
A Guide to the Ergodic Decomposition Theorem: Ergodic Measures as Extreme Points

JORGE ALBARRÁN

September 22, 2004

1 Introduction

In the study of dynamical systems *invariant measures* play a very important role, being the *ergodic measures* the simplest invariant measures of a dynamical system. The question motivating these notes is the following:

Is it possible to decompose an invariant measure into ergodic measures in such a way that the study of the former reduces to the study of the second ones?

Two remarkable theorems give an affirmative answer to this: *Krein-Milman's* and *Choquet's* Theorems.

The basic idea lies in a classic theorem of *Minkowski*, who says that the points in a convex set in finite dimension are a *convex sum* of its *extreme points*, e.g. it is possible to get some information of the points in the convex set via some special points: extreme points.

One can easily show that ergodic measures are extreme points of the convex set of invariant measures of some given continuous dynamical system, that is the reason of why Minkowski's idea is useful.

The goal of these notes is to provide a basic almost self-content path through the necessary theory to understand the ergodic decomposition

theorem by a 3rd grade undergraduate student. Most of the proofs will be only referred.

The last section is an extract of the author's graduate thesis [1] and it illustrates an application of ergodic decomposition to the study of *transverse measures* and *foliated cycles*¹ in one-dimensional foliations.

Acknowledgements: I would like to thank Alberto Verjovsky for his explanations on the subject and whose conversations brought the idea of writing these notes, to the valuable help of Manuel Cruz, Daniel Massart and Ana García for reading these notes and making useful comments.

Keywords: *Convex, Minkowski, Krein-Milman, Choquet, measure theory, integral representations, ergodic theory, ergodic decomposition, foliated cycles, invariant transverse measures, one-dimensional foliations.*

2 Convex Sets, Extreme Points and Minkowski's Theorem

The theory of convex sets in \mathbb{R}^n is a classical very well studied area of mathematics, there are many introductory books on this subject, we cite [3, 20] in the references.

Let us begin with the basic definitions on convexity.

Let \mathbb{E} be a finite dimensional vector space. A subset $C \subset \mathbb{E}$ is called a **convex set** if for every $a, b \in C$, the segment line determined by a and b is contained in C , i.e.

$$\{\lambda a + (1 - \lambda)b : \lambda \in [0, 1]\} \subset C. \quad (1)$$

For every $A \subset \mathbb{E}$ we define the **convex hull of A** as the set:

$$\text{conv}(A) = \bigcap \{C \subset \mathbb{E} \text{ convex} : A \subset C\}. \quad (2)$$

In words: $\text{conv}(A)$ is the smallest convex set containing A .

¹These two being the same by a theorem of Dennis Sullivan [19].

A point x in a convex set C is called an **extreme point** if the following condition is satisfied: given $x_1, x_2 \in C$, there exists $\lambda \in]0, 1[$ such that if $x = \lambda x_1 + (1 - \lambda)x_2$ then $x = x_1 = x_2$.

In words: *there exist no segment containing x in its interior that is entirely contained in C .*

We denote by $\text{ex}(C)$ to the set of extreme points of the convex set C .

Example 2.1

The extreme points of a plane polygon are its vertices.

Example 2.2

The closed n -dimensional disk $\mathbb{D}^n \subset \mathbb{R}^n$ has the $(n - 1)$ -dimensional sphere \mathbb{S}^{n-1} as its set of extreme points.

Example 2.3

In general, the set $\text{ex}(C)$ of extreme points of C is not a closed or convex set. In fact, if we let $D := \{(x, y, z) \in \mathbb{R}^3 : x^2 + (y - 1)^2 \leq 1, z = 0\}$ and we define:

$$C := D \cup \{(0, 0, 1), (0, 0, -1)\}, \tag{3}$$

the set $\text{conv}(C)$ (see figure 1) is the union of two convex closed cones: one with vertex at the point $(0, 0, 1)$ and base D and the other with vertex at $(0, 0, -1)$ and base D (see exercises). Then $\text{conv}(C)$ is compact, however (see exercises)

$$\text{ex}(\text{conv}(C)) = (\partial D \setminus \{(0, 0, 0)\}) \cup \{(0, 0, 1), (0, 0, -1)\}, \tag{4}$$

where ∂D is the boundary of D . Therefore $\text{ex}(\text{conv}(C))$ is not a closed neither convex set.

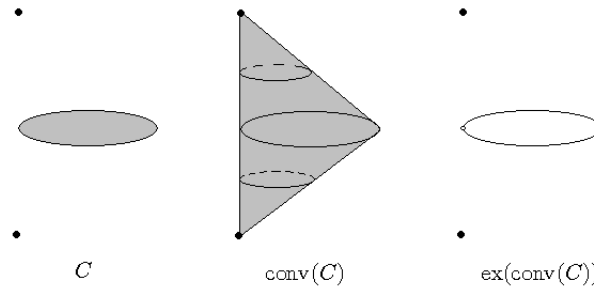


Figure 1: The set $\text{ex}(\text{conv}(C))$ is not closed even when C is compact.

The next definition is the way we can represent the points of a convex set via its extreme points. Given k points $x_1, \dots, x_k \in \mathbb{E}^n$, we call a **convex combination** of them to a finite linear combination $\sum_{i=1}^k \alpha_i x_i$, with all the coefficients positive and $\sum_{i=1}^k \alpha_i = 1$.

We state now the main theorem of this section.

Theorem 2.4 (Minkowski) *Let $C \subset \mathbb{E}^n$ be a convex, compact set. Then each $x \in C$ can be written as a convex combination of elements in $\text{ex}(C)$.*

Proof. See [8] p. 36. □

The moral of this theorem is the following: *each point of a compact, convex set C in finite dimension can be represented as a convex combination of extreme points of C .*

This classic result is also known as the finite-dimensional version of Krein-Millman's theorem. Carathéodory shows a stronger result: *the convex combination in the conclusion in Minkowski's theorem has at most $n + 1$ terms (see exercises 2).*

Exercises 1

1. Complete the details of example 2.3, e.g. prove that the sets $\text{conv}(C)$ and $\text{ex}(\text{conv}(C))$ are indeed the sets mentioned above.
2. Let C be a convex set. Show that a point x belongs to the set $\text{ex}(C)$ if and only if $C \setminus \{x\}$ is a convex set.
3. Prove that the conclusion in Minkowski's theorem (Theorem 2.4) is equivalent to:

$$C = \text{conv}(\text{ex}(C)).$$

4. If $C \subset \mathbb{R}^n$ is a compact convex set and if $x \in \text{ex}(C)$ then there exists a hyperplane $H \subset \mathbb{R}^n$ such that $x \in H \cap C$, and if $h \in H$ the sign of the canonical inner product in \mathbb{R}^n is constant for every $y \in C$, e.g. C is entirely contained in one of the halfspaces defined by H . Such a hyperplane H is called a **supporting hyperplane**.
5. Show the result of Carathéodory mentioned above by induction on the dimension of C . (Hint: Use exercise 4).

