Ergodic Theory

Implications for classical mechanical systems

Stilianos Louca

July 1, 2009

Student research paper

University:	Friedrich Schiller University, Jena
Institute:	PAF, Theoretisch-Physikalisches Institut
Advising Professor:	Prof. Dr. rer. nat. habil. Gerhard Schäfer
Period:	March 1, 2009 - July 1, 2009

Contents

1 Abstract **1**

2 Introduction **2**

2.0.1 Definition: Measure spaces 2
2.0.2 Definition: Invariant measure 2
2.0.3 Definition: Automorphism 3
2.0.4 Definition: τ-induced operator 3
2.0.5 Definition: G-semiflow 3
2.0.6 Definition: Invariant functions and sets 4
2.0.7 Definition: Metrically isomorphic systems 4

3 Ergodic concepts **5**

3.1 Recurrence 5
3.1.1 Poincaré recurrence theorem 5
3.1.2 Corollary: Trajectories of automorphisms 5
3.1.3 Corollary: Average return time 6
3.1.4 Corollary: Approximating trajectories 7
3.1.5 The trapped gas 7
3.1.6 The Poincaré map 8
3.2 Birkhoff's ergodic theorem 10
3.2.1 Maximal ergodic theorem 10
3.2.2 Corollary: Maximal ergodic inequality 10
3.2.3 Birkhoff-Khinchin-Ergodic Theorem 11
3.2.4 Birkhoffs ergodic theorem for time-semi-flows 11
3.2.5 Corollary: Mean sojourn time 11
3.2.6 Von Neumann mean ergodic theorem 12
3.3 Ergodicity 14
3.3.1 Definition: Ergodic semi-flows 14
3.3.2 Lemma about ergodic semi-flows 14
3.3.3 Theorem: Invariant functions and ergodic semi-flows 14
3.3.4 Theorem: Characterization of ergodic G-flows 15
3.3.5 Corollary about ergodic G-flows in topological spaces 17
3.3.6 Theorem: Characterization of ergodicity for time- and iteration-semi-flows 17
3.3.7 Lemma: Ergodicity in metrically isomorphic systems 20
3.3.8 Notes on measure-invariance of ergodicity 21
3.4 Strict ergodicity in topological spaces 21
3.4.1 Theorem: Existence of invariant Borel-measures 22
3.4.2 Theorem: Existence of ergodic Borel measures 23
3.4.3 Definition: Strictly ergodic semi-flow 23
3.4.4 Theorem: Strict ergodicity of homeomorphisms 24
3.4.5 Corollary for strictly ergodic homeomorphisms 25
3.5 Mixing 25
3.5.1 Definition: Relaxing systems 25
3.5.2 Lemma about system relaxations 26
3.5.3 Definition: Mixing G-semi-flows 26
3.5.4 Theorem: Characterization of mixing G-semi-flows 27
3.5.5 Lemma about mixing systems 28
3.5.6 Lemma: Mixing in metrically isomorphic systems 29
3.6 Ergodicity in random variables 29
3.6.1 Preconsiderations 29
3.6.2 Lemma: Trajectories for random variables 30
3.7 System families and leafs 31
3.7.1 Lemma: Measure construction from leaf-measures 31
3.7.2 Theorem about ergodic decompositions 32
3.7.3 Definition: Product space 33
3.7.4 Lemma about ergodicity in product spaces 33

3.7.5 Lemma: Connection between separable Hamilton functions and product spaces 33

4 Flows on manifolds **35**
4.1 Measures on manifolds 35
4.1.1 Definition: Lebesgue-measure on $\mathcal{C}^\infty$-manifolds 35
4.1.2 Notes on metrizability of manifolds 35
4.1.3 Measure-evolution under flows 36
4.1.4 Liouville's Theorem 36
4.1.5 Corollary to the Liouville-Theorem 36
4.1.6 Specialization for Hamiltonian flows 37
4.1.7 Lemma: Equilibrium densities for Hamilton flows 38
4.1.8 Invariant measure on energy-leafs 38
4.2 Flows on the torus 39
4.2.1 The torus as a manifold 39
4.2.2 The translation group on the torus 39
4.2.3 Theorem: Ergodicity of $\{\tau_\omega^n\}$ 40
4.2.4 Theorem: Ergodicity of $\{\tau_\omega^t\}$ 42
4.2.5 Lemma: Properties of ergodic translation-flows 42
4.2.6 Notes on translation flows of the torus 43
4.3 Flows on Liouville-integrable systems 43
4.3.1 Identification with the torus 43
4.3.2 Harmonic oscillators 44
4.3.3 Bowl-travel 45
4.4 Flows on nearly Liouville-integrable systems 46
4.4.1 Definition: Diophantine vectors 46
4.4.2 Arnold's theorem 47

5 Billiards **48**
5.1 Construction of Billiards 48
5.1.1 Definition: Billiard table 48
5.1.2 Construction of the billiard flow 49
5.1.3 Construction of an invariant measure 49
5.1.4 The return map 50
5.1.5 Lemma about pathological points 51
5.1.6 Lemma: Mean free path of a billiard 51
5.2 Billiards in ellipses 52
5.2.1 Proposition: Convergence of focal trajectories in ellipses 53
5.2.2 Proposition: Restriction of non-focal trajectories in ellipses 54
5.2.3 Corollary: Non-ergodicity of ellipse-billiards 55

6 Summary **57**

A Appendix **58**
A.1 Proof of Birkhoff's ergodic theorem 58
A.1.1 Maximal ergodic theorem 58
A.1.2 Corollary: Maximal ergodic inequality 59
A.1.3 Birkhoff-Khinchin Ergodic Theorem 60
A.2 General topology & geometry 62
A.2.1 Lemma about continuous functions 62
A.2.2 Lemma: Induced volume forms on leafs 63
A.2.3 Lemma about reflections in ellipses 63
A.3 Measure theory 64
A.3.1 Corollary to the Poincaré recurrence theorem 64
A.3.2 Lemma: Characterization of measure preserving maps 64
A.3.3 Lemma about invariant sets 65
A.3.4 Lemma: Topological aspects of measure families 66
A.3.5 Lemma about mixing systems 67
A.3.6 Choquet theorem 68

Symbol reference

W.l.o.g.**:** **W**ithout **l**oss **o**f **g**enerality.

A.e.**:** **A**lmost **e**verywhere.

z^***:** Complex conjugate of $z \in \mathbb{C}$.

$\mathbb{R}^*$**:** Extended real axis, $\mathbb{R}^* := \mathbb{R} \cup \{\pm\infty\}$.

$\mathbb{C}^*$**:** Extended complex plane, $\mathbb{C}^* := \{x + iy : x, y \in \mathbb{R}^*\}$.

$\mathbb{R}_+ := [0, \infty)$.

Id**:** Identity map.

$(G, \circ, e)$**:** Group or monoid with neutral element e.

$(\tau^g)_{g \in G}$**:** G-semi-flow or G-flow for some (semi-)group G, see 2.0.5.

$\tau^G(x)$**:** Trajectory of point x under G-semi-flow $(\tau^g)_{g \in G}$, $\tau^G(x) := \{\tau^g(x)\}_{g \in G}$.

$\tau^G(A)$**:** Trajectory of set A under G-semi-flow $(\tau^g)_{g \in G}$, $\tau^G(A) := \bigcup_{g \in G} \tau^g(A)$.

Π**:** Natural projection on phase-space, $\Pi(\mathbf{q}, \mathbf{p}) := \mathbf{q}$.

$\Pi\left(\tau^G x\right)$**:** Spatial trajectory of x under the G-semi-flow $(\tau^g)_{g \in G}$.

$d(\cdot, \cdot)$**:** Metric.

$\langle \cdot, \cdot \rangle$**:** Scalar product.

P_V**:** Projector on closed subspace V.

f^+**:** Positive part of a function f, $f^+ := \max\{f, 0\}$.

f^-**:** Negative part of a function f, $f^- := \max\{-f, 0\}$.

$(M, \mathcal{M}, \mu)$**:** Measure space with σ-algebra $\mathcal{M} \subset \mathscr{P}(M)$ and measure μ.

$\sigma(\mathcal{A})$**:** Smallest σ-algebra containing the family of sets $\mathcal{A}$.

$\delta(\mathcal{A})$**:** Smallest Dynkin-system containing the family of sets $\mathcal{A}$.

$\mathcal{M} \cap A$**:** Induced σ-algebra on set $A \in \mathcal{M}$, $\mathcal{M} \cap A := \{B \cap A : B \in \mathcal{M}\}$.

$\mu \otimes \nu$**:** Product measure of μ and ν.

$\mathcal{M} \otimes \mathcal{N}$**:** Product σ-algebra of $\mathcal{M}$, $\mathcal{N}$: $\mathcal{M} \otimes \mathcal{N} := \sigma\left(\{A \times B : A \in \mathcal{M}, B \in \mathcal{N}\}\right)$.

$\frac{d\mu}{d\nu}$**:** Density of measure μ with regard to measure ν.

$\nu \ll \mu$**:** Measure ν is *absolutely continuous* to μ, that is, $\mu(A) = 0$ implies $\nu(A) = 0$.

μ_τ**:** Image measure of map τ under measure μ, $\mu_\tau(A) := \mu\left(\tau^{-1}(A)\right)$.

$\mu(f)$**:** For measure μ and measurable function f: $\mu(f) := \int f \, d\mu$.

$\langle f \rangle_\mu$**:** Same as $\mu(f)$, but mainly for probability-measures μ.

δ_E**:** Dirac measure at point E : $\delta_E(A) = 1 \;\Leftrightarrow\; E \in A$.

1_A**:** Indicator function for set A, that is, $1_A(x) = 1$ for $x \in A$ and $1_A(x) = 0$ for $x \notin A$.

A^c**:** Complement for some set $A \in M$, $A^c := M \setminus A$.

$A \triangle B$**:** Symmetrical difference between sets A, B : $A \triangle B := (A \cup B) \setminus (A \cap B)$.

$\mathrm{cl}(A)$**:** Closure of set A.

$V^\perp$: Orthogonal space to linear subspace V.

p a.e.: Almost everywhere, $\mu(\{x : \neg p(x)\}) = 0$.

p for a.e. $x \in A$: There exists a set $\tilde{A} \subset A$, $\tilde{A} = A(\text{mod}0)$ such that p holds on $\tilde{A}$.

$\{f = g\} := \{x : f(x) = g(x)\}$.

$\mathscr{O}(M)$: Topology (open sets) of topological space M.

$\mathscr{O}(M) \cap A$: Induced topology on subset $A \subset M$ of a topological space M, $\mathscr{O}(M) \cap A := \{B \cap A : B \in \mathscr{O}(M)\}$.

$\mathscr{B}(M)$: σ-algebra of Borel sets of topological space M, $\mathscr{B} := \sigma(\mathscr{O}(M))$.

$A = B(\text{mod}0)$: $1_A = 1_B$ almost everywhere, that is, $\mu(A \triangle B) = 0$.

$\mathcal{C}(M)$: Linear space of continuous complex functions on M.

$\|\cdot\|_\infty$: Supremum-norm, $\|f\|_\infty := \sup\{f(x) : x \in M\}$.

$\|\cdot\|_p$: L_p-norm.

$L_p(M, \mathcal{M}, \mu)$: Linear space of L_p-integrable functions on $(M, \mathcal{M}, \mu)$.

$B^o_\varepsilon(x)$: Open ε-ball, $B^o_\varepsilon(x) := \{y : d(x, y) < \varepsilon\}$.

B^n_ε: Closed ε-ball in $\mathbb{R}^n$.

S^n: n-dimensional unit-circle, that is, $S^n = \partial B^{n+1}_1$.

S^1: 1-dimensional unit-circle, $S^1 \simeq \mathbb{R}/\mathbb{Z}$.

T^n: n-dimensional torus, $T^n = S^1 \times \cdots \times S^1 \simeq \mathbb{R}^n/\mathbb{Z}^n$.

λ_{T^n}: Standard Lebesgue measure on torus T^n.

$\lambda_{\mathbb{R}^n}$: Lebesgue measure in $\mathbb{R}^n$.

$\lambda_{\mathbf{x}}$: Lebesgue measure on some manifold, induced by coordinates $\mathbf{x}$.

$d_{T^n}(\cdot, \cdot)$: Standard metric on torus T^n, see 4.2.1.

$\omega(\cdot, \cdot)$: Symplectic form.

T_qM: Tangent-space of manifold M at point q.

T^*_qM: Co-tangent-space of manifold M at point q.

TM: Tangential bundle of manifold M.

T^*M: Co-tangential bundle of manifold M.

(M, g): Riemannian manifold with metric g.

(M, g, J): Riemannian manifold with the isomorphism $J : TM \to T^*M$, defined by $JX := g(X, \cdot)$.

$(M, g, \tilde{g})$: Riemannian manifold with inverse metric $\tilde{g}$, that is, $\tilde{g}(a, b) = g(J^{-1}a, J^{-1}b)$ for $a, b \in T^*M$.

V^M: Volume-form on M induced by Riemann-metric g.

$V^{\partial M}$: Volume form on ∂M induced by Riemann-metric g.

$\text{vol}_{V^M}(A)$: Volume of set A under volume-form V^M: $\text{vol}_{V^M}(A) := \int_A V^M$.

∂_i: Coordinate vector field on manifold.

(M, ω): Symplectic manifold with symplectic form ω.

$(M, d\boldsymbol{\vartheta} \wedge d\mathbf{s})$: Symplectic manifold with symplectic coordinates $\boldsymbol{\vartheta}, \mathbf{s}$.

H: Hamilton-function.

X_H: Hamilton-vector field, defined by $\omega(X_H, \cdot) = dH$.

φ_X^t: (Semi-)flow for some vector field X.

$\mathcal{R}_X$: Poincaré return-map for vektor-field X, see 3.1.6.

$A_n f$: Iterated average of some function f under some map τ, $A_n f := \sum_{k=0}^{n-1} f \circ \tau^k$.

$M_n^A f$ $:= \max\{A_1 f, ..., A_n f\}$.

$S_n f$ $:= n \cdot A_n f$.

$M_n^S f$ $:= \max\{S_1 f, .., S_n f\}$.

$M \simeq N$: M diffeomorphic to N.

1 Abstract

The theory of dynamical systems has in the past 60 years evolved from an empirical, observation-based science, into an abstract, powerful mathematical tool to describe problems ranging from physics, biology up to number and information theory. Particularly in physics, dynamical systems are typically described by some set of differential equations, which if solved, describe the evolution of the system up to any time in the future and not seldomly in the past.

Questions arise about, not only the explicit solutions, but also *holist* aspects like distribution and denseness of trajectories in phase space, existence of periodic behavior and limiting of solutions to certain *time-invariant state-sets*. The mathematical framework built around these questions is coined *ergodic theory*, and incorporates disciplines from all branches of mathematics, including group theory, measure theory and differential geometry.

This paper is meant as an introduction to current developments in this field and an elaboration on deeper implications for Hamiltonian systems. It begins with an introduction and characterization of key-concepts like ergodicity, strict-ergodicity, mixing and relaxing systems and the existence of ergodic equilibrium measures. We proceed with elaborating on certain concepts emerging in systems described on smooth manifolds (e.g. symplectic), restricting the analysis mainly to conservative Hamiltonian systems. Finally, an example application of the developed concepts is given for billiards, a popular model attracting the attention of mathematicians as well as physicists.

2 Introduction

In the theory of dynamical systems, a system's state is typically abstracted as a point in *phase space*[8], in general some topological, and not seldomly, measurable space. Its evolution is described by *flows*, which, sometimes depending on some free parameter (i.e. time), *act* upon the phase-space's states, *moving* them as *time passes*[17]. In case of autonomous systems, i.e. described by some time-invariant law, this flow is abstracted as a group action, that is, a family of mappings within the phase-pace with group-like properties[7].

Dynamical systems arising in nature, usually come with the burden of *uncertainty*. Though the laws governing a system might be explicitly given, its *initial state* is often fully or partially hidden from observers, either due to lack of means to identify it, or simply due to its sheer complexity and huge number of parameters needed to describe it. A so called *probability distribution* is thus additionally introduced to the phase space, expressing in a sense the information, or better, the lack of it, given for the system's state. As the system naturally evolves, this probability distribution for the system's states might change, following the *flow* of the system within the phase space, at all times expressing what information we may have about the system's state. This evolution of the probability distribution is naturally given by the image-measure, induced by the flow-action within the phase-space.[8]

Equilibrium distributions are special distributions, characterized by their *invariance* to the system's flow. As time passes, they remain constant, thus resulting in all expectations one might have about the system being constant with time. It is with respect to these distributions, that questions like *ergodicity* and *mixing* arise[3].

We shall begin with a brief introduction to the terminology and basic definitions to be used in the rest of the article.

2.0.1 Definition: Measure spaces

Let $M \neq \emptyset$ be some set and $\mathcal{M} \subset \mathscr{P}(M)$ some σ-algebra[1] on M. Then $(M, \mathcal{M})$ is called a *measurable space*.
If further $\mu : \mathcal{M} \to [0, \infty]$ is a measure[2] on $(M, \mathcal{M})$, we call $(M, \mathcal{M}, \mu)$ a *measure space*. We call μ *finite*, if $\mu(M) < \infty$ and *σ-finite*, if there exists a countable partition $(A_n)_n \subset \mathcal{M}$ of M with $\mu(A_n) < \infty$.
We call a map $f : M \to N$ between measurable spaces $(M, \mathcal{M})$, $(N, \mathcal{N})$ *measurable*, if for $B \in \mathcal{N}$ it follows $f^{-1}(B) \in \mathcal{M}$.
We say $A \subset M$ is *measurable* if $A \in \mathcal{M}$ and call A a *nullset*, if $\mu(A) = 0$.
For two measurable sets $A, B \in \mathcal{M}$ we define:

$$A = B(\text{mod}0) \quad :\Leftrightarrow \quad \mu(A \triangle B) = 0$$

and say A is *μ-almost equal to B*.
For two measurable functions f, g on M we define:

$$f = g(\text{mod}0) \quad :\Leftrightarrow \quad \mu\left(\{x : f(x) \neq g(x)\}\right) = 0$$

and say f is *μ-almost equal to g*.
We say $(M, \mathcal{M})$ is *topological*, if M is a topological space and $\mathcal{M}$ contains the topology $\mathscr{O}(M)$ of M. By $\mathscr{B}(M)$ we denote the Borel-σ-algebra of M. See more in [6].

2.0.2 Definition: Invariant measure

Let $(M, \mathcal{M}, \mu)$ be a measure space and $\tau : M \to M$ measurable. Then μ is called *τ-invariant* if the image-measure

$$\mu_\tau(A) := \mu(\tau^{-1}(A))$$

[1]A family of sets $\mathcal{M} \subset \mathscr{P}(M)$ is called a σ-algebra if:

1. $M \in \mathcal{M}$
2. From $A \in \mathcal{M}$ follows $A^c \in \mathcal{M}$
3. For any sequence $(A_n) \subset \mathcal{M}$ one has $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$

[2]A non-negative function $\mu : \mathcal{M} \to M$ on a σ-algebra $\mathcal{M}$ is called a *measure*, if $\mu(\emptyset) = 0$ and for disjoint $A_n \in \mathcal{M}$ follows $\mu\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n)$

is equal to μ. The map τ is called *measure preserving*, or an *endomorphism* on $(M, \mathcal{M}, \mu)$.
Note: For bijective $\tau : M \to M$, τ-invariance of μ is equivalent to τ^{-1}-invariance.

2.0.3 Definition: Automorphism

Let $(M, \mathcal{M}, \mu)$ be a measure space and $\tau : M \to M$. Then τ is called an *automorphism*, if:

- it is invertible,
- both τ and τ^{-1} are measurable,
- $\forall\ A \in \mathcal{M} : \mu(\tau(A)) = \mu(A)$.

Note that all automorphisms are measure preserving.[1]

2.0.4 Definition: τ-induced operator

Let M be an arbitrary set and $\tau : M \to M$. Then the operator T on the space of (real or complex) functions on M defined as

$$Tf := f \circ \tau$$

is called the *τ-induced operator*.[1]

Note: If $\tau : M \to M$ is measure preserving on the measure space $(M, \mathcal{M}, \mu)$, then by lemma A.3.2 the operator $T : L_p \to L_p$ is an isometry[3] for any $0 < p < \infty$. If τ is additionally bijective, then $T : L_2 \to L_2$ is unitary.

2.0.5 Definition: G-semiflow

Let $(G, \circ)$ be a monoid[4] and $(\tau^g)_{g \in G}$ a family of maps on the set M with $\tau^g \circ \tau^h = \tau^{g \circ h}$ for any $g, h \in G$ and $\tau^0 = \mathrm{Id}$. Then $(\tau^g)_{g \in G}$ is called a *G-semi-flow*.
Note:

- $(\{\tau^g\}_{g \in G}, \circ)$ is a monoid.
- If G is abelian, then $\{\tau^g\}$ is abelian.
- If G is a group, then $(\{\tau^g\}, \circ)$ is a group, all τ^g are bijections and $\tau^{-g} = (\tau^g)^{-1}$. We then call $(\tau^g)_{g \in G}$ a *G-flow*.

Further conventions:

- In the context of measurable spaces $(M, \mathcal{M})$ and $(G, \mathcal{G})$ we further demand that $\tau^g : G \times M \to M$ are measurable in $(G \times M, \mathcal{G} \otimes \mathcal{M})$, as well as measurability of any set of the form $\{\tau^g A\}_{g \in G'}$, $G' \in \mathcal{G}$, $A \in \mathcal{M}$.
- A G-semi-flow is *measure preserving* if all its elements are measure preserving.
- The set of points $\{\tau^g(x)\}_{g \in G}$ for some $x \in M$ is called the *trajectory* of x along the G-semi-flow.
- If G is an ordered monoid[5], we call (τ^g) an *ordered G-(semi-)flow*. We call $\{\tau^g x\}_{g \geq 0}$ the *future (trajectory)* of x and $\{\tau^g x\}_{0 \geq g}$ the *history (trajectory)* of $x \in M$.
- If τ^g is only defined for $g \in A \subsetneq G$, we call $(\tau^g)_{g \in A}$ a *partial* G-semi-flow. In that case, the rule $\tau^g \circ \tau^h = \tau^{g \circ h}$ shall hold wherever it makes sense.
- If $G = \mathbb{R}_+$ or $G = \mathbb{R}$ we sometimes call (τ^g) a *time-(semi)-flow*.

[3] An operator $T : V \to W$ between two normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ is called an isometry $:\Leftrightarrow\ \|Tv\|_W = \|v\|_V\ \ \forall\, v \in V$
[4] Semi-group with neutral element.
[5] A monoid $(G, \circ)$ is called *ordered*, if it is equipped with a *translation invariant* order $\geq$, that is, for $a, b, g \in G$ and $a \leq b$ follows $a + g \leq b + g$.

- If $G = \mathbb{N}$ or $G = \mathbb{Z}$ and (τ^n) is generated by iterating over some map $\tau : M \to M$, we call (τ^g) an *iteration-(semi)-flow*.
- For vector fields X on some smooth manifold M, we write $(\varphi_X^t)_t$ for the induced flow.

2.0.6 Definition: Invariant functions and sets

Let M be an arbitrary set and $\tau : M \to M$. A function f on M is τ*-invariant*, if $f \circ \tau = f$. A set $A \subset M$ is called τ*-invariant*, if 1_A is τ-invariant, that is, $\tau^{-1}(A) = A$.
Now let $\{\tau^\alpha\}_{\alpha\in\mathcal{A}}$ be a set of maps $\tau^\alpha : M \to M$. A function f (set A) is $\{\tau^\alpha\}$*-invariant*, if for any $\alpha \in \mathcal{A}$ it is τ^α-invariant.
Note: Any τ-invariant function f (set), where τ bijective, is also τ^{-1}-invariant.

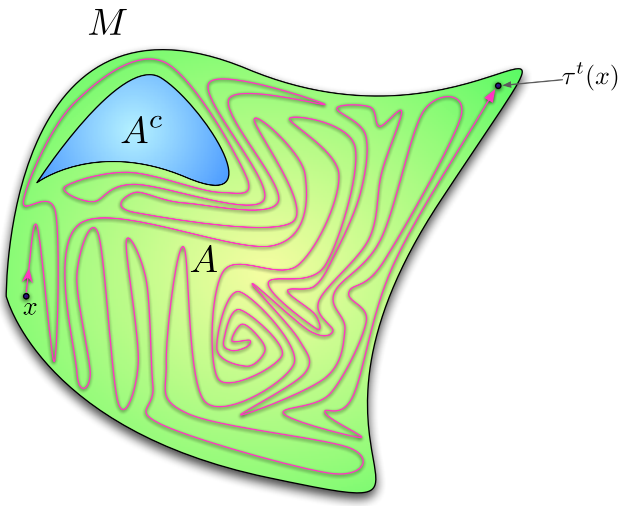

Figure 1: On invariant sets of flows. The set A above is (τ^t)-invariant, meaning that every trajectory is either completely included in or completely excluded from A.

2.0.7 Definition: Metrically isomorphic systems

Let $(M, \mathcal{M}, \mu, (\tau^g)_{g\in G})$ and $(N, \mathcal{N}, \nu, (\lambda^g)_{g\in G})$ be measure spaces with G-semi-flows and $\varphi : M \to N$ a measure preserving ($\nu = \mu_\varphi$) bijection that commutes with the semi-flows, that is,

$$\lambda^g = \varphi \circ \tau^g \circ \varphi^{-1} \quad , \quad g \in G \ .$$

Then the systems are called *metrically isomorphic (by* φ*)*.[3]

Note:

- (τ^g) is measure preserving $\Leftrightarrow$ (λ^g) is measure preserving, since $\nu_\lambda = (\mu_\tau)_\varphi$.
- A function f on M is $\{\tau^g\}$-invariant $\Leftrightarrow$ $f \circ \varphi^{-1}$ is $\{\lambda^g\}$-invariant, since

$$\left(f \circ \varphi^{-1}\right) \circ \lambda^g = f \circ \varphi^{-1} \circ \varphi \circ \tau^g \circ \varphi^{-1} = f \circ \tau^g \circ \varphi^{-1} \ .$$

3 Ergodic concepts

3.1 Recurrence

One of the first concepts studied in dynamical systems and later in ergodic theory, is that of recurrence, a question originating in celestial mechanics and the study of periodic orbits. One of the first milestones is the recurrence theorem of Poincaré about measure preserving maps, to be discussed below. Its implications are intriguing as well as fundamental, and at first presented a contradiction to the *irreversibility* of many physical processes. See more in [28].

3.1.1 Poincaré recurrence theorem

Let $(M, \mathcal{M}, \mu)$ be a finite measure space and $\tau : M \to M$ measure preserving. Then for every set $A \in \mathcal{M}$, almost all points $x \in A$ *return* through the trajectory $\{\tau^n(x)\}_{n\in\mathbb{N}_0}$ infinitely often to A. In other words:

$$\mu\left(\{x \in A \mid \exists\, n_x \in \mathbb{N} : \forall\, n \geq n_x : \tau^n x \notin A\}\right) = 0 \quad \forall\, A \in \mathcal{M}\ . \tag{3.1.1.1}$$

Proof: We shall adopt the proof found in [26]. Let

$$A_n := \bigcup_{k\geq n} \overbrace{\tau^{-k}(A)}^{(\tau^k)^{-1}(A)} \quad,\ n \in \mathbb{N}_0$$

be the set of all points that come to A after at least n iterations. Then $A_m \subset A_n$ for $m \geq n$ and $\tau^{-n}(A_0) = A_n$, which implies

$$\mu(A_0) \overset{\mu_\tau = \mu}{=} \mu\left(\tau^{-n}(A_0)\right) = \mu(A_n)\ . \tag{3.1.1.2}$$

Therefore

$$\mu\left[A \setminus \left(\bigcap_{n\in\mathbb{N}_0} A_n\right)\right] = \mu\left[\bigcup_{n\in\mathbb{N}_0} (A \setminus A_n)\right] \leq \sum_{n=0}^{\infty} \mu(\underbrace{A \setminus A_n}_{\substack{\subset A_0\setminus A_n \\ \text{since} \\ A\subset A_0}}) \leq \sum_{n=0}^{\infty} \underbrace{\mu(A_0 \setminus A_n)}_{\substack{\mu(A_0)-\mu(A_n) \\ \text{since} \\ A_n\subset A_0 \\ \text{and } \mu \text{ finite}}} \overset{(3.1.1.2)}{=} 0\ .$$

□

3.1.2 Corollary: Trajectories of automorphisms

Let $(M, \mathcal{M}, \mu)$ be a finite measure space and $\tau : M \to M$ an automorphism. Then the *history*

$$\tau^{-\mathbb{N}_0}(A) := \bigcup_{n\in\mathbb{N}_0} \tau^{-n}(A)$$

is almost equal to the *future*

$$\tau^{\mathbb{N}_0}(A) := \bigcup_{n\in\mathbb{N}_0} \tau^{n}(A)$$

of any set $A \in \mathcal{M}$.

Proof: Due to symmetry, it suffices to show only one of the two inclusions. Set $K := \tau^{N_0}(A) \setminus \tau^{-\mathbb{N}_0}(A)$ and suppose $\mu(K) > 0$. Then for some $n \geq 1$:

$$\mu\Big[\underbrace{\tau^n(A) \setminus \tau^{-\mathbb{N}_0}(A)}_{K_n}\Big] > 0$$

since otherwise K would be a countable union of null-sets and thus a null-set. By Poincaré we know that a.e. point $x \in K_n$ (since $\mu(K_n) > 0$, at least one) returns to K_n infinitely often, e.g at some time $m > n$. But then x also visits A, since $\tau^{m-n}(x) \in A$. This is a contradiction to the construction of K_n!
□

3.1.3 Corollary: Average return time

Let $(M, \mathcal{M}, \mu)$ be a finite measure space and $\tau : M \to M$ measure preserving. For any set $A \in \mathcal{M}$ let

$$R_A(x) := \inf\left(\{n \in \mathbb{N} : \tau^n(x) \in A\} \cup \{\infty\}\right)$$

be the *next visit time* of x in A. Then

$$\int_A R_A \, d\mu = \mu\Bigg[\underbrace{\bigcup_{n\in\mathbb{N}_0} (\tau^n)^{-1}(A)}_{\tau^{-\mathbb{N}_0}(A)}\Bigg] .$$

See more in [27].

Example: In the special case of a bijection τ the *history* $\tau^{-\mathbb{N}_0}(A)$ is almost equal to the future $\tau^{\mathbb{N}_0}(A)$ (corollary 3.1.2). Thus the *average next visit time* of A to A is the ratio between the measure of its future to its own.

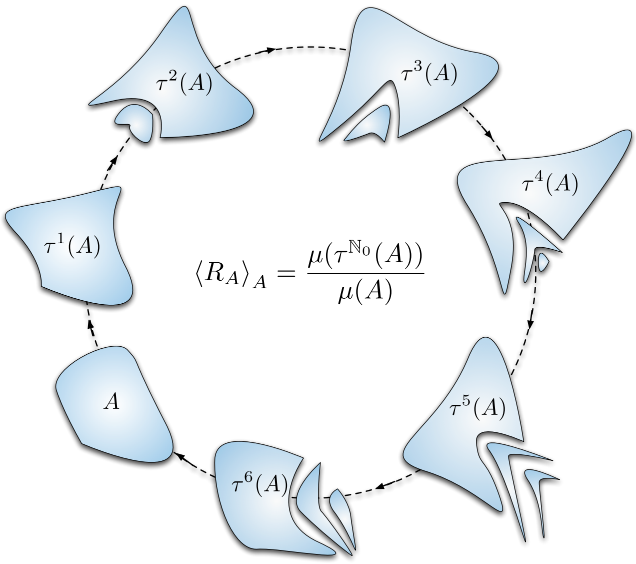

Figure 2: Trajectory of set A under measure preserving flow. Average return time is ratio between future measure and own.

Proof: Set $\mathcal{G}_0 := A$ and

$$\mathcal{G}_k := \tau^{-1}(\mathcal{G}_{k-1}) \cap A^c \quad , \quad \mathcal{R}_k := \tau^{-1}(\mathcal{G}_{k-1}) \cap A$$

with $\mathcal{G}_k$ the points visiting A for the first time k and $\mathcal{R}_k$ the points of A returning to A for the first time k. Then

$$\mu(\mathcal{G}_k) \overset{\mu_\tau = \mu}{=} \mu\left[\tau^{-1}(\mathcal{G}_k)\right] = \mu[\underbrace{\tau^{-1}(\mathcal{G}_k) \cap A^c}_{\mathcal{G}_{k+1}}] + \mu[\underbrace{\tau^{-1}(\mathcal{G}_k) \cap A}_{\mathcal{R}_{k+1}}] = \mu(\mathcal{G}_{k+1}) + \mu(\mathcal{R}_{k+1}) ,$$

which implies

$$\mu(\mathcal{G}_n) = \sum_{k=n+1}^{\infty} \mu(\mathcal{R}_k) + \lim_{k\to\infty} \mu(\mathcal{G}_k) .$$

By Poincaré a.e. point of A returns to A at some time, thus

$$A = \left(\biguplus_{k\in\mathbb{N}} \mathcal{R}_k\right) (\text{mod}0) \quad \Rightarrow \quad \mu(A) = \sum_{k=1}^{\infty} \mu(\mathcal{R}_k) .$$

By setting $n=0$ we get

$$\lim_{k\to\infty}\mu(A_k)=0\ ,$$

hence

$$\mu\left[\tau^{-\mathbb{N}_0}(A)\right]=\sum_{n=0}^{\infty}\mu(\mathcal{G}_n)=\sum_{n=0}^{\infty}\sum_{k=n+1}^{\infty}\mu(\mathcal{R}_k)=\sum_{k=1}^{\infty}k\cdot\mu(\mathcal{R}_k)=\int_A R_A\ d\mu\ .$$

□

3.1.4 Corollary: Approximating trajectories

Let $(M,\mathcal{M},\mu)$ be a second-countable[6] topological, finite measure space so that $\mathcal{M}$ contains the topology of M and $\tau:M\to M$ measure preserving. Then for almost all $x\in M$, the trajectory $\{\tau^n x\}_{n\in\mathbb{N}}$ comes arbitrary *close* to x. In other words, for almost all $x\in M$ and any open neighborhood V of x: $\tau^n x\in V$ for some $n\in\mathbb{N}$.

