A Borel–Cantelli lemma for non-stationary dynamical systems

Teodor Åberg

June 20, 2023

Abstract

In this thesis we present a Borel–Cantelli lemma for dynamical systems such that the dynamics driving the system varies with time. These results are largely based on a decomposition of transfer operators produced by Rychlik. Some limitations on the structure of sequences that is necessary for the results is given some discussion. Two examples of systems that could be considered are presented, one of which is accompanied by a method by Liverani for bounding rates of decay.

The function is one of the most fundamental, and recognizable, objects of study in mathematics. Being as widespread as they are the amount of questions that one could ask about their behaviour seems endless. If one has a function f, such that the output $y = f(x)$ is a valid input to the function one could examine what happens if one applies the funciton to its output, over and over again.

In the study of discrete dynamical systems questions of this sort are considered. What one can say about the evolution of the system varies drastically depending on what function is used. For instance one can find functions where the iterates approach a single point, no matter where one starts. On the other hand, one can also find functions where, even for points that are initially very close to one another, behaviour can vary wildly.

This thesis handles systems that fall into the second category. For these types of systems it is difficult to make statements for the iterates starting at some specific point. Instead one can try to find properties that are fulfilled very commonly. The property that is examined here is the so called Borel–Cantelli property, which can be understood as finding conditions such that one can say something about the value of

$$\underbrace{f(\dots f}_{n\,\text{times}}(x)\dots),$$

for infinitely many different values of n and for basically any initial point x.

Some known results of this type are presented, as well as a new result where, instead of a single function used when iterating, functions from a given family are used instead. Some adjacent previous results are also mentioned, as they are useful when examining the conditions that are necessary for the new result to be valid.

Acknowledgments

First and foremost I would like to thank my supervisor, Tomas Persson, for sharing his expertise on the subject and for the advice and tips that have been very helpful over the course of writing this thesis.

I would also like to thank my friends and coursemates for the support, taking many different forms, that they have offered.

Finally I would like to thank my parents, Jan and Eva-Ann, whose support and encouragement has been a great source of comfort over this period.

Contents

1 Introduction

The Borel–Cantelli lemma in probablity theory gives a condition that allows one to say if events from a given sequence are rare, in the sense that only a finite number of these events occur, or common. One can also pose a similar question for a dynamical system, namely if a member of a sequence of sets will be visited at some specific iteration an infinite number of times, and if this is a property that is fulfilled for almost the whole domain of the system.

There exists results which ensure that this property is fulfilled for some dynamical systems and sequences of sets. The results presented here will prove a sufficient condition for some dynamical systems that are not based on iterating one single map but instead the map that is used changes after some number of iterations. To produce these results the primary tool that is used will be the Perron–Frobenius operator, more specifically a decomposition of this operator that is presented by Rychlik in [9]. A result by Liverani in [7] is also presented, and used in order to bound coefficients, which are needed to check the conditions.

2 Ergodic systems and Borel–Cantelli lemmas

Before going into the details of the tools and main results of this thesis we will here give a brief introduction to ergodic theory, as well as to Borel–Cantelli lemmas which will be the primary focus.

While it is necessary to have some knowledge of measure theory when discussing ergodic thoery, as the results are based on this framework, we assume that the reader is already familiar with this in order to be able to stay on topic. For the remainder of this project λ will be used to denote the Lebesgue measure. More detailed explainations of the following concepts and examples from ergodic theory could most likely be found in some introductory textbook on the subject, for example [10].

In ergodic theory the subject of study is some measure space $(X, \mathcal{S}, \mu)$ and a μ-measurable map $T : X \to X$, that is measure-preseving with respect to μ, which is defined as the following.

Definition 2.1. A map $T : X \to X$ is called measure-preserving with respect to a measure μ if for any measurable set E

$$\mu(T^{-1}(E)) = \mu(E).$$

With this condition fulfilled we can alternatively say that μ is T-invariant.

Here $\mathcal{S}$ is a (σ-)algebra on X. More information on this structure and the connection to measurability can be found in the textbook mentioned above. As a first example we can consider the following.

Example 2.2 (The Doubling Map)**.** For $([0, 1], \mathcal{B}([0, 1]), \lambda, T)$ with

$$T(x) = 2x \mod 1 = \begin{cases} 2x & \text{for } x \in [0, \frac{1}{2}) \\ 2x - 1 & \text{for } x \in [\frac{1}{2}, 1] \end{cases}$$

where mod 1 means that only fractional part is considered. Then for any measurable set E

$$T^{-1}(E) = \frac{1}{2}(E) \cup \left(\frac{1}{2}(E) + \frac{1}{2}\right)$$

and so the Lebesgue measure is invariant with respect to T.

Let us also consider a case that is slightly more complicated.

Example 2.3 (Gauss map)**.** Here take $T : [0,1) \to [0,1)$ with

$$T(x) = \begin{cases} \frac{1}{x} \mod 1 & \text{for } x \in (0,1) \\ 0 & \text{for } x = 0 \end{cases},$$

and instead of λ as measure take

$$\mu = \frac{1}{\ln 2}\frac{1}{1+x}\lambda.$$

We wish to show that T is invariant with respect to $([0,1), \mathcal{B}([0,1)), \mu)$. It is sufficient to show that T is measure preserving for sets on the form

$$E = [0, a).$$

Then for any $a \in (0,1)$, the measure of E is $\mu(E) = \frac{1}{\ln(2)}\int_0^a \frac{1}{1+x}d\lambda(x) = \frac{\ln(1+a)}{\ln(2)}$. Each set $(\frac{1}{n+a}, \frac{1}{n}]$ maps to E, and without much additional work one can conclude that the preimage of E is

$$T^{-1}(E) = \bigcup_{n=1}^{\infty}(\frac{1}{n+a}, \frac{1}{n}],$$

the measure of which is

$$\begin{aligned}
\mu(T^{-1}(E)) &= \frac{1}{\ln(2)}\sum_{n=1}^{\infty}\mu((\frac{1}{n+a}, \frac{1}{n}]) \\
&= \frac{1}{\ln(2)}\sum_{n=1}^{\infty}\ln\left(\frac{1+\frac{1}{n}}{1+\frac{1}{n+a}}\right) \\
&= \frac{1}{\ln(2)}\sum_{n=1}^{\infty}(\ln(n+1) - \ln(n) + \ln(n+a) - \ln(n+1+a)) \\
&= \frac{1}{\ln(2)}\lim_{n\to\infty}(\ln\left(\frac{n+1}{n+1+a}\right) + \ln(1+a)) = \frac{\ln(1+a)}{\ln(2)}
\end{aligned}$$

meaning that μ is T-invariant.

This map is of interest as it relates to continued fraction expressions. Any irrational $x \in (0,1)$ will have an infinite continued fraction expression

$$x = \cfrac{1}{U_0 + \cfrac{1}{U_1 + \cfrac{1}{\ddots}}}$$

and applying by applying the Gauss map to this x gives

$$T(x) = \cfrac{1}{U_1 + \cfrac{1}{U_2 + \cfrac{1}{\ddots}}}.$$

Thus, with this perspective the Gauss map can be seen as a shift in the continued fraction expression. Additionally, one can identify k:th coefficient U_k of this expansion for some x by finding the interval $(\frac{1}{n+1}, \frac{1}{n})$ that contains $T^k(x)$.

Definition 2.4. A probability space $(X, \mathcal{S}, \mu)$ along with a μ-presering transform $T : X \to X$ is called ergodic if all strictly invariant sets A of T, meaning $A = T^{-1}(A)$ fulfill either $\mu(A) = 1$ or $\mu(A^c) = 1$.

