arXiv:2305.05028v1 [math.DS] 8 May 2023

随机动力系统遍历定理的非平稳版本

ANTON GORODETSKI 和 Victor Kleptsyn

谨以此文献给我们的老师尤利·谢尔盖耶维奇·伊利亚申科教授

Abstract.我们证明了非平稳随机动力系统的逐点遍历定理的一个版本。此外，我们还讨论了该结果适用的两个具体例子：非平稳迭代函数系统和非平稳随机矩阵乘积。

1. Introduction

伯克霍夫遍历定理是动力系统理论的关键工具之一。在拓扑动力系统的背景下，它指出，对于一个连续动力系统 $f\colon\ X \to X$ 在紧度量 X 上，若存在一个 f 不变测度ν 定义在 X上，给定一个“测试函数” $\varphi \in C(X)$，时间平均

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(x))$$

对 ν几乎处处的 $x \in X$收敛；若测度 ν 是遍历的，则几乎处处的极限等于空间平均 $\int_X \varphi\, d\nu$。

该定理有一个自然推广，适用于随机动力系统。具体而言，设 X 为紧度量空间， μ 为连续映射空间 $C(X,X)$上的概率测度。相应随机动力系统的迭代是由复合构成的序列

$$f_1,\ f_2 \circ f_1,\ \dots,\ f_n \circ \dots \circ f_1,\ \dots,$$

其中 $f_i\colon\ X \to X$ 是相对于测度 μ随机且独立选取的。

不变测度这一概念的类似概念是平稳测度。具体来说， μ-平稳测度 ν 是 X 上的概率测度，满足 $\mu * \nu = \nu$；其中，卷积 $\mu * \nu$ 是相互独立且分别依照 μ 和 ν 选取的 $f(x)$ 与 x 的分布。这样的测度总是存在的（可通过类似克雷洛夫-博戈柳博夫平均过程的方法证明），但一般不唯一（正如经典确定性动力系统中的不变测度通常也不唯一）。

Date：2023年5月10日。A. G. 的部分研究工作获得了美国国家科学基金会资助，项目编号为 DMS–1855541。A. G. 在瑞典于尔斯霍尔姆的米塔-莱夫勒研究所驻留期间完成了部分工作，驻留时间为2023年春季学期，并获得瑞典研究理事会项目（编号 2016-06596）的资助。V.K. 的部分研究工作获得了 ANR Gromeov（ANR-19-CE40-0007）和亨利·勒贝格中心（ANR-11-LABX-0020-01）的资助。

下面的定理对应于平稳随机动力系统的伯克霍夫遍历定理：

定理 1.1（随机伯克霍夫遍历定理）。对于 X 上的任意遍历平稳测度 ν，对于任意 $\varphi \in C(X,\mathbb{R})$，对 ν-几乎处处的 $x \in X$，$\mu^{\mathbb{N}}$-几乎必然有

$$\frac{1}{n}\sum_{k=0}^{n-1} \varphi(f_k \circ \ldots \circ f_1(x)) \to \int_X \varphi(x) d\nu(x),$$

其中 f_1、f_2、.....是根据分布 μ 随机且独立选取的。

然而，这一推广在很大程度上依赖于 f_n 具有相同分布这一事实。如果 f_n 仍然是独立选取的，但每个 f_n 都服从不同的测度 μ_n，那么会有什么结论？这可以用下面的图示来表示：

$$X \xrightarrow[f_1]{\mu_1} X \xrightarrow[f_2]{\mu_2} X \to \cdots \to X \xrightarrow[f_n]{\mu_n} X \to \ldots$$

乍一看，许多概念和工具都不再适用，首先就是平稳测度的概念：不同步骤上的测度 μ_n 可能各不相同，而且对所有 n 而言，方程组 $\nu = \mu_n * \nu$ 通常彼此不相容。因此，即使只是找到一个恰当的表述也并非易事，更不用说证明了。

与此同时，这种情形自然会出现在某些问题中。例如，在存在背景势的情况下研究一维安德森局域化问题 [GK3]，会引出对独立非同分布矩阵乘积的研究；非平稳富斯滕贝格定理 [GK2], 属于非平稳遍历理论的范畴。这也促使人们研究给定初始测度的平均像的 H¨older 正则性，见 [GKM，定理 2.8]。

本文旨在探讨非平稳情形下遍历定理的一种范式，既给出（反）例，也提出一个受限版本的遍历定理，见下文定理 3.1。后者只是一个“概念验证”；它显然可以大幅推广，我们也希望不久便能实现：我们的目标是激励这一方向上的进一步研究。

2. Preliminaries

在本文中，我们令 X 为一个紧度量空间，并假设给定了一个“测试函数” $\varphi \in C(X)$。此外，对于每个 n，我们假设在 $C(X,X)$ 上给定一个测度μ_n（如果假设动力系统是可逆的，则也可以改为在 X 的同胚集合上给定测度）。于是我们记

$$\mathbb{P} := \prod_{n=1}^{\infty} \mu_n, \tag{1}$$

我们的目标是描述沿着 $\mathbb{P}$-几乎每个迭代序列的时间平均。

我们用 $\mathcal{M} = \mathcal{M}(X)$ 表示紧度量空间 X 上的博雷尔概率测度空间。我们认为该空间配备了 Wasserstein距离（这是在概率测度空间上度量弱-*拓扑的方法之一）：

定义 2.1. 设 ν_1、$\nu_2 \in \mathcal{M}$ 是 X 上的两个概率测度。ν_1 与 ν_2 之间的 Wasserstein 距离定义为

$$\operatorname{dist}_{\mathcal{M}}(\nu_1, \nu_2) = \inf_{\gamma} \iint_{X\times X} d_X(x,y)\, d\gamma(x,y), \tag{2}$$

其中，下确界取遍所有定义在 $X \times X$ 上、边缘概率测度分别为 ν_1 和 ν_2 的概率测度 γ，即 $(\pi^j)_*\gamma = \nu_j$, $j = 1, 2$，其中 $\pi^1,\ ^2\colon\ X \times X \to X$ 分别是到第一和第二个因子的投影。

接下来，注意对于一个随机轨道

$$(x_n)_{n\in\mathbb{N}}, \quad x_n = f_n(x_{n-1})$$

动力系统的非平稳性意味着 x_n 的分布可能发生相当显著的变化，因此 $\varphi(x_n)$ 的分布及其期望值也会如此。因此，我们不应预期时间平均 $\frac{1}{n}\sum_{k=0}^{n-1}\varphi(x_k)$ 会收敛。相反，我们应预期它们具有确定性的行为，这与非平稳大数定律的情形类似。具体而言，在对独立但不同分布的随机变量 ξ_n 施加合理假设的条件下，有

$$\frac{1}{n}\left|\sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n}\mathbb{E}\xi_k\right| \to 0, \quad n\to\infty.$$

既然伯克霍夫遍历定理推广了大数定律，我们就应该预期

$$\frac{1}{n}\left|\sum_{k=1}^{n}\varphi(x_k) - \sum_{k=1}^{n}\int_X \varphi(x)d\nu_k(x)\right| \to 0, \quad n\to\infty \tag{3}$$

对于某些合理的非随机测度 ν_k。事实上，结果确实如此：见下文的定理 3.1。

3. Statement of the main result

我们的主要结果将在以下假设下陈述。从某种意义上说，该假设是对唯一遍历性的推广（并略有加强）。后者要求平稳测度是唯一的，因此，这一设定比“简单的”遍历性更具限制性；然而，在此设定下，这样的假设似乎是恰当的，参见下文第 6.2 节中的讨论及其中的例 6.3。

