Towards an Asymptotic Double Ergodicity Formula for Measure Class-Preserving Transformations

Axel Liu-Bai, Anush Tserunyan

October 19, 2026

Abstract

This paper investigates characterizations of double ergodicity (DE) of invertible nonsingular Borel automorphisms of a standard probability space with the goal of generalizing the asymptotic formula from the probability measure-preserving (pmp) to the measure class-preserving (mcp) setting. Inspired by the pmp case, a formula for DE is proposed using Radon-Nikodym cocycles, and a sufficient condition on orbit weight ratios is identified for its validity. This condition is shown to fail for two counterexamples, a Type I translation and a Type III odometer action. Additional work is required to determine whether this condition is satisfied by DE actions. Consequently, whether the formula holds in that setting remains an open question.

1 Introduction

When considering an action of $\mathbb{Z}$ on a measure space (X, μ), ergodicity is a concept that provides information about the structure of T-invariant sets. A system is ergodic if every T-invariant set is null or conull. A strengthening of this idea is double ergodicity (DE), which is the ergodicity of the product transformation $T \times T$ on the product space $(X \times X, \mu \times \mu)$.

A special type of transformations are measure preserving ones. This means, for any measurable set $A \subseteq X$, we have $\mu(T^{-1}A) = \mu(A)$. In probability spaces, we call them probability measure-preserving (pmp). Invertible pmp transformations have been well studied, and double ergodicity for them can be characterized by a simple asymptotic formula: For all $A, B \in \mathcal{B}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\right| = 0 \qquad (*)$$

However, many transformations are not pmp. Instead, they preserve only the measure class: null sets are mapped to null sets, and positive measure sets are mapped to positive measure sets. These are called measure class preserving (mcp).

In this paper, we explore how to extend the formula for double ergodicity to mcp transformations using the Radon-Nikodym cocycles. We propose an mcp analog and a sufficient condition for the formula to hold. Finally, we show that the condition does not hold generally through specific examples of mcp transformations.

Acknowledgements

A. Liu-Bai was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [USRA - 616360 - 2026]. A. Liu-Bai a été financé par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) [USRA - 616360 - 2026]. A. Liu-Bai was supported by the Fonds de recherche du Québec – Nature et technologies (FRQNT) (#2012716).

Disclosure of generative-AI tool use

While the specific example of Type III action was our own, the initial verification and computation of the limit behaviour of θ_n in Section 5.2 were assisted by ChatGPT.

2 Preliminaries (pmp)

In this section, we present the proof from [Wal82] (Theorem 1.24) for the formula characterizing double ergodicity in the pmp setting. It provides inspiration for our proposed formula and its proof. The following theorem states the formula for pmp transformations:

Theorem 1. *Let T be an invertible pmp transformation of a standard probability space $(X, \mathcal{B}, \mu)$, so $\mathcal{B}$ is the Borel σ-algebra. Then, TFAE:*

1. *For all $A, B \in \mathcal{B}$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\right| = 0 \tag{1}$$

2. *For all $A \times C, B \times D \in \mathcal{B} \times \mathcal{B}$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu^2((T \times T)^{-i}(A \times C) \cap (B \times D)) - \mu^2(A \times C)\mu^2(B \times D)\right| = 0 \tag{2}$$

3. *$T \times T$ is ergodic, i.e, T is doubly ergodic (DE)*

To prove this, we will use the following theorem and the corollary of Birkhoff's pointwise ergodic theorem, both stated without proof.

Theorem 2. *Let T be a pmp transformation of a standard probability space $(X, \mathcal{B}, \mu)$. Then, TFAE:*

1. *For all $A, B \in \mathcal{B}$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\right| = 0$$

2. *For all $A, B \in \mathcal{B}$, there is a subset $J \subseteq \mathbb{N}$ of density zero s.t*

$$\lim_{\substack{n\to\infty \\ n\notin J}} \mu(T^{-i}(A) \cap B) - \mu(A)\mu(B) = 0$$

3. For all $A, B \in \mathcal{B}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\right|^2 = 0$$

Corollary 1. *Let T be an ergodic pmp transformation on a probability space $(X, \mathcal{B}, \mu)$. Then, for all $A, B \in \mathcal{B}$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \mu(T^{-i}(A) \cap B) = \mu(A)\mu(B)$$

We are now ready to prove the theorem.

Proof of theorem 1. **1** $\implies$ **2:** Let $A, B, C, D \in \mathcal{B}$. Then, by theorem 2 (1 $\iff$ 2), there are $J_1, J_2 \subseteq \mathbb{N}$ of density zero s.t

$$\lim_{\substack{n\to\infty \\ n\notin J_1}} \mu(T^{-i}(A) \cap B) = \mu(A)\mu(B)$$
$$\lim_{\substack{n\to\infty \\ n\notin J_2}} \mu(T^{-i}(C) \cap D) = \mu(C)\mu(D)$$

Consider $J = J_1 \cup J_2$, which also has density zero. Then,

$$\begin{aligned}\lim_{\substack{n\to\infty \\ n\notin J}} \mu^2((T\times T)^{-i}(A\times C) \cap (B\times D) &= \lim_{\substack{n\to\infty \\ n\notin J}} \mu(T^{-i}(A) \cap B)\mu(T^{-i}(C) \cap D) \\ &= \mu(A)\mu(B)\mu(C)\mu(D) \\ &= \mu^2(A\times C)\mu^2(B\times D)\end{aligned}$$

Thus, by theorem 2 (1 $\iff$ 2), we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu^2((T\times T)^{-i}(A\times C) \cap (B\times D)) - \mu^2(A\times C)\mu^2(B\times D)\right| = 0$$

By approximating sets in $\mathcal{B}\times\mathcal{B}$ by rectangles, we have that Equation (2) holds on $\mathcal{B}\times\mathcal{B}$.