Theorem 2.1 (Birkhoff Ergodic Theorem)**.** *Let $(X, \mathcal{S}, \mu)$ be a probablity space with μ-preserving transform T. If T is ergodic then for any integrable function $f : X \to \mathbb{R}$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int f(x) d\mu(x).$$

2.1 On Borel–Cantelli Lemmas

Let us start by stating the original formulation of the Borel–Cantelli lemma.

Lemma 2.2. *For a probablity space (Ω, C, P) and sequence of C-measurable events $\{E_k\}$ in Ω. If $\sum_{k=1}^{\infty} P(E_k) < \infty$ then*

$$P(\limsup E_k) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k\right) = 0.$$

If instead $\sum_{k=1}^{\infty} P(E_k) = \infty$ and all E_k are independent then

$$P(\limsup E_k) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k\right) = 1.$$

While expressed in this manner it may seem a touch cryptic though it has a rather intuitive interpretation, being that in the first case it is true with probability 1 that only a finite number of the events E_k will occur, and similarly for the second that with probability 1 infinitely many for the events E_k will occur.

A similar idea can be introduced when discussing dynamical systems, where one can talk of Borel–Cantelli sequences as well as strong Borel–Cantelli sequences. In this case we for some probability space $(X, \mathcal{S}, \mu)$ with μ-preserving maps T then we say that a sequence of sets $\{A_n\}_{n=0}^{\infty}$ with $S_n = \sum_{i=0}^{n-1} \mu(A_i) \to \infty$ is a Borel–Cantelli sequence if

$$\sum_{i=0}^{n-1} \mathbb{1}_{A_i} \circ T^i(x) = E_n(x) \to \infty \quad \mu\text{-a.e.}$$

and a sequence is called a strong Borel–Cantelli sequence if

$$\frac{E_n(x)}{S_n} \to 1 \quad \mu\text{-a.e.}$$

and so in this case $E_n(x)$ asymptotically grows like S_n, giving a quantitative estimate on how often $T^n(x) \in A_n$ occurs. Even more generally one can take a sequence of measurable functions $f_k(x)$ and instead letting $S_n = \sum_{i=0}^{n-1} \int f d\mu$ and $E_n(x) = \sum_{i=0}^{n-1} f_i(T^i(x))$ and again considering the limit above. Worth mentioning here is that for any ergodic T a constant sequence will be a strong Borel–Cantelli sequence as a direct consequence of Birkhoffs ergodic theorem, as it says that

$$\lim_{n\to\infty} E_n(x) = \lim_{n\to\infty} n\mu(A) = \lim_{n\to\infty} S_n.$$

It is known, from [8], that for the Gauss map any sequence of intervals of non-summable measure are strong Borel–Cantelli sequences. This then means, with μ being the previously introduced measure which is invariant with respect to the Gauss map, that the continued fraction expansion of μ-a.e. x will have $U_k \geq k$ for infinitely many k. This follows as the sequence $\{A_n\}_{n=1}^{\infty} = \left\{\left(0, \frac{1}{n}\right)\right\}_{n=1}^{\infty}$ fulfills

$$\sum_{n=1}^{N} \mu\left((0, \frac{1}{n})\right) > \frac{1}{2\ln(2)} \sum_{n=1}^{N} \frac{1}{n} \to \infty$$

meaning the sequence $\{A_n\}$ is Borel–Cantelli, which is an equivalent formulation of the desired property.

While it will not be of major necessity for the main results here let us also introduce the following result of Sprindžuk.

Lemma 2.3 ([11])**.** *Let $(X, \mathcal{S}, \mu)$ be a measure space, let $f_k(x)$ for $k \in \mathbb{N}^+$ be a sequence of nonnegative μ-measurable functions, let r_k, ϕ_k be sequences of real numbers such that $0 \leq r_k \leq \phi_k \leq M$ Suppose that*

$$\int \left(\sum_{n<k\leq m} f_k(x) - \sum_{n<k\leq m} r_k \right)^2 d\mu \leq C \sum_{n<k\leq m} \phi_k$$

for any integers $n < m$. Then

$$\sum_{k=1}^{n} f_k(x) = \sum_{k=1}^{n} r_k + O\left(\Phi^{\frac{1}{2}}(n) \ln^{\frac{3}{2}+\varepsilon}(\Phi(n))\right)$$

with $\Phi(n) = \sum_{k=1}^{n} \phi_k$.

This theorem serves as a powerful tool, which will most likely be encountered if one looks into some of the existing literature, as it gives a condition that directly implies that a sequence is strongly Borel–Cantelli.

3 Variation and the transfer operator

We begin by introducing the notion of the variation of a function as defined in [9].

Definition 3.1. The variation of a function $f : X \to \mathbb{R}$ on some subset C is given by

$$\mathrm{Var}_C f = \sup \left\{ \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| \right\},$$

where the supremum is take over sequences $(x_0, x_1, \ldots, x_k)$ with all $x_i \in C$ fulfilling $x_i \leq x_{i+1}$. Here $\mathrm{Var}\, f$ means $\mathrm{Var}_X\, f$.

With the variation defined one can also defined the function space BV as defined by Rychlik as

$$BV = \{f \in L_\infty \text{ and } f \text{ has version with bounded variation}\},$$

which can be shown to be a Banach-space with norm

$$\|f\|_{BV} = \max\left(\|f\|_1, \inf\{\mathrm{Var}\, \tilde{f}\} \text{ where } \tilde{f} \text{ is a version of } f\right),$$

where being a version here means that $f = \tilde{f}$ except for on some set of Lebesgue measure 0.

3.1 Approximations from variation

We here introduce two approximations that one can obtain using variation. The following lemmas will all assume that $f : X \to \mathbb{R}$ is μ-measurable and $\mu(X) = 1$. We start with a upper bound.

Lemma 3.1. *Let $f : X \to \mathbb{R} \in BV$, then*

$$\|f\|_\infty \leq |f(x)| + \mathrm{Var}\, f \quad \forall x \in X.$$

Proof. Assume the contrary. If there exists some $x^* \in X$ such that $|f(x^*)| = \|f\|_\infty$ then we have

$$\exists x \in X \text{ such that } \quad |f(x^*)| - |f(x)| > \mathrm{Var}(f)$$

But this is obviously a contradiction by the reverse triangle inequality and the fact that $\mathrm{Var}(f) \geq |f(x) - f(y)|$ for all x and y. If there is no such point x^* instead consider a sequence of points $\{x_n\}$ such that $|f(x_n)| \to \|f\|_\infty$. Then similarly

$$\exists x \in X \text{ such that } \quad \lim_{n\to\infty} |f(x_n)| - |f(x)| > \mathrm{Var}(f)$$

but this is again clearly a contradiction. $\square$

Similarly the variation of a function can be used to bound integrals from below by the following.

Lemma 3.2. *Let $f : X \to \mathbb{R} \in BV$, then*

$$\inf_X f \geq f(x) - \operatorname{Var} f \quad \forall x \in X.$$

Proof. Identical to proof for Lemma 3.1 □

For the cases that will be encountered working with the mean value of f will be preferrable to working with f directly though this is no problem due to the following lemma.

