, h'\le h\le H-1\\ &(III) \quad [H-H_0-1, H-1]^2. \end{aligned}$$

于是按照此三部分得到

$$\begin{aligned} &\frac{1}{H^2}\sum_{(h,h')\in[0,H-1]^2}s_{|h-h'|}\\ =&\frac{1}{H}\sum_{h=0}^{H-H_0-2}\frac{1}{H}\sum_{h'=h}^{H-1}s_{h'-h}+\frac{1}{H}\sum_{h'=0}^{H-H_0-2}\frac{1}{H}\sum_{h=h'}^{H-1}s_{h-h'}+\frac{1}{H^2}\sum_{(h,h')\in[H-H_0-1,H-1]^2}s_{|h-h'|}. \end{aligned}$$

因为$\{u_n\}_{n\in\mathbb{Z}_+}$ 有界, 所以$\{s_n\}_{n\in\mathbb{Z}_+}$ 有界, 取M 使得$|s_n|\le M, \forall n\in\mathbb{Z}_+$. 对于上面第一项, 根据(9.3.1), 对于任何$0\le h\le H-H_0-2$,

$$\frac{1}{H}\sum_{h'=h}^{H-1}s_{h'-h}<\varepsilon,$$

所以

$$\frac{1}{H}\sum_{h=0}^{H-H_0-2}\frac{1}{H}\sum_{h'=h}^{H-1}s_{h'-h}<\varepsilon.$$

类似的, 第二项也有

$$\frac{1}{H}\sum_{h'=0}^{H-H_0-2}\frac{1}{H}\sum_{h=h'}^{H-1}s_{h-h'}<\varepsilon.$$

对于第三项,

$$\frac{1}{H^2}\sum_{(h,h')\in[H-H_0-1,H-1]^2}s_{|h-h'|}\le\frac{H_0^2}{H^2}M,$$

取$H>\sqrt{\frac{M}{\varepsilon}}H_0$ 使之小于ε. 综上, 当$H>\sqrt{\frac{M}{\varepsilon}}H_0$ 时,

$$\limsup_{N\to\infty}\left\|\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}\right\|^2 \le \frac{1}{H^2}\sum_{h,h'=0}^{H-1}s_{|h-h'|}<3\varepsilon.$$

结合(9.3.2), 我们有

$$\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n\right\|\le 4\varepsilon.$$

由ε 任意就得到:

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n\right\|=0.$$

证毕! □

§9.3.2 Furstenberg弱混合多重遍历定理

定理 9.3.2 (Furstenberg) 设$(X,\mathcal{X},\mu,T)$ 为弱混合系统, $k\in\mathbb{N}$. 那么对于任何$f_1,f_2,\dots,f_k\in L^\infty(X,\mu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2\dots T^{kn}f_k\xrightarrow{L^2}\int_X f_1d\mu\int_X f_2d\mu\dots\int_X f_kd\mu,\ N\to\infty. \tag{9.3.4}$$

证明. 我们归纳证明. 当$k=1$ 是直接由遍历定理即可. 我们假设命题对于k 成立. 下面证明$k+1$ 时情况. 设$f_1,f_2\dots,f_{k+1}\in L^\infty(X,\mu)$. 不妨设$\int_X f_1d\mu=0$, 否则用$f_1-\int_X f_1d\mu$ 替代它. 令

$$u_n=T^nf_1T^{2n}f_2\dots T^{(k+1)n}f_{k+1},\ \forall n\in\mathbb{Z}_+.$$

我们需要证明

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n\right\|_2=0.$$

为此我们运用定理9.3.1.

$$\begin{aligned}s_h&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h},u_n\rangle\\&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X T^{n+h}f_1T^{2n+2h}f_2\dots T^{(k+1)n+(k+1)h}f_{k+1}T^nf_1T^{2n}f_2\dots T^{(k+1)n}f_{k+1}d\mu\\&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X(f_1T^hf_1)T^n(f_2T^{2h}f_2)T^{2n}(f_3T^{3h}f_3)\dots T^{kn}(f_{k+1}T^{(k+1)h}f_{k+1})d\mu\end{aligned}$$

由归纳假设, 我们得到

$$\begin{aligned}s_h&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X(f_1T^hf_1)T^n(f_2T^{2h}f_2)T^{2n}(f_3T^{3h}f_3)\dots T^{kn}(f_{k+1}T^{(k+1)h}f_{k+1})d\mu\\&=\int_X f_1T^hf_1d\mu\int_X f_2T^{2h}f_2d\mu\dots\int_X f_{k+1}T^{(k+1)h}f_{k+1}d\mu\end{aligned}$$

于是

$$\begin{aligned}\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}s_h &= \frac{1}{H}\sum_{h=0}^{H-1}\int_X f_1T^hf_1d\mu\int_X f_2T^{2h}f_2d\mu\dots\int_X f_{k+1}T^{(k+1)h}f_{k+1}d\mu\\ &=\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}\int_{X^{k+1}}\Big(\bigotimes_{i=1}^{k+1}f_i\Big)\Big((T\times T^2\times\dots\times T^{k+1})^h\bigotimes_{i=1}^{k+1}f_i\Big)d\mu^{k+1}\\ &=\int_{X^{k+1}}\Big(\bigotimes_{i=1}^{k+1}f_i\Big)\Big(\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}(T\times T^2\times\dots\times T^{k+1})^h\bigotimes_{i=1}^{k+1}f_i\Big)d\mu^{k+1}\\ &=\left(\int_{X^{k+1}}\bigotimes_{i=1}^{k+1}f_id\mu^{k+1}\right)^2=\prod_{i=1}^{k+1}\Big(\int_X f_id\mu\Big)^2=0.\end{aligned}$$

上面运用了$(X^{k+1},\mathcal{X}^{k+1},\mu^{k+1},T\times T^2\times\dots\times T^{k+1})$ 为遍历的.

于是根据定理9.3.1, 我们得到

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n\right\|_2=0.$$

由此根据归纳就完成了证明. □

§9.4 Furstenberg-Sárközy定理

定理 9.4.1 (Furstenberg-Sárközy 定理数论版本) 设$E\subseteq\mathbb{N}$ 具有正上*Banach*密度, $p(n)$ 为常系数等于*0*的整系数多项式(即$p(0)=0$). 那么一定存在$x,y\in E$ 以及$n\in\mathbb{N}$ 使得

$$x-y=p(n).$$

定理 9.4.2 (Furstenberg-Sárközy 定理动力系统版本) 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $p(n)$ 为常系数等于*0*的整系数多项式(即$p(0)=0$). 那么对于任何$A\in\mathcal{X}$, $\mu(A)>0$, 存在$n\in\mathbb{N}$ 使得

$$\mu(A\cap T^{-p(n)}A)>0.$$

证明. 设$\mathcal{H}=L^2(X,\mu)$. 对于$m\in\mathbb{N}$, 令

$$\mathcal{H}_m=\{f\in\mathcal{H}:T^mf=f\},$$

$$\mathcal{V}_m=\{f\in\mathcal{H}:\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}T^{mn}f\right\|_2=0\}.$$

于是我们有分解

$$\mathcal{H}=\mathcal{H}_m\bigoplus\mathcal{V}_m.$$

令

$$\mathcal{H}_{rat} = \overline{\bigcup_{m=1}^{\infty} \mathcal{H}_m} = \overline{\{f \in \mathcal{H} : \text{存在 } m \in \mathbb{N} \text{ 使得 } T^m f = f\}}.$$

以及

$$\mathcal{V} = \bigcap_{m=1}^{\infty} \mathcal{V}_m = \{f \in \mathcal{H} : \lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} T^{mn} f \right\|_2 = 0, \forall m \in \mathbb{N}\}.$$

容易验证

$$\mathcal{H} = \mathcal{H}_{rat} \bigoplus \mathcal{V}. \tag{9.4.1}$$

于是根据分解我们得到

$$\mathbf{1}_A = f + g, \quad f \in \mathcal{H}_{rat}, g \in \mathcal{V}.$$

对$m \in \mathbb{N}$, 设$f_m = \mathbb{E}(\mathbf{1}_A|\mathcal{I}(T^m))$. 因为$\mathbf{1}_A \ge 0$, 所以$f_m \ge 0$ 并且

$$\int_X f_m d\mu = \int_X \mathbf{1}_A d\mu = \mu(A) > 0.$$

因为$f_{m!} \to f, m \to \infty$, 根据鞅定理, 我们得到$f \ge 0$, 并且$\int_X f d\mu = \int_X f_m d\mu = \mu(A) > 0$. 因为分解(9.4.1)为$T$ 不变的, 所以

$$\mu(A \cap T^{p(n)}A) = \int_X (f+g) T^{p(n)}(f+g) d\mu = \int_X f T^{p(n)} f d\mu + \int_X g T^{p(n)} g d\mu. \tag{9.4.2}$$

下面我们分别处理两项.

(1) 我们证明:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f T^{p(n)} f d\mu > 0. \tag{9.4.3}$$

如果$f \in \mathcal{H}_m$, 那么$\{T^{p(n)}f\}$ 为周期序列[[这是因为$\{p(n) \pmod m\}$ 为周期的.]] 于是(9.4.3) 极限显然成立. 所以对于$\mathcal{H}_{rat}$ 的稠密集合$\bigcup_{m=1}^{\infty} \mathcal{H}_m$ 极限(9.4.3) 成立. 通过简单的逼近分析, 我们知道对于$f \in \mathcal{H}_{rat}$, (9.4.3) 中极限是存在的, 需要验证的是极限大于0.

对于$\varepsilon = \frac{1}{4}\mu(A)^2$, 取$m \in \mathbb{N}$ 使得$\|f - f_m\|_2 < \varepsilon$. 由$m|p(mn)$, 所以$T^{p(mn)} f_m = f_m$. 于是由Cauchy不等式,

$$\int_X f_m T^{p(mn)} f_m d\mu = \int_X f_m^2 d\mu \ge \left(\int_X f_m \right)^2 = \mu(A)^2.$$

由此

$$\begin{aligned}
\int_X f T^{p(mn)} f d\mu &= \langle f, T^{p(mn)} f \rangle \\
&= \langle f_m, T^{p(mn)} f_m \rangle + \langle f_m, T^{p(mn)} f - T^{p(mn)} f_m \rangle + \langle f - f_m, T^{p(mn)} f \rangle \\
&\ge \mu(A)^2 - 2\varepsilon = \frac{1}{2}\mu(A)^2 > 0
\end{aligned}$$

前面已证$f \geq 0$, 所以$T^{p(n)}f \geq 0$. 于是

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\int_X fT^{p(n)}fd\mu \geq \frac{1}{m}\int_X fT^{p(mn)}fd\mu \geq \frac{1}{2m}\mu(A)^2 > 0.$$

(2) 我们证明：

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\int_X gT^{p(n)}gd\mu = 0. \tag{9.4.4}$$

设g 的谱测度为μ_g 那么(9.4.4) 等价于

$$\lim_{N\to\infty}\int_{\mathbb{S}^1}\frac{1}{N}\sum_{n=0}^{N-1} z^{p(n)}d\mu_g(z) = 0. \tag{9.4.5}$$

如果$z = e^{2\pi i\theta} \in \mathbb{S}^1$ 不为单位根, 那么根据Weyl 等分布定理,

$$\frac{1}{N}\sum_{n=0}^{N-1} z^{p(n)} = \frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi\theta p(n)} \to \int_0^1 e^{2\pi ix}dx = 0,\ \mathbb{N}\to\infty.$$

下证对于单位根z, $\mu_g(\{z\}) = 0$. 否则设存在$z_0 = e^{2\pi i\frac{a}{b}}$, $a \in \mathbb{Z}, b \in \mathbb{N}$, 使得$\mu_g(\{z_0\}) > 0$. 于是$L^2(\mathbb{S}^1, \mu_g)$ 包含非零函数

$$f(z) = \begin{cases} 1, & \text{如}z = z_0; \\ 0, & \text{否则}. \end{cases}$$

注意

$$z^b f(z) = z_0^b f(z_0) = f(z),\ \forall z \in \mathbb{S}^1.$$

根据命题 g 的生成空间$Z(g)$ 酉等价于$L^2(\mathbb{S}^1, \mu_g)$:

$$\begin{tikzcd} Z(g) \arrow[r, "U_T"] \arrow[d, "W_x"'] & Z(g) \arrow[d, "W_x"] \\ L^2(\mathbb{S}^1, \mu_g) \arrow[r, "V"'] & L^2(\mathbb{S}^1, \mu_g). \end{tikzcd}$$

其中

$$V : L^2(\mathbb{S}^1, \sigma_x) \to L^2(\mathbb{S}^1, \sigma_x),\ q(z) \mapsto zq(z).$$

设$h \in L^2(X, \mu)$ 为$f(z)$ 对应的函数, 于是

$$\langle g, h\rangle_{L^2(X,\mu)} = \langle 1, f(z)\rangle_{L^2(\mathbb{S}^1,\mu_g)} = \mu_g(\{z_0\}) > 0.$$

但是$h \in \mathcal{H}_b \subseteq \mathcal{H}_{rat}$, $g \perp h$, 与上面矛盾! □

注记 9.4.3 根据上面的证明, 我们可以证明: 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $p(n)$ 为整系数多项式（不需要$p(0)=0$ 这个条件）, 那么对于任何$f\in L^\infty(X,\mu)$,

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}T^{p(n)}f-\frac{1}{N}\sum_{n=0}^{N-1}T^{p(n)}\mathbb{E}(f|\mathcal{H}_{rat})\right\|_2\longrightarrow 0,\ N\to\infty.$$

习　题

1. 运用van der Corput 引理给出Furstenberg-Sárközy 定理的另一个证明.

§9.5　$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2$ 平均收敛定理以及Roth定理

在这一节中我们介绍遍历平均$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2$ 收敛性, 并且给出Roth 定理的动力系统证明.

§9.5.1　Roth定理的陈述

定理 9.5.1 (Roth定理数论版本) 设$E\subseteq\mathbb{N}$ 具有正上*Banach*密度, 那么一定存在$a,b\in\mathbb{N}$ 使得

$$a,a+b,a+2n\in E.$$

定理 9.5.2 (Roth定理的动力系统版本) 设$(X,\mathcal{X},\mu,T)$ 为保测系统. 如果$A\in\mathcal{X}$, $\mu(A)>0$, 那么

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A\cap T^{-2n}A)>0. \tag{9.5.1}$$

§9.5.2　Furstenberg-Weiss的$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2$平均收敛定理

引理 9.5.3 设$\{a_i\},\{b_i\}\subseteq\mathbb{C}$. 那么

$$\prod_{i=1}^k a_i-\prod_{i=1}^k b_i=(a_1-b_1)b_2\dots b_k+a_1(a_2-b_2)b_3\dots b_k+a_1\dots a_{k-1}(a_k-b_k). \tag{9.5.2}$$

引理 9.5.4 设$(X,\mathcal{X},\mu,T)$ 为保测系统. $K:L^2(X,\mu)\to L^2(X,\mu)$ 为紧自伴算子, 并且$KU_T=U_TK$. 那么K 任何非零特征值的特征空间是有限维的并且为U_T 不变的, 并且由U_T 的特征函数张成.

定理 9.5.5 (Furstenberg-Weiss) 设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 那么对于任何$f_1,f_2\in L^\infty(X,\mu)$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2 \tag{9.5.3}$$

在L^2 中存在.

具体讲, 设$(Z,\mathcal{K},m,R_a)$ 为$(X,\mathcal{X},\mu,T)$ *Kronecker* 因子, 其中$R_a:Z\to Z, z\mapsto a+z$. 设$\pi:X\to Z$ 为因子映射, $\widetilde{f_1}=\mathbb{E}(f_1|\mathcal{K}),\widetilde{f_2}=\mathbb{E}(f_2|\mathcal{K})$. 那么我们有

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2-\frac{1}{N}\sum_{n=0}^{N-1}T^n\mathbb{E}(f_1|\mathcal{K})T^{2n}\mathbb{E}(f_2|\mathcal{K})\right\|_2\to 0,N\to\infty.$$

上面极限等于

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}R_a^n\widetilde{f_1}(\pi(x))R_a^{2n}\widetilde{f_2}(\pi(x))\overset{L^2}{=}\int_Z\widetilde{f_1}(\pi(x)+\theta)\widetilde{f_2}(\pi(x)+2\theta)dm(\theta).\tag{9.5.4}$$

证明. 我们分三步走:

(1) 首先我们证明

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2-\frac{1}{N}\sum_{n=0}^{N-1}T^n\mathbb{E}(f_1|\mathcal{K})T^{2n}\mathbb{E}(f_2|\mathcal{K})\right\|_2\to 0,N\to\infty.$$

(2) 对于Kronecker 系统$(Z,\mathcal{K},m,R_a)$ 证明$\lim\limits_{N\to\infty}\frac{1}{N}\sum\limits_{n=0}^{N-1}T^nf_1T^{2n}f_2$ 存在.

(3) 由(1)(2)得到(9.5.3).

下面我们开始每个步骤的证明.

(1) 根据引理9.5.3, 我们需要证明, 当$\mathbb{E}(f_1|\mathcal{K})=0$ 或者$\mathbb{E}(f_2|\mathcal{K})=0$ 时有

$$\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2\xrightarrow{L^2}0,\ N\to\infty.\tag{9.5.5}$$

为了证明上式, 我们运用van der Corput 引理. 令

$$u_n=T^nf_1T^{2n}f_2.$$

于是

$$\begin{aligned}\langle u_{n+h},u_n\rangle&=\int_X T^{n+h}f_1T^{2n+2h}f_2T^nf_1T^{2n}f_2d\mu &(9.5.6)\\&=\int_X(f_1T^hf_1)T^n(f_2T^{2h}f_2)d\mu &(9.5.7)\\&=\int_X T^{-n}(f_1T^hf_1)(f_2T^{2h}f_2)d\mu &(9.5.8)\end{aligned}$$

当$\mathbb{E}(f_1|\mathcal{K})=0$ 时我们运用(9.5.8); 当$\mathbb{E}(f_2|\mathcal{K})=0$ 时我们运用(9.5.7).

先设$\mathbb{E}(f_1|\mathcal{K})=0$，运用(9.5.8) 以及遍历定理我们得到

$$\begin{aligned}
s_h &= \limsup_{N\to\infty}\left|\int_X\left(\frac{1}{N}\sum_{n=0}^{N-1}T^{-n}(f_1T^hf_1)\right)f_2T^{2h}f_2d\mu\right| \\
&\le \int_X\left|\left(\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{-n}(f_1T^hf_1)\right)f_2T^{2h}f_2\right|d\mu \\
&= \int_X\left|\int_X f_1T^hf_1d\mu f_2T^{2h}f_2\right|d\mu \\
&\le \|f_2\|_\infty^2\left|\int_X f_1T^hf_1d\mu\right|
\end{aligned}$$

我们希望证明

$$\frac{1}{H}\sum_{h=0}^{H-1}s_h\le\|f_2\|_\infty^2\frac{1}{H}\sum_{h=0}^{H-1}\left|\int_X f_1T^hf_1d\mu\right|\to 0, H\to\infty. \tag{9.5.9}$$

为此我们需要证明

$$\begin{aligned}
\frac{1}{H}\sum_{h=0}^{H-1}\left|\int_X f_1T^hf_1d\mu\right|^2 &= \frac{1}{H}\sum_{h=0}^{H-1}\int_{X^2}f_1\otimes\bar{f}_1(T\times T)^hf_1\otimes\bar{f}_1d\mu d\mu \\
&= \int_{X^2}f_1\otimes\bar{f}_1\left(\frac{1}{H}\sum_{h=0}^{H-1}(T\times T)^hf_1\otimes\bar{f}_1\right)d\mu d\mu \\
&\to \int_{X^2}f_1\otimes\bar{f}_1\mathbb{E}\big(f_1\otimes\bar{f}_1|\mathcal{I}(T\times T)\big)d\mu d\mu, N\to\infty.
\end{aligned}\tag{9.5.10}$$

令

$$F_H=\frac{1}{H}\sum_{h=0}^{H-1}(T\times T)^hf_1\otimes\bar{f}_1$$
$$F=\mathbb{E}\big(f_1\otimes\bar{f}_1|\mathcal{I}(T\times T)\big).$$

以及设

$$K_H: L^2(X,\mu)\to L^2(X,\mu),\quad g\mapsto\int_X F_H(x,y)g(y)d\mu(y),$$
$$K: L^2(X,\mu)\to L^2(X,\mu),\quad g\mapsto\int_X F(x,y)g(y)d\mu(y).$$

那么K_H,K 为$L^2(X,\mu)$ 上紧自伴算子，在$L^2(X^2,\mu\times\mu)$ 中$F_H\longrightarrow F, H\to\infty$, 因而在算子模下$K_H\to K, H\to\infty$. 注意$F$ 满足

$$(T\times T)F=F,\ F(y,x)=\overline{F(x,y)},$$

所以K 满足引理9.5.4 条件. 所以K 任何非零特征值的特征空间是有限维的并且为U_T 不变的, 并且由U_T 的特征函数张成.

因为$\mathbb{E}(f_1|\mathcal{K})=0$, 所以$\mathbb{E}(T^hf_1|\mathcal{K})=0,\forall h\in\mathbb{Z}_+$. 于是对于$\xi\in L^2(X,\mathcal{K},\mu)$,

$$K_H\xi=\int_X F_H(x,y)\xi(y)d\mu(y)=\int_X \frac{1}{H}\sum_{h=0}^{H-1}T^h f_1(x)T^h\bar{f}_1(y)\xi(y)d\mu(y)=0.$$

所以令$H\to\infty$ 得到

$$K\xi=0.$$

结合引理9.5.4, 我们由此得到$K=0$, 所以$F=0$.

于是(9.5.10) 成为

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}\left|\int_X f_1T^hf_1d\mu\right|^2=0,$$

从而

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}s_h=0.$$

由van der Corput 引理, (9.5.5) 成立.

下设$\mathbb{E}(f_2|\mathcal{K})=0$, 运用(9.5.7) . 此时(9.5.9) 换为:

$$\frac{1}{H}\sum_{h=0}^{H-1}s_h\le\|f_1\|_\infty^2\frac{1}{H}\sum_{h=0}^{H-1}\left|\int_X f_2T^{2h}f_2d\mu\right|\to 0, H\to\infty. \tag{9.5.11}$$

而(9.5.10) 换为

$$\frac{1}{H}\sum_{h=0}^{H-1}\left|\int_X f_2T^{2h}f_2d\mu\right|^2\to\int_{X^2} f_2\otimes\bar{f}_2\mathbb{E}\big(f_2\otimes\bar{f}_2|\mathcal{I}(T^2\times T^2)\big)d\mu d\mu, N\to\infty. \tag{9.5.12}$$

此时注意到$\mathcal{K}(T^2)=\mathcal{K}(T)$ 就可类似完成证明. $\mathcal{K}(T)\subseteq\mathcal{K}(T^2)$ 是显然的. 设$f\in L^2(X,\mu)$ 为λ 对应的T^2 特征函数. 令$\eta=\sqrt{\lambda}$. 于是

$$f=\frac{f+\overline{\eta}Tf}{2}+\frac{f-\overline{\eta}Tf}{2},$$

以及

$$T\frac{f\pm\overline{\eta}Tf}{2}=\frac{Tf\pm\overline{\eta}T^2f}{2}=\frac{Tf\pm\overline{\eta}\lambda f}{2}=\pm\eta\frac{f\pm\overline{\eta}Tf}{2}$$

所以$\mathcal{K}(T^2)\subseteq\mathcal{K}(T)$.

