arXiv:2305.02730v1 [math.NT] 4 May 2023

PSL2(R)商空间中的多项式序列的等分布

LAURITZ STRECK

Abstract. 本文证明，对于每个格 $\Gamma \subset PSL_2(\mathbb{R})$，都存在一个 $c > 0$，使得对于任意 $0 \leq \gamma < c$，序列 $ph(n^{1+\gamma})$ 对任意 $p \in \Gamma\backslash PSL_2(\mathbb{R})$都等分布，其中 h 是 horocycle 流。这在 Shah 的一个猜想方面取得了些许进展，并推广了 Venkatesh 的一个结果（“稀疏等分布问题、周期界与次凸性”，2005年）；该结果证明了对于余紧格，相同的等分布性质成立。证明利用了这样一种二分性：要么有良好的等分布估计，要么 $\{ph(t)\ t \leq T\}$可由周期较小的闭合 horocycle 逼近。

1. Introduction

考虑带有哈尔测度 μ_G 的（乘法）群 $G := PSL_2(\mathbb{R})$。格 $\Gamma \subset G$ 是一个离散子群，使得商空间 $X := \Gamma\backslash G$ 在 G 中有一个哈尔测度有限的基本域。于是，哈尔测度诱导出有限测度 μ_X。我们定义矩阵

$$h(x) := \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \quad a(y) := \begin{pmatrix} y^{\frac{1}{2}} & 0 \\ 0 & y^{-\frac{1}{2}} \end{pmatrix}.$$

$p \in X$ 在时刻 t 的测地流定义为 $g_t(p) := pa(e^t)$，而在时刻 t 的 horocycle 流定义为 $h_t(p) := ph(t)$。

虽然轨道 $g_t(p)$ 对于 $t \to \infty$ 的行为可能因初始点而相当不规则，但 horocycle 轨道 $h_t(p)$ 的行为则已知要规整得多。在详细介绍已知结果之前，我们先明确一些记号。我们称轨道 $h_t(p)$ 关于 μ_X 等分布，若对于 X上的任意紧支撑连续函数 f，

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(ph(t))\, dt \to \int f\, d\mu_X.$$

日期：2023年5月5日。作者获得了欧盟研究与创新计划地平线2020项下欧洲研究理事会（ERC）的资助（资助协议编号：803711）。

类似地，我们称该轨道沿序列 $a_n \in \mathbb{R}$关于 μ_X 等分布，若

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(ph(a_n))\, dt \to \int f\, d\mu_X.$$

最后，若存在一个 $t_0 \in \mathbb{R}$ 使得 $p = ph(t_0)$，则称点 $p \in X$ 为周期点。在这种情况下，horocycle 轨道将被困在周期轨道中，永远不会相对于 μ_X等分布；此时，系统 $t \mapsto ph(t)$ 与环面 $\mathbb{R}/\mathbb{Z}$上的圆周旋转 $x \mapsto x + t_0^{-1}$同构。下文中，我们用“$ph(a_n)$ 等分布”作为“对于所有非周期的 $p \in X$， $ph(a_n)$ 都相对于 μ_X等分布”的简写。Dani 和 Smillie 证明了，对于$t \in \mathbb{R}$， $ph(t)$ 等分布；对于 $n \in \mathbb{N}$， $ph(n)$ 也等分布。

随后，人们提出了这样的问题：对于 $\mathbb{N}$以外的序列，情况会如何。Margulis 猜想， p_n 为第 n 个素数时， $ph(p_n)$ 也应当等分布。Shah 猜想，对于任意 $\gamma \geq 0$， $ph(n^{1+\gamma})$ 都会等分布。我们指出，根据 Bourgain 在更一般背景下的研究 [1]，这些结果对于几乎所有 μ_X 都成立。真正的挑战在于证明对所有非周期的 $p \in X$ 都成立等分布。

Venkatesh 通过证明以下事实推进了 Shah 的猜想：对于余紧的 Γ，存在一个小的 $c = c(\Gamma) > 0$，使得对于所有 $0 \leq \gamma < c$ 和所有 $p \in X$， $ph(n^{1+\gamma})$ 都等分布于 [8]。他的证明通过控制形如 $ph(sn)$ 的算术数列来实现，其中 $n \in \{0, \ldots, N-1\}$，且 s 相对于 n 很小。控制这些稀疏序列还意味着，对于余紧的 Γ，几乎素数也是等分布的；也就是说，当 R 足够大时， $ph(q)$ 等分布，其中 q 遍历至多含有 R 个素因子的所有数。这种对稀疏序列的控制足以控制几乎素数，这一点既可以通过筛法看出，也可以通过 Goldston 和 Yilmaz 引入、随后被 Green 和 Tao 用于证明素数包含任意长等差数列的伪随机测度 ν 看出 [3],[4] （参见 [5]，其中使用筛法证明了几乎素数的等分布；另见 [6]，其中使用伪随机测度 ν 给出了证明）。

Sarnak 和 Ubis 证明，对于$\Gamma = PSL_2(\mathbb{Z})$，几乎素数是等分布的，而该对象并非余紧 [5]。随后，作者证明了几乎素数对于 Γ在 $PSL_2(\mathbb{R})$中的所有格 [6]都是等分布的。