Lemma 3.3. *Let $f : X \to \mathbb{R} \in BV$, then there exists x_1, x_2 such that*

$$f(x_1) \leq \int_X f d\mu.$$
$$f(x_2) \geq \int_X f d\mu.$$

Proof. We only prove the first inequality. If this is false for all x

$$f(x) > \int f d\mu \Rightarrow \int f d\mu > \mu(X) \int f d\mu.$$

Obviously a contradiction. The second inequality is proven in an identical manner. □

3.2 Decomposition of Perron–Frobenius Operator

Before introducing the main tool that will be utilized we introduce the class of maps that is to be considered. The considered maps T are defined on $(0, 1)$ such that T' exists on a countable set of open disjoin intervals (a_i, b_i) fulfilling $\overline{\cup_i(a_i, b_i)} = [0, 1]$. It will be assumed that $\|\frac{1}{T'}\|_\infty < 1$ as well as $\operatorname{Var} \frac{1}{T'} < \infty$. The formulation of these conditions is similar to those in [6]. Let (A) denote these conditions.

The Perron–Frobenius operator, sometimes also called a transfer operator, P corresponding to some map T fulfilling (A) is defined as

$$P(f)(x) = \sum_{y=T^{-1}(x)} \begin{cases} \frac{f(y)}{|T'(y)|} & \text{Where } T'(y) \text{ is defined} \\ 0 & \text{Otherwise} \end{cases},$$

for $f \in L_1$. This operator has the following properties.

1. $P(f \circ T \cdot g) = f \cdot P(g)$
2. $\int_X P(f) d\lambda = \int f d\lambda$,

as well as P being linear.

Rychlik shows in [9] that in this case P can be decomposed as

$$P = \sum_i e_i Q_i + R$$

where e and Q are eigenvalues and projections onto eigenspaces to P respectivly, all e_i lie on the unit circle, and R as an operator $BV \to BV$ is a contraction such that $\inf_N \|R^N\|_{BV}^{\frac{1}{N}} \leq r < 1$. Considered here will only be the case of a single eigenvalue $e_1 = 1$, for which we have

$$Q(f) = \int f d\lambda \cdot h$$

with $h \in BV$ fulfilling $h \geq 0$ and $\int_X h d\lambda = 1$. Further more $QR = RQ = 0$. Note that as both P and Q are linear operators this must also be true for R. An equivalent and slightly more wieldy form of the bound on R expressed above, which will be the prefered form going forward, is

$$\|R^N(f)\|_{BV} \leq M \cdot r^N \|f\|_{BV} \tag{1}$$

for some positive M and $0 < r < 1$. We also note the that

$$\int R(f) d\lambda = 0 \quad \forall f \in BV.$$

Proof. Combining the fact that P preserves the integral with respect to the Lebesgue measure of f and $\int h d\lambda = 1$

$$\int f d\lambda = \int P(f) d\lambda = \int f d\lambda \cdot \int h d\lambda + \int R(f) d\lambda$$
$$\Rightarrow \int R(f) d\lambda = 0.$$

□

As a final note we bring up a theorem and corresponding proof by Kim to exemplify the usefulness of this decomposition when examining Borel–Cantelli properties.

Theorem 3.4. *[6] Let T be a interval map fulfilling (A). Assume that T has uniquely invariant measure which is absolutely continuous with respect to the Lebesgue measure $\mu = h d\lambda$ and h is bounded away from 0. If f_n is a sequence of nonnegative functions with $\sum \mu(f_n) = \infty$ and $\|f_n\|_{BV} < M$ then for a.e. x*

$$\lim_{N \to \infty} \frac{\sum_{n=1}^N f \circ T^n(x)}{\sum_{n=1}^N \mu(f_n)} = 1$$

Proof. For a sequence f_k fulfilling the conditions above then for $j > i$

$$\begin{aligned}
&\left| \int f_j \circ T^j \cdot f_i \circ T^i d\mu - \int f_j d\mu \cdot \int f_i d\mu \right| \\
&= \left| \int f_j \circ T^{j-i} \cdot f_i d\mu - \int f_j d\mu \cdot \int f_i d\mu \right| \\
&= \left| \int f_j \left(P^{j-i}(f_i h) - \int f_i h d\lambda h \right) d\lambda \right| \\
&\leq CMr^{j-i} \int |f_j| d\lambda \\
&\quad \leq DMr^{j-i} \int |f_j| d\mu,
\end{aligned}$$

giving an exponetial decay of correlations. With this

$$\begin{aligned}
&\int \left(\sum_{n<k\leq m} f_k \circ T^k(x) - \sum_{n<k\leq m} \mu(f_k) \right)^2 d\mu \\
&= \sum_{n<i,j\leq m} \int f_i \circ T^i(x) f_j \circ T^j(x) - \mu(f_i)\mu(f_j) d\mu \\
&\quad \leq (M + 2MD\frac{r}{1-r}) \sum_{n<i<m} \mu(f_i).
\end{aligned}$$

Applying Lemma 2.3 with $r_k = \phi_k = \mu(f_k)$ finishes the proof. $\square$

4 The considered problem

With this established we can consider the general formulation of the problem that will be considered. We will now instead consider a family of interval maps

$$T_i : X \to X \text{ for } i \in \mathcal{I},$$

where all T_i fulfill the conditions (A) described previously. As is natural we let P_i be the transfer operator corresponding to T_i. For a given sequence $\{i_k\}_{k=1}^{\infty}$ the corresponding system will under iteration act as

$$x \to T_{i_1}(x) \to T_{i_2} \circ T_{i_1}(x),$$

and so on. That is to say that the element of the sequence at index k defines the map that is applied at iteration k.

The reason that one could wish to study a system of this type could for example be to model behaviour where some perturbation leads to dynamics that change over time. Thus studying systems as described above could be used to identify structure that would be necessary for the system to be well behaved. Another situation where one could imagine encountering a situation like this would be to have a dynamical system that is dependent on some input which could be controlled, where this formulation would be connected to how one can find properties that are fulfilled by certain sequence.

As compared to systems that are traditionally considered when examining Borel–Cantelli properties there is no obvious choice of invariant measure that would be natural to use. From Rychliks results it is known that for the separate maps T_i fulfilling (A) there is an invariant measure μ_i, but since these are generally different this is not of assistance. We choose to use the Lebesgue measure λ when examining the iterates of the map above.

The sets that will be considered are going to be of the form

$$E_n = T_{i_1}^{-1} \circ T_{i_2}^{-1} \circ \cdots \circ T_{i_n}^{-1}(A_n)$$

where A_n are intervals such that $\sum_{n=1}^{\infty} \lambda(A_n) = \infty$. This formulation is helpful as it means that $x \in E_n \iff T_{i_n} \circ \cdots \circ T_{i_1}(x) \in A_n$, and so $\limsup E_n$ is the set such that the n:th iterate of x lies in A_n for infinitely many n. Since it is the measure of this set $\limsup E_n$ we are interested in this is natural.

Using this formulation one can simplify the measure of E_n to

$$\begin{aligned}\lambda(E_n) &= \int_X \mathbb{1}_{E_n}(x) d\lambda \\ &= \int_X \mathbb{1}_{T_{i_1}^{-1}\circ\cdots\circ T_{i_n}^{-1}(A_n)}(x) d\lambda \\ &= \int_X \mathbb{1}_{T_{i_2}^{-1}\circ\cdots\circ T_{i_n}^{-1}(A_n)} \circ T_1(x) d\lambda \\ &= \int_X \mathbb{1}_{A_n} \circ T_{i_n} \circ \cdots \circ T_{i_1}(x) d\lambda,\end{aligned}$$

from which the Perron-Frobenius operator can be used in order to simplify expressions of this sort further.