Example: Any compact subspace of $\mathbb{R}^{2n}$ ($x^1,..,x^{2n}$ symplectic), equipped with the Lebesgue measure, and Hamilton-flow $\tau:=\varphi_{X_H}^{t_0}$.

Proof: Let $U=\{U_1,U_2,..\}$ be a countable base of the topological space M. For any set $A\in\mathcal{M}$ let

$$\tilde{A}:=\{x\in A\mid\tau^n x\notin A\ \ \forall\ n\in\mathbb{N}\}$$

be the points in A that never return to A. From the recurrence theorem we know $\mu(\tilde{A})=0$. Thus

$$\mu\left(\bigcup_{i\in\mathbb{N}}\tilde{U}_i\right)\le\sum_{i=1}^{\infty}\mu\left(\tilde{U}\right)=0\ .$$

But for each

$$x\in M\setminus\bigcup_{i\in\mathbb{N}}\tilde{U}_i$$

and each open neighborhood $\emptyset\neq V$ of x there exists an $\underbrace{U_i}_{\ni x}\subset V$ (since $V=\bigcup_{i\in I}U_i$ for some $I\subset\mathbb{N}$) and by construction

$$\tau^n x\in U_i\subset V$$

for some $n\in\mathbb{N}$.
□

3.1.5 The trapped gas

Consider N gas-particles trapped within a bounded volume $B\subset\mathbb{R}^d$, described by the Hamiltonian

$$H(\mathbf{p},\mathbf{q})=U(\mathbf{q})+\frac{\|\mathbf{p}\|^2}{2m}$$

in symplectic coordinates $\mathbf{p},\mathbf{q}$, so that U is bounded below in B. The measure $\lambda_{\mathbf{q},\mathbf{p}}$ described by the volume-form

$$V:=dq^1\wedge\ldots dq^n\wedge dp^1\wedge\cdots\wedge dp^N$$

is strict positive (with regard to the topology induced by open sets in $\mathbb{R}^{dN}$) and preserved by the Hamilton-flow $\left(\varphi_{X_H}^t\right)$ (see section 4). Furthermore, any set

$$M:=\{H_{\text{low}}\le H(\mathbf{q},\mathbf{p})\le H_{\text{high}}\}$$

is due to the structure of H compact, which implies the finiteness of the Lebesgue-measure $\lambda_{\mathbf{q},\mathbf{p}}$ on the system's manifold M. Note that actually $\varphi_{X_H}^t:M\to M$ since $\varphi_{X_H}^t$ conserves H.

[6] There exists a countable family $U=\{U_1,U_2,..\}$ of open sets such that any open set V in M is a union of elements of U.

Now consider the system in a certain state 1 (see fig. 3). By corollary 3.1.4 the system will (almost surely) eventually come **arbitrarily close** (state 3) to its initial configuration.

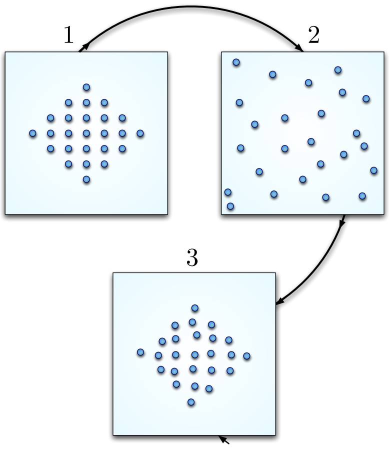

Figure 3: Gas trapped within a bounded volume, at 3 different times.

Note that the corollary does not say much about the time needed for this process to happen. Nonetheless, corollary 3.1.3 allows us to estimate the average return time $\langle R_A \rangle_A$ of configurations from and to a certain start-set A. Consider the case of a two-part box of volume B filled with N gas particles, initially located at one of two halves $B_{1,2}$ of the box.

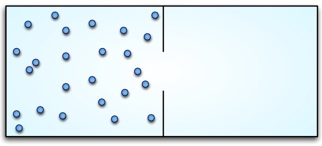

Figure 4: Gas initially located at one of the two halves of a box.

Observing the system only on discrete times $\{n \cdot t_0\}_{n \in \mathbb{N}_0}$, $t_0 > 0$, we ask for the *average observations count* needed for the whole particles to be observed in the initial half again, taken *over all possible configurations starting in that half.* For simplicity we assume the gas to be ideal, that is, the potential to be $\lambda_{\mathbf{q}}$-almost everywhere zero and on some null-set ∞ (the *collision-set*). With

$$\lambda_{\mathbf{q},\mathbf{p}}(A) = \lambda_{\mathbf{q},\mathbf{p}} \left(\left\{ \mathbf{q} \in B_1^N, \; H_{\text{low}} \leq H(\mathbf{q},\mathbf{p}) \leq H_{\text{high}} \right\} \right)$$

$$= \lambda_{\mathbf{q},\mathbf{p}} \left(\left\{ \mathbf{q} \in B_1^N, mH_{\text{low}} \leq \|\mathbf{p}\| \leq mH_{\text{high}} \right\} \right) = B_1^N \cdot \underbrace{\frac{(m\pi)^{\frac{dN}{2}}}{\Gamma\left(1+\frac{dN}{2}\right)} \cdot \left[H_{\text{high}}^{\frac{dN}{2}} - H_{\text{low}}^{\frac{dN}{2}} \right]}_{C_N} \; ,$$

$$\lambda_{\mathbf{q},\mathbf{p}}(M) = (B_1 + B_2)^N \cdot C_N$$

we get the upper bound:

$$\langle R_A \rangle_A = \frac{1}{\lambda_{\mathbf{q},\mathbf{p}}(A)} \int\limits_A R_A \, d\lambda_{\mathbf{q},\mathbf{p}} \leq \left[\frac{B_1 + B_2}{B_1} \right]^N = 2^N \; .$$

3.1.6 The Poincaré map

Let M be a n-dimensional $\mathcal{C}^\infty$ manifold and $(\varphi_X^t)_{t \in \mathbb{R}}$ the flow induced by the $\mathcal{C}^1$ vector field X on M. Let $p \in M$ be a periodic point of (φ_X^t), that is, $\varphi_X^{t_0}(p) = p$ for some minimal $t_0 > 0$. Finally, let $X_p \neq 0$ and N be a $(n-1)$-dimensional sub-manifold containing p and transversal to X in some open neighborhood of p.

Then there exists some local homeomorphism $\mathcal{R}_X : U_p \subset N \to N$ on an open neighborhood U_p of p such that $\mathcal{R}_X(x)$ is the *first* intersection of the future trajectory $\{\varphi_X^t x\}_{t>0}$ with N. This map is called *Poincaré (return) map* induced by (φ_X^t) on U_p.

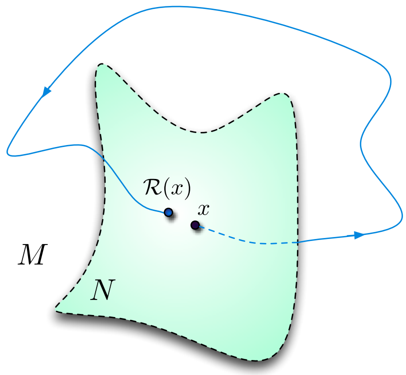

Figure 5: Poincaré recurrence map for 2D-submanifold N in $\mathbb{R}^3$.

Interpretation: Using the Poincaré map it is possible to reduce the study of continuous-time, smooth dynamical systems $(M, (\varphi_X^t)_{t\in\mathbb{R}})$ to discrete-time ones on lower-dimensional manifolds $(N, (R_X^n)_{n\in\mathbb{N}_0})$. Though the explicit calculation of the return map is in most cases impossible, observed properties of this map can deliver great insights into the structure of the *original* system (see section 5.1.4). See more in [2].

Proof: We shall assume $\{\varphi_X^t\}$ to be at least diffeomorphisms. Choose coordinates x^i around p such that $x^1\big|_N = 0,\ \partial_1 = X$ and $x^2, .., x^n\big|_N$ are coordinates in N. This is always possible since X is transversal to N. Then by construction

$$\varphi_X^t(0, x^2, .., x^n) = (t, x^2, .., x^n)$$

at least in some open neighborhood $\underbrace{U_0 \times U_p}_{\ni(0,p)} \subset \mathbb{R}\times N$. Thus

$$\underbrace{\varphi_X^{t_0+t}(0, x^2, .., x^n)}_{\Phi(t_0+t, x^2, .., x^n)} = \underbrace{\varphi_X^{t_0}(t, x^2, .., x^n)}_{\substack{\text{diffeomorphism}\\ \text{in } (t,x^2,..,x^n)}} ,$$

hence

$$\Phi : \underbrace{\overbrace{(U_{t_0} \times U_p)}^{\text{open}}}_{\ni(t_0,p)} \subset (\mathbb{R}\times N) \to M$$

is a diffeomorphism to its image $\tilde{U}_p \ni p$. Note that Φ essentially maps every point $y \in \tilde{U}_p$ to some point $x \in N$ and time t_x such that $\varphi_X^{t_x}(x) = y$.

Define the projection $\Pi : \mathbb{R}\times N \to N$ by

$$\Pi(t, x^2, .., x^n) = (x^2, .., x^n)$$

and consider the map $T := \Pi \circ \Phi^{-1}\big|_{\tilde{U}_p\cap N}$. We shall show that $T : (\tilde{U}_p \cap N) \to N$ is (locally) a diffeomorphism. Continuity (and differentiability) is given, as Π and Φ^{-1} are continuous and differentiable. By the implicit function theorem, it suffices to show that dT maps $(n-1)$ linear independent vectors to $(n-1)$ linear independent vectors. Consider the vectors $\partial_2, .., \partial_n \in T_pN$. Since $d_{(t_0,p)}\Phi(\partial_1) = \partial_1$ we know $d_p\Phi^{-1}(\partial_1) = \partial_1$, thus $d_p\Phi^{-1}(\partial_i),\ i = 2, .., n$ are such that, their *projection* $d\Pi(d_p\Phi^{-1}\partial_i)$ on T_pN is a basis in T_pN. Otherwise, with $d_p\Phi^{-1}(\partial_i) =: a_i^j\partial_j$, there would exist some $0 \neq (k^2, .., k^n) \in \mathbb{R}^{n-1}$ such that

$$\sum_{i=2}^n k^i \cdot \overbrace{\sum_{j=2}^n a_i^j \partial_j}^{d\Pi\ d_p\Phi^{-1}(\partial_i)} = 0$$

and thus for $k^1 := -\sum_{i=2}^n k^i a_i^1$:

$$k^1 \cdot d_p\Phi^{-1}(\partial_1) + \sum_{i=2}^n k^i \cdot d_p\Phi^{-1}(\partial_i) = \underbrace{k^1 \cdot \partial_1 + \sum_{i=2}^n k^i a_i^1 \partial_1}_{0} + \underbrace{\sum_{i=2}^n k^i \cdot \sum_{j=2}^n a_i^j \partial_j}_{0} = 0 \ ,$$

which is a contradiction to the fact that $d_p\Phi^{-1}$ maps linear independent vectors to linear independent ones.

Thus T is at least a local homeomorphism, mapping points $y \in \tilde{U}_p \cap N$ to a *previous intersection point* $x \in N$. As by assumption $T(p)$ is the *most recent* previous intersection of p with N, continuity of T implies the same at least for some open neighborhood of p. Setting $\mathcal{R}_X := T^{-1}$ yields what was to be shown.
□

3.2 Birkhoff's ergodic theorem

One key question arising in the study of dynamical systems with physical origin, is the concept of *average values* of functions defined on the phase-space. One distinguishes between two kinds of averages: time average, corresponding to collections of values taken by iterating along the flow, and phase-averages, taken through some probability measure imposed on the phase space. While the later is well defined, the existence of the former is in no way trivial and was only established in the 1930s by David Birkhoff.[31] The following section sketches the rather lengthy proof (found in appendix A.1) of this important theorem and outlines some immediate consequences. For more on this theorem see [1],[3],[4],[17].

3.2.1 Maximal ergodic theorem

Let $(M, \mathcal{M}, \mu)$ be a measure space and $T : L_1(M, \mathcal{M}, \mu) \to L_1(M, \mathcal{M}, \mu)$ a positive[7]. contraction[8]. For any $f \in L_1(M, \mathcal{M}, \mu)$ let

$$S_n f := \sum_{k=0}^{n-1} T^k f \quad , \quad A_n f := \frac{S_n f}{n} \quad ,$$

$$M_n^S f := \max\{S_1 f, .., S_n f\} \ , \ M_n^A f := \max\{A_1 f, .., A_n f\} \quad ,$$

$$P_n f := \underbrace{\{M_n^S f \geq 0\}}_{\{M_n^A f \geq 0\}} \ , \ P_\infty f := \bigcup_{n\in\mathbb{N}} P_n f \ .$$

Then

$$\int_{P_n f} f d\mu \geq 0 \quad , \quad \int_{P_\infty f} f \ d\mu = 0 \ .$$

Proof: See appendix A.1.1.

3.2.2 Corollary: Maximal ergodic inequality

Let $(M, \mathcal{M}, \mu)$ be a σ-finite measure space[9] and $T : L_1(M, \mathcal{M}, \mu) \to L_1(M, \mathcal{M}, \mu)$ the operator induced by the measure preserving map $\tau : M \to M$. Then the inequality

$$\mu\left(\{M_n^A f \geq \alpha\}\right) \leq \frac{\|f\|_1}{\alpha} \tag{3.2.2.1}$$

[7] An operator $T : V \to W$ between vector spaces with partial order, is *positive* $(T \geq 0) :\Leftrightarrow$

$$T\{v \in V : v \geq 0\} \subset \{w \in W : w \geq 0\} \ .$$

[8] A bounded, linear operator $T : V \to V$ in the normed vector space $(V, \|\cdot\|)$ is called a *contraction* $:\Leftrightarrow \ \|T\| \leq 1$.

[9] A measure space $(M, \mathcal{M}, \mu)$ is *σ-finite* $:\Leftrightarrow \ \exists \ U_1, U_2, .. \in \mathcal{M} : \bigcup_{n\in\mathbb{N}} U_n = M \ \wedge \ \mu(U_n) < \infty$.

holds for any real valued $f \in L_1$ and $\alpha > 0$.
Note: For $T = \mathrm{Id}$ this reduces to the known Markov-inequality.

Proof: See appendix A.1.2.

3.2.3 Birkhoff-Khinchin-Ergodic Theorem

Let $(M, \mathcal{M}, \mu)$ be a finite measure space, $\tau : M \to M$ measure preserving and $f \in L_1$ (real or complex). Then for almost all $x \in M$ the averages

$$A_n f(x) := \frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau^k(x)$$

converge pointwise to some τ-invariant $\overline{f} \in L_1$ with $\left\|\overline{f}\right\|_1 \leq \|f\|_1$. For each τ-invariant $A \in \mathcal{M}$:

$$\int_A \overline{f}\ d\mu = \int_A f\ d\mu\ .$$

Proof: See appendix A.1.3.

3.2.4 Birkhoffs ergodic theorem for time-semi-flows

Let $(M, \mathcal{M}, \mu)$ be a finite measure space, (τ^t) a measure preserving $\mathbb{R}_+$-semi-flow or $\mathbb{R}$-flow on M and $f \in L_1$ (real or complex). Then for almost all $x \in M$ the average

$$\frac{1}{2T} \int_{-T}^{T} f\left(\tau^t x\right)\ dt$$

in case of a flow and

$$\frac{1}{T} \int_0^T f\left(\tau^t x\right)\ dt$$

in case of a semi-flow, converges for $T \to \infty$ to some (τ^t)-invariant $\overline{f} \in L_1$ with $\left\|\overline{f}\right\|_1 \leq \|f\|_1$. For each (τ^t)-invariant $A \in \mathcal{M}$:

$$\int_A \overline{f}\ d\mu = \int_A f\ d\mu\ .$$

Proof: See [1] and [2].

3.2.5 Corollary: Mean sojourn time

Let $(M, \mathcal{M}, \mu)$ be a measure space, $\tau : M \to M$ measure preserving or $(\tau^t)_{t \geq 0}$ a measure preserving $\mathbb{R}_+$-semi-flow. For a set $A \in \mathcal{M}$ define the *sojourn time* $\mathcal{S}_A(x)$ of a point $x \in M$ as

$$\mathcal{S}_A(x) := \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} 1_A(\tau^n x)$$

(where possible) for τ, and

$$\mathcal{S}_A(x) := \lim_{T \to \infty} \frac{1}{T} \int_0^T 1_A(\tau^t x)\ dt$$

for the semi-flow (τ^t). Then $\mathcal{S}_A(x)$ is well-defined a.e., is (mod0) equal to some semi-flow - invariant function $\mathcal{S}_A : M \to [0, \infty]$ and

$$\int_M \mathcal{S}_A\ d\mu = \mu(A)\ .$$

Interpretation: Realizing that (in the case of τ)

$$\mathcal{S}_A(x) = \lim_{n\to\infty} \frac{1}{n} \# \{0 \le m < n : \tau^m x \in A\}$$

is really the *average frequency* of visits of the future trajectory $\{\tau^n x\}_{n\in\mathbb{N}_0}$ of x in A, this rather trivial consequence of the Birkhoff theorem justifies the more or less intuitive notion that the expected *average residence time* of a trajectory within A is proportional to its measure.

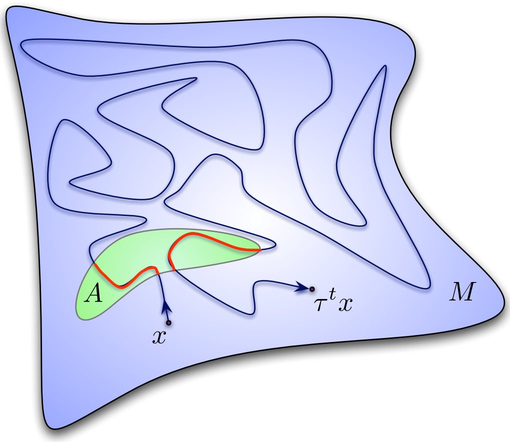

Figure 6: On the definition of the sojourn time. Red trajectory-section *accounts* for sojourn time of x in A.

Proof: By the Birkhoff ergodic theorem the above limits exist a.e. and there exists some semi-flow-invariant $\tilde{\mathcal{S}}_A : M \to [0, \infty)$ such that $\tilde{\mathcal{S}}_A = \mathcal{S}_A(\text{mod}0)$. Furthermore

$$\int_M \mathcal{S}_A d\mu = \int_M \tilde{\mathcal{S}}_A \, d\mu = \int_M 1_A \, d\mu = \mu(A) \ ,$$

which proves what was to be shown.
□

3.2.6 Von Neumann mean ergodic theorem

Let T be a contraction in a Hilbert-Space $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ and P_T the projector on the 1-Eigenspace of T:

$$\mathcal{H}_T := \{f \in \mathcal{H} : Tf = f\} \ .$$

Then the averages

$$A_n f := \frac{1}{n} \sum_{k=0}^{n-1} T^k f$$

converge (in norm $\|\cdot\|_{\mathcal{H}}$) to $P_T f$ for any $f \in \mathcal{H}$.

Note the connection to Birkhoffs ergodic theorem: Setting $\mathcal{H} = L_2(M, \mathcal{M}, \mu)$, $Tf := f \circ \tau$ (since τ is measure preserving T is an isometry) yields the L_2-convergence of the averages $A_n f$ to a τ-invariant $P_T f =: \overline{f}$ with $\|\overline{f}\|_2 \le \|f\|_2$. The average $\overline{f}$ thus *contains only the τ-invariant parts of f*.

Proof: We shall adopt the proof found in [1].

- Let K be a contraction in $\mathcal{H}$. Then $g = Kg \;\Leftrightarrow\; g = K^*g$ for any $g \in \mathcal{H}$.
Proof:

$$\langle g, Kg\rangle = \|g\|^2 \;\Rightarrow\; \langle g, Kg\rangle \in \mathbb{R} \;\Rightarrow\; \langle g, Kg\rangle = \langle Kg, g\rangle$$

$$\Rightarrow\; \|Kg - g\|^2 = \|Kg\|^2 + \|g\|^2 - 2\langle g, Kg\rangle \overset{\|K\|\le 1}{\le} 2\|g\|^2 - 2\|g\|^2 = 0$$

$$\Rightarrow\; g = Kg$$

Thus

$$g = Kg \;\Leftrightarrow\; \|g\|^2 = \langle g, Kg\rangle = \langle K^*g, g\rangle \;\Leftrightarrow\; g = K^*g$$

- Let $\mathcal{K}$ be a family of contractions in $\mathcal{H}$ and

$$\mathcal{H}_{\mathcal{K}} := \bigcap_{K\in\mathcal{K}} H_K = \{f \in \mathcal{H} : Kf = f \;\forall\; K \in \mathcal{K}\} \;.$$

Then

$$\mathcal{H}_{\mathcal{K}}^{\perp} = \mathrm{cl}\,\underbrace{\mathrm{span}\,\{Kh - h : h \in \mathcal{H},\; K \in \mathcal{K}\}}_{N} \;.$$

Proof: We write

$$g \perp N \;\Leftrightarrow\; \langle g, (K - \mathrm{Id})h\rangle = 0 \;\; \forall\; h \in \mathcal{H},\; K \in \mathcal{K}$$

$$\Leftrightarrow\; \langle K^*g - g, h\rangle = 0 \;\; \forall\; h \in \mathcal{H},\; K \in \mathcal{K}$$

$$\Leftrightarrow\; K^*g = g \;\; \forall\; K \in \mathcal{K} \quad\Leftrightarrow\quad Kg = g \;\; \forall\; K \in \mathcal{K}$$

$$\Leftrightarrow\; g \in \mathcal{H}_{\mathcal{K}}$$

and get $N^{\perp} = \mathcal{H}_{\mathcal{K}}$. Note that $\mathcal{H}_{\mathcal{K}}$ is closed since all $K \in \mathcal{K}$ are bound and thus continuous, which implies

$$\mathcal{H}_{\mathcal{K}}^{\perp} = (N^{\perp})^{\perp} = \mathrm{cl}(N) \;.$$

- Now let $f \in \underbrace{(T - \mathrm{Id})\mathcal{H}}_{N}$, that is, $f = (T - \mathrm{Id})h$ for some $h \in \mathcal{H}$. Then

$$\|A_n f\| = \frac{1}{n}\left\|Th - h + T^2h - Th + \cdots + T^n h - T^{n-1}h\right\| = \frac{1}{n}\|T^n h - h\| \overset{\|T\|\le 1}{\le} \frac{2}{n}\|h\| \overset{n\to\infty}{\longrightarrow} 0 \;.$$

For $f \in \mathrm{cl}\,N$, that is, $f = \lim_{k\to\infty} f_k,\; f_k = (T - \mathrm{Id})h_k$:

$$\limsup_{n\to\infty} \overbrace{\|A_n f\|}^{\le\|f\|} = \limsup_{n\to\infty}\lim_{k\to\infty}\|A_n f_k\| \le \limsup_{n\to\infty}\frac{2}{n}\sup_k\{\|h_k\|\} = 0 \;.$$

Let now $f \in \mathcal{H}$. Since $\mathcal{H} = \underbrace{\mathrm{cl}\,\mathcal{H}_K}_{\mathcal{H}_K} \oplus \mathcal{H}_K^{\perp}$ we can split $f = \overbrace{f_{\mathcal{H}_K}}^{\in\mathcal{H}_K} + \overbrace{f_{\mathcal{H}_K^{\perp}}}^{\in\mathcal{H}_K^{\perp}}$ and get

$$A_n f = \underbrace{A_n f_{\mathcal{H}_K}}_{f_{\mathcal{H}_K}} + A_n f_{\mathcal{H}_K^{\perp}} \overset{\mathcal{H}_K^{\perp} = \mathrm{cl}\,N}{=} f_{\mathcal{H}_K} + A_n f_{\mathrm{cl}\,N} \underset{\|\cdot\|_{\mathcal{H}}}{\overset{n\to\infty}{\longrightarrow}} f_{\mathcal{H}_K} \;,$$

with $f_{\mathcal{H}_K} = P_T f$.

□

3.3 Ergodicity

While Birkhoff's *mean ergodic* theorem secured the existence of *time means* along trajectories, it did not answer the serious question initially raised by Boltzmann's work on the equivalence of phase-averages and time-averages in dynamical systems. The importance of such a connection can be understood, by considering the fact that, typical studies of certain processes are conducted through the observation of one or just a few *copies* of the system, that is, along just a few *trajectories* through the phase-space (see also [8]). Hence, time averaging is the tool usually available to experimentalists, while the development and refinement of theoretical models often requires knowledge of certain phase averages, ideally the underlying probability distribution.

It turns out that these two concepts are connected through a new qualitative trait, characterizing certain dynamical systems: *Ergodicity.* We shall in the following section address this important idea with its implications, and attempt to develop a framework of sufficient conditions for ergodicity.

3.3.1 Definition: Ergodic semi-flows

Let $(M, \mathcal{M}, \mu)$ be a measure space with a measure preserving family of maps $\{\tau^g\}_{g\in G}$. If every $\{\tau^g\}$-invariant set $A \in \mathcal{M}$ has the property $\mu(A) = 0$ or $\mu(A^c) = 0$, then $\{\tau^g\}$ is called *ergodic* and (in case of a G-semi-flow) $(M, \mathcal{M}, \mu, (\tau^g))$ an *ergodic system.*[3]

3.3.2 Lemma about ergodic semi-flows

Let $(M, \mathcal{M}, \mu)$ be a measure space and $(\tau^g)_{g\in G}$ a G-semi-flow of bijections. Then $(\tau^g)_{g\in G}$ is ergodic $\Leftrightarrow$ $\left\{\tau^g, (\tau^g)^{-1}\right\}_{g\in G}$ is ergodic.
Example: The ergodicity of an $\mathbb{R}$-flow $(\tau^t)_{t\in\mathbb{R}}$ is equivalent to the ergodicity of $(\tau^t)_{t\geq 0}$.

Proof: Note that measure invariance of τ^g is equivalent to measure invariance of $(\tau^g)^{-1}$. Thus, any $\{\tau^g\}$-invariant set $A \in \mathcal{M}$, is $\left\{\tau^g, (\tau^g)^{-1}\right\}$-invariant as well. Obviously, the converse is also true. But this implies what was to be shown.
□

3.3.3 Theorem: Invariant functions and ergodic semi-flows

Let $(M, \mathcal{M}, \mu)$ be a measure space with a measure preserving G-semi-flow (τ^g). Then the following statements are equivalent:

1. (τ^g) is ergodic.
2. Any measurable, (τ^g)-invariant function $f : M \to \mathbb{C}^*$ is constant almost everywhere.
3. Any bounded, measurable, (τ^g)-invariant function $f : M \to \mathbb{R}^*$ is constant almost everywhere.
4. In case of a finite measure space, and some arbitrary $0 < p < \infty$: Every bounded, measurable, (τ^g)-invariant, L_p-integrable function $f : M \to \mathbb{R}$ is constant a.e.

Proof: We shall elaborate on the proof outlined in [3].

$1 \to 2$: Let f be flow-invariant and w.l.o.g. real. Then the set $\{f > c\}$ is flow invariant for any $c \in \mathbb{R}$, thus $\mu(\{f > c\}) = 0$ or $\mu(\{f \leq c\}) = 0$. W.l.o.g let $\mu(\{f > c\}) = 0$ for some c. Assume that the number

$$c_0 := \inf\{c \in \mathbb{R} : \mu(\{f > c\}) = 0\}$$

is real. If it does not, then

$$\forall\, c \in \mathbb{R} : \mu(\{f > c\}) = 0 \quad\Rightarrow\quad \mu\Bigg[\underbrace{\bigcup_{n\in\mathbb{N}} \{f > -n\}}_{\{f=-\infty\}^c}\Bigg] \leq \sum_{n=1}^{\infty} \underbrace{\mu\left(\{f > -n\}\right)}_{0} = 0$$

which implies $f = -\infty$ almost everywhere.
Since c_0 is infimum, this means that $\forall\, \varepsilon > 0$:

$$\mu(\{f > c_0 + \varepsilon\}) = 0 \quad \wedge \quad \mu(\{f \leq c_0 - \varepsilon\}) = 0$$

which implies $\mu(\{f \neq c_0\}) = 0$.

$2 \to 3$: Trivial.

$3 \to 1$: Let $A \in \mathcal{M}$ be flow invariant. Then by definition 1_A is flow invariant and thus constant (0 or 1) almost everywhere. But this precisely means $\mu(A) = 0$ or $\mu(A^c) = 0$.

$3 \to 4$: Trivial.

$4 \to 1$: Same as $3 \to 1$.

□

3.3.4 Theorem: Characterization of ergodic G-flows

Let $(M, \mathcal{M}, \mu)$ be a measure space and $(\tau^g)_{g\in G}$ a measure preserving G-flow. Then the following statements are equivalent:

1. (τ^g) is ergodic.
2. For any $A \in \mathcal{M}$ with $\mu(A) > 0$:
$$\underbrace{\bigcup_{g\in G} \tau^g(A)}_{\tau^G(A)} = M(\mathrm{mod} 0)$$
that is, the trajectory of A covers almost all of M.
3. For $\mu(A) > 0$ almost every trajectory visits A, that is:
$$\mu[\{x \in M : \underbrace{\{\tau^g(x)\}_{g\in G}}_{\tau^G(x)} \cap A = \emptyset\}] = 0\ .$$
4. If μ is σ-finite: Any (τ^g)-invariant measure $\nu \ll \mu$ [10], is equal to μ up to a multiplicative constant.

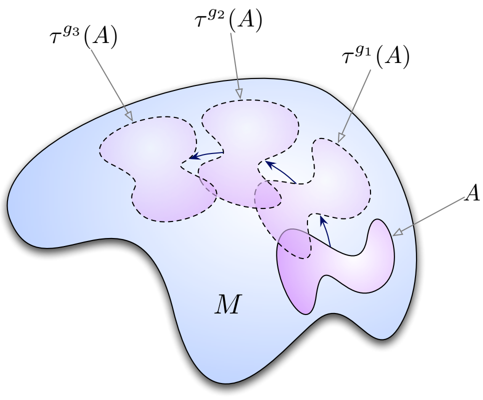

Figure 7: Trajectory of set A in M.

[10] A measure ν is called *absolutely continuous* with respect to μ, if $\mu(A) = 0$ implies $\nu(A) = 0$. We write $\nu \ll \mu$. Note that by the Radon-Nikodym theorem, if μ is σ-finite, this implies the existence of a density $\frac{d\nu}{d\mu}$.

Proof:

1 → 2: Since $\tau^G(A)$ is (τ^g)-invariant and $\tau^G(A) \supset A$ one has

$$\mu(\tau^G(A)) \geq \mu(A) > 0 \ ,$$

which implies

$$\mu\left[\left(\tau^G(A)\right)^c\right] = 0 \ .$$

2 → 3: Since $\tau^G(A)$ is (τ^g)-invariant:

$$\tau^G(x) \cap A = \emptyset \ \Leftrightarrow \ \tau^G(x) \cap \tau^G(A) = \emptyset \ \Leftrightarrow \ x \notin \tau^G(A) \ ,$$

which implies

$$\mu[\{x \in M : \tau^G(x) \cap A = \emptyset\}] = \mu\left[\left(\tau^G(A)\right)^c\right] = 0 \ .$$

3 → 1: Let $A \in \mathcal{M}$ be (τ^g)-invariant and $\mu(A) > 0$. Then $A = \tau^G(A)$ and similar to above:

$$\tau^G(x) \cap A = \emptyset \ \Leftrightarrow \ x \notin \tau^G(A) = A \ ,$$

which implies

$$0 = \mu[\{x \in M : \tau^G(x) \cap A = \emptyset\}] = \mu(A^c) \ .$$

1 → 4: For any $A \in \mathcal{M}$:

$$\int_A \frac{d\nu_{\tau^g}}{d\mu} \ d\mu = \nu_{\tau^g}(A) = \int_M 1_{(\tau^g)^{-1}A} \ d\nu = \int_M \frac{d\nu}{d\mu} \cdot 1_A \circ \tau^g \ d\mu = \int_{(\tau^g)^{-1}M} \frac{d\nu}{d\mu} \cdot (1_A \circ \tau^g) \ d\mu$$

$$= \int_M 1_A \cdot \left[\frac{d\nu}{d\mu} \circ \tau^{-g}\right] \ d\mu_{\tau^g} \overset{\mu_{\tau^g} = \mu}{=} \int_A \frac{d\nu}{d\mu} \circ \tau^{-g} \ d\mu$$

$$\Rightarrow \ \frac{d\nu_{\tau^g}}{d\mu} = \frac{d\nu}{d\mu} \circ \tau^{-g} \quad \text{a.e. (w.l.o.g. everywhere)}$$

Since ν is (τ^g)-invariant:

$$\frac{d\nu}{d\mu} = \frac{d\nu}{d\mu} \circ \tau^{-g} \ ,$$

that is, the density $\frac{d\nu}{d\mu}$ is (τ^g)-invariant. But by theorem 3.3.3 this means that $\frac{d\nu}{d\mu}$ is constant a.e., hence $\nu = c \cdot \mu$ for some $c \in [0, \infty]$.