3 Convexity in Infinite Dimension: Krein-Milman's Theorem

Many properties that we usually have in finite-dimensional vector spaces are lost when passing to infinite-dimensional ones. Continuing our

study of the set of extreme points we illustrate this point with the following examples appearing in [17].

Example 3.1

A generalization of example 2.2 is the following.

Let $(X, \|\cdot\|)$ be a normed vector space and let $B_X := \{x \in X : \|x\| \leq 1\}$ be the closed unit ball, then

$$\text{ex}(B_X) \subset S_X := \{x \in X : \|x\| = 1\},$$

since $x = \lambda(x/\|x\|) + (1 - \lambda)(0)$ is true for $\lambda = \|x\|$, i.e. if x is an extreme point we must have $\|x\| = 1$. However the other contention is not true in general as we see in the next example.

Example 3.2

In the space $L^1[0, 1]$ (e.g. the real-valued functions in the interval $[0, 1]$ for which $\|f\| := \int_0^1 |f(t)|dt < \infty$) the ball B_X doesn't have any extreme point.

In fact, by the example 3.1 above the extreme points of B_X lie on S_X . The next discussion shows that any $f \in S_X$ cannot be an extreme point in B_X .

Let $f \in S_X$, the map $s \mapsto \int_0^s |f(t)|dt$ is continuous in s , then, by the mean value theorem, there exists $s_0 \in]0, 1[$ such that $\int_0^{s_0} |f(t)|dt = 1/2$.

Let us define $f_1(t) := 2\chi_{[0, s_0]}(t)f(t)$ and $f_2(t) := 2\chi_{]s_0, 1]}(t)f(t)$, where $\chi_A(t)$ is the indicating function of the set A (e.g. $\chi_A(t) = 1$ if $t \in A$ and $\chi_A(t) = 0$ otherwise). An easy computation shows that $\|f_i\| = 1$ for $i = 1, 2$ and $f = \frac{1}{2}f_1 + \frac{1}{2}f_2$, but $f_1 \neq f_2$, therefore f cannot be an extreme point of B_X .

In general, the set of extreme points is not even a borel set as Bishop and Phelps [2] show; however, it is possible to give sufficient conditions for avoiding this as the next section shows.

Let us state some basic notions of the theory of topological vector spaces, for details on this theory we refer the reader to [11, 16].

Let \mathbb{K} be a field, a **(\mathbb{K} -)topological vector space** is a (\mathbb{K} -)vector space with a Hausdorff topology such that translation and scalar multiplication are homeomorphisms. A topological vector space is **locally convex** if every open set contains an open convex set. The following examples are in [16].

Example 3.3

Every normed vector space is a topological vector space with the induced topology by the norm, moreover, it is a locally convex vector space.

Example 3.4

Let A be a non-empty set and define \mathbb{R}^A to be the set of all functions from A to \mathbb{R} . Then \mathbb{R}^A is an \mathbb{R} -vector space with the addition defined by $(f + g)(x) = f(x) + g(x)$.

It is possible to define a basis for a topology over \mathbb{R}^A that makes it a topological vector space. Define the sets

$$V_{H,\varepsilon} := \{f : A \rightarrow \mathbb{R} : |f(\alpha)| \leq \varepsilon \text{ if } \alpha \in H\}, \text{ with } H \subset A \text{ finite and } \varepsilon \geq 0.$$

A neighborhood of $f \in \mathbb{R}^A$ is defined as $S_f = f + S_0$, where $V_{H,\varepsilon} \subseteq S_0$ for some H and ε , in this case we say S_0 contains a **0-neighborhood** in \mathbb{R}^A . We then define an open W of \mathbb{R}^A as a set for which $f \in W$ implies that there exists a neighborhood $S_f \subseteq W$ of f .

Example 3.5

Let $V \neq \emptyset$ be a topological vector space and let us define

$$\mathcal{C}_{\mathbb{R}}(V) := \{f : V \rightarrow \mathbb{R} : \sup_{t \in V} |f(t)| < \infty\} \subset \mathbb{R}^V.$$

The set $\mathcal{C}_{\mathbb{R}}(V)$ is a vector space with the operations of addition and scalar multiplication induced by \mathbb{R}^V . A basis for a topology for $\mathcal{C}_{\mathbb{R}}(V)$ is given by the set

$$\mathcal{U}_n := \{f : \sup_{t \in V} |f(t)| \leq n^{-1}\}, \text{ for } n \in \mathbb{N}.$$

The topology induced by these sets as in the previous example makes $\mathcal{C}_{\mathbb{R}}(V)$ a topological vector space.

Example 3.6

Let $\mathbb{R}[t]$ be the ring of polynomials with coefficients in \mathbb{R} , this is a vector space with the usual addition and scalar multiplication. A basis for a topology that makes it a topological vector space is given by the sets

$$V_\varepsilon := \{p \in \mathbb{R}[t] : \sum |p(t)|^r \leq \varepsilon\}, r \in]0, 1].$$

The main theorem of this section, published in [12], is the following.

Theorem 3.7 (Krein-Milman) *Let E be a topological locally convex vector space. If $C \subset E$ is a compact convex set then*

$$C = \overline{\text{conv}(\text{ex}(C))}. \tag{5}$$

Proof. [5] Vol. II, p. 106, or [11] p. 335. □

Remark that we have to take the closure of the set $\text{conv}(\text{ex}(C))$ in the equality 5, which is not necessary in the finite-dimensional version, namely Minkowski’s theorem 2.4. This is the reason why we cannot, in general, write every point of C as a convex combination of some of its extreme points, but an infinite approximation of these. However, it is possible to generalize the concept of convex combinations using some integral representations which is one of the points in the next section.

