

Sistemi dinamici e misure invarianti

Andrea Giacobbe

A.A. 2008-2009

Indice

1	Dinamica e misure invarianti, definizioni ed esempi	2
1.1	Generalità su misure	2
1.2	Sistemi dinamici misurabili	6
1.3	Primi esempi di sistema dinamico misurabile	6
1.4	Secondo esempio: misura invariante ed insiemi di livello	7
1.4.1	Applicazione: l'oscillatore armonico	8
1.4.2	Applicazione: sistemi 1D conservativi	9
1.4.3	Caso con più integrali primi	10
1.5	Terzo esempio: la mappa di Poincarè	10
1.6	Quarto esempio: I biliardi	12
1.6.1	Biliardi a bordo circolare	13
1.6.2	Biliardi a bordo qualsiasi	13
1.7	Schemi di Bernoulli	15
2	Sistemi Ergodici	16
2.1	Medie temporali e teorema di Birkhoff-Kinchin	16
2.2	Definizione di sistema ergodico	18
2.3	Esempi di sistemi Ergodici e non	19
2.4	Sistemi iperbolici ed ergodicità	20
2.5	Misure ergodiche	21
3	Sistemi Mescolanti	22
4	Entropia	25
4.1	L'entropia delle partizioni	25
4.2	L'entropia di Kolmogorov-Sinai	27
4.2.1	L'entropia dei sistemi continui	29
4.2.2	Prime stime di Entropia	29
4.3	Partizioni generanti e calcolo dell'entropia	30
4.4	Entropia come invariante dinamico	31
4.5	gli schemi di Bernoulli e l'entropia	31
4.6	Entropia e sistemi classici	31
4.7	Entropia ed esponenti di Liapunov	33
4.8	Entropia condizionata e distanza sullo spazio delle partizioni misurabili	33
4.8.1	Dimostrazione del teorema sulle partizioni generanti	36

Introduzione

L'ossatura di queste note è quella delle note di Teoria Ergodica di G. Benettin [B], salvo alcune parti che sono state emendate. Nella prima metà le differenze consistono in una più accurata descrizione della teoria della misura soggiacente, con particolare attenzione a quelle regolari rispetto alla misura di Lebesgue, nella seconda metà è parzialmente diversa la descrizione della regolarità statistica, la quale è stata scritta prendendo a spunto la trattazione di G. Gallavotti [G].

In queste note si è cercato di scrivere in modo preciso le definizioni, e di fornire dimostrazioni il più possibile rigorose, per adeguarsi alla forma mentis degli studenti del Corso di Laurea in Matematica.

Per riassumere in una riga: si discute di come estrarre informazioni da un sistema dinamico che possiede anche una misura conservata dalla dinamica.

1 Dinamica e misure invarianti, definizioni ed esempi

Prima di cominciare lo studio dei sistemi dinamici con misura, richiamiamo alcuni fondamenti di teoria della misura, presentiamo alcuni modi pratici per descrivere e studiare una classe di misure che ci interessa, ed investighiamo svariati esempi di sistemi dinamici con misure invarianti, utili per guidare l'intuizione.

1.1 Generalità su misure

Sia M uno spazio, che chiameremo **spazio delle fasi**. Una σ -algebra è un sottoinsieme $\mathcal{M}$ dell'insieme delle parti di M chiusa a unioni numerabili ed a complementari, gli elementi di $\mathcal{M}$ si dicono **insiemi misurabili**. Una **misura** μ è una funzione da $\mathcal{M}$ in $[0, +\infty]$ tale che

$$(m1) \quad \mu(\emptyset) = 0 \text{ e}$$

$$(m2) \quad \text{numerabilmente addittiva, ovvero tale che } \mu(\sqcup_{n \in \mathbb{N}} A_n) = \sum_n \mu(A_n). \text{ (Il simbolo usato indica l'unione disgiunta.)}$$

Conseguenza immediata di queste due condizioni è che una misura è monotona, ovvero che se $A \subset B$ allora $\mu(A) \leq \mu(B)$ ed è sub-addittiva, ovvero $\mu(\cup_{n \in \mathbb{N}} A_n) \leq \sum_n \mu(A_n)$. Vi sono esempi tipici di σ -algebre e misure.

Esempio 1 *Lo spazio $\mathbb{R}^n$ con la misura di Lebesgue.*

Svolgimento. Lo spazio delle fasi è lo spazio vettoriale $\mathbb{R}^n$ dotato di un prodotto scalare (una forma bilineare simmetrica e definita positiva). La σ -algebra è quella generata dagli aperti di $\mathbb{R}^n$ e la misura si definisce come segue.

