

Some results on linear dynamical systems

by

George Lee

MSc, University of Glasgow, 2017

BA, University of Cambridge, 2016

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mathematics and Statistics

© George Lee, 2023
University of Victoria

This work is licensed under a Creative Commons “Attribution 4.0 International” license.



Some results on linear dynamical systems

by

George Lee

MSc, University of Glasgow, 2017

BA, University of Cambridge, 2016

Supervisory Committee

Dr. Anthony Quas, Supervisor
(Department of Mathematics and Statistics)

Dr. Chris Bose, Departmental Member
(Department of Mathematics and Statistics)

Dr. Sean Chester, Outside Member
(Department of Computer Science)

ABSTRACT

A linear cocycle is an object that arises naturally in the study of dynamical systems and statistics. Oseledets' multiplicative ergodic theorem [22] guarantees a decomposition of a linear space of states into equivariant subspaces that grow logarithmically at rates corresponding to the Lyapunov exponents. Theorem 66 is the main result of this thesis, a semi-invertible version of this theorem: the ergodic system is invertible, the state space is a separable Banach space and the cocycle is strongly measurable and forward integrable, but with no invertability or injectivity assumptions. Past results have implicitly or explicitly made extra assumptions on the underlying cocycle, but in this work no injectivity or separability of the dual is assumed, and no prior version of the result or other overly unconstructive machinery is used in obtaining the direct sum decomposition of the state space.

Table of Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	iv
Acknowledgements	vii
Dedication	viii
Chapter 1 Ergodic and linear background	1
1.1 Ergodic theory	1
1.1.1 Introduction	1
1.1.2 Banach space: justifications and notation	2
1.1.3 Measurable spaces	2
1.1.4 Measures	4
1.1.5 Null sets	5
1.1.6 Measure preservation	5
1.1.7 Ergodicity	6
1.1.8 Standard and Lebesgue probability spaces	7
1.1.9 Discretising continuous time systems	9
1.2 Stochastic or deterministic dynamics?	10
1.2.1 Generalising from shifts and iid statistics: the ergodic assumption	11
1.2.2 Examples from topology and geometry	11
Chapter 2 Towers and subadditive theorems	14
2.1 Towers	14
2.1.1 Overview	14
2.1.2 Representation by return time	14
2.1.3 Partitions and processes	15

2.1.4	A Rokhlin lemma	16
2.1.5	The backward Vitali lemma	18
2.2	A balanced subadditive ergodic theorem	19
2.2.1	Subadditive families	19
2.2.2	Kingman theorems	19
Chapter 3	The Grassmannian	27
3.1	Definition	27
3.1.1	A familiar pattern in the construction of measurable maps in a separable setting	30
3.2	Measurable statistics and the Grassmannian	31
3.2.1	A motivating example: singular value decomposition	31
3.2.2	Measurable statistics for linear operators	31
3.2.3	Equivalence of growth statistics	35
Chapter 4	Cocycles	37
4.1	Motivation	37
4.1.1	What is a cocycle?	37
4.1.2	Differential examples	39
4.2	Defining the Lyapunov exponents	39
4.2.1	Applying statistics from the Grassmannian to cocycles	39
4.2.2	Properties for functions on Ω	42
4.2.3	Other approaches?	43
Chapter 5	Constructing fast and slow spaces	44
5.1	Decomposing a cocycle	44
5.1.1	Searching for a decomposition	44
5.1.2	Quasicompactness: does an interesting Oseledets decomposition exist?	44
5.1.3	Finding a slow space: a splitting lemma	48
Chapter 6	Oseledets' theorem	54
6.1	History and overview	54
6.1.1	The fully invertible result	54
6.1.2	The non-invertible case: weaker premises and conclusions	55
6.1.3	Continuous cases and generalisations to infinite dimensions	56
6.1.4	A further refinement: semi-invertibility	56

6.1.5	A slow space measurability disaster	57
6.1.6	Statement of main result	58
6.2	Proof of main result	59
Bibliography		62

ACKNOWLEDGEMENTS

It's hard to imagine better luck than to have had Anthony as a supervisor, our meetings are invariably an enlightening and motivating experience, and I'm very grateful to him for it. Thanks are due to Chris Bose for his excellent advice and guidance, as well as Sean Chester for agreeing to act as an outside member of the supervisory committee, and Ian Morris for agreeing to act as an external examiner.

I feel truly lucky to have met some really wonderful people in the last few years. I am extremely grateful to innumerable many fellow students and staff in the dynamics group and the department for their enthusiasm, humour and support. Between my friends made in the department and through other channels, these last few years have been an enlightening and exciting time.

My family, for everything.

DEDICATION

To Juliet.

Chapter 1

Ergodic and linear background

1.1 Ergodic theory

1.1.1 Introduction

This work sets out to present some novel results in measurable linear dynamics and discount the entry cost to the field in the process. An infinite dimensional notion of linear dynamical systems is the main object of study, in a general setting applicable to several quantitative disciplines.

The main result is the semi-invertible multiplicative ergodic theorem of Oseledets on a separable Banach space. An overview of different versions of this theorem with their respective conditions and conclusions is given in section 6.1.1.

As a mathematical construction, while in many cases the study may be said to be of a random linear dynamical system, and we adopt some probabilistic notation and language throughout, there is no significant distinction between a deterministic and stochastic process: the underlying space on which dynamics occurs may have a more geometric or statistically motivated construction, but the results obtained are the same. For the most part there is no smoothness required of the dynamics, but a particular notion of "strong" measurability of a linear cocycle is sufficient and satisfyingly general.

This work has a soft mathematical analysis approach to estimate finding. When taking limits of long term logarithmic averages though, there is little issue, and work here isn't optimised, as there is already enough need for precision to extract the main results. If, for example, two quantities q_1 and q_2 that depend on the behaviour of a system from time 0 to time n can be related by $q_1 < cq_2 < Cq_1$ for c and C possibly huge but independent of n , then taking log and dividing by n to get an average growth rate, the difference between the resulting quantities is less than $\frac{\log C}{n} \rightarrow 0$ as $n \rightarrow \infty$.

1.1.2 Banach space: justifications and notation

In general we will focus on a dynamical system where the linear state space is a separable Banach space X . This is the source of some extra hassle - where there are sharp constructions in Hilbert space, in Banach space there may be penalties: while we can always pick the orthogonal projection of a closed subspace of Hilbert space, in a Banach space it is some work to show any bound. Generally, the smallest projection that can be guaranteed for a k dimensional space is \sqrt{k} as per theorem 2.1.18 of [2]. The dual space may fail to be separable: an influential precursor to the main result of this thesis is [13], which assumes that the dual is separable - in this case, it is easier to study the dual cocycle A^* and the result is slightly stronger: in this case finite co-dimensional spaces are better behaved, and so there is a stronger notion of measurability of the resulting decomposition.

This extra effort should be worthwhile, though. Plenty of objects of perennial interest to analysts can be dealt with if we drop the Hilbert space requirement - amongst the most obvious candidates, we may take the state space $X = C(K)$ or $L^1[0, 1]$, for example.

We write $\mathcal{B}(X)$ for the space of bounded linear operators on X . If X is finite dimensional then $\mathcal{B}(X)$ may be identified with the d by d real matrices $\mathbb{R}^{d \times d}$ where convenient by choosing any basis.

1.1.3 Measurable spaces

There are countless introductions to this subject, see [26, 17], with as many points of view as to how to introduce or think about it. Here we start with a set, and an appropriate collection of subsets to which we wish to assign measures. It is possible to start with an algebraic notion of measures, but these seem to still require analytic conditions, and measurable sets are worked with throughout this work. Therefore we are content to write down things in an algebraic way when convenient, without going so far as to become pointless - physically, it is still important and physically intuitive to be able to assume you can at least to some degree keep track of a quantity $f(\omega)$ at a point $\omega \in \Omega$. Being able to drop certain assumptions like commutivity of the algebra on which probabilities are measured leads one in the direction of the C^* or Von Neumann - algebraic programmes. For us though, noncommutivity will instead arise at the level of the linear cocycles we will introduce over the probability space.

Definition 1. *Let Ω be a set.*

Let $\mathcal{F} \subseteq 2^\Omega$ be a collection of subsets that contains \emptyset and is closed under countable

union and complementation:

$$U_i \in \mathcal{F} \text{ for all } i \in \mathbb{N} \implies \bigcup_{i=1}^{\infty} U_i \in \mathcal{F} \text{ and } U \in \mathcal{F} \implies U^c = \Omega \setminus U \in \mathcal{F}.$$

Then \mathcal{F} is called a sigma algebra and (Ω, \mathcal{F}) is a measurable space.

There is a natural representation $\chi : \mathcal{F} \rightarrow \{0, 1\}^\Omega$ given by

$$\chi(A)(\omega) = \chi_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise.} \end{cases}.$$

One can check that by this definition a sigma algebra must be closed under countable intersection and difference, and form a boolean ring - sigma algebra may actually be defined in a few different ways by exchanging definitions and lemmas. Closure under countable basic set operations then is their important distinguishing feature. The notion of sub-sigma algebra $\mathcal{F} \subseteq \mathcal{F}'$ is that of inclusion in 2^{2^Ω} . Under inclusion, the smallest and largest sigma algebras on Ω are $\{\emptyset, \Omega\}$ and 2^Ω , and when Ω is infinite the sigma algebras of interest will with an application of the axiom of choice strictly lie between these.

Definition 2. Given a collection $\mathcal{F}_0 \subseteq 2^\Omega$ write

$$\Sigma(\mathcal{F}_0) = \bigcap \{ \mathcal{F} \subseteq 2^\Omega : \mathcal{F}_0 \subseteq \mathcal{F} \text{ and } \mathcal{F} \text{ is a sigma algebra} \}.$$

Since 2^Ω is a sigma algebra the family over which the intersection is taken is nonempty, and it is straightforward to check that the intersections of families of sigma algebras over Ω are themselves sigma algebras. One may say that $\Sigma(\mathcal{F}_0)$ is the smallest collection of sets containing some starting family that itself forms a sigma algebra.

Definition 3. Let (Ω, τ) be a compact Hausdorff space, where $\tau \subseteq 2^\Omega$ denotes the collection of open sets. Then the Borel sigma algebra is that generated by the topology: $\mathcal{B}_\Omega = \Sigma(\tau)$.

Where there is no ambiguity as to which topology a space is endowed with, we refer to a function as measurable when it is measurable with respect to this sigma algebra. Otherwise we must specify which notion of measurability is being employed. Like continuous functions, the condition for a map to be measurable may be expressed in terms of preimages:

Definition 4. Let (Ω, \mathcal{F}) and (Ω', \mathcal{F}') be measurable spaces. A measurable map is a function $\sigma : \Omega \rightarrow \Omega'$ such that $U \in \mathcal{F}' \implies \sigma^{-1}U \in \mathcal{F}$.

Given any function f , we may define $f^{-1}|_{\mathcal{F}'} : \mathcal{F}' \rightarrow 2^\Omega$: measurability is equivalent to the condition that the range of $f^{-1}|_{\mathcal{F}'}$ is contained in \mathcal{F} : and $f^{-1}\mathcal{F}'$ is contained in (ie, a "sub-sigma algebra of") \mathcal{F} . We write $f \in \mathcal{M}(\Omega \rightarrow \Omega')$.

1.1.4 Measures

The next important ingredient is the measure itself. I use \sqcup to denote disjoint union.

Here, we work with traditional measures, which assign nonnegative numbers to the sizes of sets in a way that commutes with countable measurable partitioning:

Definition 5. Let $\nu : \mathcal{F} \rightarrow \mathbb{R}$. We call ν

- countably additive if for disjoint countable collections $U_i \in \mathcal{F}$,

$$\nu \left(\bigsqcup_i U_i \right) = \sum_i \nu(U_i),$$

- positive if $\nu(U) \geq 0$ for all $U \in \mathcal{F}$
- and finite if $\nu(\Omega) < \infty$.
- Call ν a measure if it is both additive and positive.
- Call a finite measure $\nu = \mathbb{P}$ a probability measure if $\mathbb{P}(\Omega) = 1$.

The prototypical example, and it turns out the only one of importance as a base measure space in this thesis if one discards the geometric properties, is the following:

Theorem 6. There is a unique measure \mathbb{P} on $[0, 1]$ such that for all $0 \leq a \leq b \leq 1$,

$$\mathbb{P}(a, b) = b - a.$$

Defined this way, the measure \mathbb{P} is uniform on the unit interval, but any other measurable space that is isomorphic to $[0, 1]$ may also be endowed with a measure called Lebesgue, defined precisely in definition 11 in the general sense relevant to our purposes.

One natural construction given these new concepts is

Definition 7. Let $f \in \mathcal{M}((\Omega, \mathcal{F}) \rightarrow (\Omega', \mathcal{F}'))$ and (Ω, ν) a measure space. Then the pushforward measure is given by $f_*\nu(U) = \nu(f^{-1}U)$ for $U \in \mathcal{F}'$.

Given a measurable $M \subseteq \Omega$ with $\mathbb{P}(M) > 0$, we may form the conditional measure written

$$\mathbb{P}(S|M) = \frac{\mathbb{P}(S \cap M)}{\mathbb{P}(M)}.$$

It is easy to check this is a probability space. This represents the proportion of a time when $\omega \in M$ that we also have $\omega \in S$.

1.1.5 Null sets

Once a measure is fixed, one may check that $\mathcal{F}_0 = \{U \in \mathcal{F} : \nu(U) \text{ or } \nu(U^c) = 0\}$ is a sub sigma-algebra of \mathcal{F} . On the level of the measure, throughout this thesis, when a measurable set $U \subseteq \Omega$ is discussed, we are mainly concerned with equality up to a null set. Thus we may write $U \subseteq U'$ if $\mathbb{P}(U \setminus U') = 0$ and $U = U'$ if that and the opposite condition are met. Similarly, for a measurable $f : (\Omega, \mathcal{F}) \rightarrow (\Omega', \mathcal{F}')$, we may form the functions that agree up to null sets.

The most relevant mode of convergence for a collection $(f_n)_{n=1}^\infty \in \mathcal{M}(\Omega \rightarrow \mathbb{R})^\mathbb{N}$ in this setting is that of pointwise almost everywhere convergence. If $f : \Omega \rightarrow \mathbb{R}$ satisfies $\mathbb{P}(f_n(\omega) \rightarrow f(\omega)) = 1$, then f is the measurable almost everywhere pointwise limit of f_n , and in this work write $f_n \rightarrow f$. As with sets, here any limit f may be replaced by an f' differing on a null set.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, call $\mathcal{P} \subseteq \mathcal{F}$ a partition if the members of \mathcal{P} are disjoint and have full measure, that is, $\sqcup \mathcal{P} = \Omega$. We will always further assume that a partition is countable, which is a less alarming scenario, allowing us to write, for example, $1 = \sum_{P \in \mathcal{P}} \mathbb{P}(P)$. Just as a measurable set M has an associated indicator χ_M , we can choose a labelling of a partition $\mathcal{P} \rightarrow \mathbb{N}_{\leq N}$ for $N = \#\mathcal{P}$.

We will discuss the precise class of measure space we work over, after we introduce ergodicity and some examples.

1.1.6 Measure preservation

For a huge variety of physical systems, some form of generalised volume conservation may be found:

- If a fluid can be approximated as incompressible and we follow the fluid initially contained in some domain \mathcal{D} , then at later times the region \mathcal{D}' occupied by the fluid that was in \mathcal{D} initially should have the same volume.
- Generalising the notion of a fluid to a phase space, in many circumstances volume in phase space may be taken constant.

- The classical harmonic oscillator has this property: for initial conditions $(x_0, p_0) \in \mathbb{R}^2$ and appropriately normalised dimensions, the solution is given by

$$\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = R_t \begin{pmatrix} x_0 \\ p_0 \end{pmatrix},$$

where R_θ denotes a rotation of the plane by angle θ , and preserves the volume of regions of \mathbb{R}^2 .

- The example of an oscillator may be generalised to Hamiltonian systems, where the volume conservation is referred to as Liouville’s theorem, recognised first by Gibbs[11].

In our case the notion of interest is of $\sigma \in \mathcal{M}((\Omega, \mathcal{F}, \nu) \rightarrow (\Omega, \mathcal{F}, \nu))$ leaving the measure invariant: if for any $U \in \mathcal{F}$ we have $\mathbb{P}(U) = \mathbb{P}(\sigma^{-1}U)$ then $(\Omega, \mathbb{P}, \sigma)$ is called a measure preserving system. This is equivalent to the algebraic equation $\sigma_*\mathbb{P} = \mathbb{P}$.

A measure preserving system may be written $(\Omega, \mathbb{P}, \sigma)$ and is the space over which the cocycles that will be introduced later are based.

Definition 8. *A morphism of measure preserving systems is a map $f \in \mathcal{M}((\Omega, \mathbb{P}, \sigma) \rightarrow (\Omega', \mathbb{P}', \sigma'))$ which is measure preserving and commutes with the action of the systems, that is to say $f_*\mathbb{P} = \mathbb{P}'$ and $f \circ \sigma = \sigma' \circ f$. If f^{-1} is also a morphism of measure preserving systems then call f an isomorphism.*

1.1.7 Ergodicity

Ergodicity of a measure preserving system $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ is a condition on the behaviour of the measurable sets with respect to σ .

Definition 9. *Given $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ a measure preserving system, call $U \in \mathcal{F}$ σ -invariant, or just invariant where clear, if $U = \sigma^{-1}U$.*

It is a common theme in mathematics to come across a notion of decomposability. A finite abelian group can be written as a direct sum of cyclic groups, a matrix into Jordan blocks, a case which will arise in the discussion of cocycles. Such a decomposition is into a collection of objects which are themselves in some sense indecomposable - in the case of abelian finite groups, these are copies of $\mathbb{Z}/p^k\mathbb{Z}$ for p prime and $k \in \mathbb{N}$, which have the property that they cannot be written as a direct sum of smaller groups. Meanwhile, Jordan blocks represent a splitting of \mathbb{R}^m that a matrix acts on into the distinct possible growth rates of vectors.

What, then, is an appropriate notion of indecomposability for measure preserving systems? Given $(\Omega, \mathbb{P}, \sigma)$ a measure preserving system, if there is an invariant measurable set Ω' with $0 < \mathbb{P}(\Omega') < 1$, then so is Ω'^c , and we could instead study the measure preserving systems generated by restricting σ to each of these sets and renormalising \mathbb{P} as appropriate. Ergodicity is the condition that this scenario cannot occur: an invariant set Ω' cannot have $\mathbb{P}(\Omega') \in (0, 1)$:

Definition 10. *Call a measure preserving $\sigma \in \mathcal{M}((\Omega, \mathbb{P}) \rightarrow (\Omega, \mathbb{P}))$ ergodic if for any $U \in \mathcal{F}$ that is invariant, either U has full measure or is null:*

$$U = \sigma^{-1}U \implies \mathbb{P}(U) \in \{0, 1\}.$$

Given ergodic systems $(\Omega, \mathbb{P}, \sigma)$ and $(\Omega', \mathbb{P}', \sigma')$ one can readily construct a measure preserving but not ergodic space $\Omega'' = \Omega \sqcup \Omega', \sigma'' = \sigma \sqcup \sigma'$ and $\mathbb{P}''(U) = \frac{1}{2}(\mathbb{P}(U \cap \Omega) + \mathbb{P}'(U \cap \Omega'))$. A measure formed from a nontrivial convex combination of ergodic measures can never be ergodic, in this case Ω and Ω' are invariant measure $\frac{1}{2}$ sets. Geometrically, the invariant measures form a convex set whose extreme points $\mathcal{E}(\Omega, \sigma)$ are the ergodic measures. Where there are infinitely many ergodic measures, by the Choquet decomposition[5] we may represent a measure preserving transformation as an integral over the ergodic measures:

$$\mathbb{P}(U) = \int_{\mathcal{E}(\Omega, \sigma)} \nu(U) d\tau(\nu),$$

1.1.8 Standard and Lebesgue probability spaces

It doesn't seem accidental that in the last example, a huge class of possibilities may be modelled by simply generating bits: though the above examples of Ω were presented differently, in fact we may view the space to be the same object on the level of measure. To illustrate this, the notions of standard probability space and Lebesgue probability space are discussed for the reader's convenience. As mentioned, though $\{0, 1\}^{\mathbb{Z}}$ may not seem to geometrically or topologically match $[0, 1]$, both spaces have this property.

In Rudolph's introduction to this subject, the notion of a generating tree $\mathcal{P} = (\mathcal{P}_i)_{i=1}^{\infty}$ is a sequence of partitions that is nested and separates the points in the underlying space: taking $\mathcal{P}_0 = \{\Omega\}$,

$$U \in \mathcal{P}_i \text{ for } i > 0 \implies \exists U \subseteq V \in \mathcal{P}_{i-1} \text{ and } \omega \neq \omega' \in \Omega \implies \exists i : \omega \in U \in \mathcal{P}_i, \omega' \notin U.$$

is used to make precise these notions: that the closure under null sets of \mathbb{P} of the sigma algebra is always unique is then a theorem.

Definition 11. *The largest sigma algebra which is consistent with this notion of null sets is formed by closing the Borel sigma algebra under null sets. Precisely, the Lebesgue sigma algebra is the collection*

$$\mathcal{L} = \{U \subseteq U' \in \mathcal{B}_{[0,1]} : U' \setminus U \subseteq V \in \mathcal{F}_0\}.$$

Though on the level of measure any two Lebesgue spaces look alike, they may be associated with, for example, quite different topological spaces - any separable complete metric or Polish space is associated with an isomorphic Lebesgue sigma algebra.

To illustrate the generality of a standard probability space, we give a definition of standardness involving a general topological space:

Definition 12. • [6] *Say that $(\Omega, \mathcal{F}, \mathbb{P})$ is standard if there is a metrisable topology*

τ on Ω with $\mathcal{F} = \Sigma(\tau)$ and for every $\epsilon > 0$ there is a compact set K with $\mathbb{P}(K) > 1 - \epsilon$.