本文在一般格的情形下，证明了 $ph(n^{1+\gamma})$对于较小的γ 是等分布的。这将 Venkatesh 关于余紧 Γ 的结果推广到了 $PSL_2(\mathbb{R})$中的所有格 Γ ，并在 Shah 的猜想方面取得了一些进展。

我们将在下文的结果中对这一点作出精确表述，该结果也是本文的主要结论。为此，我们还需要引入一些记号，并首先定义】【。

度量 d_X。群 $G = PSL_2(\mathbb{R})$ 配备了一个自然的左不变度量 d_G (例如，参见 [2]中的第9章)。该度量通过 $d_X(\Gamma g, \Gamma h) := \inf_{\gamma\in\Gamma} d_G(g, \gamma h)$ 诱导出 X 上的度量。我们还固定一个点$p_0 \in X$，并定义 $\mathrm{dist}(p) := d_X(p, p_0)$。

对于两个函数 $f, g\colon U \to \mathbb{R}$，如果存在常数 C，使得对所有 $x \in U$，都有 $|f(x)| \le C|g(x)|$，其中 U是某个定义域，我们记作 $f \ll g$或 $f = O(g)$。本文中，该定义中隐含的常数 C 始终允许依赖于格 Γ 和 γ的选取，但不依赖于其他任何因素。若 $f \ll g$ 且 $g \ll f$，我们记作 $f \sim g$。

对于函数 $f \in C^4(X)$，令 $\|f\|_{W^4}$为其在希尔伯特空间 $W^{4,\,2}$中的索博列夫范数，其中涉及四阶导数；令$\|f\|_{\infty,\,j}$为 j 阶导数的上确界范数。定义

$$\|f\| := \|f\|_{W^4} + \|f\|_{\infty,1} + \|f\|_{\infty,0};$$

这是 Str¨ombergsson 用来证明其等分布结果 [7]时所使用的同一个范数。令 β 为定理 2 中的常数；它最终来源于有效混合的速率。定理 1 中的常数可取为 $c = \frac{\beta}{600}$。

定理 1. 对于任意格 $\Gamma \subset PSL_2(\mathbb{R})$，存在常数 $c = c(\Gamma) > 0$，使得对于任意 $0 \le \gamma \le c$、任意非周期的 $p \in X$ 以及任意满足 $\|f\| = 1$ 的函数 $f \in C^4(X)$,

$$\left|\frac{1}{T}\sum_{n\le T} f\left(ph\left(n^{1+\gamma}\right)\right) - \int f\, d\mu_X\right| \ll r^{-\frac{\beta}{4}},$$

其中 $r = T^{1+\gamma}\exp(-\mathrm{dist}(g_{\log T^{1+\gamma}}(p)))$。由于 $r \to \infty$ 当 $T \to \infty$ 时，序列 $ph(n^{1+\gamma})$ 等分布。

为了证明定理 1，我们将把取值范围划分为不同的区间，并在每个区间上使用泰勒展开。在区间 $[T_0, T_1]$ 上，只要 T_0 不太小，且取值范围不太长，函数 $t^{1+\gamma}$ 就近似等于 $T_0^{1+\gamma} + (1+\gamma)T_0^{\gamma}(t - T_0)$。因此，问题就转化为研究当 $s \sim T^{\gamma}$ 时， $ph(ns)$ 的等分布情况。为了控制这些稀疏算术序列，我们需要两个结果。

第一个是以下定理，它是将 Str¨ombergsson 的等分布结果 [7]与 Venkatesh 的方法 [8], 相结合后直接得到的，Zheng [9]就曾采用这种做法。

定理 2（[9], 定理 1.2）。设 Γ 是 G 中的一个非紧格。设 $f \in C^4$（X），其中 $\|f\| < \infty$ 且 $1 \le s < T$。那么

$$\left|\frac{s}{T}\sum_{1\le j\le T/s} f(ph(sj)) - \int f\, d\mu_X\right| \ll s^{\frac{1}{2}} r^{-\frac{\beta}{2}}\|f\|$$

对于任意初始点 $p \in X$，其中 $r = T\exp(-\mathrm{dist}(g_{\log T}(p)))$。参数 $\frac{1}{6} > \beta > 0$ 和隐含常数仅取决于 Γ。

当 r 相较于 T 足够大时（例如，对于某个绝对常数 ϵ，有 $r \geq T^{\varepsilon}$ ），这一结果本身就足以证明序列 $ph(n^{1+\gamma})$ 的等分布。

下面的结果将用于处理等分布较差的情形。作者在 [6] 中证明了这一结果，以展示几乎素数的等分布。其证明运用了 Sarnak 和 Ubis [5] 的思想，并与 [7], 有相似之处；而 [7], 的证明则运用了可追溯至 Marina Ratner 的思想。该结果涵盖了摘要中提到的二分性。

引理 3（ [6] 中的引理 1.3）。设 $G = PSL_2(\mathbb{R})$ 中的 Γ 是一个格，并设 $X = \Gamma\backslash G$。设 $p \in X$ 和 $T \geq 0$。设 $\delta > 0$ 和 $K \leq T$。

存在一个大小为 $|I_0| \leq \delta^{-1}K^2$ 的区间 $I_0 \subset [0, T]$，使得：对于所有 $t_0 \in [0, T]\backslash I_0$，都存在一段闭测地线 $\{ph(t)\ t \leq K\}$，以 δ 阶的精度逼近 $\{ph(t_0 + t)\ 0 \leq t \leq K\}$，其含义是

$$\forall 0 \leq t \leq K: \quad d_X\left(ph(t_0 + t), ph(t)\right) \leq \delta.$$

这条闭测地线的周期 $P = P(t_0, p)$ 至多为 $P \ll r$，其中 $r = T\exp(-\text{dist}(g_{\log T}(p)))$.