Sia $\mathcal{P}$ il parallelepipedo con vertici in $P_I = P + \sum_{i \in I} v_i$, dove $v_1, \dots, v_n$ sono un insieme di vettori linearmente indipendenti ed $I \subset \{1, \dots, n\}$. A tale n -upla di vettori si associa una matrice $n \times n$ detta **matrice Gramiana**, definita come $Gram(v_1, \dots, v_n) = (v_i \cdot v_j)_{i,j}$. Rimane così definito anche un numero reale positivo detto determinante **gramiano** che è la radice quadrata del determinante della matrice Gramiana, è facile osservare che il determinante della matrice Gramiana è necessariamente positivo, in quanto si ottiene come $Gram(v_1, \dots, v_n) = V^t V$ dove V è la matrice ottenuta mettendo in colonna le componenti dei vettori $v_1, \dots, v_n$. Il volume di $\mathcal{P}$ per definizione è

$$\lambda(\mathcal{P}) = \text{gram}(v_1, \dots, v_n) = \sqrt{\det(Gram(v_1, \dots, v_n))}.$$

Rimane quindi definita una misura λ sulla σ -algebra generata dai parallelepipedi. Questa definizione di misura permette di definire una integrazione che formalmente associa ad una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ed a un insieme misurabile A il numero reale $\int_A f d\lambda$ e che, praticamente, comporta un passaggio al limite. In un corso di Analisi tale integrale verrebbe scritto dopo aver scelto coordinate naturali $x^1, \dots, x^n$ come $\int_A f dx^1 \cdots dx^n$, in geometria differenziale come $\int_A f dx^1 \wedge \cdots \wedge dx^n$ (entreremo nei dettagli di queste scritture nelle prossime pagine). $\square$

Esempio 2 *Lo spazio S^1 con la misura di Lebesgue.*

Svolgimento. Questo secondo esempio è la più facile estensione della misura di Lebesgue ad uno spazio compatto. Si considera S^1 come $\mathbb{R}/2\pi\mathbb{Z}$. Questo spazio è dotato di una sommersione naturale $\rho : \mathbb{R} \rightarrow S^1$, $x \mapsto [x] = \vartheta$. Si definiscono come misurabili gli insiemi A tali che $\rho^{-1}(A)$ è misurabile. La misura di A è la misura

di $\rho^{-1}(A) \cap [0, 2\pi]$. Anche in questo caso, l'integrazione associa ad una funzione $f : S^1 \rightarrow \mathbb{R}$, che è una funzione periodica $\tilde{f}$ di $\mathbb{R}$ in $\mathbb{R}$, e ad un insieme misurabile A l'integrale $\int_A f d\lambda$, che è $\int_A \tilde{f} \chi_{[0, 2\pi]} d\lambda$ e che si scriverà a volte come $\int_A f(\vartheta) d\vartheta$. $\square$

Definizione 3 Una k -forma differenziale τ su di una varietà m -dimensionale M associa ad ogni punto $q \in M$ una forma k -multilineare alternante τ_q sullo spazio $T_q M$.

Una forma k -multilineare alternante si valuta su k vettori e porge un numero reale. In uno spazio vettoriale $\mathbb{R}^n$, una forma k -multilineare alternante è combinazione lineare delle $dx^{i_1} \wedge \dots \wedge dx^{i_k}$, che a k vettori $v_1, \dots, v_k$ associa il determinante della matrice $k \times k$

$$\begin{pmatrix} v_1^{i_1} & \dots & v_1^{i_k} \\ \vdots & & \vdots \\ v_k^{i_1} & \dots & v_k^{i_k} \end{pmatrix}.$$

Segue immediatamente che le k -forme differenziali sono tutte triviali quando k eccede m , la dimensione dello spazio.

Ogni k -forma differenziale τ , una volta scelte delle coordinate locali $u_1, \dots, u_m$ per M , ha espressione del tipo $\tau_u = \sum_{I=(i_1 < \dots < i_k)} f_I(u) du^{i_1} \wedge \dots \wedge du^{i_k}$. Un cambio di coordinate $u = u(v)$ determina una riscrittura della forma secondo la regola $\tau_v = \sum_{I=(i_1 < \dots < i_k)} f_I(u(v)) \left(\sum_j \partial_{v_j} u^{i_1} dv_j \right) \wedge \dots \wedge \left(\sum_j \partial_{v_j} u^{i_k} dv_j \right)$. Su quest'ultima si possono usare le regole dell'algebra multilineare ed ottenere

$$\tau_v = \sum_{I=(i_1 < \dots < i_k), J=(j_1 < \dots < j_k)} f_I(u(v)) \text{jac}(u^I, v^J) dv^{j_1} \wedge \dots \wedge dv^{j_k}.$$

dove $\text{jac}(u^I, v^J)$ è il determinante della matrice Jacobiana $(\partial_{v_j} u^{i_i})$. Due dati locali $\sum_I f_I(u) du^{i_1} \wedge \dots \wedge du^{i_k}$ in coordinate u e $\sum_J g_J(v) dv^{j_1} \wedge \dots \wedge dv^{j_k}$ in coordinate v si dicono **compatibili per cambio di coordinate** se vale $g_J(v) = \sum_I f_I(u(v)) \text{jac}(u^I, v^J)$.

Definizione 4 Una k -forma differenziale è il dato di una famiglia di espressioni locali

$$\tau_u = \sum_{I=(i_1 < \dots < i_k)} f_I(u) du^{i_1} \wedge \dots \wedge du^{i_k}$$

compatibili per cambi di coordinate.

Dato un diffeomorfismo $\varphi : N \rightarrow M$ ed una forma τ in M , si definisce il **pull-back** $\varphi^* \tau$. Scelte coordinate locali v per N ed u per M , il pull-back di τ ha espressione in coordinate v

$$\varphi^*(\tau) = \sum_{I=(i_1 < \dots < i_k)} f_I(u(v)) \left(\sum_j \partial_{v_j} u^{i_1} dv_j \right) \wedge \dots \wedge \left(\sum_j \partial_{v_j} u^{i_k} dv_j \right).$$

Esempio 5 Calcolare il pull-back della forma differenziale $dx^1 \wedge dx^2$ sotto il cambiamento di coordinate lineare $x^1 = \alpha y^1 + \beta y^2$, $x^2 = \gamma y^1 + \delta y^2$.