综上我们完成了(1)的证明.

(2) 设Z 是一个紧致可度量交换群和$a\in Z$, 群运算记为加号. 设遍历群旋转

$$R_a\colon Z\to Z, x\mapsto a+x,$$

其中m 为Haar 测度. 研究乘积系统

$$R_a\times R_a^2: Z^2\to Z^2,\ \ (x,y)\mapsto(x+a,y+2a)$$

设$x\in Z$, 令

$$Z_x=\overline{\mathcal{O}((x,x),R_a\times R_a^2)}.$$

因为$(Z^2,R_a\times R_a^2)$ 为群旋转, 易证$(Z_x,R_a\times R_a^2)$ 为唯一遍历的, 其测度为ψ_*m, 其中

$$\psi:Z\to Z^2,\quad z\mapsto(x+z,x+2z)$$

设$f_1,f_2\in C(Z)$ 为连续函数. 我们就有对于$x\in Z$

$$\begin{aligned}&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f_1(R_a^nx)f_2(R_a^{2n}x)\\ =&\int_{Z_x}f_1(y_1)f_2(y_2)d\psi_*m(y_1,y_2)\\ =&\int_Z f_1(x+z)f_2(x+2z)dm(z)\end{aligned}$$

接着处理$f_1,f_2\in L^\infty(Z,m)$, 通过标准的处理, 可以证明

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f_1(R_a^nx)f_2(R_a^{2n}x)=\int_Z f_1(x+z)f_2(x+2z)dm(z), m\ a.e.\ x\in Z.$$

这就完成了(2) 的证明.

也可通过Fourier展开证明(2). 设$\gamma,\eta\in\widehat{Z}$. So

$$\begin{aligned}\frac{1}{N}\sum_{n<N}\gamma(z+nt)\eta(z+2nt)&=\frac{1}{N}\sum_{n<N}\gamma(z)\gamma(t)^n\eta(z)\eta(t)^{2n}\\ &\to\begin{cases}0 & ;\text{ if }\gamma(t)\eta(t)^2\neq 1\\ 1 & ;\text{ if }\gamma(t)\eta(t)^2=1\end{cases}\end{aligned}$$

接着对于f_1,f_2 运用Fourier展开, 计算可得到

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f_1(R_a^nx)f_2(R_a^{2n}x)=\int_Z f_1(x+z)f_2(x+2z)dm(z).$$

(3) 根据(1)(2), 对于$(X,\mathcal{X},\mu,T)$ 为遍历系统. 那么对于任何$f_1,f_2\in L^\infty(X,\mu)$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2$$

在L^2 中存在. 定理证毕. □

注记 9.5.6 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $a\neq b\in\mathbb{Z}\setminus\{0\}$. 那么对于任何$f_1,f_2\in L^\infty(X,\mu)$, 在$L^2$ 中

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^{an}f_1T^{bn}f_2=\int_Z\widetilde{f}_1(\pi(x)+a\theta)\widetilde{f}_2(\pi(x)+b\theta)dm(\theta).$$

§9.5.3 Kronecker 系统的多重回复性

Roth 定理9.5.2 将由定理9.5.5 以及下面定理直接得到.

定理 9.5.7 设$(Z,\mathcal{K},m,R_a)$ 为Kronecker 系统, 即Z 为紧致度量交换群, $a\in Z$, $R_a(z)=az$. 那么对于任何$A\in\mathcal{K}$, $m(A)>0$ 以及$k\in\mathbb{N}$, 有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}m(A\cap R_a^{-n}A\cap R_a^{2n}A\cap\ldots\cap R_a^{-kn}A)>0. \tag{9.5.13}$$

证明. 设$f\in L^1(Z,m), g\in Z$, 记

$$f^g(h)=f(gh).$$

首先我们证明对于任何$f\in L^\infty(Z,m)$,

$$Z\to L^1(Z,m),\quad g\mapsto f^g$$

为连续映射. 设d 为Z 的度量. 对于任何$\varepsilon>0$, 取$\widetilde{f}\in C(Z)$ 使得$\|f-\widetilde{f}\|_1<\varepsilon$, 再根据紧致性, 取$\delta>0$ 使得$d(g_1,g_2)<\delta$ 蕴含$|\widetilde{f}(g_1h)-\widetilde{f}(g_2h)|<\varepsilon,\forall h\in G$. 于是当$d(g_1,g_2)<\delta$ 时,

$$\|f^{g_1}-f^{g_2}\|_1\le\|f^{g_1}-\widetilde{f}^{g_1}\|_1+\|\widetilde{f}^{g_1}-\widetilde{f}^{g_2}\|_1+\|\widetilde{f}^{g_2}-f^{g_2}\|_1<3\varepsilon.$$

所以$g\mapsto f^g$ 为连续的.

取定$f\in L^\infty(Z,m)$ 以及$k\in\mathbb{N}$, 我们下面证明

$$G\to L^1(Z,m),\ g\mapsto\int_Z f(h)f(gh)\ldots f(g^kh)dm(h)$$

为连续的. 因为$g\mapsto f^{g^j}, 1\le j\le k$ 为连续的, 所以对于任何$\varepsilon>0$, 取$\delta>0$ 使得$d(g_1,g_2)<\delta$ 蕴含

$$\|f^{g_1^j}-f^{g_2^j}\|_1<\varepsilon,\quad 1\le j\le k.$$

根据引理9.5.3,

$$\begin{aligned}
&\left\|\int_Z f(h)f(g_1h)\ldots f(g_1^kh)dm(h)-\int_Z f(h)f(g_2h)\ldots f(g_2^kh)dm(h)\right\|_1\\
=&\left\|\sum_{i=1}^k\int_Z f(h)f(g_1h)\ldots f(g_1^{i-1}h)\big(f(g_1^ih)-f(g_2^ih)\big)f(g_2^{i+1}h)\ldots f(g^kh)dm(h)\right\|_1\\
\le&\sum_{i=1}^k\left\|\int_Z f(h)f(g_1h)\ldots f(g_1^{i-1}h)\big(f(g_1^ih)-f(g_2^ih)\big)f(g_2^{i+1}h)\ldots f(g^kh)dm(h)\right\|_1\\
\le& k\varepsilon\|f\|_\infty^k.
\end{aligned}$$

最后, 我们设$f\in L^\infty(Z,m)$ 为非负的, 并且$f\neq 0, a.e.$. 我们证明

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_Z f(h)f(R_a^nh)\ldots f(R_a^{kn}h)dm(h)>0. \tag{9.5.14}$$

根据上面证明, 函数

$$\phi(a)=\int_Z f(h)f(ah)\dots f(a^kh)dm(h)$$

为连续的. 因为(Z,R_a) 为唯一遍历的, 所以

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_Z f(h)f(R_a^nh)\dots f(R_a^{kn}h)dm(h)\\
=&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_Z f(h)f(a^nh)\dots f(a^{kn}h)dm(h)\\
=&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\phi(a^n)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\phi(R_a^ne_Z)\\
=&\int_Z\phi(h)dm(h),
\end{aligned}$$

其中e_Z 为Z 单位元. 因为$\phi(e_Z)=\int_Z(f(h))^{k+1}dm(h)>0$ 以及ϕ 连续, 所以

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_Z f(h)f(R_a^nh)\dots f(R_a^{kn}h)dm(h)>0.$$

□

§9.6 Furstenberg-Zimmer结构定理

§9.6.1 $\mathbb{Z}$作用下的Furstenberg-Zimmer结构定理

收到Furstenberg 极小distal系统结构定理的启发, Parry 于1967 年引入了测度distal的概念. Zimmer 在1976年左右对于一般的局部紧群作用下动力系统发展了测度distal 系统的理论, 并且给出了遍历系统的结构定理. 于此同时, Furstenberg 也在独立的给出了一样的结论来作为工具证明Szemerédi 定理. 所以一般而言, 将这个遍历系统的结构定理称为Furstenberg-Zimmer 结构定理, 这个定理指出任何系统都是测度distal系统的弱混合扩充, 而测度distal系统是由若干紧扩充得到的.

首先我们介绍测度distal, 然后给出Furstenberg-Zimmer结构定理的陈述.

定义 9.6.1 (Parry) *设$(X,\mathcal{X},\mu,T)$ 为遍历系统, 一个$\mathcal{X}$ 中可测集列$A_1\supset A_2\supset\dots$ 称为***分离筛** *(separating sieve) 是指, $\mu(A_n)>0$, $\mu(A_n)\to 0, n\to\infty$, 并且存在子集$X_0\in\mathcal{X}$, $\mu(X_0)=1$ 使得对于任何$x,x'\in X_0$ 它们满足条件"对于任何$n\in\mathbb{N}$ 存在$k\in\mathbb{Z}$ 使得$T^kx,T^kx'\in A_n$", 那么就有$x=x'$.*

*我们称遍历系统$(X,\mathcal{X},\mu,T)$ 为***测度*distal*** *是指要么X 为周期系统, 要么它有一个分离筛.*

Parry 证明了对于极小拓扑系统(X,T) 为拓扑distal 的, 那么对于任何$\mu \in \mathcal{M}(X,T)$, $(X,\mathcal{B}_X,\mu,T)$ 为测度distal的. 并且任何测度distal系统具有零熵.

设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,T)$ 为两个保测系统间的因子映射, 且设

$$\mu=\int_Y \mu_y d\nu(y)$$

为μ 相对于ν 的积分分解.

定义 9.6.2 (相对几乎周期函数) 一个函数$f\in L^2(X,\mathcal{X},\mu)$ 称为**相对Y 几乎周期的** *(almost periodic over Y)*, 是指对于任何$\varepsilon>0$, 存在$g_1,\dots,g_l\in L^2(X,\mathcal{X},\mu)$ 使得对于任何$n\in\mathbb{Z}$

$$\min_{1\le j\le l}\left\|T^nf-g_j\right\|_{L^2(\mu_y)}<\varepsilon$$

对ν 几乎处处$y\in Y$ 成立.

定义 9.6.3 (紧扩充与弱混合扩充) 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,T)$ 为两个保测系统间的因子映射.

1. 我们称X 为相对于Y 的**紧扩充***(compact extension)* 或者**等距扩充***(isometric extensions)*, 是指相对Y 的几乎周期函数在$L^2(X,\mu)$ 中稠密.
2. 我们称X 为相对于Y 的**弱混合扩充** *(relatively weak mixing extension)*, 如果$(X\times X,\mathcal{X}\times\mathcal{X},\mu\times_Y\mu,T\times T)$ 为遍历的.

设$K(X|Y,T)$ 为$L^2(X)$ 中由相对$\mathcal{Y}$ 几乎周期函数张成的闭子空间. 当$\mathcal{Y}$ 平凡时, $K(X|Y,T)$ 为$H_c=L^2(X,\mathcal{K}_\mu,\mu)$, T 特征函数张成的空间.

注记 9.6.4 根据$K(X|Y,T)$ 我们可以用之来定义紧扩充和弱混合扩充. 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,T)$ 为两个保测系统间的因子映射. 我们称X 为相对于Y 的**紧扩充***(compact extension)* 或者**等距扩充***(isometric extensions)*, 如果$K(X|Y,T)=L^2(X)$; 称X 为相对于Y 的**弱混合扩充** *(relatively weak mixing extension)*, 如果$K(X|Y,T)=L^2(Y)$. 我们会看到这些定义是等价的.

注记 9.6.5 可以证明一个遍历系统X 为(Y,ν,S) 的紧扩充当且仅当X 同构于斜积$X'=Y\times M$, 其中$M=G/H$, G 为紧致度量空间, H 为闭子群以及$\mu'=\nu\times m_M$ (m_M 为M 由G 诱导的*Haar* 测度). 此处X' 上作用T' 由下式给出

$$T'(y,gH)=(Sy,\alpha(y)gH),$$

其中$\alpha:Y\to G$ 为*cocycle*. 一般记X',T' 为$Y\times_\alpha G/H$ 和T_α. 当H 为平凡的, 我们称$Y\times_\rho G$ 为Y 的群扩充.我们用不到这些结果, 感兴趣读者请参见*[76]*.

定义 9.6.6 (Zimmer) 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, 那么$(X,\mathcal{X},\mu,T)$ 为测度*distal* 的当且仅当存在可数序数η 以及有序因子列$(X_\theta,\mathcal{X}_\theta,\mu_\theta,T),\theta\le\eta$ 使得

1. $X_0=\{pt\}$ 为平凡的, 且$X_\eta=X$.
2. 对于$\theta<\eta$ 扩充$\pi_\theta:X_{\theta+1}\to X_\theta$ 为紧的且非平凡的(即不为同构).
3. 对于极限序数$\lambda\le\eta$, $X_\lambda=\varprojlim_{\theta<\lambda} X_\theta$ (即$\mathcal{X}_\lambda=\bigvee\mathcal{X}_\theta$).

$$\{pt\}=X_0\xleftarrow{\pi_0}X_1\xleftarrow{\pi_1}\cdots\xleftarrow{\pi_{\theta-1}}X_\theta\xleftarrow{\pi_\theta}X_{\theta+1}\xleftarrow{\pi_{\theta+1}}\cdots\longleftarrow X_\eta=X.$$

注记 9.6.7 我们在本书中不会用到*Parry*关于测度*distal*的定义, 我们不会给出上面定理的证明, 而直接把上面定理作为*distal* 的定义.

下面是$\mathbb{Z}$ 作用下的Furstenberg-Zimmer结构定理, 对于一般局部紧群作用下的结构定理表述是类似的. 对于$\mathbb{Z}^l$ 作用下非遍历保测系统的结构定理我们在本节后面介绍.

定理 9.6.8 (Furstenberg-Zimmer结构定理) 任何遍历的保测系统$(X,\mathcal{X},\mu,T)$ 都是测度*distal*系统的弱混合扩充:

$$\underbrace{X_0\xleftarrow{\pi_0}X_1\xleftarrow{\pi_1}\cdots\xleftarrow{\pi_{n-1}}X_n\xleftarrow{\pi_n}X_{n+1}\xleftarrow{\pi_{n+1}}\cdots\longleftarrow X_\eta}_{\text{distal 因子}}\xleftarrow{\pi_\eta}X.$$

§9.6.2 Furstenberg-Zimmer结构定理的证明

定理 9.6.9 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为两个遍历系统间的因子映射. 那么下面之一成立:

1. X 为相对于Y 的弱混合扩充;
2. 存在X 的因子$(Z,\mathcal{Z},\eta,R)$ 使得图表交换

$$\begin{tikzcd} X \arrow[d, "\pi"'] \arrow[dr, "\sigma"] & \\ Y & Z, \arrow[l, "\rho"] \end{tikzcd}$$

其中$\rho:Z\to Y$ 为非平凡紧扩充.

证明. 设

$$\widetilde{X}=X\times X,\ \widetilde{\mu}=\mu\times_Y\mu,\ \widetilde{T}=T\times T.$$

设

$$\mu=\int_Y\mu_y d\nu(y)$$

为μ 相对于ν 的积分分解. 那么

$$\widetilde{\mu} = \mu \times_Y \mu = \int_Y \mu_y \times \mu_y d\nu(y).$$

在证明中, 我们经常会将$\mu_y = \mu_{\pi(x)}$ 记为μ_x.

如果$\pi : X \to Y$ 不是弱混合扩充, 那么存在非常值的$\widetilde{T}$-不变函数$H \in L^\infty(\widetilde{X}, \widetilde{\mu})$ 对于$\phi \in L^2(X, \mu)$ 定义算子

$$\begin{aligned} H * \phi(x) &= \int_X H(x, x')\phi(x')d\mu_x(x') \\ &= \int_X H(x, x')\phi(x')d\mu_{\pi(x)}(x'). \end{aligned} \tag{9.6.1}$$

首先我们有

$$\begin{aligned} T(H * \phi)(x) &= H * \phi(Tx) \\ &= \int_X H(Tx, x')\phi(x')d\mu_{Tx}(x') \\ &= \int_X H(Tx, Tx')\phi(Tx')d\mu_x(x') \\ &= H * (T\phi)(x). \end{aligned} \tag{9.6.2}$$

如果$\phi \in L^\infty(X, \mu)$, 那么对于取定$y \in Y$, $\{T^n\phi : n \in \mathbb{Z}\} \subseteq L^\infty(X, \mu_y)$. 因为

$$L^2(X, \mu_y) \to L^2(X, \mu_y),\ \phi \mapsto H * \phi$$

为紧算子, 所以对于任何$y \in Y$ 和$\phi \in L^\infty(X, \mu)$, 集合

$$\{T^n(H * \phi) = H * (T^n\phi) : n \in \mathbb{Z}\} \subseteq L^2(X, \mu_y)$$

为完全有界的.

我们需要证明$H*\phi$ 是相对于Y 的几乎周期函数. 对于任何$\varepsilon > 0, y \in Y$. 存在$M(y) \in \mathbb{N}$ 使得

$$\{T^j(H * \phi) : |j| \le M(y)\}$$

为$\{T^n(H * \phi) : n \in \mathbb{Z}\}$ 相对于μ_y 的ε-网. 设$M(y)$ 为使得上面条件满足的最小的数值. 于是定义了可测函数

$$M : Y \to \mathbb{N},\ y \mapsto M(y).$$

这可由下面证明：注意$\{y \in Y : M(y) \le K\}$ 等于

$$\{y \in Y : \forall m \in \mathbb{Z}, \exists |j| \le K\ s.t.\ \|T^m(H * \phi) - T^j(H * \phi)\|_{L^2(\mu_y)} < \varepsilon\}.$$

于是$\{y \in Y : M(y) \le K\}$ 为可测集, 进而M 为可测函数. 所以存在$M \in \mathbb{N}$ 使得

$$A = \{y \in Y : M(y) \le M\} \in \mathcal{Y},$$

且$\nu(A)>0$.

对于每个$j\in\mathbb{Z}$, $|j|\le M$, 定义g_j 如下：

$$g_j(x)=\begin{cases} T^j(H*\phi)(x)=H*(T^j\phi)(x), & \text{如}A\in\mathcal{A};\\ g_j(T^m x), & y=\pi(x),Sy,\dots,S^{m-1}y\notin A \text{ 并且} S^m y\in A.\end{cases}$$

根据遍历性, $\nu(\bigcup_{m\in\mathbb{N}}S^{-m}A)=1$, 所以$g_j$ 是几乎处处可定义的.

对于$y\in A$, 根据假设

$$\min_{-M\le j\le M}\|T^n(H*\phi)-g_j\|_{L^2(\mu_y)}<\varepsilon,\ \forall n\in\mathbb{Z}.$$

如果存在$m\in\mathbb{N}$ 使得$y,Sy,\dots,S^{m-1}y\notin A$, 但是$S^m y\in A$. 对于$S^m y$ 运用上式

$$\min_{-M\le j\le M}\|T^n(H*\phi)-g_j\|_{L^2(S^m\mu_y)}<\varepsilon,\ \forall n\in\mathbb{Z}.$$

注意到

$$\|T^n(H*\phi)-g_j\|_{L^2(S^m\mu_y)}=\|T^mT^n(H*\phi)-T^mg_j\|_{L^2(\mu_y)}$$

以及$T^mg_j(x)=g_j(T^mx)=g_j(x)$, 我们得到

$$\min_{-M\le j\le M}\left\|T^{m+n}(H*\phi)-g_j\right\|_{L^2(\mu_y)}<\varepsilon,\ \forall n\in\mathbb{Z}.$$

上式也可写为

$$\min_{-M\le j\le M}\|T^n(H*\phi)-g_j\|_{L^2(\mu_y)}<\varepsilon,\ \forall n\in\mathbb{Z}.$$

综上, 我们得到对于ν几乎处处的y,

$$\min_{-M\le j\le M}\|T^n(H*\phi)-g_j\|_{L^2(\mu_y)}<\varepsilon,\ \forall n\in\mathbb{Z}.$$

换而言之, $H*\phi$ 为相对于Y 的几乎周期函数.

下面我们证明断言：

断言： *存在$\phi\in L^\infty(X,\mu)$ 使得$H*\phi\notin L^2(Y,\nu)$.*

断言证明： 否则我们假设对于任何$\phi\in L^\infty(X,\mu)$, $H*\phi\in L^2(Y,\nu)$.

我们需要用到下面事实：设$\mathcal{N}=\{\emptyset,X\}$

$$H*\phi(x)=\mathbb{E}\left(H(x,x')\phi(x')|\mathcal{X}\times\mathcal{N}\right). \tag{9.6.3}$$

这个事实是由下式得到的

$$\mu\times_Y\mu=\int_Y\mu_y\times\mu_y d\nu(y)=\int_X\delta_x\times\mu_{\pi(x)}d\mu(x).$$

取X 有限剖分$\mathcal{P}_n$ 使得

$$\sigma(\mathcal{P}_n) \nearrow \mathcal{X},\ n \to \infty.$$

设$x_2 \in P \in \mathcal{P}_n$, 根据条件期望定义,

$$\mathbb{E}(H|\mathcal{X} \times \sigma(\mathcal{P}_n))(x_1, x_2) = \frac{\mathbb{E}(H \cdot \mathbf{1}_{X \times P}|\mathcal{X} \times \mathcal{N})(x_1, x_2)}{\mathbb{E}(\mathbf{1}_{X \times P}|\mathcal{X} \times \mathcal{N})(x_1, x_2)}. \tag{9.6.4}$$

根据(9.6.3) 与我们的假设, $\mathbb{E}(H \cdot \mathbf{1}_{X \times P}|\mathcal{X} \times \mathcal{N})$ 为$\mathcal{Y} \times \mathcal{N}$ 可测. 同样的根据

$$\mathbb{E}(\mathbf{1}_{X \times P}|\mathcal{X} \times \mathcal{N})(x_1, x_2) = \mu_{x_1}(P)$$

知道$\mathbb{E}(\mathbf{1}_{X \times P}|\mathcal{X} \times \mathcal{N})(x_1, x_2)$ 也是$\mathcal{Y} \times \mathcal{N}$ 可测的. 于是$\mathbb{E}(H|\mathcal{X} \times \sigma(\mathcal{P}_n))$ 为$\mathcal{Y} \times \mathcal{N}$ 可测的.

根据鞅定理, 以及

$$\mathcal{X} \times \sigma(\mathcal{P}_n) \nearrow \mathcal{X} \times \mathcal{X},$$

我们得到H 为$\mathcal{Y} \times \mathcal{N}$ 可测的. 注意到

$$\mathcal{Y} \times \mathcal{N} = \mathcal{N} \times \mathcal{Y} \mod \widetilde{\mu}.$$

所以H 为$\mathcal{N} \times \mathcal{Y}$ 可测的. 即$H(x_1, x_2)$ 只是第二个坐标的函数. 但是H 是非常值不变的, 这与T 遍历矛盾. □

令

$$\mathcal{F} = \{f \in L^\infty(X, \mu) : f \text{ 为相对于} Y \text{ 几乎周期的}\}.$$

根据断言, $\mathcal{F}$ 包含非$\mathcal{Y}$ 可测的函数. 下面证明$\mathcal{F}$ 为代数. 我们仅需验证$\mathcal{F}$ 对于乘法封闭.