- Say that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless if $\mathbb{P}(\{\omega\}) = 0$ for all $\omega \in \Omega$.

In spite of this seemingly general condition, there is a simple classification. Atomlessness guarantees that Ω is uncountable, since otherwise we would have $0 = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 1$, but in fact the space is even more familiar:

Definition 13. *Call a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a Lebesgue space if $(\Omega, \mathcal{F}, \mathbb{P})$ is standard and atomless.*

This notion of Lebesgue space is more liberal in appearance than the common characterisation as a translation-invariant measure on \mathbb{T} , for example, but if we pay attention only to equivalence up to isomorphism of measure spaces, then the definitions coincide.

In this setting, though one may have imagined $\{0, 1\}^{\mathbb{Z}}$ to be somehow distinct from $[0, 1]^{\mathbb{Z}}$, as they are in topology where a notion of connectedness distinguishes the two spaces, each of the (Ω, \mathbb{P}) described are Lebesgue:

Theorem 14 (Uniqueness of Lebesgue space). *The following are equivalent:*

- $(\Omega, \mathcal{F}, \mathbb{P})$ is isomorphic to the Lebesgue measure on the unit interval together with the Lebesgue completion of the associated Borel sigma algebra \mathcal{L} defined above.
- There exists a tree \mathcal{P} of partitions such that \mathcal{F} is the closure under \mathbb{P} of the sigma algebra generated by \mathcal{P} and $(C_1, C_2, \dots) \in \mathcal{P} \implies \mathbb{P}(\cap_i C_i) = 0$.
- $(\Omega, \mathcal{F}, \mathbb{P})$ is standard and atomless.

Proof. These equivalences are established in the second chapter of Rudolph's book [26]. □

Since in some texts Lebesgue probability spaces are allowed to contain atoms, so we may repeat the non-atomic condition for emphasis while stating some propositions. Under the ergodic assumption, since the orbit of a point is invariant, any atom ω must have $\mathbb{P}(\cup_i \sigma^i \omega) \in \{0, 1\}$. As such, an ergodic map on a finite set is periodic and the main results in this thesis become much easier, so throughout we assume that the measure is atomless. Since we are already working with a standard probability space, for just about all Ω considered in this thesis the reader will lose little generality in imagining Ω to be their favourite probability space, such as the interval or sequence spaces above.

Thus it would seem nonpathological measure spaces aren't themselves objects to which there is a rich and complicated classification theory, but instead a single very

general object with many useful properties. From a constructive point of view, these spaces might be seen to describe the kinds of spaces where statisticians, physicists and computer scientists work, where we may imagine points being described by discrete or real values coordinates, and have some well behaved notion of size or probability.

What is absolutely not unique is the choice of transformation σ . Even when imposing invertibility, there is a huge variety of possibilities.

1.1.9 Discretising continuous time systems

In proving the first version of the theorem of the main result of this thesis, Oseledets started with a 1 parameter group of automorphisms of a Lebesgue probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\theta_t\}_{t \in \mathbb{G}})$ for $\mathbb{G} = \mathbb{R}$ or \mathbb{Z} . These two cases correspond to a continuous and discrete time dynamical system respectively.

Given a topological group $(\mathbb{G}, +)$ a measure-preserving \mathbb{G} -action on $(\Omega, \mathcal{F}, \mathbb{P})$ is a measurable map

$$\theta : \Omega \times \mathbb{G} \rightarrow \Omega, \theta(\omega, t) = \theta_t \omega$$

such that $\theta_{s+t} \omega = \theta_s(\theta_t \omega)$ and $\mathbb{P}(\theta_t \Omega') = \mathbb{P}(\Omega')$ for all $\Omega' \in \mathcal{F}$.

The discrete and continuous time cases divide as follows:

- In the discrete time case, $\mathbb{G} = \mathbb{Z}$, and just as \mathbb{Z} is the free group on a single generator, θ is determined by the invertible measure preserving system $\sigma = \theta_1$.
- In the continuous time case, an \mathbb{R} -action has no natural association with a single map, but one can still study $\sigma = \theta_\delta$.

It is a common approach to discretise a continuous time problem, for example, a differential equation: for example, given an $a \in L^1(\Omega \rightarrow \mathbb{R}^{d \times d})$, consider the initial value problem of finding $z : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$ such that

$$z'(t) = a(\theta_t \omega) z(t) \text{ for } t > 0, x(0) = u \in \mathbb{R}^n.$$

One simple case would be where $\Omega = \mathbb{T}$ and θ represents constant flow around a circle $\theta_t \omega = \omega + vt$, which is the central object of study in the Floquet theory.

According to Carathéodory's existence and uniqueness theorem this problem has a unique operator solution $U_{t\omega} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ for $t > 0$ such that setting $z(t) = U_{t\omega} u$ we have that z solves the initial value problem above. Setting $A_\omega = U_{\delta\omega}$ for a time step parameter δ , in this discretised picture we are really studying the composition of sequences of these A s, since

$$U_{(n\delta)\omega} = A_{\theta_{(n-1)\delta}\omega} \circ \dots \circ A_{\theta_\delta\omega} \circ A_\omega.$$

For a fixed starting state $x_0 = u \in \mathbb{R}^d$, the state from one discrete timestep to the next is correspondingly given by $x_{n+1} = A_{\sigma^n \omega} x_n$. This example is actually a construction "over" the system $(\Omega, \mathbb{P}, \theta)$, and will be discussed in more detail when linear cocycles, of which U and A are continuous and discrete time examples respectively, are introduced in chapter 4.

The system $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ is a classical object of study and the background upon which dynamical results in this thesis will be built. Many results in this setting may have analogous results where invertibility or ergodicity or invertibility may be dropped - where appropriate some discussion of the differences in conclusions will be had, and generalisations alluded to, but for most work having to hand an invertible ergodic system on Lebesgue space is extremely convenient and general.

1.2 Stochastic or deterministic dynamics?

Common examples of the ergodic system may be rotations or shifts, or a particular diffeomorphism over a manifold.

Each of these have quite different interpretations based on the academic field of study: on one hand, statisticians may often work with sequences of random variables, which in many cases are assumed to be independent and identically distributed (iid). On the other, many ergodic systems arise from a perfectly "deterministic" setting, for example as some diffeomorphism or homeomorphism with an associated ergodic measure.

Beyond these views, the pseudorandom number generators ubiquitously relied upon for *in silico* research, will typically be dependent upon a finite amount of history plus a seed, and quite explicitly deterministic, though it is standard to assume they can be used as a source of iid statistics.

We will remain somewhat agnostic as to whether our underlying system is either deterministic or random: though hidden variable theories of physics seem unpopular or inconsistent with certain additional assumptions when developing a theory, it remains unclear to the author how anyone could ever prove determinism or its converse, or even if there should be any immediate real world consequences.

Results in this area promise to hold in more general situations than iid statistics. This common assumption throughout a huge variety of statistical methods is very often not valid, but researchers apply these methods with successes nonetheless. It should be seen as worthwhile to understand with greater precision when these kinds of assumptions can be weakened, and results on general ergodic systems can achieve this.

For those considering the analysis of data, some form of this ergodicity is a weaker, but physical assumption for processes.

Here I outline some common iid, or possibly non-iid, cases.

1.2.1 Generalising from shifts and iid statistics: the ergodic assumption

Given a Lebesgue probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$, one may study the infinite product $\Omega = \Omega_0^{\mathbb{Z}}$ together with the product measure defined by the cylinder set condition $\mathbb{P}(\{\omega : \omega_i \in M\}) = \mathbb{P}_0(M)$, along with the shift map $\sigma(\omega)_i = \omega_{i+1}$. This is the space of bi-infinite sequences of independent identically distributed random variables. By this process, results regarding iidrvs may be recovered as special cases of the problems considered in this thesis: where the statistician talks of iid variables F_0, F_1, F_2, \dots sampled according to $f_*\mathbb{P}_0$ we may instead study the sequence

$$\dots, f(\omega_{-1}) = f(\sigma^{-1}\omega), f(\omega_0), f(\omega_1) = f(\sigma(\omega)_0), f(\omega_2) = f((\sigma^2\omega)_0), \dots$$

For example, the following form important statistically motivated examples:

- For $\Omega_0 = \{0, 1\}, \mathbb{P}_0$ the Bernoulli measure $\mathbb{P}_0(1) = p$ and f the identity, F_i represents a sequence of random coin tosses.
- For $\Omega_0 = [0, 1], \mathbb{P}_0$ the Lebesgue measure and

$$f(x) = \begin{cases} 1 & \text{if } x < p \\ 0 & \text{otherwise,} \end{cases},$$

F_i also represents a sequence of coin tosses - but this construction may allow one to derive results about dynamical behaviour of different parameters p at once.

- As mentioned, a computer scientist wishing to approximate a random variable computationally, Ω might be modelled by generating random bit strings based on a finite random seed, and treating the result as being a reasonable iid string of bits that may be coerced to approximate whichever random variable is desired. Excepting a computer that is connected to some external input, however, since the state of a computer is finite, the space of possible generated sequences is finite and the measure may not align precisely with the idealised situations considered here.

1.2.2 Examples from topology and geometry

Aside from examples motivated by statistics and computation, we might think of much studied cases arising from an attempt to describe dynamics on a space that had some smooth structure, for example:

- An irrational rotation: $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ given by rotating by an irrational angle, a perennial example

- As an example of a pseudorandom number generator, one may consider the measure on $\{0, 1\}^{\mathbb{Z}}$ induced by the image of

$$\varphi : ([0, 1], \mathcal{L}, \sigma : \omega \mapsto \omega + \alpha \bmod 1) \rightarrow \{0, 1\}, \quad \varphi = \chi_{[0, \frac{1}{2}]}$$

Write $X_i(\omega) = \varphi(\sigma^i \omega)$. One should understand that this could be extremely non-optimal. If α is very close to $\frac{1}{2}$, for example, then typically a sequence of sampled bits might look like alternating 1s and 0s for the most part. However, since this entirely deterministic source of bits is generated by an ergodic system, it will at least still have some useful statistical properties, for example, we have $\mathbb{P}(X_i = 1) = \frac{1}{2}$, and as a consequence $\mathbb{P}(\frac{1}{n} \sum_{i=0}^{n-1} X_i \rightarrow \frac{1}{2}) = 1$. This example of the strong law of large numbers is an intuitively obvious statement for physicists but needs Birkhoff's theorem, introduced later. It is an example of why making some form of ergodic assumption may be a natural one to make in certain physical or experimental scenarios.

- Generalising the case of an irrational rotation, $F : M \rightarrow M$ a volume preserving diffeomorphism with respect to \mathbb{P} .
- If Ω is a compact metric space and σ a homeomorphism, we know that by the Krylov-Bogolyubov theorem that measure preserving and ergodic measures exist. For example, if $\sigma : \mathbb{T}^2 \rightarrow \mathbb{T}^2, (x, y) \mapsto (x + \alpha, y) \bmod 1$ then an ergodic measure is just Lebesgue measure conditional on a fixed y level set.
- As a non-example by the lack of invertibility, we may model 1-dimensional bread mixing by a collection of ingredients distributed along $[0, 1]$ at any moment in time. The baker's transformation $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ is to divide $[0, 1]$ into two halves, roll out each half to make it length 1, then put each infinitely thin half beside the other to create a distribution of ingredients on $[0, 1]$ at the next moment. This simplified baker's transformation is equivalent to multiplication by 2 mod 1 on the unit interval which may be thought of as a continuous map on the circle \mathbb{T} . It's easy to check that measure is preserved in this case, as the preimage of any set is two contracted by a factor of two and translated copies of that set, for example $\mathbb{P}(\sigma^{-1}[a, b]) = \mathbb{P}([\frac{1}{2}a, \frac{1}{2}b]) + \mathbb{P}([\frac{1}{2} + \frac{1}{2}a, \frac{1}{2} + \frac{1}{2}b]) = \frac{2}{2}\mathbb{P}([a, b])$. The map is ergodic: we can see this by proving that Birkhoff's theorem holds for indicators of subintervals of the unit interval of the form $[0.a_1a_2 \cdots a_k 0, 0.a_1a_2 \cdots a_k 1]$, which is equivalent to the statement that almost all numbers are normal in base 2. The problem here is that the map is 2 to 1, so that only a weaker non-invertible multiplicative ergodic theorem holds - this result that holds when invertibility of the base space is dropped is discussed in chapter 6.

As an indication of how one may recover theorems over such a transformation, we

observe that writing down points on the interval as the space of forward-infinite binary strings (up to the issue of nonunique decimal representation, which is on a set of zero measure), we may consider the measure preserving factor onto shift space defined above:

$$\varphi : \{0, 1\}^{\mathbb{Z}} \rightarrow [0, 1], (\dots, b_{-1}, b_0, b_1, b_2, \dots) \mapsto 0.b_1b_2b_3 \dots .$$

As a means of applying the results of this thesis to the case of a non-invertible ergodic system, we construct a φ as above, then instead of studying a function $A : (\Omega, \mathcal{F}, \sigma) \rightarrow Y$ we can form a map from an invertible system $A' : (\Omega', \mathcal{F}', \sigma') \rightarrow Y$ where A' is given by $A'(\omega) = A(\varphi(\omega))$.

Chapter 2

Rokhlin tower lemmas and Kingman sub-additive ergodic theorems

2.1 Towers

2.1.1 Overview

Though there are many related concepts and directions generalising the setup, given an invertible ergodic system $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$, tower type constructions divide the underlying space Ω into a collection of sets that may be thought of as "floors", and σ moves one floor to the next.

This is really useful for certain constructions: in the work of [4] they are used as a good starting point from which to construct perturbations showing the instability of nonuniformly hyperbolic finite dimensional systems, and a variation on the concept will be used to prove the main result in this thesis.

2.1.2 Representation by return time

Given a set of positive measure $\mathbb{P}(\Omega_0) > 0$, we may wish to investigate the times at which a randomly sampled point $\omega \in \Omega$ lands in Ω_0 , that is to say: for which i is $\sigma^i \omega \in \Omega_0$? It is useful to define and think about the return time

$$r : \Omega \rightarrow \mathbb{N} \cup \{\infty\}, r(\omega) = \inf\{i > 0 : \sigma^i \omega \in \Omega_0\}.$$

Since σ is ergodic and $\mathbb{P}(\Omega_0) > 0$, it follows that r is finite almost everywhere, so we may decompose Ω by return time:

$$\Omega = \bigsqcup_{t=1}^{\infty} r^{-1}\{t\}.$$

We have obtained a suggestive first example of a partition forming a "tower", where each element of the partition forms a level, and $\sigma^{\pm 1}$ shift us up or down a level, except when we're sent all the way to the bottom or top again, depending on convention. In this case if we imagine Ω_0 to be the base and pick the convention that σ raises the floor by one, actually the structure consists of a series of levels beneath Ω_0 whose measure decreases the further down we go. Finally, upon reaching level 0, another application of σ may send a point ω to any selection of levels.

We may sometimes, for example, choose an Ω_0 on which there is some particular statistical behaviour. From this decomposition of a space and the pigeonhole principle, we can notice that the height of a return time tower associated with $M \subseteq \Omega$ with $\mathbb{P}(M) \in (0, \frac{1}{N})$ must have height at least N . We also obtain the fact that we can always make a tower which has fixed size levels for as long as we want:

Lemma 15. *Let $(\Omega, \mathbb{P}, \sigma)$ be an invertible ergodic system on a non-atomic Lebesgue probability space and let $T > 0$. Then there is a positive measure set Ω_r for which the $\sigma^i \Omega_r$ for $0 \leq i < T$ are disjoint.*

Proof. Choose a measurable Ω_0 with $0 < \mathbb{P}(\Omega_0) < \frac{1}{T}$. Then with r the return time as above, since $\mathbb{P}(r^{-1}\{t\}) \leq \mathbb{P}(\Omega_0) < \frac{1}{T}$ for all t and $\mathbb{P}(\Omega) = 1$, the set $\Omega_r = r^{-1}(T)$ has positive measure and is disjoint from $\sigma^i \Omega_r \subseteq r^{-1}(T - i)$ for $0 \leq i < T$. \square

Closely related constructions include Kakutani, Rokhlin or Alpern towers, which are prevalent throughout dynamically oriented mathematics and highlighted in [7, 16].

2.1.3 Partitions and processes

Let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ be an invertible ergodic system and let a collection of measurable sets $\mathcal{P} \subseteq \mathcal{F}$ be a partition. There is an associated inclusion map $\Omega \rightarrow \mathcal{P}$, and with the addition of dynamics, there is an induced measure space by the map that simply records where a given base point $\omega \in \Omega$ is at each point in time:

$$\varphi : (\Omega, \sigma) \rightarrow (\mathcal{P}^{\mathbb{Z}}, \tau)$$

gives a measure $\mathcal{P}_{\mathcal{P}} = \varphi_* \mathbb{P}$. If \mathcal{P} has N elements then we have a mapping into the space of bi-infinite strings of N -digits.

One may construct the matrix $(A_{PQ})_{P, Q \in \mathcal{P}}$ describing transitions in one step, given by

$$A_{PQ} \mathbb{P}(P) = \mathbb{P}(\sigma P \cap Q).$$

If $\sigma^t \omega$ encodes the point of a particle in a system at time t , A_{PQ} is the probability that $\sigma^{t+1} \omega \in Q$ given that $\sigma^t \omega \in P$.

The resulting process is Markov when the probability of a particular series of visits to different elements of the partition can be calculated by taking products of the entries

of this matrix:

$$\mathbb{P}(\omega = P_0, \dots, \sigma^n \omega = P_n) = \prod_{i=1}^n \mathbb{P}(\sigma^i \omega \in P_i | \sigma^{i-1} \omega \in P_{i-1}) = \prod_{i=1}^n \frac{\mathbb{P}(\sigma^{-1} P_{i-1} \cap P_i)}{\mathbb{P}(P_{i-1})},$$

which allows transition probabilities to be calculated by exponentiating A_{PQ} . On the other hand, an irreducible Markov chain may be associated with an ergodic system in a natural manner:

Definition 16. *Given a matrix associated with an irreducible Markov chain,*

$$A = (A_{ij})_{i,j=1}^N \text{ satisfies } \forall i, j : \sum_k A_{kj} = 1 \text{ and } \exists t > 0 : (A^t)_{ij} > 0,$$

we may form the ergodic system on $\Omega = \{1, \dots, N\}^{\mathbb{Z}}$ of the shift and the probability measure satisfying the cylinder set condition

$$\mathbb{P}(\omega_{t+1} = i | \omega_t = j) = \mathbb{P}(\omega_{t+1} = i | \omega_t = j_t, \omega_{t-1} = j_{t-1}, \dots, \omega_{t-k} = j_{t-k}) = A_{ij} \mathbb{P}(\omega_t = j)$$

for all t .

A very simple example would be the case when $A_{ij} = \frac{1}{N}$, so that the state at one moment is independent of state at the next and this measure is just uniform on infinite strings with N symbols. There is little you need other than A then to study the properties of A^n for large n . The Perron-Frobenius theorem guarantees, for example, that there is a unique distribution π on $\{1, \dots, N\}$ with the property that $\mathbb{P}(\omega_t = i) = \pi_i$ and $A\pi = \pi$.

2.1.4 A Rokhlin lemma

The kind of construction originally discussed by Kakutani and Rokhlin has an elegant theory. The most recent and miraculous statements on this subject can be found in the work of Kalikow [16]. While it contains no Markovian free lunch - though there exists no result to reduce an arbitrary ergodic system to a Markov chain, it for example tells us the remarkable fact that any ergodic system on Lebesgue space may be taken to have Ω isomorphic to the subset of $\mathbb{N}^{\mathbb{Z}}$ such that there are functions $f_i : \mathbb{N}^{i-1} \rightarrow \mathbb{N}$ such that $\omega_{t-1} = i \implies \omega_t = f_i(\omega_{t-i}, \dots, \omega_{t-2})$ together with the shift.

Here I present a very easily obtained result, without going so far as the powerful reduction to symbolic dynamics as Kalikow gives us. It uses a very simple fact about numbers:

Lemma 17. *Let $N \geq n^2$. Then there exist $a, b \in \mathbb{N}_0$ with $N = an + b(n+1)$ and $0 \leq b < n$.*

Proof. We may write $N = sn + t$ with $0 \leq t < n$ and $s \geq n > t$, which can be rewritten $N = (s - t)n + t(n + 1)$ as required. \square

Lemma 18. *Let $n \geq 1$, $(\Omega, \sigma, \mathbb{P})$ an invertible ergodic system on a Lebesgue probability space and $\epsilon > 0$. Then there are measurable sets Ω_g and Ω_ϵ such that*

$$\Omega = \left(\bigsqcup_{i=0}^{n-1} \sigma^i \Omega_g \right) \sqcup \Omega_\epsilon, \quad \sigma \Omega_\epsilon \subseteq \Omega_g \quad \text{and} \quad \mathbb{P}(\Omega_\epsilon) < \epsilon.$$

If $\mathbb{P}(M) > 0$ then we can always guarantee that $\mathbb{P}(M \cap \Omega_g) > 0$.