Svolgimento. Basta scrivere $dx^1 \wedge dx^2 = d(\alpha y^1 + \beta y^2) \wedge d(\gamma y^1 + \delta y^2)$ e poi distribuire $= d\alpha y^1 \wedge \gamma y^1 + d\alpha y^1 \wedge \delta y^2 + d\beta y^2 \wedge \gamma y^1 + d\beta y^2 \wedge \delta y^2$. Si usa poi linearità ed antisimmetria e si ottiene $= (\alpha\delta - \beta\gamma) dy^1 \wedge dy^2$. $\square$

Una forma volume in una varietà M di dimensione m è una forma differenziale di **grado massimo** (ovvero di grado m) e **mai nulla**. Una forma differenziale di grado massimo associa ad ogni punto p di M una densità di volume infinitesimo ed un segno: scelti $v_1, \dots, v_m$ vettori tangenti ad un punto p il volume associato al parallelepipedo infinitesimo è $|\tau(v_1, \dots, v_m)|$, il segno determina una **orientazione**, ed è $\text{sgn}(\tau(v_1, \dots, v_m))$.

Esempio 6 La misura in una varietà M dotata di forma volume τ .

Svolgimento. Supponiamo ora data una forma volume τ in una varietà m -dimensionale M . Scelto un punto q di M , esiste sempre un intorno U di q detto **aperto di carta** ed m -funzioni $(x^1, \dots, x^m) : U \rightarrow \mathbb{R}^m$ dette **sistema di coordinate**. In ogni sistema di coordinate, la forma differenziale τ ha una scrittura locale del tipo $h(x^1, \dots, x^m) dx^1 \wedge \dots \wedge dx^m$ con h funzione mai nulla. La forma differenziale τ permette di integrare le funzioni

nel seguente modo: si sceglie un quasi-atlante, ovvero un ricoprimento $\mathcal{U}$ fatto di aperti di carta la cui unione è densa (ne esiste uno fatto di un unico aperto)¹ e si definisce

$$\int_M f \tau = \sum_{U \in \mathcal{U}} \int_{x \in U} f(x) |h(x)| dx^1 \cdots dx^m.$$

Volendo ci si sarebbe potuti restringere ai sistemi di coordinate orientate positivamente. In quel caso h sarebbe stata sempre positiva e si sarebbe evitato il valore assoluto. $\square$

Esempio 7 La misura di Lebesgue λ_M di M , una varietà immersa in $\mathbb{R}^n$.

Svolgimento. Il prodotto scalare permette di definire un volume m -dimensionale associato ad ogni parallelepipedo m -dimensionale immerso in $\mathbb{R}^n$, per ogni $m \leq n$. Siano dati m vettori indipendenti $v_1, \dots, v_m$, e sia $\mathcal{P}$ il parallelepipedo m -dimensionale di vertici $P_I = P + \sum_{i \in I} v_i$ dove $I \subset \{1, \dots, m\}$. la matrice Gramiana associata ai v_i è la matrice $m \times m$ definita come $\text{Gram}(v_1, \dots, v_m) = (v_i \cdot v_j)_{i,j}$, il volume di $\mathcal{P}$ per definizione è

$$\text{gram}(v_1, \dots, v_m) = \sqrt{\det(\text{Gram}(v_1, \dots, v_m))}.$$

Supponiamo ora data una varietà m -dimensionale M immersa in $\mathbb{R}^n$. Vogliamo definire una σ -algebra di misurabili ed una misura in M . Parametizziamo M in un aperto denso U con un sistema di coordinate locali scelte a piacere

$$\mathbb{R}_u^m \hookrightarrow \mathbb{R}_x^n, \quad (u_1, \dots, u_m) \mapsto (x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m)).$$

In ogni punto $\mathbb{R}^n \supset M \ni \bar{q} = (\bar{u}_1, \dots, \bar{u}_m)$ della varietà il sistema di coordinate definisce m vettori tangenti $\partial_{u_i} x(\bar{u})$. Resta quindi definito un volume in M : dato A un sottoinsieme misurabile di M contenuto nella parametrizzazione locale, il volume di A è dato dall'integrale m -dimensionale

$$\lambda(A) = \int_{u \in A} \text{gram}\left(\frac{\partial x}{\partial u^i}(u)\right) du^1 \cdots du^m.$$

Indicheremo con λ la misura indotta in M dalla metrica Euclidea, e chiameremo anche questa misura di Lebesgue (eventualmente specificando *definita su M varietà immersa in $\mathbb{R}^n$*). Tale misura permette di integrare una funzione $f : M \rightarrow \mathbb{R}$ in un misurabile A usando la definizione

$$\int_A f d\lambda := \int_{u \in A} f(u) \text{gram}\left(\frac{\partial x}{\partial u^i}(u)\right) du^1 \cdots du^m.$$

Per dimostrare proposizioni, torna utile usare il linguaggio delle forme differenziali. Dimostriamo che gli stessi oggetti definiscono non solo una misura, ma una forma volume su M . Dato quindi uno spazio delle fasi M orientabile ed immerso in $\mathbb{R}^n$, la metrica Euclidea induce in M una forma volume che, preso un aperto di carta U ed un sistema di coordinate orientate positivamente u , ha espressione locale

$$\tau|_U(u) = \text{gram}\left(\frac{\partial x}{\partial u}(u)\right) du^1 \wedge \cdots \wedge du^m.$$