$f_1, f_2 \in \mathcal{F}$. 对于$\varepsilon > 0$, 取$g_1, \ldots, g_J \in L^2(X, \mu)$, $h_1, \ldots, h_K \in L^2(X, \mu)$ 使得$\|g_j\|_\infty \leq \|f_1\|_\infty, 1 \leq j \leq J$ 对ν $a.e.$ $y \in Y$

$$\min_{1 \leq j \leq J} \|T^n f_1 - g_j\|_{L^2(\mu_y)} < \varepsilon, \tag{9.6.5}$$

以及

$$\min_{1 \leq k \leq K} \|T^n f_2 - h_k\|_{L^2(\mu_y)} < \varepsilon. \tag{9.6.6}$$

于是由

$$\begin{aligned}
&\|T^n(f_1 f_2) - g_j h_k\|_{L^2(\mu_y)} \\
\leq\ &\|T^n f_2 - (T^n f_1 - g_j)\|_{L^2(\mu_y)} + \|g_j(T^n f_2 - h_k)\|_{L^2(\mu_y)} \\
<\ &\|f_2\|_\infty \varepsilon + \|g_j\|_\infty \varepsilon \leq (\|f_1\|_\infty + \|f_2\|_\infty)\varepsilon.
\end{aligned}$$

于是

$$\min_{1 \leq j \leq J, 1 \leq k \leq K} \|T^n(f_1 f_2) - g_j h_k\|_{L^2(\mu_y)} < (\|f_1\|_\infty + \|f_2\|_\infty)\varepsilon.$$

所以$f_1 f_2 \in \mathcal{F}$.

根据定理2.5.15, 存在X 的因子$(Z,\mathcal{Z},\eta,R)$ 使得

$$\overline{\mathcal{F}} = L^2(Z,\mathcal{Z},\eta).$$

并且存在交换图表:

$$\begin{tikzcd} X \arrow[d, "\pi"'] \arrow[dr, "\sigma"] & \\ Y & Z, \arrow[l, "\rho"] \end{tikzcd}$$

其中$\rho: Z \to Y$ 为非平凡紧扩充. □

§9.6.3 $\mathbb{Z}^l$保测系统的结构定理

设Γ为与$\mathbb{Z}^l$同构的群, 其中$l \in \mathbb{N}$. 设$X = (X,\mathcal{B},\mu,\Gamma)$ 为保测系统,$Y = (Y,\mathcal{D},\nu,\Gamma)$为其因子, 而$\phi: (X,\mathcal{B},\mu,\Gamma) \to (Y,\mathcal{D},\nu,\Gamma)$ 为相应的因子映射.

定义 9.6.10 *称系统X为Y相对于$T \in \Gamma$的***遍历扩充***是指$\mathcal{B}$的全体T不变子集恰为$\mathcal{D}$的T不变子集在ϕ^{-1}下的像（指模去零测集的意义下）. 记为$X \to Y$遍历（rel. T).*

定义 9.6.11 *称系统X为Y相对于$T \in \Gamma$的***弱混合扩充***是指$X \times_Y X \to Y$ 为Y相对于T的遍历扩充. 记为$X \to Y$弱混合(rel. T). 如果ϕ相对于每个$T \in \Gamma$为弱混合的, 那么称X为Y相对于Γ的弱混合扩充.*

定义 9.6.12 *称X为Y的***紧扩充***是指存在稠密函数集族$\mathcal{F} \subseteq L^2(X)$满足: 对任意$f \in \mathcal{F}$及$\delta > 0$, 存在有限个函数$g_1, g_2, \ldots, g_k \in L^2(X)$使得对每个$S \in \Gamma$, 不等式*

$$\min_{1 \le j \le k} ||Sf - g_j||_{L^2(X,\mu_y)} < \delta, \nu \ a.e. \ y \in Y.$$

下面我们陈述Γ作用下的Furstenberg-Zimmer定理

定理 9.6.13 (Furstenberg-Zimmer结构定理) *设$X = (X,\mathcal{B},\mu,\Gamma)$为可分的保测系统,其中$\Gamma$为有限秩的自由交换群. 那么存在序数$\eta$使得对任意序数$\xi \le \eta$存在$X$的因子$X_\xi$, $\pi_\xi: X \to X_\xi$满足:*

(i) *X_0为平凡系统,而$X_\eta = X$.*

(ii) *如果$\xi \le \xi'$, 则存在同态$\pi_\xi^{\xi'}: X_{\xi'} \to X_\xi$使得$\pi_\xi = \pi_\xi^{\xi'}\pi_{\xi'}$.*

(iii) *对每个$\xi < \eta$, $X_{\xi+1} \to X_\xi$为本原的,即Γ为两个子群的直积$\Gamma = \Gamma_c \times \Gamma_w$,其中$X_{\xi+1} \to X_\xi$相对于$\Gamma_c$为紧扩充,而$X_{\xi+1} \to X_\xi$ 相对于Γ_w为弱混合扩充.*

(iv) *如果$\xi \le \eta$为极限序数,则X_ξ为因子$\{X_{\xi'}, \xi' < \xi\}$的极限.*

当$T=\mathbb{Z}$时,我们可以做到除最后一步$X=X_\eta\to X_{\eta'}(\eta=\eta'+1)$可能为弱混合扩充外其余各扩充均为紧扩充.

通过上面结论可以得到有限个交换作用下的多重回复定理. 保测系统的多重回复定理为下定理的推论:

定理 9.6.14 设$T_1,T_2,\ldots,T_l$为概率空间$(X,\mathcal{B},\mu)$上的相互可交换的保测变换,$A\in\mathcal{B}$满足$\mu(A)>0$. 那么

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(T_1^{-n}A\cap T_2^{-n}A\cap\ldots\cap T_l^{-n}A)>0. \tag{9.6.7}$$

证明定理9.6.14的基本思路为:先对$(X,\mathcal{B},\mu,\Gamma)$为$(Y,\mathcal{D},\nu,\Gamma)$的紧扩充和弱混合扩充时分别证明,如果$(Y,\mathcal{D},\nu,\Gamma)$满足(9.6.7),那么$(X,\mathcal{B},\mu,\Gamma)$也满足(9.6.7);然后用结构定理归纳证明结论.根据定理9.6.14,我们有

定理 9.6.15 设$T_1,T_2,\ldots,T_l$为概率空间$(X,\mathcal{B},\mu)$上的相互可交换的保测变换,$A\in\mathcal{B}$满足$\mu(A)>0$. 那么存在$n\geq 1$使得

$$\mu(T_1^{-n}A\cap T_2^{-n}A\cap\ldots\cap T_l^{-n}A)>0.$$

习 题

1. 根据定理9.6.9 完成Furstenberg-Zimmer定理的证明.

§9.7 Poincaré多重回复定理的证明

§9.7.1 SZ性质

定义 9.7.1 设$(X,\mathcal{X},\mu,T)$ 为保测系统, 我们称它满足***SZ*性质**, 是指对于任何$k\in\mathbb{N}$, $A\in\mathcal{X}$, $\mu(A)>0$, 我们有

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A\cap T^{-2n}A\cap\ldots\cap T^{-kn}A)>0. \tag{9.7.1}$$

根据自然扩充, 我们可以不妨设所涉及的系统均为可逆的. 下面说明, 我们可以约化到遍历系统. 设

$$\mu=\int_\Omega \mu_y d\nu(y)$$

为遍历分解.

设$A\in\mathcal{X}$, $\mu(A)>0$. 则

$$\mu\big(\{x\in X:\mu_y(A)>0\}\big)>0.$$

由Fatou引理,

$$\begin{aligned}
&\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A\cap\ldots\cap T^{-kn}A)\\
=&\liminf_{N\to\infty}\int_{\Omega}\frac{1}{N}\sum_{n=0}^{N-1}\mu_y(A\cap T^{-n}A\cap\ldots\cap T^{-kn}A)d\nu(y)\\
\geq&\int_{\Omega}\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu_y(A\cap T^{-n}A\cap\ldots\cap T^{-kn}A)d\nu(y)\\
>&\,0.
\end{aligned}$$

最后的正号是因为如果对于遍历系统SZ成立, 那么对于正测集的y, 积分号里面值大于0. 所以我们把命题约化到遍历系统.

我们分为紧扩充、弱混合扩充以及逆极限三种情况证明结论.

§9.7.2 紧扩充

引理 9.7.2 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为两个遍历系统间的紧扩充. 设$B\in\mathcal{X}$, $\mu(B)>0$. 那么存在$\widetilde{B}\subseteq B,\mu(\widetilde{B})>0$ 使得$\mathbf{1}_{\widetilde{B}}$ 为相对Y 几乎周期函数, $A=\pi(\widetilde{B})\in\mathcal{Y}$ 具有正测度, 并且

$$\mu_y(\widetilde{B})>\frac{1}{2}\mu(\widetilde{B}),\ \forall y\in A \text{ 以及 } \mu_y(\widetilde{B})=0,\ \forall y\in Y\setminus A.$$

证明. 设

$$B'=\{x\in B:\mu_{\pi(x)}(B)>\frac{1}{2}\mu(B)\}$$

如果$x\in B\setminus B'$, 设$y=\pi(x)$, 则有

$$\mu_y(B\setminus B')\leq\mu_y(B)\leq\frac{1}{2}\mu(B).$$

对y 做积分, 就得到

$$\mu(B\setminus B')\leq\frac{1}{2}\mu(B).$$

因为$\mu(B)>0$. 所以由此我们得到$\mu(B')>0$, 并且对于任何$x\in B'$,

$$\mu_{\pi(x)}(B')>\frac{1}{2}\mu(B)\geq\frac{1}{2}\mu(B').$$

下面需要微调之使得它成为几乎周期的.

取递减序列$\{\varepsilon_l\}_{l=1}^{\infty}$ 使得$\varepsilon_l>0,\forall l\in\mathbb{N}$ 且

$$\sum_{l=1}^{\infty}\varepsilon_l<\frac{1}{2}\mu(B').$$

对于每个$l\in\mathbb{N}$, 取Y 几乎周期函数f_l 使得

$$\|\mathbf{1}_{B'}-f_l\|_{L^2(\mu)}^2=\int_X|\mathbf{1}_{B'}-f_l|^2d\mu<\varepsilon_l^2.$$

令

$$A_l=\{y\in Y:\|\mathbf{1}_{B'}-f_l\|^2_{L^2(\mu_y)}\geq\varepsilon_l\}.$$

则

$$\mu(\pi^{-1}(A_l))=\nu(A_l)\leq\frac{1}{\varepsilon_l}\int_{A_l}\|\mathbf{1}_{B'}-f_l\|^2_{L^2(\mu_y)}d\nu(y)\leq\varepsilon_l.$$

设$A=\bigcup_{l=1}^\infty A_l$ 以及

$$\widetilde{B}=B'\setminus\bigcup_{l=1}^\infty\pi^{-1}(A_l).$$

首先我们有

$$\mu(\widetilde{B})=\mu(B')-\mu(\bigcup_{l=1}^\infty\pi^{-1}(A_l))\geq\mu(B')-\sum_{l=1}^\infty\varepsilon_l>\frac{1}{2}\mu(B').$$

根据$\widetilde{B}$ 定义,

$$\mu_y(\widetilde{B})=\begin{cases}\mu_y(B')>\frac{1}{2}\mu(B')\geq\frac{1}{2}\mu(\widetilde{B}), & x\in\widetilde{B},y=\pi(x)\in A;\\ 0, & \text{否则}.\end{cases}$$

下说明$\mathbf{1}_{\widetilde{B}}$ 为几乎周期的.

设$\varepsilon>0$. 取$\varepsilon_l<\frac{1}{2}\varepsilon$. 如果$T^ny\notin\bigcup_{l\in\mathbb{N}}A_l$,

$$\begin{aligned}\|T^n\mathbf{1}_{\widetilde{B}}-T^nf_l\|_{L^2(\mu_y)}&=\|\mathbf{1}_{\widetilde{B}}-f_l\|_{L^2(\mu_{T^ny})}\\&=\|\mathbf{1}_{B'}-f_l\|_{L^2(\mu_{T^ny})}\\&<\varepsilon_l<\frac{1}{2}\varepsilon.\end{aligned}$$

如果$T^ny\in\bigcup_{l\in\mathbb{N}}A_l$, 则

$$\|T^n\mathbf{1}_{\widetilde{B}}\|_{L^2(\mu_y)}=\|\mathbf{1}_{\widetilde{B}}\|_{L^2(\mu_{T^ny})}=0.$$

因为f_l 为Y几乎周期的, 所以存在函数$g_1,\ldots,g_m$ 使得

$$\min_{1\leq j\leq m}\|T^nf_l-g_j\|_{L^2(\mu_y)}\leq\frac{1}{2}\varepsilon,\ \forall y\ a.e.$$

设$g_0=0$, 则

$$\min_{0\leq j\leq m}\left\|T^n\mathbf{1}_{\widetilde{B}}-g_j\right\|_{L^2(\mu_y)}\leq\varepsilon,\ \forall y\ a.e.$$

即$\mathbf{1}_{\widetilde{B}}$ 为Y 几乎周期的. □

命题 9.7.3 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为两个遍历系统间的紧扩充. 如果$(Y,\mathcal{Y},\nu,S)$ 满足SZ性质, 那么$(X,\mathcal{X},\mu,T)$ 也满足SZ 性质.

证明. 设$B \in \mathcal{X}$, $\mu(B) > 0$. 我们要证它具有SZ 性质. 根据引理9.7.2, 我们假设$f = \mathbf{1}_B$ 为Y 几乎周期的, $A = \pi(B) \in \mathcal{Y}$ 具有正测度，并且

$$\mu_y(B) > \frac{1}{2}\mu(B),\ \forall y \in A \text{ 以及 } \mu_y(B) = 0\ \forall y \in Y \setminus A.$$

我们将用$A = \pi(B)$ 的SZ 性质推出所求.

设$k \in \mathbb{N}$. 因为f 为Y 几乎周期的, 对于$\varepsilon = \frac{\mu(B)}{12(k+1)}$ 存在$g_1, \dots, g_r$ 使得

$$\min_{1 \le s \le r} \|T^n f - g_s\|_{L^2(\mu_y)} < \varepsilon,\ \forall n \in \mathbb{Z}$$

对于几乎每个$y \in Y$ 成立. 我们可以假设$\|g_s\|_\infty \le 1, 1 \le s \le r$. 根据van der Waerden 定理, 取K 使得$[1, K] = \{1, 2, \dots, K\}$ 的任何r 染色中包含$k+1$ 长的等差数列单色. 取$c_1 > 0$ 使得

$$R_K = \{n \in \mathbb{Z}_+ : \nu(A \cap S^{-n}A \cap \dots \cap S^{-Kn}A) > c_1\}$$

具有正的下密度. [[这是因为对A 用SZ 性质，存在$c_0 > 0$ 使得$\liminf\limits_{N\to\infty} \dfrac{1}{N}\sum\limits_{n=0}^{N-1} \nu(A \cap S^{-n}A \cap \dots \cap S^{-Kn}A) \ge c_0 > 0$. 于是对于充分大$N$ 以及仅依赖于c_0 的c_1, 成立$\dfrac{1}{N}|R_K \cap [0, N-1]| > c_1$. 由此得到$R_K$ 具有正的下密度.]]

设$n \in R_K$. 对于$y \in A \cap S^{-n}A \cap \dots \cap S^{-Kn}A$, 我们有

$$\min_{1 \le s \le r} \|T^{in} f - g_s\|_{L^2(\mu_y)} < \varepsilon,\quad i \in \{1, 2, \dots, K\}.$$

利用此给$[1, K]$ 进行r 染色, 对于$i \in [1, K]$, 取$c(i) \in \{1, 2, \dots, r\}$ 使得

$$\|T^{in} f - g_{c(i)}\|_{L^2(\mu_y)} < \varepsilon.$$

即$c : [1, K] \to \{1, 2, \dots, r\}$ 为染色, 它仅依赖于y 与n. 由K 的选取, 存在i, d 以及$t \in \{1, 2, \dots, r\}$使得

$$i, i+d, \dots, i+kd \in c^{-1}(t).$$

尤其有

$$\|T^{(i+jd)n} f - g_t\|_{L^2(\mu_y)} < \varepsilon,\ \forall\ 0 \le j \le k.$$

设$\widetilde{g} = T^{-in} g_t$, 则

$$\|T^{jdn} f - \widetilde{g}\|_{L^2(\mu_{S^{in}y})} < \varepsilon,\ \forall\ 0 \le j \le k.$$

因为在上式中可以取$j = 0$, 所以由三角不等式有

$$\|T^{jdn} f - f\|_{L^2(\mu_{S^{in}y})} < 2\varepsilon,\ \forall\ 1 \le j \le k. \tag{9.7.2}$$

综上, 对于$n \in R_K$, 集合$A \cap S^{-n}A \cap \dots \cap S^{-Kn}A$ 有剖分

$$A \cap S^{-n}A \cap \dots \cap S^{-Kn}A = D_{n,1} \cup \dots \cup D_{n,M},$$

其中M 为$[1,K]$ 中长为$k+1$ 的等差数列的总个数, 在每个$y\in D_{n,h}$ 上面分析中的$i,d,\widetilde{g}$ 都保持相同. 因为$\nu(A\cap S^{-n}A\cap\ldots\cap S^{-Kn}A)>c_1$, 存在某个$D=D_{n,h}$ 使得

$$\nu(D)>\frac{c_1}{M}.$$

我们设D 对应的等差数列仍记为$i,i+d,\ldots,i+kd$.

因为$D\subseteq A\cap S^{-n}A\cap\ldots\cap S^{-Kn}A$, 我们有

$$\mu_{S^{in}y}(B)>\frac{1}{2}\mu(B).$$

于是

$$\begin{aligned}
&\mu_{S^{in}y}(B\cap T^{-dn}B\cap\ldots\cap T^{-k(dn)}B)\\
=&\int_X f\cdot T^{dn}f\cdots T^{kdn}fd\mu_{S^{in}y}\\
>&\int_X f^{k+1}d\mu_{S^{in}y}-2(k+1)\varepsilon\quad(\text{由 }(9.7.2))\\
>&\frac{1}{2}\mu(B)-2(k+1)\varepsilon=\frac{1}{3}\mu(B).
\end{aligned}$$

上式对于所有$y\in D$ 成立. 由于在D 上, i 为固定的, 所以对于$S^{in}y\in S^{in}D$ 做积分得到

$$\mu(B\cap T^{-dn}B\cap\ldots\cap T^{-k(dn)}B)>\frac{1}{3}\mu(B)\nu(D)\geq\frac{c_1}{3M}\mu(B).$$

对于任何$n\in R_K$, 我们都有上式. 注意d 可能会依赖n 而变化. 设

$$R'=\Big\{n\in\mathbb{Z}_+:\mu(B\cap T^{-n}B\cap\ldots\cap T^{-kn}B)\geq\frac{c_1}{3M}\mu(B).\Big\}.$$

综上我们就有, 对于任何$n\in R_K$, 存在$d\in[1,K]$ 使得$dn\in R'$. 于是(作为习题验证)

$$\liminf_{N\to\infty}\frac{|R'\cap[0,N-1]|}{N}\geq\frac{c_1}{2K^2}.\tag{9.7.3}$$

尤其,

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(B\cap T^{-n}B\cap T^{-2n}B\cap\ldots\cap T^{-kn}B)>0.$$

即$(X,\mathcal{X},\mu,T)$ 满足SZ 性质. □

§9.7.3 弱混合扩充

定理 9.7.4 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为两个遍历系统间的弱混合扩充. 那么对于任何$k\in\mathbb{N}$, $B_0,B_1,\ldots,B_k\in X$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_Y\left|\mu_y(B_0\cap T^{-n}B_1\cap\ldots\cap T^{-kn}B_k)-\mu_y(B_0)\mu_y(T^{-n}B_1)\ldots\mu_y(T^{-kn}B_k)\right|^2d\nu(y)=0\tag{9.7.4}$$

等价地, 对于任意$\varepsilon > 0$ 存在$N_0 \in \mathbb{N}$ 使得对于$N \geq N_0$,

$$\left|\mu_y(B_0 \cap T^{-n}B_1 \cap \ldots \cap T^{-kn}B_k) - \mu_y(B_0)\mu_y(T^{-n}B_1)\ldots\mu_y(T^{-kn}B_k)\right| < \varepsilon, \forall (n,y) \in \big([0,N-1]\times Y\big)\setminus G, \tag{9.7.5}$$

其中$m \times \nu(G) < \varepsilon$, m 为$[0, N-1]$ 上计数测度的规范化概率测度.

注记 9.7.5 我们把(9.7.5) 简记为: 对于充分大N 有

$$\mu_y(B_0 \cap T^{-n}B_1 \cap \ldots \cap T^{-kn}B_k) \stackrel{\varepsilon}{\approx} \mu_y(B_0)\mu_y(T^{-n}B_1)\ldots\mu_y(T^{-kn}B_k)$$

命题 9.7.6 设$\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ 为两个遍历系统间的弱混合扩充. 如果$(Y, \mathcal{Y}, \nu, S)$ 满足SZ 性质, 那么$(X, \mathcal{X}, \mu, T)$ 也满足SZ 性质.

证明. 设$B \in \mathcal{X}$, $\mu(B) > 0$. 对$a > 0$ 令$A = \{y \in Y : \mu_y(B) > a\}$. 取$a > 0$ 使得$\nu(A) > 0$.

根据定理9.7.4, 对于$\varepsilon > 0$, 对于充分大的N 有

$$\mu_y(B_0 \cap T^{-n}B_1 \cap \ldots \cap T^{-kn}B_k) \geq \mu_y(B_0)\mu_y(T^{-n}B_1)\ldots\mu_y(T^{-kn}B_k) - \varepsilon, \forall (n,y) \in \big([0,N-1]\times Y\big)\setminus G,$$

其中$m \times \nu(G) < \varepsilon$. 对于$(n, y) \in G$, 我们有

$$\mu_y(B_0 \cap T^{-n}B_1 \cap \ldots \cap T^{-kn}B_k) \geq 0 \geq \mu_y(B_0)\mu_y(T^{-n}B_1)\ldots\mu_y(T^{-kn}B_k) - \varepsilon - 1.$$

于是我们得到:

$$\begin{aligned}
&\frac{1}{N}\sum_{n=0}^{N-1}\mu(B \cap T^{-n}B \cap \ldots \cap T^{-kn}B) \\
=& \frac{1}{N}\sum_{n=0}^{N-1}\int_Y \mu_y(B \cap T^{-n}B \cap \ldots \cap T^{-kn}B)d\nu(y) \\
\geq& \frac{1}{N}\sum_{n=0}^{N-1}\int_Y \Big(\mu_y(B)\mu_y(T^{-n}B)\ldots\mu_y(T^{-kn}B) - \varepsilon\Big)d\nu(y) - \varepsilon
\end{aligned} \tag{9.7.6}$$

注意到对于$y \in A \cap S^{-n}A \cap \ldots \cap S^{-kn}A$, 我们有

$$\mu_y(T^{-in}B) = \mu_{S^{in}y}(B) > a,\ 0 \leq i \leq k.$$

于是从(9.7.6), 我们得到

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(B \cap T^{-n}B \cap \ldots \cap T^{-kn}B) \geq (a^{k+1} - \varepsilon)\frac{1}{N}\sum_{n=0}^{N-1}\nu(A \cap S^{-n}A \cap \ldots \cap S^{-kn}A) - \varepsilon.$$

于是

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(B \cap T^{-n}B \cap \ldots \cap T^{-kn}B) \geq a^{k+1}\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\nu(A \cap S^{-n}A \cap \ldots \cap S^{-kn}A) > 0.$$

所以X 具有SZ 性质. □

命题 9.7.7 设$\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ 为两个遍历系统间的弱混合扩充. 那么对于任何$f, g \in L^\infty(X, \mu)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \|\mathbb{E}(fT^n g|\mathcal{Y}) - \mathbb{E}(f|\mathcal{Y})T^n\mathbb{E}(g|\mathcal{Y})\|_2 = 0 \tag{9.7.7}$$

等价的, 对于任意$\varepsilon > 0$ 存在$N_0 \in \mathbb{N}$ 使得对于$N \geq N_0$,

$$|\mathbb{E}(fT^n g|\mathcal{Y}) - \mathbb{E}(f|\mathcal{Y})T^n\mathbb{E}(g|\mathcal{Y})| < \varepsilon, \forall (n, y) \in \big([0, N-1] \times Y\big) \setminus G, \tag{9.7.8}$$

其中$m \times \nu(G) < \varepsilon$, m 为$[0, N-1]$ 上计数测度的规范化概率测度.