Proof. Choose $N_0 > \max\{\frac{1}{\epsilon}, n^2\}$. Choose a positive measure set S such that the collection $\{\sigma^i S : 0 \leq i < N_0\}$ is disjoint. Index S by return time to S , $r(\omega) = \inf\{i > 0 : \sigma^i \omega \in S\} \geq N_0$. Denote the level set $R_N = r^{-1}N$ for $N \geq N_0$, write the union of the first l iterates as $R_{lN} = \bigsqcup_{i=0}^{l-1} \sigma^i R_N$ and write $N = s_N n + k_N(n + 1)$ with $0 \leq k_N < n$. We may then divide the iterations of R_N up to time N by

$$\bigsqcup_{i=0}^{N-1} \sigma^i R_N = \left(\bigsqcup_{a=0}^{s_N-1} \sigma^{an} R_{nN} \right) \sqcup \left(\bigsqcup_{b=0}^{k_N-1} \sigma^{s_N n + b(n+1)} R_{(n+1)N} \right).$$

Then

$$\Omega = \bigsqcup_{N \geq N_0} \bigsqcup_{i=0}^{N-1} \sigma^i R_N = \left(\bigsqcup_{N \geq N_0} \bigsqcup_{a=0}^{s_N-1} \sigma^{an} R_{nN} \right) \sqcup \left(\bigsqcup_{N \geq N_0} \bigsqcup_{b=0}^{k_N-1} \sigma^{s_N n + b(n+1)} R_{(n+1)N} \right),$$

so setting

$$\Omega_m = \left(\bigsqcup_{N \geq N_0} \bigsqcup_{a=0}^{a_N-1} \sigma^{am} R_N \right) \sqcup \left(\bigsqcup_{N \geq N_0} \bigsqcup_{b=0}^{k_N-1} \sigma^{s_N n + b(n+1)} R_N \right), \quad \Omega_\epsilon = \Omega \setminus \bigsqcup_{i=0}^{n-1} \sigma^i \Omega_m,$$

one may observe that Ω_ϵ is simply the ends of $(n + 1)$ -blocks,

$$\Omega_\epsilon = \bigsqcup_{N \geq N_0} \bigsqcup_{b=0}^{k_N-1} \sigma^{s_N n + b(n+1) + n} R_N$$

of measure at most $\mathbb{P}(S) < \epsilon$ and $\sigma \Omega_\epsilon \subseteq \Omega_g$. \square

When invoking this form of Rokhlin's lemma for a fixed n , it may be that ϵ doesn't need to be stipulated - Ω_g is necessarily small, with $\mathbb{P}(\Omega_\epsilon) < \mathbb{P}(\Omega_g) < \frac{1}{n}$.

We may use this to demonstrate that an appropriately chosen partition can never correspond to a Markov chain as described above: by choosing a tower of height greater than $\frac{N}{\epsilon}$ and error set size less than ϵ^2 one may pick a new tower of height N and error set size ϵ such that, for example, visiting the error set at the end of one tower ascent

guarantees no further visits to the error set for 100 tower ascents, which is not behaviour that occurs with a Markov chain. With this being said, while ascending the tower one needs no knowledge of the history to know which level the next step will be contained in, a very convenient decomposition for the dynamicist, and by Birkhoff's theorem 24 there is a "stationary" distribution that describes the propotion of the time $\mathbb{P}(P)$ that $\sigma^i\omega \in P$.

As a further example of an application of this result that serves to build intuition about the nature of ergodic systems, we can use this result to show that any ergodic system can be modified on an arbitrarily small set to obtain a periodic map:

Lemma 19. *Let σ be an invertible ergodic system. Then for any $\epsilon > 0$ there is an σ' which agrees with σ on Ω' with $\mathbb{P}(\Omega') > 1 - \epsilon$ and is periodic.*

Proof. Choose a Rokhlin tower of base Ω_g with height $N > \frac{2}{\epsilon}$ and error set $\mathbb{P}(\Omega_b) < \frac{1}{2}\epsilon$. Let

$$\sigma'(\omega) = \begin{cases} \sigma(\omega) & \text{for } \omega \in \sigma^i\Omega_g \text{ with } 0 \leq i < N \\ \omega & \text{for } \omega \in \Omega_b \text{ and} \\ \sigma^{-N}\omega & \text{for } \omega \in \sigma^N\Omega_g. \end{cases}$$

σ' is periodic, and equal to σ except on $\Omega_b \sqcup \sigma^N\Omega_g$ which has measure $< \epsilon$, so the conclusion is clear. \square

While this tower result is already useful and leads us to interesting conclusions about ergodic systems, it is just one simple example of a family of possible results.

2.1.5 The backward Vitali lemma

The backward Vitali lemma is another result providing a tower-like decomposition of the underlying dynamical space. We make use of this result to prove the novel Kingman theorem in the next section. In particular, this will allow us to decompose the underlying space σ acts on into orbits over an interval of time in which some particularbehaviour is observed.

The following is a slight simplification of the result that can be found in [26]: there are two families i_k and $j_k : \Omega \rightarrow \mathbb{N}_0$ but all we need is the case where $i = j$.

Lemma 20 (Backward Vitali). *[26] Let $(\Omega, \mathbb{P}, \sigma)$ be an invertible ergodic system on a Lebesgue probability space and $\epsilon > 0$. Suppose that there is a sequence of integer valued functions $j_k \in \mathcal{M}(\Omega \rightarrow \mathbb{N}_0)$ such that $j_k(\omega) \rightarrow \infty$ pointwise almost everywhere as $k \rightarrow \infty$. Then there is a measurable $A \subseteq \Omega$ and a measurable $j \in \mathcal{M}(A \rightarrow \mathbb{N}_0)$ such that for every $\omega \in A$ there is a k with $j_\omega = j_k(\omega)$ and writing $I_\omega = \{\sigma^i\omega : i \in \{-j_\omega, \dots, j_\omega\}\}$, the I_ω are disjoint and*

$$\mathbb{P}\left(\bigcup_{\omega \in A} I_\omega\right) > 1 - \epsilon.$$

2.2 A balanced subadditive ergodic theorem

2.2.1 Subadditive families

In this section, we think of subintervals of \mathbb{Z} as periods of discrete time over which particular dynamical phenomena may occur.

Definition 21. Let $\mathcal{F} = \{f_n\}_{n \in \mathbb{N}_0} \subseteq \mathcal{M}(\Omega \rightarrow \mathbb{R})$ satisfy

$$f_{m+n}(\omega) \leq f_m(\sigma^n \omega) + f_n(\omega) \quad \text{almost everywhere.}$$

Then \mathcal{F} is referred to as a subadditive family of measurable functions.

The following view is useful:

Definition 22. A subadditive family $\{f_n\}_{n \in \mathbb{N}_0}$ generates a stationary subadditive process $\{f_{a \rightarrow b} : a < b, a, b \in \mathbb{Z}\}$ and vice versa via the relation

$$f_{a \rightarrow b} := f_{b-a} \circ \sigma^a.$$

A subadditive process is a collection $\{f_{a \rightarrow b}\}_{a < b \in \mathbb{Z}} \subseteq \mathcal{M}(\Omega \rightarrow \mathbb{R})$ such that for all $a < c < b \in \mathbb{Z}$,

$$f_{a \rightarrow b} \leq f_{a \rightarrow c} + f_{c \rightarrow b} \quad \text{almost everywhere,}$$

and the stationarity condition is

$$f_{a+l \rightarrow b+l} = f_{a \rightarrow b} \circ \sigma^l.$$

As such both notations may be interchanged as appropriate. A similar formalism is outlined in [19].

Given a $f \in L^1\Omega$, the obvious first example of a subadditive family is that of the partial sums of the translations of f : if we set $g_n = \sum_{i=0}^{n-1} \varphi \circ \sigma^i$, then it is easily verified that the subadditivity condition holds, in fact this is exactly an additive family: $g_{m+n} = g_m \circ \sigma^n + g_n$.

2.2.2 Kingman theorems

The Kingman theorem may be seen as a generalisation of Birkhoff's theorem to the case of subadditive families of functions.

Theorem 23 (Kingman[18]). Let $(\Omega, \sigma, \mathbb{P})$ be an ergodic system and let $\{f_n\}_{n \in \mathbb{N}} \subset L^1\Omega$ be subadditive. Then there is a constant $C \in (-\infty, \infty)$ such that

$$\frac{1}{n} f_n(\omega) \rightarrow C = \lim_{n \rightarrow \infty} \frac{1}{n} \int f_n$$

pointwise almost everywhere.

Considering the additive example of partial sums, one obtains the Birkhoff theorem from this result:

Theorem 24 (Birkhoff). *Let $(\Omega, \sigma, \mathbb{P})$ be an ergodic system and let $f \in L^1\Omega$. Then there is a constant $C \in (-\infty, \infty)$ such that*

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ \sigma^i \rightarrow C = \int f$$

pointwise almost everywhere.

Proof. Set g_n to be the partial sums as described above, which is a subadditive family, so that

$$\frac{1}{n} g_n = \frac{1}{n} \sum_{i=0}^{n-1} f \circ \sigma^i \rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \int \sum_{i=0}^{n-1} f \circ \sigma^i = \lim_{n \rightarrow \infty} \int f = \int f$$

by σ -invariance of integration. □

With this stronger condition of being partial sums, all results in this section are similarly rendered easy. This more general case will be crucial in the proof of the multiplicative ergodic theorem.

In the case where σ is invertible, one easily obtains the following useful corollary:

Corollary 25. *Let $(\Omega, \sigma, \mathbb{P})$ be an invertible ergodic system on a Lebesgue probability space and let $\{f_n\}_{n \in \mathbb{N}} \subset L^1\Omega$ be subadditive with $f_1^+ \in L^1(\Omega)$. Then*

$$\lim_{n \rightarrow \infty} \frac{1}{n} f_n(\sigma^{-n}\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} f_n(\omega)$$

pointwise almost everywhere.

Proof. One may view this as following from the invertibility of σ . We can then apply Kingman's theorem to the family $g_n = f_n \circ \sigma^{-n}$. Then g_n is subadditive with respect to σ^{-1} :

$$\begin{aligned} g_{m+n}(\omega) &= f_{m+n}(\sigma^{-(m+n)}\omega) \leq f_m(\sigma^{-(m+n)}\omega) + f_n(\sigma^{m-(m+n)}\omega) \\ &= g_m(\sigma^{-n}\omega) + g_n(\omega), \end{aligned}$$

so that certainly there exists some L such that $\frac{1}{n} f_n \circ \sigma^{-n} \rightarrow L$ pointwise almost everywhere. It remains to check that this limit and $\frac{1}{n} f_n(\omega) \rightarrow L'$ coincide. Let $\epsilon > 0$. Since pointwise convergence almost everywhere implies convergence in measure, there exists an N such that

$$\mathbb{P}(\omega : \frac{1}{n} f_n(\omega) \in (L' - \epsilon, L' + \epsilon)) > \frac{1}{2},$$

and

$$\mathbb{P}(\omega : \frac{1}{n}g_n(\omega) = \frac{1}{n}f_n(\sigma^{-n}\omega) \in (L - \epsilon, L + \epsilon)) > \frac{1}{2},$$

so that the set

$$\{\omega : \frac{1}{n}f_n(\sigma^n\omega) \in (L - \epsilon, L + \epsilon)\} \cap \{\omega : \frac{1}{n}f_n(\omega) \in (L' - \epsilon, L' + \epsilon)\}$$

has positive measure, whence $|L - L'| < 2\epsilon$. ϵ was arbitrary though, so that $L = L'$. \square

We now modify Kingman's original result in a few useful directions. One may very rapidly obtain the following, which may be found in [8]:

Lemma 26. *Let $\{f_n\}_{n \in \mathbb{N}} \subseteq L^1\Omega$ be a subadditive family of functions. Then*

$$\frac{1}{n}f_{n \rightarrow 2n} \rightarrow C = \lim_{n \rightarrow \infty} \frac{1}{n} \int f_n \quad \text{pointwise almost everywhere.}$$

Proof. The lower bound is immediately clear, since

$$\frac{1}{n}f_{n \rightarrow 2n}(\omega) \geq \frac{1}{n}(f_{0 \rightarrow 2n} - f_{0 \rightarrow n}) \rightarrow 2C - C = C \text{ as } n \rightarrow \infty.$$

On the other hand, let $\epsilon > 0$, then since $\frac{1}{n} \int f_n \rightarrow C$ there exists some N such that for all $n \geq N$, $\frac{1}{n} \int f_n \leq C + \epsilon$. Given some fixed $n \geq N$ and some $j < N$, set $j = \lfloor \frac{n}{N} \rfloor$ and break up the interval $[0, n) = [n, n+j) \sqcup [n+j, n+j+N), \dots \sqcup [n+j+(t-1)N, n+j+\lfloor \frac{n}{N} \rfloor N) \sqcup [n+j+\lfloor \frac{n}{N} \rfloor N, 2n)$. Applying subadditivity in this manner

$$\begin{aligned} f_{n \rightarrow 2n} &\leq f_{n \rightarrow n+j} + \sum_{i=0}^{t-1} f_{n+j+iN \rightarrow n+j+(i+1)N}(\omega) + f_{n+j+tN \rightarrow n} \\ &\leq S_N |f_1| \circ \sigma^n(\omega) + \sum_{i=0}^{t-1} f_N \circ \sigma^{n+j+iN}(\omega) + S_N |f_1| \circ \sigma^{2n-N}(\omega). \end{aligned}$$

The above is valid for all $j \in [0, N)$, so summing and dividing by nN we obtain

$$\begin{aligned} \frac{1}{n}f_{n \rightarrow 2n} &\leq \frac{1}{n}(S_N |f_1| \circ \sigma^n + S_N |f_1| \circ \sigma^{2n-N}) + \frac{1}{nN} S_n f_N \circ \sigma^n \\ &\leq 2\epsilon + \frac{1}{Nn} n(N(C + \epsilon)) \end{aligned}$$

for sufficiently large n . Since ϵ was arbitrary, the result is proven. \square

A similar result for $f_{-n \rightarrow n}$ will be crucial to the main result, but it doesn't appear to be so easily obtained. As such, the proof here is novel, and uses the Backward Vitali 20 result.

Theorem 27 (Kingman's theorem for balanced intervals). *Let $(\Omega, \sigma, \mathbb{P})$ be an invertible ergodic system on a Lebesgue probability space, and let $(f_n)_{n \in \mathbb{N}}$ be subadditive sequence*

of measurable functions from Ω to \mathbb{R} with $f_1^+ \in L^1(\Omega)$. Then

$$\frac{1}{2n}f_{2n}(\sigma^{-n}\omega) \rightarrow C = \lim_{n \rightarrow \infty} \frac{1}{n} \int f_n$$

pointwise almost everywhere.

Proof. Write

$$\underline{C}_\omega = \liminf_{n \rightarrow \infty} \frac{1}{2n}f_{2n}(\sigma^{-n}\omega) \quad \text{and} \quad \bar{C}_\omega = \limsup_{n \rightarrow \infty} \frac{1}{2n}f_{2n}(\sigma^{-n}\omega).$$

Kingman's theorem immediately yields an upper bound: since

$$\frac{1}{2n}f_{2n}(\sigma^{-n}\omega) \leq \frac{1}{2n}(f_n(\sigma^{-n}\omega) + f_n(\omega)) \rightarrow \frac{1}{2}C + \frac{1}{2}C \quad \text{almost everywhere,}$$

it is clear that

$$\bar{C}_\omega = \limsup_{n \rightarrow \infty} \frac{1}{2n}f_{2n}(\sigma^{-n}\omega) \leq C.$$

If $C = -\infty$ then there is nothing more to show. Otherwise,

$$\begin{aligned} \underline{C}_{\sigma^{-1}\omega} &= \liminf_{n \rightarrow \infty} \frac{1}{2n}f_{2n} \circ \sigma^{-n}(\sigma^{-1}\omega) \\ &\geq \liminf_{n \rightarrow \infty} \left(\frac{1}{2n}(f_{2(n+1)} \circ \sigma^{-(n+1)}(\omega) - f_2 \circ \sigma^{n-1}(\omega)) \right) \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{2n}f_{2n} \circ \sigma^{-n}(\omega) - \limsup_{n \rightarrow \infty} \frac{1}{2n}f_2 \circ \sigma^{n-1}(\omega) \\ &= \underline{C}_\omega - 0, \end{aligned}$$

whence $\underline{C} \leq \bar{C}$ is constant. Fix $\epsilon > 0$. For almost every $\omega \in \Omega$ the set of j such that $f_{-j \rightarrow j} < 2j(\underline{C} + \epsilon)$ is infinite, so denote the ordered elements

$$\{j : f_{-j \rightarrow j}(\omega) < 2j(\underline{C} + \epsilon)\} = \{j_k(\omega) : k \in \mathbb{N}\}.$$

Since f is integrable, it is uniformly integrable and so there exists a $\delta > 0$ such that if $\mathbb{P}(A) < \delta$ then $\int_A f < \epsilon$. Apply Backward Vitali to the j_i s and the $-j_i$ s with parameter $\epsilon' := \min\{\epsilon, \delta\}$ to obtain an $A \subseteq \Omega$ and a $j : A \rightarrow \mathbb{N}$ according to Theorem 20 such that writing $B = \Omega \setminus \bigcup_{\omega \in A} I_\omega$, we have $\mathbb{P}(B) < \delta$ and so

$$\int_{\Omega \setminus \bigcup_{\omega \in A} I_\omega} f_1^+ < \epsilon.$$

While every $j(\omega) = j_i(\omega)$, for some i so that for every $\omega \in A$,

$$f_{-j(\omega) \rightarrow j(\omega)}(\omega) \leq 2j(\underline{C} + \epsilon).$$

Write $\Lambda_\pm = \{\sigma^{\pm j_i(\omega)}(\omega) : \omega \in A\}$. We denote the order-guaranteed return times for

Λ_{\pm} by $r_{\pm} \in \mathcal{M}(\Omega \rightarrow \mathbb{N})$, that is:

$$\begin{aligned} r_-(\omega) &= \min\{t \in \mathbb{N}_0 : \sigma^t \omega \in \Lambda_-\}, \\ r_+(\omega) &= \min\{t > r_-(\omega) : \sigma^t \omega \in \Lambda_+\}. \end{aligned}$$

r_- is the earliest nonnegative time such that $\sigma^{r_-(\omega)} \omega$ is the start of a period at growth rate guaranteed close to \underline{C} , r_+ by contrast seeks the first time after r_- which was the end of such an interval. Each are almost surely finite since $\mathbb{P}(\Lambda_{\pm}) = \mathbb{P}(A) > 0$.

The idea of the next step is that $[0, n)$ may be measurably broken up into intervals of slow growth, with small gaps in-between them. Set $b_0 = 0$ and recursively iterate along the orbit for $i > 0$:

$$\begin{aligned} a_i(\omega) &= r_-(\sigma^{b_{i-1}(\omega)} \omega), \\ b_i(\omega) &= r_+(\sigma^{b_{i-1}(\omega)} \omega). \end{aligned}$$

For any $m \in \mathbb{N}_0$ we may set $t_m(\omega) = \max\{t : b_t(\omega) \leq m\}$. Since j and r_{\pm} are measurable and the $t_m(\omega) \rightarrow \infty$ as $m \rightarrow \infty$, it is possible to pick an N such that for all $n \geq N$:

- $\mathbb{P}(j \geq N) < \frac{1}{4}\epsilon'$,
- $\mathbb{P}(t_n = 0) < \frac{1}{4}\epsilon'$ and
- $\mathbb{P}(r_{\pm} \geq N) < \frac{1}{4}\epsilon'$.