此外，通过将 I_0 的界放宽至 $|I_0| \leq \max(\delta^{-1}K^2, \eta T)$，可以确保某个 $\eta > 0$ 满足 $P \gg \eta^2 r$ 。

致谢。本文是 [6], 的后续论文，后者基于我于 2020 年在耶路撒冷希伯来大学完成的硕士论文。因此，我感谢我的论文导师 Tamar Ziegler 和 Elon Lindenstrauss 对我的支持；Elon Lindenstrauss 还建议，本文的结果应当可以用 [6] 中的思路实现。感谢我的博士导师 P´eter Varj´u，在我攻读博士学位期间给予我自由，使我得以完成这两篇论文的工作。最重要的是，我感谢 Adri´an Ubis；他在评审上一篇论文时提出了断言 7 的证明中所用的论证，大大简化了 [6] 中的证明。若不是两年后从新的角度看待这些材料，我甚至不会想到重新研究本文所解决的问题。

2. On the behaviour of the equidistribution parameter in Theorem 2

除了引理 3 本身之外，为了证明定理 1，我们还需要用到 [6] 第 4 章中的一些其他内容。我们回顾其中部分内容，并稍微扩展 [6] 中所呈现的内容。

众所周知，$G \cong T_1\mathbb{H}$，其中 $\mathbb{H}$ 是带有双曲度量的上半平面。$X = \Gamma\backslash G$ 的基本域是一个集合 T_1F，其中 F 是 $\mathbb{H}$ 中的一个测地多边形——也就是说，它是一个具有有限个顶点、边由测地线片段组成的多边形[2]。这个基本多边形 F 有有限个顶点与

上半平面的边界相接，这些顶点或者位于实部等于零的轴上，或者位于无穷远处。将处于 Γ作用下同一轨道中的顶点进行识别后，可得到 X的尖点，我们将其记为 r_1、.....、r_n。任意这样的尖点 r_i都与某个元素 $\gamma_i \in \Gamma$ 一一对应，并满足： γ_i 固定 r_i ，且 γ_i 与 $h(1)$共轭（见 [6]中的引理 3.1）。对于每个尖点，都存在一些元素 $\sigma_i \in G$ ，使得 $\sigma_i r_i = \infty$ 和 $\sigma_i \gamma_i \sigma_i^{-1} = h(1)$。

对于 $g \in G$，我们定义 $Y_i^0(g) := \mathrm{Im}(\sigma_i g)$，其中

$$\mathrm{Im}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) := \frac{1}{c^2 + d^2}$$

是投影到 $\mathbb{H}$ 上的矩阵的虚部。我们还对$p = \Gamma g_p \in X$设定 $y_i^0(p) := \max_{\gamma \in \Gamma} Y_i^0(\gamma g_p)$。

引理 4.1 在 [6] 中证明了，存在每个尖点 r_i 的互不相交邻域 $C_i \subset X$，且 $K = X \backslash \bigcup C_i$ 是紧的，使得对于任意 $p \in C_i$，$\exp(\mathrm{dist}(p)) \sim y_i^0(p)$（当然，当 $p \in K$ 时，$\exp(\mathrm{dist}(p)) \sim 1$ ）。仿照引理 4.1 中第 1 点的证明，还可以看出，若 $p = \Gamma g_p \in C_i$ 且 g_p 满足 $Y_i^0(g_p) = y_i^0(p)$，则对于任意 $\gamma \in \Gamma$，要么 $Y_i^0(\gamma g_p) \ll 1$，要么 $Y_i^0(\gamma g_p) = Y_i^0(g_p)$（后一种情形下，$\sigma_i \gamma g_p = h(n)\sigma_i g_p$ 且对某个 n，有 $\gamma = (\gamma_i)^n$ ）。特别地，这意味着存在一个绝对常数 $C = C(\Gamma)$，使得若 g_p 满足 $Y_i^0(g_p) \geq C$，则

$$Y_i^0(g_p) \sim y_i^0(p) \sim \exp(\mathrm{dist}(p)),$$

其中，第二个等价关系成立，是因为 $y_i^0(p) \geq C$ 蕴含$p \in C_i$ 对 C 充分大时成立。

在轨道长度变化时，我们将在定理 1的陈述中使用等分布参数 r ，因此定义

$$r(q, K) := K \exp(-\mathrm{dist}(g_{\log K}(q))).$$

观察 4. 存在一个绝对的 $c_0 = c_0(\Gamma) > 0$ ，使得对于任意T 和任意 p，若存在 p 的一个代表元 g_p 以及一个 i ，使得对$\sigma_i g_p =: \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，$\max(T^2c^2, d^2) \leq c_0 T$，则 $r(p, T) \sim \max(T^2c^2, d^2)$ 。