基本假设：如果$\mu_1, \mu_2, \mu_3, \ldots$ 是 $C(X, X)$上的一列分布，并且对于任意 $\delta > 0$，都存在 $m \in \mathbb{N}$，使得任意两个初始测度在任意初始时刻 n 之后，经过m个随机步骤取平均后的像彼此 δ-接近，那么我们称这列分布满足基本假设：

$$\forall \nu, \nu' \in \mathcal{M}, \quad \forall n \in \mathbb{N}, \quad \operatorname{dist}_{\mathcal{M}}(\mu_{n+m} * \ldots * \mu_{n+1} * \nu, \mu_{n+m} * \ldots * \mu_{n+1} * \nu') < \delta.$$

注意，如果经过给定的步数 m 后，任意两个初始测度 ν、ν' 的平均像彼此接近，那么可以选取其中任意一个像，并指出其他所有像都与它接近。因此，可以选取任意测度 ν_0（例如，给定初始点 x_0处的狄拉克测度），并定义

$$\nu_n := \mu_n * \nu_{n-1}, \quad n = 1, 2, \ldots. \tag{4}$$

那么，对于任意给定的测度 ν，卷积 $\mu_{n+m} * \ldots * \mu_{n+1} * \nu$与$\nu_{n+m}$接近，这也解释了公式 (3) 中测度 ν_n 的出现。

以下是本文的主要结果：

定理 3.1. 假设分布序列 $\mu_1, \mu_2, \mu_3, \ldots$ 满足上述基本假设。给定 X 上的任意博雷尔概率测度 ν_0，定义

$$\nu_n := \mu_n * \nu_{n-1}, \quad n = 1, 2, \ldots.$$

那么，对于任意 $\varphi \in C(X, \mathbb{R})$ 和任意 $x \in X$，几乎必然地

$$\frac{1}{n}\left|\sum_{k=1}^{n} \varphi(f_k \circ \ldots \circ f_1(x)) - \sum_{k=1}^{n} \int_X \varphi \, d\nu_k\right| \to 0 \quad \text{as} \quad n \to \infty,$$

其中，测度 ν_n 由（4）定义。

此外，大偏差定理的一个类似结论也成立。具体而言，对于任意$\varepsilon > 0$，存在 C、$\delta > 0$，使得对于任意 $x \in X$

$$(5) \quad \forall n \in \mathbb{N}, \quad \mathbb{P}\left(\frac{1}{n}\left|\sum_{k=1}^{n} \varphi(f_k \circ \ldots \circ f_1(x)) - \sum_{k=1}^{n} \int_X \varphi \, d\nu_k\right| > \varepsilon\right) < C \exp(-\delta n).$$

注 3.2. 尽管上述基本假设看起来相当严格，但在一些有趣的情形下，它确实自然成立。下一节将证明，在两种情形下该假设成立：一是由压缩映射定义的随机动力系统，也称为随机迭代函数系统；二是由随机矩阵乘积生成的射影空间上的随机动力系统。

注 3.3. 在平稳情形下，当 $\mu_i = \mu$ 对所有 $i \in \mathbb{N}$都成立时，基本假设意味着由分布 μ定义的随机动力系统是唯一遍历的，即具有唯一的平稳测度。不过请注意，基本假设实际上比唯一遍历性更强（即限制更多）。具体而言，唯一遍历性等价于以下陈述：任意初始测度 ν经卷积作用后的像的时间平均值 $\overline{\nu}$ 均匀地收敛到唯一的平稳测度，且该收敛关于 ν均匀：

$$\frac{1}{m}\sum_{j=0}^{m} \mu^{(*j)} * \nu \to \overline{\nu}, \quad m \to \infty,$$

其中 $\mu^{(*j)}$ 是 j 次卷积幂。例 6.3 说明了为什么非平稳情形下的基本假设确实应该更强。

注 3.4. 我们想指出，遍历定理的非平稳版本已经在多种不同的情形下建立。例如，[BB] 考虑了保持同一固定测度的映射序列。[CR、HNTV] 研究了所有映射相对于勒贝格测度均为非奇异映射的情形，因此可以应用传递算子技术。在这些语境中，有时会使用“序列动力系统”这一术语。我们强调，这些研究中的情形和底层机制都与我们的不同。

4. Examples of applicability

我们将给出两个满足基本假设的非平稳随机动力系统示例，因此定理 3.1 适用。

4.1. 随机迭代函数系统。迭代函数系统由完备度量空间上的有限个压缩映射定义，通常是 $\mathbb{R}^d$ 或 $\mathbb{R}^d$ 的紧子集。在许多情况下，它会产生一个分形吸引子。例如，标准康托集可以通过单位区间上的两个压缩映射定义，$x \mapsto \frac{x}{3}$ 和 $x \mapsto \frac{x}{3} + \frac{2}{3}$。许多自相似分形都可以通过这种方式生成。若为每个压缩映射赋予概率，就可以考虑相应的随机动力系统。已知它必定存在唯一的平稳测度 [Hu, Sz]。这种平稳测度的性质一直是深入研究的主题。关于这种测度的精确维数性和分形维数，已有大量文献；这里仅提及一篇近期综述 [FS]。文献 [FL, Proposition 2.2] 和 [GKM, Section 1.2] 研究了这种测度的 Hölder 正则性。关于同样属于这一框架的伯努利卷积理论综述，见 [PSS]，或较近期的 [Va]。

可以考虑一个由压缩映射生成的非平稳随机动力系统，其中每一步都可以在压缩映射空间上选取不同的分布。例如，[GKM] 中研究了这种情形；研究表明，在一些温和且自然的条件下，任意给定初始测度的平均迭代会以指数速度收敛到空间中的 H¨older 正则测度。

很容易看出，在这种情形下基本假设成立，因此定理 3.1 适用：

命题 4.1. 设 X 是一个紧度量空间，$\lambda \in (0, 1)$ 是一个常数。假设 $\{\mu_i\}_{i\in\mathbb{N}}$ 是压缩映射空间 $X \to X$ 上的概率分布序列，且这些压缩映射的利普希茨常数至多为 λ。那么基本假设成立，即对于 X 上任意两个博雷尔概率测度 ν、ν'，我们有

$$\mathrm{dist}_{\mathcal{M}}(\mu_{n+m} * \ldots * \mu_{n+1} * \nu, \mu_{n+m} * \ldots * \mu_{n+1} * \nu') \to 0 \quad \text{as} \quad m \to \infty,$$

关于 (ν, ν') 和 $n \in \mathbb{N}$均匀地。

证明。事实上，对于任意 μ_i，任意映射 $f \in \operatorname{supp} \mu_i$ 都是压缩映射，其利普希茨常数至多为 λ。因此，对于任意 $\nu_1, \nu_2 \in \mathcal{M}$，我们有

$$\mathrm{dist}_{\mathcal{M}}(\mu * \nu_1, \mu * \nu_2) \leq \lambda \mathrm{dist}_{\mathcal{M}}(\nu_1, \nu_2),$$

由此命题 4.1 得证。 □

4.2. 随机矩阵乘积。现在我们证明，在某些非退化性假设下，由射影映射定义的非平稳随机动力系统必定满足基本假设。

假设 K 是 $SL(2, \mathbb{R})$ 上概率分布空间中的一个紧集，并且对任意 $\mu \in$ K，都满足测度条件（我们采用 [GKM] 中的术语），也就是说，不存在 $\mathbb{RP}^1$ 上的概率测度 ν_1, ν_2，使得对于所有 $A \in \operatorname{supp} \mu$，均有 $(f_A)_*(\nu_1) = \nu_2$。稍微滥用一下记号，我们也将 μ视为射影映射空间 $f_A : \mathbb{RP}^1 \to \mathbb{RP}^1$ 上的测度。