证明. (9.7.8) 蕴含(9.7.9) 是容易给出的, (9.7.9) 蕴含(9.7.8) 可以根据Chebyshev 不等式给出. Chebyshev 不等式是指在空间$(Z, \mathcal{Z}, \eta)$ 上$F \in L^1(Z, \eta)$ 有

$$\eta(\{z \in Z : F(z) > \varepsilon\}) < \frac{\|F\|_1}{\varepsilon}.$$

对于$[0, N-1] \times X$ 上概率测度对于函数

$$F(n, x) = |\mathbb{E}(fT^n g|\mathcal{Y}) - \mathbb{E}(f|\mathcal{Y})T^n\mathbb{E}(g|\mathcal{Y})|^2$$

运用Chebyshev 不等式就得到(9.7.8).

下面我们证明(9.7.9). 因为

$$f = \mathbb{E}(f|\mathcal{Y}) + (f - \mathbb{E}(f|\mathcal{Y})),$$

我们我们分$f \in L^\infty(X, \mathcal{Y})$ 和$\mathbb{E}(f|\mathcal{Y}) = 0$ 两种情况证明即可.

如果$f \in L^\infty(X, \mathcal{Y})$, 那么

$$\mathbb{E}(fT^n g|\mathcal{Y}) = f\mathbb{E}(T^n g|\mathcal{Y}) = \mathbb{E}(f|\mathcal{Y})T^n\mathbb{E}(g|\mathcal{Y}).$$

此时(9.7.9) 自然成立.

如果$\mathbb{E}(f|\mathcal{Y}) = 0$, 则

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mathbb{E}(fT^n g|\mathcal{Y})|^2 \, d\mu &= \lim_{N\to\infty} \int_{X\times X} f \otimes f \frac{1}{N} \sum_{n=0}^{N-1} (T \times T)^n \big(g \otimes g\big) d\mu \times_Y \mu \\
&= \int_{X\times X} f \otimes f \Big(\int_{X\times X} g \otimes g d\mu \times_Y \mu \Big) d\mu \times_Y \mu \\
&= \int_X \mathbb{E}(f|\mathcal{Y})^2 d\mu \int_X \mathbb{E}(g|\mathcal{Y})^2 d\mu = 0.
\end{aligned}$$

证毕! □

定理 9.7.8 设$\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ 为两个遍历系统间的弱混合扩充. 那么对于任何$k \in \mathbb{N}$, $f_1, f_2, \ldots, f_k \in L^\infty(X, \mu)$,

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} T^n f_1 T^{2n} f_2 \ldots T^{kn} f_k - \frac{1}{N} \sum_{n=0}^{N-1} T^n \mathbb{E}(f_1|\mathcal{Y}) T^{2n} \mathbb{E}(f_2|\mathcal{Y}) \ldots T^{kn} \mathbb{E}(f_k|\mathcal{Y}) \right\|_2 = 0. \tag{9.7.9}$$

证明. 我们对k 归纳证明. $k = 1$ 由遍历定理直接得到. 下面证明k 的情况. 我们需要用van der Corput 引理（定理9.3.1）.

根据引理9.5.3, 我们不妨设存在l 使得$\mathbb{E}(f_l|\mathcal{Y}) = 0$. 下面证明

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} T^n f_1 T^{2n} f_2 \ldots T^{kn} f_k \right\|_2 = 0.$$

设

$$u_n = T^n f_1 T^{2n} f_2 \ldots T^{kn} f_k, \ \forall n \in \mathbb{Z}_+.$$

则

$$\begin{aligned}
s_h &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \langle u_{n+h}, u_n \rangle \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X T^{n+h} f_1 T^{2n+2h} f_2 \ldots T^{kn+kh} f_k T^n f_1 T^{2n} f_2 \ldots T^{kn} f_k d\mu \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X (f_1 T^h f_1) T^n (f_2 T^{2h} f_2) T^{2n} (f_3 T^{3h} f_3) \ldots T^{(k-1)n} (f_k T^{kh} f_k) d\mu
\end{aligned}$$

由归纳假设, 我们得到

$$\begin{aligned}
s_h &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X (f_1 T^h f_1) T^n (f_2 T^{2h} f_2) T^{2n} (f_3 T^{3h} f_3) \ldots T^{(k-1)n} (f_k T^{kh} f_k) d\mu \\
&= \int_X \mathbb{E}(f_1 T^h f_1|\mathcal{Y}) \mathbb{E}(f_2 T^{2h} f_2|\mathcal{Y}) \ldots \mathbb{E}(f_k T^{kh} f_k|Y) d\mu \\
&\leq \|\mathbb{E}(f_l T^{lh} f_l|\mathcal{Y})\|_2 \prod_{i\neq l} \|f_i\|_\infty^2.
\end{aligned}$$

于是

$$\begin{aligned}
\lim_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} s_h &\leq \frac{1}{H} \sum_{h=0}^{H-1} \|\mathbb{E}(f_l T^{lh} f_l|\mathcal{Y})\|_2 \prod_{i\neq l} \|f_i\|_\infty^2 \\
&= 0 \quad (\text{命题9.7.7})
\end{aligned}$$

根据van der Corput 引理(定理9.3.1), 我们得到

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} u_n \right\|_2 = 0.$$

于是根据归纳完成了证明. □

定理 9.7.9 设$\pi:(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ 为两个遍历系统间的弱混合扩充. 那么

$$\widetilde{\pi}:(X\times X,\mathcal{X}\times\mathcal{X},\mu\times_Y\mu,T\times T)\to(Y,\mathcal{Y},\nu,S)$$

也为弱混合扩充.

证明. 设

$$\widehat{\mu}=\mu\times_Y\mu,\ \widetilde{\mu}=\widehat{\mu}\times_Y\widetilde{\mu}.$$

设

$$F=f_1\otimes f_2\otimes f_3\otimes f_4,\quad G=g_1\otimes g_2\otimes g_3\otimes g_4,$$

其中$f_i,g_i\in L^\infty(X,\mu),1\le i\le 4$. 对于$\varepsilon>0$, 取充分大的$N$, 我们运用命题9.7.7,

$$\begin{aligned}
&\frac{1}{N}\sum_{n=0}^{N-1}\int_{X^4}F(T^{(4)})^nGd\widetilde{\mu}\\
&=\frac{1}{N}\sum_{n=0}^{N-1}\int_X\mathbb{E}(f_1T^ng_1|\mathcal{Y})\mathbb{E}(f_2T^ng_2|\mathcal{Y})\mathbb{E}(f_3T^ng_3|\mathcal{Y})\mathbb{E}(f_4T^ng_4|\mathcal{Y})d\mu\\
&\overset{O(\varepsilon)}{\approx}\frac{1}{N}\sum_{n=0}^{N-1}\int_X\mathbb{E}(f_1|\mathcal{Y})T^n\mathbb{E}(g_1|\mathcal{Y})\mathbb{E}(f_2|\mathcal{Y})T^n\mathbb{E}(g_2|\mathcal{Y})\mathbb{E}(f_3|\mathcal{Y})T^n\mathbb{E}(g_3|\mathcal{Y})\mathbb{E}(f_4|\mathcal{Y})T^n\mathbb{E}(g_4|\mathcal{Y})d\mu\\
&=\frac{1}{N}\sum_{n=0}^{N-1}\int_X\mathbb{E}(f_1|\mathcal{Y})\mathbb{E}(f_2|\mathcal{Y})\mathbb{E}(f_3|\mathcal{Y})\mathbb{E}(f_4|\mathcal{Y})T^n\mathbb{E}(g_1|\mathcal{Y})T^n\mathbb{E}(g_2|\mathcal{Y})T^n\mathbb{E}(g_3|\mathcal{Y})T^n\mathbb{E}(g_4|\mathcal{Y})d\mu\\
&\to\int_X\mathbb{E}(f_1|\mathcal{Y})\mathbb{E}(f_2|\mathcal{Y})\mathbb{E}(f_3|\mathcal{Y})\mathbb{E}(f_4|\mathcal{Y})d\mu\int_X\mathbb{E}(g_1|\mathcal{Y})\mathbb{E}(g_2|\mathcal{Y})\mathbb{E}(g_3|\mathcal{Y})\mathbb{E}(g_4|\mathcal{Y})d\mu\\
&=\int_{X^4}Fd\widetilde{\mu}\int_{X^4}Gd\widetilde{\mu}.
\end{aligned}$$

于是$(X^4,\widetilde{\mu},T^{(4)})$ 为遍历的. □

定理9.7.4证明. 我们根据k 来归纳证明下式：对于任何$f_0,f_1,\dots,f_k\in L^\infty(X,\mu)$, 有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X\left|\mathbb{E}(f_0T^nf_1\dots T^{kn}f_k|\mathcal{Y})-\mathbb{E}(f_0|\mathcal{Y})T^n\mathbb{E}(f_1|\mathcal{Y})\dots T^{kn}\mathbb{E}(f_k|\mathcal{Y})\right|^2d\mu=0.\quad(9.7.10)$$

类似于命题9.7.7 的证明, (9.7.10) 等价于：对于任何$\varepsilon>0$, 存在$N_0\in\mathbb{N}$ 使得对于任何$N>N_0$,

$$\left|\mathbb{E}(f_0T^nf_1\dots T^{kn}f_k|\mathcal{Y})-\mathbb{E}(f_0|\mathcal{Y})T^n\mathbb{E}(f_1|\mathcal{Y})\dots T^{kn}\mathbb{E}(f_k|\mathcal{Y})\right|<\varepsilon,\forall(n,y)\in\big([0,N-1]\times Y\big)\setminus G.\quad(9.7.11)$$

其中$m\times\nu(G)<\varepsilon$, m 为$[0,N-1]$ 上计数测度的规范化概率测度. 我们简记为

$$\mathbb{E}(f_0T^nf_1\dots T^{kn}f_k|\mathcal{Y})\overset{\varepsilon}{\approx}\mathbb{E}(f_0|\mathcal{Y})T^n\mathbb{E}(f_1|\mathcal{Y})\dots T^{kn}\mathbb{E}(f_k|\mathcal{Y}).$$

现在假设$k-1$ 是命题成立, 下证k. 先设$f_k \in L^\infty(Y,\mathcal{Y})$, 此时

$$\mathbb{E}(f_0 T^n f_1 \dots T^{kn} f_k|\mathcal{Y}) = \mathbb{E}(f_0 T^n f_1 \dots T^{k-1n} f_{k-1}|\mathcal{Y}) T^{kn} f_k.$$

对于$f_0, f_1, \dots, f_{k-1}$, 根据(9.7.11), 我们得到充分大N 时

$$\begin{aligned}
&\mathbb{E}(f_0 T^n f_1 \dots T^{kn} f_k|\mathcal{Y}) \\
&= \mathbb{E}(f_0 T^n f_1 \dots T^{k-1n} f_{k-1}|\mathcal{Y}) T^{kn} f_k \\
&\overset{\varepsilon}{\approx} \mathbb{E}(f_0|\mathcal{Y}) T^n \mathbb{E}(f_1|\mathcal{Y}) \dots T^{k-1n} f_{k-1}|\mathcal{Y}) T^{kn} \mathbb{E}(f_k|\mathcal{Y}).
\end{aligned}$$

于是根据

$$f_k = \mathbb{E}(f_k|\mathcal{Y}) + (f_k - \mathbb{E}(f_k|\mathcal{Y}))$$

我们不妨假设$\mathbb{E}(f_k|\mathcal{Y}) = 0$.

根据定理9.7.9, $(X\times X, \mathcal{X}\times\mathcal{X}, \mu\times_Y \mu, T\times T) \to (Y,\mathcal{Y},\nu,S)$ 为弱混合扩充. 对于$f_1 \otimes f_1, \dots, f_k \otimes f_k$ 运用定理9.7.8,以及$\mathbb{E}(f_k\otimes f_k|\mathcal{Y}) = \mathbb{E}(f_k|\mathcal{Y}) \otimes \mathbb{E}(f_k|\mathcal{Y}) = 0$ 有

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} (T\times T)^n f_1 \otimes f_1 \cdot (T\times T)^{2n} f_2 \otimes f_2 \dots (T\times T)^{kn} f_k \otimes f_k \right\|_2 = 0. \tag{9.7.12}$$

根据$\mu\times_Y \mu$ 定义, 把(9.7.12) 内积上$f_0 \otimes f_0$ 后得到

$$\begin{aligned}
&\int_{X\times X} \left(\frac{1}{N} \sum_{n=0}^{N-1} f_0\otimes f_0 \cdot (T\times T)^n f_1 \otimes f_1 \cdot (T\times T)^{2n} f_2 \otimes f_2 \dots (T\times T)^{kn} f_k \otimes f_k \right) d\mu \underset{Y}{\times} \mu \\
&= \frac{1}{N} \sum_{n=0}^{N-1} \int_X \mathbb{E}(f_0 T^n f_1 \dots T^{kn} f_k|\mathcal{Y})^2 d\mu \\
&\to 0, \quad N\to\infty.
\end{aligned}$$

于是k 是命题成立, 根据归纳法我们完成了证明. □

§9.7.4 逆极限的情况

定理 9.7.10 设$(X,\mathcal{X},\mu,T)$ 为可逆保测系统, 设$\mathcal{Y}_1 \subseteq \mathcal{Y}_2 \subseteq \dots$ 为递增的因子列, 即均为$\mathcal{X}$ 的T 不变子σ代数. 设每个$\mathcal{Y}_n$ 都满足SZ 性质, 那么因子$\mathcal{Y} = \sigma(\bigcup_{n\geq 1} \mathcal{Y}_n)$ 也满足SZ 性质.

证明. 不妨设$\mathcal{X} = \mathcal{Y}$. 将$\mathcal{Y}_m$ 对应系统记为$(Y_m,\mathcal{Y}_m,\nu_m,T)$. 积分分解记为$\mu = \int_{Y_m} \mu_{m,y} d\nu_m(y)$. 设$\pi_m : X \to Y_m$ 为相应的因子映射. 设$A\in\mathcal{X}$, $k\in\mathbb{N}$. 令$\eta = \frac{1}{2(k+1)}$, 那么对于$\varepsilon = \frac{1}{4}\eta\mu(A) > 0$ 存在$m\in\mathbb{N}$ 以及$A_1 \in \mathcal{Y}_m$ 使得$\mu(A\Delta A_1) < \varepsilon$.

令

$$A_0 = \{x \in A_1 : \mu_{\pi_m(x)}(A) \geq 1-\eta\}.$$

则$A_0 \in \mathcal{Y}_m$. 我们断言$\mu(A_0) > \frac{1}{2}\mu(A)$. 然后我们将$\mathcal{Y}_m$ 的SZ 性质诱导到$\mathcal{Y}$ 上.

$$\begin{aligned}\varepsilon = \frac{1}{4}\eta\mu(A) &> \mu(A_1 \setminus A) \\ &= \int_{A_1} \mu_{m,y}(A_1 \setminus A) d\nu_m(y) \\ &\geq \int_{A_1 \setminus A_0} (1 - \mu_{m,y}(A)) d\nu_m(y) \\ &\geq \eta\mu(A_1 \setminus A_0),\end{aligned}$$

于是

$$\mu(A_1 \setminus A_0) < \frac{1}{4}\mu(A).$$

所以

$$\mu(A_0) = \mu(A_1) - \mu(A_1 \setminus A_0) > \frac{3}{4}\mu(A) - \frac{1}{4}\mu(A) = \frac{1}{2}\mu(A).$$

下面我们证明：

$$\mu(A \cap T^{-n}A \cap \ldots \cap T^{-kn}A) \geq \frac{1}{2}\mu(A_0 \cap T^{-n}A_0 \cap \ldots \cap T^{-kn}A_0) \tag{9.7.13}$$

设$y \in A_0 \cap T^{-n}A_0 \cap \ldots \cap T^{-kn}A_0$. 由$y \in A_0$

$$\nu_{m,y}(A) \geq 1 - \eta.$$

类似地, $y \in T^{-jn}A_0, j \in \{1, \ldots, k\}$, $\nu_{m,T_m^{jn}y}(A) \geq 1 - \eta$, 即

$$\nu_{m,y}(T^{-jn}A) \geq 1 - \eta.$$

于是

$$\nu_{m,y}(A \cap T^{-n}A \cap \ldots \cap T^{-kn}A) \geq 1 - (k+1)\eta = \frac{1}{2}.$$

由此得到(9.7.13).

根据$\mathcal{Y}_n$ 满足SZ, 所以根据(9.7.13), 我们有

$$\begin{aligned}&\liminf_{N\to\infty} \frac{1}{N}\mu(A \cap T^{-n}A \cap \ldots \cap T^{-kn}A) \\ \geq &\frac{1}{2}\liminf_{N\to\infty} \frac{1}{N}\mu(A_0 \cap T^{-n}A_0 \cap \ldots \cap T^{-kn}A_0) > 0.\end{aligned}$$

证毕! □

习 题

1. 验证式子(9.7.3)

§9.8 Host-Kra定理以及陶哲轩定理

在本节中, 我们给出Austin关于陶哲轩定理的证明.

§9.8.1　Host-Kra定理以及陶哲轩定理的陈述

关于多重遍历定理, 一个重要的突破是Host和Kra的工作. 他们证明了

定理 9.8.1 (Host-Kra) *如果$(X,\mathcal{X},\mu,T)$ 为保测系统, $d \geq 1$ 为整数以及$f_1, f_2, \cdots, f_d \in L^\infty(X,\mu)$, 那么多重遍历平均*

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x) f_2(T^{2n} x) \cdots f_d(T^{dn} x) \tag{9.8.1}$$

在$L^2(\mu)$中收敛.

另外Ziegler 独立的给出了这个定理的不同的证明[202]. 之后陶哲轩[188]证明了

定理 9.8.2 (陶哲轩) *设$T_1, \ldots, T_d$ 为概率空间$(X,\mathcal{X},\mu)$ 上可交换的保测变换并且$f_1, f_2, \cdots, f_d \in L^\infty(X,\mu)$. 那么遍历平均*

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n x) f_2(T_2^n x) \cdots f_d(T_d^n x) \tag{9.8.2}$$

在$L^2(\mu)$ 中收敛.

陶哲轩的证明采用有限遍历理论(finitary ergodic theory), 并没有运用经典遍历理论中的知识. 下面我们给出Austin 的证明.

§9.8.2　幂等族以及饱足系统

设Γ 为可数离散半群（或群）, 下面我们主要指$\Gamma = \mathbb{Z}_+^d, \mathbb{Z}^d$.

定义 9.8.3 (幂等系统) *设$\mathcal{C}$ 为一族Lebesgue空间上Γ-保测系统. 称$\mathcal{C}$ 为***幂等的***(idempotent), 是指它对于同构、可数交下封闭.*

*一个幂等族$\mathcal{C}$ 称为***可遗传的***(hereditary), 是指它还在取因子下封闭.*

*$\mathcal{C}$ 中系统我们称为***$\mathcal{C}$-系统***.*

注记 9.8.4 *1. 在Austin 的原始定义中, 族$\mathcal{C}$ 为幂等的还要求$\mathcal{C}$ 对于逆极限封闭, 因为逆极限也是一种交, 所以我们这里不再加入此条件.*

2. 全体零熵系统, 离散谱系统都是可遗传的幂等族. 存在不是可遗传的幂等族[12].

引理 9.8.5 *设$\mathcal{C}$ 为一个Γ 作用下动力系统幂等族. 那么对于任何Γ- 保测系统存在唯一的最大$\mathcal{C}$因子, 其中$\mathcal{C}$因子指在$\mathcal{C}$ 中的因子.*

证明. 设$(X,\mathcal{X},\mu,\Gamma)$ 为动力系统. 设

$$\mathcal{F} = \{\mathcal{A} : \mathcal{A} \text{ 为}\mathcal{X} \text{ 不变子}\sigma\text{代数, 且}\mathcal{A} \in \mathcal{C}\}.$$

因为$\mathcal{N} = \{X, \emptyset\} \in \mathcal{F}$, 所以$\mathcal{F} \neq \emptyset$. 容易验证满足Zorn 引理的条件, 所以$\mathcal{F}$ 存在最大元. □

定义 9.8.6 (最大$\mathcal{C}$-因子) 设$\mathbf{X}=(X,\mathcal{X},\mu,\Gamma)$ 为保测系统, 我们记它的最大$\mathcal{C}$-因子为$\mathbf{X}_\mathcal{C}=(X_\mathcal{C},\mathcal{X}_\mathcal{C},\mu_\mathcal{C},\Gamma)$. 我们将因子映射记为

$$\zeta^X_\mathcal{C}:\mathbf{X}\to\mathbf{X}_\mathcal{C}.$$

设$\pi:\mathbf{X}=(X,\mathcal{X},\mu,\Gamma)\to\mathbf{Y}=(Y,\mathcal{Y},\nu,\Gamma)$ 为因子映射. 那么我们自然会得到因子映射

$$C_\pi:\mathbf{X}_\mathcal{C}\to\mathbf{Y}_\mathcal{C}.$$

即, 有如下交换图表:

$$\begin{tikzcd} & \mathbf{X} \arrow[dl, "\zeta^X_\mathcal{C}"'] \arrow[dr, "\pi"] & \\ \mathbf{X}_\mathcal{C} \arrow[dr, "C_\pi"'] & & \mathbf{Y} \arrow[dl, "\zeta^Y_\mathcal{C}"] \\ & \mathbf{Y}_\mathcal{C} & \end{tikzcd}$$

定义 9.8.7 设$\mathcal{C}_1,\mathcal{C}_2$ 为两个幂等的族, 我们记$\mathcal{C}_1\bigvee\mathcal{C}_2$ 为全体由$\mathcal{C}_1$-系统和$\mathcal{C}_2$-系统交系统组成的集合. 易见$\mathcal{C}_1\bigvee\mathcal{C}_2$ 仍为幂等的. 我们称$\mathcal{C}_1\bigvee\mathcal{C}_2$ 为$\mathcal{C}_1$ 和$\mathcal{C}_2$ 的**交***(join)*.