证明。我们有

$$2\max(T^2c^2, d^2) \geq T(c^2T + d^2T^{-1}) = Y_i^0(g_{\log T}(g_p))^{-1}T.$$

因此， $Y_i^0(g_{\log T}(g_p)) \geq \frac{1}{2}c_0^{-1}$，这表明

$$\exp(\mathrm{dist}(g_{\log T}(p))) \sim Y_i^0(g_{\log T}(g_p))$$

根据上述论证，只要 c_0 足够小即可。 □

3. Proof of Theorem 1

我们首先用稀疏算术序列逼近 $t^{1+\gamma}$。更准确地说，我们写作

$$t^{1+\gamma} = T_0^{1+\gamma} + (1+\gamma)T_0^{\gamma}(t-T_0) + O(T^{-\frac{1}{6}})$$

在 $[T_0, T_0+T^{\frac{1}{3}}]$ 上使用泰勒展开来处理 $T_0 \geq T^{\frac{5}{6}}$。

我们将分成几种情况。为了确定我们属于哪种情况，固定某个 $\varepsilon > 0$，并要求 $\gamma < \frac{\varepsilon\beta}{6}$。最后我们会看到，$\varepsilon$ 取什么值才能使一切成立（结果将是 $\varepsilon = \frac{1}{100}$）。

为了应用有关稀疏等分布的结果，我们因此需要计算如下形式的表达式。

$$\left| \frac{1}{K} \sum_{n \leq K} f\left(qh\left((1+\gamma)T_0^{\gamma} n\right)\right) - \int f \, d\mu_X \right|$$

对于 $q = ph(T_0^{1+\gamma})$ 和 $T^{\frac{1}{6}} \leq K \leq T^{\frac{1}{3}}$，给定某个 $T_0 \leq T$。在 $r(q,K) \geq T^{\varepsilon}$ 的情况下，定理 2 足以推出良好的等分布性。

若 $r(q,K) \leq T^{\varepsilon}$，则如上一节所述，$g_{\log K}(q)$ 必须位于某个尖点 r_i 的邻域 C_i 中。在这种情况下，存在一个（本质上唯一的）q 的代表元 g_q，使得 $r(q,K) \sim \max(K^2c^2, d^2)$，其中我们定义

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} := \sigma_i g_q,$$

此后以及接下来的几页中，

接下来需要再分成两种情况。这两种情况的区别取决于

$$W_q := \left| \frac{d}{c} \right|.$$

这个 W_q 的重要性在于，它衡量了从等分布不佳到再次获得良好等分布所需的时间。更确切地说，根据观察 4，

$$r(q,K) \sim \begin{cases} d^2, K \leq W_q \\ d^2 \frac{K^2}{W_q^2}, K \geq W_q \end{cases} \tag{1}$$

只要 $r(q,K) \leq c_0 K$。

这意味着，即使 q 和 K 满足 $r(q,K) \leq T^{\varepsilon}$，也有 $r(q, T^{\varepsilon}W_q) \geq T^{2\varepsilon}$。结合定理 2，这足以证明在除下文命题 5 的假设之外的所有假设下，都有有效等分布。在这些假设下（它们涵盖了最有趣的情形），几乎整个 horocycle 轨道$\{ph(t)\ t \leq T^{1+\gamma}\}$都接近周期较短的周期 horocycle 轨道。此时，需要用引理 3 来得出结论。

命题 5. 设 Γ 和 $\gamma < c$ 如定理 1 中所述，并设 $\varepsilon = \frac{1}{100}$。设 $p \in X$ 和 T 满足 $r(p, T^{1+\gamma}) \leq T^{4\varepsilon}$ 和 $W_p \geq T^{1-\varepsilon}$。则对于定理 1 中所述的 f，

$$\left| \frac{1}{T} \sum_{n \leq T} f(ph(n^{1+\gamma})) - \int f \, d\mu_X \right| \ll r^{-\frac{\beta}{4}}.$$

为证明定理 1，我们首先将利用观察 4 和定理 2，说明如何把定理 1 的证明归约为命题 5。随后我们将证明命题 5。

在假设命题 5 的情况下证明定理 1。设给定某个 t_0，并令 $q = ph(t_0^{1+\gamma})$。如果 $r := r(q, T^{\frac{1}{6}}) \geq T^{\varepsilon}$，那么根据定理 2，对于任意满足 $\|f\| \leq 1$ 的 f，

$$\left| \frac{1}{T^{\frac{1}{6}}} \sum_{n \leq T^{\frac{1}{6}}} f\left(qh\left((1+\gamma)t_0^{\gamma} n\right)\right) - \int f \, d\mu_X \right| \ll T^{\frac{\gamma}{2}} r^{-\frac{\beta}{2}} \leq r^{-\frac{\beta}{4}},$$