2 $\implies$ **3:** If $E \in \mathcal{B}\times\mathcal{B}$ was a $T\times T$-invariant set, then for any $i \in \mathbb{N}$, we have $(T\times T)^{-i}E = E$. Then, using the assumption on E, we have

$$\frac{1}{n} \sum_{i<n} \left|\mu^2((T\times T)^{-i}(E) \cap E) - \mu^2(E)\mu^2(E)\right| = 0$$

By invariance, we have

$$0 = \frac{1}{n} \sum_{i<n} \left|\mu^2(E) - (\mu^2(E))^2\right| = \left|\mu^2(E) - (\mu^2(E))^2\right|$$

So $\mu^2(E) = (\mu^2(E))^2$ and so $\mu^2(E) \in \{0, 1\}$. Thus, $T\times T$ is ergodic.

3 $\implies$ **1:** Let $A, B \in \mathcal{B}$. By theorem 2 ($1 \iff 3$), it is equivalent to show that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i<n} \left|\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\right|^2 = 0$$

First, we compute

$$\begin{aligned}\frac{1}{n} \sum_{i<n} \mu(T^{-i}(A) \cap B) &= \frac{1}{n} \sum_{i<n} \mu^2((T \times T)^{-i}(A \times X) \cap (B \times X)) \\ &\to \mu^2(A \times X)\mu^2(B \times X) \\ &= \mu(A)\mu(B)\end{aligned}$$

Similarly, we have

$$\begin{aligned}\frac{1}{n} \sum_{i<n} \mu(T^{-i}(A) \cap B)^2 &= \frac{1}{n} \sum_{i<n} \mu^2((T \times T)^{-i}(A \times A) \cap (B \times B)) \\ &\to \mu^2(A \times A)\mu^2(B \times B) \\ &= \mu(A)^2\mu(B)^2\end{aligned}$$

Thus, combining these, we obtain

$$\begin{aligned}&\frac{1}{n} \sum_{i<n} \left|\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\right|^2 \\ &\quad= \frac{1}{n} \sum_{i<n} \left(\mu(T^{-i}(A) \cap B)^2 - 2\mu(T^{-i}(A) \cap B)\mu(A)\mu(B) + \mu(A)^2\mu(B)^2\right) \\ &\quad\to \mu(A)^2\mu(B)^2 - 2\mu(A)\mu(B)\mu(A)\mu(B) + \mu(A)^2\mu(B)^2 \\ &\quad= 0\end{aligned}$$

Thus, T satisfies Equation (1). □

3 Preliminaries (mcp)

Our goal is to adapt this proof in the mcp case. The primary obstacle is that Birkhoff's ergodic theorem applies specifically to pmp transformations. In this section, we will introduce Radon-Nikodym cocycles to account for changes in measure, and present an mcp analog of the ergodic theorem. We follow the Introduction section of [MT24].

Let $T : X \to X$ be an invertible mcp transformation of a probability space (X, μ), i.e, $\mu \sim T_*\mu$. Define the Radon-Nikodym cocycle $w : \mathbb{N} \times X \to [0, \infty)$ by

$$w^x(T^k x) := \frac{d\mu}{d(T_*^k\mu)}(T^k x) = \frac{d(T_*^{-k}\mu)}{d\mu}(x)$$

Intuitively,

$$w^x(T^k x) \approx \frac{\text{weight of } T^k x}{\text{weight of } x}$$

For $f : X \to \mathbb{R}$ and $n \in \mathbb{N}$, we define the nth weighted sum of f as follows: For each $x \in X$,

$$S_n^w f(x) = \sum_{k<n} f(T^k x) w^x(T^k x)$$

This is the sum of the values of f over $\{x, Tx, \ldots, T^{n-1}x\}$ weighted by the corresponding values of $w^x(T^k x)$. For $g : X \to \mathbb{R}^+$, define the ratio

$$R_n^w(f, g) := \frac{S_n^w f}{S_n^w g}$$

Define the nth total weight as $t_n(x) = \sum_{k<n} w^x(T^k x)$ and define the w-weighted average of f over $\{x, Tx, \ldots, T^{n-1}x\}$ to be

$$R_n^w f(x) := \frac{S_n^w f(x)}{S_n^w \mathbb{1}(x)} = \frac{S_n^w f(x)}{\sum_{k<n} w^x(T^k x)} = \frac{S_n^w f(x)}{t_n(x)}$$

The mcp variation of Birkhoff's pointwise ergodic theorem is the following:

Theorem 3 (Ratio ergodic theorem, Hurewicz [Hur44] and Halmos [Hal46]). *Let $T : X \to X$ be an invertible ergodic mcp transformation of a probability space (X, μ) and w be the associated Radon-Nikodym cocycle. Then, for $f \in L^1(X, \mu)$ and for g non-negative, measurable and with $\lim_{n\to\infty} S_n^w g = \infty$ a.e,*

$$R_\infty^w(f, g) = \frac{\int f \, d\mu}{\int g \, d\mu}$$

For the remainder of the paper, we take $g = \mathbb{1}$. We prove an analog of the corollary of Birkhoff's ergodic theorem.