定义 9.8.8 (饱足系统) 取定一个幂等的族$\mathcal{C}$. 称一个系统$\mathbf{X}=(X,\mathcal{X},\mu,\Gamma)$ 为$\mathcal{C}$-**饱足的***(sated)*, 是指对于$\mathbf{X}$ 的任何扩充$\widetilde{\mathbf{X}}=(\widetilde{X},\widetilde{\mathcal{X}},\widetilde{\mu},\Gamma)$, $\widetilde{\mathbf{X}}_\mathcal{C}$ 与$\mathbf{X}$ 为相对于$\mathbf{X}_\mathcal{C}$ 相对独立的. 即在$\widetilde{\mu}$ 下, $\widetilde{\mathbf{X}}_\mathcal{C}$ 和$\mathbf{X}$ 相对于它们的公因子$\mathbf{X}_\mathcal{C}$ 为相对独立的.

$$\begin{tikzcd} & \widetilde{\mathbf{X}} \arrow[dl, "\zeta^{\widetilde{X}}_\mathcal{C}"'] \arrow[dr, "\pi"] & \\ \widetilde{\mathbf{X}}_\mathcal{C} \arrow[dr, "C_\pi"'] & & \mathbf{X} \arrow[dl, "\zeta^Y_\mathcal{C}"] \\ & \mathbf{X}_\mathcal{C} & \end{tikzcd}$$

回顾逆极限的定义. 设$\{\mathbf{X}_i=(X_i,\mathcal{X}_i,\mu_i,\Gamma)\}_{i=0}^\infty$ 为保测系统族, 对于任何$i\geq i$ 存在因子映射$\psi^i_j:\mathbf{X}_i\to\mathbf{X}_j$ 使得$\psi^i_i=\mathrm{id}$ 并且

$$\psi^i_j\psi^j_k=\psi^i_k,\ \forall i\geq j\geq k.$$

$$\begin{tikzcd} X_i \arrow[dr, "\psi^i_j"] \arrow[dd, "\psi^i_k"'] & \\ & X_j \arrow[dl, "\psi^j_k"] \\ X_k & \end{tikzcd}$$

定义$\prod_{i=0}^{\infty} X_i$ 的子集X 为

$$X=\left\{x=(x_i)_{i=0}^{\infty}\in\prod_{i=0}^{\infty}X_i:\psi_j^i x_i=x_j,\ \forall i\geq j\right\}.$$

设$\psi_i:X\to X_i,\psi_i(x)=x_i$ 为投射. 易见它满足$\psi_j^i\psi_i=\psi_j,\forall i\geq j$.

$$\begin{tikzcd}
& X \arrow[dl, "\psi_i"'] \arrow[dr, "\psi_j"] & \\
X_i \arrow[rr, "\psi_j^i"] & & X_j
\end{tikzcd}$$

设$\mathcal{X}$ 为由$\{\psi_i^{-1}\mathcal{X}_i\}_{i=0}^{\infty}$ 生成的σ-代数. 定义$\mu:\bigcup_{i=0}^{\infty}\psi_i^{-1}(\mathcal{X}_i)\to\mathbb{R}_+$ 满足$\mu(\psi_i^{-1}E)=\mu_i(E),\forall E\in\mathcal{X}_i$. 再将$\mu$ 延拓到$\mathcal{X}$ 上. 对于$\gamma\in\Gamma$, 定义

$$\gamma:X\to X,\quad \gamma(x_i)_i=(\gamma x_i)_i.$$

明显的,

$$\psi_i:(X,\mathcal{X},\mu,\Gamma)\to(X_i,\mathcal{X}_i,\mu_i,\Gamma)$$

为因子映射. $\mathbf{X}=(X,\mathcal{X},\mu,T)$ 称为$\{(X_i,\mathcal{X}_i,\mu_i,T_i)\}_{i=0}^{\infty}$ 的逆极限系统，记为

$$(X,\mathcal{X},\mu,\Gamma)=\varprojlim_{i\to\infty}(X_i,\mathcal{X}_i,\mu_i,\Gamma)$$

定理 9.8.9 (Austin) *设$\{\mathcal{C}_i\}_{i\in I}$ 为至多可数个幂等族, 那么对于任何系统$\mathbf{X}_0=(X_0,\mathcal{X}_0,\mu_0,\Gamma)$, 存在一个系统$\mathbf{X}=(X,\mathcal{X},\mu,\Gamma)$ 为它的扩充, $\pi:X\to X_0$, 它满足:*

1. *$\mathbf{X}$ 为$\mathcal{C}_i$-饱足的, $\forall i\in I$;*
2. *$\mathcal{X}_0$ 以及$\mathcal{X}_{\mathcal{C}_i}$ 生成了$\mathcal{X}$.*

证明. 首先我们证明I 为一个元素的情况. 设$\mathcal{C}$ 为幂等族. 设系统$\mathbf{X}_0=(X_0,\mathcal{X}_0,\mu_0,\Gamma)$ 为保测系统. 设$\{f_r\}_{r\geq 1}$ 为在$\{f\in L^\infty(X_0,\mu_0):\|f\|_\infty\leq 1\}$ 中稠密（在$L^2(X_0,\mu_0)$）的可数子集. 设$\{r_i\}_{i\in\mathbb{N}}\in\mathbb{N}^{\mathbb{N}}$ 使得每个数都出现无穷次.

我们从X_0 出发, 构造一个逆极限族$\{\mathbf{X}_m=(X_m,\mathcal{X}_m,\mu_m,\Gamma)\}_{m\geq 0},\{\psi_k^m\}_{m\geq k\geq 0}$, 使得每个$\mathbf{X}_{m+1}$ 为$\mathbf{X}_m$ 的$\mathcal{C}$-交（$\mathcal{C}$- 交是指与某个$\mathcal{C}$-系统的交）.

假设对于某个$m_1\geq 0$, 我们已经取好了系统$\{\mathbf{X}_m\}_{m_1\geq m\geq 0},\{\psi_k^m\}_{m_1\geq m\geq k\geq 0}$ 使得

$$\mathcal{X}_{m_1}=(\mathcal{X}_{m_1})_{\mathcal{C}}\vee\mathcal{X}_0,$$

其中$(\mathcal{X}_{m_1})_{\mathcal{C}}$ 为$\mathbf{X}_{m_1}$ 的最大$\mathcal{C}$-因子. 我们分两种情况考虑:

（1）存在扩充$\pi:\widetilde{\mathbf{X}}\to\mathbf{X}_{m_1}$ 使得

$$\|\mathbb{E}_{\widetilde{\mu}}(f_{r_{m_1}}|\widetilde{X}_{\mathcal{C}})\|_2^2>\|\mathbb{E}_{\mu_{m_1}}(f_{r_{m_1}}|(X_{m_1})_{\mathcal{C}})\|_2^2+\frac{1}{2^{m_1}},$$

即

$$a_{m_1} = \sup\{\|\mathbb{E}_{\widetilde{\mu}}(f_{r_{m_1}}|\widetilde{X}_{\mathcal{C}})\|_2^2 - \|\mathbb{E}_{\mu_{m_1}}(f_{r_{m_1}}|(X_{m_1})_{\mathcal{C}})\|_2^2 : \pi : \widetilde{\mathbf{X}} \to \mathbf{X}_{m_1}\} > \frac{1}{2^{m_1}}.$$

取$\pi : \widetilde{\mathbf{X}} \to \mathbf{X}_{m_1}$ 使得

$$\|\mathbb{E}_{\widetilde{\mu}}(f_{r_{m_1}}|\widetilde{X}_{\mathcal{C}})\|_2^2 - \|\mathbb{E}_{\mu_{m_1}}(f_{r_{m_1}}|(X_{m_1})_{\mathcal{C}})\|_2^2 > \frac{a_{m_1}}{2},$$

并且$\widetilde{\mathbf{X}}$ 为$\mathbf{X}_{m_1}$ 的$\mathcal{C}$-交, 以及$\widetilde{\mathcal{X}} = \mathcal{X}_{m_1} \vee \widetilde{\mathcal{X}}_{\mathcal{C}}$ 令

$$\mathbf{X}_{m_1+1} = \widetilde{\mathbf{X}}, \quad \psi_{m_1}^{m_1+1} = \pi.$$

（2）对于任何扩充$\pi : \widetilde{\mathbf{X}} \to \mathbf{X}_{m_1}$ 都有

$$\|\mathbb{E}_{\widetilde{\mu}}(f_{r_{m_1}}|\widetilde{X}_{\mathcal{C}})\|_2^2 \le \|\mathbb{E}_{\mu_{m_1}}(f_{r_{m_1}}|(X_{m_1})_{\mathcal{C}})\|_2^2 + \frac{1}{2^{m_1}},$$

即

$$a_{m_1} = \sup\{\|\mathbb{E}_{\widetilde{\mu}}(f_{r_{m_1}}|\widetilde{X}_{\mathcal{C}})\|_2^2 - \|\mathbb{E}_{\mu_{m_1}}(f_{r_{m_1}}|(X_{m_1})_{\mathcal{C}})\|_2^2 : \pi : \widetilde{\mathbf{X}} \to \mathbf{X}_{m_1}\} \le \frac{1}{2^{m_1}}.$$

令

$$\mathbf{X}_{m_1+1} = \mathbf{X}_{m_1}, \quad \psi_{m_1}^{m_1+1} = \mathrm{id}_{X_{m_1}}.$$

最后令

$$\mathbf{X}_\infty = \overleftarrow{\lim_{i\to\infty}} \mathbf{X}_m.$$

则

$$\mathcal{X}_\infty = \bigvee_{m=0}^{\infty} (\mathcal{X}_m)_{\mathcal{C}} \vee \mathcal{X}_0.$$

于是$\mathbf{X}_\infty$ 仍为$\mathbf{X}_0$ 的$\mathcal{C}$-交. 下证明$\mathbf{X}_\infty$ 为饱足的.

设$\pi : \widetilde{\mathbf{X}} \to \mathbf{X}_\infty$ 为扩充, $f \in L^\infty(X_\infty, \mu_\infty)$. 下面我们证明

$$\mathbb{E}_{\widetilde{\mu}}(f|\widetilde{\mathbf{X}}_{\mathcal{C}}) = \mathbb{E}_{\mu_\infty}(f|(\mathbf{X}_\infty)_{\mathcal{C}}),$$

从而完成证明.

因为$\mathbf{X}_\infty$ 为$\mathbf{X}_0$ 的$\mathcal{C}$-交, 所以在$L^2(X_\infty, \mu_\infty)$ 中f 由形如

$$\sum_i g_i h_i, g_i \in L^\infty((X_\infty)_{\mathcal{C}}, \mu_\infty), h_i \in L^\infty(X_0, \mu_0)$$

的函数逼近. 因为$\{f_r\}_{r\ge1}$ 为在$\{f \in L^\infty(X_0, \mu_0) : \|f\|_\infty \le 1\}$ 中稠密（在$L^2(X_0, \mu_0)$）的可数子集, 所以我们仅需对于$f = gf_r, g \in L^\infty((X_\infty)_{\mathcal{C}}, \mu_\infty)$ 验证

$$\mathbb{E}_{\widetilde{\mu}}(f|\widetilde{\mathbf{X}}_{\mathcal{C}}) = \mathbb{E}_{\mu_\infty}(f|(\mathbf{X}_\infty)_{\mathcal{C}}).$$

因为$g \in L^\infty((X_\infty)_{\mathcal{C}}, \mu_\infty)$, 所以仅需验证

$$\mathbb{E}_{\widetilde{\mu}}(f_r|\widetilde{\mathbf{X}}_{\mathcal{C}}) = \mathbb{E}_{\mu_\infty}(f_r|(\mathbf{X}_\infty)_{\mathcal{C}}), \forall r.$$

根据鞅定理,

$$\mathbb{E}_{\mu_m}(f_r|(\mathbf{X_m})_{\mathcal{C}}) \xrightarrow{m\to\infty} \mathbb{E}_{\mu_\infty}(f_r|(\mathbf{X}_\infty)_{\mathcal{C}}).$$

于是如果

$$\mathbb{E}_{\widetilde{\mu}}(f_r|\widetilde{\mathbf{X}}_{\mathcal{C}}) > \mathbb{E}_{\mu_\infty}(f_r|(\mathbf{X}_\infty)_{\mathcal{C}}),$$

那么对于充分大m, 我们有$r_m = r$ (因为在$\{r_i\}_{i\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ 中每个数都出现无穷次), 于是

$$\begin{aligned}
&\|\mathbb{E}_{\mu_{m+1}}(f_r|(\mathbf{X}_{m+1})_{\mathcal{C}})\|_2^2 - \|\mathbb{E}_{\mu_m}(f_r|(\mathbf{X}_m)_{\mathcal{C}})\|_2^2 \\
\le\ &\|\mathbb{E}_{\mu_\infty}(f_r|(\mathbf{X}_\infty)_{\mathcal{C}})\|_2^2 - \|\mathbb{E}_{\mu_m}(f_r|(\mathbf{X}_m)_{\mathcal{C}})\|_2^2 \\
<\ &\frac{1}{2}\left(\|\mathbb{E}_{\widetilde{\mu}}(f_r|\widetilde{\mathbf{X}}_{\mathcal{C}})\|_2^2 - \|\mathbb{E}_{\mu_m}(f_r|(\mathbf{X}_m)_{\mathcal{C}})\|_2^2\right)
\end{aligned}$$

以及

$$\|\mathbb{E}_{\widetilde{\mu}}(f_r|\widetilde{\mathbf{X}}_{\mathcal{C}})\|_2^2 \ge \|\mathbb{E}_{\mu_m}(f_r|(\mathbf{X}_m)_{\mathcal{C}})\|_2^2 + \frac{1}{2^m}.$$

这与$\mathbf{X}_{m+1} \to \mathbf{X}_m$ 的定义矛盾. 所以我们有

$$\mathbb{E}_{\widetilde{\mu}}(f_r|\widetilde{\mathbf{X}}_{\mathcal{C}}) = \mathbb{E}_{\mu_\infty}(f_r|(\mathbf{X}_\infty)_{\mathcal{C}}).$$

下面我们证明一般情况. 设$\{\mathcal{C}_i\}_{i\in I}$ 为至多可数个幂等族. 取序列$\{i_m\}_{m\ge 1} \in I^{\mathbb{N}}$ 使得每个I 中元素出现无穷次. 我们从$\mathbf{X}_0$ 开始构造逆极限系统. 取$\{X_m\}_{m\ge 0}, \{\psi_k^m\}_{m\ge k\ge 0}$ 使得对于每个$m \ge 1$ 使得$\mathbf{X}_m$ 为$\mathcal{C}_{i_m}$-饱足的. 则逆极限系统满足我们所需条件. □

§9.8.3 定理9.8.2的证明

我们对d 进行归纳证明. $d = 1$ 就是von Neumann 平均遍历定理.

假设定理9.8.2对于$d-1$ 已经成立. 我们于是可以定义测度

$$\begin{aligned}
&\mu^F(A_1 \times A_2 \times \ldots \times A_d) \\
=\ &\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\int_X \mathbf{1}_{A_1}(T_1^n)\mathbf{1}_{A_2}(T_2^n)\ldots\mathbf{1}_{A_d}(T_d^n)d\mu(x) \\
=\ &\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\mu(T_1^{-n}A_1 \cap T_2^{-n}A_2 \cap \ldots \cap T_d^{-n}A_d),
\end{aligned}$$

其中$A_1,\ldots,A_d \in \mathcal{X}$. 称$\mu^F$ 为**Furstenberg 自交**(Furstenberg self-joining). 令

$$T^\Delta = T_1 \times T_2 \times \ldots \times T_d,$$

$$\widetilde{T}_i = T_i \times T_i \times \ldots \times T_i, \quad 1 \le i \le d.$$

X^d 上的对角测度为

$$\Delta_\mu^d(A_1 \times A_2 \times \ldots \times A_d) = \mu(A_1 \cap A_2 \cap \ldots \cap A_d),\ \forall A_1,\ldots,A_d \in \mathcal{X}.$$

那么我们有

$$\mu^F = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} (T^\Delta)_*^n \Delta_\mu^d.$$

根据定义我们得到：

- μ^F 是$\mathbf{X}$ 的d 阶自交.
- μ^F 为T^Δ 不变的.
- μ^F 为$\widetilde{T}_i$ 不变的, $i=1,2,\dots,d$.

根据定义, 我们也有对于任何$f_1,\dots,f_d \in L^\infty(X,\mu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1}\int_X T_1^n f_1 T_2^n f_2 \dots T_d^n f_d d\mu \xrightarrow{N\to\infty} \int_{X^d} f_1\otimes f_2 \otimes \dots \otimes f_d d\mu^F. \tag{9.8.3}$$

我们需要证明：对于任何$f_1,f_2,\cdots,f_d\in L^\infty(X,\mu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n x) f_2(T_2^n x)\cdots f_d(T_d^n x) \tag{9.8.4}$$

在$L^2(\mu)$中收敛.

我们需要定义一个幂等族$\mathcal{C}$. 我们定义$(X,\mathcal{X},\mu,T_1,\dots,T_d)\in\mathcal{C}$ 当且仅当

$$\mathcal{X} = \mathcal{I}(T_1)\vee\mathcal{I}(T_2T_1^{-1})\vee\dots\vee\mathcal{I}(T_dT_1^{-1}).$$

易验证$\mathcal{C}$ 为幂等的. 我们分几步进行证明.

第一步： 如果$\mathbf{X}=(X,\mathcal{X},\mu,T_1,\dots,T_d)\in\mathcal{C}$, 那么形如

$$\sum_n g_1g_2\dots g_d,\quad g_1\in L^\infty(X,\mathcal{I}(T_1)), g_2\in L^\infty(X,\mathcal{I}(T_2T_1^{-1})),\dots,g_d\in L^\infty(X,\mathcal{I}(T_dT_1^{-1}))$$

的函数在$L^2(X,\mathcal{X})$ 中稠密. 于是我们不妨设

$$f_1 = g_1g_2\dots g_d,$$

其中$g_1\in L^\infty(X,\mathcal{I}(T_1)), g_2\in L^\infty(X,\mathcal{I}(T_2T_1^{-1})),\dots,g_d\in L^\infty(X,\mathcal{I}(T_dT_1^{-1}))$. 于是

$$f_1(T_1^nx) = g_1(T_1^nx)g_2(T_1^nx)\dots g_d(T_1^nx) = g_1(x)g_2(T_2^nx)g_3(T_3^nx)\dots g_d(T_d^nx)$$

从而

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^nx)f_2(T_2^nx)\cdots f_d(T_d^nx)\\ =&\frac{1}{N}\sum_{n=0}^{N-1} g_1(x)g_2(T_2^nx)g_3(T_3^nx)\dots g_d(T_d^nx)f_2(T_2^nx)\cdots f_d(T_d^nx)\\ =&\ g_1(x)\frac{1}{N}\sum_{n=0}^{N-1}(g_2f_2)(T_2^nx)(g_3f_3)(T_3^nx)\dots(g_df_d)(T_d^nx).\end{aligned}$$

根据归纳假设, 上式在$L^2(X,\mu)$ 中收敛.

第二步: 设$\mathbf{X} = (X,\mathcal{X},\mu,T_1,\dots,T_d)$, $\mathbf{X}_{\mathcal{C}} = (X_{\mathcal{C}},\mathcal{X}_{\mathcal{C}},\mu_{\mathcal{C}},T_1,\dots,T_d)$ 为其最大的$\mathcal{C}$-因子. 如果$f_1 \in L^\infty(X_{\mathcal{C}},\mu_{\mathcal{C}})$, 那么根据第一步的证明方法, 我们仍有对于任何$f_2,f_3,\cdots,f_d \in L^\infty(X,\mu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n x)f_2(T_2^n x)\cdots f_d(T_d^n x)$$

在$L^2(\mu)$中收敛.

第三步: 设$\mathbf{X} = (X,\mathcal{X},\mu,T_1,\dots,T_d)$, $\mathbf{X}_{\mathcal{C}} = (X_{\mathcal{C}},\mathcal{X}_{\mathcal{C}},\mu_{\mathcal{C}},T_1,\dots,T_d)$ 为其最大的$\mathcal{C}$-因子. 根据定理9.8.9, 我们设$\mathbf{X}$ 为$\mathcal{C}$-饱足的.

设$f_1,f_2,\cdots,f_d \in L^\infty(X,\mu)$. 根据第二步, 我们不妨设

$$\mathbb{E}_\mu(f_1|X_{\mathcal{C}}) = 0.$$

下面我们运用von der Corput 引理9.3.1 来完成归纳证明. 设

$$u_n = T_1^n f_1 T_2^n f_2 \dots T_d^{dn} f_d,\ \forall n \in \mathbb{Z}_+.$$

我们需要证明如果$\mathbb{E}(f_1|X_{\mathcal{C}}) = 0$, 那么

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} u_n\right\|_2 = 0.$$

为此我们运用定理9.3.1.

$$\begin{aligned}
s_h &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h},u_n\rangle \\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X T_1^{n+h}f_1T_2^{n+h}f_2\dots T_d^{n+h}f_dT_1^nf_1T_2^nf_2\dots T_d^nf_d d\mu \\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X T_1^n(f_1T_1^hf_1)(T_2^nf_2T_2^hf_2)T_3^n(f_3T_3^hf_3)\dots T_d^n(f_dT_d^hf_d)d\mu \\
&= \int_{X^d}(f_1T_1^hf_1)\otimes(f_2T_2^hf_2)\otimes\dots\otimes(f_dT_d^hf_d)d\mu^F(x_1,\dots,x_d) \quad (\text{根据}(9.8.3)) \\
&= \int_{X^d}(f_1\otimes\dots\otimes f_d)(T^\Delta)^h(f_1\otimes\dots\otimes f_d)d\mu^F
\end{aligned}$$

于是

$$\begin{aligned}
\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}s_h &= \lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}\int_{X^d}(f_1\otimes\dots\otimes f_d)(T^\Delta)^h(f_1\otimes\dots\otimes f_d)d\mu^F \\
&= \lim_{H\to\infty}\int_{X^d}(f_1\otimes\dots\otimes f_d)\frac{1}{H}\sum_{h=0}^{H-1}(T^\Delta)^h(f_1\otimes\dots\otimes f_d)d\mu^F \\
&= \int_{X^d}(f_1\otimes\dots\otimes f_d)\mathbb{E}_{\mu^F}((f_1\otimes\dots\otimes f_d)|\mathcal{I}(T^\Delta))d\mu^F.
\end{aligned}$$

下面我们需要证明

$$\mathbb{E}_\mu(f_1|X_\mathcal{C})=0 \implies \mathbb{E}_{\mu^F}((f_1\otimes\ldots\otimes f_d)|\mathcal{I}(T^\Delta))=0.$$

令

$$\widetilde{\mathbf{X}}=(X^d,\mathcal{X}^d,\mu^F,T^\Delta,\widetilde{T_2},\widetilde{T_3},\ldots,\widetilde{T_d}).$$

那么到一个坐标投射的映射π 为因子映射, 即

$$\pi:\widetilde{\mathbf{X}}\to\mathbf{X},\quad (x_1,x_2,\ldots,x_d)\mapsto x_1$$

为因子映射.