我们在此回顾 $\gamma \leq \frac{\varepsilon\beta}{6}$。因此，除非存在某个 q 使得 $r = r(q, T^{\frac{1}{6}}) \leq T^{\varepsilon}$，否则我们已经完成证明。正如我们在第 2 节中所见，令 c 和 d 如前一页所定义，则

$$r \sim \max\left(T^{\frac{2}{6}} c^2, d^2\right). \tag{2}$$

如果 $c^2 T^{\frac{2}{6}}$ 在（2）式中取得最大值，或者等价地，如果 $W_q \leq T^{\frac{1}{6}}$，那么根据（1）式，$r(q, T^{\frac{1}{4}}) \sim T^{\frac{1}{6}} r \geq T^{\frac{1}{6}}$，由定理 2 即得结论。因此我们可以假设 $W_q \geq T^{\frac{1}{6}}$。下面的断言说明了如何进一步改进关于 W_q 的下界。

断言 6. 设 $q = ph(t_0^{1+\gamma})$ 满足 $r \leq T^{\varepsilon}$。令 $W := W_q$。如果 $W \leq T^{1-\varepsilon}$，那么对于 $K = W^{1+\varepsilon}$ 以及对于满足 $\|f\| \leq 1$ 的 f，

$$\left| \frac{1}{K} \sum_{0 \leq n \leq K} f\left(ph\left((t_0+n)^{1+\gamma}\right)\right) - \int f \, d\mu_X \right| \ll r^{-\frac{\beta}{4}}.$$

断言 6 的证明。固定某个 $W^{1+\varepsilon} \geq s \geq W^{1+\frac{\varepsilon}{2}}$，并注意到此时 $c^2 s^2 \sim W^{-2} s^2 d^2 \gg d^2$。因此，

$$\begin{aligned} r(qh(s), T^{\frac{1}{3}}) &\sim \max\left(T^{\frac{2}{6}} c^2, (d+cs)^2\right) \sim \max\left(T^{\frac{2}{6}} c^2, c^2 s^2\right) \\ &= c^2 s^2 \sim \left(\frac{s}{W}\right)^2 r \geq r W^{\varepsilon} \geq r T^{\frac{\varepsilon}{6}}, \end{aligned}$$

其中，第一个等价关系由观察 4 得出，其适用条件是 $\left(\frac{s}{W}\right)^2 r \ll T^{3\varepsilon}$。应用定理 2 可知

$$\left|\frac{1}{T^{\frac{1}{3}}}\sum_{n\leq T^{\frac{1}{3}}} f\left(ph(t_0^{1+\gamma}+s)h\left((1+\gamma)(t_0+s)^{\gamma}n\right)\right) - \int f\, d\mu_X\right| \ll T^{\frac{\gamma}{2}}T^{-\frac{\epsilon\beta}{12}}r^{-\frac{\beta}{2}} \leq r^{-\frac{\beta}{2}}.$$

现在，我们如上所述使用泰勒近似，将 $(t_0+n)^{\gamma}$ 的轨道按 $n \leq K$ 划分为不同范围 $[s, s+T^{\frac{1}{3}}]$，并注意到，对于除比例为 $W^{-\frac{\varepsilon}{2}}T^{\gamma}$ 的 s 之外的所有情形，均有 $(t_0+s)^{1+\gamma} - t_0^{1+\gamma} \geq W^{1+\frac{\varepsilon}{2}}$。由于 $W^{-\frac{\varepsilon}{2}}T^{\gamma} \leq T^{-\frac{\varepsilon}{12}} \leq r^{-\frac{\beta}{4}}$，结论得证。□

因此，除非存在一个$q = ph(t_0)$，使得 $r(q, T^{\frac{1}{6}}) \leq T^{\epsilon}$ 且 $W_q \geq T^{1-\varepsilon}$，否则我们已经证明了定理 1 的结论。我们令 c和 d如上所定义，并注意到在所讨论的情形下，根据 W_q 的定义， $r(q, T^{\frac{1}{6}}) \sim \max(c^2T^{\frac{1}{3}}, d^2) = d^2$。由 (1) 可知，这意味着

$$r(q, T^{1+\gamma}) \ll d^2\frac{T^{2(1+\gamma)}}{W_q^2} \ll T^{4\varepsilon}.$$

最后，为了将误差项写成 $r(p, T^{1+\gamma})$ 而非 $r(q, T^{1+\gamma})$，注意到 $g_q h(-t_0^{1+\gamma})$ 是 p 的一个代表元，并且由于 $(d-ct_0^{1+\gamma})^2 \ll d^2T^{2(1+\gamma)}W^{-2} \ll T^{4\varepsilon}$，

$$r(p, T^{1+\gamma}) \sim \max\left(c^2T^{2(1+\gamma)}, (d-ct_0)^2\right) \ll T^{4\varepsilon},$$

其中，第一个等价关系源于观察 4。因此，我们已将定理 1 的证明化归为满足命题 5 的假设。□

现在只需证明命题 5，它是定理 1 证明的主要部分。

命题 5 的证明。设给定 p 和 T，使得 $r := r(p, T^{1+\gamma}) \leq T^{4\varepsilon}$ 且 $W := W_p \geq T^{1-\varepsilon}$。这里， $W = \left|\frac{d}{c}\right|$，其中 c 和 d 按观察 4 中的定义。我们还令 $g := g_p$，并令 σ_i 如观察 4 中所述。我们使用引理 3 将轨道 $[0, T^{1+\gamma}]$ 分割成长度为 $K = T^{\frac{1}{3}}$ 的若干段。与 [6], 第 4 章引理 1.3 的证明一样，现在我们利用方程对该轨道进行参数化

$$\sigma_i gh(W+s) = lh(s) = h\left(\alpha - \frac{Rs}{s^2+1}\right) a\left(\frac{R}{s^2+1}\right) k(-\operatorname{arccot} s)$$