Let $n \geq N$. The orbit of ω is considered over the following intervals of times:

$$\begin{aligned} [0, n) &= [0, a_1(\omega)) \cup [a_1(\omega), b_1(\omega)) \cup \\ &\quad [b_1(\omega), a_2(\omega)) \cup [a_2(\omega), b_2(\omega)] \cup \\ &\quad \dots \\ &\quad [b_{t_n(\omega)-1}(\omega), a_{t_n(\omega)}(\omega)) \cup [a_{t_n(\omega)}(\omega), b_{t_n(\omega)}(\omega)) \cup \\ &\quad [b_{t_n(\omega)}(\omega), n). \end{aligned}$$

Time $a_i \rightarrow b_i$ is guaranteed to have low growth:

$$f_{a_i(\omega) \rightarrow b_i(\omega)}(\omega) \leq (b_i(\omega) - a_i(\omega))(\underline{C}_{\omega} + \epsilon),$$

and in addition there is a good chance that the gaps between the a_i s and the b_i s won't be too large:

$$\mathbb{P}(a_1(\omega) \geq N) = \mathbb{P}(r_-(\omega) \geq N) < \frac{1}{4}\epsilon'.$$

Since r_- corresponds to a time at the endpoint of the period of guaranteed growth, it

holds that either $r_-(\sigma^n\omega) = n - b_{t_n(\omega)}(\omega)$ or $t_n(\omega) = 0$. Therefore,

$$\begin{aligned}\mathbb{P}(n - b_{t_n(\omega)}(\omega) \geq N) &\leq \mathbb{P}(r_-(\sigma^n\omega) \geq N) + \mathbb{P}(t_n(\omega) = 0) \\ &< \mathbb{P}(r_- \geq N) + \frac{1}{4}\epsilon' < \frac{1}{2}\epsilon' .\end{aligned}$$

Set

$$\Lambda = \{a_1 \geq N \text{ or } n - b_{t_n}(\omega) \geq N\},$$

so that it is immediate from above that $\mathbb{P}(\Lambda) < \epsilon'$. Given some measurable g write $S_p g = \sum_{i=0}^{p-1} g \circ \sigma^i$. We then repeatedly apply subadditivity to $f_{0 \rightarrow n}$:

$$\begin{aligned}f_n(\omega) &= f_{0 \rightarrow n}(\omega) \\ &\leq (f_{0 \rightarrow a_1(\omega)} + f_{a_1(\omega) \rightarrow b_1(\omega)} + \cdots + f_{a_t(\omega) \rightarrow b_{t_n(\omega)}(\omega)} + f_{b_{t_n(\omega)}(\omega) \rightarrow n})(\omega) \\ &\leq \sum_{i=1}^{t_n(\omega)} f_{a_i(\omega) \rightarrow b_i(\omega)}(\omega) + \sum_{i=1}^{t_n(\omega)} f_{b_{i-1}(\omega) \rightarrow a_i(\omega)}(\omega) + f_{b_{t_n(\omega)}(\omega) \rightarrow n}(\omega) \\ &\leq \sum_{i=1}^{t_n(\omega)} (b_i(\omega) - a_i(\omega))(C + \epsilon) + \\ &\quad \sum_{i=a_1(\omega)}^{b_{t_n(\omega)}(\omega)} 1_B(\sigma^i\omega) f_1 \circ \sigma^i + f_{b_{t_n(\omega)}(\omega) \rightarrow n}(\omega).\end{aligned}$$

In the case where $t_n(\omega) = 0$ the first two terms are empty sums. Outside Λ it is guaranteed that $[0, a_1(\omega)) \subseteq [0, M)$ and $[b_{t_n(\omega)}(\omega), n) \subseteq [n - M, n)$. Therefore

$$\begin{aligned}f_n(\omega) &\leq \sum_{i=0}^{n-1} (C + \epsilon) 1_{B^c}(\sigma^i\omega) + \sum_{i=0}^{n-1} 1_B(\sigma^i\omega) f_1(\sigma^i\omega) + \\ &\quad \sum_{i=0}^M (|f_1|(\sigma^i\omega) + |C| + \epsilon) + \sum_{i=n-M}^{n-1} (|f_1|(\sigma^i\omega) + |C| + \epsilon) + \\ &\quad 1_\Lambda(\omega) S_n |f_1|(\omega) \\ &= (C + \epsilon) S_n 1_{B^c}(\omega) + S_n (f_1 1_B)(\omega) + \\ &\quad S_M |f_1|(\omega) + S_M |f_1 \circ \sigma^{n-M}|(\omega) + 1_\Lambda(\omega) S_n |f_1|(\omega).\end{aligned}$$

It remains to check that each of these terms are small enough to yield the result when

we divide through by n and integrate:

$$\begin{aligned}
\frac{1}{n} \int (C + \epsilon) S_n 1_{B^c}(\omega) + S_n(1_B f_1)(\omega) &= \int_{\Omega} ((C + \epsilon) 1_{B^c} + f_1 1_B) \\
&= (C + \epsilon) \mathbb{P}(B^c) + \int_B |f_1| \\
&\leq \max\{(C + \epsilon)(1 - \epsilon), C + \epsilon\} + \epsilon \\
&\leq (C + \epsilon) + \epsilon(|C| + \epsilon) + \epsilon,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{n} \int (S_M |f_1|(\omega) + S_M |f_1 \circ \sigma^{n-M}|(\omega) + 1_{\Lambda}(\omega) S_n |f_1|(\omega)) \\
&= \frac{1}{n} \int (M |f_1| + M |f_1|) + \frac{1}{n} \int_{\Lambda} S_n |f_1| \\
&\leq \frac{2M\epsilon}{n} + \frac{1}{n} \sum_{i=0}^{n-1} \int_{\sigma^i \Lambda} |f_1| \\
&\leq \frac{2M\epsilon}{n} + \epsilon.
\end{aligned}$$

Putting these together we obtain

$$\begin{aligned}
(C - \epsilon) &\leq \frac{1}{n} \int_{\Omega} f_n \leq (C + \epsilon) + \epsilon(|C| + \epsilon) + \epsilon + \frac{2M\epsilon}{n} + \epsilon \\
&\rightarrow C + (3 + |C|)\epsilon \text{ as } n \rightarrow \infty.
\end{aligned}$$

Letting $\epsilon \rightarrow 0$ we obtain $C \leq \underline{C}$ as required. \square

Once this Kingman type theorem is established, some extensions are almost immediate. Even if a family f_n isn't subadditive, if weaker inequalities can be written down these results can still be applied. Quantities that can be written as the differences of subadditive families also have such quantities converge, by the convergence of the difference of two convergent sequences. In our case, this means that we may exploit the subadditivity up to a scalar constant of k dimensional volume growth for every K to conclude that the averages of singular values also converge as required.

Lemma 28. *Let sequences of measurable f_n and g_n with f_1 and $g_1 \in L^1(\Omega)$ satisfy the following weak subadditivity condition:*

$$f_{m+n}(\omega) \leq f_m(\sigma^n \omega) + f_n(\omega) + C, g_{m+n}(\omega) \leq g_m(\sigma^n \omega) + g_n(\omega) + D.$$

Then writing $h_n = f_n + tg_n$,

$$\frac{1}{n}h_n \circ \sigma^n, \frac{1}{2n}h_{2n} \circ \sigma^{-n}, \frac{1}{n}h_n \rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \int f_n + t \lim_{n \rightarrow \infty} \frac{1}{n} \int g_n.$$

Proof. If $t > 0$ then $f + tg$ is easily verified to be subadditive itself, we're mainly worried here about $t = -1$, but it's thankfully still easy to see. The map $f'_n = f_n - nC$ can be seen to be subadditive by subbing $f'_n + nC$ into the inequality f_n satisfies. Then $\frac{1}{n}f'_n$ tends to some L , and thus $\frac{1}{n}f(n) \rightarrow L + C$. Additivity and scalar multiplicity follow as standard properties of limits. \square

Chapter 3

The Grassmannian

3.1 Definition

A first course in linear algebra is likely to cover the notion of a subspace, and those associated with a particular operator, such as the domain, range, image, kernel or (generalised) eigenspaces. Though definitions vary depending upon the literature, a Grassmannian is invariably some set of subspaces of a given vector space.

A familiar object is the set of 1 dimensional subspaces, which may be identified in our case with the real projective space $\mathbb{RP}(X) = (X \setminus \{0\}) / \{x \sim -x\}$ of lines through the origin.

In this context we will refer to "the" Grassmannian $\mathcal{G}X$ as a generalisation of projective space, to include subspaces of the underlying vector space of arbitrary dimension. Though a line may be defined with no reference to a norm or other extra structure, we will work with the case where X is a separable Banach space.

Definition 29. We write $V \leq X$ if V is a vector subspace of X . Write \oplus for the (algebraic) direct sum of two vector spaces. When we write $X = E \oplus V$, this means that there is a unique linear isomorphism

$$X \xrightarrow{\Pi \times P} E \times V, x \mapsto (\Pi x, Px)$$

with inverse given by $+$: $(u, v) \mapsto u + v$. E and V are said in the literature to be complemented. When emphasising dependence of this projection upon E and V , we use the notation

$$\Pi_{E\|V} = P \quad \text{and} \quad \Pi_{V\|E} = \Pi.$$

In general the maps P and Π may not be bounded, but if they are, then E and V are said to be topologically complemented to each other.

Oseledets used the Grassmannian in [22] in order to establish measurability of a map which takes values in the subspaces of \mathbb{R}^d .

Definition 30. *The Grassmannian is defined as the set of subspaces of X which have a topological complement:*

$$\mathcal{G}X = \{V \leq X : \text{there exists a projection } X \xrightarrow{\Pi} V \text{ with } \|\Pi\| < \infty\}.$$

In addition, for $N \in \mathbb{N} \cup \{\infty\}$ and $k \in \mathbb{N}_0$ write

$$\begin{aligned} \mathcal{G}_k X &= \{E \in \mathcal{G}X : \dim E = k\}, & \mathcal{G}_{<N} X &= \sqcup_{0 \leq i < N} \mathcal{G}_i X, \\ \mathcal{G}^k X &= \{E \in \mathcal{G}X : \text{codim } E = k\} & \text{and } \mathcal{G}^{<N} X &= \sqcup_{0 \leq i < N} \mathcal{G}^i X. \end{aligned}$$

Lemma 31. *We can also define the set*

$$\tilde{\mathcal{G}}(X) = \{V \text{ closed } \leq X : \text{there exists } W \text{ closed } \leq X \text{ with } X = V \oplus W\},$$

which is in fact another way of writing down the Grassmannian, $\tilde{\mathcal{G}}(X) = \mathcal{G}(X)$.

Proof. To show that $\tilde{\mathcal{G}}(X) \supseteq \mathcal{G}X$, suppose that $V \in \mathcal{G}X$, so that there is a bounded projection $X \xrightarrow{\Pi} V$ onto the subspace V of X . Write $P = 1 - \Pi$ which is also bounded and set $E = PX$, we need to check E and V are closed. Suppose that $w_n \in E$ has $w_n \rightarrow w_\infty$ in the norm topology on X . Then since Π is continuous,

$$\|\Pi(w)\| = \lim_{n \rightarrow \infty} \|\Pi w_n\| = 0,$$

whence $w \in E$, so E is closed. Applying the same to a sequence in V and P , we conclude that $V \in \tilde{\mathcal{G}}(X)$.

To show that $\tilde{\mathcal{G}}(X) \subseteq \mathcal{G}(X)$, suppose that $V \in \tilde{\mathcal{G}}(X)$, so that there is a closed subspace W with $X = V \oplus W$, but the associated projection $X \xrightarrow{\Pi} V$ needs to be verified to be bounded. As closed subspaces V and W are Banach spaces in their own rights. $V \oplus W$ can be endowed with the structure of a Banach space by letting $\|(u, v)\|_{\oplus} = \|u\| + \|v\|$. The graph G of $P \times \Pi$ may be viewed as a subspace then of $X \times V \oplus W$. Suppose that the sequence $(x_n, v_n, w_n) \in G$ for all $n \in \mathbb{N}$ and $(x_n, v_n, w_n) \rightarrow (x_*, v_*, w_*)$. We have that $v_* \in V$ and $w_* \in W$ by closedness of V and W . $x_n = v_n + w_n \rightarrow v_* + w_*$, so $(x_*, v_*, w_*) \in G$ and G is closed. By the Closed Graph Theorem, $P \times \Pi$ must be continuous, whence each component P and Π is bounded and $V \in \mathcal{G}(X)$. \square

So far, we just have a set, but we will endow it with more structure. Write \mathbb{S}_X for the unit ball of a normed space X .

Definition 32. *Let (Γ, d) be a metric space. Given a point $\gamma \in \Gamma$ and a subset $A \subseteq \Gamma$, we define*

$$d(A, \gamma) = d(\gamma, A) = \begin{cases} \inf_{a \in A} d(a, \gamma) & \text{if } A \neq \emptyset, \\ 1 & \text{otherwise.} \end{cases}$$

Given two subsets of a metric space $A, B \subseteq \Gamma$, we define the Hausdorff distance

$$d(A, B) = \max\{\sup\{d(A, b) : b \in B\} \cup \{0\}, \sup\{d(a, B) : a \in A\} \cup \{0\}\}.$$

Lemma 33. $\mathcal{G}X$ may be metrised by the Hausdorff distance between the unit balls of spaces:

$$d(U, V) = \max\left(\sup_{x \in \mathbb{S}_U} \inf_{y \in \mathbb{S}_V} \|x - y\|, \sup_{x \in \mathbb{S}_V} \inf_{y \in \mathbb{S}_U} \|x - y\|\right)$$

In particular, when X is a separable Banach space, $\mathcal{G}_{<\infty}X$ is a separable complete metric space.

Proof. Since the Hausdorff distance is a metric on closed sets, and distinct closed subspaces have distinct unit spheres, d is certainly a metric. Separability and completeness of $\mathcal{G}_{<\infty}X$ is shown in [12] and can be seen by considering the surjective and continuous map

$$\text{span} : \{(x_1, \dots, x_k) \in X^k : \dim \text{span}\{x_i\} = k\} \rightarrow \mathcal{G}_k X.$$

□

Example 34. In \mathbb{R}^d the space may be taken to be the disjoint union of $d + 1$ smooth manifolds, each connected component corresponding to subspaces of a fixed dimension. If two subspaces have different dimensions, their distance apart must be 1, by observing the nonemptiness of the intersection of the orthogonal complement of the lower dimensional space with the higher dimensional space.

Closed finite-codimensional subspaces $\mathcal{G}^{<\infty}X$ will also feature, but just as the dual of a separable Banach space may fail to be separable, this space may not be as well behaved, and some extra care is needed to obtain results. Given $T \in \mathcal{B}(X)$ define the slowest growth of vectors in a given subspace under T :

$$g(T, V) = g_T(V) = \inf_{x \in \mathbb{S}_V} \|Tx\|.$$

By taking the supremum over $g(T, V)$ for a fixed T and the dimension of V fixed and finite, we obtain a quantity that encodes the fastest possible rate of guaranteed growth in k dimensional spaces associated with T :

$$\rho_k(T) = \sup_{V \in \mathcal{G}_k X} g(T, V).$$

The ρ_k are called Bernstein numbers in the work of Pietsch[23], which gives an overview of similar kinds of statistics in Banach spaces.

Lemma 35. Let $T \in \mathcal{B}(X, Y)$ and $S \in \mathcal{B}(Y, Z)$. Then

$$g_T(X)g_S(TX) \leq g_{S \circ T}(X) \leq \min\{\|T\|g_S(TX), g_T(X)\|S\|\}.$$

Proof. Let $\epsilon > 0$ and choose $x \in X \setminus \{0\}$ with $\|STx\| < g_{S \circ T}(X) + \epsilon$. Then

$$g_T(X)g_S(TX) \leq \frac{\|STx\|}{\|Tx\|} \frac{\|Tx\|}{\|x\|} < g_{S \circ T}(X) + \epsilon,$$

and the second inequality follows from the fact that $g_T(X) \leq \|T\|$ for any such T and X . \square

Lemma 36. Define $D_k \in \mathcal{M}(\mathcal{B}(X) \rightarrow \mathbb{R})$ by

$$D_k(T) = \sup_{u_1, \dots, u_k \in \mathbb{S}_X} \prod_{i=1}^k d(\text{span}\{Tu_j\}_{j < i}, Tu_i).$$

Then D_k is submultiplicative, and there are constants $c_k < C_k$ such that

$$c_k D_k \leq \prod_{i=1}^k \rho_i \leq C_k D_k.$$

Proof. See [13]. \square

3.1.1 A familiar pattern in the construction of measurable maps in a separable setting

Here we mention a notion that seems to be a very natural means of building measurable maps. The idea is very simple, such that its mentioning may seem barely worthwhile, and a more elucidating statement may be possible, but we give it as a lemma whose implicit use may be forgotten or re-explained in a sentence or so. We make use of the following straightforward construction:

Lemma 37. Let Ω be a measurable space. Suppose that $\{x_n\}_{n \in \mathbb{N}} \subseteq X$ for some measurable space X . Then if we can write down a measurable set $S_n \subseteq \Omega$ for each $n \in \mathbb{N}$ such that $\bigcup_n S_n = \Omega$ then there is an associated measurable $N: \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ given by $N_\omega = \inf\{n : \omega \in S_n\}$. In fact N is almost everywhere finite and may be taken \mathbb{N} -valued since

$$\mathbb{P}\left(\bigcup_{m < n} S_m\right) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Further, the map $\omega \mapsto x_{N_\omega}$ is measurable.

This principle is applied to check measurability in the context of spaces which are separable.

3.2 Measurable statistics and the Grassmannian

3.2.1 A motivating example: singular value decomposition

Finite dimensional or Hilbert space dynamics confer lots of useful tools to keep track of behaviour of elements of the space.

In finite dimensions, any matrix may be written $A = U_A D_A V_A^T$, where U_A and V_A are orthogonal and D_A nonnegative and diagonal. U, V and $D : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ may be taken to be measurable maps. They may be taken smooth in a neighbourhood of any A such that the singular values are disjoint and positive. Geometrically, D describes the principal axes of the ellipsoid $A\mathbb{S}^{d-1}$, that are dilated by a factor of the singular value. The orthogonal matrices U and V specify two decompositions of $\mathbb{R}^d = \bigoplus_{i=1}^k E_k = \bigoplus_{i=1}^k E'_k$ into a series of orthogonal spaces growing at distinct rates $g_1 > \dots > g_k \geq 0$ with $TE_k = E'_k$ if $g_i > 0$.

One may use singular value decomposition in the course of the proof of some versions of the multiplicative ergodic theorem, for example, that of Raghunathan[24], and even by itself it represents a motivating result. With this being said, the multiplicative ergodic theorem may perhaps be more directly related with the Jordan normal form as per[1].

Here we formulate weaker statements about the rates of growth of operators on X with the aid of the Grassmannian: even in Hilbert space, the resulting decomposition may no longer be into orthogonal spaces, but will describe distinct growth rates.

3.2.2 Measurable statistics for linear operators

Lemma 38. *Let X be a normed space. Then $g : \mathcal{B}(X) \times \mathcal{G}_k X \rightarrow [0, \infty)$ is continuous.*

Proof. Let $\epsilon > 0$. Here use the following metric inducing the product topology:

$$d((S, V), (T, W)) = \|S - T\| + d(V, W).$$

First note that g is continuous at $(0, V)$ for any V since $g(0, V) = 0$ and if $d(V, W) + \|T - 0\| < \epsilon$ then $g(T, W) < \epsilon$. Otherwise, the projection onto the sphere $p_S T = \frac{1}{\|T\|} T$ is continuous. Let $\epsilon > 0$. Since $g(T, V) = \|T\|g(p_S T, V)$ for nonzero T , it suffices to check g is continuous on $\mathbb{S}_{\mathcal{B}(X)} \times \mathcal{G}_k X$. Let

$$(S, V), (T, W) \in \mathbb{S}_{\mathcal{B}(X)} \times \mathcal{G}_k X$$

be such that

$$d((S, V), (T, W)) = \|S - T\| + d(V, W) < \epsilon.$$

Since V and W are finite dimensional their spheres are compact and we may choose

an $x \in \mathbb{S}_V$ with $\|Sx\| = g(S, V)$. Choose a $y \in \mathbb{S}_W$ with $\|x - y\| < \epsilon$. Then

$$\begin{aligned} g(T, W) &\leq \|Ty\| \leq \|T - S\| + \|S\| \|y - x\| + \|Sx\| \\ &< \epsilon + \epsilon + g(S, V). \end{aligned}$$

By the same argument swapping (S, V) and (T, W) then, $g(S, V) < g(T, W) + 2\epsilon$ and continuity (in fact, uniform continuity) on $\mathbb{S}_{\mathcal{B}(X)} \times \mathcal{G}_k X$ is clear. Thus g is continuous on $\mathcal{B}(X) \times \mathcal{G}_k X$ as required. \square

Corollary 39. *The quantities $\rho_k : \mathcal{B}(X) \rightarrow \mathbb{R}$ are measurable functions, where measurability is with respect to the Borel σ -algebra on $\mathcal{B}(X)$ with the strong operator topology.*

Proof. The ρ_k may now be written as the supremum of the functions $\{\varphi_V : T \mapsto g(T, V) : V \in \mathcal{G}_k X\}$, since choosing a collection $\{V_i\}_{i \in \mathbb{N}}$ which is dense in $\mathcal{G}_k(X)$, we may write ρ as the supremum of a countable family

$$\rho_k(T) = \sup_{i \in \mathbb{N}} \varphi_{V_i}(T),$$

since $\mathcal{G}_k X$ is separable. \square

We make liberal use of a simple geometric fact about k dimensional fast growing spaces when $\rho_{k+1}(T)$ is smaller than $\rho_k(T)$. We will be mainly concerned with cases where $\rho_{k+1}(T)$ is much smaller than $\rho_k(T)$, which forces images of any two sufficiently fast growing spaces to be close together.

Lemma 40. *Grassmannian contraction estimates: Let $T \in \mathcal{B}(X)$ with $\rho_{k+1}(T) < \rho_k(T)$ and $\Theta \in (\rho_{k+1}T, \rho_k T)$. Suppose that $V, W \in \mathcal{G}_k X$ are choices of fast growing spaces:*

$$g_T(V), g_T(W) > \Theta.$$

Then

$$d(TV, TW) < \frac{2}{1 - \frac{\rho_{k+1}T}{\Theta}} \frac{\rho_{k+1}T}{\Theta}.$$

In particular, if $\Theta > 2\rho_{k+1}T$ then

$$d(TV, TW) < 4 \frac{\rho_{k+1}T}{\Theta}.$$

Proof. If $V = W$ there is nothing to prove so assume otherwise. Let $a \in V \setminus W$. Then $W \oplus \text{span}\{a\} \in \mathcal{G}_{k+1} X$ so that there is some $a - b \in W \oplus \text{span}\{a\}$ with $b \in W$ and $g_T(a - b) \leq \rho_{k+1}T$. Set $c = a - b$. By the uniformly large growth of the spaces, $\|a\| \leq \frac{\|Ta\|}{\Theta}$ and $\|b\| \leq \frac{\|Tb\|}{\Theta}$.

$$\begin{aligned}
\|Tc\| &\leq \|c\|\rho_{k+1}T \leq (\|a\| + \|b\|)\rho_{k+1}T \\
&\leq \frac{\|Ta\| + \|Tb\|}{\Theta}\rho_{k+1}T \\
&\leq \frac{2\|Ta\| + \|Tc\|}{\Theta}\rho_{k+1}T,
\end{aligned}$$

So that

$$\|Tc\| < \frac{2}{1 - \frac{\rho_{k+1}T}{\Theta}} \frac{\rho_{k+1}T}{\Theta} \|Ta\|.$$

Since a isn't in the kernel, $d\left(\frac{Ta}{\|Ta\|}, TW\right) \leq \frac{\|Tc\|}{\|Ta\|} < \frac{2}{1 - \frac{\rho_{k+1}T}{\Theta}} \frac{\rho_{k+1}T}{\Theta}$. Since the choice of a was arbitrary,

$$d(TV, TW) < \frac{2}{1 - \frac{\rho_{k+1}T}{\Theta}} \frac{\rho_{k+1}T}{\Theta}.$$

□

We have made precise the notion that under appropriately spaced growth statistics, fast growing subspaces are forced close together under a linear map.

Lemma 41. *Any n dimensional normed space X has a basis of unit vectors $(x_i)_{i=1}^n$ with*

$$d(\text{span}\{x_i : i < j\}, x_j) = 1.$$

Proof. Pick a basis $(y_i)_{i=1}^n$, and suppose that for some $1 \leq l < n$,

- we have obtained x_i for $i < l$ and
- $\text{span}\{x_i : i < l\} = \text{span}\{y_i : i < l\}$.