Perchè questa sia la definizione di una forma differenziale, bisogna e basta controllare la compatibilità per cambio di coordinate, bisogna cioè mostrare che qualora si scelga un nuovo aperto di carta V ed un sistema di coordinate v orientate positivamente, allora nell'intersezione $U \cap V$ vale

$$\text{gram}\left(\frac{\partial x}{\partial v}\right) = \text{gram}\left(\frac{\partial x}{\partial u}\right) \det\left(\text{Jac}\left(\frac{\partial u}{\partial v}\right)\right).$$

Un semplice conto dimostra che

$$\text{Gram}\left(\frac{\partial x}{\partial v}\right) = \text{Jac}\left(\frac{\partial x}{\partial v}\right)^T \text{Jac}\left(\frac{\partial x}{\partial v}\right) = \text{Jac}\left(\frac{\partial u}{\partial v}\right)^T \text{Jac}\left(\frac{\partial x}{\partial u}\right)^T \text{Jac}\left(\frac{\partial x}{\partial u}\right) \text{Jac}\left(\frac{\partial u}{\partial v}\right)^T,$$

quindi, se vale $\det \text{Jac}(\partial_v u) > 0$, si ottiene che

$$\text{gram}\left(\frac{\partial x}{\partial v}\right) = \det\left(\text{Jac}\left(\frac{\partial u}{\partial v}\right)\right) \text{gram}\left(\frac{\partial x}{\partial u}\right),$$

che è quello che bisognava dimostrare. $\square$

¹In alternativa si considera un atlante e si usano delle funzioni chiamate *partizioni dell'unità*.

Esempio 8 La misura di Lebesgue di una curva immersa in $\mathbb{R}^n$.

Svolgimento. La definizione appena data di misura indotta prescrive di scegliere una parametrizzazione della curva $\gamma : \mathbb{R}_t \rightarrow \mathbb{R}_x^n$, $t \mapsto \gamma(t)$. La matrice Jacobiana associata a questo sistema di coordinate è $Jac\left(\frac{\partial x}{\partial t}\right) = (\gamma'_1 \cdots \gamma'_n)^T$, il determinante gramiano è quindi $gram\left(\frac{\partial x}{\partial t}\right) = \sqrt{\gamma'^2_1 + \cdots + \gamma'^2_n}$, e quindi la forma volume è $\tau = |\gamma'(t)|dt$. Segue che se $A = [t_0, t_1]$ è un intervallo di $\mathbb{R}$, allora $\mu(A) = \int_{t=t_0}^{t_1} |\gamma'(t)|dt$, che è la lunghezza del tratto di curva da $\gamma(t_0)$ a $\gamma(t_1)$. $\square$

Esempio 9 La misura di Lebesgue della 2-sfera $S^2 \subset \mathbb{R}^3$.

Svolgimento. Scegliamo il sistema di coordinate più naturale possibile per parametrizzare l'emisfero boreale. La parametrizzazione è data dalla mappa $D^2_u \rightarrow \mathbb{R}^3_x$, $(u_1, u_2) \mapsto (u_1, u_2, \sqrt{1 - u^2_1 - u^2_2})$. La matrice Jacobiana associata a questo sistema di coordinate è

$$Jac\left(\frac{\partial x}{\partial u}\right) = \begin{pmatrix} \frac{1}{\sqrt{1-u^2_1-u^2_2}} & \frac{0}{\sqrt{1-u^2_1-u^2_2}} \\ \frac{0}{\sqrt{1-u^2_1-u^2_2}} & \frac{1}{\sqrt{1-u^2_1-u^2_2}} \end{pmatrix},$$

la matrice Gramiana è quindi

$$Gram\left(\frac{\partial x}{\partial u}\right) = \begin{pmatrix} 1 + \frac{u^2_1}{1-u^2_1-u^2_2} & \frac{u_1 u_2}{1-u^2_1-u^2_2} \\ \frac{u_1 u_2}{1-u^2_1-u^2_2} & 1 + \frac{u^2_2}{1-u^2_1-u^2_2} \end{pmatrix},$$

da cui segue che il determinante gramiano è $gram\left(\frac{\partial x}{\partial u}\right) = \frac{1}{\sqrt{1-u^2_1-u^2_2}}$ e quindi la forma volume nell'emisfero superiore è $\tau = \frac{1}{\sqrt{1-u^2_1-u^2_2}} du_1 \wedge du_2$.

Un'altra scelta ragionevole di coordinate è quella delle coordinate cilindriche,

$$(-1, 1)_{v_1} \times (0, 2\pi)_{v_2} \rightarrow \mathbb{R}^3_x, \quad (v_1, v_2) \mapsto \left(\sqrt{1-v^2_1} \cos v_2, \sqrt{1-v^2_1} \sin v_2, v_1\right).$$