其中 $l := \sigma_i gh(W) =: (\alpha+iR, -i)$ 是 horocycle 轨道的最高点。给定一个 $M \leq T$，我们进而有 $ph(M^{1+\gamma}+t)$ $t \leq T^{\frac{1}{3}}$是

与周期性 horocycle$\xi h(t)$ $t \le T^{\frac{1}{3}}$ 上的轨道距离至多为 $O(T^{-\frac{1}{6}})$，其周期等于 y^{-1}，其中

$$y := \frac{R}{(M^{1+\gamma} - W)^2 + 1}.$$

根据引理 3 的第二个条件，我们可以假设 $r \gg y^{-1} \gg \delta^2 r$，但有一个比例为 δ 的区间除外，其中 δ 将稍后选定。因此，对 $t^{1+\gamma}$ 作泰勒近似后，我们希望给出以下界

$$\left| \frac{(1+\gamma)M^\gamma}{T^{\frac{1}{3}}} \sum_{(1+\gamma)M^\gamma n \le T^{\frac{1}{3}}} f(\xi h((1+\gamma)M^\gamma n)) - \int f d\mu_X \right|.$$

然而，这里可能会遇到问题：例如，如果 $y^{-1} = (1+\gamma)M^\gamma$，那么这些点在周期性 horocycle 中完全不等分布。为了解决这一问题及相关障碍，我们采用与 [6] 中断言 5.2 的证明类似的方法。为方便记号，令 $s := (1+\gamma)M^\gamma$。取满足 $y^{-1} \le q \le ys^{-1}T^{\frac{1}{3}}$ 的 $q \in \mathbb{N}$，使得

$$\left| sy - \frac{a}{q} \right| \le \frac{y^{-1}s}{qT^{\frac{1}{3}}}$$

其中某个 a 与 q 互素（根据抽屉原理，这样的 q 存在）。当 q 相较于 y^{-1} 较小时，就会出现问题。另一方面，如果 q 足够大，那么区间 $[0, y^{-1}]$ 中就有如此多互不相同的点，以至于它们必然足够稠密，足以逼近 $\int_0^1 f(\xi h(ty^{-1}))dt$；我们现在将证明这一点。

断言 7. 如果 $q \ge y^{-3}$，那么

$$\left| \frac{s}{T^{\frac{1}{3}}} \sum_{sn \le T^{\frac{1}{3}}} f(\xi h(sn)) - \int_0^1 f(\xi h(ty^{-1}))dt \right| \ll y \ll \delta^{-2} r^{-1},$$

其中 q、s、y 和 ξ 均依赖于 M。

断言 7 的证明。（此断言证明中的论证由 Adri´an Ubis 提出）我们令 $F(t) := f(\xi h(ty^{-1}))$，它是一个周期函数。由于函数 f 关于双曲度量是 1-利普希茨的，函数 F 是 y^{-1}-利普希茨的。我们希望证明

$$\left| \frac{s}{T^{\frac{1}{3}}} \sum_{sn \le T^{\frac{1}{3}}} F(nsy) - \int_0^1 F(t)dt \right| \ll y.$$

为此，我们注意到，和任意 n 一样，

$$\left| sny - n\frac{a}{q} \right| \le n\frac{y^{-1}s}{qT^{\frac{1}{3}}},$$

我们有

$$\begin{aligned}\frac{s}{T^{\frac{1}{3}}}\sum_{sn\leq T^{\frac{1}{3}}}F(nsy) &= O\left(\frac{y^{-2}}{q}\right)+\frac{s}{T^{\frac{1}{3}}}\sum_{sn\leq T^{\frac{1}{3}}}F\left(n\frac{a}{q}\right)\\ &= O\left(\frac{y^{-2}}{q}\right)+O\left(\frac{qs}{T^{\frac{1}{3}}}\right)+\frac{1}{q}\sum_{j=0}^{q-1}F\left(\frac{ja}{q}\right)\end{aligned}$$

根据 F 的周期性。由于 a 与 q 互素，因此它不会影响最后一个平均值，可以将其省略。此外，对于任意 $t\leq\frac{1}{q}$，

$$F\left(\frac{j}{q}\right)=O\left(\frac{y^{-1}}{q}\right)+F\left(\frac{j}{q}+t\right),$$

so

$$\begin{aligned}\frac{1}{q}\sum_{j=0}^{q-1}F\left(\frac{j}{q}\right) &= O\left(\frac{y^{-1}}{q}\right)+\frac{1}{q}\sum_{j=0}^{q-1}\int_0^1 F\left(\frac{j+t}{q}\right)dt\\ &= O\left(\frac{y^{-1}}{q}\right)+\int_0^1 F(t)dt.\end{aligned}$$