4.1 Imposed restrictions

In addition to conditions (A) that are imposed on the individual maps some restricitons are also made on the collection of maps T_i for $i \in \mathcal{I}$ that is used. We will require that there exist uniform bounds for M_i and r_i in (1). In addition to this we require that for all Q_i the corresponding functions h_i are uniformly bounded away from zero by some $1 \geq \gamma > 0$. A final restricton is that there is a uniform bound for $\|h_i\|_{BV}$. A collection T_i such that this is true is said to fulfill (B). We let $M \geq \sup_i M_i$, $1 > r \geq \sup_i r_i$ and $h_{max} > \|h_i\|_{BV}$ denote these uniform bounds if (B) is fulfilled. Additionally introduce $\Delta_{max} \geq \sup_{i \neq j} \|h_i - h_j\|_{BV}$, which will always exist by the triangle inequality.

Example 4.1. As an example of a collection of maps, closely related to the Gauss map discussed previously, that could be considered take

$$T_N(x) = \frac{N}{x} \mod 1.$$

These maps are monotone on each interval that they are continuous, and differentiable everywhere except for the points $x = \frac{n}{N}$ for all $n \geq N$, as is required. We now need to check that all maps that are considered fulfill (A). We have

$$\left\| \frac{1}{T_N'} \right\|_\infty = \frac{1}{N}$$

fulfilling the condition on the derivative as long, as $N = 1$ is excluded. For the variation the calculation is rather simple as $T_N' = -Nx^{-2}$ everywhere it is defined giving

$$\operatorname{Var} \frac{1}{T_N'} = \operatorname{Var} \frac{1}{N} x^2 = \frac{1}{N},$$

and so all T_N with $N \geq 2$ fulfill (A) as is desired. If we consider a finite family of these maps then the final condition that is needed is that all h_N are bounded away from zero by some γ. Luckily, this is indeed the case as the densities for the invariant measures are

$$h_N(x) = \frac{1}{\ln(N+1) - \ln(N)} \frac{1}{N+x}$$

as presented [5], for example. With this $h_N(x) \geq \frac{1}{(\ln(N+1)-\ln(N))} \frac{1}{N+1}$ and examining the denominator

$$\frac{\partial(\ln(N+1) - \ln(N))(N+1)}{\partial N} = \ln(1 + \frac{1}{N}) - \frac{1}{N} < 0$$

meaning the minimum value of h_N is increasing in N this collection of maps has $\gamma \geq \frac{1}{3\ln(1+\frac{1}{2})}$.

And so taking a finite family of these T_N with all $N \geq 2$ will satisfy (B). In order to have a infinite family it would require confirming existence of uniform bounds on r and M for (1) and a bound on $\|h_N\|_{BV}$. The latter is in this case rather simple $\|h_N\|_1 = 1$ always and the variation is strictly decreasing in N. Finding bounds on the coefficients in (1) on the other hand is something that rather difficult, unforunately. A brief discussion related to this is postponed until later.

Again this system of Gauss maps has a connection to continued fraction. If one instead considers a continued fraction expression, but with numerator N instead of 1, this gives for T_N

$$x = \cfrac{N}{U_0 + \cfrac{N}{U_1 + \cfrac{N}{\ddots}}} \to T_N(x) = \cfrac{N}{U_1 + \cfrac{N}{U_2 + \cfrac{N}{\ddots}}}.$$

Then T_N acts as a shift when considering these expansions, just as when $N = 1$. So one could then study a dynamical system taking compositions on different T_{N_i} in order to find properties of continues fraction expressions where the chosen numerators are not constant but instead vary. If any such system where T_{N_i}

is used at iteration i would fulfill a Borel–Cantelli lemma then one has similar results as for the case when $N \equiv 1$. Since $T_{N_i} \circ \dots \circ T_{N_1}(x) \in (\frac{N_i}{K+1}, \frac{N_i}{K})$, for some integer $K \geq N_i$ means that the ith coefficient U_i, of the continued fraction expansion of x with numerators defined by $\{N_i\}$, is larger than K. Then with $A_n = [0, \frac{1}{n})$, one could say that $U_i > N_i \cdot i$ occurs for inifinitely many i. Unfortunately the results that will be obtained are not quite as general as this, as some requirements on the structure of $\{N_i\}$ and $\{A_n\}$ will be imposed.

Another assumption will be that the sequence $\{i_k\}$ is constructed of blocks of length $l \leq l_n$ where all i in a block are the same,

$$i_1 = i_2 = \dots = i_{l_1} \qquad i_{l_1+1} = i_{l_1+2} = \dots = i_{l_1+l_2},$$

and so on. Though similar results may be possible to produce without an assumption like this it will be exemplified later why some assumption of this sort is necessary when using Rychliks decomposition of the transfer operator.

As was also noted in the example for requirements set on $\mathcal{I}$, all except one is fulfilled if the family is assumed to be finite. For this case the only condition that remains to be checked is that all functions h_i are bounded away from zero.

The requirement of long sequences of the same transfer operators is a technical requirement in order to get convergence when using the inequality (1) exemplified by the following. If no requirement is set on the length of blocks one will have to handle bounds like

$$\begin{aligned}\|R_{i_m} \circ R_{i_{m-1}} \circ \dots R_{i_1}(f)\|_{BV} &\leq M_m r_m \|R_{i_{m-1}} \circ \dots \circ R_{i_1}(f)\|_{BV} \\ &\leq (Mr)^m \|f\|_{BV}.\end{aligned}$$

From this we can see that if $Mr > 1$ this bound will grow expontentially and prove unusable for the results that are desired. With the requirement that each map is used in blocks of length $l_m > l$ these bounds will instead become

$$\begin{aligned}\|R_{i_m}^{l_m} \circ R_{i_{m-1}}^{l_{m-1}} \circ \dots R_{i_1}^{l_1}(f)\|_{BV} &\leq M_m r_m^{l_m} \|R_{i_{m-1}}^{l_{m-1}} \circ \dots \circ R_{i_1}^{l_1}(f)\|_{BV} \\ &\leq (Mr^l)^m \|f\|_{BV},\end{aligned}$$

meaning for any uniform bounds M and r there exists a minimum block size such that compositons of R_i of this sort can be bounded by an exponentially decreasing funciton, with rate of decay depending on this block size.

4.2 Other assumptions producing sufficient conditions for B.C-lemmas

For clarity we make sure to note that this does not imply that long block lengths would be a necessity in order to obtain analagous results with other or potentially no conditions on the structure of the sequence. For example take maps of the type considered by Rychlik but with the additional restriction that they only have a finite number of intervals of monotonicity. For these T_i we have

$$T_i : [0,1] \to [0,1] \; i \in \mathcal{I}, \qquad \left\| \frac{1}{T_i'} \right\|_\infty < 1, \qquad \operatorname{Var}\left(\frac{1}{T_i'}\right) < \infty.$$

Now one can consider a composition of these maps such that

$$T = T_{i_n} \circ \cdots \circ T_{i_1},$$

corresponding to periodic sequences. Then we have

$$T' = T_{i_1}' \cdot \prod_{k=2}^{n} T_{i_k}' \circ T_{i_{k-1}} \circ \cdots \circ T_{i_1},$$

where it is rather straight forward to see that

$$\left\| \frac{1}{T'} \right\|_\infty \leq \prod_{k=1}^{n} \left\| \frac{1}{T_{i_k}'} \right\|_\infty < 1.$$