命题 4.2. 假设每个 $\mu \in$K 都满足测度条件。那么对于任意序列 $\{\mu_i\} \in$ K$^{\mathbb{N}}$ 以及任意概率测度 $\nu, \nu' \in \mathcal{M}$，我们有：

$$\mathrm{dist}_{\mathcal{M}}(\mu_{n+m} * \ldots * \mu_{n+1} * \nu, \mu_{n+m} * \ldots * \mu_{n+1} * \nu') \to 0 \quad \text{as} \quad m \to \infty,$$

关于 (ν)、(ν') 和 $n \in \mathbb{N}$ 均匀地。

我们将从以下陈述推导出这一命题，该陈述本身也具有独立的意义：

命题 4.3。假设对每个 $\mu \in \mathbf{K}$ 都满足测度条件。那么对于任意 $\varepsilon > 0$，存在 m，使得对于任意测度 $\mu_1, \dots, \mu_m \in \mathbf{K}$ 以及任意点 $x, y \in \mathbb{RP}^1$，都有

$$(\mu_m * \dots * \mu_1)\{F \mid d(F(x), F(y)) < \varepsilon\} > 1 - \varepsilon,$$

其中，复合映射 $F = f_m \circ \dots \circ f_1$ 由独立随机映射 f_i 组成，这些映射服从 μ_i。换言之，随机复合映射 F 以至少 $1 - \varepsilon$ 的概率，将 x 和 y 之间的距离缩短至小于 ε。

这一推论确实是显而易见的：

命题 4.2 的证明。首先注意到

$$\operatorname{dist}_{\mathcal{M}}(\mu_{n+m} * \dots * \mu_{n+1} * \nu, \mu_{n+m} * \dots * \mu_{n+1} * \nu') \le \iint_{\mathbb{RP}^1} \operatorname{dist}_{\mathcal{M}}(\mu_{n+m} * \dots * \mu_{n+1} * \delta_x, \mu_{n+m} * \dots * \mu_{n+1} * \delta_y)\, d\nu(x) d\nu'(y);$$

这是由传输距离的凸性得出的：对耦合进行混合，就得到了平均测度的一个耦合。

与此同时，我们还有

$$\operatorname{dist}_{\mathcal{M}}(\mu_{n+m} * \dots * \mu_{n+1} * \delta_x, \mu_{n+m} * \dots * \mu_{n+1} * \delta_y) \le \le \int d(F(x), F(y))\, d(\mu_{n+m} * \dots * \mu_{n+1})(F). \tag{6}$$

取任意一个 $\varepsilon > 0$；一旦 m 取得足够大，使得命题 4.3 的结论成立，(6) 式的右端便不超过

$$\varepsilon \cdot (\mu_{n+m} * \dots * \mu_{n+1})\{F \mid d(F(x), F(y)) < \varepsilon\} + \\ + 1 \cdot (\mu_{n+m} * \dots * \mu_{n+1})\{F \mid d(F(x), F(y)) \ge \varepsilon\} \le \varepsilon \cdot 1 + 1 \cdot \varepsilon = 2\varepsilon.$$

由于 $\varepsilon > 0$ 是任意的，我们得到了所需的收敛性，且该收敛性关于 n、ν 和 ν' 是一致的。□

现在让我们回到命题 4.3。在继续证明之前，请注意，如果动力学是平稳的，且 $\log \|A\|$ 的期望有界，那么该命题的结论将由著名的 Furstenberg 定理推出。这促使我们通过 Furstenberg 定理的非平稳版本来处理该命题。此类版本最近在 [GK2] 和 [G] 中得到了研究。

命题 4.3 的证明。假设 $\|A\|$ 的期望关于 $\mu \in \mathbf{K}$ 有界：

$$\exists C_{\mathbf{K}} : \quad \forall \mu \in \mathbf{K} \quad \int \|A\|\, d\mu(A) < C_{\mathbf{K}}. \tag{7}$$

那么定理 1.1 和定理 1.4 的条件适用，其中 [GK2]。具体而言，记

$$L_n := \mathbb{E} \log \|A_n \dots A_1\|.$$

来自 [GK2] 的定理 1.1 进而推出了 L_n 的线性下界：存在 $\lambda_{\mathbf{K}} > 0$ 使得

$$L_n \ge n\lambda_{\mathbf{K}};$$

此外， $\lambda_{\mathbf{K}}$ 不取决于对 μ_i 的具体选择，而仅取决于紧集 K 本身。

现在， [GK2] 中的定理1.4指出，大偏差类型的估计成立：对于任意固定的 $\delta > 0$，存在 δ'，使得对于任意 $v_0 \in \mathbb{R}^2$、 $|v_0| = 1$，以及任意充分大的 n （即，对于所有 n 大于某个 n_0的情形），都有

$$\mathbb{P}(|\log |A_n \dots A_1 v_0| - L_n| > n\delta) < e^{-n\delta'}.$$

此外（参见 [GK2，注 1.4]），常数 δ' 和 n_0 可以对所有可能的 $\mu_i \in$K 序列均匀地选取。

为在这一特定情形下证明命题的结论，即 (7) 成立时，取 $\delta := \frac{1}{2}\lambda_{\mathbf{K}} > 0$ 并考虑相应的 n_0, δ'。然后，一旦不等式

$$|\log |A_n \dots A_1 v_0| - L_n| < n\delta \tag{8}$$

成立时，它意味着

$$\log |A_n \dots A_1 v_0| > L_n - n\delta \geq \frac{1}{2}\lambda_{\mathbf{K}} n.$$

特别地，对于任意给定的 $\varepsilon > 0$，取 n_1 为 $\log_{\frac{1}{2}\lambda_{\mathbf{K}} n}^{2\varepsilon^{-1}}$ 的整数部分，我们有

$$|A_n \dots A_1 v_0| > 2\varepsilon^{-1}$$

一旦 $n > n_1$ 且 (8) 成立。进而，这个不等式意味着，像 $A_n \dots A_1 v_0$ 与扩张程度最大的单位向量的像之间的夹角不超过 $\frac{\varepsilon}{2}$。因此，对对应于两个点 x、 $y \in \mathbb{RP}^1$ 的两个向量 v_0、 v_0' 都有这一性质，便可推出

$$\operatorname{dist}(F(x), F(y)) < \varepsilon, \tag{9}$$

其中 $F = f_{A_n \dots A_1}$。另一方面，对于任意 $n > n_2 := \lceil \frac{1}{\delta'} \log 2\varepsilon^{-1} \rceil$，有

$$e^{-n\delta'} < \frac{\varepsilon}{2}.$$

因此，对于 $n >$ max(n_0, n_1, n_2)，我们以至少为.....的概率得到所需的 (9)

$$1 - 2e^{-n\delta'} > 1 - 2\frac{\varepsilon}{2} = 1 - \varepsilon$$

对于任意的 μ_1、.....、 $\mu_n \in$K 以及 x、 $y \in \mathbb{RP}^1$。

现在，我们转向一般情形。以下工具有助于研究圆上的随机动力系统。考虑一个逆平稳测度，即满足以下条件的测度 ν^-

$$\nu^- = \mathbb{E}(f^{-1})_* \nu^-.$$

区间经过随机正向迭代后所得的测度构成一个鞅：

$$\mathbb{E}\nu^-(f(J)) = \mathbb{E}((f^{-1})_* \nu^-)(J) = \nu^-(J).$$

在非平稳情形下，单个平稳测度的概念不再适用。因此，与定理 3.1 的表述类似，我们将以一个通过卷积关联的测度序列来取代这一概念。具体而言，对于给定的每个n 以及测度 $\mu_1, \dots, \mu_n \in$K，我们定义

$$\nu_n^- := \mathrm{Leb}, \quad \forall i = 1, \dots, n \quad \nu_{i-1}^- := \mathbb{E}_{\mu_i}(f^{-1})_* \nu_i^- = \mu_i^- * \nu_i^-,$$