Pick a $z \in \{a \in \text{span}\{x_i : i < j\} \mid \|y_j - a\| \text{ is minimized}\}$ which minimises $\|y_j - z\|$ - this is attained since the minimiser must be in the compact ball of vectors of norm at most $2\|y_j\| = 2$, and set

$$x_j = \frac{y_j - z}{\|y_j - z\|}.$$

Then adding x_j to the span is equivalent to adding y_j to the span, and for $v \in \text{span}\{x_j : j < i\}$ we have

$$\|v - x\| = \frac{1}{\|y_j - z\|} \|y_j - (z - \frac{1}{\|y_j - z\|}v)\| \geq \frac{1}{\|y_j - z\|} \|y_j - z\| = 1,$$

whence $d(\text{span}\{x_i : i < j\}, x_j) = 1$. Iterate to obtain all $(x_i)_{i=1}^n$ with the desired property. □

Lemma 42. *Let X be a separable Banach space and $k \in \mathbb{N}$. Then there is a Borel measurable $U : \mathcal{G}_k X \rightarrow \mathcal{G}_{k-1} X$ such that $U(V) < V$ for all $V \in \mathcal{G}_k X$.*

Proof. Let $\{U_1, U_2, \dots\}$ be dense in $\mathcal{G}_{k-1} X$. Suppose $V \in \mathcal{G}_k X$. Let $s_0(V) = 1$ and

$$s_i(V) = \inf\{s : \sup_{v \in \mathbb{S}_{U_s}} d(v, V) < 2^{-i} \text{ and } d(U_s, U_{s_{i-1}}(V)) < 2^{1-i}\} \text{ for } i > 0.$$

In other words, $U_{s_i(V)}$ is the first member of the dense collection to fall in the nonempty open set

$$B_{2^{1-i}}(U_{s_{i-1}}) \cap \bigcup_{W \in \mathcal{G}_{k-1} X : W < V} B_{2^{-i}}(W).$$

Then U_{s_i} is pointwise Cauchy and so the pointwise limit of measurable functions $U_{s_i(V)}$ is convergent to some measurable $U(V) \in \mathcal{G}_{k-1} X$. Further, $\sup_{v \in \mathbb{S}_{U(V)}} d(v, V) \leq 2^{-i}$ for any i , so $U(V) < V$. \square

Lemma 43. *Let X be a separable Banach space and $k \in \mathbb{N}$. Let $U \in \mathcal{M}(\mathcal{G}_k X \rightarrow \mathcal{G}_{k-1} X)$ (where recall measurability is taken to be with respect to the Borel sigma algebras) with $U(V) < V$ as guaranteed by the previous lemma. Then there is a Borel to norm-topology Borel measurable $b : \mathcal{G}_k X \rightarrow \mathbb{S}_X$ such that $U \oplus \text{span}\{b\}$ is the identity on $\mathcal{G}_k X$.*

Proof. Let $\{x_1, x_2, \dots\}$ be dense in \mathbb{S}_X . We iteratively define a family of measurable maps $t_i : \mathcal{G}_k X \rightarrow \mathbb{N}$ by $t_0(V) = 1$ and

$$t_i(V) = \inf\{t : d(U(V) \oplus \text{span}\{x_t\}, V) < 3^{-t}, d(U(V), x_t) > \frac{1}{2} \text{ and } \|x_t - x_{t_{i-1}}\| < 2^{1-i}\}$$

for $i > 0$. Each t_i is the first time when running through the countable dense set that we fall into a desired open set, so finite. The sequence of functions $V \mapsto x_{t_i(V)}$ is pointwise Cauchy so converges pointwise to some measurable b with the desired properties. \square

Applying lemmas 43 and 44 recursively we obtain the following:

Lemma 44. *Let X be a separable Banach space. Then for every $k > 0$ there is a measurable map $b : \mathcal{G}_k X \rightarrow \mathbb{S}_X^k$ such that for each $V \in \mathcal{G}_k X$,*

$$V = \text{span}\{b_i(V)\}_{i=1}^k.$$

Lemma 45. *Let $V \in \mathcal{G}^k X$ with bounded projection $\Pi : X \rightarrow V$. Let $T \in \mathcal{B}(X)$. Then $\rho_{k+l}(T) \leq \rho_l(T \circ \Pi)$.*

Proof.

$$\begin{aligned}
\rho_{k+l}(T) &= \sup_{U \in \mathcal{G}_{k+l}X} g_T(U) \leq \sup_{U \in \mathcal{G}_{k+l}X} g_T(U \cap V) \\
&= \sup_{U \in \mathcal{G}_{\leq k+l, \geq l}V} g_T(U) = \sup_{U \in \mathcal{G}_lV} g_{T|_V}(U) \\
&= \sup_{U \in \mathcal{G}_lV} g_{T \circ \Pi}(U) \leq \rho_l(T \circ \Pi).
\end{aligned}$$

□

Lemma 46. *Let $X = E \oplus V = E' \oplus V'$ with corresponding projections $\Pi : X \rightarrow V \in \mathcal{G}^k X$ and $\Pi' : X \rightarrow V' \in \mathcal{G}^k X$. Let $T \in \mathcal{B}(X)$ such that $\Pi' \circ T = T \circ \Pi$, and suppose further that $g_T(E) > \|T|_V\|$. Then $\rho_l(T \circ \Pi) \leq 4\|\Pi\|\|\Pi'\|\rho_{l+k}(T)$.*

Proof. Let $U \in \mathcal{G}_lV$ with $g_{T \circ \Pi}(U) > e^{-\epsilon}\rho_l(\Pi' \circ T|_V) = e^{-\epsilon}\rho_l(T|_V)$. Let $x \in \mathbb{S}_{U \oplus E}$. Observe that $\|\Pi'Tx\| \leq \|\Pi'\|\|Tx\|$ and similar for $1 - \Pi'$. Then

$$\begin{aligned}
\|Tx\| &\geq \max\left\{\frac{\|\Pi'Tx\|}{\|\Pi'\|}, \frac{\|(1-\Pi')Tx\|}{\|1-\Pi'\|}\right\} \geq \max\left\{\frac{g_{T\Pi}(U)\|\Pi x\|}{\|\Pi'\|}, \frac{g_T(E)\|(1-\Pi)x\|}{\|1-\Pi'\|}\right\} \\
&\geq \frac{1}{1+\|\Pi'\|} \max\{g_{T\Pi}(U)\|\Pi x\|, g_T(E)\|(1-\Pi)x\|\} \\
&\geq \frac{e^{-\epsilon}\rho_l(T|_V)}{1+\|\Pi'\|} \max\{\|\Pi x\|, \|(1-\Pi)x\|\} \geq \frac{e^{-\epsilon}\rho_l(T|_V)}{4\|\Pi'\|} \geq \frac{e^{-\epsilon}\rho_l(T \circ \Pi)}{4\|\Pi'\|\|\Pi\|},
\end{aligned}$$

where on the last line we use the fact that $\rho_l(T \circ \Pi) \leq \|\Pi\|\rho_l(T|_V)$. Since x was arbitrary, we may then conclude that $g_T(U \oplus E) \geq \frac{e^{-\epsilon}\rho_l(T \circ \Pi)}{4\|\Pi'\|\|\Pi\|}$ with $\epsilon > 0$ arbitrary, so that

$$\rho_l(T \circ \Pi) \leq 4\rho_{l+k}(T \circ \Pi)\|\Pi'\|\|\Pi\|$$

as required. □

3.2.3 Equivalence of growth statistics

The Gelfand numbers may be defined by $s_k(T) = \inf_{V \in \mathcal{G}^{k-1}X} \|T|_V\|$. Throughout this section we assume $T \in \mathcal{B}(X)$.

Lemma 47. $\rho_k \leq s_k$.

Proof. Let $\epsilon > 0$. Choose $V \in \mathcal{G}_k X$ with $g_T(V) \geq \rho_k(T) - \epsilon$. Then since any $W \in \mathcal{G}^{k-1}X$ intersects V nontrivially, suppose x is a unit vector in the intersection. Then $\rho_k(T) - \epsilon \leq \|Tx\| \leq \|T|_W\|$. W was arbitrary so $\rho_k(T) \leq s_k(T)$. □

A better bound than that in the following lemma may be found in the work of Pietsch on s -numbers[23], who axiomatised and gave a few examples of these kinds of easily written down statistics that turn out to agree up to some constant, fixing dimension:

Lemma 48. For all k , $s_k \leq 4^{k-1}((k-1)!)^{\frac{1}{2}}\rho_k$.

Proof. For $k = 1$ each quantity is just the norm of T , so equality holds. Write $c_i = 4^{1-i}((i-1)!)^{-\frac{1}{2}}$. Suppose that the proposition holds for each $l < k$ with $k \geq 2$. Choose $U \in \mathcal{G}_{k-1}X$ with $g_T(U) \geq (1-\epsilon)\rho_{k-1}(T)$, so that in particular

$$g_T(U) \geq (1-\epsilon)c_{k-1}s_{k-1}(T) \geq (1-\epsilon)c_{k-1}s_k(T).$$

We may by the theorem of Kadets appearing as theorem 10 in section III.B of [29] choose a complement $TU \oplus V = X$ with $\Pi = \Pi_{TU \parallel V}$ such that $\|\Pi\| \leq (k-1)^{\frac{1}{2}} + \epsilon$. Since $T^{-1}V \in \mathcal{G}^{k-1}X$, we may choose $x \in \mathbb{S}_{T^{-1}V}$ with $\|Tx\| \geq s_k(T)(1-\epsilon)$. Set $W = U \oplus \text{span}\{x\} \in \mathcal{G}_k X$ and let $a = b + tx \in \mathbb{S}_W$. Applying T ,

$$\begin{aligned} \|Ta\| &\geq \max \left\{ \frac{\|Tb\|}{\|\Pi\|}, \frac{\|tTx\|}{\|1-\Pi\|} \right\} \\ &\geq (1 + \sqrt{k-1} + \epsilon)^{-1} \max\{(1-\epsilon)c_{k-1}s_k(T)\|b\|, s_k(T)(1-\epsilon)|t|\} \\ &\geq (2\sqrt{k-1} + \epsilon)^{-1} (1-\epsilon)c_{k-1} \max\{\|b\|, |t|\} \\ &\geq \frac{(1-\epsilon)c_{k-1}}{4(\sqrt{k-1} + \epsilon)} s_k(T). \end{aligned}$$

a was arbitrary so the final line is a lower bound for $g_T(W)$. ϵ was also arbitrary, so that $\rho_k(T) \geq \frac{1}{4c_{k-1}\sqrt{k-1}}s_k(T) = c_k s_k(T)$. \square

The operator norm on $\mathcal{B}(X)$ has already been exploited extensively. Another useful piece of information about an operator is its index of compactness seminorm

$$\|T\|_c = \inf\{\|T - T'\| : T \in \mathcal{B}(X), \dim(T(X)) < \infty\}.$$

Lemma 49. The Gelfand numbers satisfy

$$s_k(T) \geq k^{-\frac{1}{2}}\|T\|_c.$$

Proof. Let $V \in \mathcal{G}^{k-1}X$. Similarly to the proof of the last lemma, we may choose a complement $V \oplus W = X$ with $\|\Pi_{V \parallel W}\| \leq \sqrt{k-1} + \epsilon$. Then $T \circ \Pi_{W \parallel V}$ is finite rank, so $\|T\|_c \leq \|T - T \circ \Pi_{W \parallel V}\| = \|T \circ \Pi_{V \parallel W}\| \leq \|T|_V\| \|\Pi_{V \parallel W}\| \leq \|T|_V\|(\sqrt{k-1} + \epsilon)$. Taking the inf over such V and letting $\epsilon \rightarrow 0$ we obtain the bound. \square

Chapter 4

Cocycles

4.1 Motivation

4.1.1 What is a cocycle?

The term cocycle is part of the vocabulary across a wide range of areas of mathematics. In the discrete time, linear setting considered here, a cocycle corresponds to a linear operator-valued function:

Definition 50. *Let $(\Omega, \sigma, \mathbb{P})$ be a measure preserving system, let X be a vector space and suppose that $A : \Omega \times X \rightarrow X$. Equivalently, A may be thought of as a map $\Omega \rightarrow X^X$.*

- *If for fixed ω A_ω is a linear map, then call A linear cocycle. Equivalently, a cocycle is a map $\Omega \times \mathbb{N}_0 \times X \rightarrow X$ written $(\omega, t, x) \mapsto A_\omega^{(t)}x \in X$ that satisfies the cocycle condition*

$$A_\omega^{(m+n)} = A_{\sigma^n \omega}^{(m)} \circ A_\omega^{(n)}.$$

- *Suppose in addition that X is a Banach space and $\mathcal{V} \subseteq \mathcal{B}(X)$ is a Banach space of bounded linear operators on X . The strong operator topology on \mathcal{V} is the smallest making the maps*

$$\varphi_x : \mathcal{V} \rightarrow X, T \mapsto Tx$$

continuous (with X endowed with its norm topology). If A is measurable as a map from Ω to \mathcal{V} with the strong operator topology on \mathcal{V} , then call A strongly measurable, writing $A \in \mathcal{SM}(\Omega \rightarrow \mathcal{V})$.

- *if X is a normed space and $\log^+ \|A\| \in L^1\Omega$ then say A is forward integrable.*
- *if X is a normed space write $GL(X)$ for the group of bounded invertible maps. If $A_\omega \in GL(X)$ for almost every $\omega \in \Omega$ then call A an invertible cocycle, denoting this subset of $\mathcal{SM}(\Omega \rightarrow \mathcal{B}(X))$ by $\mathcal{SM}(\Omega \rightarrow GL(X))$.*

- If X is a normed space, $(A_\omega)^{-1}$ exists for pointwise almost every $\omega \in \Omega$ and is forward integrable then say A is backward integrable.

One may keep in mind that we may define a cocycle for continuous time as a measurable $A : \Omega \times [0, \infty) \rightarrow \Omega$ such that $A_\omega^{(s+t)} = A_{\theta_t \omega}^{(s)} A_\omega^{(t)}$ for all $s, t > 0$ but we will focus on discrete time.

Lemma 51. *We have the following characterisation of the space of strongly measurable functions:*

$$\mathcal{SM}(\Omega \rightarrow \mathcal{B}(X)) = \left\{ A : \Omega \rightarrow \mathcal{B}(X) : \begin{array}{l} \text{for each } x \in X, \\ (\omega \mapsto A_\omega x) \in \mathcal{M}(\Omega \rightarrow X) \end{array} \right\}.$$

Proof. Lemma A.4 of [12] checks the equivalence of these conditions. \square

In his paper [22] Oseledets points out that with the framework he develops, one may not only study systems like those arising from differential equations, but also to problems arising in probability theory.

Some examples include

- In the case of a constant map $A_\omega = A_0$ we obtain a constant cocycle corresponding to the application of the same linear map at each time step.
- Extensively studied is the case where the base system $\sigma : \Omega \rightarrow \Omega$ is also a diffeomorphism of a compact manifold with \mathbb{P} some invariant or ergodic measure. One may consider the derivative of σ^n to be a map on \mathbb{R}^n up to a choice of atlas: fix a finite covering of patches $\varphi : \mathbb{R}^n \rightarrow U_i$ open in M and let $V_i = U_i \setminus \bigcup_{j < i} U_j$. The derivative $A|_{V_i \cap \sigma^{-1} V_j} = d(\varphi_i \sigma \varphi_j^{-1})$ is a linear cocycle where $A^{(n)} = d\sigma^n$.
- A simple function $A : \Omega \rightarrow \mathcal{B}(X)$ is one defined piecewise constant on the elements of a partition: that is, there exists some countable family of measurable sets \mathcal{P} and with $\bigsqcup \mathcal{P} = \Omega$ and an associated family of bounded operators $(A_P)_{P \in \mathcal{P}}$ such that $A_\omega = A_P$ for $\omega \in P \in \mathcal{P}$. These are simple to understand, but one can easily be designed to demonstrate various forms of problematic behaviour.
- A cocycle on \mathbb{R}^n of iid or stationary stochastic operators is the case considered by Furstenberg and Kesten[10].

Lemma 52. *Suppose that X is a Banach space and Ω a measurable space. If $E \in \mathcal{M}(\Omega \rightarrow \mathcal{G}_k X)$ and $A \in \mathcal{SM}(\Omega \rightarrow \mathcal{B}(X))$ then the pointwise pushforward $\omega \mapsto A(\omega)E(\omega)$ is measurable.*

Proof. The pushforward result is established in corollary B.13 of [12]. \square

The emphasis here will be on establishing the long term behaviour of elements of X . These techniques may be used to study continuous time systems, and since they are another huge source of examples, I elucidate a little here:

4.1.2 Differential examples

Recall the initial value problem discussed in the introduction:

$$\dot{x}(t) = a(\theta_t\omega)x(t), x(0) = x_0,$$

which has a unique solution given by the operator

$$A_{0\omega}^{(t)} \in \mathbb{R}^{d \times d}.$$

In this cocycle formulation, for example, the Euler scheme for numerical integration can be seen to hinge on the assumption that in the right circumstances, for small δ , we can make the approximation $A_\omega \approx 1 + \delta a(\omega)$.

The first differential equation is finite dimensional and well studied, but an infinite dimensional state space may also arise naturally in the continuous time space: with A as above and $b \in \mathcal{SM}(\Omega \rightarrow \mathbb{R}^{d \times d})$ such that $\omega \mapsto \int_0^1 \|b_{\theta_t\omega}\| dt \in L^1(\Omega)$, consider the initial value delay differential equation:

$$\dot{x}(t) = a(\theta_t\omega)x(t) + b(\theta_t\omega)x(t-1) \text{ for } t > 1 \text{ and } x|_{[0,1]} = u \in X = C([0,1] \rightarrow \mathbb{R}^n).$$

Whereas $A_{0\omega} = A_{0\omega}^{(1)} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ depended on the finite dimensional datum $x_0 \in \mathbb{R}^d$, the solution operator here $A \in \mathcal{SM}(\Omega \rightarrow \mathcal{B}(X))$ depends on the infinite dimensional datum $u \in X$ - even a Hilbert space result such as that of Ruelle may be insufficient here. This evolution of a unit timestep is entirely determined by the last unit of time, and can be written in terms of A_0, a and b . Writing $A_\omega : C[0,1] \rightarrow C[0,1]$ for the unit time evolution operator composed with the shift by one unit time step is computationally convenient. In fact, in [15] it is shown that regardless of whether one focuses on the problem in a continuous or L^p setting, the evolution operator over a duration greater than 1 or 2 is compact. As such, when studying long term behaviour one may focus on a compact discrete time system acting on X and apply results proved in that context.

4.2 Defining the Lyapunov exponents

4.2.1 Applying statistics from the Grassmannian to cocycles

We shall see that the growth statistics of a given cocycle can then be checked to converge.

Given a cocycle $A : \Omega \rightarrow \mathcal{B}(X)$, for fixed ω and each $b \geq a \in \mathbb{Z}$ we may define the map $A_{a \rightarrow b}(\omega) := A_{\sigma^a \omega}^{(b-a)}$, which may be thought of as the evolution rule for X from time a to time b . Under this notation it is easy to see that for any $a < c < b \in \mathbb{Z}$, $A_{c \rightarrow b} \circ A_{a \rightarrow c} = A_{a \rightarrow b}$. The inequality $\|A_{a \rightarrow b}\| \leq \|A_{a \rightarrow c}\| \|A_{c \rightarrow b}\|$ holds. Taking log of

each side, one obtains the following triangle inequality-like bound on the growth of points in X from time a to b :

$$f_{a \rightarrow b}(\omega) := \log \|A_{a \rightarrow b}\| \leq \log \|A_{a \rightarrow c}\| + \log \|A_{c \rightarrow b}\| = f_{a \rightarrow c}(\omega) + f_{c \rightarrow b}(\omega).$$

This is an example of a subadditive cocycle, exactly the object that we can apply Kingman theorems to:

Definition 53. *Given a forward-integrable linear dynamical system $\mathcal{R} = (\Omega, \mathbb{P}, \sigma, X, A)$ with $(\Omega, \mathbb{P}, \sigma)$ an invertible ergodic transformation over an atom-free Lebesgue probability space, write μ_i for the constant pointwise almost everywhere limit*

$$\mu_i = \lim_{n \rightarrow \infty} \frac{1}{n} \log \rho_i(A_\omega^{(n)}) \quad \text{for almost every } \omega \in \Omega$$

where it exists, as shall be shown below. The numbers $(\mu_i)_{i=1}^\infty$ are nonincreasing. We may write λ_i for these numbers counted without repeats, and m_i for the multiplicity of λ_i , that is, if λ_{i+1} is defined then m_i is defined as the largest $m \in \mathbb{N}$ such that

$$\mu_{m_1 + \dots + m_{i-1} + m} = \lambda_i.$$

Call the λ_i s the Lyapunov exponents, λ_1 being the top Lyapunov exponent. Let

$$\lambda_\omega(x) = \limsup_n \frac{1}{n} \log \|A_\omega^{(n)}x\| \in [-\infty, \infty)$$

for the Lyapunov exponent associated with $x \in X$.

Finally, define the limit $\nu = \lim_{i \rightarrow \infty} \mu_i \in [-\infty, \infty)$: we shall demonstrate in lemma 55 below that this corresponds to the more standard notion of the index of compactness of a cocycle.

Lemma 54. *Let \mathcal{R} be as above. The quantities λ_i, μ_i and ν are almost everywhere constants. Further, λ_2 exists if and only if $\lambda_1 > \nu$.*

Proof. Consider the following choices of sequences:

$$\begin{aligned} f_n^k(\omega) &= \log \rho_k(A_\omega^{(n)}), \\ g_n(\omega) &= \log \|A_\omega^{(n)}\|_c. \end{aligned}$$

In [13] ρ_k is defined (written there as F_k) and it is shown that there exist numbers C_k such that for any operators $T, S \in \mathcal{B}(X)$,

$$\prod_{i=1}^k \rho_i(S \circ T) \leq C_k \prod_{i=1}^k \rho_i(S) \rho_i(F),$$

whence the existence of the desired limits is guaranteed by noting that

$$f_n^k(\omega) = \log \prod_{i=1}^k \rho_i(A_\omega^{(n)}) - \log \prod_{i=1}^{k-1} \rho_i(A_\omega^{(n)})$$

converges with an application of 28. The second statement is a consequence of the definitions $\nu = \lim_{n \rightarrow \infty} \mu_n$ and $\lambda_2 = \mu_{\min\{t: \mu_t < \mu_1\}}$. \square

There is a large degree of choice in the definitions of the Lyapunov exponent, that used here happens to be convenient for our purposes, and can be checked to agree with other notions. As an example which arises in this Banach setting, the quantity κ measures a notion of compactness of the system. Earlier versions of the main result assume, for example, that A acts on a finite dimensional space or more generally is compact after iterating forward in time enough - these would just correspond to $\kappa = -\infty$.