Left is to verify that $\operatorname{Var}\left(\frac{1}{T'}\right) < \infty$. The first thing to note for $f, g \in BV$ and g with finite intervals of monotonicity

$$\operatorname{Var}(f \circ g) \leq n_{mon} \operatorname{Var}(f)$$

where n_{mon} is the number of intervals of monotonicity for g. This can be seen by examining the variation of $f \circ g$ over some interval of monotonicity of g, say (a,b). If it is the case that $g((a,b)) = (0,1)$, then

$$\operatorname{Var}(f) = \operatorname{Var}_{(a,b)}(f \circ g),$$

and if on the other hand some interval of monotonicity does not span the entire interval, so $g((a,b)) = (c,d)$ instead, it will be that case that

$$\operatorname{Var}_{(a,b)}(f \circ g) = \operatorname{Var}_{(c,d)}(f) \leq \operatorname{Var}(f).$$

Summing over all intervals of monotonicity of g then gives

$$\operatorname{Var}(f \circ g) = \operatorname{Var}(f) n_{mon}.$$

By repeating this argument this can then be generalized to

$$\operatorname{Var}(f \circ g_m \circ \cdots \circ g_1) \leq \operatorname{Var}(f) \prod_{k=1}^{m} n_{mon_k},$$

which when combined with

$$\operatorname{Var}(fg) \leq \|g\|_\infty \operatorname{Var} f + \|f\|_\infty \operatorname{Var} g$$

can be used in order to establish the bound

$$\mathrm{Var}\left(\frac{1}{T'}\right) \leq \sum_{k=1}^{m} \mathrm{Var}\left(\frac{1}{T_k'}\right) \prod_{j=1}^{k-1} n_{mon_j},$$

meaning T is a uniformly expanding map whose with inverse derivative of bounded variation. So in this specific case the problem can be handled as if only a single map is considered and perticularly if one is able to find a lower bound for the invariant density of this map it would be possible to apply the result of Kim for this map.

5 Main results

We are now ready to state the main result.

Theorem 5.1. *Let $T_i \in \mathcal{I}$ be a family of maps each fulfilling condition (A) individually and (B) as a collection. Then for l that fulfills $Mr^l < 1$ and*

$$\frac{Mr^l}{1-(Mr^l)} < \frac{\gamma}{\Delta_{max}}$$

and maps T_i appear in blocks of length $l \leq l_n < \infty$ for all $n \in \mathbb{N}$ then for any sequence of intervals $\{A_n\}$ such that $\sum_{n=1}^{\infty} \lambda(A_{\sum_{k=1}^{n} l_k}) = \infty$ is a Borel–Cantelli sequence.

It is worth noting here that the divergence required of the measures subsequence of $\{A_n\}$ is implied by the divergence of the measures of the entire sequence, if we also assume that there is some uniform upper bound on l_k and that $\lambda(A_n)$ is decreasing.

As a point of comparison take Proposition 5.2 in [1]. This states that Random Dynamical Systems fulfill a Strong Borel–Cantelli property for almost every sequence of maps T_i and initial argument x. The difference between these results is that in the theorem presented here one is able to draw conclusions about behaviour of specific sequences, though it comes with the downside of requiring specific structure for these sequences.

To prove the desired result we will use two known results. The first of which being the following lemma.

Lemma 5.2 (Chung–Erdős inequality [2])**.** *If $\{E_n\}_{n=1}^{N}$ is a sequence of sets in a probability space X with $\mu\left(\bigcup_{n=1}^{N} E_n\right) > 0$ then*

$$\mu(\bigcup_{n=1}^{N} E_n) \geq \frac{\left(\sum_{n=1}^{N} \mu(E_n)\right)^2}{2\sum_{1 \leq n < m \leq M} \mu(E_n \cap E_m) + \sum_{n=1}^{N} \mu(E_n)}.$$

As well as the following theorem.

Theorem 5.3 (Lebesgues density theorem [3]). *Let E be a Lebesgue measureable subset of $\mathbb{R}^d$ then λ-almost every point of E is a point of density of E, and λ-almost every point of E^c is a point of dispersion of E.*

That x is a point of density of E means

$$\lim_{r\to 0} \frac{\lambda(E \cap B_r(x))}{\lambda(B_r(x))} = 1,$$

and similariy x being a point of dispersion means

$$\lim_{r\to 0} \frac{\lambda(E \cap B_r(x))}{\lambda(B_r(x))} = 0.$$

With these two results as motivation, the first step is to find appropriate upper and lower bounds for $\lambda(I \cap E_n)$, as well as a lower bound on $\lambda(I \cap E_n \cap E_m)$ for arbitrary intervals I.

We note here that while the full sequence of E_n will generally be on the form

$$E_n = T_{i_1}^{-l_1} \circ \cdots \circ T_{i_m}^{-l_m} \circ T_{i_{m+1}}^{-k}(A_n)$$

where $\sum_{i=1}^m l_i + k = n$ it will be sufficient to consider only the elements that end consecutive applications of the same map under the assumptions set up in theorem 5.1, as if this subsequence is Borel–Cantelli this will be the case for the complete sequence too. Thus we introduce

$$\begin{cases} A_n^* &= A_{\sum_{k=1}^n l_k} \\ E_n^* &= T_{i_1}^{-l_1} \circ \cdots \circ T_{i_n}^{-l_n}(A_n^*) \end{cases}$$

which will be used when the following bounds are being constructed. The fact that it is sufficient to consider a subsequence is due to $\limsup E_n^*$ being a subset of $\limsup E_n$ and so $\lambda(\limsup E_n^*) = 1$ implies $\lambda(\limsup E_n) = 1$.

We note here that the result can also be formulated by allowing A_n^* be any element is atleast l iterations in the nth block. The proof essentially stays identical, and the required restriction on the diverging sum of measures is slightly less restrictive.

5.1 Evaluating composition of Transfer-Operators

Before finding the necessary bounds needed it is worthwhile to take a quick detour in order to establish how composition of transfer operators works for their decomposition. Starting with the simplest case of the composition of two transfer operators as an illustative example for $f \in BV$

$$\begin{aligned} P_2^{l_2} \circ P_1^{l_1}(f) &= P_2^{l_2}\left(\int f d\lambda h_1 + R_1^{l_1}(f)\right) \\ &= Q_2\left(\int f d\lambda h_1 + R_1^{l_1}(f)\right) + R_2^{l_2}\left(\int f d\lambda h_1 + R_1^{l_1}(f)\right). \end{aligned}$$

We know that Q_i acts as a projection onto h_i. Introducing $\Delta_{1,2} = h_1 - h_2$ and using the fact that $\int h_i d\lambda = 1$ and $\int R_i(f) d\lambda = 0$ for all i means that the first term becomes

$$Q_2\left(\int f d\lambda \cdot (h_2 + \Delta_{1,2}) + R_1^{l_1}(f)\right) = \int f d\lambda \cdot h_2.$$

As we also know that operators R_i and Q_i fulfill $R_i(Q_i) = 0$ this means the second term can be simplified as

$$R_2^{l_2}\left(\int f d\lambda \cdot h_1 + R_1^{l_1}(f)\right) = R_2^{l_2}(\int f d\lambda \cdot \Delta_{1,2} + R_1^{l_1}(f)).$$

In order to keep the following expressions a more compact we introduce

$$\hat{\rho}_n^m(f) = R_m^{l_m} \circ \cdots \circ R_n^{l_n}(f).$$

With this in mind it would be useful if one could make similar simplifications for longer compositions of these For longers compositions of transfer operators we introduce the following lemma.