La matrice Jacobiana associata a tali coordinate è

$$Jac\left(\frac{\partial x}{\partial v}\right) = \begin{pmatrix} -\frac{v_1}{\sqrt{1-v^2_1}} \cos v_2 & -\sqrt{1-v^2_1} \sin v_2 \\ -\frac{v_1}{\sqrt{1-v^2_1}} \sin v_2 & \sqrt{1-v^2_1} \cos v_2 \\ 1 & 0 \end{pmatrix},$$

la matrice Gramiana è quindi

$$Gram\left(\frac{\partial x}{\partial v}\right) = \begin{pmatrix} 1 + \frac{v^2_1}{1-v^2_1} & 0 \\ 0 & 1-v^2_1 \end{pmatrix},$$

da cui segue che il determinante gramiano è $gram\left(\frac{\partial x}{\partial v}\right) = 1$ e quindi la forma volume nell'emisfero superiore è $\tau = dv_1 \wedge dv_2$. $\square$

Definizione 10 Dato uno spazio con misura (M, μ) ed un automorfismo φ di M in se, il pull-back di μ rispetto a φ è la misura $\varphi^* \mu$ definita sulla stessa σ -algebra tale che $(\varphi^* \mu)(A) = \mu(\varphi^{-1}(A))$.

La misura μ si dice φ -invariante se $\varphi^* \mu = \mu$. A volte, quando è chiaro quale sia l'automorfismo φ , si omette la sua indicazione e si dice semplicemente che la misura è **invariante**.

Se la misura μ proviene da una forma differenziale τ , che relazione intercorre tra $\varphi^* \mu$ e $\varphi^* \tau$

Proposizione 11 Data una forma volume τ in una varietà M ed un automorfismo $\varphi : M \rightarrow M$, se la misura μ su M è associata ad una forma volume τ , allora la misura $\varphi^* \mu$ è associata alla forma volume $\varphi^* \tau$.

Dimostrazione. La dimostrazione di questo fatto non è altro che la formula di cambiamento di variabili negli integrali multipli [DM]. $\square$

1.2 Sistemi dinamici misurabili

Noi non studieremo gli spazi con misura (M, μ) per se, bensì investigheremo i sistemi dinamici ambientati in uno spazio con misura compatibile con la dinamica.

Definizione 12 Uno **shift** è un automorfismo $\Phi : M \rightarrow M$. Un **flusso** è uno spazio M dotato di una famiglia 1-parametrica di automorfismi $\Phi : \mathbb{R} \times M \rightarrow M$ che godono della proprietà $\Phi_0 = \mathbb{I}$, $\Phi_{t+s} = \Phi_t \circ \Phi_s$. Chiameremo **sistema dinamico** (continuo o discreto se necessario specificarlo) la coppia (M, Φ) dove Φ è uno shift od un flusso.

I flussi appaiono nell'ambito delle equazioni differenziali ordinarie, e sono associati a degli oggetti chiamati **campi vettoriali**. Essi sono uno dei modi con cui si formalizza la dinamica. Molti shift sono ottenuti a partire da campi vettoriali attraverso una opportuna discretizzazione del tempo, ma anche in modi meno immediati.

Come spesso succede, l'ipotesi che le mappe Φ siano automorfismi è una ipotesi detta di **reversibilità**. Questa ipotesi è a volte non verificata, ma semplifica notevolmente alcune dimostrazioni. In queste note assumeremo che sistemi dinamici siano reversibili.

Mettiamo ora assieme i concetti di sistema dinamico (M, Φ) e di spazio con misura (M, μ) .

Definizione 13 Un **sistema dinamico misurabile** è una tripletta (M, μ, Φ) in cui (M, μ) è uno spazio con misura, (M, Φ) è un sistema dinamico, e la misura μ è Φ -invariante.

Abbiamo definito la categoria in cui lavoreremo. È una categoria ragionevole, ovvero ci sono esempi di sistemi misurabili? Se esistono, in che modo vengono studiati? Useremo questo primo Capitolo a rispondere alla prima domanda. Nei Capitoli successivi affronteremo la seconda.

1.3 Primi esempi di sistema dinamico misurabile

Cominciamo con il piano Euclideo, una qualsiasi mappa $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ che sia una roto-traslazione (un elemento di $SE(2)$) preserva la misura di Lebesgue. In generale ogni shift che sia trasformazione affine-ortogonale (elemento di $SE(n)$) dello spazio Euclideo è, tautologicamente, un sistema dinamico misurabile di $\mathbb{R}^n$ con misura di Lebesgue. Passando a sistemi dinamici meno ovvi si ha

Esempio 14 Lo shift $\Phi(x, y) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (\frac{1}{2}x, 2y)$ preserva la misura di Lebesgue di $\mathbb{R}^2$. Il flusso $\Phi^t(x, y) = \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (e^{-t}x, e^ty)$ preserva la misura di Lebesgue di $\mathbb{R}^2$.

Dimostrazione. Senza usare il linguaggio delle forme differenziali, l'area di un romboide di lati v_1, v_2 è il valore assoluto del determinante della matrice $(v_1 \ v_2)$ (che coincide con $\text{gram}(v_1, v_2)$ con l'unica indicazione di avere cautela con il segno). Tale romboide viene mutato nel romboide di lati $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix} (v_1 \ v_2)$, che ha area uguale. Usando invece il linguaggio delle forme differenziali, si nota che il pull-back della forma $dx \wedge dy$ è la forma $(\frac{1}{2}dx) \wedge (2dy)$, che rimane quindi invariata.

Per il secondo esempio valgono considerazioni analoghe. □

Esempio 15 Lo shift del 2-toro in $\mathbb{S}^2$ $\Phi(\varphi, \psi) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = (\varphi + \psi, \varphi + 2\psi)$ preserva la misura di T^2 associata alla forma volume $d\varphi \wedge d\psi$, la cui misura associata è la misura di Lebesgue. Tale mappa si chiama gatto di Arnol'd.