4 → 1: Let A be (τ^g)-invariant and $\mu(A) > 0$. Then define the (τ^g)-invariant measure

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

Clearly $\nu \ll \mu$, which implies

$$\nu = c \cdot \mu$$

for some constant $c \in [0, \infty]$ (actually $c > 0$, since $\nu \neq 0$), hence:

$$\mu(A^c) = \frac{\nu(A^c)}{c} = \frac{\mu(A^c \cap A)}{c \cdot \mu(A)} = 0 \ .$$

□

3.3.5 Corollary about ergodic G-flows in topological spaces

Let $(M, \mathcal{M}, \mu)$ be a topological, second-countable, strict-positive[11] measure space so that $\mathcal{M}$ contains the topology of M and $(\tau^g)_{g\in G}$ an ergodic G-flow. Then almost every trajectory $\tau^G(x)$ is dense in M, that is,

$$\mu\left\{x \mid \exists \text{ open } U \neq \emptyset : \tau^G(x) \cap U = \emptyset\right\} = 0$$

Example: Any complete, ergodic Hamilton-flow $(\varphi_t^{X_H})_{t\in\mathbb{R}}$ on a $\mathcal{C}^\infty$ manifold with an induced Lebesgue-measure, almost surely leads to dense trajectories.

Proof: Let $\emptyset, U_1, U_2, ..$ be the topological basis of M ($U_i \neq \emptyset$). Then

$$\underbrace{\left\{x \mid \exists \text{ open } U \neq \emptyset : \tau^G(x) \cap U = \emptyset\right\}}_{\Omega} = \left\{x :\mid \exists\, i \in \mathbb{N} : \tau^G(x) \cap U_i = \emptyset\right\}$$

$$= \left\{x :\mid \exists\, i \in \mathbb{N} : x \notin \tau^G(U_i)\right\} = \underbrace{\bigcup_{i\in\mathbb{N}} \left[\tau^G(U_i)\right]^c}_{\in\mathcal{M}}$$

and thus

$$\mu(\Omega) \leq \sum_{i=1}^{\infty} \underbrace{\mu\left[\left(\underbrace{\tau^G(U_i)}_{\substack{M(\text{mod}0)\\(3.3.4)}}\right)^c\right]}_{0} = 0 \quad .$$

□

3.3.6 Theorem: Characterization of ergodicity for time- and iteration-semi-flows

Let $(M, \mathcal{M}, \mu)$ be a probability space and $\tau : M \to M$ a measure preserving map[12] or (τ^t) a measure preserving $\mathbb{R}_+$-semi-flow. Then the following statements are equivalent:

1. The system is ergodic.
2. For any real or complex $f, g \in L_2$:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_M f\left(\tau^n(x)\right) \cdot g(x)\; d\mu = \int_M f\; d\mu \cdot \int_M g\; d\mu$$

   in case of τ, and

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt \int_M f\left(\tau^t(x)\right) \cdot g(x)\; d\mu = \int_M f\; d\mu \cdot \int_M g\; d\mu$$

   in case of an $\mathbb{R}_+$-semi-flow.

3. For any real or complex $f \in L_1$ and almost all $x \in M$:

$$\underbrace{\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(\tau^n(x)\right)}_{A_n f(x)} = \underbrace{\int_M f\; d\mu}_{\langle f\rangle_\mu} \tag{3.3.6.1}$$

   in case of τ, and

$$\underbrace{\lim_{T\to\infty} \frac{1}{T} \int_0^T f\left(\tau^t(x)\right) dt}_{A_T f(x)} = \langle f\rangle_\mu$$

[11] For any open, non-empty set $U : \mu(U) > 0$.

[12] We identify any map τ with its iteration-semi-flow $(\tau^n)_{n\in\mathbb{N}_0}$ or iteration-flow $(\tau^n)_{n\in\mathbb{Z}}$ (if τ bijective). Notice that measure invariance of (τ^n) is equivalent to measure invariance of τ.

in case of an $\mathbb{R}_+$-semi-flow.
Note that this expresses exactly the often sought-after notion *time averages equal phase averages* for state variables f in ergodic, dynamical systems.

4. For any set $A \in \mathcal{M}$ and almost all $x \in M$:

$$\underbrace{\lim_{n\to\infty} A_n 1_A(x)}_{\mathcal{S}_A(x)} = \mu(A) \tag{3.3.6.2}$$

in the case of τ, and

$$\lim_{T\to\infty} A_T 1_A(x) = \mu(A)$$

in the case of an $\mathbb{R}_+$-semi-flow.

Note:

- The above limes is exactly the *sojourn time* of $x \in M$ in A (compare to corollary 3.2.5). Almost every trajectory is thus in some sense *equidistributed* within M!
- Applying this statement to open sets, yields the denseness of almost every future trajectory $\{\tau^n x\}$ (or $\{\tau^t x\}$) in any topological, strict-positive, finite measure space (compare to corollary 3.3.5).

5. For any $A, B \in \mathcal{M}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu\left[(\tau^n)^{-1}(A) \cap B\right] = \mu(A) \cdot \mu(B)$$

in the case of τ, and

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mu\left[(\tau^t)^{-1}(A) \cap B\right] = \mu(A) \cdot \mu(B)$$

in case of an $\mathbb{R}_+$-semi-flow.

6. In case of τ: For any $A \in \mathcal{M}$: Let

$$R_A(x) := \inf\left[\{n \in \mathbb{N} : \tau^n(x) \in A\} \cup \{\infty\}\right]$$

be the *next visit time* of A by the future trajectory of $x \in M$ and

$$R_A^0(x) := R_A(x) \quad , \quad R_A^{i+1}(x) := \begin{cases} R_A\left[\tau^{R_A^i(x)} x\right] + R_A^i(x) & : R_A^i(x) < \infty \\ \infty & : R_A^i(x) = \infty \end{cases}$$

the *visiting times* of A by x in increasing order. Then for almost all $x \in M$:

$$\lim_{n\to\infty} \frac{R_A^n(x)}{n} = \frac{1}{\mu(A)} \ .$$

Note that the above limes is exactly the *average visiting period* of A by the trajectory $\{\tau^n x\}$.

7. In case of τ: For any $\mu(A) > 0$:

$$\underbrace{\bigcup_{n\in\mathbb{N}_0} (\tau^n)^{-1}(A)}_{\tau^{-\mathbb{N}_0}(A)} = M(\text{mod}0) \ .$$

8. In case of τ: For any $\mu(A) > 0$: Almost every trajectory $\underbrace{\{\tau^n(x)\}_{n\in\mathbb{N}_0}}_{\tau^{\mathbb{N}_0}}$ visits A, that is,

$$\mu\left(\left\{x : \tau^{\mathbb{N}_0}(x) \cap A = \emptyset\right\}\right) = 0$$

(compare to theorem 3.3.4 on G-flows).

9. In case of τ: For any $\mu(A) > 0$:

$$\int_A R_A \; d\mu = 1 \; .$$

In other words: The *average return time* for points in A is $1/\mu(A)$.

See also [3],[4].

Note: Many of the above properties, like (4) and (9) about average sojourn & return times, are in accordance with what simple intuition often leads to assume. But as this theorem proves, these are generally not given and actually only hold for ergodic systems.

Proof:

$1 \to 2$: Using Neumann's ergodic theorem (3.2.6) one obtains the convergence of

$$A_n f := \frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau^n$$

to $\int_M f \; d\mu$ with $n \to \infty$ in $\|\cdot\|_2$, for $f \in L_2$. Now consider the Hilbert-space $(L_2, \langle\cdot,\cdot\rangle)$. Since $\langle\cdot,\cdot\rangle$ is continuous in $\|\cdot\|_2$:

$$A_n f \underset{\|\cdot\|_2}{\overset{n\to\infty}{\longrightarrow}} \int_M f \; d\mu \quad \Rightarrow \quad \int_M g \cdot A_n f \; d\mu = \langle g^*, A_n f\rangle \overset{n\to\infty}{\longrightarrow} \left\langle g^*, \int_M f \; d\mu \right\rangle = \int_M g \; d\mu \cdot \int_M f \; d\mu \; .$$

The convergence in the case of $A_T f$ can be reduced to the discrete case $A_{T_i} f$ by the very definition of convergence $T \to \infty$.

$1 \to 3$: With Birkhoff's ergodic theorem we have

$$\lim_{n\to\infty} A_n f(x) = \overline{f}(x)$$

almost everywhere and

$$\int_M \overline{f} \; d\mu = \int_M f \; d\mu$$

for some τ-invariant $\overline{f} \in L_1$. But by theorem 3.3.3 $\overline{f}$ is constant $f(x) = f_0$ a.e., which implies

$$\lim_{n\to\infty} A_n f(x) = f_0 = \int_M \overline{f} \; d\mu = \int_M f \; d\mu = \langle f\rangle_\mu$$

almost everywhere. The case of (τ^t) is similar.

$3 \to 4$: By setting $f = 1_A$ we obtain:

$$\lim_{n\to\infty} A_n 1_A(x) = \int_M 1_A \; d\mu = \mu(A)$$

almost everywhere. The case of the (τ^t) is analogous.

$4 \to 1$: For any τ-invariant $A \in \mathcal{M}$ one has

$$1_A(x) = \lim_{n\to\infty} \overbrace{A_n 1_A}^{1_A}(x) = \mu(A)$$

for almost all $x \in M$ and thus $\mu(A) = 0$ or $\mu(A) = 1$. The case of the semi-flow is analogous.

2 → 5: Set $f := 1_A$ and $g = 1_B$, to obtain

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}\mu\left[(\tau^n)^{-1}(A)\cap B\right] = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}\int_M (\overbrace{1_A\circ\tau^n}^{1_{(\tau^n)^{-1}(A)}})\cdot 1_B\; d\mu = \mu(A)\cdot\mu(B)\ .$$

The case of a (τ^t) is analogous.

5 → 1: For any τ-invariant $A\in\mathcal{M}$ set $B := A$ and so

$$\mu(A) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu[\underbrace{(\tau^n)^{-1}(A)}_{A}\cap A] = [\mu(A)]^2$$

which implies $\mu(A) = 0$ or $\mu(A) = 1$. The case of (τ^t) is similar.

4 ↔ 6: By construction

$$\lim_{n\to\infty} A_n 1_A(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} 1_A(\tau^k x) = \lim_{n\to\infty}\frac{(n+1)+1_A(x)}{R_A^n(x)+1} = \lim_{n\to\infty}\frac{n}{R_A^n(x)}\ .$$

1 → 7: Obviously

$$\tau^{-1}\left(\tau^{-\mathbb{N}_0}(A)\right)\subset\tau^{-\mathbb{N}_0}(A)\ .$$

By Lemma A.3.3 there exists a τ-invariant set B such that $B = \tau^{-\mathbb{N}_0}(A)(\mathrm{mod}0)$. Since $\mu(B)\geq\mu(A)>0$ this implies

$$0 = \mu(B^c) = \mu\left[\left(\tau^{-\mathbb{N}_0}(A)\right)^c\right]\ .$$

7 → 8: Let $\mu(A)>0$. Since

$$\tau^{\mathbb{N}_0}(x)\cap A = \emptyset\ \Leftrightarrow\ \tau^{\mathbb{N}_0}(x)\cap\tau^{-\mathbb{N}_0}(A) = \emptyset\ \Leftrightarrow\ x\notin\tau^{-\mathbb{N}_0}(A)$$

one has

$$\mu\left(\left\{x : \tau^{\mathbb{N}_0}(x)\cap A = \emptyset\right\}\right) = \mu\left[\left(\tau^{-\mathbb{N}_0}(A)\right)^c\right] = 0\ .$$

8 → 1: Let $A\in\mathcal{M}$ be τ-invariant and $\mu(A)>0$. Then $\tau^{-\mathbb{N}_0}(A) = A$ and thus

$$0 = \mu\left(\left\{x : \tau^{\mathbb{N}_0}(x)\cap A\right\}\right) = \mu\left[\left(\tau^{-\mathbb{N}_0}(A)\right)^c\right] = \mu(A^c)\ .$$

7 → 9: By corollary 3.1.3 of the Poincaré recurrence theorem:

$$1 \overset{7.}{=} \mu\left(\tau^{-\mathbb{N}_0}(A)\right) \overset{(3.1.3)}{=} \int_A R_A\; d\mu\ .$$

9 → 1: Let A be τ-invariant and $\mu(A)>0$, then since $R_A\big|_A = 1$:

$$\mu(A) = \int_A R_A\; d\mu = 1\ .$$

□

3.3.7 Lemma: Ergodicity in metrically isomorphic systems

Let $(M,\mathcal{M},\mu,(\tau^g)_{g\in G})$ and $(N,\mathcal{N},\nu,(\lambda^g)_{g\in G})$ be metrically isomorphic through $\varphi : M\to N$. Then (τ^g) is ergodic ⇔ (λ^g) is ergodic.

Interpretation: Metrically isomorphic systems arise naturally under coordinate transformations on manifolds describing dynamical systems. In that case this lemma expresses nothing other than the coordinate-invariance of ergodicity.

An important conclusion is, that a.e. trajectory of an ergodic flow on n-dimensional $\mathcal{C}^1$-manifolds can not be recurrent, as the image of a.e. trajectory would be dense in some open subset of $\mathbb{R}^n$. But recurrent trajectories have finite *length* (as continuous images of compact time-interval), which precludes their denseness.

Proof: It suffices to show one of the two directions, since metrical isomorphicity is symmetric between the two systems. Note that the measure invariance of (τ^g) is equivalent to the measure invariance of (λ^g). For any set $A \in \mathcal{M}$ the (τ^g)-invariance of A is equivalent to the (λ^g)-invariance of $\varphi(A)$, since $1_{\varphi(A)} = 1_A \circ \varphi^{-1}$ (see notes in 2.0.7). Now let (τ^g) be ergodic and $B \in \mathcal{N}$ (λ^g)-invariant. Then $\varphi^{-1}(B)$ is (τ^g)-invariant, which implies

$$\mu(\varphi^{-1}(B)) = 0 \quad \vee \quad \mu\big[\underbrace{\left(\varphi^{-1}(B)\right)^c}_{\varphi^{-1}(B^c)}\big] = 0 \ .$$

But since φ is measure preserving, this means $\nu(B) = 0$ or $\nu(B^c) = 0$.
□

3.3.8 Notes on measure-invariance of ergodicity

Let $(M, \mathcal{M})$ be a measurable space and $(\tau^g)_{g\in G}$ a G-semi-flow on M. Let $\nu \ll \mu$ be (τ^g)-invariant measures. Then:

1. If (τ^g) is ergodic to μ, then it is also ergodic to ν.
2. If $\frac{d\nu}{d\mu} > 0$ μ-almost everywhere and (τ^g) is ergodic to ν, then it is also ergodic to μ.

Notes:

- Recall that any measure ν with some density $\frac{d\nu}{d\mu}$, is absolutely continuous with respect to μ.
- For any (τ^g)-invariant, equivalent[13] measures μ, ν, ergodicity to the one implies ergodicity to the other.
- Compare statement 1. to theorem 3.3.4, which in case of (τ^g) being a G-flow would actually imply $\nu = \text{const} \cdot \mu$.

Proof:

1. Follows directly from definition of ergodicity.
2. Let $A \in \mathcal{M}$ such that
$$0 = \nu(A) = \int_A \frac{d\nu}{d\mu} \, d\mu \ .$$
Then $\frac{d\nu}{d\mu} = 0$ μ-almost everywhere in A. But this implies $\mu(A) = 0$, hence $\mu \ll \nu$. Statement 1. implies what was to be shown.

□

3.4 Strict ergodicity in topological spaces

Up until now, we did not require any specific structure for the phase-space other than that of a measurable space. Assuming some extra structure (e.g. topological), allows for an even deeper insight into ergodic systems and the development of more ideas, akin to that of ergodicity.

An important result for compact, topological spaces, is the existence of invariant, ergodic measures, to be addressed below. We shall furthermore introduce a special case of ergodicity, so called *strict ergodicity*, which allows for even stronger statements for dynamical systems.

[13] Two measures are called *equivalent*, if they have the same nullsets.

3.4.1 Theorem: Existence of invariant Borel-measures

Let $(M, \mathcal{M})$ be a compact, metrizable[14], measurable space and $\mathcal{M}$ the Borel-σ-algebra on M. Let $(\tau^g)_{g\in G}$ be a 1-parameter G-semi-flow (i.e. $G \subset \mathbb{R}$ measurable) and every $\tau^g : M \to M$ continuous[15]. Then there exists a (τ^g)-invariant, probability Borel measure.

Proof: I shall generalize the Krylov-Bogoliubov theorem[17].

- In the following proof identify each finite measure μ with the induced positive, linear functional

$$\mu(f) := \int_M f \, d\mu \tag{3.4.1.1}$$

  on the Banach-space $(\mathcal{C}(M), \|\cdot\|_\infty)$. Recall that by the Frigyes-Riesz representation theorem for metrizable spaces, any positive, linear functional μ on $(\mathcal{C}(M), \|\cdot\|_\infty)$ is identifiable with a unique, inner-regular[16] Borel measure satisfying eq. 3.4.1.1.

- By the existence theorem of André Weil, there exists a non-trivial Haar measure γ on the Borel-σ-algebra $\mathcal{G} := \mathscr{B}(\mathbb{R}) \cap G$ such that $\gamma(G \cap [0, g_0]) < \infty$ for any $g_0 \in G$. Clearly $\gamma(G \cap [0, g_0]) > 0$ for $0 \neq g_0 \in G$ and thus $\gamma(G) = \infty$. Within the context of this proof we shall only consider set intersections with G and omit "$G\cap$".

- Choose $g_0 \in G$, w.l.o.g. $g_0 > 0$. Now given some arbitrary, *initial* probability measure μ on $(M, \mathcal{M})$, consider the sequence of probability measures

$$\mu^n(f) := \frac{1}{\gamma([0, n \cdot g_0])} \cdot \int_0^{n\cdot g_0} \mu\left(f \circ \tau^g\right) \, d\gamma(g) \quad , \quad n \in \mathbb{N}$$

  on $(M, \mathcal{M})$. Since M is compact, by lemma A.3.4 the set $\mathfrak{M}$ of probability measures on M is weakly* compact[17], therefore the sequence μ^n has an accumulation point $\nu \in \mathfrak{M}$, with $\mu^{n_k} \overset{k\to\infty}{\underset{\text{weak*}}{\longrightarrow}} \nu$ for some subsequence $(\mu^{n_k})_k \subset (\mu^n)_n$. Furthermore for any $h \in G$:

$$\begin{aligned}
\left|\mu^{n_k}(f \circ \tau^h) - \mu^{n_k}(f)\right| &= \left| \frac{1}{\gamma([0, n_k g_0])} \int_0^{n_k g_0} \mu\left(f \circ \tau^{h+g}\right) \, d\gamma(g) - \frac{1}{\gamma([0, n_k g_0])} \int_0^{n_k g_0} \mu\left(f \circ \tau^g\right) \, d\gamma(g) \right| \\
&= \frac{1}{\gamma([0, n_k g_0])} \left| \int_h^{h+n_k g_0} \mu(f \circ \tau^g) \, d\gamma(g) - \int_0^{n_k g_0} \mu(f \circ \tau^g) \, d\gamma(g) \right| \\
&\leq \frac{1}{\gamma([0, n_k g_0])} \left[\left| \int_{n_k g_0}^{h+n_k g_0} \mu(f \circ \tau^g) \, d\gamma(g) \right| + \left| \int_0^h \mu(f \circ \tau^g) \, d\gamma(g) \right| \right] \\
&\leq \frac{2 \cdot \gamma([0, h])}{\gamma([0, n_k g_0])} \underbrace{\|f\|_\infty}_{<\infty} = \frac{\|f\|_\infty}{n_k} \cdot \frac{2 \cdot \gamma([0, h])}{\gamma([0, g_0])} \overset{k\to\infty}{\longrightarrow} 0 \, .
\end{aligned}$$

---

[14] A topological space T is called *metrizable*, if it admits a metric $d(\cdot, \cdot)$ such that d produces its topology. Note that any compact, second-countable Hausdorff space is metrizable.

[15] A map $\tau : M \to M$ is *continuous*, if the preimages of open sets are open.

[16] The measure μ is inner-regular if for any $A \in \mathcal{M} : \mu(A) = \sup\{\mu(K) : K \subset A, \ K \text{ compact}\}$.

[17] A sequence $\{\mu_n\}_n$ of measures on a topological space M with the Borel σ-algebra, converges *weakly** to μ, if for any bounded, continuous function $f : M \to \mathbb{C}$ the following holds: $\int_M f \, d\mu_n \overset{n\to\infty}{\longrightarrow} \int_M f \, d\mu$. A set $\mathfrak{M}$ of measures on M is *weakly* compact*, if every sequence of measures in $\mathfrak{M}$ contains a weak* convergence subsequence.[5][10]

By choice of ν this implies

$$\underbrace{\nu(f \circ \tau^h)}_{\nu_{\tau^h}(f)} = \lim_{k\to\infty} \mu^{n_k}\left(f \circ \tau^h\right) = \lim_{k\to\infty} \mu^{n_k}(f) = \nu(f) \quad , \quad f \in \mathcal{C}(M) \ .$$

But by Riesz, this implies $\nu = \nu_\tau$. By construction, $\nu(M) = 1$.

□

3.4.2 Theorem: Existence of ergodic Borel measures

Let $(M, \mathcal{M})$ be a metrizable, compact, measurable space, such that $\mathcal{M} = \mathscr{B}(M)$. Let $(\tau^g)_{g\in G}$ be a 1-parameter G-semi-flow (i.e. $G \subset \mathbb{R}$ measurable) on $(M, \mathcal{M})$ and every $\tau^g : M \to M$ continuous. Then there exists a (τ^g)-invariant, **ergodic** probability Borel measure.

Interpretation: Consider some dynamical system described by the measurable space $(M, \mathscr{B}(M))$ and the time-flow $(\tau^t)_{t\in\mathbb{R}}$. Then among all possible equilibrium states, which by the Krylov-Bogoliubov lemma exist, at least one actually turns the flow into an ergodic one! Of course, nothing is said about the structure, let alone the physical meaning, of this measure. In the not-unusual case of the existence of a fixed-point x_0 of the flow, the Dirac-measure δ_{x_0} would be an ergodic, equilibrium probability measure, expressing the dull triviality of the system's evolution.

Proof: Using the generalized Krylov-Bogoliubov theorem, I shall generalize the proof found in [17]. Consider some countable, dense set of functions $\{f_i\}_{i\in\mathbb{N}} \subset \mathcal{C}(M)$ in the Banach-space $(\mathcal{C}(M), \|\cdot\|_\infty)$ and define the monotonically decreasing sequence of measure-families

$$\mathfrak{M}_{i+1} := \left\{\mu \in \mathfrak{M}_i : \mu(f_{i+1}) = \max_{\nu\in\mathfrak{M}_i} \nu(f_{i+1})\right\}$$

with $\mathfrak{M}_0 := \mathfrak{M}((\tau^g))$ as the set of (τ^g)-invariant probability Borel measures. By definition of weak* convergence of measures, the map $\nu \mapsto \nu(f_{i+1})$ is continuous, hence, compactness and non-emptiness of $\mathfrak{M}_i$ implies compactness and non-emptiness[18] of $\mathfrak{M}_{i+1}$. Since $\mathfrak{M}_0$ is compact (see lemma A.3.4) and non-empty by the Krylov-Bogoliubov theorem, these properties hold for all $\mathfrak{M}_i$. By construction, $\mathfrak{M}_i$ are convex, thus the intersection $\mathfrak{D} := \bigcap_i \mathfrak{M}_i$ is also convex and non-empty[19].

We show that $\mathfrak{D} \subset \mathfrak{M}((\tau^g))$ contains only extreme points of $\mathfrak{M}((\tau^g))$. Suppose

$$\mu = t \cdot \mu_1 + (1-t) \cdot \mu_2 \ \in \mathfrak{D}$$

with $\mu_1, \mu_2 \in \mathfrak{M}_0,\ t \in [0,1]$, then for all $f \in \mathcal{C}(M)$:

$$\mu(f) = t \cdot \mu_1(f) + (1-t) \cdot \mu_2(f) \ .$$

But since $\mu \in \mathfrak{M}_1$, it follows $\mu(f_1) \geq \mu_j(f),\ j = 1,2$ and thus $\mu(f_1) = \mu_1(f_1) = \mu_2(f_1)$, hence $\mu_j \in \mathfrak{M}_1$. By induction it follows that $\mu(f_i) = \mu_j(f_i)$ and $\mu_j \in \mathfrak{M}_i,\ i \in \mathbb{N},\ j = 1,2$. By construction the linear functionals $\mu, \mu_j : \mathcal{C}(M) \to \mathbb{C}$ are bounded and thus continuous. Consequently, as $\{f_i\}$ is dense in $\mathcal{C}(M)$, it follows that $\mu(f) = \mu_j(f) \ \ \forall\ f \in \mathcal{C}(M)$ (see lemma A.2.1). By the Frigyes Riesz representation theorem, one has $\mu = \mu_j$, that is, μ is extremal in $\mathfrak{M}((\tau^g))$.

By lemma A.3.4 this implies that (τ^g) is ergodic to μ.

□

3.4.3 Definition: Strictly ergodic semi-flow

Let $(M, \mathcal{M})$ be a topological, measurable space with the Borel-σ-algebra $\mathcal{M}$. The G-semi-flow $(\tau^g)_{g\in G}$ is *strictly ergodic*, if it has precisely one invariant, probability Borel measure μ. [3]

[18] Recall that any continuous map $f : K \to \mathbb{R}$ on a compact set $K \neq \emptyset$ is bounded and attains its supremum. Since f is continuous, the set of points $\{x \in K : f(x) = \max_{x\in K} f(x)\}$ is closed. Furthermore, any closed subset of a compact set is compact.

[19] Note that the intersection of a monotonically decreasing sequence of compact, non-empty sets $\mathfrak{M}_i \supset \mathfrak{M}_{i+1}$ in a topological space is non-empty. Otherwise $\{\mathfrak{M}_i^c\}_{i\in\mathbb{N}}$ would be an open cover of $\mathfrak{M}_1$. Consequently, finitely many $\{\mathfrak{M}_i^c\}_{i\in I}$ would still be a cover of $\mathfrak{M}_1$, thus $\mathfrak{M}_1 \cap \bigcap_{i\in I} \mathfrak{M}_i = \emptyset$, a contradiction!

Interpretation: Any strictly-ergodic semi-flow is actually ergodic with regard to its invariant probability measure μ. Otherwise, for some (τ^g)-invariant set $A \in \mathcal{M}$ with $0 < \mu(A) < 1$, the probability measure

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

would also be (τ^g)-invariant but $\mu_A \neq \mu$.
Now consider a dynamical system, described by the manifold M and some $\mathbb{R}_+$-flow (τ^t) (e.g. Hamilton-flow). If (τ^t) is uniquely ergodic, then there exists exactly one probability (Borel) measure describing the system in equilibrium. With regards to that equilibrium measure, the system is actually ergodic!

3.4.4 Theorem: Strict ergodicity of homeomorphisms

Let $(M, \mathcal{M}, \mu)$ be a compact metric, probability space and $\mathcal{M} = \mathscr{B}(M)$. Let $\tau : M \to M$ be a measure preserving homeomorphism[20] on M. Then the following statements are equivalent:

1. $(\tau^n)_{n\in\mathbb{Z}}$ is strict ergodic.
2. For any $f \in \mathcal{C}(M)$ and **every** point $x \in M$:

$$\underbrace{\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(\tau^k x\right)}_{A_n f(x)} = \underbrace{\int_M f \, d\mu}_{\mu(f)} \ .$$

3. For any $f \in \mathcal{C}(M)$ the time means $A_n f$ converge in $\|\cdot\|_\infty$ to $\mu(f)$.
4. In case that μ is strictly positive and τ a contraction[21]: $(\tau^n)_{n\in\mathbb{Z}}$ is ergodic.

Proof:

$1 \leftrightarrow 2 \leftrightarrow 3$**:** See [3] Chapter 1, §8 and [17] chapter 4.

$1 \to 4$**:** Every strictly ergodic semi-flow is ergodic.

$4 \to 2$**:** We shall generalize the proof found in [3] for the torus-flow. Let $f \in \mathcal{C}(M)$ and $\varepsilon > 0$. As f is uniformly continuous (since M compact) there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \frac{\varepsilon}{2} \ \text{ for } \ d(x, y) < \delta \ .$$

Due to ergodicity of (τ^n) the relation

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(\tau^k x\right) = \int_M f \, d\mu$$

(see theorem 3.3.6) holds on some full-measure set $A \subset M$. As μ is strictly positive this set A is dense in M and thus contains a finite δ-net[22] $x_1, .., x_r \in A$. By construction there exists a $n_0 \in \mathbb{N}$ such that

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} f\left(\tau^k x^i\right) - \int_M f \, d\mu \right| < \frac{\varepsilon}{2}$$

[20] A bijection f, with f and f^{-1} continuous, is called a *homeomorphism*. In context of measure spaces, we further demand measurability of f. Note that continuous bijections from compact to Hausdorff spaces (which is the case here) are homeomorphisms.

[21] A map $f : X \to Y$ between two metric spaces (X, d_X) and (Y, d_Y) is a *contraction*, if $d_Y(f(a), f(b)) \leq d_X(a, b)$ for $a, b \in X$.

[22] A set of points $\{x_i\}_{i\in I}$ with I a directed set, in a metric space X is called a δ-net, if $X = \bigcup_{i\in I} B^o_\delta(x_i)$. Note that every compact metric space has a finite δ-net for any $\delta > 0$. Actually, every dense subset $A \subset X$ contains a finite δ-net. To see this, choose a $\delta/2$-net $x_1, .., x_n$ in X, and for all points $x_i \notin A$, a point $x'_i \in B^o_{\delta/2}(x_i)$.

for all $n \geq n_0$ and $1 \leq i \leq r$. But for any $x \in M$ we have $d(x, x^i) < \delta$ for some $1 \leq i \leq r$ and thus

$$d\left(\tau^k x, \tau^k x^i\right) < \delta$$

for any $k \geq 0$. This implies

$$\left|\frac{1}{n}\sum_{k=0}^{n-1} f\left(\tau^k x\right) - \int\limits_M f\ d\mu\right| \leq \left|\frac{1}{n}\sum_{k=0}^{n-1}\left[f\left(\tau^k x\right) - f\left(\tau^k x^i\right)\right]\right| + \left|\frac{1}{n}\sum_{k=0}^{n-1} f\left(\tau^k x^i\right) - \int\limits_M f\ d\mu\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $n \geq n_0$.

□

3.4.5 Corollary for strictly ergodic homeomorphisms

Let $(M, \mathcal{M}, \mu)$ be a compact metric space with the Borel-σ-algebra $\mathcal{M} = \mathscr{B}(M)$ and $\tau : M \to M$ strictly ergodic to the τ-invariant probability Borel measure μ. If μ is strict positive, then **every** trajectory $\{\tau^n(x)\}_{n\in\mathbb{Z}}$ is dense in M.

Proof: See [3] Chapter 1, §8

3.5 Mixing

Though ergodicity was very quickly accepted as an important property of certain dynamical systems, it later became evident, that even *stronger* traits might be existent in some systems, in particular the ones emerging in physics. In an attempt to describe such properties as physical mixing, *relaxation* and sensitivity to initial conditions, a new concept has been introduced, so called *mixing*.

Mixing systems, as it turns out, exhibit some very interesting properties including ergodicity and sensitive dependence on initial conditions, typically found in chaotic systems [29].

3.5.1 Definition: Relaxing systems

Let $(M, \mathcal{M}, \mu)$ be a probability space and $(\tau^g)_{g\in G}$ an ordered, μ-preserving G-semi-flow, such that, for any probability measure $\mu_0 \ll \mu$ with $\frac{d\mu_0}{d\mu} \in L_2(M, \mathcal{M}, \mu)$, the image measures $\mu_g := (\mu_0)_{\tau^g}$ converge setwise[23] to μ as $g \to \infty$ [24]. Then the system $(M, \mathcal{M}, \mu, (\tau^g))$ is called *relaxing* with *relaxation measure* μ.

Interpretation: Relaxation describes systems, which *independently* of their *initial* probability distribution μ_0 *tend* to pass over to a certain *relaxation distribution* μ. In a sense, such systems (if described on topological spaces) display a *weaker* form of strict ergodicity, since μ is the only equilibrium measure of the class of measures absolutely continuous to μ.

[23] The family of measures $\{\mu_g\}_{g\in G}$ *converges setwise* to μ, if for every set $A \in \mathcal{M}$: $\lim\limits_{g\to\infty} \mu_g(A) = \mu(A)$.

[24] The limes should be taken as a generalization in the following sense: We write $\lim\limits_{g\to\infty} a_g = a$ for some family $\{a_g\}_{g\in G} \subset T$ (T a topological space, $a \in T$) if:

$$\forall \text{ open} \underbrace{U}_{\ni a} : \exists\ g_0 \in G : \forall\ g \geq g_0 : a_g \in U\ .$$

Note that uniqueness as well as linearity of the limes are still preserved if T is Hausdorff. If $T = \mathbb{R}$ then monotony is also preserved.

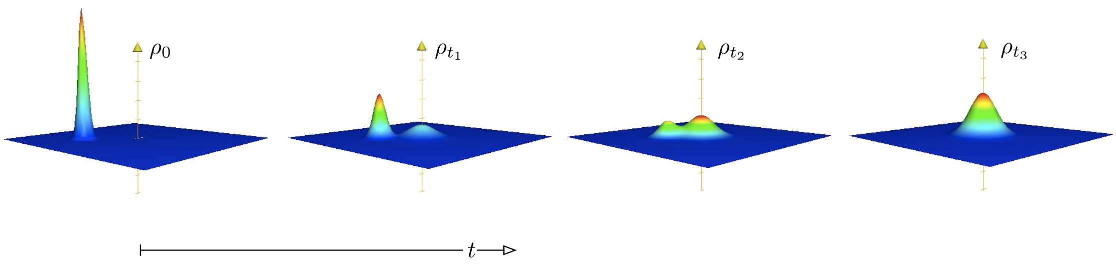

Figure 8: Evolution of density ρ_t of image measure $(\mu_0)_{\tau^t}$ in a 2D relaxing system. Relaxation measure is standard-normal distribution.