Lemma 5.4. *With P_i being transfer operators corresponding to T_i such that both (A) and (B) are fulfilled. Then for $f \in BV$*

$$P_m^{l_m} \circ \cdots \circ P_1^{l_1}(f) = \int f d\lambda \left(h_m + \sum_{n=2}^{m} \hat{\rho}_n^m(\Delta_{n-1,n})\right) + \hat{\rho}_1^m(f).$$

Proof. The base $N = 2$ has been shown above and its validity for $N = 1$ is a direct consequence of the decomposition of the transfer operator. So now we examine

$$\begin{aligned}
&P_{m+1}^k\left(\int f d\lambda \left(h_m + \sum_{n=2}^{m} \hat{\rho}_n^m(\Delta_{n-1,n})\right) + \hat{\rho}_1^m(f)\right) \\
&\quad = Q_{m+1}\left(\int f d\lambda (h_{m+1} + \Delta_{m,m+1})\right) \\
&\quad + R_{m+1}^k(\int f d\lambda (h_{m+1} + \Delta_{m,m+1} + \sum_{n=2}^{m} \hat{\rho}_n^m(\Delta_{n-1,n})) + \hat{\rho}_1^m(f)).
\end{aligned}$$

Again using the same properties used for finding the expression for the base case is further simplified to

$$\int f d\lambda \left(h_{m+1} + R_{m+1}^k(\Delta_{m,m+1} + \sum_{n=2}^{m} \hat{\rho}_n^m(\Delta_{n-1,n}))\right) + R_{m+1}^k(\hat{\rho}_1^m(f))$$

and with $k = l_{m+1}$

$$\int f d\lambda \left(h_{m+1} + \sum_{n=2}^{m+1} \hat{\rho}_n^{m+1}(\Delta_{n-1,n})\right) + \hat{\rho}_1^{m+1}(f). \qquad \square$$

5.2 Lower bound for $\lambda(E_n^* \cap I)$

From the definitions of E_n^* means that the measure first can be rewritten as

$$\lambda(E_n^* \cap I) = \int \mathbb{1}_{E_n^*} \cdot \mathbb{1}_I d\lambda = \int \mathbb{1}_{A_n^*} \circ T_{i_n}^{l_n} \circ \cdots \circ T_{i_1}^{l_1}(x) \cdot \mathbb{1}_I(x) d\lambda.$$

By applying the composed operators $P_{i_n}^{l_n} \circ \cdots \circ P_{i_1}^{l_1}$ to the integrand we have

$$\begin{aligned}
\lambda(E_n^* \cap I) &= \int \mathbb{1}_{A_n^*} \cdot P_{i_n}^{l_n} \circ \cdots \circ P_{i_1}^{l_1}(\mathbb{1}_I) d\lambda \\
&= \int \mathbb{1}_{A_n^*} \left(\lambda(I) h_n + \lambda(I) \sum_{k=2}^{n} \hat{\rho}_k^n \left(\Delta_{k-1,k}\right) + \hat{\rho}_1^n(\mathbb{1}_I) \right) d\lambda \\
&\geq \gamma \lambda(I) \int \mathbb{1}_{A_n^*} d\lambda - \int \mathbb{1}_{A_n^*} d\lambda \\
&\qquad \cdot \operatorname{Var} \left(\hat{\rho}_1^n(\mathbb{1}_I) + \lambda(I) \sum_{k=2}^{n} \hat{\rho}_k^n \left(\Delta_{k-1,k}\right) \right) \\
&\geq \lambda(A_n^*) \left(\lambda(I)\gamma - \left\| \lambda(I) \sum_{k=2}^{n} \hat{\rho}_k^n (\Delta_{k-1,k}) + \hat{\rho}_1^n(\mathbb{1}_{A_n^*}) \right\|_{BV} \right).
\end{aligned}$$

Where the first inequality is due to fact that we have a uniform lower bound $h_i > \gamma$, $\int R_i(f) d\lambda = 0$ for all $f \in BV$ and Lemma 3.2.

Finally using the bound (1) and the triangle inequality gives

$$\lambda(E_n^* \cap I) \geq \lambda(A_n^*) \left(\lambda(I) \left(\gamma - \Delta_{max} \sum_{k=1}^{n-1} (Mr^l)^k \right) - \|\mathbb{1}_I\|_{BV} (Mr^l)^n \right).$$

While this bound is generally true, it is not of any use unless the bounding coefficient is greater than 0. Now under the assumptions on Mr^l in Theorem 5.1 the limit of the geometric sum can be taken, resulting in

$$\begin{aligned}
\lambda(E_n^* \cap I) &\geq \lambda(A_n^*) \left(\lambda(I) \left(\gamma - \frac{Mr^l \|\Delta_{max}\|_{BV}}{1 - Mr^l} \right) - \|\mathbb{1}_I\|_{BV} (Mr^l)^n \right) \\
&\geq \lambda(A_n^*) \left(\lambda(I)\varepsilon - (Mr^l)^n 2 \right),
\end{aligned} \tag{2}$$

for some $\varepsilon > 0$. This bound means that as long as n is chosen large enough this will always yield a bound that is strictly greater than zero as is desired. If one could find a tighter bound than that which is used here it may be possible to weaken the restriction set on Mr^l, allowing the result to be applied to a larger set of sequences.

5.3 Upper bounds for $\lambda(E_n^* \cap I)$ and $\lambda(E_n^* \cap E_m^* \cap I)$

Starting exactly the same way as when constructing the lower bounds gives

$$\begin{aligned}
\lambda(E_n^* \cap I) &= \int \mathbb{1}_{A_n^*} \cdot P_{i_n}^{l_n} \circ \cdots \circ P_{i_1}^{l_1}(\mathbb{1}_I) d\lambda \\
&= \int \mathbb{1}_{A_n^*} \left(\lambda(I) \left(h_n + \sum_{k=2}^{n} \hat{\rho}_k^n(\Delta_{k-1,k}) \right) + \hat{\rho}_{\mathbb{1}}^n(\mathbb{1}_I) \right) d\lambda \\
&\leq \lambda(A_n^*) \left(\lambda(I) + \mathrm{Var}(\lambda(I)(h_n + \sum_{k=2}^{n} \hat{\rho}_k^n(\Delta_{k-1,k})) + \hat{\rho}_1^n(\mathbb{1}_I)) \right) \\
&\leq \lambda(A_n^*) \Bigg(\lambda(I)(1 + \|h_{max}\|_{BV} \\
&\qquad + \|\Delta_{max}\|_{BV} \sum_{k=1}^{n-1} (Mr^l)^k) + (Mr^l)^n 2 \Bigg) \\
&\leq \lambda(A_n^*) \left(\lambda(I)(1 + \|h_{max}\|_{BV} + \gamma - \varepsilon) + (Mr^l)^n 2 \right).
\end{aligned}$$

Here the first inequality follows by using lemma 3.1 and the fact that all h_i have meanvalue 1. A result which then for $m > n$ can be extended to include the intersection between E_n, E_m and an arbitrary interval I.