由于 $y^{-2}q^{-1}$ 和 $qsT^{-\frac{1}{3}}$ 都是 $O(y)$，这就证明了该结论。 □

根据 Str¨ombergsson 的结果 [7],

$$\left|y\int_0^{y^{-1}}f(\xi h(t))dt-\int f\,d\mu_X\right|\ll y^{\beta}\ll(\delta^{-2}r^{-1})^{\beta},$$

因此我们从断言 7 得知

$$\left|\frac{(1+\gamma)M^{\gamma}}{T^{\frac{1}{3}}}\sum_{n\leq T^{\frac{1}{3}}}f(\xi h((1+\gamma)M^{\gamma}n))-\int f d\mu_X\right|\ll(\delta^2 r)^{-\beta}$$

除非存在与 q 互素的 $q\leq y^{-3}\leq r^3$ 和 a，使得

$$\left|(1+\gamma)M^{\gamma}y-\frac{a}{q}\right|\ll M^{\gamma}y^{-1}T^{-\frac{1}{3}}\leq rT^{-\frac{1}{3}+\gamma}.$$

为了完成定理 1 的证明，我们只需说明这种情况是极为罕见的。

幸运的是，这正是我们所预期的：如果我们令

$$I_{q,a}:=\left\{v\in\mathbb{R}:\left|v-\frac{a}{q}\right|\leq rT^{-\frac{1}{3}+\gamma}\right\}$$

将 $q\leq r^3$ 和 $(a,q)=1$的问题区间记为 $rT^{-\frac{1}{3}+\gamma}$ 的比例。进一步地，给定互不相同的区间$I_{q_1、a_1}, I_{q_2、a_2}$，它们之间的间隔至少为 r^{-6} 量级，因为

$$\left|\frac{a_1}{q_1}-\frac{a_2}{q_2}\right|\geq\frac{1}{q_1q_2}\geq r^{-6}.$$

由于 $r \ll T^{4\varepsilon}$，这意味着集合 $E := \bigcup_{q \le r^3,\ (a,\ q)=1} I_{q,\ a}$ 仅占整个范围中极小的一部分。除非函数

$$G(t) := \frac{t^\gamma R}{(t^{1+\gamma} - W)^2 + 1} = t^\gamma y$$

高度集中于其取值范围的一小部分，因此，我们的问题情形$\{t \le T : (1+\gamma)G(t) \in E\}$ 只会出现在 $[0, T]$中可忽略的比例上。下面的断言表明， G 不会以这种不同寻常的方式表现。

断言 8. 对于除 $O(\delta + \delta^{-5} r^7 T^{-\frac{1}{3}+\gamma})$比例的 $t \le T$之外的所有情形，不存在 $q \le r^3$ 使得

$$\left| (1+\gamma)G(t) - \frac{a}{q} \right| \le r T^{-\frac{1}{3}+\gamma}.$$

在证明该断言之前，我们先说明它如何推出命题 5。该断言表明，在应用泰勒近似时，我们将 $[0, T]$ 划分成的区间中，只有很小一部分是坏区间；对于其余区间，根据断言 7，我们知道等分布成立。将所有不同的误差项汇总起来，

$$\left| \frac{1}{T} \sum_{n \le T} f(ph(n^{1+\gamma})) - \int f \, d\mu_X \right| \ll \delta + \delta^{-5} r^7 T^{-\frac{1}{3}+\gamma} + (\delta^{-2} r^{-1})^\beta,$$

其中，误差项依次来自引理 3 和断言 8、序列$\xi(1+\gamma)M^\gamma n$在周期性 horocycle 中不等分布的那些问题区间$I_{q,\ a}$，以及在好区间上与 $\int f \, d\mu_X$ 的比较。令 $\delta = r^{-\frac{1}{10}}$ 即可处理第一项和第三项；另一方面，回想 $r \ll T^{4\varepsilon}$，令 $\varepsilon = \frac{1}{100}$即可控制第二项。至此，命题 5（因而也包括定理 1）的证明完成，只剩下断言 8 尚待证明。□

断言 8 的证明。为了证明这一断言，我们使用以下简单引理，其证明留给读者作为练习。

引理 9. 设 $I \subset \mathbb{R}$ 是一个开区间，且 $G\colon I \to \mathbb{R}$ 是连续可微的，并且对于所有 $t \in I$，都有 $0 < c \le |G'(t)| \le C$。设 $\theta > 0$，并设 $a_1 < b_1 < a_2 < \cdots < a_{n-1} < b_{n-1} < a_n$ 为满足以下性质的实数：对于所有 $1 \le i < n$，都有 $b_i - a_i \le \theta(a_{i+1} - b_i)$。那么对于 $E := (a_1, b_1) \cup \cdots \cup (a_{n-1}, b_{n-1})$，

$$|\{t \in I : G(t) \in E\}| \le 2\theta C c^{-1} |I|$$

前提是 $|I| \ge \theta C c^{-1}$。

为了将此应用于我们

$$G(t) = \frac{t^\gamma R}{(t^{1+\gamma} - W)^2 + 1} = t^\gamma y$$

感兴趣的函数，我们需要计算其导数。我们可以看到，

$$\frac{dy}{dt}(t) = -\frac{2(1+\gamma)t^{\gamma}(t^{1+\gamma}-W)R}{((t^{1+\gamma}-W)^2+1)^2} = -\frac{2y(1+\gamma)t^{\gamma}(t^{1+\gamma}-W)}{(t^{1+\gamma}-W)^2+1}$$