Exercises 2

1. Show the affirmations of example 3.3, i.e. if V is a normed vector space, then it is a locally convex topological vector space with the topology induced by the norm.
2. Show that in the statement of Krein-Milman’s theorem (Theorem 3.7) it is not possible to replace *compact* by *closed and bounded*.
3. Show with detail that Minkowski’s theorem (Theorem 2.4) follows from Krein-Milman’s theorem (Theorem 3.7).

4 Measure Representation: integral form of Krein-Milman’s theorem and Choquet’s theorem.

We start with some basic definitions from measure theory, we follow in part sections 1 and 4 of [21].

Recall a vector space is **locally compact** if for every point of $x \in X$ there exists a compact subset of X containing x .

Let X be a topological locally compact space and \mathcal{B} its Borel σ -algebra (the σ -algebra generated by the open subsets of X). A **borel measure** on X is a σ -aditive map $\mu : \mathcal{B} \rightarrow \mathbb{R}_+ \cup \{\infty\}$, i.e. $\mu(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mu(A_i)$ if $A_i \cap A_j = \emptyset$ for every $i, j \in \mathbb{N}, i \neq j$.

☐ **NOTE:** In the rest of this section X is supposed to be a locally compact topological space.

A measure μ on X is said to be:

- a) A **probability measure**: if $\mu(X) = 1$.
- b) A **finite measure**: if $\mu(X) < \infty$
- c) A **regular measure**: if for every $B \in \mathcal{B}$ we have

$$\mu(B) = \sup\{\mu(C) \mid C \subset B, C \text{ compact}\}, \text{ and}$$

$$\mu(B) = \inf\{\mu(A) \mid A \supset B, A \text{ open}\}$$

- d) A **Radon measure**: if it is finite and regular.

Define the **support of a measure** μ as the set $\text{supp}(\mu) \subset X$ characterized by the condition:

$$x \in \text{supp}(\mu) \Leftrightarrow \mu(B) > 0, \quad \forall B \in \mathcal{B}, \text{ such that } x \in B.$$

We denote the set of **Radon probability measures** by $\mathcal{M}^1(X)$.

A more general definition is the one of **signed measure**, this is a σ -aditive map $\mu : \mathcal{B} \rightarrow \mathbb{R}$ such that $\mu(\emptyset) = 0$. If μ is a signed measure we define the maps:

$$\mu^+(B) := \sup\{\mu(E) \mid E \subset B\} \text{ and } \mu^-(B) := \sup\{-\mu(E) \mid E \subset B\},$$

for $B \in \mathcal{B}$. We then define a **Radon signed measure** as a signed measure μ , such that μ^+ y μ^- are Radon measures.

We denote by $\mathcal{M}(X)$ to the set of all Radon signed measures on X . Note that $\mathcal{M}^1(X) \subset \mathcal{M}(X)$.

The set $\mathcal{M}(X)$ is a natural vector space² and, moreover, if we define the norm

$$\|\mu\| := \mu^+(X) + \mu^-(X), \quad (6)$$

then the space $\mathcal{M}(X)$ with this norm is a complete normed vector space, i.e. a Banach space. Even more, we can topologize this space in order to turn it into a topological vector space such that the subspace $\mathcal{M}^1(X)$ is a compact set. We illustrate this as follows.

Let $C(X)$ be the vector space of continuous functions $f : X \rightarrow \mathbb{R}$ equipped it with the norm:

$$\|f\|_\infty := \sup_{x \in X} |f(x)|.$$

Its dual vector space $C(X)^*$, defined as the set of linear functionals $\varphi : C(X) \rightarrow \mathbb{R}$ which are bounded with the norm:

$$\|\varphi\| := \sup_{\|f\|_\infty=1} \|\varphi(f)\|,$$

is also a normed vector space³.

Let us state the famous Riesz representation theorem:

Theorem 4.1 (Riesz) *Let X be a compact Hausdorff space. The relation*

$$\varphi(f) = \int_X f d\mu,$$

with $f \in C(X)$ defines a biunivoque correspondence between $C(X)^$ and $\mathcal{M}(X)$.*

Proof. [9] p. 111. □

This theorem gives us the important identification:

$$\boxed{C(X)^* \cong \mathcal{M}(X)}$$

Now, a very interesting and important topology on $\mathcal{M}(X)$, called the **weak* topology** (weak-star topology), is characterized by the following condition:

$$\varphi_i \rightarrow 0 \Leftrightarrow \varphi_i(f) \rightarrow 0, \quad \forall f \in C(X), \quad (7)$$

²Addition is pointwise addition and scalar multiplication is the usual one.

³Hahn-Banach's theorem (see [9, p. 135]) assure the existence of linear functionals in normed spaces and therefore $C(X)^* \neq \emptyset$.

that is: the topology of pointwise convergence in $C^*(X)$.

Intepreting the weak* topology via the Riesz representation theorem in the space $\mathcal{M}(X)$, we obtain a topology characterized by:

$$\mu_i \rightarrow \mu \Leftrightarrow \int_X f d\mu_i \rightarrow \int_X f d\mu, \quad \forall f \in \mathcal{M}(X). \quad (8)$$

This topology makes $\mathcal{M}(X)$ a topological vector space, and an easy diagonal argument shows that it contains $\mathcal{M}^1(X)$ as a compact set. This is a very useful and important result which is the key in the rest of this paper, that's why it worths to restate it as a proposition.

Proposition 4.2 *The set $\mathcal{M}(X)$, equipped with the weak* topology, is a topological vector space and it contains the set $\mathcal{M}^1(X)$ as a weak*-compact vector subspace.*

Now, for starting the discussion on measure representation we start with the finite-dimensional case.