Corollary 2. *Let $T : X \to X$ be an invertible mcp transformation of a probability space (X, μ). Then, T is ergodic iff*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k<n} \int \mathbb{1}_B(x) \mathbb{1}_A(T^k x) w^x(T^k x) \frac{n}{t_n(x)} \, d\mu(x) = \mu(A)\mu(B)$$

Proof. $\Longrightarrow$: Let $A, B \subseteq X$ be positive measure sets. By the ratio ergodic theorem, we have for a.e $x \in X$,

$$R_n^w(\mathbb{1}_A, \mathbb{1})(x) = \frac{\sum_{k<n} \mathbb{1}_A(T^k x) w^x(T^k x)}{t_n(x)} \to \mu(A)$$

We will use DCT with

$$f_n = \mathbb{1}_B R_n^w(\mathbb{1}_A, \mathbb{1}) \qquad\qquad g = 1$$

We obtain

$$\lim_{n\to\infty} \int \mathbb{1}_B(x) \frac{\sum_{k<n} \mathbb{1}_A(T^k x) w^x(T^k x)}{t_n(x)} \, d\mu(x) = \mu(A)\mu(B)$$

Thus,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k<n}\int \mathbb{1}_B(x)\mathbb{1}_A(T^kx)w^x(T^kx)\frac{n}{t_n(x)}\,d\mu(x)=\mu(A)\mu(B)$$

$\Longleftarrow$: Let A be T-invariant. Then,

$$\begin{aligned}\mu(A)^2&=\lim_{n\to\infty}\frac{1}{n}\sum_{k<n}\int \mathbb{1}_A(x)\mathbb{1}_A(T^kx)w^x(T^kx)\frac{n}{t_n(x)}\,d\mu(x)\\&=\lim_{n\to\infty}\int \mathbb{1}_A(x)\mathbb{1}_A(x)\frac{\sum_{k<n}w^x(T^kx)}{t_n(x)}\,d\mu(x)\\&=\mu(A)\end{aligned}$$

so $\mu(A)\in\{0,1\}$ and T is ergodic. □

4 Proposed Formula

Following the pmp characterization of DE, we present our proposed formula:

Formula 1. *Let T be DE. We want to show that for $A,B\in\mathcal{B}$,*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k<n}\left|\int_X \mathbb{1}_B(x)\mathbb{1}_A(T^kx)w^x(T^kx)\frac{n}{t_n(x)}\,d\mu(x)-\mu(A)\mu(B)\right|=0$$

Our proof follows the one for the pmp case. We first start with reduction.

Reduction 1. *It is enough to show that*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k<n}\left|\int_X \mathbb{1}_B(x)\mathbb{1}_A(T^kx)w^x(T^kx)\frac{n}{t_n(x)}\,d\mu(x)-\mu(A)\mu(B)\right|^2=0$$

Proof of reduction. First, let $a_{n,k}=\left|\int_X \mathbb{1}_B(x)\mathbb{1}_A(T^kx)w^x(T^kx)\frac{n}{t_n(x)}\,d\mu(x)-\mu(A)\mu(B)\right|$. By Cauchy-Schwarz, we have

$$\begin{aligned}\left(\sum_{k<n}a_{n,k}\cdot 1\right)^2\le\left(\sum_{k<n}a_{n,k}^2\right)\left(\sum_{k<n}1^2\right)&\implies\left(\sum_{k<n}a_{n,k}\right)^2\le n\sum_{k<n}a_{n,k}^2\\&\implies\left(\frac{1}{n}\sum_{k<n}a_{n,k}\right)^2\le\frac{1}{n}\sum_{k<n}a_{n,k}^2\end{aligned}$$

Since $\frac{1}{n}\sum_{k<n}a_{n,k}^2\to 0$, then

$$0\le\frac{1}{n}\sum_{k<n}a_{n,k}\le\sqrt{\frac{1}{n}\sum_{k<n}a_{n,k}^2}\to 0$$

□

In attempting to prove this formula, we identify a sufficient condition.

Proof of formula. We will denote the product measure $\mu \times \mu$ as μ^2 and for a set A, denote $A \times A$ as A^2. Let $A, B \in \mathcal{B}$. Let

$$P_{n,k}(x) := \mathbb{1}_B(x)\mathbb{1}_A(T^k x)w^x(T^k x)\frac{n}{t_n(x)}$$
$$P_{n,k}(x,y) := \mathbb{1}_{B^2}(x,y)\mathbb{1}_{A^2}(T^k x, T^k y)w^{(x,y)}(T^k x, T^k y)\frac{n}{t_n(x,y)}$$

By the reduction, it is enough to show that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k<n}\left|\int_X P_{n,k}(x)\ d\mu(x) - \mu(A)\mu(B)\right|^2 = 0$$