3.5.2 Lemma about system relaxations

Let $(M, \mathcal{M}, \mu)$ be a measure space and $(\tau^g)_{g \in G}$ a 1-parameter G-flow, that is, $G \subset \mathbb{R}$ (e.g. $\mathbb{Z}, \mathbb{R}$). Let μ_0 be some *initial* measure, such that the sequence of image-measures $\mu_g := (\mu_0)_{\tau^g}$ converges setwise to μ as $g \to \infty$. Then $\mu_0 \ll \mu$.

Interpretation: Consider a dynamical system described by the measurable space $(M, \mathcal{M})$, the $\mathbb{R}$-flow $(\tau^t)_{t \in \mathbb{R}}$ (e.g. Hamilton-flow) and the equilibrium probability distribution μ. Suppose the initial probability distribution μ_0 is not absolutely continuous to μ. An example would be that μ_0 is *concentrated* on some μ-nullset. Then the system's probability distribution can not possibly setwise converge to the equilibrium μ.

Proof: We first show the statement for automorphisms τ. Suppose $\mu(A) = 0 \neq \mu_0(A)$ for some $A \in \mathcal{M}$. Then the trajectory $\tau^{\mathbb{Z}}(A)$ is a countable union of μ-nullsets, thus its self a μ-nullset. Furthermore, it is τ-invariant, and thus

$$\overbrace{(\mu_0)_{\tau^n}}^{\mu_n}(\tau^{\mathbb{Z}} A) = \mu_0(\tau^{\mathbb{Z}} A) \geq \mu_0(A) > 0$$

which implies

$$\lim_{n \to \infty} \mu_n(\tau^{\mathbb{Z}} A) = \mu_0(\tau^{\mathbb{Z}} A) \neq 0 = \mu(\tau^{\mathbb{Z}} A) \ .$$

Now consider the general case $G \subset \mathbb{R}$, w.l.o.g. $G \neq \{0\}$ (otherwise the statement is trivial). Then for some $0 < t_0 \in G$, the automorphism τ^{t_0} induces a $\mathbb{Z}$-flow $((\tau^{t_0})^n)_{n \in \mathbb{Z}} \subset (\tau^t)_{t \in G}$ similar to above. Suppose $(\mu_0)_{\tau^t} \underset{\text{setwise}}{\overset{t \to \infty}{\longrightarrow}} \mu$, then in particular

$$(\mu_0)_{(\tau^{t_0})^n} \underset{\text{setwise}}{\overset{n \to \infty}{\longrightarrow}} \mu \ .$$

But this implies $\mu_0 \ll \mu$.
□

3.5.3 Definition: Mixing G-semi-flows

Let $(M, \mathcal{M}, \mu)$ be a probability space and $(\tau^g)_{g \in G}$ an ordered, measure preserving G-semi-flow. Then (τ^g) is called *mixing*, if for any two functions $f_1, f_2 \in L_2$ the relation

$$\lim_{g \to \infty} \int_M f_1^* \cdot (f_2 \circ \tau^g) \, d\mu = \int_M f_1^* \, d\mu \cdot \int_M f_2 \, d\mu \qquad (3.5.3.1)$$

holds[25]. See also [3] and [17] for an alternative definition.

[25] Note that by the Hölder inequality we have $\|f_1 \cdot f_2\|_1 \leq \|f_1\|_2 \cdot \|f_2\|_2$ for any $f_1, f_2 \in L_2$. Since $\mu(M) = 1$ this further implies $\|f_1\|_1 \leq \|f_1\|_2$.

3.5.4 Theorem: Characterization of mixing G-semi-flows

Let $(M, \mathcal{M}, \mu)$ be a probability space and $(\tau^g)_{g\in G}$ a measure preserving, ordered G-semi-flow. Then the following statements are equivalent:

1. The system $(M, \mathcal{M}, \mu, (\tau^g))$ is relaxing.
2. The semi-flow (τ^g) is mixing.
3. For any sets $A_1, A_2 \in \mathcal{M}$ the relation
$$\lim_{g\to\infty} \mu\left[A_1 \cap (\tau^g)^{-1} A_2\right] = \mu(A_1)\cdot\mu(A_2)$$
holds.
4. If (τ^g) is a G-flow, then for any sets $A_1, A_2 \in \mathcal{M}$ the relation
$$\lim_{g\to\infty} \mu\left[A_1 \cap \tau^g A_2\right] = \mu(A_1)\cdot\mu(A_2)$$
holds.

Interpretation: Consider a mixing dynamical system described by the measurable space $(M, \mathcal{M})$, the ordered G-flow $(\tau^g)_{g\in G}$ and the relaxation probability distribution μ. Then the flow *mixes* any set $A \in \mathcal{M}$, $\mu(A) > 0$ *throughout the whole space*, in the sense that, for any finite number of sets with positive measure, there exists some time, after which each one is always intersected by $\tau^g A$ (see lemma 3.5.5 below).

Similarly, the *probability*[26], that a point x of A intersects some arbitrary set B at time g, converges to $\mu(B)$ as $g \to \infty$.

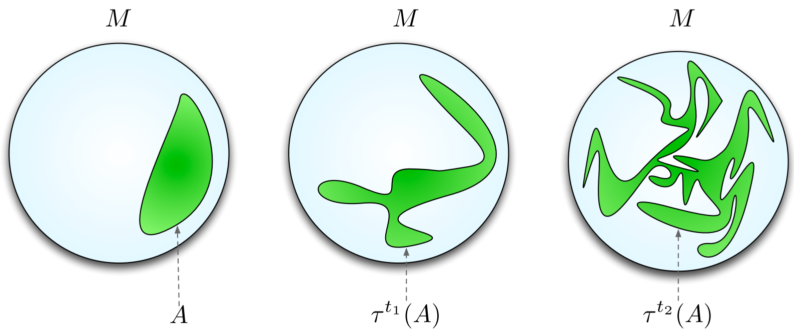

Figure 9: Evolution of some set A under the flow-action in mixing systems, with the Lebesgue-measure as relaxation measure.

Proof:

1 → 3: Let $A_1, A_2 \in \mathcal{M}$ and w.l.o.g. $\mu(A_1) > 0$ (otherwise the statement is trivial). Then the probability measure
$$\mu_0(B) := \frac{\mu\left[A_1 \cap B\right]}{\mu(A_1)}$$
is absolutely continuous with respect to μ with density
$$\frac{d\mu_0}{d\mu} = \frac{1}{\mu(A_1)} \cdot 1_{A_1} \;\in\; L_2(M, \mathcal{M}, \mu)$$
Thus
$$\frac{\mu\left[A_1 \cap (\tau^g)^{-1} A_2\right]}{\mu(A_1)} = \mu_0\left[(\tau^g)^{-1} A_2\right] = \mu_g(A_2) \overset{g\to\infty}{\longrightarrow} \mu(A_2)\ .$$
Note: Alternatively, setting $f_1 := A_1,\ f_2 := A_2$ in eq. 3.5.3.1 yields what is to be shown.

[26] Formally given by $\frac{\mu(B\cap\tau^g A)}{\mu(A)}$.

$2 \to 1$: We shall adopt the proof in [3]. Let $\mu_0 \ll \mu$ be a probability measure with $\frac{d\mu_0}{d\mu} \in L_2(M, \mathcal{M}, \mu)$. Then for any $A \in \mathcal{M}$:

$$\lim_{g\to\infty} \mu_g(A) = \lim_{g\to\infty} \int_M 1_A \, d\mu_g = \lim_{g\to\infty} \int_{(\tau^g)^{-1}(M)} 1_A \circ \tau^g \, d\mu_0 \overset{(\tau^g)^{-1}(M)=M}{=} \lim_{g\to\infty} \int_M \frac{d\mu_0}{d\mu} \cdot (1_A \circ \tau^g) \, d\mu$$

$$= \underbrace{\int_M \frac{d\mu_0}{d\mu} \, d\mu}_{\mu_0(M)=1} \cdot \underbrace{\int_M 1_A \, d\mu}_{\mu(A)} = \mu(A) \; .$$

$3 \leftrightarrow 4$: Since (τ^g) is measure preserving:

$$\mu \left[A_1 \cap \tau^g A_2 \right] \overset{\mu_{\tau^g}=\mu}{=} \mu \left[\tau^{-g} \left(A_1 \cap \tau^g A_2 \right) \right] = \mu \left[\left(\tau^{-g} A_1 \right) \cap A_2 \right] \; .$$

Setting $A_1' := A_2, A_2' := A_1$ on the right-hand-side yields what was to be shown.

$3 \to 2$: Clearly for any $A_1, A_2 \in \mathcal{M}$ the functions $f_1 := 1_{A_1}$ and $f_2 := 1_{A_2}$ satisfy eq. 3.5.3.1. As indicator functions of measurable sets form a complete set[27] in L_2, lemma A.3.5 implies its validity for any $f_1, f_2 \in L_2$.

3.5.5 Lemma about mixing systems

Let $(M, \mathcal{M}, \mu)$ be a probability space and $(\tau^g)_{g\in G}$ a mixing G-semi-flow. Then:

1. The semi-flow (τ^g) is ergodic.
2. Let (τ^g) be a G-flow. Then for any *start-set* $A_s \in \mathcal{M}$, $\mu(A_s) > 0$ and finite family of *target-sets* $\{A_i\}_{i\in I}$ with positive measure, there exists some $g_0 \in G$, such that $\tau^g \cap A_i \neq \emptyset$ for all $i \in I$, $g \geq g_0$.
3. Let (τ^g) be a G-flow, (M, d) a metric space such that $\mathcal{M}$ contains the topology of M and μ strict positive. Then (τ^g) is sensitively dependent on initial conditions[28].

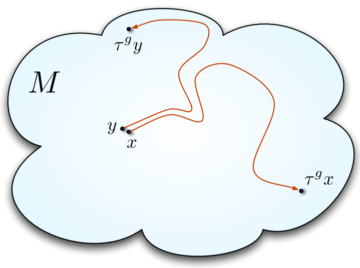

Figure 10: Sensitive dependency on initial conditions in a mixing system in $\mathbb{R}^2$. Arbitrarily close points eventually always *drift apart*.

Proof:

1. Let $A \in \mathcal{M}$ be (τ^g)-invariant. Then by theorem 3.5.4:

$$\mu(A) = \lim_{g\to\infty} \mu[A \cap \overbrace{(\tau^g)^{-1} A}^{A}] = [\mu(A)]^2$$

which is only possible if $\mu(A) = 0$ or $\mu(A) = 1$.

[27] By construction of the integral.

[28] An ordered semi-flow $(\tau^g)_{g\in G}$ is *sensitively dependent on initial conditions* if there exists some $\delta > 0$ such that for any $x \in M$ and open neighborhood $\underbrace{U_x}_{\ni x}$ there exists some $y \in U_x \setminus \{x\}$ and $0 \leq g \in G$ such that $d(\tau^g y, \tau^g x) > \delta$.

2. Choose
$$\varepsilon < \min_{i\in I} \{\mu(A_i)\} \cdot \mu(A_s) \ .$$
Then by theorem 3.5.4 there exists some $g_0 \in G$ such that
$$\mu\left(A_i \cap \tau^g A_s\right) \geq \mu(A_i) \cdot \mu(A_s) - \varepsilon > 0 \quad \forall\, i \in I,\ g \geq g_0 \ ,$$
hence $\tau^g A_s \cap A_i \neq \emptyset$.

3. Choose some $a \neq b \in M$ and set $\delta := \frac{1}{8} d(a,b)$. Now let $U_x \in \mathcal{M}$ be open and $x \in U_x$. Then by statement 2. at some time $g \in G$:
$$\tau^g(U_x) \cap B^o_\delta(a) \neq \emptyset \quad \wedge \quad \tau^g(U_x) \cap B^o_\delta(b) \neq \emptyset \ .$$
But this implies that there exists **no** $c \in M$ such that
$$\tau^g U_x \subset B^o_\delta(c)$$
since otherwise there would exist $x_a \in B^o_\delta(a) \cap B^o_\delta(c)$ and $x_b \in B^o_\delta(b) \cap B^o_\delta(c)$ and so
$$d(a,b) \leq d(a,c) + d(c,b) \leq d(a,x_a) + d(x_a,c) + d(c,x_b) + d(x_b,b) \leq 4\delta \ ,$$
a contradiction! In particular $\tau^g U_x \not\subset B^o_\delta(\tau^g x)$.

□

3.5.6 Lemma: Mixing in metrically isomorphic systems

Let $(M, \mathcal{M}, \mu, (\tau^g)_{g\in G})$ and $(N, \mathcal{N}, \nu, (\rho^g)_{g\in G})$ be metrically isomorphic systems by $\varphi : M \to N$. Then: (τ^g) is mixing, if and only if (ρ^g) is mixing.

Proof: Since metrical isomorphicity is symmetric, it suffices to show only one direction. Let $(M, \mathcal{M}, \mu, (\tau^g))$ be relaxing, $\nu_0 \ll \nu$ and $B \in \mathcal{N}$. Then $(\nu_0)_{\varphi^{-1}} \ll \mu$ and thus
$$(\nu_0)_{\rho^g}(B) = \nu_0\left[(\rho^g)^{-1}B\right] = \nu_0\left[\varphi \circ (\tau^g)^{-1}\varphi^{-1}B\right] = \left[(\nu_0)_{\varphi^{-1}}\right]_{\tau^g}\left[\varphi^{-1}B\right] \overset{g\to\infty}{\longrightarrow} \mu\left[\varphi^{-1}B\right] = \nu(B) \ ,$$
hence the system $(N, \mathcal{N}, \nu, (\rho^g))$ is relaxing. □

Notes on mixing in physical systems

When considering dynamical systems arising in nature, one is typically confronted with systems described by a disproportionately large number of parameters, with time-flows impossible to explicitly calculate. It is thus not an unusual approach, to assume some sort of ergodicity of the flow with respect to some, preferably already known, equilibrium probability distribution.

As a typical example shall be named the classical ideal gas, trapped within some bounded region. The energy leaf $\mathscr{L}_h = \{H = h\}$ is in this case compact, and easily allows for the formulation of an equilibrium distribution μ_h on it (see section 4.1.8), expressing in a sense the typical presence of absolute uncertainty. It is with respect to this measure, that ergodicity is often assumed. An empirical justification comes from the observed fact, that existing information (i.e. some initial probability distribution $\mu_0 \neq \mu_h$, typically $\mu_0 \ll \mu_h$) tends to *blur out* as time passes and any expectations for the system's state converge to the equilibrium distribution μ_h. This is exactly the behavior describing relaxing systems (3.5.1), thus implying the mixing-property and ergodicity.

3.6 Ergodicity in random variables

3.6.1 Preconsiderations

We have so far studied the evolution of dynamical systems and their probability distributions with respect to their flow in abstract measure spaces. Often though, systems can never possibly be observed in their complete detail, let alone described. In many cases, one can *observe* or *measure* a finite set of variables describing the system, so called random variables, which in the most general case correspond to some measurable map X from the phase-space $(M, \mathcal{M}, \mu)$ into some other measurable space $(N, \mathcal{N})$. Their probability distribution is nothing else than the image measure μ_X induced by the map $X : M \to N$.

Typical examples of random variables are mappings into $\mathbb{R}^n$, like the temperature of some gas, the velocity of one particle etc. The question arises, as to what behavior is to be expected for these variables and if they exhibit any of the notions described above, in particular some sort of *ergodicity*.

Note that as a rule, a system is not *completely* described by the random variables (i.e. *observables*) at one's disposal. Thus the introduction of a flow into N is somewhat pointless, since a single value in N may correspond to several sates of the system, each with its own particular future.

3.6.2 Lemma: Trajectories for random variables

Let $(M, \mathcal{M}, \mu)$ be a probability space, $(N, \mathcal{N})$ a measurable space and $X : M \to N$ a random variable. Then:

1. In case that $(M, \mathcal{M})$ and $(N, \mathcal{N})$ are topological and X continuous, surjective: If any trajectory $\tau^G x$ is dense in M, then its image $X\tau^G x$ is dense in N as well.

2. In case that $(M, \mathcal{M})$ and $(N, \mathcal{N})$ are topological[29], M second countable and X continuous: For almost all $x \in M$, the trajectory $X\tau^{\mathbb{N}} x$ comes arbitrarily close to Xx (compare: Poincaré recurrence theorem).

3. If a trajectory $\tau^G x$ of a G-semi-flow $(\tau^g)_{g \in G}$ has the property $A \cap \tau^G x \neq \emptyset$ for any $\mu(A) > 0$, then the *trajectory* $X\tau^G x$ has the similar property: $B \cap (X\tau^G x) \neq \emptyset$ for any $\mu_X(B) > 0$.

4. In case of an ergodic $\tau : M \to M$:

   a) For any $f \in L_1(N, \mathcal{N}, \mu_X)$ and almost all $x \in M$:

   $$\lim_{n\to\infty} \underbrace{\frac{1}{n} \sum_{k=0}^{n-1} f\left(X\tau^n(x)\right)}_{A_n f(Xx)} = \int_N f \, d\mu_X \ .$$

   b) For any set $B \in \mathcal{N}$ and almost all $x \in M$:

   $$\lim_{n\to\infty} A_n 1_B(Xx) = \mu_X(B) \ .$$

   Interpretation: The *sojourn time* of the X*-value* of the *traveling* x in B, is equal to $\mu_X(B)$.

   Similar statements (to a. and b.) hold for ergodic $\mathbb{R}_+$-semi-flows $(\tau^t)_{t\geq 0}$ (compare to theorem 3.3.6).

Proof:

1. Trivial, as surjective, continuous functions map dense sets to dense sets.

2. By corollary 3.1.4 for a.e. $x \in M$, the trajectory $\tau^{\mathbb{N}} x$ comes arbitrarily close to x. Thus, for any open set $\underbrace{U}_{\ni Xx} \in \mathcal{N}$, there exists some $n \in \mathbb{N}$ such that $\tau^n x \in \underbrace{X^{-1}U}_{\text{open}}$, hence $X\tau^n x \in U$.

3. Let $\mu_X(B) > 0$, then by definition $\mu(X^{-1}B) > 0$. Thus there exists some $y \in (X^{-1}B) \cap (\tau^G x)$, hence

   $$Xy \in \left(XX^{-1}B\right) \cap \left(X\tau^G x\right) = B \cap X\tau^G x \ .$$

4. a) Consider the function $fX \in L_1(M, \mathcal{M}, \mu)$. Then theorem 3.3.6 (eq. 3.3.6.1) yields what was to be shown.

   b) By theorem 3.3.6:

   $$A_n 1_B(Xx) \overset{\text{def}}{=} \frac{1}{n} \sum_{k=0}^{n-1} 1_B(X\tau^n(x)) = \frac{1}{n} \sum_{k=0}^{n-1} 1_{X^{-1}B}(\tau^n(x)) \underset{\text{a.e.}}{\overset{(3.3.6.2)}{=}} \mu(X^{-1}B) \overset{\text{def}}{=} \mu_X(B) \ .$$

□

[29] That is, $\mathcal{M}$ and $\mathcal{N}$ contain the topologies of M and N respectively.

3.7 System families and leafs

It is often the case that a dynamical system can be *decomposed* into independent or invariant sub-systems, with the system's flow constructed by some sort of combination of the sub-flows in each sub-system. This kind of decomposition often promises to ease the analysis of the system and reduce the problem to a lower-dimensional one. The question about the existence of ergodic behavior on these sub-systems is not exempt from this reductionism.

We shall in the following examine systems composed of so called *leafs*, each equipped with some measure, and address the relation to existing *global* measures of the space. Moreover, a short connection is established between so called *product-spaces* and separable Hamiltonians.

3.7.1 Lemma: Measure construction from leaf-measures

Let $(M, \mathcal{M})$ be a measurable space, $(N, \mathcal{N}, \nu)$ a measure space and $\{\mathscr{L}_f\}_{f\in N}$ a family of disjoint, measurable subsets of M. Let $(\mu_f)_{f\in N}$ be a family of measures, such that μ_f is a measure on the leaf $\mathscr{L}_f$ (with σ-algebra $\mathcal{M} \cap \mathscr{L}_f$) and the function $\mu_f(A \cap \mathscr{L}_f)$ measurable (in f) for any $A \in \mathcal{M}$. Then:

1. The function

$$\mu(A) := \int_N \mu_f\left(A \cap \mathscr{L}_f\right)\, d\nu(f) \quad , \quad A \in \mathcal{M} \tag{3.7.1.1}$$

defines a measure on M.

2. If all μ_f and ν are probability measures, then μ is a probability measure.

3. Let $(\tau^g)_{g\in G}$ be a G-semi-flow and every leaf $\mathscr{L}_f$ (τ^g)-invariant. If ν-almost every μ_f is (τ^g)-invariant, then μ is (τ^g)-invariant.

4. Let $(N, \mathcal{N})$ be topological and ν strictly positive. Let $(\tau^g)_{g\in G}$ be a G-flow, such that every leaf $\mathscr{L}_f$ is (τ^g)-invariant. Let there exist a *bi-measurable*[30] map $G : M \to K$ to a measurable space $(K, \mathcal{K})$, such that:
   - $G\big|_{\mathscr{L}_f}$ is injective
   - $\mu_f\left(\mathscr{L}_f \cap G^{-1}B\right)$ is continuous in f for all $B \in \mathcal{K}$.
   - $\mu_f\left(\mathscr{L}_f \cap \tau^{-g}G^{-1}B\right)$ continuous in f for all $B \in \mathcal{K}$.

   Then: If μ is (τ^g)-invariant, all μ_f are (τ^g)-invariant.

Proof:

1. Clearly $\mu(\emptyset) = 0$ and $\mu : \mathcal{M} \to [0, \infty]$. Now let $A_1, A_2, .. \in \mathcal{M}$ be disjoint sets, then:

$$\mu\left(\biguplus_{n\in\mathbb{N}} A_n\right) \overset{\text{def}}{=} \int_N \mu_f\Bigg(\underbrace{\mathscr{L}_f \cap \biguplus_{n\in\mathbb{N}} A_n}_{\biguplus_{n\in\mathbb{N}} A_n\cap\mathscr{L}_f}\Bigg)\, d\nu(f) = \int_N \sum_{n=1}^{\infty} \mu_f(A_n \cap \mathscr{L}_f)\, d\nu(f)$$

$$\overset{\text{B. Levi}}{=} \sum_{n=1}^{\infty} \int_N \mu_f(A_n \cap \mathscr{L}_f)\, d\nu(f) = \sum_{n=1}^{\infty} \mu(A_n)\ .$$

2. Trivial.

3. Let ν-almost all μ_f be (τ^g) invariant, that is, there exists a ν-nullset $J \in \mathcal{N}$ such that for all $f \in J^c$ the measure μ_f is (τ^g)-invariant. Then for any $A \in \mathcal{M}$:

$$\mu\left[(\tau^g)^{-1}(A)\right] = \int_N \mu_f\left[(\tau^g)^{-1}(A) \cap \mathscr{L}_f\right]\, d\nu(f) \overset{\substack{\nu(J)=0\\ (\tau^g)^{-1}(\mathscr{L}_f)=\mathscr{L}_f}}{=} \int_{J^c} \underbrace{\mu_f\left[(\tau^g)^{-1}(A \cap \mathscr{L}_f)\right]}_{\mu_f(A\cap\mathscr{L}_f)}\, d\nu(f) = \mu(A)\ .$$

[30]We call a map $G : (M, \mathcal{M}) \to (K, \mathcal{K})$ *bi-measurable* if for any $A \in \mathcal{M}, B \in \mathcal{K}$: $G^{-1}(B) \in \mathcal{M},\ G(A) \in \mathcal{K}$.

4. We shall write
$$F^{-1}(B) := \bigcup_{f\in B} \mathscr{L}_f$$
for $B \in \mathcal{N}$. Suppose there exists some measurable $A \subset \mathscr{L}_{f_0}$ and $g \in G$ such that
$$\mu_{f_0}(A) \neq \mu_{f_0}(\tau^{-g}A) \quad , \text{w.l.o.g.} \quad \mu_{f_0}(A) > \mu_{f_0}(\tau^{-g}A) \ .$$
Then the set $\tilde{A} := G^{-1}(\underbrace{G(A)}_{\in\mathcal{K}})$ satisfies $\tilde{A} \cap \mathscr{L}_{f_0} = A$ and
$$\mu_f(\mathscr{L}_f \cap \tilde{A}) \quad \text{and} \quad \mu_f\left[\mathscr{L}_f \cap \tau^{-g}\tilde{A}\right]$$
are continuous in f. Thus there exists an open neighborhood $\underbrace{U_{f_0}}_{\ni f_0} \in \mathcal{N}$ such that
$$\mu_f\left[\mathscr{L}_f \cap \tilde{A}\right] > \mu_f\left[\mathscr{L}_f \cap \tau^{-g}\tilde{A}\right]$$
for all $f \in U_{f_0}$. But then
$$\mu\left[\tau^{-g}\left(\tilde{A} \cap F^{-1}U_{f_0}\right)\right] = \mu\left[(\tau^{-g}\tilde{A}) \cap F^{-1}U_{f_0}\right] = \int\limits_{U_{f_0}} \mu_f\left[\mathscr{L}_f \cap \tau^{-g}\tilde{A}\right]\, d\nu(f)$$
$$\overset{\nu(U_{f_0})>0}{<} \int\limits_{U_{f_0}} \mu_f\left[\mathscr{L}_f \cap \tilde{A}\right]\, d\nu(f) = \mu\left[\tilde{A} \cap F^{-1}U_{f_0}\right] \ .$$
Thus, μ is not τ^g-invariant.

□

3.7.2 Theorem about ergodic decompositions

Let $(M, \mathcal{M}, \mu)$ be a metrizable, compact probability space with $\mathcal{M} = \mathscr{B}(M)$. Let $(\tau^g)_{g\in G}$ be a measure preserving 1-parameter G-semi-flow on $(M, \mathcal{M})$ and every $\tau^g : M \to M$ continuous. Then there exists a partition $\{\mathscr{L}_f\}_{f\in N}$ (with $(N, \mathcal{N}, \nu)$ a measure space) of (τ^g)-invariant sets $\mathscr{L}_f$, each carrying an ergodic (τ^g)-invariant measure μ_f, such that
$$\mu(A) = \int\limits_N \mu_f(A \cap \mathscr{L}_f)\, d\nu(f) \quad , \quad A \in \mathcal{M} \ .$$

Proof: Use lemma A.3.4 and the Choquet Theorem A.3.6. See also [17].

Interpretation: Consider any dynamical system described by the symplectic manifold (M, ω) and the Hamilton-flow $\left(\varphi^t_{X_H}\right)_t$. By theorem 3.3.3 the existence of a non-a.e.-constant $\left(\varphi^t_{X_H}\right)$-invariant function $\mathbf{F} : M \to \mathbb{R}^k$ (e.g. the Hamiltonian H) usually[31] implies the non-ergodicity of the system in M, as any leaf $\mathscr{L}_{\mathbf{f}} := \{\mathbf{F} = \mathbf{f}\}$ would be flow-invariant. But since the system evolves on some leaf $\mathscr{L}_{\mathbf{f}}$, which can be turned into a probability space, the question of ergodicity on $\mathscr{L}_{\mathbf{f}}$ remains open.

Lemma 3.7.1 shows that, given some equilibrium (i.e. flow-invariant) probability measure $\mu_{\mathbf{f}}$ on each leaf $\mathscr{L}_{\mathbf{f}} := \{\mathbf{F} = \mathbf{f}\}$, where usually $\mathbf{F} : M \to \mathbb{R}^n$ is a set of flow-invariants (i.e. *integrals of motion*), any *probability measure* $\nu_{\mathbb{R}^n}$ *among the leafs* induces a new equilibrium probability measure μ on M as in eq. 3.7.1.1.

As a special case recall the *microcanonical distribution*, where $F = H$ and $\nu = \delta_E$ (see also [8]). If no other integrals exist, the flow might as well be ergodic on $\mathscr{L}_E$ under some suitable measure μ_E. In that case, it is even ergodic in M to μ defined in eq. 3.7.1.1.

On the other hand, the decomposition theorem can be considered a justification for the analysis of properties of ergodic (sub-)systems, as any equilibrium measure of a typical dynamical system admits the reduction of the system to invariant, ergodic components. Often these components correspond to lower-dimensional sub-spaces, as is the case for ergodic energy leafs in Hamiltonian, conservative systems.

[31] Actually depending on the measure imposed.

3.7.3 Definition: Product space

Mechanical systems in nature can often be modeled as sets of independently evolving sub-systems, each one described as a probability space $(M^i, \mathcal{M}^i, \mu^i),\ 1 \leq i \leq N$ with a G-semi-flow $(\tau_i^g)_{g\in G}$. In order to investigate such a system as a whole, we introduce the *product space*

$$(M, \mathcal{M}, \mu) := \bigotimes_{i=1}^{N} (M^i, \mathcal{M}^i, \mu^i)$$

equipped with the product σ-algebra

$$\mathcal{M} := \bigotimes_i \mathcal{M}^i := \sigma\Big(\big\{ \times_i A^i : A^i \in \mathcal{M}^i \big\}\Big)$$

on

$$M := \times_i M^i$$

and the *product* probability measure μ.

In case of stochastically independent sub-systems, the measure $\mu := \bigotimes_i \mu^i$ is the uniquely defined probability measure on $(M, \mathcal{M})$ with

$$\mu(\times_i A^i) = \prod_i \mu^i(A^i) \quad , \quad A^i \in \mathcal{M}^i \ .$$

The G-semi-flows $(\tau_i^g)_{g\in G}$ induce the *product*-G-semi-flow $(\tau^g)_{g\in G} := (\tau_1^g \otimes \cdots \otimes \tau_N^g)_{g\in G}$ on M defined by

$$\tau^g x := \tau_1^g \otimes \cdots \otimes \tau_N^g(x) := \big(\tau_1^g x^1, .., \tau_N^g x^N\big) \quad , \quad x = (x^1, .., x^N) \in M \ .$$

Clearly, (τ^g) is μ-preserving, if an only if every (τ_i^g) is μ_i-preserving. The question arises, how ergodicity of the sub-systems is connected to ergodicity of the whole.

3.7.4 Lemma about ergodicity in product spaces

Let $(M^i, \mathcal{M}^i, \mu^i)$ be (*independent*) probability spaces with the G-semi-flows $(\tau_i^g)_{g\in G}$. Let $(M, \mathcal{M}, \mu)$ the induced product space with the induced G-semi-flow $(\tau^g)_{g\in G}$. If (τ^g) is ergodic, then all sub-flows (τ_i^g) are ergodic.

Note: The converse is not generally true! Consider for example two tori $T^1 \cong S^1$ with translation flows (see section 4.2) $\tau_i^t \vartheta := \vartheta + t$. Both systems are ergodic, but the product system is periodic on T^2 and thus non-ergodic.

Proof: Let (τ^g) be ergodic and $A^k \subset \mathcal{M}^k$ $\{\tau_k^g\}$-invariant for some k. Then clearly the set

$$A := M^1 \times \cdots \times M^{k-1} \times A^k \times M^{k+1} \times \cdots \times M^N$$

is (τ^g)-invariant, hence $\mu(A) = 0$ or $\mu(A) = 1$. But this implies $\mu(A^k) = 0$ or $\mu(A^k) = 1$.
□

3.7.5 Lemma: Connection between separable Hamilton functions and product spaces

Let (M, ω) be a symplectic manifold with symplectic coordinates $\mathbf{q}, \mathbf{p}$ and Hamilton function with the structure

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} H_i(\mathbf{q}^i, \mathbf{p}^i) \ ,$$

that is, separable into Hamilton functions depending on different, *independent* symplectic coordinate groups $\mathbf{q}^i, \mathbf{p}^i$. Then the Hamilton flow is given by

$$\varphi_{X_H}^t = \varphi_{X_{H_1}}^t \circ \cdots \circ \varphi_{X_{H_n}}^t \ . \qquad (3.7.5.1)$$

Proof: It suffices to show that the flow in eq. 3.7.5.1 is actually a solution to

$$\frac{d}{dt}\varphi^t_{X_H} = X_H \ .$$

Since H_i depends only on $\mathbf{q}^i, \mathbf{p}^i$, it follows that X_{H_i} only contains entries at indices belonging to $\mathbf{q}^i, \mathbf{p}^i$. Consequently, all other coordinates of $\varphi^t_{X_{H_i}}$ do not change with t, and thus

$$\frac{\partial \varphi^t_{X_{H_i}}}{\partial(\mathbf{q}, \mathbf{p})}$$

is the identity matrix except for entries for index combinations belonging to $\mathbf{q}^i, \mathbf{p}^i$. This implies

$$\frac{\partial \varphi^t_{X_{H_i}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \underbrace{\frac{\partial \varphi^t_{X_{H_j}}}{\partial t}}_{X_{H_j}} \overset{j \neq i}{=} X_{H_j}$$

which leads to

$$\begin{aligned}
\frac{d}{dt}\varphi^t_{X_{H_1}} \circ \cdots \circ \varphi^t_{X_{H_n}} &= \underbrace{\frac{\partial \varphi^t_{X_{H_1}}}{\partial t}}_{X_{H_1}} + \frac{\partial \varphi^t_{X_{H_1}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \frac{d}{dt}\varphi^t_{X_{H_2}} \circ \cdots \circ \varphi^t_{X_{H_n}} \\
&= X_{H_1} + \underbrace{\frac{\partial \varphi^t_{X_{H_1}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \frac{\partial \varphi^t_{X_{H_2}}}{\partial t}}_{X_{H_2}} + \frac{\partial \varphi^t_{X_{H_1}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \frac{\partial \varphi^t_{X_{H_2}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \frac{d}{dt}\varphi^t_{X_{H_3}} \circ \cdots \circ \varphi^t_{X_{H_n}} \\
&= \cdots = X_{H_1} + \cdots + X_{H_{n-1}} + \underbrace{\frac{\partial \varphi^t_{X_{H_1}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \cdots \cdot \frac{\partial \varphi^t_{X_{H_{n-1}}}}{\partial(\mathbf{q}, \mathbf{p})} \cdot \underbrace{\frac{d}{dt}\varphi^t_{X_{H_n}}}_{X_{H_n}}}_{X_{H_n}} \\
&= X_{\sum_i H_i} = X_H \ .
\end{aligned}$$

□

Interpretation: The above proof makes evident, that the coordinates of any trajectory $\varphi^t_{X_H}(\mathbf{q}_0, \mathbf{p}_0)$ can be separated into coordinate groups $\{(\mathbf{q}^i, \mathbf{p}^i)\}_i$, in which $(\mathbf{q}^i, \mathbf{p}^i)$ only depend on the initial $(\mathbf{q}^i_0, \mathbf{p}^i_0)$. The representation in eq. 3.7.5.1 is thus equivalent to the formalism introduced in section 3.7.3 about product spaces, whereas the sub-flows, are exactly the $\left(\varphi^t_{X_{H_i}}\right)_t$ with regards to the i-th coordinate group.