Lemma 55. *The index of compactness defined here is just the usual index of compactness*

$$\lim_{n \rightarrow \infty} \mu_n = \nu = \kappa = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_\omega^{(n)}\|_c \quad \text{for pointwise almost every } \omega \in \Omega.$$

Proof. Since ρ dominates the compactness seminorm up to a multiplicative constant, $\kappa \leq \mu_i$ for every $i \in \mathbb{N}$, so certainly $\kappa \leq \nu$. It remains to verify that $\kappa \geq \nu$. Suppose that $k \in \mathbb{N}$. There exists a $\delta > 0$ such that for all $\mathbb{P}(\Lambda) < \delta$, $\int_\Lambda \log \|A\| < \frac{\epsilon}{2}$. Choose N sufficiently large that

$$\mathbb{P}(\|A_\omega^{(N)}\|_c \geq e^{N(\kappa+\epsilon)}) < \frac{1}{2}\delta.$$

Choose r sufficiently large that

$$\mathbb{P}(G) = \mathbb{P}\left(A_\omega^{(N)}\mathbb{B}_X \text{ may be covered by at most } e^{rN} e^{N(\kappa+\frac{1}{2}\epsilon)}\text{-balls}\right) > 1 - \delta.$$

Set $f_n(\omega) = \log \inf\{t > 0 : A_\omega^{(n)}\mathbb{B}_X \text{ is covered by } e^{rn} t\text{-balls}\}$. In this case

$$\begin{aligned} f_{m+n}(\omega) &\leq \log \inf\{ab : A_\omega^{(n)}\mathbb{B}_X \text{ is covered by } e^{rn} a\text{-balls, and } A_{\sigma^n \omega}^{(m)}\mathbb{B}_X \text{ by } e^{rm} b\text{-balls}\} \\ &= f_n(\omega) + f_m(\sigma^n \omega) \end{aligned}$$

so that the family is subadditive, $\int_\Omega f_1 \leq \int_\Omega \log \|A_\omega\| < \infty$ and

$$\frac{1}{N} \int_\Omega f_N \leq \kappa + \frac{1}{2}\epsilon + \frac{1}{N} \int_{G^c} \log \|A_\omega^{(N)}\| < \kappa + \frac{1}{2}\epsilon + \frac{1}{2}\epsilon.$$

Thus we may apply Kingman again to obtain that $\frac{1}{n}f_n(\omega) \rightarrow C < \kappa + \epsilon$ pointwise almost everywhere. Almost surely, for sufficiently large n we may guarantee the following:

- $\|A_\omega^{(n)}\|_c < e^{n(\kappa + \epsilon)}$
- There is a $V \in \mathcal{G}_k X$ with $g_{A_\omega^{(n)}}(V) > e^{n(\mu_k - \epsilon)}$.
- $A_\omega^{(n)}\mathbb{B}_X$ is covered by at most $e^{rn} e^{n(\kappa + \epsilon)}$ -balls
- $e^{n\epsilon} > (2k)^k$.

By Lemma 41 we may choose a basis of unit vectors x_i for V with $d(x_i, \text{span}\{x_j\}_{j < i}) = 1$. Write

$$\Lambda = \left\{ \sum_{i=1}^k a_i x_i : a_i \in \left\{ 0, \pm \frac{2e^{n(\kappa + \epsilon)}}{\|A_\omega^{(n)} x_i\|}, \pm \frac{4e^{n(\kappa + \epsilon)}}{\|A_\omega^{(n)} x_i\|}, \dots \right\}, |a_i| < \frac{1}{k} \right\} \subseteq \mathbb{B}_X$$

If two members $a = \sum_i a_i x_i$ and $b = \sum_i b_i x_i$ are distinct then there is a maximal $j \leq k$ with $a_j \neq b_j$. Then

$$\|a - b\| \geq |a_j - b_j| d(A_\omega^{(n)} x_j, A_\omega^{(n)} \text{span}\{x_i : i \leq j\}) > e^{n(\kappa + \epsilon)} > \|A_\omega^{(n)}\|_c.$$

The points in $A_\omega^{(n)}\Lambda$ are then of distance at least $2e^{n(\kappa + \epsilon)} > \|A_\omega^{(n)}\|_c$ apart, and there are at least

$$|\Lambda| = \prod_{i=1}^k \frac{1}{2} \left\lfloor \frac{\|A_\omega^{(n)} x_i\|}{e^{n(\kappa + \epsilon)k}} \right\rfloor \geq (2k)^{-k} \prod_{i=1}^k e^{n(\mu_k - \epsilon) - n(\kappa + \epsilon)} = e^{nk(\mu_k - \kappa - 3\epsilon)}$$

of them. Each member of a cover of $A_\omega^{(n)}\mathbb{B}_X$ by $e^{n(\kappa + \epsilon)}$ -balls contains at most one element of Λ , so the cover has cardinality at least $e^{nk(\mu_k - \kappa - 3\epsilon)}$. On the other hand, this quantity is bounded by e^{rn} :

$$e^{rn} \geq e^{nk(\mu_k - \kappa - 4\epsilon)}.$$

Taking log and rearranging we obtain

$$k \leq \frac{r}{\mu_k - \kappa - 4\epsilon}.$$

In the case that $\mu_k > \kappa + 5\epsilon$, it must hold that $k < \frac{r}{\epsilon}$. This bound on k shows the number of μ s greater than $\kappa + 5\epsilon$ is finite, whence $\mu_k \downarrow \kappa = \nu$ as $k \rightarrow \infty$. \square

4.2.2 Properties for functions on Ω

Here we define useful properties that will help characterise the splitting. Let $(\Omega, \mathbb{P}, \sigma)$ be an ergodic invertible system on Lebesgue probability space throughout.

Definition 56. Say that a function $f \in \mathcal{M}(\Omega \rightarrow \mathbb{R})$ is Lyapunov if

$$\frac{1}{n}f(\sigma^n\omega) \rightarrow 0.$$

as $n \rightarrow \infty$.

Definition 57. Let $A \in \mathcal{SM}(\Omega \rightarrow \mathcal{B}(X))$ and $E \in \mathcal{M}(\Omega \rightarrow \mathcal{G}X)$.

- If $\|A\|$ is Lyapunov then call A tempered.
- If $A_\omega E(\omega) \subseteq E(\sigma\omega)$ then call E equivariant with respect to A . If the same holds with equality instead of mere containment then call the equivariance strict.

4.2.3 Other approaches?

There is some freedom of choice in the approach to a proof of Oseledets' theorem. Raghunathan in [24], for example, uses exterior algebra applied to \mathbb{R}^d , studying the growth rate of m -dimensional volumes being relatable to $\wedge^m X$. A single cocycle will also act on the tensor product by letting $x_1 \otimes \cdots \otimes x_m \mapsto (A_\omega x_1) \otimes \cdots \otimes (A_\omega x_m)$ for some basis $\{x_i\}$ and extending linearly, which can be shown to uniquely define $\otimes^n X$. Raghunathan uses the norm on the exterior algebra - not all the details of how this would be generalised to a separable Banach space have materialised: in the overview of [25] it is pointed out that there is no single settled definition of the tensor or exterior product.

As such, we stick to working on the Grassmannian.

Chapter 5

Constructing fast and slow spaces

5.1 Decomposing a cocycle

5.1.1 Searching for a decomposition

Having introduced the necessary background material, we now seek out the Oseledets decomposition. The first part of the proof will be in obtaining the top fast space $E_1 : \Omega \rightarrow \mathcal{G}(X)$, which is finite dimensional and seemingly presents less of a challenge than obtaining the slow space - as evidenced by the problems with the possible bad behaviour in the case where X^* is not separable.

5.1.2 Quasicompactness: does an interesting Oseledets decomposition exist?

In finite dimensions, every operator is compact, but there is no reason to expect a nontrivial Oseledets decomposition - the constant cocycle consisting of the identity on \mathbb{R}^n is an example: there is still a well defined $E_1 : \omega \mapsto \mathbb{R}^n$.

If X is infinite dimensional there is no reason to expect a finite dimensional unique E_1 - if, for example, the cocycle takes values in the isometries of X , then there is no unique finite dimensional fast space. With the convention that in the finite dimensional case we have $V_N = \{0\}$ we can write down a straightforward condition for nontriviality of an Oseledets decomposition by stipulating that the norm $\|A_\omega^{(n)}\|$ and compactness seminorm $\|A_\omega^{(n)}\|_c$ grow at different asymptotic exponential rates. The crucial condition for a cocycle to have any nontrivial decomposition into fast and slow spaces is that of quasicompactness. One may investigate the growth of each of $\|A_\omega^{(n)}\|$ and $\|A_\omega^{(n)}\|_c$: as a consequence of Lemma 55, these quantities growing at different exponential rates is equivalent to the main condition we develop for quasicompactness of a cocycle.

Definition 58. *Say A is quasicompact if it satisfies $\nu(A) < \lambda_1(A)$.*

If $\nu(A) = \lambda_1(A)$ then we can find a space of arbitrarily large finite dimension of

growth rate λ_1 , and there's no hope of finding an E_1 which is finite dimensional. It remains, then, to check that something interesting happens otherwise.

Lemma 59. *Let $\mathcal{R} = (\Omega, \mathbb{P}, \sigma, X, A)$ be a strongly measurable linear dynamical system with invertible ergodic base on a separable Banach space, with (Ω, \mathbb{P}) an atom-free Lebesgue probability space. Suppose that A is quasicompact. Then there is a unique measurable choice of fast space $E : \Omega \rightarrow \mathcal{G}_{m_1} X$ (where the multiplicity of the top space m_1 was introduced in definition 53) for which the following hold almost surely:*

- *equivariance:* $A_\omega E(\omega) = E(\sigma\omega)$
- $\lambda(x) = \lambda_1$ for every $x \in E(\omega) \setminus \{0\}$
- $\lim_{n \rightarrow \infty} \frac{1}{n} \log g(A_\omega^{(n)}, E(\omega)) = \lambda_1$.

Proof. Since X is separable, by lemma 33 $\mathcal{G}_{m_1} X$ is separable: choose some dense $\{E_i\}_{i \in \mathbb{N}} \subseteq \mathcal{G}_{m_1} X$. Let $\epsilon \in (0, \frac{1}{6}(\lambda_1 - \lambda_2))$. The sets

$$\Gamma_i = \left\{ \omega \in \Omega : g(A_{\sigma^{-n}\omega}^{(2n)}, E_i) > e^{-\epsilon} \rho_k(A_{\sigma^{-n}\omega}^{(2n)}) \right\}$$

are measurable and cover Ω , since for fixed ω , density of the E_i s and continuity of $g_{A_{\sigma^{-n}\omega}^{(2n)}}$ means that we may find an i with $g(A_{\sigma^{-n}\omega}^{(2n)}, E_i)$ as close to $\rho_k(A_{\sigma^{-n}\omega}^{(2n)})$ as we please. Then lemma 37 above provides measurable functions

$$\begin{aligned} \iota^{(n)}(\omega) &= \inf \{ i \in \mathbb{N} : g(A_{\sigma^{-n}\omega}^{(2n)}, E_i) > e^{-\epsilon} \rho_k(A_{\sigma^{-n}\omega}^{(2n)}) \}, \\ \tilde{E}^{(n)}(\omega) &:= E_{\iota^{(n)}(\omega)}, \end{aligned}$$

and the pushforward

$$E^{(n)}(\omega) = A_{\sigma^{-n}\omega}^{(n)} \tilde{E}^{(n)}(\omega)$$

is also then measurable by lemma 33.

First, we establish that this sequence is Cauchy, and therefore convergent, to a family of spaces $E \in \mathcal{M}(\Omega \rightarrow \mathcal{G}_{m_1} X)$.

For almost every $\omega \in \Omega$ the fastest m_1 dimensional growth rate

$$\begin{aligned} \frac{1}{n} \log \rho_k(A_\omega^{(n)}) &\rightarrow \lambda_1 \text{ by Kingman's theorem 23,} \\ \frac{1}{n} \log \rho_k(A_{\sigma^{-n}\omega}^{(n)}) &\rightarrow \lambda_1 \text{ again by theorem 23 applied to } \sigma^{-1}, \\ \frac{1}{2n} \log \rho_k(A_{\sigma^{-n}\omega}^{(2n)}) &\rightarrow \lambda_1 \text{ by the balanced version, theorem 27,} \end{aligned}$$

and

$$\rho_k \left(A_{\sigma^{-n}\omega}^{(n)} \right) \geq g_{A_{\sigma^{-n}\omega}^{(n)}} \left(\tilde{E}^{(n)}(\omega) \right) > e^{-\epsilon} \rho_k \left(A_{\sigma^{-n}\omega}^{(n)} \right).$$

Thus for each ω in this full measure set we may choose M_ω such that for $n \geq M_\omega$ we can usefully estimate growth under A in a few cases:

$$\begin{aligned} \|A_{\sigma^{-n}\omega}^{(n)}\|, \|A_\omega^{(n)}\| &\in (e^{n(\lambda_1-\epsilon)}, e^{n(\lambda_1+\epsilon)}), \\ g(A_{\sigma^{-n}\omega}^{(2n)}, \tilde{E}^{(n)}(\omega)) &\in (e^{2n(\lambda_1-\epsilon)}, e^{2n(\lambda_1+\epsilon)}), \\ \|A_{\sigma^{-(n+1)}\omega}\| &< e^{n\epsilon} \end{aligned}$$

and for convenience

$$e^{n\epsilon} > 4.$$

Applying the inequalities in lemma 35 then,

$$g(A_{\sigma^{-n}\omega}^{(n)}, \tilde{E}^{(n)}(\omega)) \geq \frac{g(A_{\sigma^{-n}\omega}^{(2n)}, \tilde{E}^{(n)}(\omega))}{\|A_\omega^{(n)}\|} \geq e^{n(\lambda_1-3\epsilon)}$$

and

$$g\left(A_\omega^{(n)}, A_{\sigma^{-n}\omega}^{(n)} \tilde{E}^{(n)}(\omega)\right) = g(A_\omega^{(n)}, E^{(n)}(\omega)) \geq \frac{g(A_{\sigma^{-n}\omega}^{(2n)}, \tilde{E}^{(n)}(\omega))}{\|A_{\sigma^{-n}\omega}^{(n)}\|} \geq e^{n(\lambda_1-3\epsilon)}.$$

In addition, $A_{\sigma^{-(n+1)}\omega} \tilde{E}^{(n+1)}(\omega)$ is also guaranteed to have fast growth under $A_{\sigma^{-n}\omega}^{(n)}$:

$$\begin{aligned} g\left(A_{\sigma^{-n}\omega}^{(n)}, A_{\sigma^{-(n+1)}\omega} \tilde{E}^{(n+1)}(\omega)\right) &\geq \frac{g(A_{\sigma^{-(n+1)}\omega}^{(n+1)}, \tilde{E}^{(n+1)}(\omega))}{\|A_{\sigma^{-(n+1)}\omega}\|} \\ &> e^{(n+1)(\lambda_1-3\epsilon)} e^{-n\epsilon} \geq e^{n(\lambda_1-4\epsilon)}. \end{aligned}$$

$E^{(n)}(\omega)$ consists then of the image of vectors that were fast from time $-n$ to 0, and will grow fast from time 0 to n . Then by lemma 40 with $\Theta = e^{n(\lambda_1-4\epsilon)}$ we have

$$\begin{aligned} d(E^{(n)}(\omega), E^{(n+1)}(\omega)) &= d\left(A_{\sigma^{-n}\omega}^{(n)} \tilde{E}^{(n)}(\omega), A_{\sigma^{-n}\omega}^{(n)} (A_{\sigma^{-(n+1)}\omega} \tilde{E}^{(n+1)}(\omega))\right) \\ &< 4 \frac{e^{n(\lambda_2+\epsilon)}}{e^{n(\lambda_1-4\epsilon)}} < e^{-n(\lambda_1-\lambda_2-6\epsilon)}. \end{aligned}$$

Thus $E^{(n)}(\omega)$ is Cauchy and convergent since $\mathcal{G}_{m_1} X$ is complete, say to $E(\omega)$.

To prove equivariance, observe that for $n \geq \max\{M_\omega, M_{\sigma\omega}\}$, we find $\tilde{E}^{(n+1)}(\sigma\omega)$ is uniformly fast growing under $A_{\sigma^{-n}\omega}^{(n)}$:

$$g(A_{\sigma^{-n}\omega}^{(n)}, \tilde{E}^{(n+1)}(\sigma\omega)) \geq \frac{g(A_{\sigma^{-n}\omega}^{(n+1)}, \tilde{E}^{n+1}(\sigma\omega))}{\|A_\omega\|} \geq e^{(n+1)(\lambda_1-\epsilon)} e^{-n\epsilon} > e^{n(\lambda_1-2\epsilon)}.$$

Then, once again, by closeness of images of fast spaces,

$$d(A_{\sigma^{-n}\omega}^{(n)} \tilde{E}^{(n+1)}(\sigma\omega), E^{(n)}(\omega)) < e^{-n(\lambda_1 - \lambda_2 - 3\epsilon)},$$

whence

$$\begin{aligned} d(E^{(n+1)}(\sigma\omega), A_\omega E^{(n)}(\omega)) &= d(A_{\sigma^{-(n+1)}\omega}^{(n+1)} \tilde{E}^{(n+1)}(\sigma\omega), A_\omega E^{(n)}(\omega)) \\ &\leq e^{n\epsilon} d(A_{\sigma^{-(n+1)}\omega}^{(n)} \tilde{E}^{(n+1)}(\sigma\omega), E^{(n)}(\omega)) \\ &< e^{-n(\lambda_1 - \lambda_2 - 4\epsilon)}, \end{aligned}$$

so that $A_\omega E(\omega) = \lim_{n \rightarrow \infty} A_\omega E^{(n)}(\omega) = E(\sigma\omega)$.

To check that $E(\omega)$ is uniformly fast growing, choose $x \in \mathbb{S}_{E(\omega)}$. Then since for $n \geq M_\omega$ we have $d(E^{(n)}(\omega), E(\omega)) < e^{-n(\lambda_1 - \lambda_2 - \epsilon)}$, for each such n we may choose an $x_n \in \mathbb{S}_{E^{(n)}(\omega)}$ with $\|x - x_n\| < e^{-n(\lambda_1 - \lambda_2 - \epsilon)}$. Since $x_n \in E^{(n)}(\omega)$ and $g(A_\omega^{(n)}, E^{(n)}(\omega)) \geq e^{n(\lambda_1 - 3\epsilon)}$ it then follows that,

$$\begin{aligned} \|A_\omega^{(n)} x\| &\geq \|A_\omega^{(n)} x_n\| - \|A_\omega^{(n)}(x - x_n)\| \\ &\geq e^{n(\lambda_1 - 3\epsilon)} - e^{-n(\lambda_1 - \lambda_2 - \epsilon)} \|A_\omega^{(n)}\| \\ &\implies \|A_\omega^{(n)} x\| \geq \frac{1}{2} e^{n(\lambda_1 - \epsilon)}. \end{aligned}$$

Thus we may conclude that as well as being equivariant, $E(\omega)$ is fast growing for all sufficiently large n ; since the choice of x was arbitrary the growth is uniform:

$$g(A_\omega^{(n)}, E(\omega)) > e^{n(\lambda_1 - \epsilon)}.$$

On the other hand, for n sufficiently large we also have

$$g(A_\omega^{(n)}, E(\omega)) \leq \rho_k(A_\omega^{(m)}) < e^{n(\lambda_1 + \epsilon)},$$

whence $\frac{1}{n} \log g(A_\omega^{(n)}, E(\omega)) \rightarrow \lambda_1$. Finally, we check uniqueness: Suppose that $E(\omega)$ and $E'(\omega)$ are both equivariant and fast, so that for every ω there is some N such that for $n \geq N$,

$$g_{A_\omega^{(n)}}(E(\omega)), g_{A_\omega^{(n)}}(E'(\omega)) > e^{n(\lambda_1 - \epsilon)}.$$

Define $\varphi \in \mathcal{M}(\Omega \rightarrow [0, 1])$ by $\varphi(\omega) = d(E(\omega), E'(\omega))$. Applying lemma 40, for almost every ω we have

$$\begin{aligned} \varphi(\sigma^n \omega) &= d(E(\sigma^n \omega), E'(\sigma^n \omega)) = d(A_\omega^{(n)} E(\omega), A_\omega^{(n)} E'(\omega)) \\ &< 4 \frac{\rho_{k+1} A_\omega^{(n)}}{e^{n(\lambda_1 - \epsilon)}} \rightarrow 0. \end{aligned}$$

φ tends to zero along all orbits. Therefore the sets $\{\varphi(\omega) > \epsilon\}$ all have measure zero,

whence φ vanishes almost everywhere and $E = E'$. \square

We have found the top fast space.

5.1.3 Finding a slow space: a splitting lemma

This order of events seems natural in this setting, where working with finite dimensional spaces is easier than finite codimensional spaces. We need invertibility of σ to guarantee such a space in general, whereas the space constructed in the next subsection will exist even if σ fails to be invertible, but nonetheless this slow space seems harder to construct.