We expand

$$\begin{aligned}
&\frac{1}{n}\sum_{k<n}\left|\int_X P_{n,k}(x)\ d\mu(x) - \mu(A)\mu(B)\right|^2 \\
&= \frac{1}{n}\sum_{k<n}\left(\int_X P_{n,k}(x)\ d\mu(x)\right)^2 - \frac{1}{n}\sum_{k<n} 2\mu(A)\mu(B)\int_X P_{n,k}(x)\ d\mu(x) + \frac{1}{n}\sum_{k<n}\mu(A)^2\mu(B)^2 \\
&= \frac{1}{n}\sum_{k<n}\left(\int_X P_{n,k}(x)\ d\mu(x)\right)^2 - 2\mu(A)\mu(B)\cdot\frac{1}{n}\sum_{k<n}\int_X P_{n,k}(x)\ d\mu(x) + \mu(A)^2\mu(B)^2
\end{aligned}$$

Since T is ergodic, by Corollary 2, the second term converges:

$$-2\mu(A)\mu(B)\cdot\frac{1}{n}\sum_{k<n}\int_X P_{n,k}(x)\ d\mu(x) \to -2\mu(A)\mu(B)\cdot\mu(A)\mu(B) = -2\mu(A)^2\mu(B)^2$$

If we can show that the first term converges to $\mu(A)^2\mu(B)^2$, then we would obtain

$$\frac{1}{n}\sum_{k<n}\left|\int_X P_{n,k}(x)\ d\mu(x) - \mu(A)\mu(B)\right|^2 \to \mu(A)^2\mu(B)^2 - 2\mu(A)^2\mu(B)^2 + \mu(A)^2\mu(B)^2 = 0$$

We need to use the fact that T is doubly ergodic, which gives us by Corollary 2,

$$\frac{1}{n}\sum_{k<n}\int_{X^2} P_{n,k}(x,y)\ d\mu^2(x,y) \to \mu^2(A^2)\mu^2(B^2)$$

We work on getting the first term into this form to apply the corollary.

$$\begin{aligned}
&\frac{1}{n}\sum_{k<n}\left(\int_X P_{n,k}(x)\ d\mu(x)\right)^2 = \frac{1}{n}\sum_{k<n}\left(\int_X \mathbb{1}_B(x)\mathbb{1}_A(T^k x)w^x(T^k x)\frac{n}{t_n(x)}\ d\mu(x)\right)^2 \\
&= \frac{1}{n}\sum_{k<n}\left(\int_X\int_X \mathbb{1}_B(x)\mathbb{1}_B(y)\mathbb{1}_A(T^k x)\mathbb{1}_A(T^k y)w^x(T^k x)w^y(T^k y)\frac{n}{t_n(x)}\frac{n}{t_n(y)}\ d\mu(y)d\mu(x)\right)
\end{aligned}$$

We can combine

$$\mathbb{1}_B(x)\mathbb{1}_B(y) = \mathbb{1}_{B^2}(x,y) \qquad\qquad \mathbb{1}_A(T^k x)\mathbb{1}_A(T^k y) = \mathbb{1}_{A^2}(T^k x, T^k y)$$

By the definition of Radon-Nikodym derivatives, we have

$$w^x(T^k x)w^y(T^k y) = w^{(x,y)}(T^k x, T^k y)$$

Thus,

$$\begin{aligned}
&\frac{1}{n}\sum_{k<n}\left(\int_X P_{n,k}(x)\ d\mu(x)\right)^2 \\
&\quad = \frac{1}{n}\sum_{k<n}\left(\int_{X^2} \mathbb{1}_{B^2}(x,y)\mathbb{1}_{A^2}(T^k x, T^k y)w^{(x,y)}(T^k x, T^k y)\frac{n^2}{t_n(x)t_n(y)}\ d\mu^2(x,y)\right) \\
&\quad = \frac{1}{n}\sum_{k<n}\left(\int_{X^2} P_{n,k}(x,y)\cdot\frac{t_n(x,y)}{n}\cdot\frac{n^2}{t_n(x)t_n(y)}\ d\mu^2(x,y)\right)
\end{aligned}$$

Sufficient condition: At this point in the proof, we would like

$$\theta_n(x,y) := \frac{n^2}{t_n(x)t_n(y)}\cdot\frac{t_n(x,y)}{n} \to 1 \qquad \text{in } L^1 := L^1(\mu^2)$$

We define a condition on the Radon-Nikodym cocycles of T: We say that the Radon-Nikodym cocycle has **asymptotically equal diagonal and square averages** if $\theta_n \to 1$ in L^1.

Conclusion of the proof: Now, suppose that $\theta_n \to 1$ in L^1. Note that $\frac{1}{n}\sum_{k<n} P_{n,k}(x,y) \le 1$ for all $x, y \in X$. Indeed, using the definition of $t_n(x,y)$, we have

$$\begin{aligned}
\left|\frac{1}{n}\sum_{k<n} P_{n,k}(x,y)\right| &= \left|\frac{1}{n}\sum_{k<n}\mathbb{1}_{B^2}(x,y)\mathbb{1}_{A^2}(T^k x, T^k y)w^{(x,y)}(T^k x, T^k y)\frac{n}{t_n(x,y)}\right| \\
&\le \left|\frac{n}{t_n(x,y)}\cdot\frac{1}{n}\sum_{k<n} w^{(x,y)}(T^k x, T^k y)\right| \\
&= 1
\end{aligned}$$

Thus,

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{k<n}\left(\int_{X^2} P_{n,k}(x,y)\cdot\theta_n(x,y)\ d\mu^2(x,y)\right) - \frac{1}{n}\sum_{k<n}\int_{X^2} P_{n,k}(x,y)\ d\mu^2(x,y)\right| \\
&\le \int_{X^2}\left|\frac{1}{n}\sum_{k<n} P_{n,k}(x,y)\right|\cdot|\theta_n(x,y)-1|\ d\mu^2(x,y) \\
&\le \int_{X^2}|\theta_n(x,y)-1|\ d\mu^2(x,y) \\
&\to 0
\end{aligned}$$

and so the first term converges to the same limit as $\frac{1}{n}\sum_{k<n}\int_{X^2} P_{n,k}(x,y)\ d\mu^2(x,y)$, which is $\mu(A)^2\mu(B)^2$, as desired.