As we have mentioned before, Minkowski's theorem (Theorem 2.4) tells us that every point x in a compact convex set $C \subset \mathbb{E}$ of finite dimension, can be written as a convex combination of extreme points of C :

$$x = \sum_{i=1}^k \alpha_i x_i, \quad \text{with } \sum_i \alpha_i = 1, \alpha_i \geq 0 \text{ and } x_i \in \text{ex}(C), \quad \forall i = 1, \dots, k.$$

So, let us consider the measure:

$$\mu = \sum_i^k \alpha_i \delta_{x_i}, \quad (9)$$

where δ_{x_i} is the Dirac measure on the point x_i , i.e. for every borel set B

$$\delta_{x_i}(B) = \begin{cases} 1 & \text{if } x_i \in B \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\mu(C) = \mu(\text{ex}(C)) = \sum_{i=1}^k \alpha_i = 1,$$

and μ is a Radon measure with $\text{supp}(\mu) = \{x_1, x_2, \dots, x_k\} \subset \text{ex}(C)$, therefore:

$$\int_C z d\mu(z) = \sum_{i=1}^k \alpha_i \int_C z d\delta_{x_i}(z) = \sum_{i=1}^k \alpha_i x_i = x.$$

We can then restate Minkowski's theorem as follows

Theorem 4.3 (Integral form of Minkowski's Theorem) *Let C be a compact convex subset of a finite-dimensional vector space. Then, for every point $x \in C$, there exists a measure $\mu \in \mathcal{M}^1(\text{ex}(C))$ whose support is a finite subset of $\text{ex}(C)$ and such that*

$$x = \int_C z d\mu(z).$$

The question is now: *Can we use Krein-Milman's theorem instead of Minkowski's theorem to conclude some similar property in the infinite-dimensional case?*

Fortunately we have an affirmative answer to this question as will be explained in the rest of this section.

The previous discussion motivates us to make the following definition. Let C be a compact convex set in a locally convex topological vector space. We call a point $x \in C$ a **barycenter** of a measure $\mu \in \mathcal{M}^1(C)$ if

$$x = \int_C z d\mu(z).$$

The name barycenter is justified since in the finite dimensional case x is a convex combination of a finite set of points of A , therefore it is the barycenter of the geometric figure determined by the convex hull of A .

This generalizes the notion of a convex combination, this is the infinite-dimensional version of *representing points*. We now restate Krein-Milman's theorem.

Theorem 4.4 (Integral form of Krein-Milman's Theorem) *Let \mathbb{E} be a locally convex topological vector space and let $C \subset \mathbb{E}$ be a compact convex subset. Then every point of C is the barycenter of a measure whose support is contained in $\mathcal{M}^1(\overline{\text{ex}(C)})$.*

Proof. [21] p. 5. □

The conclusion of the theorem, following the definition of barycenter, can be also stated as:

For every $x \in C$ there exists a measure $\mu \in \mathcal{M}^1(\overline{\text{ex}(C)})$ such that

$$x = \int_{\overline{\text{ex}(C)}} z d\mu(z).$$

The moral is then: *every point of C can be represented as a “continuous sum” of points in $\overline{\text{ex}(C)}$.*

Remark that this theorem is not an exact generalization of the finite dimensional case, since the measure has support in the closure of $\text{ex}(C)$ and not in $\text{ex}(C)$ itself and we have seen in the example 2.2 that this set is not closed in general. However a famous beautiful theorem of the french mathematician Gustave Choquet is the affirmative answer to our question as an exact generalization of the finite dimensional case introducing an extra hypothesis on C .

We say a topological vector space is **metrizable** if there exists a metric on it that induces the given topology, i.e. the balls in that metric are a basis for the topology. Then, if we ask our compact convex set to be also metrizable we have the following nice result.

|| **Theorem 4.5** (Choquet) *Let \mathbb{E} be a locally convex topological vector space and $C \subset \mathbb{E}$ a compact convex metrizable set. Then every point of C is the barycenter of a measure $\mu \in \mathcal{M}^1(\text{ex}(C))$.*

Proof. [5] Vol II, p. 140-141. □

Here we can also restate the conclusion as:

For every $x \in C$ there exists a measure $\mu \in \mathcal{M}^1(\text{ex}(C))$ such that

$$x = \int_{\text{ex}(C)} z d\mu(z).$$

The moral is then: *if C is metrizable then we obtain an exact generalization of Minkowski’s theorem in the infinite-dimensional case, i.e. every point of C is representable by a measure with support contained*

in $\text{ex}(C)$.

More general and strong versions of these theorems are stated in [14, §2] and [15, §1,4,12]. A complete introduction to the weak* topology can be readed in [23].

Exercices 3

1. Show that $\|\mu\|$ defined in 6 is a norm.
2. Prove that the following conditions are equivalent:
 - i) $|\mu|$ is a Radon measure.
 - ii) μ is a signed Radon measure.
 - iii) $|\mu|(B) = \sup\{|\mu|(C) : C \subset B, C \text{ compact}, \forall B \in \mathcal{B}\}$.

3. Show that the measure

$$\mu = \sum_i^k \alpha_i \delta_{x_i},$$

defined in 9 is a Radon measure.

4. Derive Krein-Milman's theorem (Theorem 3.7) from its integral representation (Theorem 4.4).

5 Ergodic Decomposition

Grosso modo we can say that ergodic theory studies the movement in a measure space, e.g. we're interested on transformations of a measure space into itself. The definitions given here are not the more general ones, however we refer to [6, 10, 22] for a complete introduction to ergodic theory and dynamical systems. The examples given here can be found in the references above, therefore we ommit the details. Only for those examples deserving some special attention an outline of the proofs of the statements will be given as guided exercises at the end of the section. We hope the reader will reach that point motivated enough to fit in the details.

Let (M, \mathcal{B}, μ) be a measure space where M is a locally compact Hausdorff topological space and \mathcal{B} its Borel σ -algebra. Also, let $\varphi : M \rightarrow M$ be a measurable continuous map, e.g.

$$\varphi^{-1}(B) \in \mathcal{B}, \quad \forall B \in \mathcal{B}.$$

A set $A \in \mathcal{B}$ is called an **invariant set** or **μ -invariant** if

$$\mu(A) = \mu(\varphi^{-1}(A)).$$

Remark that if $A \in \mathcal{B}$ is an invariant set then $M \setminus A$ is also an invariant set, i.e.

$$\mu(\varphi^{-1}(M \setminus A)) = \mu(M \setminus A) \Leftrightarrow \mu(\varphi^{-1}(A)) = \mu(A),$$

then, we can study the map φ from the two simpler maps $\varphi|_A$ and $\varphi|_{M \setminus A}$. However, we have really gained something only when $0 < \mu(A) < 1$, since if $\mu(A) = 0$ all the information is carried on the map $\varphi|_{M \setminus A}$ and by symmetry the same is true if $M \setminus A$ when $\mu(A) = 1$. Therefore, if there exist an invariant set $A \in \mathcal{B}$ such that $0 < \mu(A) < 1$, then we are able to study φ via two simpler maps. This discussion motivates our next two definitions.