Remarks

We have in this chapter introduced important concepts like ergodicity, mixing and denseness of trajectories in phase-spaces. It is important to note that while ergodicity and mixing are measure-theoretic concepts, denseness of trajectories and sensitive dependency on initial conditions are topological ones! The connection between these two ideas is by no means trivial, and usually requires an additional structure, i.e. connection between the topology and the measure space built upon it.

4 Flows on manifolds

Due to their high level structure (e.g. Hausdorff topological, metrizability etc), manifolds allow a rich exploration of the concept of ergodicity and topological characteristics of flows. As flows and observables are often continuous or even smooth, a connection between their topological and measure-theoretical properties is much more easily established.

In the following, the Lebesgue measure is naturally introduced on manifolds using given coordinates, while all other measures will be characterized by their density with respect to the former. Following up, we examine so called translation-flows on tori and link them to so called *Liouville-integrable systems*.

4.1 Measures on manifolds

4.1.1 Definition: Lebesgue-measure on $\mathcal{C}^\infty$-manifolds

Let M be an n-dimensional $\mathcal{C}^\infty$ manifold with atlas (ψ_i, U_i) (see [11] for notation). Then any chart (ψ, U) (coordinates x^i) locally induces the *Lebesgue-measure* λ_ψ on $(U, \mathscr{B}(U))$[32]

$$\lambda_{\mathbf{x}}(B) := \lambda_\psi(B) := \lambda_{\mathbb{R}^n}\left(\psi^{-1}(B)\right) = \int\limits_{\psi^{-1}(B)} d^n x$$

with the Lebesgue-measure $\lambda_{\mathbb{R}^n}$ in $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. For any other measure μ, locally absolute continuous to $\lambda_{\mathbf{x}}$ with (local) density $\frac{d\mu}{d\lambda_{\mathbf{x}}} = \rho$, we write

$$d\mu = \rho \; d^n x \; .$$

Note:

- All measures $\lambda_{\mathbf{x}}$ induced by some coordinates $\mathbf{x}$ are equivalent.
- The density $\rho_{\mathbf{x}}$ for a measure $\mu \ll \lambda_{\mathbf{x}}$ with respect to $\lambda_{\mathbf{x}}$ is coordinate dependent! With respect to some other coordinates $\mathbf{y}$ it takes the form

$$\rho_{\mathbf{y}} = \rho_{\mathbf{x}} \cdot \left| \det\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right) \right| \; . \tag{4.1.1.1}$$

- If $\lambda_{\mathbf{x}}$ is measure invariant but non-ergodic, there can not exist an ergodic measure $\mu \ll \lambda_{\mathbf{x}}$ with $\frac{d\mu}{d\lambda_{\mathbf{x}}} > 0$ $\lambda_{\mathbf{x}}$-almost everywhere (see 3.3.8).
- If (M, ω) is symplectic and both coordinate sets $x^i,\ y^i$ are symplectic, then $\rho_{\mathbf{x}} = \rho_{\mathbf{y}}$.

4.1.2 Notes on metrizability of manifolds

Any n-dimensional manifold M can be locally transformed into a metric space, such that the resulting open sets are exactly the ones of its topology, as any chart $\psi : U \subset M \to \mathbb{R}^n$ (locally) induces the metric

$$d_M^\psi(x, y) := d_{\mathbb{R}^n}(\psi(x), \psi(y))$$

on U. Since ψ and ψ^{-1} are continuous and bijective, open sets in U correlate uniquely to open sets in $\psi(U)$ and vice versa. This furthermore implies, that all chart-metrics are equivalent[33] and that the *denseness* of any set $A \subset M$ in U (e.g. trajectory $\{\tau^g(x)\}_g$), can be defined in any arbitrary chart-metric.

[32]Recall that any manifold is in particular a Hausdorff space.
[33]Two metrics d_1, d_2 in a metric space X are set to be *equivalent*, if the produce the same open sets.

4.1.3 Measure-evolution under flows

Let M be a n-dimensional $\mathcal{C}^\infty$-manifold, (φ_X^t) the (semi)-flow belonging to a vector field X and μ_0 an *initial* measure on M. Then the flow induces the 1-parameter family of measures μ_t by:

$$\mu_t := \mu_{\varphi_X^t} \ .$$

The choice of specific coordinates x^i often allows the representation of μ_t by a $\lambda_{\mathbf{x}}$-density $\frac{d\mu_t}{d\lambda_{\mathbf{x}}} =: \rho_t$. Since we usually assume μ_t to be finite, ρ_t is finite $\lambda_{\mathbf{x}}$-almost everywhere, thus w.l.o.g. we can assume $\rho_t : M \to [0,\infty)$. In the following we shall study the *evolution* of this density under flows.

4.1.4 Liouville's Theorem

For the $\mathcal{C}^1$ density $\rho_t(x)$ the following *continuity* equation holds:

$$\frac{\partial \rho_t}{\partial t}(x) + \sum_i \frac{\partial(\rho_t X^i)}{\partial x^i}(\varphi_X^t(x)) = 0 \ .$$

Interpretation: The above equation can be compared to the known continuity equation encountered in fluid dynamics, expressing the *conservative transport* of the probability density along the flow.

Proof: By definition the density ρ_{t_0+t} is given by

$$\int_A \rho_{t_0+t}\, d\lambda \overset{!}{=} \mu_{t_0+t}(A) = \mu[\underbrace{\varphi_X^{-t_0-t}(A)}_{\varphi_X^{-t_0}(\varphi_X^{-t}(A))}] = \mu_{t_0}\left[\varphi_X^{-t}(A)\right] = \int_{\varphi_X^{-t}(A)} \rho_{t_0}\, d\lambda = \int_A \underbrace{\left|\det\left(\frac{\partial \varphi_X^{-t}}{\partial x}\right)\right| \cdot \rho_{t_0} \circ \varphi_X^{-t}}_{\rho_{t_0+t}}\, d\lambda \ .$$

Recall that the flow φ_X^{-t} preserves orientation, that is, its Jacobi-determinant is non-negative[34]. The partial time derivative $\frac{\partial \rho_t}{\partial t}(t_0)$ is thus given by

$$\left.\frac{\partial \rho_t}{\partial t}\right|_{x,t=t_0} = \left.\frac{\partial \rho_{t_0+t}}{\partial t}\right|_{x,t=0} = \frac{\partial \rho_{t_0}}{\partial x^i} \underbrace{\left.\frac{\partial (\varphi_X^{-t})^i}{\partial t}\right|_{x,t=0}}_{-X^i} \cdot \underbrace{\det\left(\frac{\partial \varphi_X^{-t}}{\partial x}\right)\Big|_{x,t=0}}_{\det(\mathrm{Id})=1} + \rho_{t_0} \circ \underbrace{\varphi_X^{-t}\big|_{x,t=0}}_{x} \cdot \frac{\partial}{\partial t} \det\left(\frac{\partial \varphi_X^{-t}}{\partial x}\right)\Big|_{x,t=0}$$

$$= -X^i \frac{\partial \rho_{t_0}}{\partial x^i} + \rho_{t_0} \cdot \frac{\partial}{\partial t} \underbrace{\det\left(\mathrm{Id} - \frac{\partial X}{\partial x} t + \mathcal{O}(t^2)\right)}_{1 - t\cdot \mathrm{trace}\left(\frac{\partial X}{\partial x}\right) + \mathcal{O}(t^2)}\Big|_{x,t=0} = -X^i \frac{\partial \rho_{t_0}}{\partial x^i} - \rho_{t_0} \cdot \underbrace{\mathrm{trace}\left(\frac{\partial X}{\partial x}\right)}_{\sum_i \frac{\partial X^i}{\partial x^i}}$$

$$= -\sum_i \frac{\partial(\rho_{t_0} X^i)}{\partial x^i} \ .$$

□

4.1.5 Corollary to the Liouville-Theorem

1. The measure μ_t with density ρ_t is φ_X^t-invariant if and only if

$$\sum_i \frac{\partial(\rho_t X^i)}{\partial x^i} = 0 \ .$$

[34]Note that $\det\left(\frac{\partial \varphi_X^t}{\partial x}\right)\Big|_{t=0} = 1$. Due to continuity of $\det\left(\frac{\partial \varphi_X^t}{\partial x}\right)$, would a negative value mean a zero value at some time t_z as well. But then $\det\left(\frac{\partial \varphi_X^t}{\partial x}\right) = \underbrace{\det\left(\frac{\partial \varphi_X^{t_z}}{\partial x}\right)}_{0} \cdot \det\left(\frac{\partial \varphi_X^{t-t_z}}{\partial x}\right) = 0 \ \ \forall\, t$, a contradiction!

2. If $\partial_i X^i = 0$ then

$$\underbrace{\frac{\partial \rho_t}{\partial t} + X\rho_t}_{\frac{d\rho}{dt}} = 0 \ , \tag{4.1.5.1}$$

with the *convective* time-derivative $\frac{d\rho_t}{dt}$ of ρ_t along the trajectory of φ_X^t.
Comparison: Since $\lambda_{\mathbf{x}}$ is an equilibrium measure, by the same argument as in the proof of theorem 3.3.4, we have

$$\rho_t = \rho_0 \circ \varphi_X^{-t}$$

(in fact this holds even for non-continuous densities). Densities in a sense *move along* the flow and thus *remain constant for an observer moving along* X.

4.1.6 Specialization for Hamiltonian flows

Consider the Hamilton-flow $\left(\varphi_{X_H}^t\right)$ of the $\mathcal{C}^1$ Hamilton vector-field X_H on the $2n$-dimensional symplectic manifold $(M, d\mathbf{q} \wedge d\mathbf{p})$ in symplectic coordinates $\mathbf{q}, \mathbf{p}$. The Hamilton equations

$$\dot{q}^i = \partial_{p^i} H \ , \ \dot{p}^i = -\partial_{q^i} H \ , \ i = 1, .., n$$

imply

$$\sum_{i=1}^{2n} \partial_i X_H^i = 0 \ ,$$

thus resulting in the known Liouville-Equation

$$\frac{\partial \rho_t}{\partial t} = -\underbrace{X_H \rho_t}_{d\rho(X_H)} = \{H, \rho_t\} = \sum_{i=1}^{n} \left[\frac{\partial H}{\partial q^i} \frac{\partial \rho_t}{\partial p^i} - \frac{\partial H}{\partial p^i} \frac{\partial \rho_t}{\partial q^i} \right] \ .$$

As a special case consider $\rho_0 \equiv 1$, which implies the invariance of the Lebesgue-measure $\lambda_{\mathbf{q},\mathbf{p}}$ to $\varphi_{X_H}^t$.

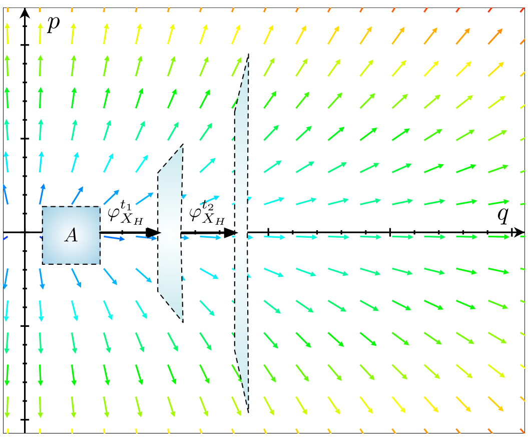

Figure 11: On the Hamilton-flow-invariance of the Lebesgue-measure $\lambda_{q,p}$ in Hamilton-systems. Notice the deformed, yet preserved volume of $\varphi_{X_H}^t(A)$. Arrows represent the vector-field X_H.

4.1.7 Lemma: Equilibrium densities for Hamilton flows

Let $(M, d\mathbf{q} \wedge d\mathbf{p})$ be a symplectic manifold with the Hamiltonian H, its induced vector-field X_H and the corresponding Hamilton-flow $\left(\varphi^t_{X_H}\right)$. Then:

1. The set of equilibrium probability $\lambda_{\mathbf{q},\mathbf{p}}$-densities is convex.
2. If ρ, ζ are equilibrium $\lambda_{\mathbf{q},\mathbf{p}}$-densities, then so are $\{\rho, \zeta\}$, $\rho \cdot \zeta$.
3. Any $\mathcal{C}^1$-function $\rho = \rho(H)$ is an equilibrium $\lambda_{\mathbf{q},\mathbf{p}}$-density.
4. For equilibrium $\lambda_{\mathbf{q},\mathbf{p}}$-densities $\zeta^1, .., \zeta^k$, any function $\rho = \rho(\zeta^1, .., \zeta^k)$ is also an equilibrium $\lambda_{\mathbf{q},\mathbf{p}}$-density.

Proof:

1. Follows directly from lemma A.3.4.
2. As we saw in section 4.1.6 about Hamilton flows, equilibrium densities are characterized by their involution with H. Thus $\{H, \rho\} = \{H, \zeta\} = 0$, consequently $\{H, \{\rho, \zeta\}\} = 0$ and $\{H, \rho \cdot \zeta\} = 0$, which was to be shown.
3. Since H is flow-invariant, it follows:
$$\{H, \rho\} = -X_H \rho(H) = -\frac{d}{dt}[\rho \circ \underbrace{H \circ \varphi^t_X}_{H}]\big|_{t=0} = 0 \ .$$
4. Follows directly from
$$X_H \rho = \frac{\partial \rho}{\partial \zeta^i} \cdot X_H \zeta^i \ .$$

□

4.1.8 Invariant measure on energy-leafs

Consider a dynamical system described by the (non-trivial) Hamiltonian H on the $2n$-dimensional symplectic manifold $(M, d\mathbf{q} \wedge d\mathbf{p})$ and the induced Hamilton-flow $\left(\varphi^t_{X_H}\right)$. As is well known, the Hamiltonian H is (τ^g)-invariant, thus the leafs $\mathscr{L}_h := \{H = h\}$ are flow-invariant. Hence, the system can not be ergodic in M.

As already mentioned in lemma 3.7.1, a flow-invariant measure could indeed be introduced on the leaf $\mathscr{L}_h$, under which the flow might as well be ergodic. It turns out that, for any flow-invariant measure $\mu \ll \lambda_{\mathbf{q},\mathbf{p}}$ with continuous density $\frac{d\mu}{d\lambda_{\mathbf{q},\mathbf{p}}}$, the measure
$$d\mu_h = \frac{d\mu}{d\lambda_{\mathbf{q},\mathbf{p}}} \cdot \frac{V^h}{\|\nabla_{\mathbf{q},\mathbf{p}} H\|}$$
is flow invariant on $\mathscr{L}_h$, where V^h is the volume form induced by the $\mathbf{q}, \mathbf{p}$-Euclidean-metric[35] on $\mathscr{L}_h$. Actually:
$$\mu(A) = \int_{\mathbb{R}} \mu_h(A \cap \mathscr{L}_h) \, d\lambda_{\mathbb{R}}(h) \quad , \quad A \in \mathcal{M} \ .$$

Note

- If $\mathscr{L}_h$ is compact, μ_h can be normed to be a probability measure.
- The measure μ_h (deduced from $\mu = \lambda_{\mathbf{q},\mathbf{p}}$) is exactly the well known *microcanonical distribution* often encountered in classical statistical mechanics! In a sense, the flow-invariance of μ_h expresses the idea that an *equidistribution* of the system on the leaf $\mathscr{L}_h$, is a stable one.

[35] $g := \sum_{i=1}^{n} dq^i \otimes dq^i + \sum_{i=1}^{n} dp^i \otimes dp^i$.

Proof: In lemma A.2.2 set

$$g := \sum_{i=1}^{n} dq^i \otimes dq^i + \sum_{i=1}^{n} dp^i \otimes dp^i$$

and compare to lemma 3.7.1. Note that

$$V^h := \sqrt{\det\left(g\left(\partial_{u^i}, \partial_{u^j}\right)\right)} \cdot du^2 \wedge \cdots \wedge du^n$$

is merely the volume-form on $\mathscr{L}_f$ induced by the volume-form $dq^1 \wedge \cdots \wedge dp^n$.
□

4.2 Flows on the torus

The torus as a manifold is an important case of a dynamical system, since many manifolds are either diffeomorphic to the torus or can be decomposed to invariant tori. Tori with linear flows, form a great part of the Liouville theory for integrable, Hamiltonian systems and KAM theory, which justifies the great attention given to their study in the recent century.

4.2.1 The torus as a manifold

We identify the torus $T^m = \underbrace{S^1 \times \cdots \times S^1}_{\times m}$ with the representation $T^m \simeq \mathbb{R}^m / \mathbb{Z}^m$ in standard coordinates

$$(\vartheta^1, .., \vartheta^m) : T^m \to [0, 1)^m \ .$$

The induced Lebesgue-measure, given by

$$\lambda_{T^m}(A) := \lambda\left[\vartheta(A)\right]$$

is essentially the Lebesgue-measure λ within $[0, 1)^m$, acting on the Borel σ-algebra on $[0, 1)^m$.

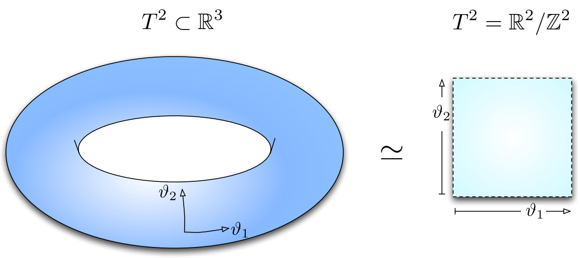

Figure 12: The Torus T^2 in two different representations.

Furthermore, we consider the *pull-back metric* on the torus

$$d_{T^m}(x, y) := d\left(\vartheta(x), \vartheta(y)\right)$$

induced by the Euclidean metric d in $[0, 1)^m$. The torus T^m is thus a compact manifold, and with the Borel-σ-algebra $\mathscr{B}(T^m)$ and probability measure λ_{T^m} a strictly-positive probability space.

4.2.2 The translation group on the torus

Consider the measure preserving $\mathbb{R}^m$-flow $(\tau_{\boldsymbol{\omega}})_{\boldsymbol{\omega} \in \mathbb{R}^m}$ on T^m, defined by

$$\tau_{\boldsymbol{\omega}}(\boldsymbol{\vartheta}) := \boldsymbol{\vartheta} + \boldsymbol{\omega} = \big((\vartheta^1 + \omega^1)\mathrm{mod}1, .., (\vartheta^m + \omega^m)\mathrm{mod}1\big) \ .$$

We call this flow the *translation group* on T^m. For a fixed $\boldsymbol{\omega} \in \mathbb{R}^m$ this induces the $\mathbb{R}$-flow $(\tau^t_{\boldsymbol{\omega}})_{t \in \mathbb{R}}$ on T^m, defined as

$$\tau^t_{\boldsymbol{\omega}} := \tau_{t \cdot \boldsymbol{\omega}}$$

and further the $\mathbb{Z}$-flow $(\tau_{\boldsymbol{\omega}}^n)_{n\in\mathbb{Z}}$ defined as

$$\tau_{\boldsymbol{\omega}}^n := (\tau_{\boldsymbol{\omega}})^n = \tau_{n\cdot\boldsymbol{\omega}}$$

In the last 2 cases, we call ω^i the *flow frequencies*.

4.2.3 Theorem: Ergodicity of $\{\tau_{\boldsymbol{\omega}}^n\}$

The following statements are equivalent[3]:

1. The flow $(\tau_{\boldsymbol{\omega}}^n)_{n\in\mathbb{Z}}$ is ergodic.
2. The flow $(\tau_{\boldsymbol{\omega}}^n)_{n\in\mathbb{Z}}$ is strictly ergodic, that is, the Lebesgue measure λ_{T^m} is the only $(\tau_{\boldsymbol{\omega}}^n)$-invariant probability measure on T^m.
3. The numbers $1, \omega^1, .., \omega^m$ are rationally independent[36].
4. Some (and thus every) future trajectory $\tau_{\boldsymbol{\omega}}^{\mathbb{N}_0}(\boldsymbol{\vartheta})$ is dense in T^m.
5. Some (an thus every) trajectory $\tau_{\boldsymbol{\omega}}^{\mathbb{Z}}(\boldsymbol{\vartheta})$ is dense in T^m.

Proof:

1 → 3**:** We shall follow the proof in [3]. Let $\mathbf{q} \in \mathbb{Z}^m$ such that

$$\sum_{k=1}^m q_k \omega^k = p$$

for some $p \in \mathbb{Z}$. Then the function

$$f(\boldsymbol{\vartheta}) := \exp\left[2\pi i \sum_{k=1}^m q_k \vartheta^k\right]$$

is $\tau_{\boldsymbol{\omega}}$-invariant, since

$$f(\tau_{\boldsymbol{\omega}}\boldsymbol{\vartheta}) = \exp\left[2\pi i \sum_{k=1}^m q_k(\vartheta^k + \omega^k)\right] = \exp\left[2\pi i \sum_{k=1}^m q_k \vartheta^k\right] \cdot \underbrace{\exp\left[2\pi i p\right]}_{1} = f(\boldsymbol{\vartheta}) \ .$$

Since $\{\tau_{\boldsymbol{\omega}}^n\}$ is ergodic, $f = \text{const}$ a.e., which implies $\mathbf{q} = 0$.

3 → 1**:** By theorem 3.3.3 it suffices to show that every $\tau_{\boldsymbol{\omega}}$-invariant, bounded, L_2-integrable function $f : T^m \to \mathbb{C}$ is a constant a.e. This function f can be expanded into a Fourier-series

$$f(\boldsymbol{\vartheta}) \overset{\|\cdot\|_2}{=} \sum_{q\in\mathbb{Z}^m} C_q \exp\left[2\pi i \sum_{k=1}^m q_k \vartheta^k\right] \ .$$

The $\tau_{\boldsymbol{\omega}}$-invariance of f yields

$$\sum_{q\in\mathbb{Z}^m} C_q \exp\left[2\pi i \sum_{k=1}^m q_k \vartheta^k\right] \overset{\|\cdot\|_2}{=} f(\boldsymbol{\vartheta})$$

$$= f(\tau_{\boldsymbol{\omega}}\boldsymbol{\vartheta}) \overset{\|\cdot\|_2}{=} \sum_{q\in\mathbb{Z}^m} C_q \exp\left[2\pi i \sum_{k=1}^m q_k(\vartheta^k + \omega^k)\right] = \sum_{q\in\mathbb{Z}^m} C_q \exp\left[2\pi i \sum_{k=1}^m q_k \vartheta^k\right] \cdot \exp\left[2\pi i \sum_{k=1}^m q_k \omega^k\right] \ .$$

The uniqueness of the Fourier-coefficients implies

$$C_q = 0 \quad \vee \quad \underbrace{\exp\left[2\pi i \sum_{k=1}^m q_k \omega^k\right] = 1}_{\Rightarrow\ \mathbf{q}\cdot\boldsymbol{\omega}\in\mathbb{Z}\ \Rightarrow\ \mathbf{q}=0} \quad \forall\ \mathbf{q} \in \mathbb{Z}^m$$

and thus $C_q = 0$ for $\mathbf{q} \neq 0$. But this essentially means $f = \text{const}$ almost everywhere.

[36] Real numbers $x^1, .., x^m$ are rationally independent $:\Leftrightarrow\ \forall\ 0 \neq \mathbf{k} \in \mathbb{Q}^m : \mathbf{k}\cdot\mathbf{x} \neq 0$. Note that using $\mathbb{Z}^m$ instead of $\mathbb{Q}^m$ does not change the definition.

$1 \to 2$**:** Holds by theorem 3.4.4, since λ_{T^m} strictly positive and $\tau_{\boldsymbol{\omega}}$ a measure preserving homeomorphism and contraction (actually an isometry).

$2 \to 1$**:** See 3.4.3.

$1 \to 4$**:** By lemma 3.3.2 the semi-flow $(\tau^n_{\boldsymbol{\omega}})_{n\in\mathbb{N}_0}$ is also ergodic. By theorem 3.3.6 this implies denseness of almost every trajectory $\tau^{\mathbb{N}_0}\boldsymbol{\vartheta}$.

$4 \to 5$**:** Suppose $\tau^{\mathbb{N}_0}_{\boldsymbol{\omega}}\boldsymbol{\vartheta}_0$ is dense in T^m for some $\boldsymbol{\vartheta}_0 \in T^m$. Now let $\boldsymbol{\vartheta} \in T^m$ be arbitrary and $U \subset T^m$ open. Then $\tilde{U} := (\boldsymbol{\vartheta}_0 - \boldsymbol{\vartheta}) + U$ is open as well, hence

$$\tau^{\mathbb{N}_0}_{\boldsymbol{\omega}}\boldsymbol{\vartheta} \cap U = \tau^{\mathbb{N}_0}_{\boldsymbol{\omega}}(\boldsymbol{\vartheta}) \cap \left(\tilde{U} + (\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0)\right) = \tau^{\mathbb{N}_0}_{\boldsymbol{\omega}}(\boldsymbol{\vartheta}) \cap \tau_{\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0}(\tilde{U}) = \tau_{\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0}\underbrace{\left[\tau^{\mathbb{N}_0}_{\boldsymbol{\omega}}(\boldsymbol{\vartheta}_0) \cap \tilde{U}\right]}_{\neq\emptyset} \neq \emptyset \ .$$

Thus, every future trajectory is dense. Obviously this implies denseness of every trajectory $\tau^{\mathbb{Z}}_{\boldsymbol{\omega}}(\boldsymbol{\vartheta})$.

$5 \to 3$**:** As above, denseness of one trajectory implies denseness of all. Suppose for some $0 \neq \mathbf{k} \in \mathbb{Z}^m$, $q \in \mathbb{Z}$ we have $\mathbf{k} \cdot \boldsymbol{\omega} = q$. Then the continuous function $\Phi : T^m \to \mathbb{C}$ defined by

$$\Phi(\boldsymbol{\vartheta}) := \exp\left[2\pi i \cdot \mathbf{k} \cdot \boldsymbol{\vartheta}\right]$$

is not constant, yet $\tau_{\boldsymbol{\omega}}$-invariant. Thus by lemma A.2.1 the flow $(\tau^n_{\boldsymbol{\omega}})$ can not admit any dense trajectory $\tau^{\mathbb{Z}}_{\boldsymbol{\omega}}\boldsymbol{\vartheta}$, as Φ would be constant on $\tau^{\mathbb{Z}}_{\boldsymbol{\omega}}\boldsymbol{\vartheta}$. [17]

□

An example: Reflecting sphere

Consider a perfectly reflecting hollow unit-sphere, inside it a *trapped* light beam, reflecting on the sphere's surface as described by Fermat's principle. As the beam moves on a plane cross-section of the sphere, we might w.l.o.g. consider the trajectory on the 2D-unit-ball. Clearly the reflection angle and thus the arc-difference $\omega \cdot 2\pi$ between consecutive reflections is constant, which admits a description of the beam's path, or better, the corresponding *reflection points* by the translation flow $(\tau^n_\omega)_{n\in\mathbb{N}_0}$.

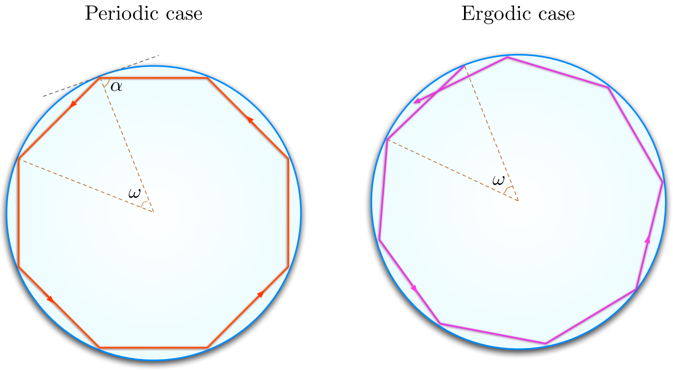

Figure 13: Reflecting light beam inside sphere.

By theorem 4.2.3 we know that the beam's reflection-path can be periodic ($\omega \in \mathbb{Q}$) as well as ergodic ($\omega \notin \mathbb{Q}$). In the later case the reflection points are dense in the circle's boundary. In fact, from this can easily be concluded, that the beam's path is dense inside the *outer-shell* M, enclosed between the reflecting circle and the inner circle (radius $R_i = \cos\omega$) tangential to the beam (see fig. 14).

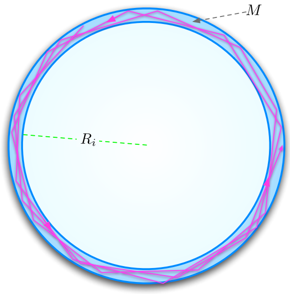

Figure 14: Denseness of beam's path in outer-most shell.

4.2.4 Theorem: Ergodicity of $\{\tau_{\boldsymbol{\omega}}^t\}$

The following statements are equivalent:

1. The flow $(\tau_{\boldsymbol{\omega}}^t)_{t\in\mathbb{R}}$ is ergodic.
2. The flow $(\tau_{\boldsymbol{\omega}}^t)_{t\in\mathbb{R}}$ is strictly ergodic.
3. The frequency vector $\boldsymbol{\omega}$ is *non-resonant*, that is, the numbers $\omega^1, .., \omega^m$ are rationally independent.
4. Some (and thus every) future trajectory $\{\tau_{\boldsymbol{\omega}}^t(\boldsymbol{\vartheta})\}_{t\geq 0}$ is dense in T^m.
5. Some (and thus every) trajectory $\tau_{\boldsymbol{\omega}}^{\mathbb{R}}(\boldsymbol{\vartheta})$ is dense in T^m.

Proof:

$1 \leftrightarrow 2$: Similar to theorem 4.2.3.

$1 \leftrightarrow 3$: See [17] section 4.2.3.

$1 \rightarrow 4$: By lemma 3.3.2 the semi-flow $(\tau_{\boldsymbol{\omega}}^t)_{t\geq 0}$ is also ergodic. By theorem 3.3.6 this implies denseness of almost every future trajectory $\{\tau_{\boldsymbol{\omega}}^t(\boldsymbol{\vartheta})\}_{t\geq 0}$.

$4 \rightarrow 5$: Trivial. Note that, similar to theorem 4.2.3, the denseness of one future trajectory implies the denseness of all future trajectories.

$5 \rightarrow 3$: Suppose $\boldsymbol{\omega}$ is resonant, that is for some $0 \neq \mathbf{k} \in \mathbb{Z}^m$ we have $\mathbf{k}\cdot\boldsymbol{\omega} = 0$. Then the function

$$\Phi(\boldsymbol{\vartheta}) := \exp\left[2\pi i \cdot \mathbf{k}\cdot\boldsymbol{\vartheta}\right]$$

is non-constant, $(\tau_{\boldsymbol{\omega}}^t)$-invariant and continuous on T^m. By lemma A.2.1, this implies that no trajectory $\tau_{\boldsymbol{\omega}}^{\mathbb{R}}\boldsymbol{\vartheta}$ can be dense in T^m.

□

4.2.5 Lemma: Properties of ergodic translation-flows

1. If $(\tau_{\boldsymbol{\omega}}^n)_{n\in\mathbb{Z}}$ is ergodic, then all trajectories $(\tau_{\boldsymbol{\omega}}^n(\boldsymbol{\vartheta}))_n$ are non-periodic[37], that is

$$\forall\ \boldsymbol{\vartheta}\ \ \forall\ n \neq 0\ :\ \tau_{\boldsymbol{\omega}}^n(\boldsymbol{\vartheta}) \neq \boldsymbol{\vartheta}\ .$$

2. If $(\tau_{\boldsymbol{\omega}}^t)_{t\in\mathbb{R}}$ is ergodic, then all trajectories $(\tau_{\boldsymbol{\omega}}^t(\boldsymbol{\vartheta}))_t$ are non-periodic.

[37] In the special case of $n \leq 2$ the converse is also true: If any trajectory is non-periodic, the flow is ergodic.

Proof:

1. We first show that $\tau_{\boldsymbol{\omega}}^n \neq \mathrm{Id} \ \forall \ n \neq 0$. Assume that $\tau_{\boldsymbol{\omega}}^n = \mathrm{Id}$ for some $n \neq 0$. Then $n \cdot \boldsymbol{\omega} = \mathbf{c}$ for some $\mathbf{c} \in \mathbb{Z}^m$, where w.l.o.g. at least one $c^i \neq 0$. Obviously, there exists $0 \neq \mathbf{q} \in \mathbb{Z}^m$ such that $\mathbf{q} \cdot \mathbf{c} = 0$ and thus

$$\mathbf{q} \cdot \boldsymbol{\omega} = \frac{1}{n} \mathbf{q} \cdot \mathbf{c} = 0 \ .$$

   Therefore, $1, \omega^i$ are not rationally independent, thus the flow is non-ergodic. Now consider individual trajectories. Let $n \neq 0$, then $\exists \ \boldsymbol{\vartheta}_0 : \tau_{\boldsymbol{\omega}}^n(\boldsymbol{\vartheta}_0) \neq \boldsymbol{\vartheta}_0$. But then for any $\boldsymbol{\vartheta}$:

$$\tau_{\boldsymbol{\omega}}^n(\boldsymbol{\vartheta}) \overset{\text{def.}}{=} \tau_{n\cdot\boldsymbol{\omega}}(\boldsymbol{\vartheta}) = \tau_{n\cdot\boldsymbol{\omega}} \circ \tau_{\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0}(\boldsymbol{\vartheta}_0) = \tau_{\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0} \circ \underbrace{\tau_{n\cdot\boldsymbol{\omega}}(\boldsymbol{\vartheta}_0)}_{\neq \boldsymbol{\vartheta}_0} \overset{\substack{\tau_{\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0} \\ \text{bijective}}}{\neq} \tau_{\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0}(\boldsymbol{\vartheta}_0) = \boldsymbol{\vartheta} \ .$$

2. Analogous to 1.

□

4.2.6 Notes on translation flows of the torus

We saw that ergodicity of the flows $(\tau_{\boldsymbol{\omega}}^t)$ and $(\tau_{\boldsymbol{\omega}}^n)$ is characterized by the flow frequency vector $\boldsymbol{\omega}$ and $(1, \boldsymbol{\omega})$ respectively. In case of rationally dependent (*resonant*) ω^i (or $1, \omega^i$ in the case of $(\tau_{\boldsymbol{\omega}}^n)$), the torus T^m decomposes into a family of sub-tori $T^k, \ 1 \leq k < m$, restricted to which the flow is again strictly ergodic[15].