Dimostrazione. Con argomento simile a quello dell'esempio precedente, l'area di un romboide (abbastanza piccolo) di lati v_1, v_2 , è il valore assoluto del determinante della matrice $(v_1 \ v_2)$. Tale romboide viene mutato nel romboide di lati $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} (v_1 \ v_2)$. Visto che il determinante del prodotto di due matrici è il prodotto dei determinanti, e che il determinante di $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ è 1, si ha che l'area del romboide non cambia.

Usando il linguaggio delle forme differenziali, Il pull-back della forma differenziale $d\varphi \wedge d\psi$ è la forma differenziale $(d\varphi + d\psi) \wedge (d\varphi + 2d\psi) = (d\varphi + d\psi) \wedge (d\varphi + 2d\psi) = d\varphi \wedge d\psi$. □

Prima di continuare, ricordiamo un teorema imparato nei corsi di Analisi.

Teorema 16 Sia dato un automorfismo Φ di $\mathbb{R}^n$, la misura di Lebesgue λ è Φ -invariante se e solo se il determinante della matrice Jacobiana di Φ rispetto ad ogni punto è 1. Sia dato un campo vettoriale di $\mathbb{R}^n$, la misura di Lebesgue è Φ_X^t -invariante se e solo se la divergenza di X è nulla.

La dimostrazione di questo teorema si trova nei testi di Analisi. Una applicazione immediata del teorema della divergenza permette di concludere che i sistemi Hamiltoniani definiscono dei sistemi dinamici che preservano una misura. Enunciamo il teorema di conservazione della misura nello spazio Euclideo con coordinate canoniche.

Teorema 17 Sia data una Hamiltoniana H nello spazio $\mathbb{R}_q^n \times \mathbb{R}_p^n$, il flusso Hamiltoniano associato al campo vettoriale $X_H = (\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q})^T$ preserva la misura di Lebesgue.

Dimostrazione. La divergenza di X_H è nulla. □

Per mantenere vivo l'interesse sulle forme differenziali, ricordiamo che l'ambiente in cui questo teorema può essere enunciato è quello delle varietà simplettiche (M, ω) con ω una 2-forma non degenera detta struttura simplettica. Il teorema in questo ambiente diventa

Teorema 18 Sia data una varietà simplettica (M, ω) ed una Hamiltoniana $H : M \rightarrow \mathbb{R}$. Il flusso Hamiltoniano associato ad H preserva la misura associata alla forma differenziale ω^n .

1.4 Secondo esempio: misura invariante ed insiemi di livello

I sistemi Hamiltoniani non solo preservano il volume, ma hanno pure un integrale primo, la funzione H . Se h è un valore regolare per H , cioè se tutti i punti del livello $M_h = \{(q, p) \in \mathbb{R}^{2n} | H(q, p) = h\}$ sono tali che $\nabla H(q, p) \neq 0$, allora $H = h$ è sottovarietà in $\mathbb{R}^{2n}$. Le sottovarietà M_h sono quindi varietà $2n - 1$ -dimensionali in cui è indotto un sistema dinamico. Esistono misure invarianti in M_h ?

Teorema 19 Sia $\mathbb{R}^{2n}$ uno spazio vettoriale simplettico ed H una Hamiltoniana. Allora gli insiemi di livello $M_h = \{x \in \mathbb{R}^{2n} | H(x) = h\}$ sono insiemi invarianti per il flusso Φ_H del campo vettoriale Hamiltoniano associato ad H ed in M_h la misura $\mu = \frac{1}{|\nabla H|} \lambda_{M_h}$ è invariante.

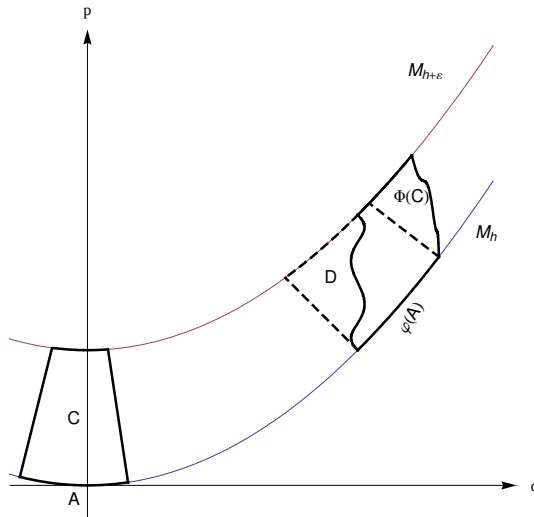
Dimostrazione. Cominciamo con una dimostrazione euristica: si userà la conservazione della misura di Lebesgue nell'intercapedine contenuta tra le superfici $M_h = \{H = h\}$ ed $M_{h+\varepsilon} = \{H = h + \varepsilon\}$. Poi, con un passaggio al limite per $\varepsilon \rightarrow 0$, si dimostrerà il teorema.

Preso un punto $(\bar{q}, \bar{p})$ in M_h , e scelto A un elemento infinitesimo di volume in M_h attorno a $(\bar{q}, \bar{p})$, costruiamo un cilindro C di base A ed altezza tracciata usando i vettori normali ∇H ai punti $(q, p) \in A$, allungati fino a che tocchino la ipersuperficie $M_{h+\varepsilon}$. A meno di infinitesimi di ordine ε^2 , il volume del cilindro C è $\lambda(C) = \lambda_{M_h}(A) \frac{\varepsilon}{|\nabla H(\bar{q}, \bar{p})|}$.