其中，测度 μ_i^- 是 μ_i 在取逆映射 $f \mapsto f^{-1}$ 下的推前像。那么，对于迭代的测度，我们也有鞅性质：对于任意 i 和任意区间 J，均有

$$\mathbb{E}_{\mu_i} \nu_i^-(f(J)) = \nu_{i-1}^-(J), \tag{10}$$

因此

$$\mathbb{E}_{\mu_n * \cdots * \mu_i} \nu_n^-(F(J)) = \nu_{i-1}^-(J). \tag{11}$$

对于任意给定的 $x, y \in \mathbb{RP}^1$，考虑它们随机迭代所构成的序列：令 f_i 相对于 μ_i 独立选取，并记

$$x_0 := x,\, y_0 := y, \quad x_i = f_i(x_{i-1}),\, y_i = f_i(y_{i-1}).$$

于是我们有以下引理。

引理 4.4. 假设对于某个 i，经过 i 次迭代后，我们有 $\nu_i^-([x_i, y_i]) < \frac{1}{2}\varepsilon^2$ 或 $\nu_i^-([y_i, x_i]) < \frac{1}{2}\varepsilon^2$。那么，以至少 $1 - \frac{\varepsilon}{2}$ 的概率，我们有

$$\operatorname{dist}(x_n, y_n) < \varepsilon.$$

证明。假设我们处于第一种情形。那么，由于

$$\mathbb{E}(\nu_n^-([x_n, y_n]) \mid f_1, \ldots, f_i) = \nu_i^-([x_i, y_i]) < \varepsilon \cdot \frac{\varepsilon}{2},$$

由于马尔可夫不等式，以至少为 $1 - \frac{\varepsilon}{2}$ 的概率我们有

$$\operatorname{dist}(x_n, y_n) = \operatorname{Leb}([x_n, y_n]) = \nu_n^-([x_n, y_n]) < \varepsilon.$$

情况 $\nu_i^-([y_i, x_i]) < \frac{1}{2}\varepsilon^2$ 可以采用相同的方法处理。 □

粗略地说，引理 4.4 表明，如果 x 和 y 的随机迭代在测度 ν_i^- 的意义下彼此趋近，那么在时刻 n，它们在通常意义下也很可能保持足够接近。接下来的观察是，测度 ν_i^- 是无原子的，并且这一性质对于 i 和测度 μ_j 的可能选择都是均匀成立的。具体而言，我们有以下引理。

引理 4.5. 存在 $\varepsilon_1 > 0$，使得对于任意 n、i 和 μ_1、.. .、μ_n，以及任意长度为 $|J| < \varepsilon_1$ 的区间 J，我们有 $\nu_i^-(J) < \frac{1}{2}\varepsilon^2$。

证明。首先注意，与逆映射对应的测度 μ^- 也满足测度条件。事实上，如果对于某个$\mu \in \mathbf{K}$，存在 μ-几乎处处的 f 满足 $f_*^{-1}\nu = \nu'$，那么我们也会有 $f_*\nu' = \nu$，这将导致矛盾。

现在，[GK2] 中的原子消散定理 1.13 适用，该定理表明，存在某个 k_0，使得对于任意 μ_1、.....、$\mu_{k_0} \in \mathbf{K}$ 以及任意 x'，$x'' \in \mathbb{RP}^1$

$$(\mu_1^- * \cdots * \mu_{k_0}^- * \delta_{x'})(\{x''\}) < \frac{1}{2}\varepsilon^2. \tag{12}$$

换句话说，任意（反向）长度为 k_0 的卷积都会使原子的权重消散，从而不超过 $\frac{1}{2}\varepsilon^2$。这意味着，对于某个 ε'，有

$$(\mu_1^- * \cdots * \mu_{k_0}^- * \delta_{x'})(J) < \frac{1}{2}\varepsilon^2 \quad \forall J,\, |J| \le \varepsilon'. \tag{13}$$

事实上，如果对于某个 $\varepsilon'_m = \frac{1}{m}$ 该陈述不成立，那么对于某些 J_m、$\mu_{1,\,m}$、.....、$\mu_{k_0,\,m}$，我们将得到这些数据的一个收敛子列，

$$J_{m_j} \to x'', \quad \mu_{i,m_j} \to \mu_i \in \mathbf{K},$$

并将与（12）产生矛盾。对 x' 关于 $\nu^-_{i+k_0}$ 取（12）的平均值，我们便得到对于所有 $i \le n-k_0$ 成立的引理结论。

最后，根据定义，所有测度 ν^-_{n-k} for $k = 1, \dots, k_0$ 都是无原子的，$\nu^-_n =$ 勒贝格测度。同样，这意味着存在 ε''，使得对于任意$\mu_{n-k+1}, \dots, \mu_n \in$ K

$$(\mu^-_{n-k+1} * \cdots * \mu^-_n * \nu^-_n)(J) < \frac{1}{2}\varepsilon^2 \quad \forall J, |J| \le \varepsilon''. \tag{14}$$

事实上，否则我们可以找到区间 J_m 的一个收敛子列以及测度$\mu^-_{n-j} \in$K，从而与上述无原子性矛盾。

取 $\varepsilon_1 := \min(\varepsilon', \varepsilon'')$，即可完成引理 4.5 的证明。 □

结合引理 4.5 和引理 4.4 的结论，可得以下推论：

推论 4.6。假设对于某个 i，经过 i 次迭代后，我们有 dist$(x_i, y_i) < \varepsilon_1$。那么，以至少 $1 - \frac{\varepsilon}{2}$ 的概率，我们有

$$\operatorname{dist}(x_n, y_n) < \varepsilon.$$

现在，我们可以通过考察两种不同情形来完成命题 4.3 的证明。首先，假设存在 $\varepsilon_2 > 0$ 和 k_2，使得对于任意x'、$y' \in \mathbb{RP}^1$ 以及任意 $\mu_1, \dots, \mu_{k_2} \in$ K，以至少 ε_2的概率，以下事件成立：

$$\exists i \le k_2 : \quad \operatorname{dist}(x'_i, y'_i) \le \varepsilon_1,$$

其中

$$x'_0 := x',\, y'_0 := y', \quad x'_i = f_i(x'_{i-1}),\, y'_i = f_i(y'_{i-1}),$$

且 f_i 是关于 μ_i独立选取的随机映射。

在这种情况下，对于任意 k_3，两个初始点 x和 y 在 $k_2 \cdot k_3$ 次迭代过程中彼此靠近的距离始终不小于 ε_1 的概率不超过 $(1-\varepsilon_2)^{k_3}$：我们有 k_3次尝试，且每次尝试成功的概率至少为 ε_2。取充分大的 k_3，使得 $(1-\varepsilon_2)^{k_3} < \frac{\varepsilon}{2}$，并定义 $n := k_2 \cdot k_3$，则以至少 $1 - \frac{\varepsilon}{2}$的概率，存在 $i \le n$ 使得

$$\operatorname{dist}(x_i, y_i) < \varepsilon_1,$$

因此，dist$(x_n, y_n) < \varepsilon$，其概率至少为 $(1 - \frac{\varepsilon}{2})^2 > 1 - \varepsilon$。

最后，如果不存在这样的 ε_2 和 k_2，我们就取一列候选对象 $\varepsilon_{2,\,(m)} = \frac{1}{m}$, $k_{2,\,(m)} = m$，并考虑点 $x'_{(m)}$, $y'_{(m)}$ 和测度 $\mu_{1,\,(m)}, \dots, \mu_{m,\,(m)} \in$K，使得所需结论不成立。提取一个收敛子列后，我们得到 x', y' 和 $\mu_1, \mu_2, \dots$，使得几乎必然地，像 x'_i, y'_i 始终至少相距 ε'。