As the non-resonant frequency vectors $\boldsymbol{\omega} \in \mathbb{R}^m$ have full Lebesgue measure[38] in $\mathbb{R}^m$, in a sense, almost all flows are strictly ergodic on T^m! It can be shown nonetheless, that the set of resonant frequencies $\boldsymbol{\omega} \in \mathbb{R}^m$ is dense in $\mathbb{R}^m$.

However, the systems $(T^m, \mathscr{B}(T^m), \lambda_{T^m}, (\tau_{\boldsymbol{\omega}}^t))$ and $(T^m, \mathscr{B}(T^m), \lambda_{T^m}, (\tau_{\boldsymbol{\omega}}^n))$ can never be mixing!

4.3 Flows on Liouville-integrable systems

4.3.1 Identification with the torus

Consider a dynamical system described by the $2m$-dimensional symplectic manifold (M, ω), Hamilton function $H \in \mathcal{C}^\infty$ and Hamilton flow $\left(\varphi_{X_H}^t\right)$. Now let $F_1, .., F_m$ be functionally independent, commuting functions, that also commute with H (*integrals* or *constants of motion*).[39]

We then know from Liouville's theorem (see [7]) that the leaf $\mathscr{L}_{\mathbf{f}} := \{\mathbf{F} = \mathbf{f}\}$ (w.l.o.g. pathwise connected) is diffeomorphic to the cylinder $T^k \times \mathbb{R}^{m-k}$ for some $k = 1, ..., m$ and that $H = H(\mathbf{F})$. Furthermore there exist standard coordinates $\vartheta^1, .., \vartheta^m$ on $\mathscr{L}_{\mathbf{f}}$ (that is, $\vartheta^1, .., \vartheta^k$ cyclic), such that in these coordinates the Hamilton-flow[40] takes the form

$$\varphi_{X_H}^t(\boldsymbol{\vartheta}) = \boldsymbol{\vartheta} + \boldsymbol{\omega} \cdot t \quad , \quad \boldsymbol{\omega} = \boldsymbol{\omega}(\mathbf{f}) \ .$$

Recall that the very existence of such an invariant $\mathbf{F}$ precludes the ergodicity of $\left(\varphi_{X_H}^t\right)$ in M or (for $m \geq 2$) even the energy leaf $\{H : \mathrm{const}\}$ (at least with respect to measures equivalent to the Lebesgue-measure)! Nonetheless, the restriction of the flow on $\mathscr{L}_{\mathbf{f}}$ (which actually describes the evolution of the system) might very well be ergodic!

In the special case that the leafs are compact in some neighborhood of $\mathbf{f_0}$ (which holds for many physical systems), these leafs are actually diffeomorphic to the torus $T^n \simeq \mathscr{L}_{\mathbf{f}}$, and there exist symplectic coordinates $\mathbf{s}, \boldsymbol{\vartheta}$, (called *action-angle variables*) such that $\mathbf{F} = \mathbf{F}(\mathbf{s})$ and $\boldsymbol{\vartheta}$ cyclic as above on every $\mathscr{L}_{\mathbf{f}}$! This ensures that the flow actually describes a conditionally periodic motion on the torus as described in section 4.2, for which the question of ergodicity has already been answered! Indeed, the frequencies $\boldsymbol{\omega}(\mathbf{s}) = \frac{\partial H}{\partial \mathbf{s}}(\mathbf{s})$ are solely dependent on

[38] Recall that the set of all $\boldsymbol{\omega} \in \mathbb{R}^m$ with rational dependent components is

$$\bigcup_{0 \neq \mathbf{k} \in \mathbb{Z}^m} \mathbf{k}^\perp \ ,$$

hence a countable union of $\lambda_{\mathbb{R}^n}$-nullsets.

[39] We call such a system *Liouville Integrable*. See [7].

[40] Since $H = H(\mathbf{f})$, the flow $\{\varphi_{X_H}^t\}$ *stays* on the leaf $\mathscr{L}_{\mathbf{f}}$.

the torus considered, and allow for periodic as well as non-resonant solutions. If furthermore, the Hamiltonian H is generic in some neighborhood $\underbrace{U}_{\text{open}} \times T^m$, that is

$$\det\left(\frac{\partial \boldsymbol{\omega}}{\partial \mathbf{s}}\right)\Big|_{U\times T^m} = \det\left(\frac{\partial^2 H}{\partial \mathbf{s}^2}\right)\Big|_{U\times T^m} \neq 0$$

then $\lambda_{\mathbb{R}^m}$-almost every torus in that neighborhood will be non-resonant, whereas the resonant tori will be dense in $U \times T^m$.[20]

We shall briefly consider a few examples of Liouville-integrable systems.

4.3.2 Harmonic oscillators

Consider n independent, harmonic oscillators, each described by the Hamilton function

$$H_i(q_i, p_i) = \frac{\omega_i^2}{2} q_i^2 + \frac{1}{2} p_i^2$$

in symplectic coordinates $\mathbf{q}, \mathbf{p}$. By lemma 3.7.5, examining the joint-system described by

$$H = \sum_{i=1}^{n} H_i$$

is equivalent to examining it as a product of independently evolving sub-systems. Since H_i : const, each sub-system moves at constant *angular velocity* about the torus[41] $T^1 \simeq \{H_i : \text{const}\}$ with the action-angle variables s^i, ϑ^i:

$$\left.\begin{aligned} s^i &= -\oint_{H_i(q_i,p_i):\text{const}} p_i \, dq_i = \frac{H_i}{\pi\omega_i}\int_0^{2\pi} \sin^2\vartheta \, d\vartheta = 2\pi\frac{H_i}{\omega_i} \\ \vartheta^i &= \int_{q_0}^{q} \frac{\partial p_i(s^i, \tilde{q}_i)}{\partial s^i} \, d\tilde{q}_i = \frac{\omega}{2\pi}\int_{q_0}^{q} \frac{d\tilde{q}_i}{\sqrt{2\omega s^i - \omega_i^2 \tilde{q}_i^2}} \overset{\substack{\text{w.l.o.g} \\ q_0=0}}{=} \frac{1}{2\pi}\arcsin\left[\frac{\omega_i q_i}{\sqrt{2H_i}}\right] \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} q_i &= \sqrt{\frac{s^i}{\pi\omega_i}}\sin\left(2\pi\vartheta^i\right) \\ p_i &= \sqrt{\frac{s^i\omega_i}{\pi}}\cos\left(2\pi\vartheta^i\right) \end{aligned}\right. ,$$

$$\dot{\vartheta}^i = \frac{\partial H_i}{\partial s^i} = \frac{\omega_i}{2\pi} \quad \rightarrow \quad \varphi^t_{X_{H_i}}(\vartheta^i, s^i) = \left(\vartheta^i + \frac{\omega_i}{2\pi}t, s^i\right) .$$

Consequently, the whole system moves on an n-dimensional sub-manifold diffeomorphic to $T^n \simeq \{\mathbf{H} : \text{const}\}$, whereas the restriction of the *flow*

$$\left\{\varphi^t_{X_H}\right\} := \left\{\varphi^t_{X_{H_1}} \otimes \cdots \otimes \varphi^t_{X_{H_n}}\right\}_{t\in\mathbb{R}}$$

on T^n is exactly the translation flow considered previously in section 4.2.2 with flow frequencies ω_i acting upon $\boldsymbol{\vartheta}$.

If any of the conditions in 4.2.4 are met, all trajectories of the system will be dense on the leaf $\{\mathbf{H} : \text{const}\}$. Recall that this denseness can be understood in some arbitrary metric like $d^{\boldsymbol{\vartheta}}_{T^n}$ or $d^{\mathbf{q}}_{T^n}$ (see 4.1.2). Figure 15 displays two arbitrary flows for different $\boldsymbol{\omega}$.

[41] Recall that for $2n$-dimensional (symplectic) dynamical systems with n independent integrals F_i, with $\{F_i, F_j\} = 0$, each leaf $L_{\mathbf{f}} := \{\mathbf{F} = \mathbf{f}\}$ (if compact) is diffeomorthetac to the torus T^n. There exist moreover symplectic coordinates $\boldsymbol{\vartheta}, \mathbf{s}$ such that $\mathbf{s} = \mathbf{s}(\mathbf{F})$, $\mathbf{s}\big|_{L_{\mathbf{f}}}$: const and $H = H(\mathbf{s})$.[7]

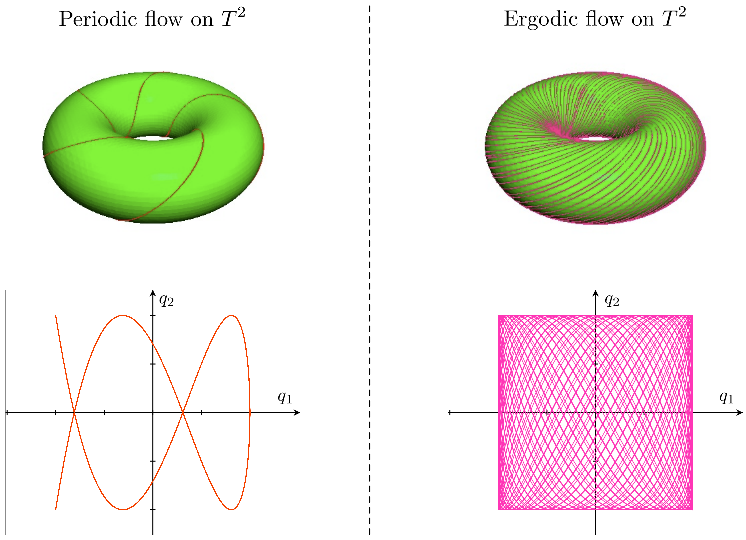

Figure 15: The evolution of two harmonic oscillators (arbitrary units), in standard-torus representation as well as **q**-coordinates, in the periodic and ergodic case. Note that only a fraction of the whole trajectory is displayed in the ergodic case.

Note that even though the denseness of the system-orbit is a coordinate-invariant property, the actual measure to which the flow is ergodic has a different representation (that is, different densities with respect to the induced Lebesgue-measures) in the two coordinate systems $\boldsymbol{\vartheta}$ and $\mathbf{q}$:

$$\rho_{\mathbf{q}} = \rho_{\boldsymbol{\vartheta}} \cdot \left| \det \left(\frac{\partial \boldsymbol{\vartheta}}{\partial \mathbf{q}} \right) \right| \overset{\rho_{\boldsymbol{\vartheta}} \equiv 1}{=} \prod_i \left| \frac{\partial \vartheta^i}{\partial q^i} \right| \overset{\substack{\text{w.l.o.g} \\ \frac{2H^i}{\omega_i^2} = 1}}{=} \prod_i \frac{1}{2\pi \sqrt{1 - q_i^2}} \ .$$

Since the flow is actually strict ergodic (see 4.2.4), this measure is also the only equilibrium probability measure of the system! In a sense, oscillators *tend* to be at their maximal swing-states (compare to fig. 15).

4.3.3 Bowl-travel

Consider a point-particle with unity mass sliding frictionless around the surface of a spherical bowl (radius 1) under influence of gravitational force.

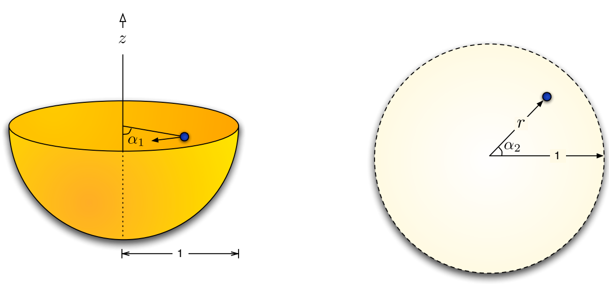

Figure 16: Particle sliding along a spherical bowl.

In symplectic coordinates $\boldsymbol{\alpha}, \mathbf{p}$ the Hamiltonian takes the form

$$H(\alpha_1, \alpha_2, p_1, p_2) = \frac{1}{2} \left[p_1^2 + \frac{p_2^2}{\sin^2 \alpha_1} \right] + \underbrace{(1 - \cos \alpha_1) \cdot g}_{U(\boldsymbol{\alpha})} \ , \quad p_1 := \dot{\alpha}_1 \ , \ p_2 := \dot{\alpha}_2 r^2 \ .$$

The system admits the integrals H and p_2. Since any leaf $\mathscr{L}_{E,P} := \{H = E, p_2 = P\}$ (we shall consider only the case $H, p_2 > 0$) is compact, the system's evolution is restricted to a Torus $T^2 \simeq \mathscr{L}_{E,P}$.

Thus the system undergoes a conditionally periodic evolution, identifiable with a translation-flow around the torus. From section 4.2 we know, that there exist periodic solutions as well as ergodic[42], depending on E, g and P. Trajectories belonging to the later are dense within $\mathscr{L}_{E,P}$! Since the natural projection

$$(\boldsymbol{\alpha}, \mathbf{p}) \overset{\Pi}{\mapsto} \boldsymbol{\alpha}$$

is continuous on $\mathscr{L}_{E,P}$, the trajectory-projection of the system (that is, the *spatial-trajectory* of the particle) is dense within the (compact[43]) projection $\Pi(\mathscr{L}_{E,P})$.[44]

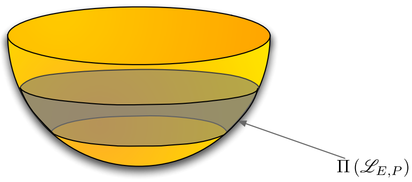

Figure 17: Natural projection of $\mathscr{L}_{E,P}$ in standard representation. The trajectory of the particle is bounded within that more or less narrow band along the bowl and conditionally periodic.

4.4 Flows on nearly Liouville-integrable systems

As we saw above, the phase-spaces of Liouville-integrable systems are composed of flow-invariant tori, on which the flow is, depending on the Hamiltonian $H(\mathbf{s})$, conditionally ergodic. The question arises, as to how perturbation-resistant this structure is, that is, as to wether these resonant or non-resonant orbits cease to exist upon some arbitrarily small modification of the Hamiltonian. Until the fifties it was actually a common belief that arbitrarily small perturbations could destroy these invariant tori and result in an ergodic flow on each energy leaf[2]. This *ergodic hypothesis* was disproved by Kolmogorov in 1954 and later on by Arnold and Moser in the 1960s. The resulting set of theorems and applications is now known as KAM theory, named after its founders. See also [30].

4.4.1 Definition: Diophantine vectors

A vector $\boldsymbol{\omega} \in \mathbb{R}^m$ is called $(\varkappa, \lambda)$-*Diophantine* if, for some $0 < \varkappa$ and $\lambda \geq m - 1$ it satisfies

$$|\boldsymbol{\omega} \cdot \mathbf{k}| \geq \frac{\varkappa}{\|\mathbf{k}\|_1^\lambda} \quad \forall\, \mathbf{k} \in \mathbb{Z}^m \setminus \{0\} \ .$$

Here $\|\mathbf{k}\|_1 := \sum_{i=1}^m \left|k^i\right|$.[18]
Notes:

- Every Diophantine $\boldsymbol{\omega}$ vector is non-resonant, that is, $\omega^1, .., \omega^m$ are rationally independent.
- Denote by $\mathcal{D}^m_{\varkappa,\lambda}$ the set of all $(\varkappa, \lambda)$-Diophantine vectors and $\mathcal{D}^m := \bigcup_{\varkappa,\lambda} \mathcal{D}^m_{\varkappa,\lambda}$. Then $\mathcal{D}^m = \mathbb{R}^m(\text{mod}0)$.

[42]Recall that for the translation flow on T^2, ergodicity is equivalent to non-periodicity.
[43]Continuous images of compact sets are compact.
[44]Recall that for any set A dense in T and continuous $f : T \to K$, the image $f(A)$ is dense in $f(T)$.

4.4.2 Arnold's theorem

Consider the Liouville-integrable Hamiltonian $H(\mathbf{s})$ on the $2m$-dimensional symplectic manifold $(M, d\boldsymbol{\vartheta} \wedge d\mathbf{s})$ in action-angle variables $\mathbf{s}, \boldsymbol{\vartheta}$ and the 1-parameter family of *nearly-integrable* Hamiltonians

$$H_\varepsilon(\boldsymbol{\vartheta}, \mathbf{s}) := H(\mathbf{s}) + \varepsilon \cdot P(\boldsymbol{\vartheta}, \mathbf{s}) \ ,$$

with H and the perturbation P real-analytic in some neighborhood $\tilde{M} \simeq \underbrace{U}_{\subset \mathbb{R}^m} \times T^m$ of the leaf $\mathscr{L}_{\mathbf{s}_0} := \{\mathbf{s} = \mathbf{s}_0\}$.

Let

$$\boldsymbol{\omega} := \left.\frac{\partial H}{\partial \mathbf{s}}\right|_{\mathbf{s}_0} \in \mathcal{D}^m$$

be the frequency vector of the unperturbated flow on $\mathscr{L}_{\mathbf{s}_0}$ and H *generic* in $\tilde{M}$, that is,

$$\det \left(\frac{\partial^2 H}{\partial \mathbf{s}^2}\right)\bigg|_{\tilde{M}} \neq 0 \ .$$

Then for $|\varepsilon|$ small enough, there exists a real analytic embedding[45]

$$\Psi_{\mathbf{s}_0, \varepsilon} : T^m \to \tilde{\mathcal{M}}$$

close to the trivial embedding $(\mathrm{Id}, \mathbf{s}_0)$, such that the torus $\mathscr{L}^\varepsilon_{\mathbf{s}_0} := \Psi_{\mathbf{s}_0,\varepsilon}(T^m)$ is flow-invariant for H_ε and

$$\varphi^t_{X_{H_\varepsilon}} \circ \Psi_{\mathbf{s}_0,\varepsilon}(\boldsymbol{\vartheta}) = \Psi_{\mathbf{s}_0,\varepsilon}(\boldsymbol{\vartheta} + t \cdot \boldsymbol{\omega}) \ .$$

Thus, the torus $\mathscr{L}_{\mathbf{s}_0}$ is only slightly deformed to $\mathscr{L}^\varepsilon_{\mathbf{s}_0}$ while its non-resonant flow is preserved!
See also [13],[14] and [18].

Note: It can even be shown, that for $|\varepsilon|$ small enough, an arbitrary big (in Lebesgue-measure $\lambda_{\boldsymbol{\vartheta},\mathbf{s}}$) measurable set $\mathcal{V}_\varepsilon \subset \tilde{M}$ exists, such that any trajectory starting from $(\boldsymbol{\vartheta}, \mathbf{s}) \in \mathcal{V}_\varepsilon$ is quasi-periodic.[14],[19]. Poincaré furthermore showed, that only finitely many periodic trajectories would survive a perturbation, the rest resulting in chaotic behavior. Thus a dense set of tori is typically destroyed. Alternatively, dropping the assumption of non-degeneracy of H, could destroy all tori and result in ergodic behavior on each energy leaf.[15]

Interpretation: This theorem completely opposes the notion that generic, multi-dimensional nonlinear Hamiltonian systems are ergodic. Most of the invariant tori are preserved provided the perturbation is sufficiently small, and the system is (conditionally) ergodic only on the *deformed* tori $\mathscr{L}^\varepsilon_{\mathbf{s},\boldsymbol{\vartheta}}$. Returning back to example 4.3.3, a sufficiently small *deformation* of the bowl would qualitatively preserve most of the trajectories, which will as a rule only slightly change, still being dense on the invariant leaf.

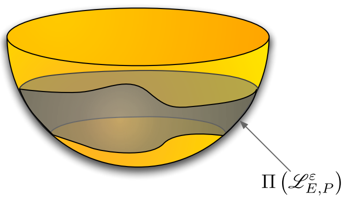

Figure 18: Slightly deformed bowl. Most non-resonant trajectories undergo only minor deviations from their original form and remain dense on the deformed leaf $\mathscr{L}^\varepsilon_{\mathbf{s},\vartheta}$.

Proof: See [7] and [16].

[45] An embedding $f : T \to K$ between to topological spaces T, K is any continuous, injective map. In the context of $\mathcal{C}^\infty$-manifolds an embedding is diffeomorphic to its image.

5 Billiards

An important concrete dynamical system studied in ergodic theory by mathematicians and physicists as well, is the so called *billiard*. The system basically consists of a single point-particle, moving *freely* at constant speed 1 within some bounded region, *bouncing of* at the boundary under the law *incidence angle equals reflection angle*. Questions about existence of periodic solutions, invariant set, denseness of trajectories and in general ergodicity of this system are of great theoretical and even practical importance, as billiards can be used to model many physical systems (e.g. free particles within a box).

One example note-worthy is the Lorentz-gas, which tries to model *freely* moving electrons *bouncing of* at heavy ions, as billiards within some domain with a number fixed, non-intersecting balls *removed*.[25]

Furthermore, many other problems, like systems of N absolutely elastic spheres in some compact domain $\subset \mathbb{R}^n$, can be reduced to billiards in some domain of $\mathbb{R}^{n\cdot N}$ [3]. This diversity of applications makes the analysis and classification of billiards an interesting and promising subject.

5.1 Construction of Billiards

5.1.1 Definition: Billiard table

Let (M, g) be a n-dimensional, pathwise-connected, compact Riemann manifold with non-degenerate smooth boundary[46] ∂M and X the vector-field corresponding to the geodesic flow on the tangent-bundle TM[47]. By construction of the geodetic flow, the unit-tangent-bundle

$$\mathfrak{B}_M := \{(q,v) \in TM : g(v,v) = 1\}$$

is flow invariant, that is, $X \in T\mathfrak{B}_M$. The system $\left(\mathfrak{B}_M, X|_{\mathfrak{B}_M}\right)$ shall be called a *billiard table* with *smooth boundary*.

Define the natural projection $\Pi : \mathfrak{B}_M \to M$ by $\Pi(q,v) = q$, then

$$\partial\mathfrak{B}_M = \Pi^{-1}(\partial M) = \{(q,v) \in \mathfrak{B}_M : q \in \partial M\} \ .$$

Now consider the *reflection-map* $R : \partial\mathfrak{B}_M \to \partial\mathfrak{B}_M$ defined by

$$R(q,v) := (q, v - 2 \cdot g(\eta_q, v) \cdot \eta_q) \ ,$$

whereas η_q denotes the inner-normal-unit-vector to $T_q\partial M$. For $v \in T_qM$ we shall sometimes write $Rv := v - 2 \cdot g(\eta_q, v) \cdot \eta_q$. Note that $R^{-1} = R$.

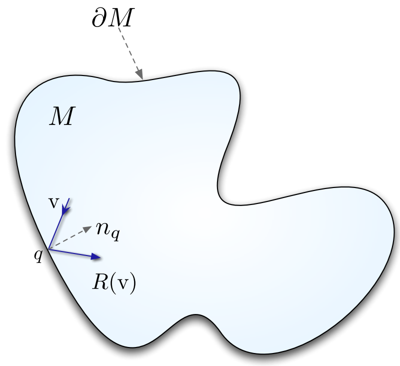

Figure 19: On the definition of billiard tables and the reflection map. Notice the connection between R and the usual principle *incidence angle equals reflection angle*.

[46] We call a subset $M \subset M_0$ of a n-dimensional Riemann manifold (M_0, g) a *compact Riemann-manifold with smooth, non-degenerate boundary* if M can be written as

$$M = \{q \in M_0 : f \geq 0\}$$

for some $f \in \mathcal{C}^\infty(M)$ such that M is compact and $\operatorname{grad} f \neq 0$ on $\partial M = f^{-1}(0)$.

[47] One can introduce the geodesic flow on TM as the flow corresponding to the vector field

$$X_{(q,v)} \cong \left(v^1, .., v^n, -\Gamma^1_{kr} v^k v^r, .., -\Gamma^n_{kr} v^k v^r\right)$$

in some coordinates $x^1, .., x^n$ and $\partial_1, ..., \partial_n \in T_xM$. Note that this flow is equivalent to the flow in T^*M with respect to the Hamiltonian $H(q,p) = \tilde{g}_q(p,p)$ under the transformation $p = Jv$.

We shall call the billiard-table *proper*, if for all $x \in \mathfrak{B}_M$ the geodesic $\varphi_X^t x$ is *incomplete* in both time-directions, that is, $\varphi_X^t x$ reaches the boundary $\partial\mathfrak{B}_M$ within a finite (positive and negative) time $|t| < \infty$. From now on, we assume all billiard-tables to be proper.[3]

5.1.2 Construction of the billiard flow

We shall now construct an $\mathbb{R}$-flow $(\tau^t)_{t\in\mathbb{R}}$ on $\mathfrak{B}_M$ as follows:

- For any point $(q, v) \in \mathfrak{B}_M$ set $\tau^t(q, v) := \varphi_X^t(q, v)$ to be the geodesic flow with respect to the proper-time t, for as long as $\varphi_X^t(q, v)$ is defined (that is, until $\varphi_X^t x$ reaches $\partial\mathfrak{B}_M$).
- For any point $(q, v) \in \partial\mathfrak{B}_M$ with $g(\eta_q, v) < 0$ (v pointing *outwards*) set $\tau^t(q, v) := \tau^t \circ R(q, v)$ for as long as it is defined, that is, *reflect* the *ball* and follow the flow *in the new direction.* Note that $R(q, v)$ *faces inwards* and since the boundary is non-degenerate, its future flow is defined at least for some positive time.
- Complete the constructed semi-flow for each point by the rule $\tau^{t_1+t_2} := \tau^{t_1} \circ \tau^{t_2}$. Since τ^t is bijective, complete the flow into negative times $\tau^{-t} := (\tau^t)^{-1}$.

We call this $\mathbb{R}$-flow *billiard flow on M* and the projection $\Pi(\tau^{\mathbb{R}} x)$ of a trajectory the *spatial* trajectory in M.[3]

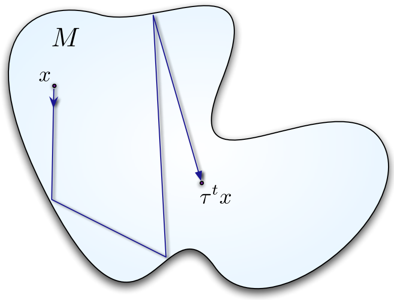

Figure 20: On the definition of the billiard flow.

Note: Actually, there exists the possibility that some point is *reflected* on the boundary infinitely often within a finite time. Though in such cases the flow is not defined beyond that finite time, we will later on show that in the defined constructed measure, these pathological points form a null-set (5.1.5).

5.1.3 Construction of an invariant measure

Define $x^- \in \partial\mathfrak{B}_M$ to be the *previous* reflection point of $x \in M$, that is $x = \tau^{t_x} x^-$ for some minimal $t_x > 0$. Since the billiard table is proper, x^- always exists. The association $x \mapsto (t_x, x^-)$ is bijective, thus we can introduce new coordinates in $\mathfrak{B}_M$ by using the already existent coordinates in $\partial\mathfrak{B}_M$ and the *time* t_x *needed* to get from $\partial\mathfrak{B}_M$ up to the given point.

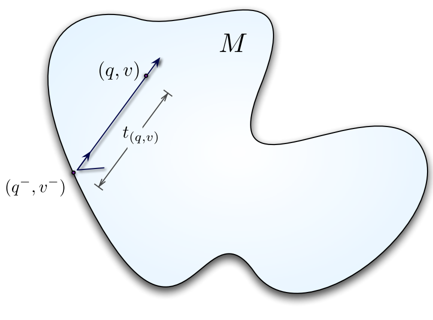

Figure 21: On the definition of time-boundary coordinates.

Let $V^{\partial M}$ be the volume-form induced by the metric g on ∂M and $V_q^{S^{n-1}}$ the standard volume form[48] on the unit-circle $S_q^{n-1} := \{v \in T_q M : g(v,v) = 1\}$. Then

$$d\mu_{(q,v)}^{\partial \mathfrak{B}_M} := |g(n_q, v)| \cdot V_q^{\partial M} \wedge V_{(q,v)}^{S^{n-1}}$$

defines a measure on the billiard-boundary $\partial \mathfrak{B}_M$. Finally, define the volume-form (measure)

$$d\mu_x := dt_x \wedge d\mu_{\tau^{-t_x}x}^{\partial \mathfrak{B}_M} \tag{5.1.3.1}$$

on $\mathfrak{B}_M$. It turns out

$$d\mu = V^M \wedge V^{S^{n-1}} \ , \tag{5.1.3.2}$$

where V^M is the volume form induced by g in M, and that μ is actually (τ^t)-invariant (see [3]).

Example: For $M \subset \mathbb{R}^n$ the volume form $V^{\partial M}$ corresponds merely to the Lebesgue-measure of the $(n-1)$-dimensional surface ∂M and $V^{S^{n-1}}$ to the $(n-1)$-dimensional solid-angle-measure. Similarly, V^M corresponds simply to the Lebesgue-measure in M.

Note: Since both μ and $\mu^{\partial \mathfrak{B}_M}$ are finite, they can be turned into an invariant probability measures

$$\tilde{\mu} := \frac{1}{\mu(\mathfrak{B}_M)} \cdot \mu \quad , \quad \mu(\mathfrak{B}_M) = \mathrm{vol}_{V^M}(M) \cdot \underbrace{\mathrm{vol}(S^{n-1})}_{\substack{\text{standard} \\ \text{unit-sphere} \\ \text{surface} \\ \text{volume}}} \ ,$$

$$\tilde{\mu}^{\partial \mathfrak{B}_M} = \frac{1}{\mu^{\partial \mathfrak{B}_M}(\partial \mathfrak{B}_M)} \cdot \mu^{\partial \mathfrak{B}_M} \quad , \quad \mu^{\partial \mathfrak{B}_M}(\partial \mathfrak{B}_M) = \mathrm{vol}_{V^{\partial M}}(\partial M) \cdot \mathrm{vol}(B_1^{n-1}) \ ,$$

whereas B_1^{n-1} is the unit-ball[49] in $\mathbb{R}^{n-1}$.

5.1.4 The return map

In direct analogy to the Poincaré return map (3.1.6), define the *inward-boundary*

$$\partial \mathfrak{B}_M^+ := \{(q,v) \in \partial \mathfrak{B}_M : g(n_q, v) > 0\}$$

and set for $(q,v) \in \partial \mathfrak{B}_M$ the *reflection time* $r(q,v) > 0$ as the minimum time needed to get (or return) to $\partial \mathfrak{B}_M^+$. The set of points with $v \in T_q \partial M$ are clearly a $\mu^{\partial \mathfrak{B}_M}$-nullset, so that $r(q,v)$ is a.e. well defined. Finally, define the return map $\mathcal{R}(x) := \tau^{r(x)} x$.

Notes:

- It can be shown, that $\mathcal{R}$ preserves the measure $\mu^{\partial \mathfrak{B}_M}$ on $\partial \mathfrak{B}_M^+$ (see [3]).
- Consider the case $M \subset \mathbb{R}^2$: Topologically $\partial \mathfrak{B}_M^+$ is a cylinder, with cyclic coordinate s the *arclength* of the boundary ∂M and *axis* coordinate the angle $\alpha \in [0, \pi]$ characterizing every *inward turned* unit-vector.

[48] Choose some orthonormal basis $b_1, .., b_n$ in $T_q M$ and associate S_q^{n-1} by the resulting coordinates with the unit-circle S^{n-1} in $\mathbb{R}^n$.

[49] Note that

$$\int_{S^{n-1}} |g(\eta, v)| \, V^{S^{n-1}} \overset{\text{w.l.o.g.}}{=} \int_{-1}^{1} dx_1 \ \ |x_1| \, \mathrm{vol}_{n-2}\left(\partial B^{n-1}_{\sqrt{1-x_1^2}}\right) = \int_0^1 du \cdot \mathrm{vol}_{n-2}\left(\partial B^{n-1}_{\sqrt{1-u}}\right) = \mathrm{vol}_{n-1}(B_1^{n-1}) \ .$$

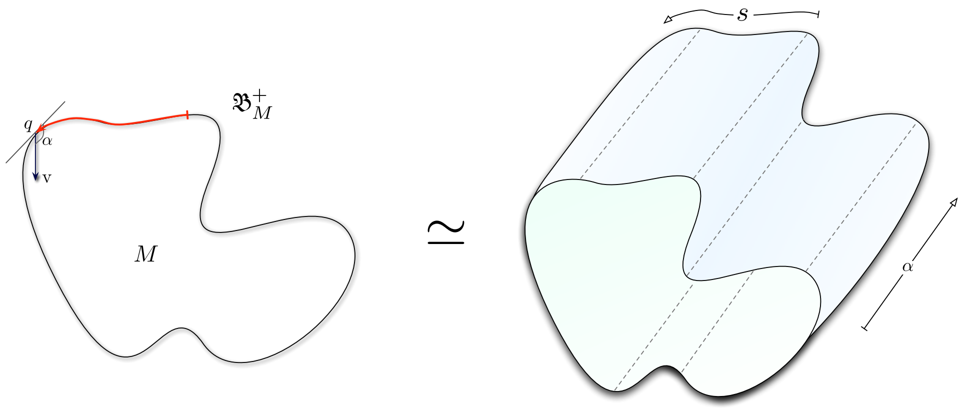

Figure 22: On the topology of the inward-boundary $\partial\mathfrak{B}_M^+$.