A measure $\mu \in \mathcal{M}^1(M)$ is said to be **invariant** or **φ -invariant** if

$$\mu(B) = \mu(\varphi^{-1}(B)), \quad \forall B \in \mathcal{B}.$$

The measure μ is said to be **ergodic** or **φ -ergodic** if the only μ -invariant sets have total or zero measure, e.g. for every $A \in \mathcal{B}$,

$$\mu(\varphi^{-1}(A)) = \mu(A) \Leftrightarrow \mu(A) = 1, 0.$$

Before going into some examples of ergodic measures we'll state some equivalences for a measure to be ergodic. Recall that a property is valid for **μ -almost every point** or **μ -a.e.p** if the set of points which do not have that property has μ -measure zero. A probability space is a measure space (M, \mathcal{B}, μ) with $\mu \in \mathcal{M}^1(M)$.

Theorem 5.1 *Let (M, \mathcal{B}, μ) be a probability space and $\varphi : M \rightarrow M$ a measurable map. Then the following are equivalent:*

i) *The measure μ is φ -ergodic.*

ii) *If $B \in \mathcal{B}$ then:*

$$\mu((\varphi^{-1}(B) \setminus B) \cup (B \setminus \varphi^{-1}(B))) = 0 \quad \text{implies} \quad \mu(B) = 0, 1.$$

iii) *If $f : M \rightarrow \mathbb{C}$ is measurable then:*

$$(f \circ \varphi)(x) = f(x) \text{ for } \mu\text{-a.e.p. implies } f \text{ is constant for } \mu\text{-a.e.p.}$$

iv) If $f \in L^2(M)$ (i.e. $\int_M |f(x)|^2 dx < \infty$) then:

$(f \circ \varphi)(x) = f(x)$ for μ -a.e.p. implies f is constant for μ -a.e.p.

Proof. [22] p. 27 y 28. □

Example 6.1

The trivial example is the identity map on M , Id. In this case every probability measure on M is Id-invariant and a measure is Id-ergodic if and only if every element of \mathcal{B} has total or zero measure.

Example 6.2

Let us consider the unit circle in the complex plane

$$\mathbb{S}^1 := \{z \in \mathbb{C} : |z| = 1\},$$

and let $\mu \in \mathcal{M}^1(\mathbb{S}^1)$ be the normalized Lebesgue measure, e.g. the measure characterized by

$$\mu(I) := \frac{\text{length}(I)}{2\pi}, \quad \forall \text{ interval } I \subset \mathbb{S}^1.$$

Given $\alpha \in \mathbb{S}^1$ we define

$$\varphi_\alpha(z) := \alpha z,$$

i.e. a rotation with angle $\arg(\alpha)$, this defines a measurable map from \mathbb{S}^1 to itself. Since it is only a rotation, length is preserved, therefore it is a μ -invariant map. The interesting fact is that it is μ -ergodic if and only if α is not a root of unity, i.e. if there is no $n \in \mathbb{N}$ such that $\alpha^n = 1$, the proof of this fact is outlined in the exercises.

Example 6.3

We can generalize the preceding example to the n -dimensional torus:

$$\mathbb{T}^n := \underbrace{\mathbb{S}^1 \times \dots \times \mathbb{S}^1}_n.$$

The measure $\hat{\mu} \in \mathcal{M}^1(\mathbb{T}^n)$ we consider is the *product measure* of the measure μ defined in the exaple 6.2 above, e.g.

$$\hat{\mu}(B_1, \dots, B_n) := \prod_{i=1}^n \mu(B_i), \quad \forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{S}^1).$$

Given $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{T}^n$ we define $\hat{\varphi}_\alpha : \mathbb{T}^n \rightarrow \mathbb{T}^n$ by

$$\hat{\varphi}_\alpha(z_1, \dots, z_n) := (\alpha_1 z_1, \dots, \alpha_n z_n),$$

which makes $\hat{\mu}$ a $\hat{\varphi}_\alpha$ -invariant measure.

For a bigger number of examples on ergodic measures in dynamical systems we refer to [10, §4.2, p. 146].

The definitions we have made so far are very general for our modest purpose, we want to understand the ergodic decomposition in the particular case of differentiable dynamical systems. So, even if our definitions are too general, later we'll restrict ourselves to dynamical systems in differentiable manifolds.

A **dynamical system** is a couple (M, φ) , where M is a topologic space and $\varphi : M \rightarrow M$ is an onto map, considering the iterates $\varphi, \varphi \circ \varphi, \dots$ one obtain a family of transformations of M called a **discrete dynamical system**.

A **continuous dynamical system** is a couple $(M, \Phi := \{\varphi^t\}_{t \in I})$, such that $I \subset \mathbb{R}$ is an open interval for every $t \in I$, $\varphi^t : M \rightarrow M$ is an onto map with the following properties:

- i) $\varphi^t \circ \varphi^s = \varphi^{t+s}, \quad \forall t, s \in \mathbb{R}$, and
- ii) $\varphi^0(x) = x, \quad \forall x \in M$.

In this case we call Φ a **semi-flow**. When the map φ^t is bijective for every $t \in I$ we call Φ a **flow**.

Example 6.4

Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ an invertible linear map, its iterates define a discrete dynamical system.

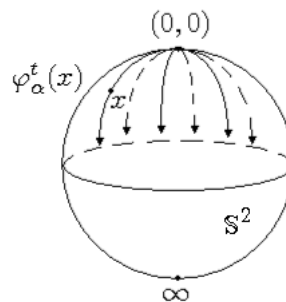


Figure 2: Flow φ_α^t on the 2-sphere \mathbb{S}^2 .

Example 6.5

Consider the euclidean plane \mathbb{R}^2 and define $\varphi^t(x) := (t+1)x$ for $t \in \mathbb{R}^+$, this is a continuous dynamical system such that the set

$$\mathcal{O}_x := \{\varphi^t(x) : t \in \mathbb{R}^+\},$$

called the **orbit** of $x \in M$, is a ray through x with origin at the point $(0, 0)$ for every $x \in M$. If we take the one point compactification of \mathbb{R}^2 , i.e. $\mathbb{R}^2 \cup \{\infty\} = \mathbb{S}^2$, defining $\varphi^t(\infty) = \infty$ for every $t \in \mathbb{R}$, we get a differentiable dynamical system on \mathbb{S}^2 whose orbits are as in the figure 2.