现在，原子溶解定理1.13 from [GK2] 与我们已经使用过的紧致性论证一起表明，对于某个 k_4, $\varepsilon_3 > 0$，对于任意长度为 $|J| \le \varepsilon_3$ 的区间 J，我们有

$$\mathbb{P}(x_{k_4} \in J), \mathbb{P}(y_{k_4} \in J) \le \frac{1}{3}.$$

特别地，对于任意 $i > k_4$，矩阵乘积 $T_{[k_4,i]} := A_i \dots A_{k_4+1}$的范数都不能高于$\frac{4}{\varepsilon_3 \varepsilon_2}$；否则，以至少 $1\,3$ 的概率，两个点x_{k_4}、y_{k_4}到其最强收缩方向的距离都至少为 $\frac{\varepsilon_3}{2}$，因此它们的像

$$x_i = f_{T_{[k_4,i]}}(x_{k_4}), \quad y_i = f_{T_{[k_4,i]}}(y_{k_4})$$

到最强扩张方向的像的距离至多为 $\frac{1}{\|T_{[k_4,i]}\|\cdot\frac{\varepsilon_3}{2}} < \frac{\varepsilon_2}{2}$，因此彼此之间的距离小于 ε_2。

因此，乘积 $A_i \dots A_{k_4+1}$ 几乎必然满足以 $\frac{4}{\varepsilon_3\varepsilon_2}$ 为上界的统一界。因此，矩阵 A_i 的范数也均匀有界。

现在考虑紧集 K′，它是 $\{\mu_{k_4+1}, \mu_{k_4+2}, \dots\}$ 的闭包。这是 K 的闭子集，因此也是一个紧的测度集，其上矩阵范数均匀有界。因此，我们再次满足非平稳 Furstenberg 定理（[GK2，定理 1.4]）的假设。于是，乘积范数的均匀界 $T_{[k_4,i]}$ 使我们得到矛盾，因此第二种情况不可能发生。

至此，命题 4.3 的证明完成。 □

注 4.7. 在高维情况下，如果没有额外假设（即有限不变平面集不存在这一条件的类似条件），测度条件不足以确保非平稳遍历性，如 [GK2, ExampleA.1] 所示。尽管如此，在一些额外假设下（这些假设足以保证“李雅普诺夫谱的简单性”的非平稳版本，或至少保证“第一李雅普诺夫指数的简单性”），本命题的结论可以推广到 $SL(d,\mathbb{R})$，$d > 2$ 的情况。

注 4.8. [A], [KN], [Mal] 已确立并详细研究了一般平稳动力系统下圆上的随机轨道收缩。尽管在命题 4.3 的证明中，我们有些地方利用了所复合的是射影映射这一事实，但我们预计，在适当假设下，该结论及其证明可以推广到圆同胚的更一般非平稳动力系统情形。

5. Proof of Theorem 3.1

我们将首先证明定理的第二部分，即大偏差估计；由博雷尔–坎泰利引理可知，简单应用该引理即可推出第一部分。

对于任意闭区间 $J \subset \mathbb{R}$，记 $\mathcal{M}(J)$ 为赋予 Wasserstein 度量的 J 上博雷尔概率测度空间。

任取 $\varphi \in C(X,\mathbb{R})$，并令 $M = \max(1, \max_{x\in X} |\varphi(x)|)$。连续映射 $\varphi\colon X \to \mathbb{R}$ 在博雷尔概率测度空间上诱导出映射 $\varphi_*\colon \mathcal{M}(X) \to \mathcal{M}([-M,M])$。

引理 5.1。映射 $\varphi_*\colon \mathcal{M}(X) \to \mathcal{M}([-M,M])$ 是连续的。

证明。事实上，由于 X 是紧的，因此 φ 在 X 上一致连续，故存在 $\delta > 0$，使得 $d_X(x,y) < \delta$ 蕴含 $|\varphi(x) - \varphi(y)| < \frac{\varepsilon'}{2}$。另一方面，取 $\varepsilon'' := \frac{\varepsilon'\delta}{4M}$。如果两个测度 ν, ν' 之间存在某个耦合 γ，使得积分 (2) 的值小于 ε''，那么根据马尔可夫不等式

$$\gamma(\{(x,y) \mid d_X(x,y) > \delta\}) < \frac{\varepsilon''}{\delta} = \frac{\varepsilon'}{4M},$$

因此，利用推前 $\gamma' := (\varphi,\varphi)_*\gamma$ 作为 $\varphi_*\nu$ 与 $\varphi_*\nu'$ 之间的耦合，我们有

$$\begin{aligned}\iint |u-v|\,d\gamma'(u,v) &= \iint_{X\times X} |\varphi(x)-\varphi(y)|\,d\gamma(x,y) \le \\ &\le \frac{\varepsilon'}{2}\cdot\gamma(\{(x,y)\mid d_X(x,y)\le\delta\}) + 2M\cdot\gamma(\{(x,y)\mid d_X(x,y)>\delta\}) \\ &< \frac{\varepsilon'}{2}\cdot 1 + 2M\cdot\frac{\varepsilon'}{4M} = \frac{\varepsilon'}{2}+\frac{\varepsilon'}{2} = \varepsilon'.\end{aligned}$$

因此，不等式 $\mathrm{dist}_{\mathcal{M}}(\nu,\nu') < \varepsilon''$ 蕴含 $\mathrm{dist}_{\mathcal{M}([-M,M])}(\varphi_*\nu,\varphi_*\nu') < \varepsilon'$，这就证明了引理 5.1。□

由于 $\mathcal{M}(X)$ 是紧的，引理 5.1 蕴含 φ_* 连续且均匀。结合基本假设可知，对于任意 $\varepsilon' > 0$，存在 m，使得对于任意 n 以及任意初始测度 $\nu\in\mathcal{M}$,

$$\mathrm{dist}_{\mathcal{M}([-M,M])}(\varphi_*(\mu_{n+m} * \ldots * \mu_{n+1} * \nu), \varphi_*\nu_{n+m}) < \varepsilon', \tag{15}$$

其中测度序列 $\{\nu_i\}_{i\ge 0}$ 由式 (4) 给出。

现在，固定 $\varepsilon > 0$，并为这个 ε 推导大偏差估计。固定一个足够大的 m，以确保式 (15) 在 $\varepsilon' = \frac{\varepsilon^2}{30M}$ 下成立。

由于 φ 有界，只需对所有能被 m 整除的 $n\ge 1$ 证明式 (5)。对于任意 $n = mq$、$q\in\mathbb{N}$，将迭代 $k = 1, \ldots, mq$ 分解为 m 个等差数列 $k = r + jm$、$j = 0$、…、$q-1$，其公差为 m，并由剩余类 $r = 1$、…、m 编号。

只需验证大偏差估计对这些子序列中的每一个都成立。确实，

$$\begin{aligned}\frac{1}{n}\left|\sum_{k=1}^{n}\varphi(f_k\circ\ldots\circ f_1(x)) - \sum_{k=1}^{n}\int_X \varphi d\nu_k\right| \le \\ \le \frac{1}{m}\sum_{r=1}^{m}\left|\frac{1}{q}\left(\sum_{j=0}^{q-1}\varphi(f_{jm+r}\circ\ldots\circ f_1(x)) - \sum_{j=0}^{q-1}\int_X\varphi d\nu_{jm+r}\right)\right|.\end{aligned}$$

因此，只需证明对于给定的 $r = 1, 2, \ldots, m$，对所有$q\ge 1$，都有

$$\mathbb{P}\left(\frac{1}{q}\left|\sum_{j=0}^{q-1}\varphi(f_{jm+r}\circ\ldots\circ f_1(x)) - \sum_{j=0}^{q-1}\int_X\varphi d\nu_{jm+r}\right| > \varepsilon\right) < \frac{C}{m}e^{-\delta n}, \tag{16}$$