5.1.5 Lemma about pathological points

Denote by $\mathfrak{T}$ the set of points in M that are reflected infinitely often within a finite time. Then $\mu(\mathfrak{T}) = 0$.

Proof: We shall elaborate on the proof in [3]. Let $r(x)$ be the time until *the next reflection* on $\partial\mathfrak{B}_M$, that is, $\tau^{r(x)}x \in \partial\mathfrak{B}_M$ for some minimum $r(x) \geq 0$. Clearly, if $x = \tau^{t_x}(x^-) \in \mathfrak{T}$ then so is $\tau^t(x^-)$ for all $0 \leq t \leq r(x^-)$. Thus

$$\mu(\mathfrak{T}) = \int\limits_{\mathfrak{T}\cap\partial\mathfrak{B}_M^+} r(x^-)\, d\mu^{\partial\mathfrak{B}_M} .$$

Since $r(x^-) > 0$ for $\mu^{\partial\mathfrak{B}_M}$-almost all $x^- \in \partial\mathfrak{B}_M^+$, by corollary A.3.1

$$\sum_{n=0}^{\infty} r\left(\mathcal{R}^n x^-\right) = \infty$$

for $\mu^{\partial\mathfrak{B}_M}$-almost all $x^- \in \partial\mathfrak{B}_M^+$. But on the other side, $\mathfrak{T} \cap \partial\mathfrak{B}_M^+$ are exactly those $x^- \in \partial\mathfrak{B}_M^+$ for which $\sum_{n=0}^{\infty} r\left(\mathcal{R}^n x^-\right)$ remains finite. Thus $\mu(\mathfrak{T}) = 0$.
□

5.1.6 Lemma: Mean free path of a billiard

Let $\mathfrak{B}_M$ be a billiard table with the billiard flow (τ^t) and $r : \partial\mathfrak{B}_M \to (0, \infty)$ the *reflection time* map (see 5.1.4). Then

$$\langle r\rangle_{\partial\mathfrak{B}_M} := \int\limits_{\partial\mathfrak{B}_M} r\, d\tilde{\mu}^{\partial\mathfrak{B}_M} = \frac{\mathrm{vol}_{V^M}(M)\cdot \mathrm{vol}(S^{n-1})}{\mathrm{vol}_{V^{\partial M}}(\partial M)\cdot \mathrm{vol}(B^{n-1})} .$$

(see also [21]).

Interpretation: Since we assumed the *speed* of a billiard to be 1 ($g(v, v) = 1$), the *time* $r(q, v)$ *needed* until the next reflection, is simply the *free path* of a billiard *bouncing of* from the boundary $q \in \partial M$ with direction v. This lemma thus points out the surprising fact that, the *mean free path* of billiards only depends on the volumes of the table and its boundary! Specifically for $M \subset \mathbb{R}^2$ this yields the known Santalo formula[22]

$$\langle r\rangle_{\partial\mathfrak{B}_M} = \frac{\pi \cdot \mathrm{Vol}(M)}{\mathrm{Vol}(\partial M)} .$$

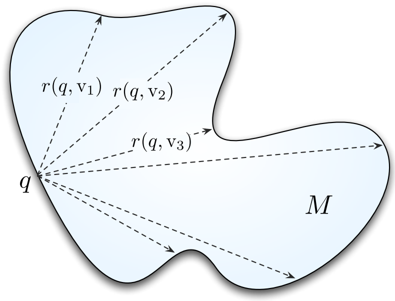

Figure 23: On the definition of the mean free path.

Note that in case of an ergodic billiard-flow on a billiard-table with r bounded, for almost any start-point $x \in \partial\mathfrak{B}_M$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} r\left(\mathcal{R}^n x\right) = \langle r \rangle_{\partial\mathfrak{B}_M}$$

(see theorem 3.3.6).

Proof:

$$\int\limits_{\partial\mathfrak{B}_M} r(x)\ d\tilde{\mu}^{\partial\mathfrak{B}_M} = \frac{1}{\mathrm{vol}_{V^M}(\partial M)\cdot \mathrm{vol}(B_1^{n-1})} \cdot \int\limits_{\partial\mathfrak{B}_M} \underbrace{r(x)}_{\int_0^{r(x)} dt}\ d\mu^{\partial\mathfrak{B}_M}$$

$$\overset{(5.1.3.1)}{=} \frac{1}{\mathrm{vol}_{V^M}(\partial M)\cdot \mathrm{vol}(B_1^{n-1})} \cdot \int\limits_{\mathfrak{B}_M} \underbrace{dt_x \wedge d\mu^{\partial\mathfrak{B}_M}_{\tau^{-t_x}x}}_{d\mu} = \frac{\mathrm{vol}_{V^M}(M)\cdot \mathrm{vol}(S^{n-1})}{\mathrm{vol}_{V^{\partial M}}(\partial M)\cdot \mathrm{vol}(B^{n-1})}\ .$$

□

5.2 Billiards in ellipses

We shall in the following section consider the ellipse

$$\mathcal{E}_c := \left\{ q \in \mathbb{R}^2 : d(q, F_1) + d(q, F_2) \le c \right\}$$

in $\mathbb{R}^2$ (standard metric) with foci at points F_1, F_2 and the induced billiard-table $\mathfrak{B}_{\mathcal{E}_c}$ with billiard flow (τ^t).[50]

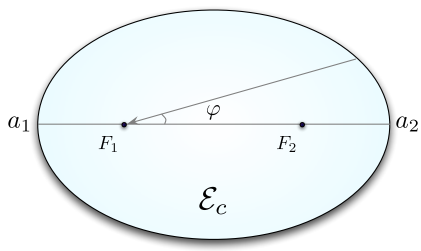

Figure 24: Ellipse with foci F_1, F_2.

[50] See also [3] and [4].

5.2.1 Proposition: Convergence of focal trajectories in ellipses

For all (q, v) such that $q \in \{F_1, F_2\}$, the sequence of reflection points of the spatial trajectory $\{\Pi\tau^t(q, v)\}_{t\geq 0}$ on $\partial\mathcal{E}_c$ can be decomposed into two sequences, each converging to one of the points a_1, a_2. Put simply, the trajectory of any billiard starting at a focus, *converges* towards the major axis (see fig. 25).

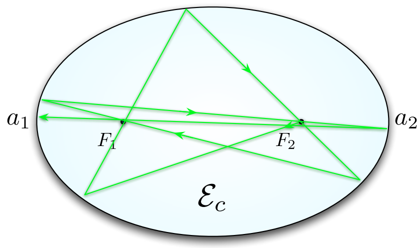

Figure 25: On the *convergence* of focal billiard trajectories towards the major axis.

Interpretation: This result might seem odd at first, since the recurrence map $\mathcal{R}$ preserves the strictly positive, finite measure $\mu^{\partial\mathfrak{B}_{\mathcal{E}_c}}$, thus by corollary 3.1.4 one would expect that almost all reflections *repeat* themselves arbitrarily accurately. But indeed, the set of all reflection points $(q, v) \in \partial\mathfrak{B}_{\mathcal{E}_c}$ leading to a transition through a focus, has measure zero.

Note that due to symmetry, the same statement can be made about the history trajectory! Focal trajectories converge to the major axis with $t \to \infty$ as well as $t \to -\infty$.

Proof: W.l.o.g. let $d(F_1, F_2) > 0$ (otherwise all focal trajectories are fixed). W.l.o.g. we shall show the proposition for all trajectories passing through F_1 at time $t = 0$. For this proof, we shall identify each *incoming* direction v of some $(q, v) \in \Pi^{-1}(F_1)$ with the angle φ between v and the positively oriented major axis (see fig. 24).

By lemma A.2.3 we know that billiards starting at a focus, reflect on the boundary and then pass through the other focus! Thus for every $(q, v) \in \Pi^{-1}(F_1)$ we can define a mapping $\Phi : S^1_{F_1} \to S^1_{F_1}$ mapping the incoming direction of a billiard at F_1 (characterized by the angle $\varphi \in [0, 2\pi)$) to the next incoming direction at F_1, after being reflected two times at $\partial\mathcal{E}_c$ (see fig. 26).

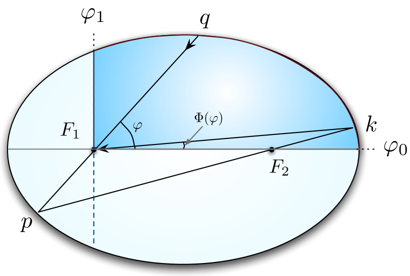

Figure 26: On the definition of $\Phi : S^1_{F_1} \to S^1_{F_1}$.

From fig. 26 it is clear, that $\Phi : (0, \pi/2] \to (0, \pi/2]$ is strongly monotonically decreasing and continuous. Thus for any direction $\varphi \in (0, \pi/2]$ the sequence $(\Phi^n\varphi)_n$ is strongly monotonically decreasing and bounded below by 0, thus convergent. Further

$$\Phi\left[\lim_{n\to\infty} \Phi^n\varphi\right] = \lim_{n\to\infty} \Phi(\Phi^n\varphi) = \lim_{n\to\infty} \Phi^n\varphi \ ,$$

that is, the limes is a fixed point of Φ. But this can only hold for $\lim\limits_{n\to\infty} \Phi^n\varphi = 0$, hence any incoming direction $\varphi \in [\varphi_0, \varphi_1] = [0, \pi/2]$ leads to convergence of the subsequent reflection points towards the major axis.

Now suppose all directions coming from $[0, \varphi_n]$, $\frac{\pi}{2} \leq \varphi_n < \pi$ subsequently lead to trajectories converging towards the major axis. Then this holds even for some bigger interval $[0, \varphi_{n+1}]$ (whereas directions within

$[0, \varphi_{n+1}]$ after 2 reflections lead to directions coming from $[0, \varphi_n]$), with $\varphi_{n+1} > \varphi_n$ given by

$$\pi - \varphi_{n+1} = \Phi(\pi - \varphi_n)$$

(see fig. 27).

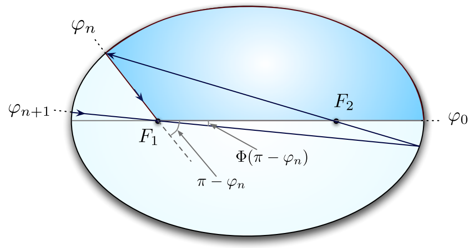

Figure 27: On the extension of *convergence directions*. Convergent of directions coming from $[0, \pm\varphi_n]$ implies convergence of directions coming from an even bigger interval $[0, \varphi_{n+1}]$.

Thus, we can construct some (monotonically increasing) sequence $(\varphi_n)_n \subset \left[\frac{\pi}{2}, \pi\right]$ such that for any $n \in \mathbb{N}$, directions coming from $[0, \varphi_n]$ lead to convergence, and $\pi - \varphi_{n+1} = \Phi^n(\pi - \varphi_1)$. But, as we saw above

$$\lim_{n\to\infty} (\pi - \varphi_{n+1}) = \lim_{n\to\infty} \Phi^n(\underbrace{\pi - \varphi_1}_{\in[0,\frac{\pi}{2}]}) = 0 \ ,$$

hence $\varphi_n \overset{n\to\infty}{\longrightarrow} \pi$. The case $\varphi \in \pi$ is trivial and symmetry of the problem implies that the above holds for $\varphi \in [\pi, 2\pi]$ as well, hence, the proposition is proved.
□

5.2.2 Proposition: Restriction of non-focal trajectories in ellipses

Let $\Pi\tau^{\mathbb{R}}x \cong \underbrace{\ldots q_1 q_2 q_3 \ldots}_{\text{reflection points}}$ be a configurational trajectory of the billiard flow in $\mathcal{E}_c$ that does not pass through any of the foci F_1, F_2. Then either all segments $q_i q_{i+1}$ are tangent to one and the same ellipse $\mathcal{E}_d$ (with foci F_1, F_2) or all segments $q_i q_{i+1}$ (i.e. their extensions) are tangent to one and the same hyperbola $\mathcal{H}_d$ (with foci F_1, F_2).[3]

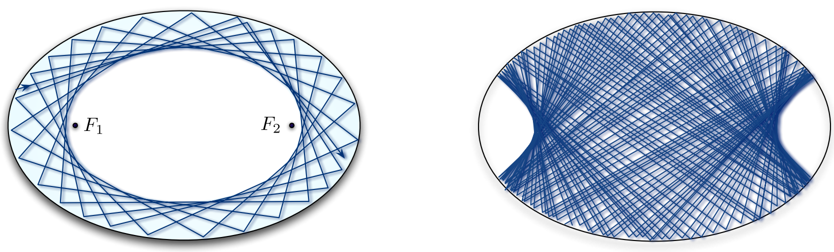

Figure 28: Restriction of non focal spatial trajectories outside some ellipse $\mathcal{E}_d$ or inside some hyperbola $\mathcal{H}_d$.

Proof: Let $q_1 q q_2$ be consecutive reflection points on $\partial\mathcal{E}_c$. By lemma A.2.3 we know that either both $\overline{q_1 q}$ and $\overline{q q_2}$ or none of them intersect the segment $\overline{F_1 F_2}$. Consider the first case, and construct two ellipses $\mathcal{E}_{c_1}, \mathcal{E}_{c_2}$ with foci F_1, F_2 such that the segments $\overline{q_1 q}$, $\overline{q q_2}$ are tangent to $\partial\mathcal{E}_{c_1}, \partial\mathcal{E}_{c_2}$ on the points p_1, p_2 respectively.

It suffices for the proposition to show that $c_1 = c_2$, i.e. both segments are tangent to one and the same ellipse. Reflect the foci F_1, F_2 on the segments $\overline{q_1 q}$ and $\overline{q q_2}$ respectively to obtain the reflections F_1', F_2' (see fig. 29).

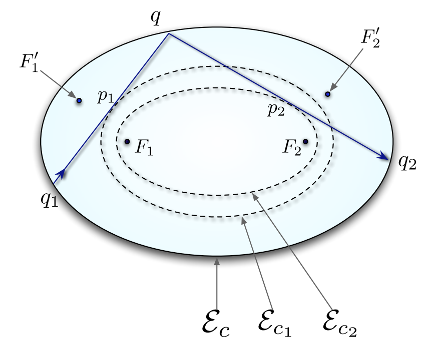

Figure 29: On the proof of proposition 5.2.2.

Then, since $p_1 \in \overline{F_2 F_1'}$ (see proof of lemma A.2.3), we have

$$c_1 = \underbrace{d(F_1, p_1)}_{d(F_1', p_1)} + d(F_2, p_1) = d(F_2, F_1')$$

and similarly $c_2 = d(F_1, F_2')$. Furthermore, since the configurational trajectory $F_1 q F_2$ would be valid, the angles $\widehat{F_1 q q_1}$, $\widehat{F_2 q q_2}$ are equal (reflection angle axiom). Since $\widehat{F_1 q q_1} = \widehat{q_1 q F_1'}$ and $\widehat{F_2 q q_2} = \widehat{q_2 q F_2'}$ it follows

$$\widehat{F_2 q F_1'} = \widehat{F_1 q F_2'}$$

(see fig. 30).

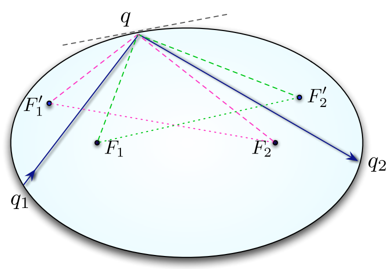

Figure 30: On the proof of proposition 5.2.2.

Moreover, $d(F_1' q) = d(F_1, q)$ and $d(F_2, q) = d(F_2' q)$, consequently the triangles $F_1' q F_2$ and $F_1 q F_2'$ are equal, hence

$$c_1 = d(F_2, F_1') = d(F_1, F_2') = c_2 \ .$$

In case where $\overline{q_1 q}$ and $\overline{q q_2}$ intersect the segment $\overline{F_1 F_2}$, considering the hyperbolas $\mathcal{H}_{c_1}, \mathcal{H}_{c_2}$ allows for the same arguments.

□

5.2.3 Corollary: Non-ergodicity of ellipse-billiards

For any ellipse $\mathcal{E}_c$ (foci $F_1, F_2,\ c > d(F_1, F_2)$), the billiard flow on $\mathfrak{B}_{\mathcal{E}_c}$ is non-ergodic.

Elaboration: Proposition 5.2.2 shows that, almost all trajectories are restricted to invariant null-sets of the type

$$\left\{ (q, v) \in \mathfrak{B}_{\mathcal{E}_c} : \Pi\left(\tau^{\mathbb{R}}(q, v)\right) \text{ tangent to } \partial\mathcal{E}_d \ \ (\partial\mathcal{H}_d) \right\}$$

for some d. The proof made actually evident, that even the spatial trajectories $\Pi(\tau^{\mathbb{R}}(q, v))$ are nowhere near to being dense in $\mathcal{E}_c$ (see fig. 28)!

Proof: For any $d(F_1, F_2) < c_1 < c_2 \le c$ the set

$$\left\{(q,v) \in \mathfrak{B}_{\mathcal{E}_c} : \Pi\left(\tau^{\mathbb{R}}(q,v)\right) \text{ tangential to some } \partial\mathcal{E}_d,\ c_1 \le d \le c_2\right\}$$

is billiard-flow-invariant and has positive, non-full measure.
□

6 Summary

In this paper, we introduced and examined so called group-flows on measurable phase spaces, vis-á-vis classical Hamiltonian flows on symplectic manifolds. Statistical uncertainty, a feature with which many a natural scientist is confronted when performing experiments, is abstracted by probability measures introduced on the phase spaces of the systems. We concentrate on so-called equilibrium distributions, that is, probability measures invariant under the flow describing the time-evolution of the system.

One of the first fruits of ergodic theory was the Poincaré recurrence theorem, securing virtually unconditional recurrent behavior in a big class of dynamical systems, especially those in statistical equilibrium. It was shown that, following from this theorem, Hamiltonian trajectories become arbitrary close to repeating them selves, provided the measure meets certain *compatibility conditions* with the existent topology of the space. Even though the Poincaré theorem provides no estimation for return-times, it nonetheless constituted an important finding, leading to hot debates regarding the irreversibility of natural processes, as prescribed by thermodynamical laws[23].

An important milestone was the Birkhoff-Khinchin ergodic theorem, securing the existence (a.e.) of time-averages along trajectories, thus providing a justification for the practice of time-averaging conducted in common experiments. One elegant result of this theorem is the intuitively already suspected connection between the sojourn-time of a trajectory within a set A and its measure $\mu(A)$.

Yet, the theorem does not provide any connection to the so sought-after phase-average, which turns out to be a non-trivial property of only a special kind of systems, namely ergodic ones. These systems are exactly those, whose flow does not admit the existence of invariant sets other than the space its self (mod0) or nullsets. Ergodicity is to be taken as a purely measure-theoretic characteristic of the flow within the phase-space, inevitably depending on the probability distribution assumed. A connection to the topology is in general only existent, if the measure its self exhibits a certain connection to the former. It was shown, that in most *mainstream* spaces, ergodic flows produce dense trajectories and that non-null-sets actually fill up the entire space under the action of the flow. Furthermore, it was demonstrated that ergodicity is invariant under equivalent measures and metrically isomorphic systems.

An important result in this paper, is the generalization of the Krylov-Bogoliubov-theorem, manifesting the existence of invariant and even ergodic probability measures, for continuous time-flows on compact, metric spaces. However, these may lack any physical meaning, and may not even be realized in usual experimental setups. In connection, the stronger concept of strict-ergodicity is briefly introduced, expressing the uniqueness of equilibrium probability distributions in certain systems, leading to an even richer family of characteristics, such as the uniform convergence of time- to phase-averages.

Following up, the promising concept of mixing is introduced, a property showing a lot of similarities with the behavior of many complex physical systems. It expresses the notion, that sets of states are under the system's flow, *whirled* around the phase space, seemingly getting dispersed as time passes. In case of strictly positive, topological measure spaces, mixing systems show a sensitivity on initial conditions, typically found in chaotic systems. It turns out, that systems described by a certain kind of relaxation tendency, are actually mixing, thus promising the presence of the mixing property in a wide class of physical systems.

The second major part of this article dealt with flows on manifolds, specifically Hamiltonian ones. We started with the continuity equation for the probability density under the flow action and continued by showing the existence of an invariant measure on energy leafs, the so called microcanonical distribution. Subsequently, we examined the translation group on the torus T^n, and outlined a characterization of its ergodicity. It turned out, that in the later case, the Lebesgue measure was actually the only invariant measure on the torus, implying the strict-ergodicity of the flow in the non-resonant case. Through the Liouville theorem for integrable systems, a connection from the torus-action to a wide class of Hamiltonian systems was presented, securing the ergodic behavior of these systems on certain flow-invariant tori, typically preserved even under small, non-linear perturbations!

In the final section we considered one of the standard models encountered in ergodic theory, the so called billiards. After demonstrating the existence of an invariant measure and determining the mean free path of a general billiard, we paradigmatically dedicated the last few pages to the ellipse-billiards, actually demonstrating its non-ergodicity. More on billiards can be found in [24].

A Appendix

A.1 Proof of Birkhoff's ergodic theorem

A.1.1 Maximal ergodic theorem

Let $(M, \mathcal{M}, \mu)$ be a measure space and $T : L_1(M, \mathcal{M}, \mu) \to L_1(M, \mathcal{M}, \mu)$ a positive[51] contraction[52]. For any $f \in L_1(M, \mathcal{M}, \mu)$ let

$$S_n f := \sum_{k=0}^{n-1} T^k f \quad , \quad A_n f := \frac{S_n f}{n} \ ,$$

$$M_n^S f := \max\{S_1 f, .., S_n f\} \ , \ M_n^A f := \max\{A_1 f, .., A_n f\} \ ,$$

$$P_n f := \underbrace{\left\{M_n^S f \geq 0\right\}}_{\{M_n^A f \geq 0\}} \ , \ P_\infty f := \bigcup_{n \in \mathbb{N}} P_n f \ .$$

Then

$$\int_{P_n f} f d\mu \geq 0 \quad , \quad \int_{P_\infty f} f \ d\mu = 0 \ .$$

Proof: Following [1], by construction

$$(M_n^S f)^+ \geq M_n^S f \geq S_n f$$

for $k = 1, .., n$ and hence

$$f + T(M_n^S f)^+ \overset{T \text{ positive}}{\geq} f + T S_k f = S_{k+1} f \ .$$

Thus

$$f \geq S_k f - T(M_n^S f)^+ \quad , \quad k = 1, .., n$$

since the case $k = 1$ is trivial and so

$$f \geq \max_{k=1,..,n} \{S_k f\} - T(M_n^S f)^+ = M_n^S f - T(M_n^S f)^+ \ .$$

Integrating over P_n yields

$$\int_{P_n f} f \ d\mu \geq \int_{P_n f} \left[M_n^S f - T(M_n^S f)^+\right] \ d\mu \overset{(M_n^S f)\big|_{P_n f} \geq 0}{=} \int_{P_n f} \left[(M_n^S f)^+ - T(M_n^S f)^+\right] \ d\mu$$

$$\overset{(M_n^S f)^+\big|_{(P_n f)^c} = 0}{=} \int_M (M_n^S f)^+ d\mu - \int_{P_n f} \underbrace{T(M_n^S f)^+}_{\geq 0} d\mu \geq \int_M (M_n^S f)^+ d\mu - \underbrace{\int_M T(M_n^S f)^+ d\mu}_{\substack{\leq \int_M (M_n^S f)^+ d\mu \\ \text{since } \|T\| \leq 1 \\ \text{and } T \geq 0}} \geq 0 \ .$$

Further, since $M_n^S f \leq M_{n+1} f$ and thus $P_n f \subset P_{n+1} f$ one has $1_{P_n f} \leq 1_{P_{n+1} f}$ and so

$$1_{P_\infty f} = 1_{\bigcup_n P_n f} = \sup_{n \in \mathbb{N}} \{1_{P_n f}\} = \lim_{n \to \infty} 1_{P_n f} \quad \text{(Pointwise)} \ .$$

[51] An operator $T : V \to W$ between vector spaces with partial order, is *positive* $(T \geq 0) :\Leftrightarrow \ T\{v \in V : v \geq 0\} \subset \{w \in W : w \geq 0\}$.
[52] A bounded, linear operator $T : V \to V$ in the normed vector space $(V, \|\cdot\|)$ is called a *contraction* $:\Leftrightarrow \ \|T\| \leq 1$.

Using the generalized Lemma of Fatou we can carry out the limit

$$\int_{P_\infty f} f \; d\mu = \int_M f \cdot 1_{P_\infty f} \; d\mu = \int_M \underbrace{\lim_{n\to\infty} \left(f \cdot 1_{P_n f}\right)}_{\limsup_{n\to\infty} (f \cdot 1_{P_n f})} \; d\mu \overset{\text{Fatou}}{\geq} \limsup_{n\to\infty} \underbrace{\int_M f \cdot 1_{P_n f} \; d\mu}_{\int_{P_n f} f \; d\mu \geq 0} \geq 0 \; .$$

□

A.1.2 Corollary: Maximal ergodic inequality

Let $(M, \mathcal{M}, \mu)$ be a σ-finite measure space[53] and $T : L_1(M, \mathcal{M}, \mu) \to L_1(M, \mathcal{M}, \mu)$ the operator induced by the measure preserving map $\tau : M \to M$. Then the inequality

$$\mu\left(\left\{M_n^A f \geq \alpha\right\}\right) \leq \frac{\|f\|_1}{\alpha} \tag{A.1.2.1}$$

holds for any real valued $f \in L_1$ and $\alpha > 0$.

Proof: We shall outline the proof developed in [1]. Let

$$P_{n,B} := \left\{M_n^A \left(f - \alpha \cdot 1_B\right) \geq 0\right\}$$

for any finite measurable set $B \subset \left\{M_n^A f \geq \alpha\right\}$. Because T is positive (since $Th = h \circ \tau \geq 0$ for $h \geq 0$) and a contraction (to be precise, an isometry), it follows from the maximal ergodic theorem (3.2.1):

$$\int_{P_{n,B}} \overbrace{\left(f - \alpha \cdot 1_B\right)}^{\in L_1} \; d\mu \geq 0 \; . \tag{A.1.2.2}$$

But for $x \in B$ one has $A_k f(x) \geq \alpha$ for some $k \leq n$, which implies

$$S_k f(x) \geq k\alpha = S_k(\alpha) \quad \Rightarrow \quad \underbrace{S_k(f - \alpha)}_{\leq S_k(f - \alpha \cdot 1_B)}(x) \geq 0 \;\Rightarrow\; x \in P_{n,B} \; .$$

It thus follows

$$\|f\|_1 \geq \int_{P_{n,B}} |f| \; d\mu \geq \int_{P_{n,B}} f \; d\mu \overset{\substack{(A.1.2.2)\\ \mu(B)<\infty}}{\geq} \alpha \cdot \underbrace{\int_{P_{n,B}} 1_B \, d\mu}_{\substack{\mu(B)\\ \text{since}\\ \mathbb{A} \subset P_{n,B}}} = \alpha \cdot \mu\left(B\right) \; .$$

Now let

$$\bigcup_{n\in\mathbb{N}} U_n = M \quad , \quad \mu(U_n) < \infty$$

for some $\{U_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$ (possible, since $(M, \mathcal{M}, \mu)$ is σ-finite). Then

$$\left\{M_n^A f \geq \alpha\right\} = \left\{M_n^A f \geq \alpha\right\} \cap M = \bigcup_{n\in\mathbb{N}} \underbrace{\left\{M_n^A f \geq \alpha\right\} \cap \left(\bigcup_{k=1}^{n} U_k\right)}_{\substack{\text{monotonically}\\ \text{increasing}\\ \text{with } n}} \; ,$$

which implies

$$\alpha \cdot \mu\left(\left\{M_n^A f \geq \alpha\right\}\right) = \lim_{n\to\infty} \underbrace{\alpha \cdot \mu\left[\left\{M_n^A f \geq \alpha\right\} \cap \left(\bigcup_{k=1}^{n} U_k\right)\right]}_{\leq \|f\|_1} \leq \|f\|_1 \; .$$

□

[53] A measure space $(M, \mathcal{M}, \mu)$ is *σ-finite* $:\Leftrightarrow \; \exists \; U_1, U_2, .. \in \mathcal{M} : \bigcup_{n\in\mathbb{N}} U_n = M \;\wedge\; \mu(U_n) < \infty$.

A.1.3 Birkhoff-Khinchin Ergodic Theorem

Let $(M, \mathcal{M}, \mu)$ be a finite measure space, $\tau : M \to M$ measure preserving and $f \in L_1$ (real or complex). Then for almost all $x \in M$ the averages

$$A_n f(x) := \frac{1}{n} \sum_{k=0}^{n-1} f \circ \tau^k(x)$$

converge pointwise to some τ-invariant $\overline{f} \in L_1$ with $\|\overline{f}\|_1 \leq \|f\|_1$. For each τ-invariant $A \in \mathcal{M}$:

$$\int_A \overline{f}\, d\mu = \int_A f\, d\mu \ .$$

Proof: We shall follow the outline given in [1].

- We first consider a real valued $f \in L_1$. From

$$A_{n+1} f = \frac{f}{(n+1)} + \frac{n}{n+1} (A_n f) \circ \tau$$

follows

$$\limsup_{n\to\infty} A_n f(x) = \lim_{n\to\infty} \sup_{k+1\geq n} \left[\frac{f(x)}{(k+1)} + \frac{k}{k+1} (A_k f \tau)(x) \right]$$

$$= \lim_{n\to\infty} \sup_{k+1\geq n} (A_k f\tau)(x) = \limsup_{n\to\infty} (A_n f \tau)(x) \quad ,$$

in other words, $f^u := \limsup_{n\to\infty} A_n f$ is τ-invariant. Same goes for $f^l := \liminf_{n\to\infty} A_n f$.

- Now consider the τ-invariant set $\{f^u > \beta\}$ for $\beta > 0$. Then, for $x \in \{f^u > \beta\}$ one finds an $n \in \mathbb{N}$ such that $A_n f(x) \geq \beta$ and thus $M_n^A f(x) \geq \beta$. In other words:

$$\{f^u > \beta\} \subset \bigcup_{n\in\mathbb{N}} \{M_n^A f \geq \beta\} \ ,$$

which by the maximal inequality corollary (3.2.2) implies

$$\mu\left(\{f^u > \beta\}\right) \leq \mu\Bigg(\bigcup_{n\in\mathbb{N}} \underbrace{\{M_n^A f \geq \beta\}}_{\substack{\text{monotonically}\\ \text{increasing}\\ \text{in } n}} \Bigg) = \lim_{n\to\infty} \underbrace{\mu\left(\{M_n^A f \geq \beta\}\right)}_{\leq \frac{\|f\|_1}{\beta}} \leq \frac{\|f\|_1}{\beta}$$

$$\Rightarrow \ \mu\left(\{f^u = \infty\}\right) \overset{\forall\ \beta>0}{\leq} \mu\left(\{f^u > \beta\}\right) \leq \frac{\|f\|_1}{\beta} \quad \overset{\beta\to\infty}{\Longrightarrow} \quad \mu\left(\{f^u = \infty\}\right) = 0 \ .$$

With

$$\mu\left(\{f^l < \alpha\}\right) = \mu\left(\{\limsup_{n\to\infty} A_n(-f) > -\alpha\} \right) \leq \frac{\|f\|_1}{|\alpha|}$$

for $\alpha < 0$, one in a similar way obtains $f^l > -\infty$ almost everywhere as well.

- Suppose $A_n f$ does not converge almost everywhere. Then

$$0 < \mu\left(\{f^l \neq f^u\}\right) = \mu\Bigg(\bigcup_{q\in\mathbb{Q}_+} \{f^l + q < f^u\} \Bigg) = \overbrace{\sum_{q\in\mathbb{Q}_+} \mu\left(\{f^l + q < f^u\}\right)}^{\text{countable sum}}$$

$$\Rightarrow \ \exists\, \alpha < \beta \in \mathbb{Q} \ : \ \mu\big(\underbrace{\{f^l < \alpha < \beta < f^u\}}_{B} \big) > 0 \ .$$

Since either $\alpha < 0$ or $\beta > 0$, by above results $\mu(B) < \infty$. The function $f' := (f - \beta) \cdot 1_B$ has the property

$$f' \circ \tau^k\big|_{B^c} \equiv 0 \ \ \forall\, k \geq 0 \quad \wedge \quad B = \{x \mid \exists\, n \geq 1 : S_n f' > 0\}$$

(since B is τ-invariant) which implies $P_\infty f' = B$ (see 3.2.1) and thus

$$\int\limits_B (f - \beta)\, d\mu = \int\limits_{P_\infty f'} f'\, d\mu \overset{(3.2.1)}{\geq} 0 \overset{\mu(B)<\infty}{\Longrightarrow} \quad \beta\mu(B) \leq \int\limits_B f\, d\mu\ .$$

In a similar way, using $f'' := (\alpha - 1) \cdot 1_B$ one gets

$$\int\limits_B f\, d\mu \leq \alpha\mu(B)\ ,$$

which is a contradiction to $\alpha < \beta$!