注意对于$2\le i\le d$,

$$\widetilde{T}_i(T^\Delta)^{-1}=T_iT_1^{-1}\times T_iT_2^{-1}\times\ldots T_iT_{i-1}^{-1}\times \mathrm{id}\times T_iT_{i+1}^{-1}\times\ldots\times T_iT_d^{-1}$$

设$p_i:X^d\to X$ 为到第i 坐标的投射, 令

$$\mathcal{X}_i=p_i^{-1}(\mathcal{X}),2\le i\le d.$$

根据上面式子我们有

$$\mathcal{X}_2\vee\mathcal{X}_3\vee\ldots\vee\mathcal{X}_d\subseteq\widetilde{\mathcal{X}}_\mathcal{C}. \tag{9.8.5}$$

尤其

$$\mathbb{E}_{\mu^F}((f_1\otimes\ldots\otimes f_d)|\widetilde{X}_\mathcal{C})=f_2\otimes\ldots\otimes f_d\mathbb{E}_{\mu^F}(f_1|\widetilde{X}_\mathcal{C})$$

因为$\mathbf{X}$ 为饱足的, 根据定义在μ^F 下, $\widetilde{\mathbf{X}}_\mathcal{C}$ 和$\mathbf{X}$ 相对于它们的公因子$\mathbf{X}_\mathcal{C}$ 为相对独立的.

$$\begin{tikzcd}
 & \widetilde{\mathbf{X}} \arrow[dl, "\zeta_\mathcal{C}^{\widetilde{X}}"'] \arrow[dr, "\pi"] & \\
\widetilde{\mathbf{X}}_\mathcal{C} \arrow[dr, "C_\pi"'] & & \mathbf{X} \arrow[dl, "\zeta_\mathcal{C}^Y"] \\
 & \mathbf{X}_\mathcal{C} &
\end{tikzcd}$$

于是

$$\mathbb{E}_{\mu^F}(f_1(x_1)|\widetilde{X}_\mathcal{C})=\mathbb{E}_\mu(f_1|X_\mathcal{C})=0.$$

从而

$$\mathbb{E}_{\mu^F}((f_1\otimes\ldots\otimes f_d)|\widetilde{X}_\mathcal{C})=f_2\otimes\ldots\otimes f_d\mathbb{E}_{\mu^F}(f_1|\widetilde{X}_\mathcal{C})=0.$$

因为

$$\mathcal{I}(T^\Delta)\subseteq\widetilde{\mathcal{X}}_\mathcal{C},$$

所以

$$\mathbb{E}_{\mu^F}((f_1\otimes\ldots\otimes f_d)|\mathcal{I}(T^\Delta))=0.$$

即有

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}s_h=0.$$

根据定理9.3.1，我们得到

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}u_n\right\|_2=0.$$

于是根据归纳完成了证明. □

§9.9　多重遍历定理的一些新进展

§9.9.1　多重遍历平均

Ramsey理论中一个典型的例子是1927年证明的van der Waerden 定理：整数集任何有限染色必有单色包含任意长的等差数列[193]. 根据对此定理的理解，Erdös 和Turán 提出猜测：如果E 为具有正上密度的整数子集, 那么它必包含任意长的等差数列[52]. 这个猜测在1975年被Szemerédi 证明是正确的, 稍后Furstenberg 运用动力系统方法于1976 年给出了新的证明. 2006 年Green 和Tao 结合Furstenberg 证明的思想解决了一个历史悠久的数论猜测：素数集包含任意长的等差数列.[2] 这也是陶哲轩获2006 年Fields 奖的主要工作之一. 目前Erdös 关于这个方面最终的猜测仍没有解决：如果$A=\{a_n\}_{n=1}^\infty\subseteq\mathbb{N}$ 满足$\sum_{n=1}^\infty\dfrac{1}{a_n}=\infty$, 那么$A$ 包含任意长的等差数列.

如前所述, 为了证明Szemerédi 定理, Furstenberg 证明了：对于保测系统$(X,\mathcal{X},\mu,T)$, 如果$f\in L^\infty(X,\mathcal{X},\mu)$, $f\geq 0$ 且不等于0, 那么对于任何k,

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int f(x)f(T^nx)\ldots f(T^{(k-1)n}x)d\mu(x)>0. \tag{9.9.1}$$

由于$\mathbb{A}_N(f,x)=\frac{1}{N}\sum_{n=0}^{N-1}f(T^nx)$ 称为遍历平均, 所以Furstenberg 将异于上述平均的相关平均称为"非传统平均"(non-conventional averages), 而现在更多是称之为"多重遍历平均"(multiple ergodic averages). 亦即, 多重遍历平均是形如下式的平均：

$$\frac{1}{N}\sum_{n=0}^{N-1}f_1(T_1^{p_1(n)}x)f_2(T_2^{p_2(n)}x)\ldots f_d(T_d^{p_d(n)}x) \tag{9.9.2}$$

其中$T_1,T_2,\ldots,T_d$ 为概率空间$(X,\mathcal{X},\mu)$ 上的保测变换, $p_1,\ldots,p_d$ 整数值多项式.

多重遍历平均问题就是指：

[2]之后Tao 和Ziegler 又推广了这一工作. 他们证明了如果$P_1,\ldots,P_k$ 为取值整数的多项式并且$P_1(0)=\ldots=P_k(0)=0$, 那么对于任意的$\epsilon>0$ 存在无限多个整数x,m 使得$1\leq m\leq x^\epsilon$ 并且$x+P_1(m),\ldots,x+P_k(m)$ 均为素数.

问题 9.9.1 设$T_1, T_2, \ldots, T_d$ 为概率空间$(X, \mathcal{X}, \mu)$ 上可逆保测变换, 它们生成一个幂零作用群. 对于整数值多项式$p_i(n)$ 以及$f_i \in L^\infty(X, \mathcal{X}, \mu)$, $i = 1, 2, \ldots, d$, 多重遍历平均

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^{p_1(n)}x) f_2(T_2^{p_2(n)}x) \ldots f_d(T_2^{p_d(n)}x) \tag{9.9.3}$$

在$L^2(\mu)$ 中是否收敛？是否几乎处处收敛？

条件中$T_1, T_2, \ldots, T_d$ 生成的群为幂零群是必需的, Bergelson和Leibman证明了如果$T_1, \ldots, T_d$ 生成的群为可解群, 那么有反例说明上述平均不收敛[17].

问题的最简单形式是线性多项式的情况($a_1, \ldots, a_d \in \mathbb{Z} \setminus \{0\}$)

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{a_1 n}x) f_2(T^{a_2 n}x) \ldots f_d(T^{a_d dn}x).$$

对于保测系统$(X, \mathcal{X}, \mu, T)$ 的L^2 意义收敛, Furstenberg 自己证明了弱混合系统的情况以及两项的情况（$d = 2$）[63]. $d = 3$ 最先得到实质性突破是法国数学家Conze 和Lesigne 在1980年代的工作[40, 41, 42], 他们首先把幂零系统引入到研究中, 为后面的工作奠定了基础. 上面平均的最终解决是由Host 和Kra 于2005年给出的[100], Ziegler 稍后给出了另一个证明[202].

陶哲轩运用组合的方法证明了$T_1, T_2, \ldots, T_d$ 生成交换群的情形[188]：

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^n x) f_2(T_2^n x) \ldots f_d(T_2^n x)$$

在$L^2(\mu)$模下收敛, 其中$T_1, \ldots, T_d$ 相互交换.

Walsh 推广了陶哲轩的结论, 他证明了

定理 9.9.2 (Walsh) *[194]* 设G 为$(X, \mathcal{X}, \mu)$ 上保测变换生成的幂零群. 那么对于任何$T_1, \ldots, T_d \in G$, 多重遍历平均

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{l} (T_1^{p_{1,j}(n)} \ldots T_d^{p_{d,j}(n)}) f_j$$

在$L^2(X, \mathcal{X}, \mu)$ 中收敛, 其中$f_1, \ldots, f_d \in L^\infty(X, \mathcal{X}, \mu)$, $p_{i,j}$ 为整值多项式.

Walsh 的证明主要沿用了陶哲轩的方法和思想, 并不是动力系统传统方法的证明. 这个新的方法只能证明极限的存在性, 不能给出收敛函数的刻画, 使得其很难在组合数论中的得到应用. 一个重要的问题是找到一种方法去得到收敛函数的结构, 以便控制这个收敛, 使得它能运用到其它具体问题中去.

§9.9.2 多重遍历平均的特征因子

在多重遍历平均问题以及多重回复定理的研究中, 特征因子的想法起了非常重要的作用. 这个想法最早是由Furstenberg 在Szemerédi 定理新证明的工作中引入的[63], 但是术语"特征因子"是在1996年他与Weiss合作的文章中第一次提出[72]. 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $(Y,\mathcal{Y},\mu,T)$ 是它的因子, 设$\{p_1,\ldots,p_d\}$ 为一组整数值多项式, $d\in\mathbb{N}$. 我们称Y 为X 关于$\{p_1,\ldots,p_d\}$的**特征因子**(characteristic factor) , 是指对于$f_1,\ldots,f_d\in L^\infty(X,\mathcal{X},\mu)$,

$$\lim_{N\to\infty}\Big\|\frac{1}{N}\sum_{n=0}^{N-1}T^{p_1(n)}f_1T^{p_2(n)}f_2\ldots T^{p_d(n)}f_d$$
$$-\frac{1}{N}\sum_{n=0}^{N-1}T^{p_1(n)}\mathbb{E}(f_1|\mathcal{Y})T^{p_2(n)}\mathbb{E}(f_2|\mathcal{Y})\ldots T^{p_d(n)}\mathbb{E}(f_d|\mathcal{Y})\Big\|_{L^2}\to 0.$$

找到合适的特征因子将会将问题约化到结构更简单的系统上.

下面我们主要讨论对于线性多项式$\{n,2n,\ldots,dn\}$ 的特征因子, 亦即我们主要考虑如下多重遍历平均:

$$\frac{1}{N}\sum_{n=0}^{N-1}f_1(T^nx)f_2(T^{2n}x)\ldots f_d(T^{dn}x). \tag{9.9.4}$$

为研究(9.9.4), 根据遍历分解定理我们仅需研究遍历的情况, 下面如不特别指出我们的系统都是遍历的. Furstenberg [63] 证明了测度distal因子为(9.9.4) 特征因子, 然后多重回复定理就转化到distal因子上处理. 但是对于研究(9.9.4) 的收敛问题, distal因子太大, 不足以控制住收敛问题. 根据Zorn引理可以证明, 在因子关系下存在最小的特征因子Z, 即Z 为特征因子, 另外如Y 也为特征因子, 那么Z 为Y 的因子.

为介绍这个特征因子, 我们需要一些概念.

定义 9.9.3 (幂零系统) *设G 为群. 对于$A,B\subseteq G$, 我们记$[A,B]$ 为由$\{[a,b]=aba^{-1}b^{-1}:a\in A,b\in B\}$ 生成的子群. 交换子子群列G_j, $j\ge 1$, 如下归纳定义:*

$$G_1=G,\ G_{j+1}=[G_j,G].$$

*设$d\ge 1$ 为整数, 称G 为d-**阶幂零群** 如果G_{d+1} 为平凡群.*

*设G 为d-阶幂零李群, Γ 为其离散余紧子群. 紧致流形$X=G/\Gamma$ 称为d-**阶幂零流形**. 群G 用左平移方式作用于X 上, 记为$(g,x)\mapsto gx$. X 上的Haar 测度μ 指唯一的平移不变的概率测度. 取$\tau\in G$, 设T 为X 上作用$x\mapsto\tau x$. 则(X,μ,T) 称为 d-**阶幂零系统**.*

Conze 和Lesigne 证明了2-阶幂零系统的逆极限为3-项多重遍历平均的特征因子, 一般情况被Host-Kra证明也是成立的. 我们简单介绍下Host-Kra 的方法和思想, 为此我们还需要一些概念.

设X 为集合, $d\ge 1$ 为自然数. 我们视$\{0,1\}^d$ 为$0,1$组成的序列$\varepsilon=\varepsilon_1\ldots\varepsilon_d$, 设$|\varepsilon|=\varepsilon_1+\varepsilon_2+\ldots+\varepsilon_d$. 我们将$X^{2^d}$ 记为$X^{[d]}$. 点$\mathbf{x}\in X^{[d]}$ 表示为$\mathbf{x}=(x_\varepsilon:\varepsilon\in\{0,1\}^d)$. 点$\mathbf{x}\in X^{[d]}$ 可以分

解为$\mathbf{x}=(\mathbf{x}',\mathbf{x}'')$, $\mathbf{x}',\mathbf{x}''\in X^{[d-1]}$, 其中$\mathbf{x}'=(x_{\varepsilon 0}:\varepsilon\in\{0,1\}^{d-1})$ 以及$\mathbf{x}''=(x_{\varepsilon 1}:\varepsilon\in\{0,1\}^{d-1})$. 这种分解自然把$X^{[d]}$ 等同为$X^{[d-1]}\times X^{[d-1]}$. 例如$X^{[2]}$ 中点形如$(x_{00},x_{10},x_{01},x_{11})$.

对于保测系统$(X,\mathcal{X},\mu,T)$, 我们用$\mathcal{I}(T)$ 记不变子集σ-代数$\{A\in\mathcal{X}:T^{-1}A=A\}$. 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $k\in\mathbb{N}$. 我们定义$X^{[k]}$ 上$T^{[k]}=T\times T\times\ldots\times T$ (2^k次) 作用下不变的测度$\mu^{[k]}$ 如下: $\mu^{[1]}=\mu\times_{\mathcal{I}(T)}\mu=\mu\times\mu$; 对于$k\geq 1$,

$$\mu^{[k+1]}=\mu^{[k]}\underset{\mathcal{I}(T^{[k]})}{\times}\mu^{[k]}.$$

对于$k\geq 1$, 设(Ω_k,P_k) 为σ-代数$\mathcal{I}^{T^{[k]}}$ 对应的系统, 设

$$\mu^{[k]}=\int_{\Omega_k}\mu^{[k]}_{\omega}\ dP_k(\omega)$$

表示$\mu^{[k]}$ 在作用$T^{[k]}$ 下的遍历分解. 那么根据定义有

$$\mu^{[k+1]}=\int_{\Omega_k}\mu^{[k]}_{\omega}\times\mu^{[k]}_{\omega}\ dP_k(\omega).$$

如果(X,μ,T) 为弱混合的, 那么$\mathcal{I}(T^{[k]})$ 平凡的, $\mu^{[k]}$ 为2^k 个μ 的乘积测度μ^{2^k}, $k\geq 1$.

测度$\mu^{[k]}$ 为$T^{[k]}$ 不变的, 并且$\mu^{[k]}$ 到2^k 坐标的自然投射为μ. 设$C:\mathbb{C}\to\mathbb{C}$ 为共轭映射$z\mapsto\overline{z}$, 容易验证对于X 上有界函数f, 积分

$$\int_{X^{[k]}}\prod_{\varepsilon\in\{0,1\}^k}C^{|\varepsilon|}f(x_\varepsilon)d\mu^{[k]}(\mathbf{x})$$

为实的且非负. 所以我们可以定义$L^\infty(\mu)$ 上半模$|||\cdot|||_k$

$$|||f|||_k=\Big(\int_{X^{[k]}}\prod_{\varepsilon\in\{0,1\}^k}C^{|\varepsilon|}f(x_\varepsilon)d\mu^{[k]}(\mathbf{x})\Big)^{1/2^k}.$$

容易验证$|||\cdot|||_k$ 为半模, 它由Host和Kra引入, 所以称为**Host-Kra 半模**. 下面式子可以根据定义以及遍历定理得到, 下面式子也可以作为半模的等价定义. 对于每个$k\geq 0$ 以及$f\in L^\infty(\mu)$, 我们有

$$|||f|||_{k+1}=\Big(\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|||f\cdot T^n\overline{f}\ |||_k^{2^k}\Big)^{1/2^{k+1}}.$$

根据半模我们可以定义因子系统$(Z_{d-1},\mathcal{Z}_{d-1},\mu_{d-1},T)$.

定义 9.9.4 *设$(X,\mathcal{X},\mu,T)$ 为遍历系统. 设$d\in\mathbb{N}$, 存在$\mathcal{X}$的T-不变的子σ-代数$\mathcal{Z}_{d-1}$ 使得对于$f\in L^\infty(\mu)$,*

$$|||f|||_d=0\quad\text{当且仅当}\quad\mathbb{E}(f|\mathcal{Z}_{d-1})=0.$$

设$(Z_{d-1},\mathcal{Z}_{d-1},\mu_{d-1},T)$ 为子σ-代数$\mathcal{Z}_{d-1}$ 对应的X 的因子系统.

*如果$X=Z_{d-1}$, 那么称X 为***阶数**$d-1$ **系统**.

运用van der Corput 引理，Host 和Kra 证明了(9.9.4) 可以由半模控制住：

定理 9.9.5 *[100] 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $d\in\mathbb{N}$. 对于满足$\|f_1\|_\infty ,\ldots, \|f_d\|_\infty \le 1$ 的f_1, $\ldots$, $f_d \in L^\infty(X,\mu)$, 我们有*

$$\limsup_{N\to\infty}\Big\|\frac{1}{N}\sum_{n=0}^{N-1}T^nf_1T^{2n}f_2\ldots T^{dn}f_d\Big\|_{L^2}\le \min_{1\le j\le d}\{j \; |||\, f_j|||_d\}$$

这个定理表明Z_{d-1} 为(9.9.4) 的特征因子. Host 和Kra 工作中最为困难的部分是刻画$(Z_{d-1},\mathcal{Z}_{d-1},\mu_{d-1},T)$ 的结构, 他们证明了它是$d-1$- 阶幂零系统的逆极限.

我们强调出第一个坐标, 记$X_*^{[d]}=X^{2^d-1}$ 以及将点$\mathbf{x}\in X^{[d]}$ 记为$\mathbf{x}=(x_{\mathbf{0}},\mathbf{x}_*)$, 其中$\mathbf{x}_*=(x_\varepsilon:\varepsilon\ne\mathbf{0})\in X_*^{[d]}$ 以及$\mathbf{0}=00\ldots0\in\{0,1\}^d$.

定理 9.9.6 (Host-Kra) *[100] 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $d\in\mathbb{N}$. 那么以下等价:*

1. *X 为阶数$d-1$系统, i.e. $(X,\mathcal{X},\mu,T)=(Z_{d-1},\mathcal{Z}_{d-1},\mu_{d-1},T)$.*
2. *系统$(X,\mathcal{X},\mu,T)$ 为$d-1$-阶幂零系统的逆极限.*
3. *$|||\cdot|||_d$ 为$L^\infty(\mu)$ 上的模, 即, $|||f|||_d=0$ 蕴含$f=0$.*
4. *存在可测函数$J:X_*^{[d]}\to X$ 使得$x_{\mathbf{0}}=J(x_\varepsilon:\mathbf{0}\ne\varepsilon\in\{0,1\}^d)$ 对$\mu^{[d]}$ 几乎处处$\mathbf{x}=(x_\varepsilon:\varepsilon\in\{0,1\}^d)\in X^{[d]}$ 成立.*

因为由此定理, 我们也称阶数$d-1$系统为$d-1$**阶拟幂零系统**($(d-1)$-step pro-nilsystem).

Z_0 平凡系统, Z_1 为Kronecker 因子(即$\mathcal{Z}_1=\mathcal{K}_\mu$). 一般的, Z_k 为Z_{k-1} 的紧致交换群扩充. 于是得到一个新的结构定理：

$$\{pt\}=Z_0\longleftarrow Z_1\longleftarrow\cdots\longleftarrow Z_n\longleftarrow Z_{n+1}\longleftarrow\cdots\longleftarrow X.$$

注意, 如果T 弱混合的, 那么对于k 所有Z_k 为平凡的.

于是(9.9.4) 的收敛问题就化归到幂零系统上, 根据幂零因子的经典结果或者Ratner理论, 这是收敛的, 从而Host 和Kra 解决了(9.9.4) 的L^2 收敛问题. 关于详细的论述和证明参见文章[100], 或者最近的专著[102].

§9.9.3　拓扑动力系统的相应问题

在拓扑动力系统中也自然需要研究特征因子. 这个问题最早是由Glasner 于1994年开始的. 设(X,T) 为拓扑动力系统, $d\in\mathbb{N}$. 设

$$\sigma_d=T\times T^2\times\ldots\times T^d.$$

设$\pi:X\to Y$ 为因子映射. (Y,T) 称为 d**阶拓扑特征因子** 是指存在X 的稠密G_δ 集Ω 使得对于任何$x\in\Omega$ 轨道闭包$L=\overline{\mathcal{O}}(x^{(d)},\sigma_d)$ is $\pi\times\ldots\times\pi$ (d 次)**饱和的**(saturated), 其

中$x^{(d)}=(x,\dots,x)$ (d 次). 亦即$(x_1,x_2,\dots,x_d)\in L$ 当且仅当$(x_1',x_2',\dots,x_d')\in L$, 对于所有$1\le i\le d$, $\pi(x_i)=\pi(x_i')$. Glasner证明了：如果(X,T) 为极小distal系统, 那么它的最大$d-1$-阶distal因子是d-阶拓扑特征因子；如果(X,T) 为极小弱混合的, 那么平凡系统是它的任意阶拓扑特征因子[74].

第二种定义特征因子的方式是由Host, Kra 和Maass 给出的[103]. 他们在Host-Kra 2005 年工作基础上建立极小distal系统涉及幂零因子的结构定理, 下面我们给出更多细节.

设$\mathbf{n}=(n_1,\dots,n_d)\in\mathbb{Z}^d$ 以及$\varepsilon\in\{0,1\}^d$, 令$\mathbf{n}\cdot\varepsilon=\sum_{i=1}^d n_i\varepsilon_i$. 设$(X,T)$ 为拓扑动力系统, $d\ge 1$. 令$\mathbf{Q}^{[d]}(X)$ 为如下元素全体在$X^{[d]}$ 中的闭包：

$$(T^{\mathbf{n}\cdot\varepsilon}x=T^{n_1\varepsilon_1+\dots+n_d\varepsilon_d}x:\varepsilon=\varepsilon_1\varepsilon_2\dots\varepsilon_d\in\{0,1\}^d),$$

其中$\mathbf{n}=(n_1,\dots,n_d)\in\mathbb{Z}^d$ 以及$x\in X$. 如果不引起混淆, 用$\mathbf{Q}^{[d]}$ 记$\mathbf{Q}^{[d]}(X)$. $\mathbf{Q}^{[d]}(X)$ 中元素称为**动力系统d维平行体** (parallelepiped of dimension d). 例如$\mathbf{Q}^{[2]}$ 为$X^{[2]}=X^4$ 中集合$\{(x,T^mx,T^nx,T^{n+m}x):x\in X,m,n\in\mathbb{Z}\}$. 的闭包. $\mathbf{Q}^{[d]}$ 可以视为$\mu^{[d]}$ 的拓扑对应.