$\diamondsuit$

5 Condition on the Asymptotic Behaviour of θ_n

In the previous section, we have identified a key condition on θ_n for our proposed formula to hold. A natural question arises: Does this condition generally hold for mcp transformations? Unfortunately, the condition that $\theta_n \to 1$ in L^1 does not arise naturally. We present a Type I and a Type III counter-example. It is important to note that these examples are not DE. More work is necessary to determine if there are DE invertible mcp transformations satisfying the condition on θ_n.

Definition 1 (Types). *For a mcp ergodic action $\Gamma \curvearrowright (X, \mu)$ on a σ-finite measure space (X, μ), we say it is*

- *Type* I *if there is a conull orbit.*
- *Type* II *if it is not Type I and there is an equivalent Γ-invariant measure.*
  - *Type* II_1 *if there is an equivalent probability Γ-invariant measure.*
  - *Type* II_∞ *if it is not Type* II_1*, but still Type* II*.*
- *Type* III *if it is not Type* I *or Type* II

5.1 Type I Counterexample

5.1.1 Setup

Consider the translation action of $\mathbb{Z} \curvearrowright \mathbb{Z}$ by $T : x \mapsto x + 1$. Equip $\mathbb{Z}$ with the following probability measure μ:

$$\mu(z) = \frac{2^{-|x|}}{3}$$

This is a Type I mcp action since the only orbit is the whole space and the action is clearly not measure preserving. Under the action of T, moving one step changes its local measure weight by a power of 2. Thus, the RN cocycle is the ratio of a weight function $h : \mathbb{Z} \to \mathbb{R}^+$ defined by

$$h(x) = 2^{-|x|}$$

Depending on the value of x, we can simplify the formula for the RN cocycle:

- If $x \leq -1$, then $w^x(Tx) = 2^1$
- If $x \geq 0$, then $w^x(Tx) = 2^{-1}$
- For a general $k \in \mathbb{N}$,
  - If $x < 0$ and $T^k x \leq 0$, then $w^x(T^k x) = 2^k$
  - If $x < 0$ and $T^k x > 0$, then $w^x(T^k x) = 2^{-2x-k}$
  - If $x \geq 0$, then $w^x(T^k x) = 2^{-k}$

5.1.2 Limit behaviour of θ_n

We first show that the t_n are bounded. Then, the factor of n in θ_n will allow us to conclude that θ_n does not converge to 1 in L^1.

Claim 1. *For fixed $x, y \in \mathbb{Z}$, we have $0 < \lim_{n\to\infty} t_n(x) < \infty$ and $0 < \lim_{n\to\infty} t_n(x, y) < \infty$.*

Proof. Clearly, the limits are positive. We have two cases: $x \geq 0$ and $x < 0$. If $x \geq 0$, then

$$t_n(x) = \sum_{k<n} w^x(T^k x) = \sum_{k<n} 2^{-k} < \infty$$

If $x < 0$, note that there are only finitely many terms where $w^x(T^k x) \geq 1$. Then, the tail decreases exponentially. Hence, $\lim_{n\to\infty} t_n(x) < \infty$.

Since the product of exponentially decaying tails is still exponentially decaying, we obtain that $\lim_{n\to\infty} t_n(x, y) < \infty$ by a similar argument. □

Claim 2. *T does not satisfy the condition that $\theta_n \to 1$ in L^1.*

Proof. Let $x, y \in \mathbb{Z}$. Then, by Claim 1, we have

$$\lim_{n\to\infty} \theta_n(x, y) = \lim_{n\to\infty} \frac{nt_n(x, y)}{t_n(x)t_n(y)} \to \infty$$

Since $\theta_n \to \infty$ pointwise a.e, then $\theta_n \not\to 1$ in L^1. □

5.2 Type III Counterexample

5.2.1 Setup

Equip $2^{\mathbb{N}}$ with the Bernoulli(p) measure, $p \neq 1/2$. Say $p = 1/3$. Consider the odometer action of $\mathbb{Z} \curvearrowright X$, where X is the Cantor space $2^{\mathbb{N}}$ with the words ending with infinitely many 0s or 1s removed, so X is conull. The action $T : X \to X$ represents adding 1 with binary carry to the right. For each word $x \in X$, starting with the leftmost bit, we flip 1 to 0 until we encounter the first 0, which we flip to 1. For example, $T(0, 0, 1, \dots) = (1, 0, 1, \dots)$ and $T(1, 1, 0, \dots) = (0, 0, 1, \dots)$. This action is Type III [DS23]. Intuitively, the RN cocycles are computed as follows:

- We use the intuition of
$$w^x(Tx) \approx \frac{\text{weight of } Tx}{\text{weight of } x}$$
where we can think of weight as the measure
- For each bit that went from 0 to 1, we multiply by
$$\frac{1/3}{2/3} = \frac{1}{2}$$
- For each bit that went from 1 to 0, we multiply by
$$\frac{2/3}{1/3} = 2$$

Thus, the RN cocycle is computed bit by bit, so we can split x into a finite sequence and a tail to compute its RN cocycle.