Agiamo ora con il flusso di Φ_H^t con t un tempo qualsiasi (per comodità da ora in poi chiamiamo la mappa semplicemente Φ). Sempre a meno di errori di ordine ε^2 , il cilindro si sposta nel cilindro D di base $\Phi(A)$ e di altezza $\frac{\varepsilon}{|\nabla H(\Phi(\bar{q}, \bar{p}))|}$, e quindi di volume $\lambda(D) = \lambda_{M_h}(\Phi(A)) \frac{\varepsilon}{|\nabla H(\Phi(\bar{q}, \bar{p}))|}$. Facendo tendere ε a zero si conclude che

$$\frac{\lambda_{M_h}(A)}{|\nabla H(\bar{q}, \bar{p})|} = \frac{\lambda_{M_h}(\Phi(A))}{|\nabla H(\Phi(\bar{q}, \bar{p}))|},$$

segue quindi che non è λ_{M_h} ad essere conservata, bensì la misura $\mu = \lambda_{M_h}/|\nabla H|$.



Una dimostrazione formale di questo teorema richiede l'uso di forme differenziali. La misura di Lebesgue di $\mathbb{R}^{2n}$ è associata alla forma volume che, in coordinate q, p , è data dall'espressione $\tau = dq_1 \wedge \dots \wedge dq_n \wedge dp_1 \wedge \dots \wedge dp_n$. L'ipotesi che h sia un valore regolare di H implica che in ogni punto (q, p) di M_h il gradiente $\nabla H(q, p) \neq 0$. Intorno ad M_h si può quindi definire il campo vettoriale $Y = \nabla H / |\nabla H|^2$.

Scegliamo una carta locale $\mathbb{R}_u^{2n-1} \rightarrow \mathbb{R}_{q,p}^{2n}$, $(u_1, \dots, u_{2n-1}) \mapsto (q(u), p(u))$ di M_h in un intorno di un punto fissato $(\bar{q}, \bar{p})$ di M_h . Il campo vettoriale Y e la carta appena scelta definiscono un nuovo sistema di coordinate in un intorno di tale punto in $\mathbb{R}^{2n}$, tale sistema di coordinate è dato dalla mappa

$$\mathbb{R}_{u,v}^{2n} \rightarrow \mathbb{R}_{q,p}^{2n}, \quad (u_1, \dots, u_{2n-1}, v) \mapsto \phi_Y^v(q(u), p(u)),$$

dove ϕ_Y^v è il flusso del campo vettoriale Y percorso per il tempo v .

Dal fatto che $Y(H) \equiv 1$ segue che l'iperpiano $\{(u, v) \in \mathbb{R}^{2n} \mid v = \varepsilon\}$ corrisponde alla ipersuperficie $M_{h+\varepsilon}$. Inoltre, la matrice Gramiana associata alle coordinate u, v calcolata nei punti di M_h è la matrice

$$\begin{pmatrix} \text{Gram}\left(\frac{\partial(q,p)}{\partial u}\right) & 0 \\ 0 & \frac{1}{|\nabla H|^2} \end{pmatrix},$$

e quindi il gramiano del cambio di coordinate è la funzione $\frac{\text{gram}\left(\frac{\partial(q,p)}{\partial u}\right)}{|\nabla H|}$. Segue che la forma volume Euclidea τ , in coordinate u, v ha la forma

$$\tau = \frac{\text{gram}\left(\frac{\partial(q,p)}{\partial u}\right)}{|\nabla H|} du^1 \wedge \dots \wedge du^{2n-1} \wedge dv.$$

Non ci resta che usare il fatto che il flusso Hamiltoniano Φ preserva la forma τ , ovvero che $\Phi^*(\tau) = \tau$ ed il fatto che Φ preserva la Hamiltoniana, ovvero che il flusso Φ , scritto in coordinate u, v , ha la forma $\Phi(u_1, \dots, u_{2n-1}, v) = (\Psi(u_1, \dots, u_{2n-1}, v), v)$. Denotiamo infine con φ il flusso Hamiltoniano ristretto ad M_h , ed osserviamo che, per definizione di φ , si ha

$$(\varphi(u_1, \dots, u_{2n-1}), 0) = \Psi(u_1, \dots, u_{2n-1}, 0).$$

La Jacobiana della mappa Φ calcolata in $v = 0$ è $Jac(\Phi(u, v)) = \begin{pmatrix} Jac(\varphi(u_1, \dots, u_{2n-1})) & * \\ 0 & 1 \end{pmatrix}$. Si ha quindi che

$$\begin{aligned} \Phi^*(\tau) &= \Phi^* \left(\frac{\text{gram}\left(\frac{\partial(q,p)}{\partial u}\right)}{|\nabla H|} du_1 \wedge \dots \wedge du_{2n-1} \wedge dv \right) = \\ &= \varphi^* \left(\frac{\text{gram}\left(\frac{\partial(q,p)}{\partial u}\right)}{|\nabla H|} du_1 \wedge \dots \wedge du_{2n-1} \right) \wedge dv. \end{aligned}$$