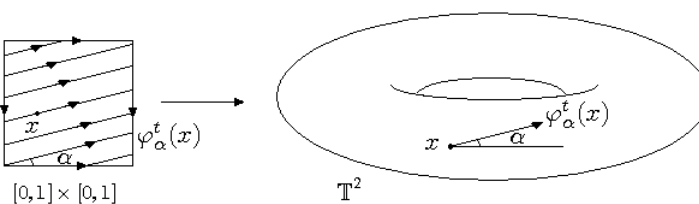


Figure 3: The flow φ_α^t on the 2-torus \mathbb{S}^2 .

Example 6.6

The 2-torus \mathbb{T}^2 can be seen as the quotient $\mathbb{R}^2/\mathbb{Z}^2$ where a system of representants can be taken in the unit square $[0, 1] \times [0, 1]$ with opposite sides identified. Then, given $\alpha \in [0, \pi/2]$ we can define a flow φ_α^t by $\varphi_\alpha^t(x, y) = (x - tx, y - t\alpha x)$ for $t \in \mathbb{R}$, this is a flow on \mathbb{R}^2 such that the orbit of a point $v \in \mathbb{R}^2$ is the line through v with slope α (see figure 3), and, when passing to the quotient $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is a finite set of parallel lines with slope α if it is rational, otherwise is a unique line dense in the unit square. In the torus we have a differentiable flow whose orbits are “twisting around” the torus.

A big amount of examples of dynamical systems can be found in [10, 6].

Given a discrete or continuous dynamical system (M, Φ) we denote by $\mathcal{M}_\Phi(M)$ the set of Φ -invariant measures on M .

As a first remark we have that $\mathcal{M}_\Phi(M)$ is weak*-compact since it is a closed subset of $\mathcal{M}^1(M)$ with the weak* topology and this last set is

a weak*-compact set as we have seen in proposition 4.2.

The second important remark is that proposition 4.2 also tells give us the identification $\mathcal{M}(M) \cong C^*(M)$ and then $\mathcal{M}(M)$ is a vector space. Therefore it makes sense to speak of convex sets, in particular on the convexity of $\mathcal{M}_\Phi(M) \subset \mathcal{M}(M)$ so we have the following proposition.

|| **Proposition 6.7** *The set $\mathcal{M}_\Phi(M)$ is a convex set and its extreme point set $\text{ex}(\mathcal{M}_\Phi(M))$ is the set of Φ -ergodic measures.*

Proof. The first part is a simple exercise, for the second part see [15, p. 75, Proposition 12.4]. The idea is proceed by contradiction, if an element of $\text{ex}(\mathcal{M}_\Phi(M))$ is not ergodic then it can be written as a convex combination of ergodic measures, none of the coefficients being zero \square

Before continuing our study we give a beautiful example of a discrete dynamical system (C, φ) in which $\text{ex}(\mathcal{M}_\varphi(C))$ is not closed in $\mathcal{M}_\varphi(C)$. The proof of the statements is a guided exercise at the end of this section.

Example 6.8

Let $C = [0, 1] \times \mathbb{S}^1$, where we identify $\mathbb{S}^1 = \mathbb{R} \pmod{1}$. Define $\varphi : C \curvearrowright$ by

$$\varphi(x, y) = (x, x + y \pmod{1}).$$

Then C is a compact Hausdorff space in which $\text{ex}(\mathcal{M}_\varphi(C))$ is not closed as a subset of $\mathcal{M}_\varphi(C)$.

Now, proposition 6.7 says really something only when $\mathcal{M}_\Phi(M) \neq \emptyset$, however, for a nice enough dynamical system not only this set, but also $\text{ex}(\mathcal{M}_\varphi(C))$ are non-empty sets. We then impose some hypothesis to our dynamical systems.

□ **NOTE:** From now on we consider only dynamical systems (M, Φ) with the following two properties:

- i) M is a compact differentiable manifold of class C^∞ and
- ii) If $\Phi =: \varphi$ or $\{\varphi^i\}_{i \in I}$ when it is discrete or continuous, respectively, then φ or the φ^i 's are diffeomorphisms.

The following proposition⁴ says what we announced above.

⁴The statement of this proposition is a simplified version of 32.2 Proposition en [5, Vol II, p. 220], however there are other statements as in [9, p. 135 Theorem 4.1.1, p. 139 Theorem 4.1.11] or [15, p. 76].

Proposition 6.9 *Let (M, Φ) be a (discrete or continuous) dynamical system. Then always exist Φ -invariant and ergodic measures in $\mathcal{M}^1(M)$, i.e. $\mathcal{M}_\Phi(M)$ and $\text{ex}(\mathcal{M}_\Phi(M))$ are always non-empty sets.*

Proof. [5, Vol II, p. 220]. □

It then makes sense to speak of invariant and ergodic measures since they always exist. However, for applying the results of the previous section we need more structure on $\mathcal{M}_\Phi(M)$. We have the following theorem.

Theorem 6.10 *Let M be a locally compact Hausdorff space. Then $\mathcal{M}(M)$ equipped with the weak* topology, is a metrizable and separable (i.e. is metrizable and contains a countable dense set) space if and only if M has a countable basis.*

Proof. [5, Vol I, p. 219]. □

Since a compact differentiable manifold always have a countable basis we have that $\mathcal{M}_\Phi(M)$ is always a metrizable space.

We can state now the main theorem of this notes.

Theorem 6.11 (Ergodic Decomposition) *Let (M, Φ) be a differentiable dynamical system then every Φ -invariant measure is the barycenter of a measure defined on $\mathcal{M}_\Phi(M)$ whose support is contained in the Φ -ergodic measures.*

Proof. Proposition 6.9 implies $\mathcal{M}_\Phi(M)$ and $\text{ex}(\mathcal{M}_\Phi(M))$ are non-empty sets. Applying proposition 6.7 and theorem 6.10 the set $\mathcal{M}_\Phi(M)$ is a metrizable weak*-compact convex set, the result follows by applying Choquet's theorem (Theorem 4.5) to this set. □

As in Choquet's theorem the conclusion of the theorem can be restated as:

For every measure $\mu \in \mathcal{M}_\Phi(M)$ there exists a measure $\eta \in \mathcal{M}^1(\text{ex}(\mathcal{M}_\Phi(M)))$ such that

$$\mu = \int_{\text{ex}(\mathcal{M}_\Phi(M))} z d\eta(z).$$

Then, since “an integral is a limit of discrete atoms”, we can see a Φ -invariant measure μ as *an approximation of convex combinations of*

ergodic measures. This is stated in the following theorem, with which we end this section.