5.2.2 Motivation

We first present the motivation behind the direction of the proof and where we will derive a contradiction. For each $n \in \mathbb{N}$, define for $k < n$,

$$\pi_n^x(T^k x) := \frac{w^x(T^k x)}{t_n(x)}$$

We can then rewrite θ_n in terms of π_n:

$$\theta_n(x,y) = \frac{n^2}{t_n(x)t_n(y)} \cdot \frac{t_n(x,y)}{n} = n \sum_{k<n} \pi_n^x(T^k x)\pi_n^y(T^k y)$$

Observe that for a fixed $x \in X$, π_n^x is a probability distribution on $\{0, \ldots, n-1\}$ by the definition of $t_n(x)$. Let

$$q_n(k) := \int \pi_n^x(T^k x)\ d\mu(x)$$

q_n is also a probability distribution on $\{0, \ldots, n-1\}$. Indeed,

$$\sum_{k<n} q_n(k) = \sum_{k<n} \int_X \pi_n^x(T^k x)\ d\mu(x) = \int_X \sum_{k<n} \pi_n^x(T^k x)\ d\mu(x) = \int_X 1\ d\mu = 1$$

We can also rewrite θ_n in terms of q_n:

$$\begin{aligned}
\int_{X^2} \theta_n(x,y)\ d\mu^2(x,y) &= \int_{X^2} n \sum_{k<n} \pi_n^x(T^k x)\pi_n^y(T^k y)\ d\mu^2(x,y) \\
&= n \sum_{k<n} \left(\int_X \pi_n^x(T^k x)\ d\mu(x) \int_X \pi_n^y(T^k y)\ d\mu(y) \right) \\
&= n \sum_{k<n} q_n^2(k)
\end{aligned}$$

Claim 3. *If $\theta_n(x,y) \to 1$ in L^1, then $(q_n(k))_{k<n}$ is asymptotically uniform.*

Proof. By Cauchy-Schwarz on $(q_n(0), \ldots, q_n(n-1))$, we obtain

$$\left(\sum_{k<n} q_n(k) \right)^2 \leq \sum_{k<n} 1 \sum_{k<n} q_n^2(k) = n \sum_{k<n} q_n^2(k)$$

Since q_n is a probability distribution, we get

$$1 \leq n \sum_{k<n} q_n^2(k)$$

with equality iff $q_n(k) = \frac{1}{n}$ for all $k < n$. Since $\theta_n(x,y) \to 1$ in L^1, then $n \sum_{k<n} q_n^2(k) \to 1$, so we have equality asymptotically. Thus, $(q_n(k))_{k<n}$ is asymptotically uniform. □

Intuitively, if q_n is asymptotically uniform, then the total probability mass in the first half of the interval, $\sum_{k<n/2} q_n(k)$, should converge to $1/2$. We will show that it does not.

5.2.3 Setup for the proof

Let $N = 2^m$ for some m. Let $x \in X$ and write $x = x_0 x_1 \dots x_{m-1} z$. Let a be the integer representation of $x_{m-1} \dots x_0$ and $c(z) := w^z(Tz)$. For $y \in 2^n$, let $[y]_n$ be the cylinder specified by y, left-padded with zeros to length n.

Since $0 \le a < N$, then for $k < N$, we have $a + k < 2N$ so there is at most 1 carry to the tail z. This carry only happens if $N \le a + k < 2N$. Then, the RN cocycle is the change in measure of the finite sequence multiplied by the RN cocycle of the tail (if there is carry over).

$$w^x(T^k x) = \frac{\mu([a+k \mod N]_m)}{\mu([a]_m)} c(z)^{\mathbb{1}\{a+k \ge N\}}$$

Define

$$H_m(x) := \sum_{k<N/2} \pi_N^x(T^k x)$$

This is the proportion of the orbit weight concentrated in the first half of the orbit of length N, so we have $0 \le H_m(x) \le 1$. Let

$$Q_m := \int H_m \, d\mu = \sum_{k<N/2} q_N(k)$$

This represents the average weight, or the total probability mass, in the first half of the orbit of length N. Our proof will analyze the asymptotic behaviour of Q_m and show that it does not converge to $1/2$.

5.2.4 Computation of H_m

To analyze Q_m, we will need to study H_m first. Let $M = N/2$ and $a = \epsilon M + u$ for $\epsilon \in \{0, 1\}$ and $0 \le u < M$. Let

$$F_{m-1}(u) := \sum_{j<u} \mu([j]_{m-1})$$

We can think of this as a discrete cumulative distribution function (CDF) for μ restricted to cylinders of length $m-1$. We have the following explicit formulas for H_m:

Claim 4. *When $\epsilon = 0$,*

$$H_m(x) = \frac{\frac{2}{3}(1 - F_{m-1}(u)) + \frac{1}{3}F_{m-1}(u)}{\frac{2}{3}(1 - F_{m-1}(u)) + \frac{1}{3} + \frac{2}{3}c(z)F_{m-1}(u)}$$

When $\epsilon = 1$,

$$H_m(x) = \frac{\frac{1}{3}(1 - F_{m-1}(u)) + \frac{2}{3}c(z)F_{m-1}(u)}{\frac{1}{3}(1 - F_{m-1}(u)) + c(z)\left(\frac{2}{3} + \frac{1}{3}F_{m-1}(u)\right)}$$

The formulas for H_m can be verified by expanding the Radon-Nikodym weights, splitting the orbit sums based on whether $u + k$ overflows M to analyze carries and evaluating the tail.