$$\begin{aligned}
&\lambda(E_n^* \cap E_m^* \cap I) \\
&= \int \mathbb{1}_{A_n^*} \circ T_{i_n}^{l_n} \circ \ldots T_{i_1}^{l_1}(x) \\
&\qquad \cdot \mathbb{1}_{A_m^*} \circ T_{i_m}^{l_m} \circ \cdots \circ T_{i_{n+1}}^{l_{n+1}} \circ T_{i_n}^{l_n} \circ \ldots T_{i_1}^{l_1}(x) \cdot \mathbb{1}_I(x) d\lambda \\
&= \int \mathbb{1}_{A_n^*}(x) \cdot \mathbb{1}_{A_m^*} \circ T_{i_m}^{l_m} \circ \cdots \circ T_{i_{n+1}}^{l_{n+1}}(x) \cdot P_{i_n}^{l_n} \circ \cdots \circ P_{i_1}^{l_1}(\mathbb{1}_I) d\lambda \\
&\leq \Bigg(\lambda(I)(1 + \|h_{max}\|_{BV} + \|\Delta_{max}\|_{BV} \sum_{k=1}^{n-1} (Mr^l)^k) \\
&\qquad + (Mr^l)^n \|\mathbb{1}_I\|_{BV} \Bigg) \cdot \int P_{i_m}^{l_m} \circ \cdots \circ P_{i_{n+1}}^{l_{n+1}}(\mathbb{1}_{A_n^*}) \cdot \mathbb{1}_{A_m^*} d\lambda \\
&\leq \lambda(A_m^*) \left(\lambda(I)(1 + \|h_{max}\|_{BV} + \|\Delta_{max}\|_{BV} \sum_{k=1}^{n-1} (Mr^l)^k) + (Mr^l)^n 2 \right) \\
&\qquad \cdot \left(\lambda(A_n^*)(1 + \|h_{max}\|_{BV} + \|\Delta_{max}\|_{BV} \sum_{k=1}^{m-(n+1)} (Mr^l)^k) + (Mr^l)^{m-n} 2 \right).
\end{aligned}$$

5.4 Full measure for $E^* = \limsup E_n^*$

All necessary inequalities needed to use the Chung–Erdős inequality have now been established, though for readability introduce

$$\begin{aligned} C =& 1 + \|h_{max}\|_{BV} + \gamma - \varepsilon \\ S_1 =& \sum_{N_0 \le n < m \le M_0} \lambda(A_n^*)\lambda(A_m^*) \\ S_2 =& \sum_{N_0 \le n < m \le M_0} (Mr^l)^{m-n}\lambda(A_m^*) \\ S_3 =& \sum_{n=N_0}^{M_0} \lambda(A_n^*), \end{aligned}$$

with ε being the same as that used in (2). Inserting all the bounds that have been found results in

$$\begin{aligned} \lambda\left(\bigcup_{n=N_0}^{M_0} (E_n^* \cap I)\right) &\ge \frac{(\lambda(I)\varepsilon - 2(Mr^l)^{N_0})^2 S_3^2}{(\lambda(I)C + 2(Mr^l)^{N_0})\big(2CS_1 + 4S_2 + S_3\big)} \\ &= \frac{(\lambda(I)\varepsilon - 2(Mr^l)^{N_0})^2}{(\lambda(I)C + 2(Mr^l)^{N_0})\left(2C\frac{S_1}{S_3^2} + 4\frac{S_2}{S_3^2} + \frac{1}{S_3}\right)}, \end{aligned}$$

where N_0 is chosen large enough such that $(Mr^l)^{N_0} < \lambda(I)\frac{\varepsilon}{4}$.

Now we want to find bounds for the fractions in the denominator. We begin with $\frac{S_2}{S_3^2}$. The sum S_2 can be approximated as

$$S_2 = \sum_{m=N_0+1}^{M_0} \lambda(A_m) \sum_{i=1}^{m-1} (Mr^l)^i \le \frac{(Mr^l)}{1 - Mr^l} S_3.$$

Hence $\frac{S_2}{S_3^2} \le \frac{Mr^l}{(1-Mr^l)S_3}$. For $\frac{S_1}{S_3^2}$ rewriting S_3^2 as

$$\begin{aligned} S_3^2 &= 2 \sum_{N_0 \le n < m \le M_0} \lambda(A_n)\lambda(A_m) + \sum_{n=N_0}^{M_0} \lambda(A_n)^2 \\ &\ge 2 \sum_{N_0 \le n < m \le M_0} \lambda(A_n)\lambda(A_m) \Rightarrow \frac{S_1}{S_3^2} \le \frac{1}{2} \end{aligned}$$

gives the bound

$$\lambda\left(\bigcup_{n=N_0}^{M_0} E_n^*\right) \ge \frac{(\lambda(I)\frac{\varepsilon}{2})^2}{\lambda(I)\left(C + \frac{\varepsilon}{2}\right)\left(C + \frac{1}{S_3}\left(\frac{4Mr^l}{1-Mr^l} + 1\right)\right)}.$$

Now using the fact that S_3 is divergent we can let M_0 go to infinity, denoting the limit as

$$k = \frac{(\lambda(I)\frac{\varepsilon}{2})^2}{\lambda(I)\left(C + \frac{\varepsilon}{2}\right)C}.$$

Thus, for any small enough $\delta > 0$ we can find M_0 such that

$$\frac{(\lambda(I)\frac{\varepsilon}{2})^2}{\lambda(I)\left(C + \frac{\varepsilon}{2}\right)\left(C + \frac{1}{S_3}\left(\frac{Mr^l}{1-Mr^l} + 1\right)\right)} > \frac{(\lambda(I)\frac{\varepsilon}{2})^2}{\lambda(I)\left(C + \frac{\varepsilon}{2}\right)C} - \delta.$$

We then set M_0 as the smallest value such that the inequality above is fulfilled with $\delta = \frac{k}{2}$. Thus for a given I some N_0 and a corresponding M_0 can be found such that

$$\frac{\lambda\left(\bigcup_{n=N_0}^{M_0}(E_n^* \cap I)\right)}{\lambda(I)} \geq \frac{\varepsilon^2}{4\left(C + \frac{\varepsilon}{2}\right)C} > 0.$$

For any I we can find a sequence of $\{N_i\}$ such that N_0 fulfills the condition set previously and $N_{i+1} - 1$ fulfills the condition on M_0 described above relative to N_i. Since all $N_i > N_0$ so the condition on the lower boundary will always be fulfilled meaning we for our sequence have

$$\frac{\lambda\left(\bigcup_{n=N_i}^{N_{i+1}-1}(E_n^* \cap I)\right)}{\lambda(I)} \geq \frac{\varepsilon^2}{4\left(C + \frac{\varepsilon}{2}\right)C} > 0.$$

With $B_m^* = \bigcup_{l=m}^{\infty}\left(\bigcup_{n=N_l}^{N_{l+1}-1} E_n^* \cap I\right)$ we have that

$$\lambda(B_m^*) \geq \frac{\varepsilon^2}{4\left(C + \frac{\varepsilon}{2}\right)C}\lambda(I),$$

and by extension

$$\frac{\lambda(E^* \cap I)}{\lambda(I)} = \frac{\lambda(\limsup E_n^* \cap I)}{\lambda(I)} = \lambda\left(\bigcap_{m=1}^{\infty} B_m\right)\frac{1}{\lambda(I)} \geq \frac{\varepsilon^2}{4\left(C + \frac{\varepsilon}{2}\right)C}.$$

This then means that E^* has non-zero density over all intervals of $(0,1)$. Now see this consider $I_n = B(x, r_n)$ where $\{r_n\}$ is a sequence approaching zero. Thus for any x

$$\lim_{n\to\infty}\frac{\lambda(E^* \cap I_n)}{\lambda(I_n)} \geq \frac{\varepsilon^2}{4\left(C + \frac{\varepsilon}{2}\right)C}.$$

But by Theorem 5.3 we know that λ-almost every $x^* \in E^c$

$$\lim_{r\to 0}\frac{\lambda(E^* \cap B_r(x^*))}{B_r(x^*)} = 0$$

and since no x^* exist such that this is true $\lambda(E^{*c}) = 0$ thus $\lambda(E^*) = 1$ and the desired result has been shown.