- By construction is $\overline{f}$ τ-invariant a.e.. Since there exists a τ-invariant $\overline{f}'$ such that $\overline{f}' = \overline{f}$ a.e.[3], we may w.l.o.g. assume $\overline{f}$ τ-invariant everywhere.
- To be shown is the inequality $\|\overline{f}\|_1 \leq \|f\|_1$. W.l.o.g. assume $f \geq 0$. In any other case

$$\|\overline{f}\|_1 = \int\limits_M |\overline{f}^+ - \overline{f}^-|\, d\mu = \int\limits_M \left[|\overline{f}^+| + |\overline{f}^-|\right]\, d\mu = \underbrace{\|\overline{f}^+\|_1}_{\|\overline{f^+}\|} + \underbrace{\|\overline{f}^-\|_1}_{\|\overline{f^-}\|_1}$$

$$\leq \|f^+\|_1 + \|f^-\|_1 = \|f\|_1\ .$$

Indeed, according to Fatou

$$\int\limits_M \overline{f}\, d\mu = \int\limits_M \liminf_{n\to\infty} A_n f\, d\mu \overset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \int\limits_M A_n f\, d\mu$$

$$= \liminf_{n\to\infty} \frac{1}{n}\left[\int\limits_M f\, d\mu + \cdots + \int\limits_M f \circ \tau^{n-1}\, d\mu\right] = \int\limits_M f\, d\mu$$

since $\int\limits_M f \circ \tau^k\, d\mu = \int\limits_M f\, d\mu$ (see eq. A.1.3.1 below).

- We now show $\int\limits_A \overline{f}\, d\mu = \int\limits_A f\, d\mu$ for any τ-invariant $A \in \mathcal{M}$. Note that

$$\int\limits_A f\tau\, d\mu = \int\limits_{\tau(A)} f\, d\mu_\tau \overset{\substack{\mu_\tau = \mu \\ \tau(A)=A(\text{mod}0)}}{=} \int\limits_A f\, d\mu \tag{A.1.3.1}$$

since $\tau(A) \subset A$ and

$$\mu(A \setminus \tau(A)) \overset{\mu(\tau(A))<\infty}{=} \mu(A) - \mu(\tau(A)) \overset{\mu_\tau=\mu}{=} \mu(A) - \mu\Big[\underbrace{\tau^{-1}(\tau(A))}_{\substack{A \\ \text{since} \\ \tau(A)\subset A}}\Big] = 0\ .$$

Hence

$$\int\limits_A A_n f\, d\mu = \int\limits_A f\, d\mu\quad ,\quad n \in \mathbb{N}\ . \tag{A.1.3.2}$$

By the decomposition $f = f^+ - f^-$ we assume $f \geq 0$. For any $\varepsilon > 0$ we can find a $c \geq 0$ such that $g_\varepsilon = f - \max\{f, c\}$ has norm $\|g_\varepsilon\|_1 < \varepsilon$. Thus

$$\int\limits_{\{A_n f \geq c\} \cap A} A_n f \; d\mu = \int\limits_A (A_n f - c)^+ d\mu \leq \int\limits_A \underbrace{(A_n f - A_n \max\{f, c\})^+}_{(A_n g_\varepsilon)^+} d\mu \overset{g_\varepsilon \geq 0}{=} \underbrace{\int\limits_A A_n g_\varepsilon \; d\mu}_{\int\limits_A g_\varepsilon d\mu} < \varepsilon \; ,$$

in other words, $\{A_n f\}_{n \in \mathbb{N}}$ is uniformly integrable (over A). As a μ-almost everywhere (pointwise) converging sequence it therefore converges in L_1-norm as well, which together with eq. A.1.3.2 implies

$$\int\limits_A \overline{f} \; d\mu = \int\limits_A f \; d\mu \; . \tag{A.1.3.3}$$

- Now consider the complex case $f = f_r + i f_i$. Since $A_n f = A_n f_r + i A_n f_i$, the a.e. convergence of $A_n f_r$ and $A_n f_i$ implies the a.e. convergence of $A_n f$ and furthermore the validity of eq. A.1.3.3. Moreover

$$\left|\overline{f}\right| = \left|\lim_{n\to\infty} A_n f\right| = \lim_{n\to\infty} |A_n f| \leq \lim_{n\to\infty} A_n |f| = \overline{|f|} \quad \Rightarrow \quad \left\|\overline{f}\right\|_1 \leq \int\limits_M \overline{|f|} \; d\mu = \int\limits_M |f| \; d\mu = \|f\|_1 \; .$$

□

A.2 General topology & geometry

A.2.1 Lemma about continuous functions

Let M be a topological space, $\mathcal{B} \subset M$ dense in M and $f : M \to \mathbb{C}$ continuous. Then:

1. If for some $\varepsilon \geq 0$ and any $x, y \in \mathcal{B}$ we have $|f(x) - f(y)| \leq \varepsilon$, then for all $\tilde{x}, \tilde{y} \in M$ we have $|f(\tilde{x}) - f(\tilde{y})| \leq \varepsilon$.
2. If f is constant on $\mathcal{B}$, then it is constant everywhere.

Proof:

1. W.l.o.g. let f be real and suppose $|f(x_1) - f(x_2)| = \underbrace{\varepsilon}_{\geq 0} + \underbrace{\delta}_{>0}$ for some $x_1, x_2 \in M$. Then, since f is continuous, there exist open sets $\underbrace{U_1}_{\ni x_1}, \underbrace{U_2}_{\ni x_2}$ such that $f(U_i) \subset B^o_{\frac{\delta}{2}}(f(x_i))$. But as $\mathcal{B}$ is dense in M, there exist points $y_i \in U_i \cap \mathcal{B}, \; i = 1, 2$, hence

   $$\varepsilon + \delta = |f(x_1) - f(x_2)| \leq \underbrace{|f(x_1) - f(y_1)|}_{<\frac{\delta}{2}} + |f(y_1) - f(y_2)| + \underbrace{|f(y_2) - f(x_2)|}_{<\frac{\delta}{2}} \; .$$

   But this implies

   $$|f(y_1) - f(y_2)| > \varepsilon \; .$$

   The statement is proved.
2. Follows from statement 1. by setting $\varepsilon = 0$.

□

A.2.2 Lemma: Induced volume forms on leafs

Let (M, g) be a n-dimensional Riemannian manifold, with the standard volume-form

$$V = \sqrt{\det(g_{ij})} \cdot dx^1 \wedge \cdots \wedge dx^n \ .$$

Let $H : M \to \mathbb{R}$ be some $\mathcal{C}^1$ function, such that $dH \neq 0$ on M. Then g induces on every leaf $\mathscr{L}_h := \{H = h\}$ a volume form V^h, such that

$$V_{\mathbf{x}} = \frac{dH_{\mathbf{x}} \wedge V_{\mathbf{x}}^{H(\mathbf{x})}}{\sqrt{\tilde{g}(dH, dH)}} \ ,$$

whereas $\tilde{g}$ is the contravariant[54] of the metric g.

Note: Identifying volume-forms with measures admits an equivalent formulation: Each leaf $\mathscr{L}_h$ is admits the measure

$$\mu_h(\mathscr{L}_h \cap A) := \int\limits_{\mathscr{L}_h \cap A} \frac{V^h}{\sqrt{\tilde{g}(dH, dH)}}$$

such that

$$\mu(A) := \int\limits_A V = \int\limits_{\mathbb{R}} \mu_h(\mathscr{L}_h \cap A)\ dh \ .$$

Proof: W.l.o.g. we may choose new coordinates $H, y^2, .., y^n$ (with same orientation), such that $\tilde{g}(dH, dy^i) = 0$. This is always possible, since $\tilde{g}$ is symmetric and positive definite. Then $\tilde{g}$ induces on $\mathscr{L}_h$ the metric $\tilde{g}^h$ with components $\tilde{g}^h_{ij} = \tilde{g}(dy^i, dy^j)$ since dy^i are 1-forms on $T\mathscr{L}_h$. Thus in these new coordinates:

$$\frac{1}{\det(g_{ij})} = \det(\tilde{g}_{ij}) = \det\left(\begin{array}{c|c} \tilde{g}(dH, dH) & 0 \\ \hline 0 & \tilde{g}(dy^i, dy^j) \end{array}\right) = \tilde{g}(dH, dH) \cdot \det\left(\tilde{g}^h_{ij}\right) = \frac{\tilde{g}(dH, dH)}{\det\left(g^h_{ij}\right)} \ ,$$

which implies

$$dV = \sqrt{\det(g_{ij})} \cdot dH \wedge dy^2 \wedge \cdots \wedge dy^n = \frac{dH}{\sqrt{\tilde{g}(dH, dH)}} \wedge \underbrace{\sqrt{\det\left(g^h_{ij}\right)} dy^2 \wedge \cdots \wedge dy^n}_{V^h} \ .$$

□

A.2.3 Lemma about reflections in ellipses

Let $\mathcal{E}_c := \left\{q \in \mathbb{R}^2 : d(q, F_1) + d(q, F_2) \leq c\right\}$ be an ellipse and $\mathcal{H}_c := \left\{q \in \mathbb{R}^2 : d(q, F_1) - d(q, F_2) \leq c\right\}$ a hyperbola in $\mathbb{R}^2$ with foci F_1, F_2. Then:

1. For any point $q \in \partial\mathcal{E}_c$ the line segments $\overline{F_1 q}$ and $\overline{F_2 q}$ form equal angles with the tangent $T_q := T_q\partial\mathcal{E}_c$ to the ellipse at the point q.
2. The same holds for any point $q \in \partial\mathcal{H}_c$.

[54] We define: $\tilde{g}(a, b) := g(J^{-1}a, J^{-1}b)$, with the isomorphism J mapping vectors to 1-forms: $Jx := g(x, \cdot)$. Note that

$$(\tilde{g}(dx^i, dx^j)) = (g(\partial_{x^i}, \partial_{x^j}))^{-1} \ .$$

See also [11].

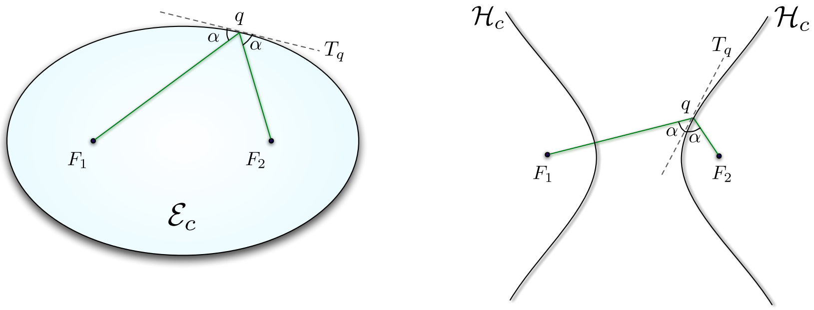

Figure 31: On angles between tangents and *rays* originating in foci of ellipses and hyperbolas.

Proof: We reproduce the proof found in [3].

1. Consider the point F_2' symmetric to F_2 with respect to T_q. Then $d(F_1, q) + d(q, F_2') = d(F_1, q) + d(q, F_2) = c$ while for all other $p \in T_q$ we have $d(F_1, p) + d(p, F_2') > c$. Thus q is the intersection of the straight line $\overline{F_1 F_2'}$ which implies what was to be shown.

2. Reflecting the focus F_2 at T_q allows for similar arguments as in 1.

□

A.3 Measure theory

A.3.1 Corollary to the Poincaré recurrence theorem

Let $(M, \mathcal{M}, \mu)$ be a finite (non-trivial) measure space, $\tau : M \to M$ measure preserving and $f : M \to \mathbb{R}^*$ such that $f > 0$ a.e. Then for almost all $x \in M$ we have

$$\sum_{n=1}^{\infty} f(\tau^n x) = \infty \ .$$

Proof: Consider the sequence of sets $A_n := \left\{ f > \frac{1}{n} \right\}$. Since almost every point $x \in A_n$ returns to A_n infinitely often, and

$$\bigcup_{n \in \mathbb{N}} A_n = \{f > 0\} = M(\mathrm{mod}0) \ ,$$

we have that a.e. point $x \in M$ visits some A_{n_x} infinitely often. This proves what was to be shown.
□

A.3.2 Lemma: Characterization of measure preserving maps

Let $(M, \mathcal{M}, \mu)$ be a measure space and $\tau : M \to M$. Then the following statements are equivalent:

1. τ is measure preserving.

2. For any measurable function $f : M \to \mathbb{C}^*$:

$$\int\limits_M f \circ \tau \ d\mu = \int\limits_M f \ d\mu \ .$$

As a special case:

$$\|f \circ \tau\|_p = \|f\| \quad \forall \ f \in L_p, \ 0 < p < \infty \ .$$

3. In case of μ finite: For any bounded, L_p-integrable ($\forall\ 0 < p < \infty$) function $f : M \to \mathbb{R}$:

$$\int_M f \circ \tau \ d\mu = \int_M f \ d\mu \ .$$

4. If M is metrizable and compact, $\mathcal{M}$ the Borel-σ-algebra of M and τ continuous:

$$\int_M f \circ \tau \ d\mu = \int_M f \ d\mu$$

   for all $f \in \mathcal{C}(M)$.

Proof:

$1 \to 2$: Since $\tau^{-1}(M) = M$:

$$\int_M f \circ \tau \ d\mu = \int_{\tau^{-1}(M)} f \circ \tau \ d\mu = \int_M f \ d\mu_\tau \overset{\mu=\mu_\tau}{=} \int_M f \ d\mu \ .$$

$2 \to 1$: For $A \in \mathcal{M}$ set $f := 1_A$:

$$\mu\left(\tau^{-1}(A)\right) = \int_M 1_{\tau^{-1}(A)} \ d\mu = \int_M 1_A \circ \tau \ d\mu = \int_M 1_A \ d\mu = \mu(A) \ .$$

$2 \to 3$: Trivial.

$3 \to 1$: Same as $2 \to 1$ since 1_A is bounded and L_p integrable $\forall\ 0 < p < \infty$.

$1 \to 4$: Trivial.

$4 \to 1$: Let $\mu(f) = \mu_\tau(f)$ for $f \in \mathcal{C}(M)$. Then by the Frigyes Riesz representation theorem $\mu = \mu_\tau$.

□

A.3.3 Lemma about invariant sets

Let $(M, \mathcal{M}, \mu)$ be a finite measure space, $\tau : M \to M$ measure preserving and $A \in \mathcal{M}$ so that

$$\tau^{-1}(A) \subset A \ .$$

Then there exists a τ-invariant set $\tilde{A} \in \mathcal{M}$ so that $A = \tilde{A}(\text{mod}0)$.[3]

Proof: Set

$$\tilde{A} := \bigcap_{n \in \mathbb{N}_0} (\tau^n)^{-1}(A) \ .$$

Then by construction $\tau(\tilde{A}) \subset \tilde{A}$. Furthermore, for $x \in \tilde{A}^c$, that is $\tau^n x \notin A$ for some $n \in \mathbb{N}_0$: $\tau^{n-1}(\tau x) \notin A$ in case $n \geq 1$ and $\tau x \notin A$ in case $n = 0$, which implies $\tau x \in A^c$. Thus $\tilde{A}$ is τ-invariant, and since $\tilde{A} \subset A$:

$$\mu\left(A \triangle \tilde{A}\right) = \mu\left[A \setminus \bigcap_{n \in \mathbb{N}_0} (\tau^n)^{-1}(A)\right] = \mu\left[\bigcup_{n \in \mathbb{N}_0} A \setminus (\tau^n)^{-1}(A)\right] \leq \sum_{n=0}^{\infty} \underbrace{\left[\mu(A) - \mu\left[(\tau^n)^{-1}(A)\right]\right]}_{\substack{0 \\ \text{since } \tau \text{ measure preserving}}} = 0 \ .$$

□

A.3.4 Lemma: Topological aspects of measure families

Let $(M, \mathcal{M})$ be a measurable space, $(\tau^g)_{g\in G}$ a G-semi-flow, $\mathfrak{M}$ the set of all probability measures and $\mathfrak{M}((\tau^g))$ the set of all (τ^g)-invariant probability measures in $(M, \mathcal{M})$. Then:

1. The sets $\mathfrak{M}$ and $\mathfrak{M}((\tau^g))$ are convex.
2. Every extreme point of $\mu \in \mathfrak{M}((\tau^g))$ is ergodic.
3. If (τ^g) is a flow, then every ergodic probability measure is extreme in $\mathfrak{M}((\tau^g))$.
4. If M is compact, metrizable and $\mathcal{M} = \mathscr{B}(M)$, then the set $\mathfrak{M}$ is sequentially compact with respect to *weak* convergence* of measures.
5. If M is compact, metrizable, $\mathcal{M} = \mathscr{B}(M)$ and every $\tau^g : M \to M$ continuous, then the set $\mathfrak{M}((\tau^g))$ is sequentially compact with respect to *weak* convergence* of measures.

Note: The weak* convergence topology on $\mathfrak{M}$ can indeed be attributed to a metric $d(\cdot,\cdot)$:

$$d(\mu,\nu) := \sum_{j=1}^{\infty} \frac{1}{2^j} \left| \int_M f_j \, d\mu - \int_M f_j \, d\nu \right|$$

for some dense, countable family of functions $\{f_i\}_{i\in\mathbb{N}} \subset \mathcal{C}(M)$ in the normed space $(\mathcal{C}(M), \|\cdot\|_\infty)$ [21]. Thus sequential compactness is equivalent to topological compactness. See also [17].

Proof: We shall here generalize the proof found in [17].

1. Let $\mu_0, \mu_1 \in \mathfrak{M}$ and $t \in [0,1]$. Then clearly $\mu_t := t\cdot\mu_1 + (1-t)\cdot\mu_0$ is a probability measure. If furthermore μ_0, μ_1 are (τ^g)-invariant, then so is μ_t.
2. Suppose $\mu \in \mathfrak{M}((\tau^g))$ is not ergodic. Choose some (τ^g)-invariant $A \in \mathcal{M}$ such that $0 < \mu(A) < 1$. Then the (τ^g)-invariant probability measures

$$\mu_A(B) := \frac{\mu(A\cap B)}{\mu(A)} \quad , \quad \mu_{A^c}(B) := \frac{\mu(A^c \cap B)}{\mu(A^c)}$$

satisfy

$$\mu = \underbrace{\mu(A)}_{\in(0,1)} \cdot \mu_A + (1-\mu(A))\cdot\mu_{A^c} \ ,$$

hence μ is not extremal in $\mathfrak{M}((\tau^g))$.

Assume now $\mu = t\cdot\mu_1 + (1-t)\cdot\mu_2$ is ergodic with $t\in[0,1]$, $\mu_i \in \mathfrak{M}((\tau^g))$. Now suppose $t \in (0,1)$, then for any μ-nullset we have $\mu_1(A) = \mu_2(A) = 0$. Thus by theorem 3.3.4, $\mu = \mu_1$ or $\mu = \mu_2$.

3.
   - Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathfrak{M}$ be a sequence of probability measures and $(f_i)_{i\in\mathbb{N}} \subset \mathcal{C}(M)$ a countable family of dense functions in $(\mathcal{C}(M), \|\cdot\|_\infty)$ (recall that the normed space $(\mathcal{C}(M), \|\cdot\|_\infty)$ is separable). Then every sequence $(\mu_n(f_i))_n$ of complex numbers is bounded, since $|\mu_n(f_i)| \le \|f_i\|_\infty$. By a diagonal argument[55] and the Bolzano-Weierstraß theorem for $(\mathbb{C}, |\cdot|)$, there exists some sub-sequence $(\mu_{n_k})_k$ such that $\mu_{n_k}(f_i)$ is convergent (in k) for every f_i.
   - Actually, $(\mu_{n_k})_k(f)$ converges for all $f \in \mathcal{C}(M)$, since: For any $\varepsilon > 0$ we can chose some f_i so that $\|f - f_i\|_\infty < \frac{\varepsilon}{3}$ and $N \in \mathbb{N}$ such that

$$|\mu_{n_k}(f_i) - \mu_{n_r}(f_i)| < \frac{\varepsilon}{3} \quad \forall\, k, r \ge N \ .$$

   Hence:

$$|\mu_{n_k}(f) - \mu_{n_r}(f)| \le \underbrace{|\mu_{n_k}(f) - \mu_{n_k}(f_i)|}_{\le \|f-f_i\|_\infty < \frac{\varepsilon}{3}} + \underbrace{|\mu_{n_k}(f_i) - \mu_{n_r}(f_i)|}_{<\frac{\varepsilon}{3}} + \underbrace{|\mu_{n_r}(f_i) - \mu_{n_r}(f)|}_{\le \|f_i - f\|_\infty < \frac{\varepsilon}{3}} < \varepsilon \quad \forall\, k,r \ge N \ ,$$

   that is, μ_{n_k} is Cauchy in $(\mathbb{C}, |\cdot|)$ and thus convergent.

---

[55] Construct a family of sub-sequences $\left\{(\mu_{n_k^m})_k\right\}_m$, $(\mu_{n_k^{m+1}})_k \subset (\mu_{n_k^m})_k$, such that $(\mu_{n_k^m})_k(f_i)$ is convergent (in k) for $1 \le i \le m$. Then the *diagonal* subsequence $(\mu_{n_k^k})_k(f_i)$ converges for all i.

- We now set $\mu(f) := \lim_{k\to\infty} \mu_{n_k}(f)$ for $f \in \mathcal{C}(M)$. Obviously the functional μ is linear, positive and normed $l(1) = 1$. By the Riesz representation theorem it corresponds to a probability Borel measure μ, such that

$$\int_M f\ d\mu \overset{\forall\ f\in\mathcal{C}}{=} \mu(f) = \lim_{n\to\infty} \int_M f\ d\mu_{n_k}\ ,$$

that is, $\mu_{n_k} \underset{\text{weak*}}{\overset{n\to\infty}{\longrightarrow}} \mu$.

4. Since $\mathfrak{M}((\tau^g)) \subset \mathfrak{M}$ and $\mathfrak{M}$ is sequentially compact, it suffices to show that $\mathfrak{M}((\tau^g))$ is sequentially closed with respect to weak* convergence. Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of measures in $\mathfrak{M}((\tau^g))$ such that $\mu_n \underset{\text{weak*}}{\overset{n\to\infty}{\longrightarrow}} \mu$ for some measure μ. Since μ_n are probability measures, so is μ, since $\int_M 1\ d\mu_n = 1\ \forall\ n$. Now let $f : M \to \mathbb{C}$ be continuous, then

$$\mu_n(\underbrace{f\circ\tau^g}_{\in\mathcal{C}(M)}) = \int_M f\circ\tau^g\ d\mu_n \overset{n\to\infty}{\longrightarrow} \int_M f\circ\tau^g\ d\mu = \mu(f\circ\tau^g) = \mu_{\tau^g}(f)$$

$$= \mu_n(f) \overset{n\to\infty}{\longrightarrow} \mu(f)\ .$$

Thus, $\mu(f) = \mu_{\tau^g}(f)$ for $f \in \mathcal{C}(M)$ by uniqueness of the limes. By the Frigyes Riesz representation theorem about metrizable, compact spaces, we have $\mu = \mu_{\tau^g}$.

□

A.3.5 Lemma about mixing systems

Let $(M, \mathcal{M}, \mu)$ be a probability space, $(\tau^g)_{g\in G}$ an ordered, measure preserving G-semi-flow and $\Phi \subset L_2$ some complete[56] system of functions in L_2. If for any $f_1, f_2 \in \Phi$ the relation

$$\lim_{g\to\infty} \int_M f_1^* \cdot (f_2 \circ \tau^g)\ d\mu = \int_M f_1^*\ d\mu \cdot \int_M f_2\ d\mu \tag{A.3.5.1}$$

holds, then (τ^g) is mixing, that is, eq. A.3.5.1 holds for all $f_1, f_2 \in L_2$.

Proof: We shall generalize the proof found in [17], chapter 4. Since both sides of the above equation are linear in f_2 and anti-linear in f_1, it holds for some dense subset $\Psi := \text{span}(\Phi)$ of L_2. Now let $f_1, f_2 \in L_2$ and $\varepsilon > 0$

[56] A subset $\Phi \subset \mathcal{H}$ of a Hilbert-space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ is called *complete*, if the set span(Φ) of finite, linear combinations of vectors in Φ is dense in $\mathcal{H}$.

and choose $\tilde{f}_i \in \Psi$ such that $\|\tilde{f}_i - f_i\|_2 < \varepsilon$. Then by the Schwarz inequality

$$\left| \int_M f_1^* \cdot (f_2 \tau^g)\, d\mu - \int_M f_1^* d\mu \cdot \int_M f_2\, d\mu \right|$$

$$= \left| \int_M (f_1 - \tilde{f}_1)^* \cdot (f_2 \tau^g) d\mu + \int_M \tilde{f}_1^* \cdot (f_2 \tau^g - \tilde{f}_2 \tau^g) d\mu + \int_M \tilde{f}_1^* \cdot (\tilde{f}_2 \tau^g) d\mu - \int_M \tilde{f}_1^* d\mu \cdot \int_M \tilde{f}_2\, d\mu \right.$$

$$\left. + \int_M (\tilde{f}_1 - f_1)^* d\mu \cdot \int_M \tilde{f}_2 d\mu + \int_M f_1^* d\mu \cdot \int_M (\tilde{f}_2 - f_2) d\mu \right|$$

$$\overset{\text{Schwarz}}{\leq} \|f_1 - \tilde{f}_1\|_2 \cdot \underbrace{\|f_2 \circ \tau^g\|_2}_{\substack{\|f_2\|_2 \\ \text{by} \\ (A.3.2)}} + \underbrace{\|\tilde{f}_1\|_2}_{\substack{\leq \|\tilde{f}_1 - f_1\|_2 \\ + \|f_1\|_2 \\ \text{Minkowski}}} \cdot \underbrace{\|(f_2 - \tilde{f}_2) \circ \tau^g\|_2}_{\substack{\|f_2 - \tilde{f}_2\|_2 \\ \text{by} \\ (A.3.2)}} + \left| \int_M \tilde{f}_1^* \cdot (\tilde{f}_2 \tau^g)\, d\mu - \int_M \tilde{f}_1^* d\mu \cdot \int_M \tilde{f}_2\, d\mu \right|$$

$$+ \underbrace{\|f_1 - \tilde{f}_1\|_1}_{\substack{\leq \|f_1 - \tilde{f}_1\|_2 \\ \text{Hölder} \\ \wedge \mu(M)=1}} \cdot \underbrace{\left| \int_M \tilde{f}_2\, d\mu \right|}_{\substack{\leq \|\tilde{f}_2\|_1 \\ \leq \|\tilde{f}_2 - f_2\|_1 + \|f_2\|_1}} + \underbrace{\|f_2 - \tilde{f}_2\|_1}_{\substack{\leq \|f_2 - \tilde{f}_2\|_2 \\ \text{Hölder} \\ \wedge \mu(M)=1}} \cdot \underbrace{\left| \int_M f_1^*\, d\mu \right|}_{\leq \|f_1\|_1}$$

$$< \varepsilon \cdot \underbrace{\left[\|f_2\|_2 + \varepsilon + \|f_1\|_2 + \varepsilon + \|f_2\|_1 + \|f_1\|_1\right]}_{\leq 2\left[\varepsilon + \|f_1\|_2 + \|f_2\|_2\right] = \text{const} < \infty} + \underbrace{\left| \int_M \tilde{f}_1^* \cdot (\tilde{f}_2 \tau^g)\, d\mu - \int_M \tilde{f}_1^* d\mu \cdot \int_M \tilde{f}_2\, d\mu \right|}_{\overset{g \to \infty}{\longrightarrow} 0} \overset{g \to \infty}{\longrightarrow} \varepsilon \cdot \text{const} \ .$$

Since $\varepsilon > 0$ was arbitrary small, eq. A.3.5.1 follows for f_1, f_2.
□

A.3.6 Choquet theorem

Let C be a metrizable, convex, compact set in a locally convex topological vector space and $x \in C$. Then there exists a probability measure μ supported[57] on the set of extrema[58] $\text{ex}(C)$, such that

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

for any affine function f on C [17].

[57] The support $\operatorname{supp} \mu$ of a measure μ on a topological space $(M, \mathscr{B}(M))$ is defined as:

$$\text{supp}(\mu) := \{x \in M \mid x \in U \in \mathcal{O}(M) \ \Rightarrow \ \mu(U) > 0\} \ .$$

We say μ is supported on some set $A \in \mathscr{B}(M)$, if $\operatorname{supp} \mu \subset A$. Note that $\operatorname{supp} \mu$ is closed, and for every $A \subset (\operatorname{supp} \mu)^c$ we have $\mu(A) = 0$.

[58] A point $x \in C$ of a convex set C is called *extreme*, if $x = t \cdot x_1 + (1-t) \cdot x_2$ with $x_1, x_2 \in C,\ t \in [0,1]$ implies $x = x_1$ or $x = x_2$. We write $\text{ex}(C)$ for the set of extreme points of C.

References

[1] *Ergodic Theorems*, U. Krengel, A. Brunel
Walter de Gruyter, Berlin 1985

[2] *Dynamical Systems II*, Y. G. Sinai
Springer, Berlin 1989

[3] *Ergodic theory*, I.P. Cornfeld, S.V. Fomin, Y.G. Sinai, A.B. Sossinski
Springer, Berlin 1982

[4] *Lecture notes on ergodicity*, T. Tao
Wordpress, California 2008
http://terrytao.wordpress.com/2008/02/04/254a-lecture-9-ergodicity/ (March 10, 2009)

[5] *On weak convergence of Measures on Hilbert Spaces*, M.J. Merkle
Elsevier, Journal of Multivariate Analysis, Vol. 29, No 2, 252-259, Amsterdam 1989
http://www.milanmerkle.com/documents/radovi/Jmva989.pdf (April 1, 2009)

[6] *Maß- und Integrationstheorie*, J. Elstrodt
Springer, Berlin 2004

[7] *Mathematische Methoden der klassischen Mechanik*, V.I. Arnold
VEB, Berlin 1988

[8] *A modern course in statistical physics*, L.E. Reichl
Wiley, New York 1998

[9] *Nonlinear stochastic evolution problems in applied sciences*, N. Bellomo, Z. Brzezniak, L.M. Socio
Kluwer, Dordrecht 1992

[10] *Recursive methods in economic dynamics*, N.L. Stokey, R.E. Lucas, E.C. Prescott
Harvard Univ. Press, Cambridge Mass. 1993

[11] *Geometrie der Raumzeit*, R. Oloff
Vieweg, Braunschweig 1999

[12] *On invariant curves of area-reserving mappings of an annulus*, J. Moser
Nachr. Akad. Wiss., Göttingen 1962

[13] *A lecture on the classical KAM-theorem*, J. Pöschel
AMS, Proceedings of Symposia in Pure Mathematics, Vol. 69, 707-732, Seattle 2001
http://www.poschel.de/pbl/kam-1.pdf (April 10, 2009)

[14] *An introduction to KAM-theory*, C.E. Wayne
Boston University, Boston 2008
http://math.bu.edu/people/cew/preprints/introkam.pdf (April 10, 2009)

[15] *Smooth ergodic theory and its applications*, A. Katok, R. de la Llave, Y. Pesin, H. Weiss
Proceedings of the AMS Summer Research institute of smooth ergodic theory and its applications, Vol 69, 107-174, Seattle 1999

[16] *Mathematical aspects of classical and celestial mechanics*, V.I. Arnold, V.V. Kozlov, A.I. Neishtadt
Springer, Berlin 2006

[17] *Introduction to modern theory of dynamical systems*, A. Katok, B. Hasselblatt, L. Mendoza
Cambridge Univ. Press, Cambridge 1997

[18] *Construction of analytic KAM surfaces and effective stability bounds*, A. Celletti, L. Chierchia
Springer, Communications in Math. Physics, Vol 118, 119-161, New York 1988
http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.cmp/1104161911 (March 13, 2009)

[19] *Some additional topics of the theory of ordinary differential equations.*, V.I. Arnold
Nauka, Moscow 1978

[20] *Dynamical Systems III*, V.I. Arnold
Springer, Berlin 1993

[21] *Introduction to the ergodic theory of chaotic billiards*, N. Chernov, R. Markarian
Pub. Mat. Rio de Janeiro: IMPA, Rio de Janeiro 2001
http://www.math.uab.edu/c̃hernov/papers/rbook.pdf (April 2, 2009)

[22] *Integral Geometry and geometrical probability*, L.A. Santalo
Cambridge University Press, Cambridge 2004

[23] *Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo*, L. Boltzmann
Annalen der Physik, Vol. 293, Issue 4, 773-784, Weinheim 1896

[24] *Chaotic billiards*, N. Chernov, R. Markarian
AMS Bookstore, Providence 2006

[25] *Hard ball systems and the Lorentz gas*, D. Szász, L.A. Bunimovich
Springer, Berlin 2000

[26] *Dynamical systems and ergodic theory*, M. Pollicott, M. Yuri
Cambridge Univ. Press, Cambridge 1998

[27] *Poincaré recurrence: old and new*, L. Barreira
14th International Congress on Mathematical Physics, Lisbon 2003

[28] *Time-reversal symmetry, Poincaré recurrence, irreversibility, and the entropic arrow of time: From mechanics to system thermodynamics*, W.M. Haddad, V.S. Chellaboina, S.G. Nersesov
Elsevier, Nonlinear Analysis: Real World Applications 9, 250-271, Amsterdam 2008
http://www.ae.gatech.edu/people/whaddad/journal/Entropic_Arrow.pdf (April 17, 2009)

[29] *Chaotic dynamics*, T. Tel, M. Gruiz, K. Kulacsy, S. Hadobás
Cambridge Univ. Press, Cambridge 2006

[30] *Ergodic hypothesis in classical statistical mechanics*, C.R. de Oliveira, T. Werlang
Revista Brasileira de Ensino de Fisica, Vol. 29, No 2, 189-201, São Paolo 2007
http://www.sbfisica.org.br/rbef/pdf/060601.pdf (April 17, 2009)

[31] *Proof of the ergodic theorem*, G.D. Birkhoff
Proceedings of the National Academy of Sciences USA, Vol. 17, 656-660, 1931
http://www.pnas.org/content/17/12/656.full.pdf (April 18, 2009)

Statement of authorship

I hereby certify that this paper has been composed by myself, and describes my own work, unless otherwise acknowledged in the text. All references and verbatim extracts have been quoted, and all sources of information have been specifically acknowledged. It has not been accepted in any previous application for a degree.