Theorem 6.12 *Let (M, Φ) be a differentiable dynamical system, then each Φ -invariant measure can be arbitrarily well approximated in the space $\mathcal{M}^1(M)$ by a finite convex combination of ergodic measures.*

Exercises 4

1. On example 6.2 show with detail that φ_α is μ -invariant. The next two incises show that if α is not a root of unity then μ is ergodic.
 - i) The set $\{f_n(x) := x^n\}_{n \in \mathbb{N} \cup \{0\}}$ is an orthonormal basis in the space L^2 .
 - ii) If $f \in L^2$ write its Fourier series $f = \sum_n a_n f_n$ and define $U(f) = f \circ \varphi_\alpha$, use this functional to prove that f is Φ -invariant if and only if f is constant on the space L^2 . Conclude using theorem 5.1

2. Define the map $\varphi : [0, 1] \circlearrowleft$ as

$$\varphi(x) = \begin{cases} 1 & \text{if } x = 0 \\ \frac{x}{2} & \text{otherwise} \end{cases}$$

Show that $\mathcal{M}_\varphi = \emptyset$.

3. Let (X, μ, \mathcal{B}) be a measure space $A \in \mathcal{B}$ such that $\mu(A) > 0$ and $\varphi : X \circlearrowleft$ be a μ -invariant map. Define

$$\mu_A(B) := \frac{\mu(A \cap B)}{\mu(A)},$$

and for each $x \in A$ define

$$n(x) := \min\{n \in \mathbb{N} : \varphi^n(x) \in A\}.$$

Show that $\varphi_A(x) := \varphi^{n(x)}(x)$ defines a map from A to itself that is μ_A -invariant. The map φ_A is called the **first return map** induced by φ on the set A .

4. In this exercise we show the statement of example 6.8.

For each $n \geq 1$ let μ_n be the measure assigning the number $1/n$ to the points $(1/n, k/n) \in C = [0, 1] \times \mathbb{S}^1$ for $k = 0, 1, \dots, n - 1$ and zero to the others. Prove the following:

- i) The measure μ_n is an extreme point of the set $\mathcal{M}_\varphi(C)$.
- ii) The sequence $\{\mu_n\}$ converges in the weak* topology to the Lebesgue measure on $\{0\} \times \mathbb{S}^1$.
- iii) Verify that $\mathcal{M}^1(\{0\} \times \mathbb{S}^1) \subset \mathcal{M}_\varphi(C)$ and therefore μ cannot belong to the set $\text{ex}(\mathcal{M}_\varphi(C))$.

7 Application: transverse invariant measures and foliated cycles in one-dimensional foliations

We finish these notes with an application of the ergodic decomposition theorem to the study of foliated cycles, which attracted the author to this subject and encouraged him to write this paper. In this section the formal definitions are mostly omitted.

In 1957 Schwartzmann [18] introduces the notion of *foliated cycles* under the name of *asymptotic cycles*, later Dennis Sullivan [19] in 1976 publishes a generalization of such asymptotic cycles and call them *structure cycles*, he shows that, in the case of foliated manifolds, these structure cycles correspond bijectively to the so called *transverse invariant measures*. In this same case of foliations, the Schwartzmann-Sullivan cycles are known under the name of **foliated cycles**. In one-dimensional foliations we can show that foliated cycles can be approximated by a “sum of circles nearly tangent to the foliation” as is stated in [19, p. 246. Proposition II.25].

In 1982 David Fried [7] introduces the concept of *homology directions* (this was also treated in Schwartzmann’s article but with a different name) which are the circles nearly tangent to the foliation of which Sullivan talked about.

Our goal in this section is to expose how the ergodic decomposition is useful to show how foliated cycles can be approximated by convex combinations of homology directions.

Some definitions are treated in abstract, however, the main reference is [1], where the author studies the theory of foliated cycles and invariant transverse measures and some of their applications.

In dimension one a (non-singular) **foliation** of a compact manifold M is a partition of M by disjoint differentiable curves, called the **leaves** of the foliation, these are given by the orbits of a **complete, non-singular** flow⁵ (just forget the time). We define a **periodic point** as a point $x \in M$ for which there exists $t_0 \in \mathbb{R}$, $t_0 > 0$ such that $\varphi^{t_0}(x) = x$, in this case

⁵A flow φ_t defined for every $t \in \mathbb{R}$ without fixed points, e.g. there is no point $x \in M$ with $\varphi_t(x) = x$ for every $t \in \mathbb{R}$.

the orbit of x is a closed curve contained in M .

i NOTE: In what follows, \mathcal{F} will be a foliation defined by a complete, non-singular flow Φ defined on a differentiable manifold M . And we speak of \mathcal{F} and Φ indistinctively.

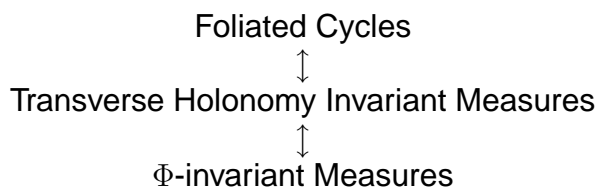
For a complete basic introduction to the theory of foliations the reader is referred to [4] and the references of [1].

Example 7.1

In the exampe 6.6 we obtain a foliation of \mathbb{R}^2 given by parallel lines with slope α , the leaves of the foliation are precisely these lines. Now, when passing to the quotient in the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, if $\alpha \in \mathbb{Q}$ this lines define a foliation of \mathbb{T}^2 by closed curves, and every point is periodic for the flow φ_α . If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ then the foliation induced on the torus is such that every leaf is an injective immersion of \mathbb{R} dense on \mathbb{T}^2 , i.e. there are no periodic points (see figure 3).

Foliated cycles⁶ as we mentioned above are identified with transverse holonomy invariant measures⁷, which, in the one-dimensional case can be identified with the set $\mathcal{M}_\Phi(M)$.

Therefore, for one-dimensional foliations we have the identifications:



If $x \in M$ is a periodic point its orbit \mathcal{O}_x defines a cycle in the singular one-homology of M , this one-homology group is isomorphic to DeRham’s homology (see [1, §1.1]) therefore is also a foliated cycle, moreover, the Φ -invariant measure associated to it is ergodic, i.e. a periodic orbit of Φ is an extreme point of $\mathcal{M}_\Phi(M)$.