因此

$$G'(t) = yt^{\gamma-1}\left(\gamma - \frac{2(1+\gamma)t^{1+\gamma}(t^{1+\gamma}-W)}{(t^{1+\gamma}-W)^2+1}\right).$$

回顾一下，在引理 3 中，我们排除了一个比例为δ 的区间 J_0 ，以确保 $r^{-1} \ll y \ll \delta^{-2}r^{-1}$。我们还排除了一个由两个区间组成的集合 J_1 ，其比例为 δ ，以确保 $t \geq \delta T$ 和 $\left|\frac{W}{t^{1+\gamma}}-1\right| \geq\delta$。这保证了在范围$[0,T]\setminus(J_0 \cup J_1)$ 上 $r^{-1}T^{\gamma-1} \ll yt^{\gamma-1} \ll \delta^{-3}r^{-1}T^{\gamma-1}$。如果我们能以类似的方式将方括号中的表达式界定在相差 δ^{-1}的幂次因子的范围内，则由引理 9 可得所述结论。

为此，我们注意到，对于 $t \in [0,T]\setminus J_1$,

$$\left|\frac{1}{(t^{1+\gamma}-W)^2+1} - \frac{1}{(t^{1+\gamma}-W)^2}\right| = O(\delta^{-4}T^{-4(1+\gamma)}),$$

这意味着

$$G'(t) = yt^{\gamma-1}\left(\gamma + \frac{2(1+\gamma)}{\frac{W}{t^{1+\gamma}}-1} + O(\delta^{-4}T^{-2})\right).$$

我们令 $J_2 := \left\{t : \left|\frac{W}{t^{1+\gamma}} - \left(1-\frac{(2+\gamma)}{\gamma}\right)\right| \geq \delta\right\}$，这是一个比例为 δ 的区间，在该区间上第二项大致抵消第一项。于是我们有

$$\delta \ll \gamma\left|\frac{W}{t^{1+\gamma}}-1\right|^{-1}\left|\frac{W}{t^{1+\gamma}}-1+\frac{(2+\gamma)}{\gamma}\right| = \left|\gamma + \frac{2(1+\gamma)}{\frac{W}{t^{1+\gamma}}-1}\right| \ll \delta^{-1}$$

在 $[0,T]\setminus(J_1 \cup J_2)$ 上，这意味着

$$\delta r^{-1}T^{\gamma-1} \ll |G'(t)| \ll \delta^{-4}r^{-1}T^{\gamma-1}$$

在 $[0,T]\setminus(J_0 \cup J_1 \cup J_2)$ 上。现在，我们可以对剩余的每个区间应用引理 9。回想一下，每个问题区间 $I_{q,\ a}$ 的长度为$rT^{-\frac{1}{3}+\gamma}$，且任意两个相邻区间之间的间隙至少为 $0.9r^{-6}$，我们得到

$$\frac{1}{T}\left|\{t \in [0,T]\setminus(J_0 \cup J_1 \cup J_2) : (1+\gamma)G(t) \in E\}\right| \ll \delta^{-5}r^{7}T^{-\frac{1}{3}+\gamma}$$

其中与前面一样，$E = \bigcup_{q\leq r^3,\ \ (a,\ q)\ =1} I_{q,\ a}$。这证明了断言 8，它是定理 1 的证明中最后缺少的一环。

□

References

[1] Jean Bourgain。《算术集合的逐点遍历定理》。载于：Inst. Hautes ´Etudes Sci. Publ. Math. 69（1989）。附录由作者、Harry Furstenberg、Yitzhak Katznelson 和 Donald S. Ornstein 撰写，第 5–45 页。[2] Manfred Einsiedler 和 Thomas Ward。《遍历理论：以数论为视角》。2011 年 1 月。[3] Daniel A. Goldston 和 Cem Yalcin Yildirim。《与素数相关的除数和的高阶相关性 I：三重相关》。载于：（2001）。arXiv: math/0111212 [math.NT].[4] Ben Green 和 Terence Tao。《素数包含任意长的等差数列》。载于：Ann Math 167（2004 年 5 月）。[5] Peter Sarnak 和 Adri´an Ubis。《素数时刻的 horocycle 流》。载于：Journal de Math ´ematiques Pures et Appliqu ´ees 103（2011 年 10 月）。[6] Lauritz Streck。《$\Gamma\backslash PSL_2(\mathbb{R})$ 中素数的非集中性》。2023年。arXiv: 2303.07781 [math.NT].[7] Andreas Str¨ombergsson。《horocycle 流的遍历平均偏差》。载于：Journal of Modern Dynamics 7.2（2013），第 291–328 页。[8] Akshay Venkatesh。《稀疏等分布问题、周期界与次凸性》。载于：Ann. of Math. (2) 172.2（2010），第 989–1094 页。[9] Cheng Zheng。《有限体积商空间中幺幂轨道的稀疏等分布》。载于：Journal of Modern Dynamics 10.1930- 53112016021（2016），第 1 页。 ISSN：1930-5311。