面变换(face transformations) $T_j^{[d]}:X^{[d]}\to X^{[d]},j=1,\dots,d$ 定义如下：对于$\mathbf{x}=(x_\varepsilon)_{\varepsilon\in\{0,1\}^d}\in X^{[d]}$

$$(T_j^{[d]}\mathbf{x})_\varepsilon=\begin{cases}Tx_\varepsilon, & 如\varepsilon_j=1;\\ x_\varepsilon, & 如\varepsilon_j=0.\end{cases}$$

也可如下归纳定义：设$T^{[0]}=T$, $T_1^{[1]}=\mathrm{id}\times T$. 如$\{T_j^{[d-1]}\}_{j=1}^{d-1}$ 已经定义好, 那么设

$$\begin{aligned}&T_j^{[d]}=T_j^{[d-1]}\times T_j^{[d-1]},\ j\in\{1,2,\dots,d-1\},\\&T_d^{[d]}=\mathrm{id}^{[d-1]}\times T^{[d-1]}.\end{aligned}$$

d维面变换群(face group of dimension d) $\mathcal{F}^{[d]}(X)$ 为由$X^{[d]}$ 上全体面变换生成的群. **d维方体群** 或者**d 维平行体群**(cube group or parallelepiped group of dimension d) $\mathcal{G}^{[d]}(X)$ 是指由对角变换$T^{[d]}$ 和$\mathcal{F}^{[d]}(X)$ 生成的群. 在不会混淆的情况下, 我们直接用$\mathcal{F}^{[d]}$ 和$\mathcal{G}^{[d]}$ 记$\mathcal{F}^{[d]}(X)$ 和$\mathcal{G}^{[d]}(X)$. 容易看出$\mathbf{Q}^{[d]}$ 为$\{Sx^{[d]}:S\in\mathcal{F}^{[d]},x\in X\}$ 在$X^{[d]}$ 中的闭包. 如果x 为X 传递点, 那么$\mathbf{Q}^{[d]}$ 为点$x^{[d]}$ 在$\mathcal{G}^{[d]}$ 作用下的轨道闭包, 其中$x^{[d]}=(x,x,\dots,x)\in X^{[d]}$.

定理 9.9.7 设(X,T) 为极小的$d\in\mathbb{N}$, 那么

- $(\mathbf{Q}^{[d]},\mathcal{G}^{[d]})$ 为极小的*[103]*.
- 对于所以$x\in X$, $(\overline{\mathcal{O}(x^{[d]},\mathcal{F}^{[d]})},\mathcal{F}^{[d]})$ 为极小的*[183]*.

定理9.9.7可以视为如下结论的拓扑对应：

定理 9.9.8 *[100]* 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $d\in\mathbb{N}$. 那么系统$(X^{[d]},\mu^{[d]},\mathcal{G}^{[d]})$ 为遍历的, 并且$(\Omega_d,P_d,\mathcal{F}^{[d]})$ 为遍历的.

对应于定理9.9.6, 下面给出了拓扑幂零系统逆极限的刻画：

定理 9.9.9 (Host-Kra-Maass) *[103] 设(X,T) 为极小系统, $d\geq 2$ 为自然数. 那么以下等价:*

1. *X 为$(d-1)$-阶极小幂零系统的拓扑逆极限.*
2. *如果$\mathbf{x},\mathbf{y}\in\mathbf{Q}^{[d]}$ 有2^d-1 个坐标吻合, 那么$\mathbf{x}=\mathbf{y}$.*
3. *如果$x,y\in X$ 满足$(x,y,\dots,y)\in\mathbf{Q}^{[d]}$, 那么$x=y$.*

*满足上面条件的极小系统称为***拓扑阶数***$d-1$***系统***(topological system of order $(d-1)$)* *或者***拓扑***$d-1$***阶拟幂零系统** *(topological $(d-1)$-step pro-nilsystem).*

注意任何d阶拟幂零系统测度同构于某个拓扑d 阶拟幂零系统[103]. 如果(X,T) 为拓扑d 阶拟幂零系统, 那么它的最大j阶拟幂零因子和j阶拓扑拟幂零因子吻合, $j\leq d$ [50].

定义 9.9.10 *[103] 设(X,T) 为拓扑动力系统, $d\in\mathbb{N}$, ρ 为度量. 点对$x,y\in X$ 称为d***阶局部渐进的***(regionally proximal of order d), 指对于任何$\delta>0$, 存在$x',y'\in X$ 以及$\mathbf{n}=(n_1,\dots,n_d)\in\mathbb{Z}^d$ 使得$\rho(x,x')<\delta,\rho(y,y')<\delta$, 且*

$$\rho(T^{\mathbf{n}\cdot\varepsilon}x',T^{\mathbf{n}\cdot\varepsilon}y')<\delta \text{ 对于任何}\varepsilon\in\{0,1\}^d\setminus\{\mathbf{0}\}.$$

*d阶局部渐进对全体记为$\mathbf{RP}^{[d]}$ 或$\mathbf{RP}^{[d]}(X,T)$. 称为d***阶渐进关系***(the regionally proximal relation of order d).*

Host-Kra-Maass 证明了, 对于极小distal系统(X,T), $\mathbf{RP}^{[d]}$ 为等价关系, 并且$(X/\mathbf{RP}^{[d]},T)$ 为(X,T) 的最大d阶拟幂零因子. 邵松和叶向东运用Ellis 半群技巧证明了他们的结果对于一般极小系统也成立[183].

定理 9.9.11 *[183, 103] 设(X,T) 为极小系统, $d\in\mathbb{N}$. 那么*

1. *$\mathbf{RP}^{[d]}(X)$ 为等价关系.*
2. *如果$\pi:(X,T)\to(Y,S)$ 为因子映射, 那么*
$$(\pi\times\pi)(\mathbf{RP}^{[d]}(X))=\mathbf{RP}^{[d]}(Y).$$
3. *$(X_d=X/\mathbf{RP}^{[d]},T)$ 为(X,T) 的最大d阶拟幂零因子.*

注意$\mathbf{RP}^{[1]}$ 为经典的局部渐进关系, 而$X_1=X_{eq}$ 为极大等度连续因子. 进一步的, 根据上面定理我们有如下结构定理:

$$\{pt\}=X_0\longleftarrow X_1\longleftarrow\cdots\longleftarrow X_n\longleftarrow X_{n+1}\longleftarrow\cdots\longleftarrow X.$$

如果(X,T) 为拓扑弱混合, 那么对于所有n, X_n为平凡的.

§9.9.4 拓扑方法在多重遍历平均问题研究中的应用

目前L^2中多重遍历平均问题得到丰富的成果, 但是几乎处处收敛意义下多重遍历平均问题结果甚少. 一个重要的结果如下：

定理 9.9.12 (Bourgain) *[34][35]* 设$(X,\mathcal{X},\mu,T)$ 为保测系统, 那么

- 对于任何$p(n)\in\mathbb{Z}[n]$ 以及$f\in L^p(X,\mathcal{X},\mu)$, $p>1$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f(T^{p(n)}x)$$

  几乎处处收敛.

- 对于$a\neq b\in\mathbb{Z}\setminus\{0\}$ 以及$f_1,f_2\in L^\infty(X,\mathcal{X},\mu)$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f_1(T^{a_1n}x)f_2(T^{a_2n}x)$$

  几乎处处收敛.

另外一个重要的结果是Derrien和Lesigne指出多重遍历平均几乎处处收敛与否取决于其Pinsker因子，即如果能证明零熵系统多重遍历平均几乎处处收敛，那么对于任何保测系统多重遍历平均几乎处处收敛[48].

设(X,T) 为拓扑动力系统, $d\in\mathbb{N}$. 设

$$\tau_d=T\times\ldots\times T\ (d\text{ 次})$$

以及

$$\sigma_d=T\times T^2\times\ldots\times T^d.$$

设$\langle\tau_d,\sigma_d\rangle$ 为由τ_d,σ_d 生成的群. 对于$x\in X$, 令

$$N_d(X,x)=\overline{\mathcal{O}((x,\ldots,x),\langle\tau_d,\sigma_d\rangle)},$$

它为$x^{(d)}$ 在群$\langle\tau_d,\sigma_d\rangle$作用下的轨道闭包. 如果$(X,T)$ 极小, 那么所以$N_d(X,x)$ 吻合, 此时直接记为$N_d(X)$.

定理 9.9.13 (Glasner) *[74]* 如果(X,T) 极小, 那么$(N_d(X),\langle\tau_d,\sigma_d\rangle)$ 极小.

于是如果$(N_d(X),\langle\tau_d,\sigma_d\rangle)$ 为唯一遍历的, 那么它为严格遍历的.

定理 9.9.14 (黄文-邵松-叶向东) *[112]* 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $d\in\mathbb{N}$. 那么它有严格遍历的模型$(\hat{X},\hat{T})$ 使得$(N_d(\hat{X}),\langle\tau_d,\sigma_d\rangle)$ 为严格遍历的.

因为唯一遍历系统具有很好的遍历平均收敛性质：对于(X,Γ), $\Gamma=\mathbb{Z}^d$, 那么(X,Γ) 唯一遍历的当且仅当对于任何连续函数$f\in C(X)$

$$\frac{1}{N^d}\sum_{\gamma\in[0,N-1]^d} f(\gamma x)$$

一致收敛到常值函数.作为定理9.9.14 的应用, 我们证明了:

定理 9.9.15 (黄文-邵松-叶向东) *[112] 设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $d\in\mathbb{N}$. 那么对于任何$f_1,\dots,f_d\in L^\infty(\mu)$, 遍历平均*

$$\frac{1}{N^2}\sum_{(n,m)\in[0,N-1]^2} f_1(T^nx)f_2(T^{n+m}x)\dots f_d(T^{n+(d-1)m}x)$$

μ-几乎处处收敛到常数.

为证明定理9.9.14, 我们发现并非所以唯一遍历模型是满足条件的, 为了我们的目的, Jewett-Krieger 定理不够, 但是Weiss 定理刚好是合适的工具.

对于$d\geq 3$, 令$\pi_{d-2}:X\to Z_{d-2}$为X 到其$d-2$-阶拟幂零因子Z_{d-2} 的因子映射. 将Z_{d-2}视为拓扑$d-2$-阶拟幂零因子, 运用Weiss 定理, 存在$(X,\mathcal{X},\mu,T)$唯一遍历模型$(\hat{X},\hat{\mathcal{X}},\hat{\mu},\hat{T})$以及因子映射$\hat{\pi}_{d-2}:\hat{X}\to Z_{d-2}$ 为$\pi_{d-2}:X\to Z_{d-2}$ 的模型. 可以证明$(\hat{X},\hat{T})$ 就是我们所需, 即$(N_d(\hat{X}),\langle\tau_d,\sigma_d\rangle)$ 为唯一遍历的.

结合拓扑模型理论, 我们证明了

定理 9.9.16 (黄文-邵松-叶向东) *[112] 设$(X,\mathcal{X},\mu,T)$ 为遍历distal 系统, $d\in\mathbb{N}$. 那么对于任何$f_1,\dots,f_d\in L^\infty(\mu)$*

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^nx)\dots f_d(T^{dn}x)$$

μ 几乎处处收敛.

根据Furstenberg-Zimmer结构定理以及定理9.9.16, 9.9.4是否逐点收敛问题就化归到处理弱混合扩充即可. 但是目前即使对于弱混合系统, 这个问题仍然没有解决. 一个部分的结果是由Assani 给出的[7], 他运用Host定理证明了如果$(X,\mathcal{X},\mu,T)$ 为具有奇异谱的弱混合系统, 那么(9.9.4) 几乎处处收敛.

最近我们对于一类重要的弱混合系统证明了(9.9.4) 几乎处处收敛. $(X_i,\mathcal{X}_i,\mu_i,T_i)$, $1\leq i\leq d$ 的交λ 称为**两两独立** 指如果它到任何两个坐标$X_i\times X_j$ 的投射为$\mu_i\times\mu_j$, $i\neq j\in\{1,2,\dots,d\}$. 系统$(X,\mathcal{X},\mu,T)$ 称为**两两独立确定的**(pairwise independently determined (PID)), 是指所以两两独立的d交($d\geq 3$) 是独立的, 即为乘积测度. PID 系统包括具有奇异谱的弱混合系统, [98]; 有限秩混合系统[178] 等.

运用拓扑方法, 可以证明

定理 9.9.17 *[90]* 设$(X,\mathcal{X},\mu,T)$ 弱混合*PID*系统, 那么对于任何$d\in\mathbb{N}$, $f_1,\ldots,f_d\in L^\infty(X,\mathcal{X},\mu)$, (9.9.4) 几乎处处收敛.

可以证明, 对于稠密G_δ 集的遍历系统(9.9.4) 几乎处处收敛. 这些给了多重遍历平均问题许多正面的信息.

对于方体群作用, Host-Kra证明了

定理 9.9.18 *[100]* 设$(X,\mathcal{X},\mu,T)$ 为保测系统, $d\in\mathbb{N}$. 那么对于$f_\varepsilon\in L^\infty(\mu)$, $\varepsilon\in\{0,1\}^d,\varepsilon\neq(0,\ldots,0)$, 平均

$$\prod_{i=1}^{d}\frac{1}{N_i-M_i}\cdot\sum_{\mathbf{n}\in[M_1,N_1)\times\ldots\times[M_d,N_d)}\ \prod_{(0,\ldots,0)\neq\varepsilon\in\{0,1\}^d}f_\varepsilon(T^{\mathbf{n}\cdot\varepsilon}x)\tag{9.9.5}$$

当$N_1-M_1,N_2-M_2,\ldots,N_d-M_d$ 趋于无穷时在$L^2(X)$ 中收敛.

设$(X,\mathcal{X},\mu,T)$ 为遍历系统, $d\in\mathbb{N}$. 运用Weiss 定理我们可以证明它存在唯一遍历模型$(\hat{X},\hat{T})$ 使得$(\mathbf{Q}^{[d]}(\hat{X}),\mathcal{G}^{[d]})$ 为严格遍历的, 从而由此可以证明(9.9.5) μ 几乎处处收敛[111]. 这个结果在依照$[0,N-1]^d$ 取平均的情形最早由Assani [8] 给出, 后来Chu 和Franzikinakis [38] 将之推广到更为一般的情况.

§9.10 注记

本章内容主要参考了著作[64, 54], 这是两本介绍动力系统在组合数论中应用的非常引人入胜的专著. 关于动力系统在组合数论中应用也可参见[15, 154]等. Szemerédi 定理的证明还可以参见[69, 89, 186, 187]. 陶哲轩定理的证明我们采用了Austin 的证明[11, 12], 另外不同的证明参见[99, 188].

在本章中Furstenberg-Zimmer 结构定理的证明我们主要沿用Furstenberg 的方法，Zimmer 的处理方法参见[203, 204].

现在幂零李群理论在动力系统中占据越来越重要的地位，关于幂零系统的基本知识可以参见[10, 44, 102, 144]等. 幂零李群在数学中应用的一个很好的例子是Green和陶哲轩等人的系列工作[85, 86, 87, 88]. 关于更多多重回复和多重遍历的信息参见[15, 16, 18, 19, 20, 66, 67, 68, 102, 135, 143].

参考文献

[1] Aaronson J. (1997). An introduction to infinite ergodic theory. Mathematical Surveys and Monographs, 50.

[2] Adler R. L., Konheim A. G. and McAndrew M. H. (1985). Topological entropy, Trans. Amer. Math. Soc., **114**: 309-319.

[3] Akin E. (1993). The general topology of dynamical systems. Graduate Studies in Mathematics, 1. American Mathematical Society, Providence, RI.

[4] Akin E. (1997). Recurrence in topological dynamics. Furstenberg families and Ellis actions. The University Series in Mathematics. Plenum Press, New York.

[5] Akin E., Glasner E., Huang W., Shao S. and Ye X.D. (2010). Sufficient conditions under which a transitive system is chaotic, Ergodic Theory and Dynamical Systems, **30**: 1277–1310.

[6] Aoki N. and Hiraide K. (1994). Topological theory of dynamical systems. Recent advances, North-Holland Mathematical Library, 52. North-Holland Publishing Co., Amsterdam.

[7] Assani I. (1998). Multiple recurrence and almost sure convergence for weakly mixing dynamical systems. Israel J. Math., **103**: 111–124.

[8] Assani I. (2010). Pointwise convergence of ergodic averages along cubes. J. Analyse Math., **110**: 241–269.

[9] Auslander J. (1988). Minimal flows and their extensions. North-Holland Mathematics Studies, **153**. Notas de Matemática [Mathematical Notes], 122. North-Holland Publishing Co., Amsterdam.

[10] Auslander L., Green L. and Hahn F. (1963). Flows on Homogeneous Spaces (Annals of Mathematical Studies, 53). Princeton University Press, Princeton, NJ.

[11] Austin T. (2010). On the norm convergence of non-conventional ergodic averages, Ergod. Th. and Dynam. Sys., **30**, 321–338.

[12] Austin T. (2010). Multiple recurrence and the structure of probability-preserving systems, PhD thesis, arXiv:1006.0491

[13] Bekka M. and Mayer M. (2000). Ergodic theory and topological dynamics of group actions on homogeneous spaces. London Mathematical Society Lecture Note Series, 269. Cambridge University Press, Cambridge.

[14] Bergelson V. (1987). Weakly mixing PET, Ergodic Theory Dynam. Systems, **7**, 337–349.

[15] Bergelson V. (2006). Combinatorial and Diophantine applications of ergodic theory. Appendix A by A. Leibman and Appendix B by Anthony Quas and Máté Wierdl, Handbook of dynamical systems. Vol. 1B, 745–869, Elsevier B. V., Amsterdam.

[16] Bergelson V., Host B. and Kra B. (2005). Multiple recurrence and nilsequences. With an appendix by Imre Ruzsa, Invent. Math. **160**(2): 261–303.

[17] Bergelson V. and Leibman S. (2002). A nilpotent Roth theorem, Invent. Math., **147**(2): 429–470.

[18] Bergelson V. and Leibman A. (1996). Polynomial extensions of van der Waerden's and Szemerédi's theorems, Journal of Amer. Math. Soc., **9**: 725–753.

[19] Bergelson V. and Leibman A. (1999). Set-polynomials and polynomial extension of the Hales–Jewett theorem, Ann. Math., **150**: 33–75.

[20] Bergelson V. and McCutcheon R. (2000). An ergodic IP polynomial Szemerédi theorem, Mem. Amer. Math. Soc., **146**: 695.

[21] Billingsley P. (1965). Ergodic theory and information. John Wiley & Sons, Inc., New York-London-Sydney.

[22] Birkhoff G. (1927). Dynamical systems, Colloquium publication IX. Ameican Mathematical Society, Providence, Rhode Island, 1927; also 1966.

[23] Birkhoff G. (1931). Proof of the ergodic theorem, Proc. Nat. Acad. Sci. USA, **17**: 656–660.

[24] Blanchard F. (1992). Fully positive topological entropy and topological mixing. Symbolic dynamics and its applications, AMS Contemp. Math., **135**:95–105.

[25] Blanchard F. (1993). A disjointness theorem involving topological entropy. Bull. Soc. Math. France, **121**: 465–478.

[26] Blanchard F., Glasner E. and Host B. (1997). A variation on the variational principle and applications to entropy pairs. Ergodic Theory Dynam. Systems, **17**: 29–43.

[27] Blanchard F, Glasner E, Kolyada S, Maass A, On Li-Yorke pairs, J. Reine Angew. Math. 547:51-68

[28] Blanchard F., Host B. and Maass A. (2000). Topological complexity. Ergod. Th. and Dynam. Sys., **20**: 641–662.

[29] Blanchard F., Host B., Maass A., Martinez S. and Rudolph D. J. (1995). Entropy pairs for a measure. Ergodic Theory Dynam. Systems, **15**(4): 621–632.

[30] Blanchard F., Host B. and Ruette S. (2002). Asymptotic pairs in positive-entropy systems, Ergod. Th. and Dynam. Sys., **22**: 671–686.

[31] Blanchard F. and Lacroix Y. (1993). Zero-entropy factors of topological flows. Proc. Amer. Math. Soc., **119**: 985–992.

[32] Bogachev V. I. (2007). Measure theory. Vol. I, II. Springer-Verlag, Berlin, Vol. I: xviii+500 pp., Vol. II: xiv+575 pp.

[33] Bourbaki N. (1966). General Topology, Hermann, Paris.

[34] Bourgain J. (1989). Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein, Inst. Hautes études Sci. Publ. Math., **69**: 5–45.

[35] Bourgain J. (1990). Double recurrence and almost sure convergence, J. Reine Angew. Math., **404**: 140–161.

[36] Bowen R. (1971). Entropy for group endomorphisms and homogeneous spaces, Trans. Amer. Math. Soc., **153**: 401–414.

[37] Buczolich Z. and Mauldin R.D. (2010). Divergent square averages, Ann. of Math. (2), **171**(3): 1479–1530.

[38] Chu Q. and Frantzikinakis N. (2012). Pointwise convergence for cubic and polynomial ergodic averages of non-commuting transformations. Ergod. Th. and Dynam. Sys. **32**: 877–897.

[39] Coronel A., Maass A. and Shao S. (2009). Sequence entropy and rigid $\sigma-$algebras, Studia Math., **194**: 207–230.

[40] Conze J. and Lesigne E. (1984). Théorèmes ergodiques pour des mesures diagonales. (French) [Ergodic theorems for diagonal measures], Bull. Soc. Math. France, **112**(2): 143–175.

[41] Conze J. and Lesigne E. (1988). Sur un théorème ergodique pour desmesures diagonales. (French) [On an ergodic theorem for diagonal measures], Probabilités, 1–31, Publ. Inst. Rech. Math. Rennes, 1987-1, Univ. Rennes I, Rennes.

[42] Conze J. and Lesigne E. (1988). Sur un théorème ergodique pour des mesures diagonales. (French) [On an ergodic theorem for diagonal measures], C. R. Acad. Sci. Paris Sér. I Math., **306**(12): 491–493.

[43] Cornfeld I., Fomin S. and Sinaǐ Y. (1982). Ergodic theory. Springer-Verlag, New York.

[44] Corwin L.J. and Greenleaf F.P. (1990). Representations of nilpotent Lie groups and their applications. Part I. Basic theory and examples, Cambridge Studies in Advanced Mathematics 18, Cambridge University Press, Cambridge, viii+269 pp.

[45] Dajani K. and Kraaikamp C. (2002). Ergodic theory of numbers, volume 29 of Carus Mathematical Monographs. Mathematical Association of America, Washington, DC.

[46] Davenport H. (1937). On some infinite series involving arithmetical functions. II, Quart. J. Math. Oxford, **8**: 313–320.

[47] Denker M., Grillenberger C. and Sigmund C. (1976). Ergodic Theory on Compact Spaces, Lecture Notes in Math. 527, Springer-Verlag, New York.

[48] Derrien J. and Lesigne E. (1996). Un théorème ergodique polynomial ponctuel pourles endomorphismes exacts et les K-systèmes, Ann. Inst. H. Poincaré Probab. Statist., **32** (6): 765–778.

[49] Devaney R. L. (1989). An introduction to chaotic dynamical systems, Second edition. Addison-Wesley Studies in Nonlinearity. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA.

[50] Dong P., Donoso S., Maass A., Shao S. and Ye X. (2013). Infinite-step nilsystems, independence and complexity. Ergodic Theory Dynam. Systems, **33**: 118–143.

[51] Eberlein E. (1975). On topological entropy of semigroups of commuting transformations, International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975), pp. 17–62.

[52] Erdös P. and Turán P. (1936). On some sequences of integers, J. London Math. Soc., **11**: 261–264.

[53] Einsiedler M., Katok A. and Lindenstrauss E. (2006). Invariant measures and the set of exceptions to Littlewood's conjecture, Ann. of Math., **164**(2): 513–560.