The space $\mathcal{G}^k X$ may fail to be separable, so finding approximants to the actual true space in the manner of the previous systems seems challenging. Instead, we can exploit separability of X to write down a nonlinear but measurable map $\Omega \times X \rightarrow X$ which decomposes any element of the state space as an element of the fast space, plus an element that grows at a lower asymptotic rate than anything in the fast space. The limit of this map may then be verified to be linear, and in particular the projection onto the fast space along the slow space.

Lemma 60. *Let $\mathcal{R} = (\Omega, \sigma, X, A)$ be a quasicompact, semi-invertible linear dynamical system. Then there exists a forward-equivariant decomposition $X = E(\omega) \oplus V(\omega)$, where $V : \Omega \rightarrow \mathcal{G}^k X$ and the corresponding projection is a strongly measurable $\Pi : \Omega \rightarrow \mathcal{B}(X)$. V is a slow growing space: $\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_\omega^{(n)}|_{V(\omega)}\| \rightarrow \lambda_2$ almost surely. Finally, Π_ω is tempered.*

Proof. By lemma 44 we may choose a measurable family of bases

$$(b_i)_{i=1}^k : \mathcal{G}_{m_1} X \rightarrow \mathbb{S}_X^k.$$

Write $v_i(\omega) = b_i(E(\omega))$ which is itself then measurable. Let $\{(q_{ij})_{i=1}^k\}_{j \in \mathbb{N}}$ be dense in \mathbb{R}^k . Let $T_j \in \mathcal{M}(\Omega \times X \rightarrow X)$ be defined by

$$T_j(\omega, x) = x - \sum_{i=1}^k q_{ij} v_i(\omega),$$

so that for each $\omega \in \Omega$, the collection $\{T(\omega, \cdot)\}$ is dense in translations of X by elements of $E(\omega)$. Let $\epsilon < \frac{1}{2}(\lambda_1 - \lambda_2)$. Define $\iota_n \in \mathcal{M}(X \setminus E(\omega) \rightarrow \mathbb{N})$ for $n \in \mathbb{N}$ by

$$\iota_n(\omega, x) = \inf\{j \in \mathbb{N} : \|A_\omega^{(n)} T_j x\| \leq e^\epsilon \rho_{k+1}(A_\omega^{(n)}) \|T_j x\|\}.$$

A sequence of nonlinear maps $\Pi^{(n)} \in \mathcal{M}(X \rightarrow X)$ may then be defined piecewise by

$$\Pi_{\omega}^{(n)}(x) = \begin{cases} 0, & \text{if } x \in E(\omega), \\ T_{\iota_n(\omega, x)}x & \text{otherwise,} \end{cases} \quad \text{and}$$

$$P_{\omega}^{(n)}(x) = x - \Pi_{\omega}^{(n)}(x).$$

To see that ι is finite on $X \setminus E(\omega)$, first note that the set

$$\{T_j x : j \in \mathbb{N}\}$$

is dense in the set $S = E(\omega) + x$. The set

$$U = \{y \in E(\omega) \oplus \text{span}\{x\} : \|A_{\omega}^{(n)}y\| < e^{\frac{1}{2}\epsilon} \rho_{k+1}(A_{\omega}^{(n)})\|y\|\}$$

is nonempty by the definition of ρ_{k+1} , open and scale invariant - for every $\theta \neq 0$ we have $\theta U = U$. Write

$$U' = (E(\omega) + x) \cap U,$$

which is open in S , and nonempty, since any direction in $E(\omega) \oplus \text{span}\{x\}$ with a nonzero x component can be found in $E(\omega) + x$. Since $\{T_j x\}_j$ is dense in this fibre, an $\iota_n(\omega, x)$ may be found in finite time. To see that ι_n is measurable, note that

$$\iota_n^{-1}\{1, \dots, m\} = \bigcup_{i=1}^m \{(\omega, x) : \|A_{\omega}^{(n)}T_j x\| \leq e^{\epsilon} \rho_{k+1}(A_{\omega}^{(n)})\|A_{\omega}^{(n)}T_j x\|\}$$

Let $x \in \mathbb{S}_X$. Immediately from the definition, $P_{\omega}^{(n)}x \in E(\omega)$. By convergence of the $\frac{1}{n} \log \rho_k$ s, for all n greater than or equal to some N_{ω} ,

$$\begin{aligned} \|A_{\omega}^{(n+1)}\Pi_{\omega}^{(n)}(x)\| &\leq e^{(n+1)(\lambda_2+\epsilon)}\|\Pi_{\omega}^{(n)}x\|, \\ \|A_{\omega}^{(n+1)}\Pi_{\omega}^{(n+1)}(x)\| &\leq e^{(n+1)(\lambda_2+\epsilon)}\|\Pi_{\omega}^{(n)}x\|, \\ \|A_{\sigma^n \omega}\| &< e^{n\epsilon} \quad \text{and} \\ g_{A_{\omega}^{(n)}}(E(\omega)) &\geq e^{n(\lambda_1-\epsilon)}. \end{aligned}$$

In addition there is an easy bound independent of $x \in \mathbb{S}_X$ on $\Pi_{\omega}^{(n)}$:

$$\begin{aligned} \|\Pi_{\omega}^{(n)}x\| &\leq \|x\| + \|P_{\omega}^{(n)}x\| \\ &\leq 1 + e^{-n(\lambda_1-\epsilon)}\|A_{\omega}^{(n)}P_{\omega}^{(n)}x\| \\ &\leq 1 + e^{-n(\lambda_1-\epsilon)}(\|A_{\omega}^{(n)}x\| + \|A_{\omega}^{(n)}\Pi_{\omega}^{(n)}x\|) \\ &\leq 1 + e^{-n(\lambda_1-\epsilon)}(e^{n(\lambda_1+\epsilon)} + e^{n(\lambda_2+\epsilon)}\|\Pi_{\omega}^{(n)}x\|), \end{aligned}$$

which rearranged yields

$$\|\Pi_\omega^{(n)}x\| \leq \frac{1 + e^{2n\epsilon}}{1 - e^{-n(\lambda_1 - \lambda_2 - 2\epsilon)}} \leq e^{3n\epsilon}.$$

Consider differences between successive approximate slow components:

$$\begin{aligned} \|\Pi_\omega^{(n+1)}(x) - \Pi_\omega^{(n)}(x)\| &= \|P_\omega^{(n+1)}(x) - P_\omega^{(n)}(x)\| \\ &\leq \frac{\|A_\omega^{(n+1)}(P_\omega^{(n+1)}x - P_\omega^{(n)}(x))\|}{g_{A_\omega^{(n)}}(E(\omega))} \\ &= \frac{\|A_\omega^{(n+1)}(\Pi_\omega^{(n+1)}x - \Pi_\omega^{(n)}(x))\|}{g_{A_\omega^{(n)}}(E(\omega))} \\ &\leq e^{-(n+1)(\lambda_1 - \epsilon)} (\|A_\omega^{(n+1)}\Pi_\omega^{(n)}(x)\| + \|A_\omega^{(n+1)}\Pi_\omega^{(n+1)}(x)\|) \\ &\leq e^{-(n+1)(\lambda_1 - \epsilon)} (\|A_{\sigma^n\omega}\| \|A_\omega^{(n)}\Pi_\omega^{(n)}(x)\| + e^{n(\lambda_2 + \epsilon)} \|\Pi_\omega^{(n+1)}(x)\|) \\ &\leq e^{-(n+1)(\lambda_1 - \lambda_2 - 3\epsilon)} (e^{3n\epsilon} + e^{3(n+1)\epsilon}) \\ &\leq e^{-(n+1)(\lambda_1 - \lambda_2 - 7\epsilon)}, \end{aligned}$$

whence gaps between subsequent points decay exponentially, so that $(\Pi_\omega^{(n)}(x))_{n \in \mathbb{N}}$ forms a Cauchy and thus convergent sequence. Set

$$\Pi_\omega(x) = \lim_{n \rightarrow \infty} \Pi_\omega^{(n)}(x) \quad \text{and} \quad P_\omega = 1 - \Pi_\omega.$$

Note that $\Pi_\omega \mathbb{S}_X$ is then bounded, since

$$\begin{aligned} \|\Pi_\omega x\| &= \lim_{m \rightarrow \infty} \|\Pi_\omega^{(m)}x\| \\ &\leq \|\Pi_\omega^{(N)}x\| + \lim_{m \rightarrow \infty} \|\Pi_\omega^{(N)}x - \Pi_\omega^{(m)}(x)\| \\ &\leq e^{3N\epsilon} + \frac{e^{-N(\lambda_1 - \lambda_2 - 7\epsilon)}}{1 - e^{-(\lambda_1 - \lambda_2 - 7\epsilon)}}, \end{aligned}$$

the final line being independent of choice of x . Not only then is $\Pi_\omega^{(n)}x$ convergent, but we have the estimate

$$\|\Pi_\omega^{(n)}x - \Pi_\omega x\| \leq \sum_{i=1}^{\infty} \|\Pi_\omega^{(n+i)}x - \Pi_\omega^{(n+i-1)}x\| < \frac{e^{-n(\lambda_1 - \lambda_2 - \epsilon)}}{1 - e^{-\lambda_1 - \lambda_2 - \epsilon}}.$$

The map Π_ω is linear: to see this let $b, c \in X$ and $t \in \mathbb{R}$, and set

$$d := \Pi_\omega^{(n)}(b + tc) - \Pi_\omega^{(n)}b - t\Pi_\omega^{(n)}c.$$

Certainly $d \in E(\omega)$, since

$$\begin{aligned} d &= \Pi_\omega^{(n)}(b + tc) - (b + tc) + b - \Pi_\omega^{(n)}b + tc - t\Pi_\omega^{(n)}c \\ &= -P_\omega^{(n)}(b + tc) + P_\omega^{(n)}b + tP_\omega^{(n)}c \in E(\omega). \end{aligned}$$

Then applying the estimate for $\Pi_\omega^{(n)}$

$$\begin{aligned} \|d\| &\leq e^{-n(\lambda_1 - \epsilon)} \|A_\omega^{(n)}d\| \\ &\leq e^{-n(\lambda_1 - \epsilon)} (\|A_\omega^{(n)}\Pi_\omega^{(n)}(b + tc)\| + \|A_\omega^{(n)}\Pi_\omega^{(n)}(b)\| + \|tA_\omega^{(n)}\Pi_\omega^{(n)}c\|) \\ &< e^{-n(\lambda_1 - \lambda_2 - 2\epsilon)} (\|\Pi_\omega^{(n)}(b + tc)\| + \|\Pi_\omega^{(n)}b\| + \|t\Pi_\omega^{(n)}c\|) \\ &< e^{-n(\lambda_1 - \lambda_2 - 2\epsilon)} (\|\Pi_\omega(b + tc)\| + \|\Pi_\omega b\| + \|t\Pi_\omega c\| + \\ &\quad (\|b + tc\| + \|b\| + \|tc\|) \frac{e^{-n(\lambda_1 - \lambda_2 - \epsilon)}}{1 - e^{-\lambda_1 - \lambda_2 - \epsilon}}) \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

Thus $d = 0$ and $\Pi_\omega \in \mathcal{B}(X)$. By construction $\Pi_\omega E(\omega) = 0$ and $P_\omega X = E(\omega)$ since for any $x \in X$ and $n \in \mathbb{N}$ we have $x - \Pi_\omega^{(n)}x \in E(\omega)$. The map Π_ω is then idempotent, since for any $x \in X$

$$\Pi_\omega^2 x - \Pi_\omega x = \Pi_\omega \circ P_\omega x \in \Pi_\omega E(\omega) = \{0\},$$

and is thus a projection. Set $V(\omega) = \Pi_\omega X$ so that $X = V(\omega) \oplus E(\omega)$. For all $n \geq N_\omega$, and any $x \in V_\omega(x)$, because of the exponential rate of convergence there is a sequence of approximants

$$\|x_n - x\| < C_\omega e^{-n(\lambda_1 - \lambda_2 - 7\epsilon)}$$

with

$$\|A_\omega^{(n)}x_n\| \in [0, e^{n(\lambda_2 + \epsilon)}).$$

Therefore,

$$\|A_\omega^{(n)}x\| < \|A_\omega^{(n)}\| \|x - x_n\| + \|A_\omega^{(n)}x_n\| < e^{n(\lambda_2 + 8\epsilon)},$$

whence

$$\|A_\omega^{(n)}|_{V(\omega)}\| < e^{n(\lambda_2 + 8\epsilon)}$$

for $n \geq N_\omega$. By the definition of λ_2 , it is possible to choose $Y_n \in \mathcal{G}_{m_1+1}(X)$ with $\frac{1}{n}g_{A_\omega^{(n)}}(Y_n) \rightarrow \lambda_2$, which means by dimension counting that there is always some sequence of vectors $x_n \in \mathbb{S}_{V(\omega) \cap Y_n}$ with $\frac{1}{n}\|A_\omega^{(n)}x_n\| \rightarrow \lambda_2$ and so

$$\frac{1}{n} \log \|A_\omega^{(n)}|_{V(\omega)}\| \rightarrow \lambda_2.$$

Further, the map $\omega \mapsto \Pi_\omega$ is strongly measurable, since for each $x \in X$ the map $\omega \mapsto \Pi_\omega x$ is the limit of a sequence of measurable functions.

To see that $V(\omega)$ is equivariant it is sufficient to show that $P_{\sigma\omega} \circ A_\omega x = 0$ for all

$x \in V(\omega)$. If this were not the case, then $A_\omega x$ would have a nonzero component in $E(\sigma\omega)$. From this it would follow that $\lambda_\omega(x) = \lambda_1$ which would contradict the fact that $\|A_\omega^{(n)}|_{V(\omega)}\| < e^{n(\lambda_2+8\epsilon)}$ for sufficiently large n .

We next check temperedness of the projections. Since Π_ω is bounded pointwise we may choose an $A \subseteq \Omega$ of positive measure on which $\|\Pi_\omega\|$ is at most some $M > 0$. Then define a new cocycle A' by

$$A'_\omega = \begin{cases} A_\omega \circ \Pi_\omega & \text{if } \omega \in A, \\ A & \text{otherwise.} \end{cases}$$

Then A'_ω is forward-integrable since

$$\int_{\Omega} \log^+ \|A'_\omega\| \leq \int_{\Omega} \log^+ \|A\| + \int_A \log^+ M < \infty.$$

Write λ' for the asymptotic growth rates associated with A' . Since A has positive measure, the hitting time of a point ω where $A'_\omega = A_\omega \circ \Pi_\omega$ is almost surely finite, say n . At time n one may apply the equivariance condition for the projection Π_ω to conclude there is almost surely an n such that $A'_\omega^{(n)} E(\omega) = \{0\}$, and so

$$A'_\omega^{(n)} X \subseteq V(\sigma^n \omega).$$

We may then conclude that for each $x \in X$, $\lambda'_\omega(x) \leq \lambda_2$ and $\lambda'_1 \leq \lambda_2$. Applying lemma 26 to the subadditive families $g_n = \log \|A_\omega^{(n)}\|$ and $g'_n = \log \|A'_\omega^{(n)}\|$ there is an N_1 such that for $n \geq N_1$,

$$\|A'_{n \rightarrow 2n}\| < e^{n(\lambda_2+\epsilon)} \text{ and } \|A_{n \rightarrow 2n}\| > e^{n(\lambda_1-\epsilon)}.$$

Clearly $A'_{n \rightarrow 2n} \neq A_{n \rightarrow 2n}$ since the latter has a greater norm. Therefore there must be some j such that

$$A'_{n \rightarrow 2n} = A'_{n \rightarrow j} \circ \Pi_{\sigma^j \omega} \circ A'_{j \rightarrow 2n} = A_{n \rightarrow 2n} \circ \Pi_{\sigma^n \omega}.$$

In addition, there exists an N_2 such that for $n \geq N_2$,

$$g_{A_{\sigma^n \omega}}^{(n)}(E(\sigma^n \omega)) \in (e^{n(\lambda_1-\epsilon)}, e^{n(\lambda_1+\epsilon)}).$$

As a final condition, because of the asymptotic growth of E , there exists some $N_3 \in \mathbb{N}$ such that for all $n \geq N_3$, $g_{A_\omega}^{(n)}(E(\omega)) \in (e^{n(\lambda_1-\epsilon)}, e^{n(\lambda_1+\epsilon)})$. Let $x \in \mathbb{S}_X$. Putting these

together, for all $n \geq \max\{N_1, N_2, N_3, \frac{1}{\epsilon}\}$,

$$\begin{aligned}
\|P_{\sigma^n \omega} x\| &\leq e^{-n(\lambda_1 - \epsilon)} \|A_{\sigma^n \omega}^{(n)} P_{\sigma^n \omega} x\| \\
&\leq e^{-n(\lambda_1 - \epsilon)} (\|A_{\sigma^n \omega}^{(n)} x\| + \|A_{\sigma^n \omega}^{(n)} \Pi_{\sigma^n \omega} x\|) \\
&\leq e^{-n(\lambda_1 - \epsilon)} (e^{n(\lambda_1 + \epsilon)} \|x\| + \|A'_{n \rightarrow 2n \omega} x\|) \\
&\leq e^{-n(\lambda_1 - \epsilon)} (e^{n(\lambda_1 + \epsilon)} + \|A'_{n \rightarrow 2n}\|) \|x\| \\
&\leq e^{2n\epsilon} + e^{-n(\lambda_1 - \lambda_2 - 2\epsilon)} \leq e^{3n\epsilon}.
\end{aligned}$$

ϵ was arbitrary and the norms of Π and P differ by at most 1 so $\frac{1}{n} \log \|P_{\sigma^n \omega}\|, \frac{1}{n} \log \|\Pi_{\sigma^n \omega}\| \rightarrow 0$ as required. \square

Corollary 61. $V(\omega) = \tilde{V}(\omega) = \{x \in X : \lambda_\omega(x) \leq \lambda_2\}$ almost everywhere.

Proof. That $\frac{1}{n} \log \|A_\omega^{(n)}|_{V(\omega)}\| \rightarrow \lambda_2$ establishes the fact that $V(\omega) \subseteq \tilde{V}(\omega)$. Conversely, any $x \in \tilde{V}(\omega) \setminus V(\omega)$ would have $P_\omega(x) \neq 0$ so that $\lambda_\omega(x) = \lambda_1$, contradicting the definition of \tilde{V} . \square

The decomposition $E \oplus V$ is referred to as a splitting, since any x in X can now be decomposed into a part that grows at least at some exponential rate, and a part that grows at most another exponential rate.

Growth of $E(\omega)$ can be bounded independently of $x \in \mathbb{S}_{E(\omega)}$, but the same cannot be said for $V(\omega)$ since the unit ball isn't compact.

Chapter 6

Oseledets' theorem

6.1 History and overview

6.1.1 The fully invertible result

Oseledets, motivated by the exponential divergence of nearby points in the study of dynamical systems, proved the following remarkable result in 1965 [22]. We will include the discrete time case for this result, but henceforth consider discrete time results for simplicity.

Theorem 62 (Fully invertible Oseledets' theorem on \mathbb{R}^d). *Let Ω be a Lebesgue probability space and $\sigma : \Omega \times \mathbb{G} \rightarrow \Omega$ be an invertible ergodic action for $\mathbb{G} = \mathbb{R}$ or \mathbb{Z} . Let the measurable map into the invertible matrices*

$$A \in \mathcal{M}(\Omega \times \mathbb{G} \rightarrow GL_d(\mathbb{R}))$$

satisfy the cocycle condition

$$A_\omega^{(s+t)} = A_{\sigma_t \omega}^{(s)} A_\omega^{(t)} \quad \text{for all } s, t \in \mathbb{G}$$

and

$$\omega \mapsto \sup_{-1 \leq t \leq 1} \log^+ \|A_\omega^{(t)}\| \in L^1 \Omega.$$

Then there are measurable numbers $\lambda_i(\omega), i \in \{1, \dots, r_\omega\}$ and a direct sum decomposition of equivariant measurable subspaces $\mathbb{R}^d = \bigoplus_i E_i(\omega)$ such that the space $E_i(\omega)$ grows asymptotically uniformly at rate λ_i as $n \rightarrow \pm\infty$, that is,

$$\frac{1}{|n|} \log g_{A_\omega^{(n)}}(E_i(\omega)) \quad \text{and} \quad \frac{1}{|n|} \log \|A_\omega^{(n)}|_{E_i(\omega)}\| \rightarrow \pm\lambda_i$$

pointwise almost everywhere as $n \rightarrow \pm\infty$. Finally, the projections $\Pi_{i,\omega}$ onto $\bigoplus_{j>i} E_j(\omega)$ parallel to $\bigoplus_{j \leq i} E_j(\omega)$ are tempered.

Note that the uniform growth limits mean that

$$x \in E_i(\omega) \setminus \{0\} \implies \frac{1}{n} \log \|A_\omega^{(n)} x\| \rightarrow \lambda_i.$$

In the discrete case that $t \in \mathbb{G} = \mathbb{Z}$ we focus on where A is generated by $A^{(1)}$, demanding that $\sup_{-1 \leq t \leq 1} \log^+ \|A_\omega^{(t)}\| \in L^1 \Omega$ just reduces to the conditions of A being both forward and backward integrable when $t \in \mathbb{Z}$. This fully invertible result is so-called because of the assumptions: σ is invertible, and A is not only forward integrable, but it has to be backward integrable (which naturally implies pointwise invertibility). At the time of his proof, the Kingman theorem had not been proven. From the work of Furstenberg and Kesten [10] it was known that λ_1 existed, but this theorem is an important cornerstone of the study of hyperbolic dynamical systems.