因为这将推出 (5)。

固定 $r\in\{1,2,\ldots,m\}$，并记

$$y_j := x_{r+mj} = f_{r+mj}\circ\cdots\circ f_1(x);$$

那么，(16) 中的第一个和为 $\sum_{j=0}^{q-1}\varphi(y_j)$。

为了建立估计式 (16)，我们将把这个和与一组独立随机变量进行比较，这些变量分别关于 $\varphi_*\nu_{jm+r}$ 分布。ξ_j 为构造上的技术原因，我们将在每次迭代中额外添加两个随机变量，它们在 $[0,1]$ 上均匀分布，并且

与各次迭代彼此独立，且相互独立。具体而言，我们考虑

$$\Omega := \{(f_n, z_n, z'_n)_{n\in\mathbb{N}}\} = C(X,X)^{\mathbb{N}} \times [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}},$$

赋予其概率测度

$$\tilde{\mathbb{P}} := \mathbb{P} \times \mathrm{Leb}^{\mathbb{N}} \times \mathrm{Leb}^{\mathbb{N}},$$

其中 $\mathbb{P}$ 由（1）定义。

引理 5.2。在空间 $(\Omega\ \tilde{\mathbb{P}})$ 上存在随机变量 ξ_j、ξ'_j，使得：

对于所有 j，• ξ_j 相互独立，且 ξ_j 的分布为 $\varphi_*\nu_{r+jm}$；
• ξ'_j 是独立同分布的伯努利变量，以概率 $\frac{\varepsilon}{8M}$ 取值 1。
•对于每个 j，如果 ξ'_j取值为 0，那么 $|\varphi(y_j) - \xi_j| \le \frac{\varepsilon}{3}$。

我们将这个引理（略微技术性的）证明推迟到本节末尾。注意，这使得我们可以完成定理 3.1 的证明。事实上，首先注意到引理 5.2 的最后一个结论推出了不等式

$$|\varphi(y_j) - \xi_j| \le \frac{\varepsilon}{3} + 2M\xi'_j; \tag{17}$$

注意，这里我们用到了 $|\varphi| \le M$ 这一事实，因此 $|\varphi(y_j) - \xi_j| \le 2M$。(16) 中的和可以估计为

$$\begin{aligned}\frac{1}{q}\left|\sum_{j=0}^{q-1}\varphi(f_{jm+r}\circ\ldots\circ f_1(x)) - \sum_{j=0}^{q-1}\int_X \varphi d\nu_{jm+r}\right| = \\ = \frac{1}{q}\left|\sum_{j=0}^{q-1}\varphi(y_j) - \sum_{j=0}^{q-1}\mathbb{E}\xi_j\right| \le \frac{1}{q}\sum_{j=0}^{q-1}|\varphi(y_j)-\xi_j| + \frac{1}{q}\left|\sum_{j=0}^{q-1}(\xi_j - \mathbb{E}\xi_j)\right| \le \\ \le \frac{\varepsilon}{3} + 2M\cdot\frac{1}{q}\sum_{j=0}^{q-1}\xi'_j + \frac{1}{q}\left|\sum_{j=0}^{q-1}(\xi_j - \mathbb{E}\xi_j)\right|\end{aligned} \tag{18}$$

现在注意，由于 $\mathbb{E}\xi'_j = \frac{\varepsilon}{8M} < \frac{\varepsilon}{6M}$，有

$$\tilde{\mathbb{P}}\left(\frac{1}{q}\sum_{j=0}^{q-1}\xi'_j > \frac{\varepsilon}{6M}\right) < C_1 e^{-\delta_1 q}$$

对于某个 $\delta_1 > 0$，这是由独立变量 ξ'_j 的大偏差估计得到的。同理，由于 ξ_j 相互独立且均被 M 所界，大偏差估计给出

$$\tilde{\mathbb{P}}\left(\left|\frac{1}{q}\sum_{j=0}^{q-1}(\xi_j - \mathbb{E}\xi_j)\right| > \frac{\varepsilon}{3}\right) < C_2 e^{-\delta_2 q}$$

现在，如果 $\frac{1}{q}\sum_{j=0}^{q-1}\xi'_j \le \frac{\varepsilon}{6M}$和 $\frac{1}{q}\left|\sum_{j=0}^{q-1}(\xi_j - \mathbb{E}\xi_j)\right| < \frac{\varepsilon}{3}$都成立，则由 (18) 可得

$$\frac{1}{q}\left|\sum_{j=0}^{q-1}\varphi(y_j) - \sum_{j=0}^{q-1}\mathbb{E}\xi_j\right| \le \frac{\varepsilon}{3} + 2M\cdot\frac{\varepsilon}{6M} + \frac{\varepsilon}{3} = \varepsilon.$$

因此，(16) 中的概率上界为

$$\mathbb{P}\left(\frac{1}{q}\left|\sum_{j=0}^{q-1}\varphi(y_j)-\sum_{j=0}^{q-1}\mathbb{E}\xi_j\right|>\varepsilon\right)\leq$$
$$\leq\tilde{\mathbb{P}}\left(\frac{1}{q}\sum_{j=0}^{q-1}\xi_j'>\frac{\varepsilon}{6M}\right)+\tilde{\mathbb{P}}\left(\left|\frac{1}{q}\sum_{j=0}^{q-1}(\xi_j-\mathbb{E}\xi_j)\right|>\frac{\varepsilon}{3}\right)$$
$$<C_1e^{-\delta_1 q}+C_2e^{-\delta_2 q},$$

因此，我们得到了所需的指数级小界 $\delta=\min(\delta_1,\delta_2)$。

由于我们已经得到了所需的界 (16)，且它对每个 $r=1,\dots,m$ 都成立，因此除了引理 5.2 外，定理 3.1 的证明已经完成。

引理 5.2 的证明。在空间 $(\Omega\ \tilde{\mathbb{P}})$ 上，定义一个 σ-代数序列

$$\mathcal{F}_j:=\sigma(\{(f_k,z_k,z_k')\mid k\leq r+jm\}).$$

注意，随机值 $\varphi(y_i)=\varphi(x_{r+mi})$，$i=0,..,j-1$ 关于 $\mathcal{F}_{j-1}$ 可测；同时，$\varphi(y_j)$ 关于这个 σ- 代数的条件分布为

$$D_{j,y_{j-1}}:=\varphi_*(\mu_{r+mj}*\dots*\mu_{r+1+m(j-1)}*\delta_{y_{j-1}}).$$

由于 m 的选取，这个条件分布与 $\varphi_*\nu_{r+jm}$ 的距离为 ε'，其中 $\varphi_*\nu_{r+jm}$ 是所需随机变量 ξ_j 的分布。这启发我们按以下方式构造 ξ_j 和 ξ_j'。假设

(i) 对每个 j，随机变量 ξ_j、ξ_j' 关于 $\mathcal{F}_j$ 可测；(ii) ξ_j 关于 $\{(f_k,z_k,z_k')\mid k\leq r+(j-1)m\}$ 的条件分布为 $\varphi_*\nu_{r+jm}$；(iii) ξ_j' 关于 $\{(f_k,z_k,z_k')\mid k\leq r+(j-1)m\}$ 的条件分布为成功概率为 $\frac{\varepsilon}{8M}$ 的伯努利分布；(iv) 若 $|\varphi(y_j)-\xi_j|>\frac{\varepsilon}{3}$，则 ξ_j' 取值为 1。

那么，这些随机变量满足引理 5.2 的条件。事实上，性质 (ii) 表明 ξ_j 与 σ-代数 $\mathcal{F}_{j-1}$ 独立。由于根据性质 (i)，ξ_0、.....、ξ_{j-1} 关于 $\mathcal{F}_{j-1}$ 可测，因此随机变量 ξ_j 与 ξ_0、.....、ξ_{j-1} 独立。