5.2.5 Limit behaviour of Q_m

Now that we know how to compute H_m, we need to study F_{m-1}. Thinking about F_{m-1} as a discrete CDF, the jumps in the function are the size of the measure of the cylinders. The largest atom is [0], the all-zero cylinder of length $m-1$, which has measure $\left(\frac{2}{3}\right)^{m-1}$. As $m \to \infty$, the largest atom shrinks to 0. Thus, F_{m-1} is asymptotically continuous. Combining this with the probability integral transform, we have that $F_{m-1}(U_{m-1}) \to U$ in distribution, where U_{m-1} is the random integer represented by the first $m-1$ bits and $U \sim Uniform(0,1)$. This will be useful when integrating H_m to compute Q_m.

Theorem 4 (Probability integral transform). *Let X be a continuous random variable with CDF F_X. Then, $F_X(X) \sim Uniform(0,1)$.*

Now, for a tail z, let R be the number of leading 1s in z, which follows a geometric distribution counting the number of failures before the first success. Then,

$$P(R = r) = \frac{2}{3}\left(\frac{1}{3}\right)^r$$

If there is a carry to the tail, adding 1 changes the r leading 1s to 0s and the following 0 to a 1. Thus, the number of 1s changes by $1 - r$ and so

$$c(z) = w^z(Tz) = \frac{1}{2^{1-r}} = 2^{r-1}$$

We have the following formulas of $\int H_m \, d\mu$ depending on the value of epsilon.

Claim 5. *Let*

$$I_0(c) = \int_0^1 \frac{2-u}{3+2(c-1)u} \, du \qquad\qquad I_1(c) = \int_0^1 \frac{1+(2c-1)u}{1+2c+(c-1)u} \, du$$

where $c := c(z)$ does not depend on u. Then, $\int H_m \, d\mu \to I_0(c)$ when $\epsilon = 0$ and $\int H_m \, d\mu \to I_1(c)$ when $\epsilon = 1$

This can be verified by using the formulas in Claim 4 and the fact that $F_{m-1}(U_{m-1}) \to U$ in distribution. We are now ready to find a formula for the limit of Q_m.

Claim 6. $\lim_{m\to\infty} Q_m = Q$, *where*

$$Q = \sum_{r\in\mathbb{N}} \frac{2}{3}\left(\frac{1}{3}\right)^r \left(\frac{2}{3}I_0(2^{r-1}) + \frac{1}{3}I_1(2^{r-1})\right)$$

Proof. To compute Q_m, we take the average of H_m over all $x \in X$. We have to average over $\epsilon \in \{0,1\}$ and the possible leading 1s configurations in the tail. Recall that $c = c(z) = 2^{r-1}$.

Thus,

$$\begin{aligned}
\lim_{m\to\infty} Q_m &= \lim_{m\to\infty} \int H_m \, d\mu \\
&= P(\epsilon = 0) \sum_{r\in\mathbb{N}} P(R = r) I_0(2^{r-1}) + P(\epsilon = 1) \sum_{r\in\mathbb{N}} P(R = r) I_1(2^{r-1}) \\
&= \frac{2}{3} \sum_{r\in\mathbb{N}} \frac{2}{3} \left(\frac{1}{3}\right)^r I_0(2^{r-1}) + \frac{1}{3} \sum_{r\in\mathbb{N}} \frac{2}{3} \left(\frac{1}{3}\right)^r I_1(2^{r-1}) \\
&= \sum_{r\in\mathbb{N}} \frac{2}{3} \left(\frac{1}{3}\right)^r \left(\frac{2}{3} I_0(2^{r-1}) + \frac{1}{3} I_1(2^{r-1})\right) \\
&= Q
\end{aligned}$$

Q is well-defined. Indeed, since $H_m(x)$ is a proportion, then $0 \leq H_m \leq 1$ so by DCT, $\lim_{n\to\infty} \int H_m \, d\mu < \infty$. □

We can now show that Q_m does not converge to $1/2$, which was our intuitive guide for the counterexample.