6.1.2 The non-invertible case: weaker premises and conclusions

At the other extreme, dropping these invertibility assumptions leads to much weaker information about the space necessarily being available. The following flag decomposition is proven in the 1978 work of Raghunathan [24]:

Theorem 63. *Let Ω be a Lebesgue probability space and $\sigma : \Omega \rightarrow \Omega$ be a (not necessarily invertible) measure preserving transformation. Let $A \in \mathcal{M}(\Omega \rightarrow \mathbb{R}^{d \times d})$ satisfy $\log^+ \|A\| \in L^1(\Omega)$. Then there are measurable numbers $\lambda_i(\omega), i \in \{1, \dots, r_\omega\}$ and a flag of equivariant measurable subspaces $\mathbb{R}^d = V_1(\omega) > V_2(\omega) > \dots > V_{r_\omega}(\omega)$ such that for almost every ω , we have that for $x \in V_i(\omega) \setminus V_{i+1}(\omega)$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_\omega^{(n)} x\| = \lambda_i(\omega).$$

The moniker flag for a descending chain of vector spaces is based upon the observation that a flag is, as we all know, a 2 dimensional surface, containing a 1 dimensional flagpole, containing a 0 dimensional base. Raghunathan [24] uses singular value decomposition as a black box, with approximants to spaces being made up of appropriate spans of singular vectors. He describes a strengthening of Oseledets' result with both these assumptions dropped - he does obtain an orthogonal splitting of the space, which can always be obtained from a flag, but such a splitting isn't the unique equivariant spaces we may have hoped for. This can be realised by bearing in mind that in general certainly E_1 and E_2 needn't be orthogonal. The result is an extension to non-invertible systems of some, but not all, of Oseledets' original conclusions.

6.1.3 Continuous cases and generalisations to infinite dimensions

These results were subsequently adapted to various settings. The work of Ruelle [27] is an example of a smooth result, where (Ω, σ) is a $C^{1+\epsilon}$ diffeomorphism on a compact manifold. He proves a non-invertible flag result and uses it to deduce the fully invertible result to find stable manifolds of points that asymptotically coalesce.

The most general separable Banach result before the semi-invertible theorems in the next section is the work of Lian and Lu [20]. Their monograph obtains the decomposition result that weakens pointwise invertibility of A_ω to pointwise injectivity of a strongly measurable cocycle in a separable Banach space, with invertible ergodic base.

Another condition that one may drop is separability of the state space. However, the notion of measurability of the cocycle which is demanded is much stronger. Mañé [21] proves the invertible version of the theorem with σ a homeomorphism, with the invertibility of A_ω being weakened to injectivity, for compact operators that satisfy a notion of p -continuity. Thiullen generalises the work of Mañé in [28] to quasicompact operators with the same p -continuity condition. A recent proof in the non-separable setting with an emphasis on volume growth is the non-invertible result of Blumenthal in [3], where a flag is obtained for uniformly measurable cocycles. Each of these non-separable results relies on a much stronger Lusin type regularity of the system, insisting that the cocycle either be continuous or constant piecewise on a countable partition: a condition that in finite dimensions again becomes equivalent to measurability.

All of the results stated thus far have either required invertibility of the base and injectivity of the cocycle to obtain a direct sum decomposition, or are fully non-invertible results. It turns out, however, that an injectivity condition isn't necessary to obtain a decomposition.

6.1.4 A further refinement: semi-invertibility

Froyland, Lloyd and Quas [9] demonstrated that as long as σ is invertible the space may still be written as a sum of fast and slow spaces with no injectivity requirements:

Theorem 64 (Semi-invertible Oseledets' theorem on \mathbb{R}^d). *Let Ω be a Lebesgue probability space and $\sigma : \Omega \rightarrow \Omega$ be an invertible ergodic system. Let the measurable matrix cocycle*

$$A \in \mathcal{M}(\Omega \rightarrow \mathbb{R}^{d \times d})$$

be forward integrable. Then there are measurable numbers $\lambda_i(\omega), i \in \{1, \dots, r_\omega\}$ and a direct sum decomposition of invariant measurable subspaces $\mathbb{R}^d = \bigoplus_i E_i(\omega)$ such that for $x \in \mathbb{S}_{E_i(\omega)}$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_\omega^{(n)} x\| = \lambda_i(\omega).$$

This limit converges uniformly in x . Finally, the projections $\Pi_{i,\omega}$ onto $\bigoplus_{j>i} E_j(\omega)$ parallel to $\bigoplus_{j\leq i} E_j(\omega)$ are tempered.

Nearly all the conclusions hold comparing to the fully invertible result! The only loss is the backward estimate, since A_ω^{-1} needn't even exist almost everywhere. The associated paper dealt with the finite dimensional case, but in [12] González-Tokman and Quas presented a generalisation of this semi-invertible result to a separable Banach space.

6.1.5 A slow space measurability disaster

Actually the claim that the slow space is measurable was made in that paper, but after Horan [14], the infinite-codimensional slow spaces may fail to be measurable, which we demonstrate in this section. González-Tokman and Quas give a subsequent shorter proof of the semi-invertible decomposition explicitly in the separable dual case, using intuitive notions of volume growth that this work partially builds upon [13]. The argument to obtain spaces for $n > 1$ relies on a notion of fibres over a measurable base space, examining cocycles whose domain varies depending on ω , a geometrically compelling argument for which not all details are supplied. For X with a separable dual, the appendix of [12] explains that it follows that $\mathcal{G}^k X$ must be separable, whence problems such as in the following lemma won't arise, and a separable dual is a sufficient condition for measurability of slow spaces.

The following shows that generally the slow space can fail to be measurable as a map from Ω to the Grassmannian of finite-codimensional spaces. Our main result instead establishes measurability of the associated projections.

Lemma 65. *Let $\Omega = \{\pm 1\}^{\mathbb{Z}}$ together with the shift $\sigma(\omega)_n = \omega_{n+1}$ with the product measure of fair coin tosses as described previously.*

Let $X = \ell^1(\mathbb{N}_0)$ and consider the cocycle given by $A_\omega(x) = (\sum_{i=1}^{\infty} \omega_i x_i, 0, 0, \dots)$. Then

- *A is strongly measurable and forward integrable.*
- *there is an equivariant decomposition with a constant fast space*

$$E_1(\omega) = \text{span}\{(1, 0, \dots)\} \quad \text{for all } \omega \in \Omega$$

and a slow space

$$V_2(\omega) = \{x \in X : \sum_{i=0}^{\infty} \omega_i x_i = 0\}$$

corresponding to exponents $\lambda_1 = 0$ and $\lambda_2 = -\infty$.

- *$V_2(\omega)$ is not measurable with respect to the Lebesgue sigma algebra on Ω .*

Proof. Write $p_i : (x_i)_{i \in I} \mapsto x_i$ for the projection onto the i th coordinate in a product space. To check strong measurability, fix an $x \in X$. We need to verify that the map

$$\varphi : \Omega \ni \omega \mapsto A_\omega x = \left(\sum_{i=0}^{\infty} \omega_i x_i, 0, \dots \right) \in X$$

is measurable. This function can be written as a single sum times a constant vector, it's sufficient to notice that $p_0 \circ \varphi(\omega) = \sum_{i=0}^{\infty} x_i \omega_i$ is measurable, as a pointwise limit of the maps $\omega \mapsto \sum_{i=0}^n x_i \omega_i$. The equivariant top fast space is straightforward:

$$E_\omega = \text{span}\{(1, 0, 0, \dots)\} \quad \text{for all } \omega \in \Omega$$

is the equivariant top space corresponding to $\lambda_1 = 0$, since for every ω we have $A_\omega(1, 0, 0, \dots) = (\pm 1, 0, 0, \dots)$. We can see that

$$V_2(\omega) = \left\{ x \in X : \sum_{i=0}^{\infty} x_i \omega_i = 0 \right\}$$

is the kernel of the functional $p_0 \circ A_\omega$, so it is an equivariant space of codimension 1 corresponding to $\lambda_2 = -\infty$. Let $\omega, \omega' \in \Omega$ so that $\omega \neq \pm \omega'$. Then there are $i \neq j$ with $\omega_i = \omega'_i$ and $\omega_j = -\omega'_j$. Then the $x \in \ell^1(\mathbb{N}_0)$ which has $x_i = 1, x_j = -1$ and $x_l = 0$ otherwise will be in exactly one of V_ω or $V_{\omega'}$ but has distance at least 1 from the other. The family $\{V_\omega : \omega \in \Omega\} \subseteq \mathcal{G}^1 X$ is then a discrete uncountable set. Taking open balls around each V_ω , any subset $S \subseteq \Omega$ with $S = -S$ is a preimage of some measurable subset of $\mathcal{G}^1 X$. any sigma algebra on Ω making V_2 measurable would contain the collection

$$\{\pm S = \{\pm \omega : \omega \in S\} : S \subseteq \Omega\},$$

whence V isn't measurable, since choosing a non-measurable set $S \subseteq \{\omega \in \Omega : \omega_0 = 1\}$, the set $\pm S$ is also not measurable, but would have to be were V_2 measurable. \square

6.1.6 Statement of main result

The main result, an extension of the semi-invertible result to separable Banach spaces, may now be stated:

Theorem 66. *Let (Ω, σ, X, A) be a strongly measurable forward-integrable linear dynamical system with (Ω, σ) an ergodic invertible map on a Lebesgue probability space and having Lyapunov exponents $(\mu_i)_{i=1}^{\infty}$ and $(\lambda_i)_{i=1}^L$, where $1 \leq L \leq \infty$. Then for each $0 \leq l < L$ there is a direct sum decomposition into equivariant spaces $X = \left(\bigoplus_{i \leq l} E_i(\omega) \right) \oplus V_{l+1}(\omega)$ with the $E_l : \Omega \rightarrow \mathcal{G}_{m_l} X$ measurable, $m_l \in \mathbb{N}$ and having*

$$\frac{1}{n} \log \|A_\omega^{(n)}|_{E_i(\omega)}\|, \frac{1}{n} \log \inf \{ \|A_\omega^{(n)} x\| : x \in \mathbb{S}_{E_i(\omega)} \} \rightarrow \lambda_i,$$

and

$$V_l(\omega) = \{x \in X : \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_\omega^{(n)} x\| \leq \lambda_l\}.$$

The projection $\Pi : X \rightarrow \bigoplus_{j < i} E_j$ parallel to V_i is strongly measurable and tempered. There is a nontrivial decomposition, $L \geq 2$, exactly when $\nu = \lim_{n \rightarrow \infty} \mu_n < \lambda_1$.

In this Banach space setting, the main theorem yields in a trichotomous classification of forward-integrable cocycles on separable Banach spaces: one of

- A fails to be quasicompact, with $\mu_i = \nu$ for all $i \in \mathbb{N}$ - no fast spaces may be detected
- A is quasicompact, with finitely many finite dimensional fast spaces growing at rates $\lambda_1 > \dots > \lambda_L > \lambda_{L+1} = \nu$
- A is quasicompact, with a countable sequence of finite dimensional fast spaces growing at rates $\lambda_1 > \lambda_2 > \dots \rightarrow \nu$.

These three possibilities just correspond to situations where there are no, finitely many or countably many fast spaces. The number ν is an alternative choice to the typical index of compactness $\kappa = \lim_{n \rightarrow \infty} \frac{1}{n} \|A_\omega^{(n)}\|_c$ that proves simpler to work with in this proof: we have seen they are equal by lemma 55.

6.2 Proof of main result

Finally, we may conclude with the main result, a well behaved decomposition of the space acted on by a linear dynamical system:

Proof of theorem 66. The decomposition is obtained inductively. At each stage it is shown that if λ_{i+1} exists, there exists an equivariant decomposition

$$X = E_{\leq i}(\omega) \oplus V_{i+1}(\omega)$$

with $E_{\leq i} \in \mathcal{SM}(\omega \rightarrow \mathcal{G}_{M_i} X)$ and bounded projections

$$\Pi_{i+1, \omega} : X \rightarrow V_{i+1}(\omega) \quad \text{and} \quad P_{i, \omega} = 1 - \Pi_{i, \omega} : X \rightarrow E_{\leq i}(\omega)$$

such that there is an equivariance type condition for the projections $A_\omega \Pi_{i, \omega} = \Pi_{i, \sigma \omega} A_\omega$ pointwise almost everywhere. The existence of the top fast space has already been established - here denote this $E_{< 2}(\omega) \oplus V_2(\omega)$ with measurable projections $\Pi_{2, \omega}$ and $P_{2, \omega}$.

Suppose that the statement is true up to $i = l - 1$. If $\lambda_l = \nu$ then we are done, so suppose otherwise - that there exists $\lambda_{l+1} \geq \nu$. The projection $\Pi_{l, \omega}$ is pointwise

bounded, so that there exists some $M > 0$ such that $A = \{\|\Pi_{l,\omega}\| < M\}$ has positive measure.

$$A'_\omega = \begin{cases} A_\omega \circ \Pi_{l,\omega}, \omega \in A, \\ A_\omega \text{ otherwise.} \end{cases}.$$

As before A' is forward integrable. Write $\lambda', \mu'_i, \lambda'_i, \nu', E_l \oplus V', m'_i, P'$ and Π' for the asymptotic growth rates, Lyapunov exponents, decomposition, fast space multiplicities, fast projection and slow projection with A' . There almost surely exists an $N \in \mathbb{N}$ such that for each $n \geq N$, the following hold:

- $A'^{(n)}_\omega = A_\omega^{(n)} \circ \Pi_{l,\omega} = \Pi_{l,\sigma^n \omega} \circ A_\omega^{(n)}$ by the equivariance condition for the projections.
- $\|\Pi_{\sigma^n \omega}\| < e^{n\epsilon}$
- $g_{A_\omega^{(n)}}(E(\omega)) > \|A_\omega^{(n)}|_{V_{l+1}(\omega)}\|$.

Therefore applying lemma 45 and 46 to $A'^{(n)}_\omega = A_\omega^{(n)} \circ \Pi_\omega$, for each k the following holds:

$$\rho_{k+M_l}(A_\omega^{(n)}) \leq \rho_k(A'^{(n)}_\omega) \leq 4\|\Pi_{\sigma^n \omega}\| \|\Pi_\omega\| \rho_{k+M_l}(A_\omega^{(n)}) \leq 4e^{n\epsilon} \|\Pi_\omega\| \rho_{k+M_l}(A_\omega^{(n)}),$$

whence $\mu_{M_l+k} = \mu'_k$ and $\nu = \nu'$. Further then, $m'_k = m_{l+k}$ and $\lambda'_k = \lambda_{k+l}$. Set $V_{l+1}(\omega) = V_l(\omega) \cap V'(\omega)$ and $\Pi_{l+1\omega} = \Pi_{l\omega} \circ \Pi'_\omega$. Write $E_{<l}(\omega) = E_{<l}(\omega) + E_l(\omega)$. The equality $A_\omega = A'_\omega$ holds on $V_l(\omega)$, whence $\lambda_\omega = \lambda'_\omega$ on $V_l(\omega)$. Since A has positive measure, there is almost surely an N such that for all $n \geq N$, $A'^{(n)}_\omega E_{<l}(\omega) = 0$, $A'^{(n)}_{\sigma^{-n}\omega} X \subseteq V_l(\omega)$, and $A'^{(n)}_\omega X \subseteq V_l(\sigma^n \omega)$. We may then conclude that for each $x \in X$, $\lambda'_\omega(x) \leq \lambda_l$ and $\lambda'_1 \leq \lambda_l$. As such, for each $i < l$, have $E_i(\omega) \subseteq V'(\omega)$. As for $E_l(\omega)$, by equivariance $E_l(\omega) = A'^{(n)}_{\sigma^{-n}\omega} E(\sigma^{-n}\omega) \subseteq V_l(\omega)$. Thus $X = E_{<l}(\omega) \oplus E_l(\omega) \oplus (V_l(\omega) \cap V'(\omega)) = E_{<l}(\omega) \oplus E_l(\omega) \oplus V_{l+1}(\omega)$.

The map Π_{l+1} is then also tempered:

$$0 \leq \frac{1}{n} \log \|\Pi_{l+1\sigma^n \omega}\| \leq \frac{1}{n} \log \|\Pi_{l\sigma^n \omega}\| + \frac{1}{n} \log \|\Pi'_{\sigma^n \omega}\| \rightarrow 0.$$

Let $\epsilon > 0$. There exists an N such that for $n \geq N$, $\|P'_{\sigma^n \omega}\|, \|\Pi'_{\sigma^n \omega}\| < e^{n\epsilon}$ and so for all $x \in X$ we have $\max\{\|P'_{\sigma^n \omega} A_\omega^{(n)} x\|, \|\Pi'_{\sigma^n \omega} A_\omega^{(n)} x\|\} \leq \|A_\omega^{(n)} x\| e^{n\epsilon}$. Suppose now $x \in E_{<l}(\omega)$ so that $\Pi' x \in E_{<l}(\omega)$. Rearranging the last inequality we obtain

$$\begin{aligned} \|A_\omega^{(n)} x\| &\geq e^{-n\epsilon} \max\{\|P'_{\sigma^n \omega} A_\omega^{(n)} x\|, \|\Pi'_{\sigma^n \omega} A_\omega^{(n)} x\|\} \\ &= e^{-n\epsilon} \max\{\|A_\omega^{(n)} P'_\omega x\|, \|A_\omega^{(n)} \Pi'_\omega x\|\} \\ &\geq e^{-n\epsilon} \max\{e^{n(\lambda_l - \epsilon)} \|P'_\omega x\|, e^{n(\lambda_1 - \epsilon)} \|\Pi'_\omega x\|\} \\ &\geq e^{-n\epsilon} e^{n(\lambda_l - \epsilon)} \max\{\|\Pi'_\omega x\|, \|P'_\omega x\|\} \geq \frac{1}{2} e^{n(\lambda_l - 2\epsilon)}. \end{aligned}$$

For $n \geq \max\{N, \frac{1}{\epsilon} \log 2\}$ it follows that $g_{A_\omega^{(n)}}(E_{\leq l}(\omega)) \geq e^{n(\lambda_l - 3\epsilon)}$. The characterisation

$$V_{l+1}(\omega) = \{x \in X : \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|A_\omega^{(n)} x\| \leq \lambda_{l+1}\}$$

holds.

□

Bibliography

- [1] L. Arnold, D. C. Nguyen, and V. Oseledets. Jordan normal form for linear cocycles. *7(4)*:303–358, 1999.
- [2] S. Artstein-Avidan, H. König, and A. Koldobsky. Asymptotic geometric analysis. *Oberwolfach Reports*, 13(1):507–565, 2016.
- [3] A. Blumenthal. A volume-based approach to the multiplicative ergodic theorem on Banach spaces. *Discrete and Continuous Dynamical Systems - A*, 36, 2016.
- [4] Jairo Bochi. Genericity of zero lyapunov exponents. *Ergodic Theory and Dynamical Systems*, 22(6):1667–1696, 2002.
- [5] G. Choquet. Existence et unicité des représentations intégrales au moyen des points extrémaux dans les cônes convexes. *Séminaire Bourbaki*, 4:33–47, 1956.
- [6] T. de La Rue. Espaces de lebesgue. *Séminaire de probabilités de Strasbourg*, 27:15–21, 1993.
- [7] S. J. Eigen and V. S. Prasad. Multiple Rokhlin tower theorem: a simple proof. *New York J. Math. A*, 3:11–14, 1997.
- [8] S. E. Ferrando. Moving ergodic theorems for superadditive processes. *Canadian Journal of Mathematics*, 47(4):728–743, 1995.
- [9] G. Froyland, S. Lloyd, and A. Quas. A semi-invertible Oseledets theorem with applications to transfer operator cocycles. *Discrete and Continuous Dynamical Systems*, 33, 01 2010.
- [10] H. Furstenberg and H. Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, 1960.
- [11] J. W. Gibbs. Elementary principles in statistical mechanics: developed with special reference to the rational foundation of thermodynamics. 1914.
- [12] C. González-Tokman and A. Quas. A semi-invertible operator Oseledets theorem. *Ergodic Theory and Dynamical Systems*, 34, 05 2011.

- [13] C. González-Tokman and A. Quas. A concise proof of the multiplicative ergodic theorem on Banach spaces. *Journal of Modern Dynamics*, 9(01):237–255, 2015.
- [14] J. A. Horan. *Spectral gap and asymptotics for a family of cocycles of Perron-Frobenius operators*. PhD thesis, University of Victoria, 2020.
- [15] S. Novo J. Mierczyński and R. Obaya. Lyapunov exponents and Oseledets decomposition in random dynamical systems generated by systems of delay differential equations. *Communications on Pure and Applied Analysis*, 19(4):2235–2255.
- [16] S. Kalikow. Infinite partitions and Rokhlin towers. *Ergodic Theory and Dynamical Systems*, 32(2):707–738, 2012.
- [17] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*. Cambridge university press, 1995.
- [18] J. F. C. Kingman. The ergodic theory of subadditive stochastic processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 30(3):499–510, 1968.
- [19] U. Krengel. *Ergodic theorems*, volume 6. Walter de Gruyter, 2011.
- [20] Z. Lian and K. Lu. *Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space*. American Mathematical Soc., 2010.
- [21] R. Mañé. Lyapounov exponents and stable manifolds for compact transformations. In *Geometric dynamics*, pages 522–577. Springer, 1983.
- [22] V. I. Oseledets. A multiplicative ergodic theorem. characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 19:179–210, 1968.
- [23] A. Pietsch. s -Numbers of operators in Banach spaces. *Studia Mathematica*, 51(3):201–223, 1974.
- [24] M. S. Raghunathan. A proof of Oseledec’s multiplicative ergodic theorem. *Israel Journal of Mathematics*, 32(4):356–362, 1979.
- [25] W. Ramasinghe. Exterior algebra of a banach space. *Bulletin des sciences mathématiques*, 131(3):291–324, 2007.
- [26] D. J. Rudolph. *Fundamentals of Measurable Dynamics: Ergodic Theory on Lebesgue Spaces*. Oxford Science Publications. Clarendon Press, 1990.
- [27] D. Ruelle. Ergodic theory of differentiable dynamical systems. *Publications Mathématiques de l’Institut des Hautes Études Scientifiques*, 50(1):27–58, 1979.

- [28] P. Thieullen. Fibrés dynamiques asymptotiquement compacts exposants de lyapounov. entropie. dimension. In *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*. Elsevier, 1987.
- [29] P. Wojtaszczyk. *Banach spaces for Analysts*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1991.