同样，由于 (i) 和 (iii)，ξ_j' 与 ξ_0'、...、ξ_{j-1}' 相互独立。引理的最后一个结论将由 (iv) 得出。

现在构造满足条件 (i)–(iv) 的随机变量 ξ_j、ξ_j'。

回顾一下，在实数轴上的两个概率测度之间，可以显式地构造最优传输。具体而言，假设给定实数轴上的两个概率测度 ρ,ρ'。取两个独立随机变量 η,ζ，其中 η 关于 ρ 服从分布，而 ζ 在 $[0,1]$ 上均匀分布。令

$$\Phi_\rho(y,z):=\rho((-\infty,y))+z\cdot\rho(\{y\});$$

于是，$\Phi(\eta,\zeta)$ 是一个均匀分布的随机变量（这就是将随机变量代入其自身分布函数的标准原理，其中使用独立随机变量来“消解原子”）。另一方面，函数

$$\Psi_{\rho'}(s)=\sup\{y:\rho'((-\infty,y))\leq s\},$$

该函数是 ρ' 的分布函数的逆函数，它将勒贝格测度映射到 ρ'。因此，它们的复合函数 $F_{(\rho,\rho')} := \Psi_{\rho'} \circ \Phi_\rho$ 在 ρ 和 ρ' 之间提供了一个耦合。也就是说，随机变量 $\xi := F_{(\rho,\rho')}(\eta,\zeta)$ 关于 ρ' 分布，而所需的耦合由联合分布 (η,ξ) 给出。该耦合使Wasserstein距离 [V] 最小。

现在，对于每个 j，考虑随机变量

$$\xi_j := F_{(D_{j,y_{j-1}},\varphi_*\nu_{jm+r})}(\varphi(y_j), z_{r+jm}).$$

由构造可知，性质 (i) 和 (ii) 均成立。

由于该传输的最优性以及关于 m 的假设，对于任意 y_{j-1}，在条件下我们有

$$\mathbb{E}(|\xi_j - \varphi(y_j)| \mid y_{j-1}) = \operatorname{dist}_{\mathcal{M}([-M,M])}(D_{j,y_{j-1}}, \varphi_*\nu_{jm+r}) \le \varepsilon'.$$

现在，考虑事件

$$A_j := \left\{ |\xi_j - \varphi(y_j)| > \frac{\varepsilon}{3} \right\}.$$

根据马尔可夫不等式，对于任意 y_{j-1}，在条件下，其概率不超过

$$\tilde{\mathbb{P}}(A_j \mid y_{j-1}) \le \frac{3}{\varepsilon} \cdot \mathbb{E}(|\xi_j - \varphi(y_j)| \mid y_{j-1}) \le \frac{3\varepsilon'}{\varepsilon} = \frac{3}{\varepsilon} \cdot \frac{\varepsilon^2}{30M} < \frac{\varepsilon}{8M}.$$

现在，定义 ξ'_j 为

$$\xi'_j = \mathbb{1}_{A_j} + \mathbb{1}_{|\xi_j - \varphi(y_j)| \le \frac{\varepsilon}{3}} \cdot \mathbb{1}_{z_{r+jm} \le \psi(j, y_{j-1})}, \tag{19}$$

其中，值

$$\psi(j, y_{j-1}) := \frac{\frac{\varepsilon}{8M} - \tilde{\mathbb{P}}(A_j \mid y_{j-1})}{1 - \tilde{\mathbb{P}}(A_j \mid y_{j-1})} \in (0,1)$$

的选取方式使得

$$\tilde{\mathbb{P}}(\xi'_j \mid y_{j-1}) = \frac{\varepsilon}{8M}$$

对于 y_{j-1} 的任意取值均成立。选取式 (19) 确保了性质 (iii)（由显式公式可知其可测性）和性质 (iv)（这可直接由构造得出）。

至此，随机变量 ξ_j、ξ'_j 的构造及其性质的验证均已完成，从而引理 5.2 得证。□

6. 结论

6.1. 不同时刻的不同相空间。一个一般性的说明是，在非平稳情形下，不同迭代时刻的相空间不再具有彼此之间的典范同一性。具体而言，考虑一条随机轨道

$$x_n = f_n(x_{n-1}).$$

应用时变变量替换 $y_n = h_n(x_n)$，其中每个h_n 都是 X的同胚，将初始系统的一个轨道映射到

$$y_n = g_n(y_{n-1}), \quad g_n = h_n \circ f_n \circ h_{n-1}^{-1}.$$

这对应于同一类别中另一个系统的随机迭代，其测度 $\mu'_n = (T_n)_*\mu_n$ 是 μ_n 在“变量替换”映射下的推前像

$$T_n : C(X,X) \to C(X,X), \quad T_n : f \mapsto h_n \circ f_n \circ h_{n-1}^{-1}.$$

然而，某些 x_n 与 x_m 相等，并不意味着它们的像 $y_n = h_n(x_n)$ 和 $y_m = h_m(x_m)$ 相等。这就是为什么可以（也应该）把不同时刻的空间 X 看作不同的纤维。

此外，在缺乏平稳性的情况下，我们不必假设相空间始终保持不变。相反，我们可以考虑一列紧致度量空间 X_0、X_1、X_2、.....，以及一列概率测度 μ_n、$n \geq 1$，其定义在 $C(X_{n-1}、X_n)$ 上：

$$X_0 \xrightarrow[f_1]{\mu_1} X_1 \xrightarrow[f_2]{\mu_2} X_2 \to \cdots \to X_{n-1} \xrightarrow[f_n]{\mu_n} X_n \to \dots.$$

然后，给定一列“可观测”函数 $\varphi_n \in C(X_n、\mathbb{R})$，我们可以提出与之前相同的问题。

要求 X_n 的直径具有一致界，并要求函数 φ_n 具有一致连续模，是合理的；于是，定理 3.1 可以推广到如下情形：

定理 6.1。假设 X_n 的直径均有界，并且基本假设（将 X 替换为相应的 X_n）成立。那么，对于任意一列 $\varphi_n \in C(X_n、\mathbb{R})$ 函数（其具有一致连续模）以及任意 $x \in X_0$，我们有

$$\frac{1}{n}\left|\sum_{k=1}^{n} \varphi_k(f_k \circ \ldots \circ f_1(x)) - \sum_{k=1}^{n} \int_{X_k} \varphi_k d\nu_k\right| \to 0 \quad \text{as} \quad n \to \infty,$$

其中，对于所有 $n \geq 1$，$\nu_n = \mu_n * \nu_{n-1}$ 均成立，而 ν_0 是 X_0 上任意的初始博雷尔概率测度。

此外，大偏差定理的一个类似结果也成立：对于任意 $\varepsilon > 0$，存在 C、$\delta > 0$，使得对于任意 $x \in X_0$

$$\forall n \in \mathbb{N}, \quad \mathbb{P}\left(\frac{1}{n}\left|\sum_{k=1}^{n} \varphi_k(f_k \circ \ldots \circ f_1(x)) - \sum_{k=1}^{n} \int_{X_k} \varphi_k d\nu_k\right| > \varepsilon\right) < C\exp(-\delta n).$$