Now we’ll see that every element of $\text{ex}(\mathcal{M}_\Phi(M))$ is a homology direction. Let $x \in M$ be a non-periodic point, then we can take a sequence of natural numbers $\{t_k\}_{k=1}^\infty$ such that $x_k := \varphi^{t_k}(x) \in U \subset M$, for every $k \geq 1$, where U is a foliated chart of M (e.g. $U \cong \mathbb{R}^{n-1} \times \mathbb{R}$ for

⁶See [1, §1.2].
⁷See [1, Apéndice A] for a brief introduction to this measures.

$n = \text{Dim}(M)$ see [4, §2] for more details). Define

$$\Gamma_k = \{\varphi^t(x) : 0 \leq t \leq t_k\}, \tag{10}$$

we can take a path γ_k joining $\varphi^{t_k}(x)$ to x obtaining then a closed curve (see figure 4).

$$\Gamma_k^* := \Gamma_k \cup \gamma_k \tag{11}$$

The length of the γ_k 's can be kept uniformly bounded, then $(1/t_k)\Gamma_k^*$ is a singular cycle.

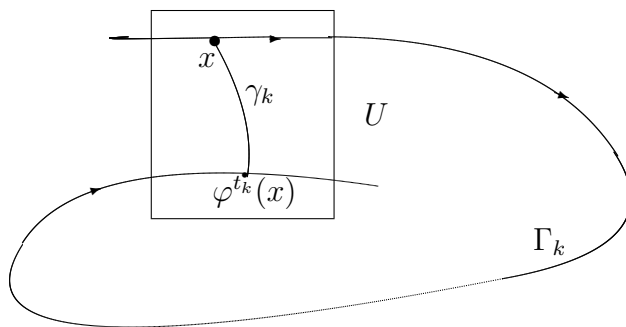


Figure 4: A long almost-closed orbit $\Gamma_k^* = \Gamma_k \cup \gamma_k$.

We then define an **homology direction** as a foliated cycle that is a closed orbit or a limit

$$\lim_{k \rightarrow \infty} \frac{1}{t_k} \Gamma_k^*,$$

where the Γ_k^* 's are defined as in 11.

The next proposition is explained in [1, §3.3].

Proposition 7.2 *If a foliated cycle corresponds to a (Φ) -ergodic measure, then it is a homology direction.*

This is,

Ergodic Measures \leftrightarrow Homology Directions,

and then, since foliated cycles can be identified to elements of the set $\mathcal{M}_\Phi(M)$, the following theorem follows immediately from theorem 9 and is the modern version of the result originally proved in [18] and then by Sullivan in [19]. Before stating it let's remark that the space $\mathcal{M}^1(M)$ can

be identified with the space of *positive distributions* (see [13, p. 20] and [1, §1.1 and §1.2]) or the space of *0-currents* which is denoted by \mathcal{D}'_0 , a foliated cycle belongs to the space D'_1 of *1-currents* when the foliation is one-dimensional.

|| **Theorem 7.3** *Every foliated cycle can be arbitrarily well approximated in the space D'_1 by a finite linear combination with positive coefficients of homology directions.*

References

- [1] **Albarrán, J.** *Ciclos para el estudio dinámico de foliaciones.* Graduate Thesis. Universidad de Guanajuato. Facultad de Matemáticas. *To appear.*
- [2] **Bishop, E.** and **Phelps, R.** *The support functionals of a convex set.* In V. Klee ed. *Convexity.* No.7 in Proc. Symposia in Pure Math.(1963) 27-35. Providence, RI. AMS.
- [3] **Bonnesen, T.** and **Fenchel, W.** *Theorie der Konvexen Korper.* Springer. Berlin, 1934.
- [4] **Camacho, C.** and **Lins Neto, A.** *Teoria Geométrica das Folheações.* Projeto Euclides, IMPA. Rio de Janeiro, 1978.
- [5] **Choquet, G.** *Lectures on Analysis.* Vols. I and II. Math. Lecture Notes Series. USA, 1969.
- [6] **Cornfeld-Fomin-Sinai.** *Ergodic Theory.* Springer. New York, 1982.
- [7] **Fried, D.** *The geometry of cross sections to flows.* *Topology* **21** (1982), 353-371.
- [8] **Holmes, R.** *Geometric Functional Analysis and its Applications.* Springer-Verlag. New York, 1975.
- [9] **Kantorovitz, S.** *Introduction to Modern Analysis.* Oxford University Press, 2003.
- [10] **Katok, A.** y **Hasselblatt, B.** *Introduction to the Modern Theory of Dynamical Systems.* Cambridge University Press, 1995.
- [11] **Köthe, G.** *Topologische Lineare Räume.* Springer. Berlin, 1966.

- [12] **Krein, M. y Milman, D.** *On extreme points of regular convex sets.* Studia Mathematica **9** (1940) 133-138.
- [13] **Lebeau, G.** *Théorie des distributions et analyse de Fourier.* Cours de L'Ecole Polytechnique. Palaiseau. France, 2001.
- [14] **Pedersen, G.** *Analysis Now.* Springer-Verlag. New York, 1989.
- [15] **Phelps, R.** *Lectures on Choquet's Theorem.* Lecture Notes in Mathematics 1757. Springer Verlag. Germany, 2001.
- [16] **Schaefer, H.** *Topological Vector Spaces.* Springer. New York, 1999.
- [17] **Schrenk, A.** *Das Theorem von Krein-Milman und schwache Topologie.* Electronic notes.
<http://www.math.ethz.ch/~theo/arne.ps>
- [18] **Schwartzmann, S.** *Asymptotic cycles.* Ann. of Math. **66** (1957), 270-284.
- [19] **Sullivan, D.** *Cycles for the dynamical study of foliated manifolds and complex manifolds.* Inv.Math. **36** (1976), 225-255.
- [20] **Valentine, F.A.** *Convex Sets.* Mc. Graw-Hill, New York, 1964.
- [21] **Vesely, L.** *Rappresentazioni Integrali.* Electronic notes.
<http://sauron.mat.unimi.it/~libor/AnSup2/choquet.ps>
- [22] **Walters, P.** *An Introduction to Ergodic Theory.* Springer. New York, 1982.
- [23] **Werner, D.** *Funktionanalysis.* Springer. Berlin, 1997.

亥

Jorge Albarrán

UNIVERSIDAD DE GUANAJUATO.
Facultad de Matemáticas.

✉ Centro de Investigación en Matemáticas, A.C.
AP. 402-36000 Guanajuato, Gto. México.

@ albarran@cimat.mx.