[54] Einsiedler M. and Ward T. (2011). Ergodic theory with a view towards number theory. Graduate Texts in Mathematics, 259.

[55] El Abdalaoui E.H., Kułaga-Przymus J., Lemańczyk M. and de la Rue T. (2017). The Chowla and the Sarnak conjectures from ergodic theory point of view, Discrete Contin. Dyn. Syst., **37**(6): 2899–2944.

[56] Ellis R. (1969). Lectures on topological dynamics, W. A. Benjamin, Inc., New York.

[57] Fogg N. P. (2002). Substitutions in dynamics, arithmetics and combinatorics, in Lecture Notes in Mathematics 1794 (Springer-Verlag, Berlin, 2002). Edited by V. Berth′ e, S. Ferenczi, C. Mauduit and A. Siegel.

[58] Folland G. B. (1999). Real Analysis: Modern Techniques and Their Applications (Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts Book 125) 2nd Edition, Kindle Edition.

[59] Friedman N. (1970). Introduction to ergodic theory. Van Nostrand Reinhold Mathematical Studies, No. 29.

[60] F61 Furstenberg H. (1961). Strict ergodicity and transformation of the torus, Amer. J. Math., **83**: 573–601.

[61] Furstenberg H. (1963). The structure of distal flows, Amer. J. Math., **85**: 477–515.

[62] Furstenberg H. (1967). Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation, Math. Systems Theory, **1**: 1–49.

[63] Furstenberg H. (1977). Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J. Analyse Math., **31**: 204–256.

[64] Furstenberg H. (1981). Recurrence in ergodic theory and combinatorial number theory. M. B. Porter Lectures. Princeton University Press, Princeton, N.J.

[65] Furstenberg H. (1984). IP-systems in ergodic theory, Conference in modern analysis and probability (New Haven, Conn., 1982), 131–148, Contemp. Math., 26, Amer. Math. Soc., Providence, RI.

[66] Furstenberg H. and Katznelson Y. (1978). An ergodic Szemerédi theorem for commuting transformations, J. Analyse Math., **34**: 275–291.

[67] Furstenberg H. and Katznelson Y. (1985). An ergodic Szemerédi theorem for IP-systems and combinatorial theory, J. Analyse Math., **45**: 117–168.

[68] Furstenberg H. and Katznelson Y. (1991). A density version of the Hales-Jewett theorem, J. Anal. Math., **57**: 64–119.

[69] Furstenberg H., Katznelson Y. and Ornstein D. (1982). The ergodic theoretical proof of Szemerédi’s theorem, Bull. Amer. Math. Soc., **7**(3): 527–552.

[70] Furstenberg H. and Weiss B. (1978). The finite multipliers of infinite ergodic transformations, The structure of attractors in dynamical systems (Proc. Conf., North Dakota State Univ., Fargo, N.D., 1977), pp. 127–132, Lecture Notes in Math., 668, Springer, Berlin.

[71] Furstenberg H. and Weiss B. (1978). Topological dynamics and combinatorial number theory, J. Analyse Math., **34**: 61–85.

[72] Furstenberg H. and Weiss B. (1996). A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$, Convergence in ergodic theory and probability (Columbus, OH, 1993), 193–227, Ohio State Univ. Math. Res. Inst. Publ., 5, de Gruyter, Berlin.

[73] Glasner S. (1976). Proximal flows. Lecture Notes in Mathematics, Vol. 517. Springer-Verlag, Berlin-New York.

[74] Glasner E. (1994). Topological ergodic decompositions and applications to products of powers of a minimal transformation, J. Anal. Math., **64**: 241–262.

[75] Glasner E. (1997). A simple characterization of the set of μ-entropy pairs and applications. Israel J. Math., **102**: 13–27.

[76] Glasner E. (2003). Ergodic theory via joinings. Mathematical Surveys and Monographs, 101. American Mathematical Society, Providence, RI.

[77] Glasner S. and Weiss B. (1983). Minimal transformations with no common factor need not be disjoint. Israel J. Math., **45**: 1–8.

[78] Glasner E. and Weiss B. (1994). Strictly ergodic, uniform positive entropy models. Bull. Soc. Math. France, **122**(3): 399–412.

[79] Glasner E. and Weiss B. (1995). Quasi-factors of zero-entropy systems. J. Amer. Math. Soc., **8**(3): 665–686.

[80] Glasner E. and Weiss B. (1995). Topological entropy of extensions. Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), 299–307, London Math. Soc. Lecture Note Ser., 205, Cambridge Univ. Press, Cambridge.

[81] Glasner E. and Weiss B. (2005). On the interplay between measurable and topological dynamics. Handbook of Dynamical Systems. Vol. 1B. Eds. Hasselblatt and Katok. North-Holland, Amsterdam, 597–648.

[82] Goodson G. R. (1999). A survey of recent results in the spectral theory of ergodic dynamical systems, J. Dynam. Control System, **5**: 173–226.

[83] Goodman T.N.T. (1974). Topological sequence entropy. Proc. London Math. Soc., **29**: 331–350.

[84] Gottschalk W. and Hedlund G. (1995). Topological Dynamics, Amer. Math. Soc., Providence.

[85] Green B.and Tao T. (2008). The primes contain arbitrarily long arithmetic progressions, Annals of Math., **167**: 481–547.

[86] Green B. and Tao T. (2010). Linear equations in primes, Annals of Math., **171**: 1753–1850.

[87] Green B. and Tao T. (2012). The quantitative behaviour of polynomial orbits on nilmanifolds, Annals of Math., **175**: 465–540.

[88] Green B.,Tao T. and Ziegler T. (2012). In inverse theorem for the Gowers $U^{s+1}[N]$-norm, Annals of Math., **176**:1231–1372.

[89] Gowers W. T. (2001). A new proof of Szemerédi’ s theorem, Geom. Funct. Anal., **11**(3): 465–588.

[90] Gutman Y., Huang W., Shao S. and Ye X.D. (2018). Almost sure convergence of the multiple ergodic average for certain weakly mixing systems, Acta Math Sinica, **34** (1): 79–90.

[91] Halmos P. R. (1950). Measure Theory. D. Van Nostrand Company, Inc., New York, N.Y.

[92] Halmos P. R. (1953). Lectures on Ergodic Theory, Chelsea, New York.

[93] Halmos P. R. and von Neumann J. (1942). Operator methods in classical mechanics, II. Ann. Math., **43**: 332–350.

[94] Hewitt E. and Ross K. A. (1979). Abstract harmonic analysis, Vol. I, in Grundlehren der Mathematischen Wissenschaften 115 (Springer-Verlag, Berlin, second ed.).

[95] Hindman N. (1974). Finite sums from sequences within cells of a partition of $\mathbb{N}$. J. Combinatorial Theory Ser. A, **17**: 1–11.

[96] Hindman N. and Strauss D. (1998). Algebra in the Stone-Čech compactification. Theory and applications. de Gruyter Expositions in Mathematics, 27. Walter de Gruyter and Co., Berlin.

[97] Hiraide K. (1990). Expansive homeomorphisms of compact surfaces are pseudo-Anosov, Osaka J. Math., **27**: 117–162.

[98] Host B. (1991). Mixing of all orders and pairwise independent joinings of systems with singular spectrum, Israel J. Math., **76**(3): 289–298.

[99] Host B. (2009). Ergodic seminorms for commuting transformations and applications, Studia Math., **195**(1): 31–49.

[100] Host B. and Kra B. (2005). Nonconventional ergodic averages and nilmanifolds. Ann. of Math. (2), **161**(1): 397–488.

[101] Host B. and Kra B. (2009). Uniformity norms on l^∞ and applications, J. Anal. Math., **108**: 219–276.

[102] Host B. and Kra B. (2018). Nilpotent Structures in Ergodic Theory, Mathematical Surveys and Monographs, Volume 236, American Mathematical Society.

[103] Host B., Kra B. and Maass A. (2010). Nilsequences and a structure theory for topological dynamical systems, Adv. Math., **224**: 103–129.

[104] Huang W., Li S.M., Shao S. and Ye X.D. (2003). Null systems and sequence entropy pairs, Ergodic Theory Dynam. Systems, **23**(5): 1505–1523.

[105] Huang W., Lian Z., Shao S., Ye X.D. (2021). Minimal systems with finitely many ergodic measures, J. Funct. Anal., **280**(12): Paper No. 109000, 42 pp.

[106] Huang W. and Lu K. (2017). Entropy, chaos, and weak horseshoe for infinite-dimensional random dynamical systems. Comm. Pure Appl. Math. **70**(10): 1987–2036.

[107] Huang W., Maass A., Romagnoli P.P., Ye X.D. (2004). Entropy pairs and a local Abramov formula for a measure theoretical entropy of open covers. Ergodic Theory Dynam. Systems, **24**: 1127–1153.

[108] Huang W., Maass A., Ye X.D. (2004). Sequence entropy pairs and complexity pairs for a measure. Ann. Inst. Fourier (Grenoble), **54**(4): 1005–1028.

[109] Huang W., Shao S., Ye X.D. (2005). Mixing via sequence entropy. Algebraic and topological dynamics. Contemp. Math., **385**: 101–122.

[110] Huang W., Shao S. and Ye X.D. (2016). Nil Bohr-sets and almost automorphy of higher order, Mem. Amer. Math. Soc. volume 241, Number 1143.

[111] Huang W., Shao S. and Ye X.D. (2017). Strictly ergodic models under face and parallelepiped group actions, Commun. Math. Stat., **5**(1): 93–122.

[112] Huang W., Shao S. and Ye X.D. (2019). Pointwise convergence of multiple ergodic averages and strictly ergodic models, J. Analyse Math., **139**(1): 265–305.

[113] Huang W., Shao S. and Ye X.D. (2019). Topological correspondence of multiple ergodic averages of nilpotent actions, J. Anal. Math., **138**(2): 687–715.

[114] Huang W., Shao S. and Ye X.D. (2020). Parallels Between Topological Dynamics and Ergodic Theory，Encyclopedia of Complexity and Systems Science.

[115] Huang W. and Ye X.D. (2006). A local variational relation and applications. Israel J. Math., **151**: 237–280.

[116] Huang W. and Ye X.D. (2009). Combinatorial lemmas and applications to dynamics. Adv. Math., **220**(6): 1689–1716.

[117] Huang W., Ye X.D. and Zhang G. (2011). Local entropy theory for a countable discrete amenable group action. J. Funct. Anal., **261**(4): 1028–1082.

[118] Hulse P. (1982). Sequence entropy and subsequence generators. J. London Math. Soc. (2), **26**: 441–450.

[119] Hulse P. (1986). Sequence entropy relative to an invariant σ-algebra. J. London Math. Soc. (2), **33**: 59–72.

[120] Jewett R.I. (1969/1970). The prevalence of uniquely ergodic systems, J. Math. Mech., **19**: 717–729.

[121] Kalikow S. (1984). Twofold mixing implies threefold mixing for rank one transformations. Ergodic Theory Dynam. Systems, **4**(2): 237–259.

[122] Katok A. (1980). Smooth non-Bernoulli K-automorphism, Invent. Math., **61**(3): 291–300.

[123] Katok A. and Hasselblatt B. (1995). Introduction to the modern theory of dynamical systems. With a supplementary chapter by Katok and Leonardo Mendoza. Encyclopedia of Mathematics and its Applications, 54. Cambridge University Press, Cambridge, xviii+802.

[124] Katznelson Y. (1976). An Introduction to Harmonic Analysis, 2nd ed., Dover Publications, Inc., New York.

[125] Katznelson Y. and Weiss B. (1982). A simple proof of some ergodic theorems, Israel J. Math., **42**(4): 291–296.

[126] Kechris A. S. (1995). Classical descriptive set theory. Graduate Texts in Mathematics, **156**. Springer-Verlag, New York.

[127] Kelley J. (1955). General Topology, Van Nostrand, Princeton.

[128] Kerr D. and Li H. (2016). Ergodic Theory: Independence and Dichotomies, Springer Monographs in Mathematics.

[129] Kerr D. and Li H. (2007). Independence in topological and C^*-dynamics. Math. Ann., **338**(4): 869–926.

[130] Kerr D., Li H. (2009). Combinatorial independence in measurable dynamics. J. Funct. Anal., **256**(5): 1341-1386

[131] Kesseböhmer M., Munday S. and Stratmann B. (2016). Infinite ergodic theory of numbers. De Gruyter Graduate. De Gruyter, Berlin, xiii+191 pp.

[132] Khinchin A. Y. (1952). Three pearls of number theory (Dover Publications Inc., Mineola, NY, 1998). Translated from the Russian by F. Bagemihl, H. Komm, and W. Seidel, Reprint of the 1952 translation.

[133] Kolmogorov A. N. (1958). A new metric invariant of transient dynamical systems and automorphisms of Lebesgue spaces, Dokl. Akad. Sci. SSSR., **119**: 861–864(Russian).

[134] Koopman B. O. and von Neumann J. (1932). Dynamical systems continuous spectra, Proc. Nat. Acad. Sci. U.S.A., **18**: 255–263.

[135] Kra B. (2006). From combinatorics to ergodic theory and back again, International Congress of Mathematicians. Vol. III, 57–76, Eur. Math. Soc., Zürich.

[136] Krengel U. (1985). Ergodic theorems, in de Gruyter Studies in Mathematics 6 (Walter de Gruyter & Co., Berlin, 1985). With a supplement by Antoine Brunel.

[137] Krieger W. (1970). On entropy and generators of measure-preserving transformations, Trans. Amer. Math. Soc., **149**: 453–464.

[138] Krieger W. (1972). On unique ergodicity, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, pp. 327–346, Univ. California Press, Berkeley, Calif., 1972.

[139] Krug E. and Newton D. (1972). On sequence entropy of automorphisms of a Lebesgue space, Z. Wahrscheinlichtkeitstheorie und Verw. Gebiete, **24**: 211–214.

[140] Kushnirenko A.G. (1967). On metric invariants of entropy type, Russion Math. Surveys, **22**: 53–61.

[141] Ledrappier F. (1978). Un champ markovien peut être d’entropie nulle et mélangeant, C. R. Acad. Sci. Paris Sér. A-B, **287**(7): A561–A563.

[142] Lehrer E. (1987). Topological mixing and uniquely ergodic systems, Israel J. Math., **57**(2): 239–255.

[143] Leibman A. (1998). Multiple recurrence theorem for measure preserving actions of a nilpotent group, Geom. Funct. Anal., **8**: 853–931.

[144] Leibman A. (2005). Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, Ergodic Theory Dynam. Systems, **25**: 201–213.

[145] Lemanczyk M. (2009). Spectral theory of dynamical systems, In Encyclopedia of Complexity and Systems Science, pages 8554–8575.

[146] Lewowicz J. (1989). Expansive homeomorphisms of surfaces, Bol. Soc. Brasil. Mat. (N.S.), **20**(1): 113–133.

[147] Lind D. and Marcus B. (1995). An introduction to symbolic dynamics and coding, Cambridge University Press.

[148] Lindenstrauss E. (2001). Pointwise theorems for amenable groups, Invent. Math., **146**(2): 259–295.

[149] Liu P. and Qian M. (1995). Smooth ergodic theory of random dynamical systems. Lecture Notes in Mathematics, 1606.

[150] Iwaniec H. and Kowalski E. (2004). Analytic number theory, in American Mathematical Society Colloquium Publications 53 (American Mathematical Society, Providence, RI, 2004).

[151] Lyons R. (1988). On measures simultaneously 2- and 3-invariant, Israel J. Math., **61**(2): 219–224.

[152] Maass A. and Shao S. (2007). Structure of bounded topological-sequence-entropy minimal systems. J. Lond. Math. Soc. (2), **76**(3): 702–718.

[153] Mañé R. (1987). Ergodic theory and differentiable dynamics. Springer-Verlag, Berlin.

[154] McCutcheon R. (1999). Elemental methods in ergodicRamsey theory. Lecture Notes in Mathematics, 1722.

[155] Nadkarni M. (1998). Basic ergodic theory. Second edition. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel.

[156] Nadkarni M. (1998). Spectral theory of dynamical systems, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel.

[157] von Neumann J. (1932). Zur Operatorenmethode in der klassischen Mechanik, Ann. of Math. (2), **33**: 587–642.

[158] O'Brien T. and Reddy W. (1970). Each compact orientable surface of positive genus admits an expansive homeomorphism, Pacific J. Math., **35**: 737–741.

[159] Ornstein D. S. (1970). On the root problem in ergodic theory, Proc. sixth Berkeley symposium Math. Statist. Probab. Univ. of California Press, 347–356.

[160] Ornstein D. S. and Shields (1973). An uncountable family of K-automorphisms, Advances in Math., **10**: 63–88.

[161] Parry W. (1981). Topics in Ergodic Theory, Cambridge Tracks in Mathematics, Cambridge University Press, Cambridge-New York.

[162] Petersen K. (1983). Ergodic theory. Cambridge Studies in Advanced Mathematics, **2**. Cambridge University Press, Cambridge.

[163] Phelps R. (1966). Lectures on Choquet's theorem, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London.

[164] Poincaré H. (1899). Les méthodes nouvelles de la mécanique céleste, vol. 3. Gauthier-Villars, Paris.

[165] Queffélec M. (1987). Substitution Dynamical systems–spectral analysis, Lecture Notes in Mathematics 1294, Springer–Verlag, Berlin.

[166] Rohlin V. A. (1949). On endomorphisms of compact commutative groups. (Russian) Izvestiya Akad. Nauk SSSR. Ser. Mat., **13**: 329–340.

[167] Rohlin V. A. (1962). On the fundament ideas of measure theory, Amer. Math. Soc. Transl. Ser. 1, **10**: 1–54.

[168] Rohlin V. A. (1967). Lectures on ergodic theory, Russian Math. Surverys, **22**: 1–52.

[169] Rohlin V. A. and Sinai Ya. G. (1967). Construction and properties of invariant measure, Usp. Mat. Nauk, **22**: 4–54 (Russian).

[170] Romagnoli P.P. (2003). A local variational principle for the topological entropy. Ergodic Theory Dynam. Systems, **23**: 1601–1610.

[171] Royden H. L. (1988). Real analysis, (Macmillan Publishing Company, New York, third ed.

[172] Rudin W. (1987). Real and complex analysis, 3rd edition.

[173] Rudin W. (2004) Functional analysis, 2nd edition.

[174] Rudolph D. (1990). $\times 2$ and $\times 3$ invariant measures and entropy, Ergodic Theory Dynam. Systems, **10**(2): 395–406.

[175] Rudolph D. J. (1990). Fundamentals of measurable dynamics. Ergodic theory on Lebesgue spaces. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York.

[176] De La Rue T. (2006). 2-fold and 3-fold mixing: why 3-dot-type counterexamples are impossible in one dimension, Bull. Braz. Math. Soc. (N.S.), **37**(4): 503–521.

[177] De La Rue T. (2009). Joinings in ergodic theory, Encyclopedia of Complexity and Systems Science, pages 5037–5051.

[178] Ryzhikov V. V. (1993). Joinings and multiple mixing of the actions of finite rank, (Russian) Funktsional. Anal. i Prilozhen. **27**(2): 63–78; translation in Funct. Anal. Appl. **27**(2): 128–140.

[179] Sarnak P. (2009). Three Lectures on the Mobius function randomness and dynamics, http://publications.ias.edu /sarnak/paper/512

[180] Shannon C. (1948). A mathematical theory of communication, Bell System Tech. J., **27**: 379–423.

[181] Saleski A. (1977). Sequence entropy and mixing. J. of Math. Anal. and Appli., **60**: 58–66.

[182] Seneta E. (1980). Non-negative Matrices and Markov Chains (Second Ed.), Springer, New York.

[183] Shao S. and Ye X.D. (2012). Regionally proximal relation of order d is an equivalence one for minimal systems and a combinatorial consequence, Adv. Math., **231**: 1786–1817.

[184] Sinai Y. (1994). Topics in ergodic theory. Princeton Mathematical Series, 44. Princeton University Press.

[185] Snoha L., Ye X.D. and Zhang R. (2020). Topology and Topological Sequence Entropy. Sci. China Math., **63**: 205–296.

[186] Szemerédi E. (1975). On sets of integers containing no k elements in arithmetic progression, Acta Arith., **27**: 199–245.

[187] Tao T. (2006). A quantitative ergodic theory proof of Szemerédi's theorem. Electron. J. Combin., **13**, no. 1, Research Paper 99, 49 pp.

[188] Tao T. (2008). Norm convergence of multiple ergodic averages for commuting transformations, Ergodic Theory Dynam. Systems, **28**: 657–688.

[189] Thouvenot J. P. (1995). Some properties and applications of joinings in ergodic theory, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), **205**: 207–235.

[190] Veech W. A. (1977). Topological systems, Bull. Amer. Math. Soc., **83**: 775–830.

[191] Viana M. and Marcelo O. (2016). Foundations of ergodic theory. Cambridge Studies in Advanced Mathematics, 151.

[192] de Vries J. (1993). Elements of Topological Dynamics, Kluwer Academic Publishers, Dordrecht.

[193] van der Waerden B. L. (1927). Beweis einer Baudetschen Vermutung, Nieuw Arch. Wisk., **15**: 212–216.

[194] Walsh M. (2012). Norm convergence of nilpotent ergodic averages, Annals of Math., **175**: 1667–1688.

[195] Walters P. (1982). An introduction to ergodic theory. Graduate Texts in Mathematics, **79**. Springer-Verlag, New York-Berlin.

[196] Weiss B. (1985). Strictly ergodic models for dynamical systems, Bull. Amer. Math. Soc. (N.S.), **13**: 143–146.

[197] Weiss B. (1989). Countable generators in dynamics – Universal minimal models, Contemporary Mathematics, **94**: 321–326.

[198] Weiss B. (1998). Multiple recurrence and doubly minimal systems, Topological dynamics and applications (Minneapolis, MN, 1995), 189–196, Contemp. Math., 215, Amer. Math. Soc., Providence, RI.

[199] Weiss B. (2000). Single Orbit Dynamics, AMS Regional Conference Series in Mathematics, vol. **95**.

[200] Zhang Q. (1992). Sequence entropy and mild mixing. Canad. J. Math., **44**: 215–224.

[201] Zhang Q. (1993). Conditional sequence entropy and mild mixing extensions. Canad. J. Math., **45**: 429–448.

[202] Ziegler T. (2007). Universal characteristic factors and Furstenberg averages, J. Amer. Math. Soc., **20**: 53–97.

[203] Zimmer R. J. (1976). Extensions of ergodic group actions, Illinois J. Math., **20**(3): 373–409.

[204] Zimmer R. J. (1976). Ergodic actions with generalized discrete spectrum, Illinois J. Math., **20**(4): 555–588.

[205] 黎景辉，冯绪宁, (2007). *拓扑群引论*, 科学出版社.

[206] 孙文祥, (2012). *遍历论*，北京大学出版社.

[207] 叶向东, 黄文, 邵松, (2008). *拓扑动力系统概论*, 科学出版社.

[208] 张筑生, (1999). *微分动力系统原理*, 科学出版社.

[209] 张景中, 熊金城, (1992). *函数迭代与一维动力系统*, 四川教育出版社.

[210] 周作领, (2001). *符号动力系统*, 科学出版社.

[211] 华罗庚, (1956). *数论导引*, 科学出版社.

索 引