5.5 Construction of bounds on r and M

In order to attain bounds on l such that we can ensure that this length of blocks is sufficient, as well as ensuring that infinite families have uniform bounds for r_i and M_i it is of importance to be able to find estimates for these.

In [7] Liverani presents a method that allows for finding numerical approximations for these constants for some maps similar to those that have been considered previously, though specifically finite partitions are considered as well as T being assumed to be twice continuously differentiable inside each interval in the partition. The bound that is established gives expressions for constants Λ and b, as well as a decreasing sequence K_n, such that for $g \in BV$ with $\int g d\lambda = 1$ and for $f \in L_1$

$$\begin{aligned}\left|\int g \cdot f \circ T^n d\lambda - \int f \cdot h d\lambda\right| \leq & \Lambda^n K_n \left(1 + b \operatorname{Var}(g)\right) \|f\|_1 \\ \leq & 2K_n \cdot \max(1, b) \|g\|_{BV} \|f\|_1.\end{aligned}$$

Since all that is required for the results of theorem 5.1 are bounds for Mr^n for $n > l$ one could take $M = 2K_l \cdot \max(b, 1)$ and $r = \Lambda$, which can give lower bounds for l, as compared to $M = 2K_1 \cdot \max(b, 1)$ which would be necessary in order to assure that the bound is valid for any n. Additionally, the bounds that are produced here are continuous with respect to the C^2 topology, which means that for some T one could find a neighbourhood of functions such that only a small change in l is necessary for all functions in this neighbourhood. We illustrate how these bounds can be construction with an example.

Example 5.1. Here we consider the map $T : [0, 1) \to [0, 1)$

$$T(x) = 3x \mod 1$$

which is twice continuously differentiable except for outside the partition $\mathcal{P}_0 = \{(0, \frac{1}{3}), (\frac{1}{3}, \frac{2}{3}), (\frac{2}{3}, 1)\}$ The first two values that are needed can be found rather straight forward

$$\begin{cases}\lambda = \inf |T'| = 3 \\ A = \sup_{x\in[0,1]} \frac{|T''(x)|}{|T'(x)|} + 2 \sup_{I\in\mathcal{P}_0} \frac{\sup_{x\in I} |T'(x)|}{|I|} = 2.\end{cases}$$

Additionally we introduce a that is required to fulfill $a > \frac{A}{1-2\lambda^{-1}} = 6$. We let $\mathcal{P}_n$ denote the partition such that T^{n+1} is monotone and differentiable on each element in this set. For this map we have

$$\mathcal{P}_n = \left\{\left(\frac{k}{3^{n+1}}, \frac{k+1}{3^{n+1}}\right) \mid k = 0, 1 \ldots 3^{n+1}\right\}$$

and need n_0 such that the partition fulfills $\min_{I\in\mathcal{P}_n} |I| \leq \frac{1}{2a}$ meaning $n_0 = 2$ could give a valid bound. $N(n_0)$ being the number of iterates of T needed such that $T^{N(n_0)}(I) = [0, 1)$ for all $I \in \mathcal{P}_{n_0}$ which without much trouble can be seen to be 3. With $a = 6.75$, then $\sigma = (2\lambda^{-1})^{N(2)} + \frac{1-(2\lambda^{-1})^{N(n_0)}}{1-2\lambda^{-1}} A a^{-1} \leq \frac{25}{27}$ we introduce

$$\Delta \leq 2 \ln \left((1 + a\sigma) \, 2 \cdot 3^3 \right),$$

resulting in the following bounds for the necessary constants

$$\begin{cases} \Lambda = \tanh\left(\frac{\Delta}{4}\right)^{(3^{-1})} \leq 0.998 \\ K_n = e^{\Delta \Lambda^{n-3}} \Lambda^{-3} \Delta \\ b = (a - \frac{A}{1-2\lambda^{-1}}) = 0.75. \end{cases}$$

The smallest possible l that may be valid must always fulfill $Mr^l < 1$, and so we can compare these bounds for the choice for $M_1 = 2K_1$ and $M_2 = 2K_l$, with $r = \Lambda$ for both situations. For this we introduce the Lambert W function, being increasing for $x > 0$ and fulfilling $W(x)e^{W(x)} = x$, for further details see, for instance, [4]. These smallest l then can be bounded by

$$l_1 > -\ln(2\Delta e^{\Delta \Lambda^{-3}} \Delta \Lambda) \ln(\Lambda)^{-1} \approx 7550,$$

and

$$\Delta \Lambda^{l_2-3} e^{\Delta \Lambda (l_2-3)} \leq \frac{1}{2} \Rightarrow \Delta \Lambda^{l_2-3} < W\left(\frac{1}{2}\right)$$

$$\Rightarrow l_2 > \ln\left(W\left(\frac{1}{2}\right) \Delta^{-1} \right) \ln(\Lambda)^{-1} + 3 \approx 1800.$$

From this example we see that the bounds that this method produces approximations for M and r and so this can be used in order so bound the necessary block length l. These bounds are still rather far away from the results that would be desired, that being $Mr < 1$. If this is fulfilled for some family, such that Δ_{max} can be made arbitrarily small, then a blocksize of $l = 1$ would be possible and so any sequence of maps would fulfill the Borel–Cantelli property for intervals. Still this is rather far away from the bounds constructed here.

6 Future work

With the established condition found it is natural to ask if one could weaken this condition further. Ultimately the question would be if any sequence of functions T_i of the type considered here always fulfills the Borel–Cantelli property or if it is possible to construct a sequence for which this would fail to hold. Another direction that is of relevance for these results would be of finding additional methods of estimating the bounds for $\|R^n(f)\|_{BV}$ as good bounds on these are necessary in order to find sequences fulfilling the conditions in Theorem 5.1.

References

[1] Romain Aimino, Matthew Nicol, and Sandro Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probability Theory and Related Fields*, 162(1-2):233–274, jul 2014.

[2] K. L. Chung and P. Erdős. On the application of the borel-cantelli lemma. *Transactions of the American Mathematical Society*, 72:179–186, 1952.

[3] Donald L. Cohn. *Measure Theory*. Springer Science and Business Media LLC, 2nd edition, 2013.

[4] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the lambert w function. *Advances in Computational Mathematics*, 5:329–359, 1996.

[5] Hiroaki Ito. Diophantine approximation by negative continued fraction, 2020.

[6] Dong Han Kim. The dynamical borel-cantelli lemma for interval maps. *Discrete and Continuous Dynamical Systems*, 17(4):891–900, 2007.

[7] Carlangelo Liverani. Decay of correlations for piecewise expanding maps. *Journal of Statistical Physics*, 78, 1995.

[8] Walter Philipp. Some metrical theorems in number theory. *Pacific Journal of Mathematics*, 20(1), 1967.

[9] Marek Rychlik. Bounded variation and invariant measures. *Studia Mathematica*, 76(1):69–80, 1983.

[10] César Ernesto Silva. *Invitation to ergodic theory*. American Mathematical Society, Providence, R.I., cop. 2008.

[11] Vladimir Gennadievič Sprindžuk. *Metric theory of diophantine approximations*. Wiley : distributed soley by Halsted Press, New York, 1979.