定理 6.1 的证明几乎完全重复了定理 3.1 的证明，因此这里省略。

6.2. 反例：无法避免的假设。本节旨在给出非平稳随机遍历定理不成立的例子。第一个例子说明了为什么基本假设必须对“扩散”初始测度所需的迭代次数至少设定某种上界：

例 6.2. 设集合 $X = \{a, b\}$ 仅由两个点组成。固定一个（增长足够快的）序列 $n_k = 10^{k^2}$，并令测度 μ_n 定义为 $\mu_n = \delta_{\mathrm{id}}$，其中 $n \neq n_k$，且

$$\mu_{n_k} = \frac{1}{2}\delta_{\mathrm{id}} + \frac{1}{2}\delta_{\sigma},$$

其中，σ 是一个对换，它交换 a 和 b。那么，一方面，对于任意初始 n、X上的任意两个测度 ν和ν'，以及任意 $n_k > n$，有

$$\mu_{n_k} * \ldots * \mu_{n+1} * \nu' = \mu_{n_k} * \ldots * \mu_{n+1} * \nu = \frac{1}{2}\delta_a + \frac{1}{2}\delta_b =: \overline{\nu}.$$

另一方面，任意单条轨道（x_n）在 n_k和 n_{k+1} 之间的每个时间区间内，要么始终处于 a，要么始终处于 b，且这两种情况发生的概率相同

，并且对于不同的 k，这些情况彼此独立。当商 $\frac{n_{k+1}-n_k}{n_k}$ 趋于无穷时，很容易得出，任意函数 φ 的时间平均几乎必然以 $\varphi(a)$ 和 $\varphi(b)$ 为聚点。特别是，这些时间平均（对于非恒定的 φ）不收敛。

下一个例子说明了为什么在基本假设中我们避免进行时间平均。事实上，不同时间的相空间之间缺乏自然的对应关系（上一节已对此进行讨论），这已经表明，对测度的像进行时间平均并不是一个适合考虑的对象：因为这要求对不同空间上的测度进行相加。然而，即使我们处理的是一个唯一遍历的确定性动力系统，从而可以进行这种相加，对于不满足基本假设的依赖于 n 的 φ_n 而言，非平稳遍历定理也可能不成立：

例 6.3. 令 $X = S^1 = \mathbb{R}/\mathbb{Z}$，且 $f(x) = x + \alpha$ 模 1 为无理旋转。那么，动力系统 (X, f) 是唯一遍历的（其唯一不变测度是勒贝格测度）；然而，基本假设并不成立（不存在随机性，且轨道之间的距离不会减小）。与此同时，圆上任意两个初始测度的像的切萨罗平均都以某个一致速率收敛到勒贝格测度。

现在，任取一个函数 $\varphi \in C(X)$，并考虑函数族 $\varphi_n = \varphi \circ f^{-n}$。这个函数族在 X 上等度连续。同时，定理 6.1 的结论对这些函数并不成立：对于任意初始点 $x \in X$，有

$$\frac{1}{n}\sum_{k=1}^{n} \varphi_k(f_k \circ \ldots \circ f_1(x)) = \frac{1}{n}\sum_{k=1}^{n} \varphi \circ f^{-k}(f^k(x)) = \varphi(x),$$

因此不存在类似常数的行为。此外，将函数的选取替换为

$$\varphi_n = \varphi \circ f^{-n+r(n)},$$

其中 $r(n) := \max\{k \mid n > n_k\}$，对于一个快速增长的序列 $n_k = 10^{k^2}$，与例 6.2 类似，每个初始点 x 对应的“非平稳”时间平均都不存在极限。

Acknowledgments

我们感谢 Andrew T¨or¨ok 对本文初稿提出的意见。

References

[A] V.A. Antonov，《周期演化过程建模：随机信号同步》。
Vestnik Leningrad. Univ. Mat. Mekh. Astronom. 1984，第2期，第67–76页。[BB] D. Berend、V. Bergelson，《变换的遍历与混合序列》，
Ergodic Theory and Dynamical Systems **4** 1984 353 366 ，（），第–页。[CR] J.-P. Conze、A. Raugi，《圆上的序列扩张动力系统极限定理》，[0, 1]，
Ergodic theory and related fields，第430卷，美国数学学会，普罗维登斯，罗得岛州，2007年，第89–121页。[DK] B. Deroin、V. Kleptsyn，《随机共形动力系统》，
Geom. Funct. Anal.（），第期，第–页。**17** 2007 4, 1043 1105
[DKN] B. Deroin、V. Kleptsyn、A. Navas，《关于圆上极小群作用的遍历性问题》，
Moscow Mathematical Journal，9（2009），第263–303页。[FL] D. Feng、K. Lau，《弱分离条件下自相似测度的多重分形形式》，*J. Math. Pures Appl.* **9**（2009），第407–428页。

[FS] D. Feng、K. Simon，《迭代函数系统与排斥子集的维数估计：综述》，
Thermodynamic formalism，第421–467页，《数学讲义》，第2290卷，CIRM Jean-Morlet 系列，施普林格，Cham，2021年。[Hu] J. Hutchinson，《分形与自相似性》，
Indiana Univ. Math. J. **30**（1981），第713–747页。[FurK] H. Furstenberg、H. Kesten，《随机矩阵乘积》，*Ann. Math. Statist.* **31**（1960），第457–469页。[FurKif] H. Furstenberg、Y. Kifer，《随机矩阵乘积与射影空间上的测度》，*Israel Journal of Mathematics* **46**（1983），第12–32页。[G] I. Goldsheid，《非平稳马尔可夫相关矩阵乘积的指数增长》，
International Mathematics Research Notices，第2022卷，第8期，2022年4月，第6310–6346页。[GK1] A. Gorodetski、V. Kleptsyn，《随机 $SL(2,\mathbb{R})$矩阵乘积的参数化 Furstenberg 定理》，
Adv. Math. **378**（2021），论文编号107522，共81页。[GK2] A. Gorodetski、V. Kleptsyn，《随机矩阵乘积 Furstenberg 定理的非平稳版本》，预印本，arXiv:2210.03805[GK3] A. Gorodetski、V. Kleptsyn，《Anderson 局域化的非平稳版本与参数化 Furstenberg 定理》，进行中的研究。[GKM] A. Gorodetski、V. Kleptsyn、G. Monakov，《平稳测度的 H¨older 正则性》，arXiv:2209.12342[KN] V. Kleptsyn、M. Nalskii，《圆上随机动力系统中轨道的收敛》，
*Funct. Anal. Appl.*第38卷（2004），第4期，第267–282页。[HNTV] N. Haydn、M. Nicol、A. T¨ör¨ok、S. Vaienti，《序列与非平稳动力系统的几乎处处不变性原理》，
Trans. Amer. Math. Soc. **369**（2017），第5293–5316页。[KS] L. Koralov、Ya. Sinai，《概率论与随机过程理论》，*Springer-Verlag Berlin Heidelberg* 2007，+353，
Universitext，xii页。[La] Thierry de La Rue，《遍历论中的联接》，*Encyclopedia of Complexity and Systems Science* 2020，施普林格，
柏林、海德堡。[Mal] D. Malicet，《Homeo（S^1）上的随机游走》，*Comm. Math. Phys.* **356**（2017），第1083–1116页。[PSS] Y. Peres、W. Schlag、B. Solomyak，《伯努利卷积六十年》，*Fractal geometry and stochastics II* 1998 39 65，（
Greifswald/Koserow，），第–页，《概率进展》，第46卷，Birkhäuser，巴塞尔，2000年。[Sz] T. Szarek，《迭代函数系统的不变测度》，*Ann. Polon. Math.* **75**（2000），第87–98页。[V] S. Vallander，《计算直线上概率分布之间的 Wasserstein距离》，
Theory of Probability and its Applications，第18卷（1974），第784–786页。[Va] P. Varju，《伯努利卷积的最新进展》，*European Congress of Mathematics* 847 867 *Eur. Math. Soc. Zürich* 2018，第–页，，，

Department of Mathematics, University of California, Irvine, CA 92697, USA
Email address: asgor@uci.edu

CNRS, Institute of Mathematical Research of Rennes, IRMAR, UMR 6625 du CNRS
Email address: victor.kleptsyn@univ-rennes1.fr