Claim 7. $Q > 1/2$

Proof. Note that all the terms are positive. Then,

$$Q = \sum_{r\in\mathbb{N}} \frac{2}{3} \left(\frac{1}{3}\right)^r \left(\frac{2}{3} I_0(2^{r-1}) + \frac{1}{3} I_1(2^{r-1})\right) > \sum_{r=0}^{1} \frac{2}{3} \left(\frac{1}{3}\right)^r \left(\frac{2}{3} I_0(2^{r-1}) + \frac{1}{3} I_1(2^{r-1})\right)$$

First term ($r = 0$): We have $c = 2^{-1}$ and

$$I_0\left(2^{-1}\right) = \int_0^1 \frac{2-u}{3-u} \, du = 1 + \ln\left(\frac{2}{3}\right) \qquad I_1\left(2^{-1}\right) = \int_0^1 \frac{1}{2 - \frac{1}{2}u} \, du = 2\ln\left(\frac{4}{3}\right)$$

Thus, the first term is

$$\frac{2}{3}\left[\frac{2}{3}\left(1 + \ln\frac{2}{3}\right) + \frac{1}{3}\left(2\ln\frac{4}{3}\right)\right] = \frac{4}{9} + \frac{4}{9}\ln\left(\frac{8}{9}\right)$$

Second term ($r = 1$): We have $c = 1$ and

$$I_0(1) = \int_0^1 \frac{2-u}{3} \, du = \frac{1}{2} \qquad I_1(1) = \int_0^1 \frac{1+u}{3} \, du = \frac{1}{2}$$

Thus, the second term is

$$\frac{2}{3}\left(\frac{1}{3}\right)^1 \left[\frac{2}{3}\left(\frac{1}{2}\right) + \frac{1}{3}\left(\frac{1}{2}\right)\right] = \frac{2}{9} \cdot \frac{1}{2} = \frac{1}{9}$$

Lower bound: Combining these, we obtain

$$Q > \left(\frac{4}{9} + \frac{4}{9}\ln\frac{8}{9}\right) + \frac{1}{9} = \frac{5}{9} + \frac{4}{9}\ln\left(\frac{8}{9}\right) \approx 0.5032075$$

Thus, $Q > 0.5032$. □

5.2.6 Conclusion of the counterexample

Finally, we can use the fact that $Q \neq 0.5$ to show the limit behaviour of θ_n.

Proposition 1. *$\theta_n \not\to 1$ in L^1.*

Proof. By Cauchy-Schwarz, we have $\left(\sum_{i<n} a_i\right)^2 \leq n \sum_{i<n} a_i^2$. Then,

$$\begin{aligned}
\int_{X^2} \theta_N(x,y)\ d\mu^2(x,y) &= N \sum_{k<N} q_N^2(k) \\
&= 2\left(\frac{N}{2} \sum_{k<N/2} q_N^2(k) + \frac{N}{2} \sum_{k=N/2}^{N-1} q_N^2(k)\right) \\
&\geq 2\left(\sum_{k<N/2} q_N(k)\right)^2 + 2\left(\sum_{k=N/2}^{N-1} q_N(k)\right)^2 \\
&= 2Q_m^2 + 2(1-Q_m)^2 \\
&= 4Q_m^2 - 4Q_m + 2 \\
&= (2Q_m - 1)^2 + 1
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
\liminf_{m\to\infty} \int_{X^2} \theta_{2^m}(x,y)\ d\mu^2(x,y) &\geq \liminf_{m\to\infty} (2Q_m^2 - 1)^2 + 1 \\
&\geq (2 \cdot 0.5032 - 1)^2 + 1 \\
&= 1.00004096
\end{aligned}$$

Thus,

$$\liminf_{m\to\infty} \int_{X^2} |\theta_{2^m} - 1|\ d\mu^2 \geq \liminf_{m\to\infty} \left| \int_{X^2} \theta_{2^m}\ d\mu^2 - 1 \right| \geq 0.00004096 > 0$$

and so $\theta_n \not\to 1$ in L^1. □

6 Conclusion

We have investigated the extension of the classical formula characterizing DE from pmp to mcp transformations. While the ratio ergodic theorem allows us to formalize an mcp analog of the ergodic formula, completing the proof using our method requires that the Radon-Nikodym cocycle has asymptotically equal diagonal and square averages. Our analysis shows that the condition for $\theta_n \to 1$ in L^1 is restrictive and can fail outside the pmp setting. Specifically, we proved that $\theta_n \not\to 1$ in L^1 for a Type I transformation due to the unbounded growth of θ_n, as well as for a Type III odometer action on the Cantor space. Because the actions considered are not DE, further investigation is required to determine whether the condition on θ_n holds for DE transformations. Finding a general characterization for DE will likely require alternative methods.

7 References

[DS23] Alexandre I. Danilenko and Cesar E. Silva. "Ergodic theory: nonsingular transformations". In: *Ergodic theory*. Encycl. Complex. Syst. Sci. Springer, New York, 2023, pp. 233–292. ISBN: 978-1-0716-2387-9. DOI: 10.1007/978-1-0716-2388-6_183.

[Hal46] Paul R. Halmos. "An ergodic theorem". In: *Proc. Nat. Acad. Sci. U.S.A.* 32 (1946), pp. 156–161. ISSN: 0027-8424. DOI: 10.1073/pnas.32.5.156.

[Hur44] Witold Hurewicz. "Ergodic theorem without invariant measure". In: *Ann. of Math. (2)* 45 (1944), pp. 192–206. ISSN: 0003-486X. DOI: 10.2307/1969081.

[MT24] Benjamin D. Miller and Anush Tserunyan. "A ratio ergodic theorem via tiling and uniformly syndetic markers". In: *Ergodic theory and dynamical systems*. De Gruyter Proc. Math. De Gruyter, Berlin, 2024, pp. 99–115. ISBN: 978-3-11-143503-9. DOI: 10.1515/9783111435503-004.

[Wal82] Peter Walters. *An introduction to ergodic theory*. Vol. 79. Graduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1982, pp. ix+250. ISBN: 0-387-90599-5.