Segue quindi che

$$\varphi^* \left(\frac{\text{gram}\left(\frac{\partial(q,p)}{\partial u}\right)}{|\nabla H|} du_1 \wedge \dots \wedge du_{2n-1} \right) = \frac{\text{gram}\left(\frac{\partial(q,p)}{\partial u}\right)}{|\nabla H|} du_1 \wedge \dots \wedge du_{2n-1}.$$

Questo dimostra la tesi. □

1.4.1 Applicazione: l'oscillatore armonico

Applichiamo il teorema appena dimostrato ad un sistema Hamiltoniano semplice, l'oscillatore armonico. Lo spazio delle fasi è $\mathbb{R}^2$ con coordinate q, p e con la sua misura di Lebesgue, ovvero con la misura associata alla 2-forma $dq \wedge dp$. Il sistema dinamico è il sistema Hamiltoniano associato alla Hamiltoniana $H = \frac{1}{2}(p^2 + q^2)$.

Il campo Hamiltoniano associato ad H è il campo vettoriale $X_H = \begin{pmatrix} p \\ -q \end{pmatrix}$. Il campo vettoriale è lineare, e quindi il flusso di X_H si scrive usando la mappa esponenziale ed è $\Phi_{X_H}^t(q, p) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$. Le orbite sono cerchi concentrici, e coincidono con le superfici di livello M_h .

La misura di Lebesgue definita nelle superfici di livello M_h si calcola facilmente usando coordinate polari. Tali coordinate sono

$$I = \frac{p^2 + q^2}{2}, \quad \varphi = \arg(q + ip).$$

Le coordinate I, φ appena definite sono coordinate azione angolo e sono quindi coordinate simplettiche, ovvero la 2-forma $dq \wedge dp$ in queste coordinate si scrive come $dI \wedge d\varphi$. Inoltre, in queste coordinate la funzione H si scrive $H(I, \varphi) = I$. (Il cambio di coordinate inverso a quello scritto è $q = \sqrt{2I} \cos \varphi$ e $p = \sqrt{2I} \sin \varphi$, da cui segue che $dq \wedge dp = \left(\frac{\cos \varphi}{\sqrt{2I}} dI - \sqrt{2I} \sin \varphi d\varphi \right) \wedge \left(\frac{\sin \varphi}{\sqrt{2I}} dI + \sqrt{2I} \cos \varphi d\varphi \right) = dI \wedge d\varphi$.)

Nelle coordinate I, φ è facile scegliere una parametrizzazione degli insiemi di livello M_h , si usa infatti la funzione $S_u^1 \rightarrow \mathbb{R}_I^+ \times S_\varphi^1$, $u \mapsto (h, u)$. La Jacobiana associata a questa parametrizzazione è $\frac{\partial(I, \varphi)}{\partial u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, quindi la Gramiana è la matrice (1) ed il gramiano è la funzione costante 1.

D'altro canto, il gradiente di H è $\nabla H = (1, 0)$, che ha sempre norma 1. Segue che la misura invariante è du . Ricordando la corrispondenza tra u e φ , si nota che la misura invariante associa ad un intervallo (u_0, u_1) la lunghezza dell'arco sotteso all'angolo $u_0 - u_1$, tale lunghezza non varia durante la dinamica.

Avremmo potuto seguire un altro approccio, usando le più naturali coordinate cartesiane. In coordinate cartesiane una parametrizzazione naturale di parte dell'insieme di livello M_h con $h > 0$ è la funzione $] -\sqrt{2h}, \sqrt{2h}[\rightarrow \mathbb{R}^2$, $v \mapsto (v, \sqrt{2h - v^2})$. Per calcolare la misura indotta sulla superficie di livello M_h scriviamo

$$\frac{\partial(q, p)}{\partial v} = \left(-\frac{1}{\sqrt{2h - v^2}} \right),$$

da cui segue che il determinante Gramiano è la funzione $\sqrt{1 + \frac{v^2}{2h - v^2}}$. Si ottiene che la misura indotta sulla circonferenza, scritta usando la coordinata v è

$$\sqrt{1 + \frac{v^2}{2h - v^2}} dv = \sqrt{\frac{2h}{2h - v^2}} dv.$$

D'altro canto, il gradiente di H è il vettore $\nabla H = \begin{pmatrix} q \\ p \end{pmatrix}$, quindi la sua norma calcolata nel punto $(v, \sqrt{2h - v^2})$ è $\sqrt{v^2 + \sqrt{2h - v^2}^2} = \sqrt{v^2 + 2h - v^2} = \sqrt{2h}$. Ne segue che la misura invariante è associata alla 1-forma $dv/\sqrt{2h - v^2}$.

1.4.2 Applicazione: sistemi 1D conservativi

Estendiamo ora l'applicazione al caso di sistemi Hamiltoniani ad un grado di libertà. Nello spazio delle fasi $\mathbb{R}_{q,p}^2$ sia data l'Hamiltoniana $H = \frac{1}{2}p^2 + V(q)$, con $V(q)$ un potenziale qualsiasi. L'equazione differenziale che governa questo sistema dinamico è

$$\dot{q} = p, \quad \dot{p} = -V'(q).$$

Fissato un livello di energia h , le curve di livello M_h si ottengono come imparato nei corsi di base di Fisica Matematica. Si usa la conservazione dell'energia $H = \frac{1}{2}p^2 + V(q) = h$ e si ricava che $p = \pm\sqrt{2}\sqrt{h - V(q)}$.

Due metà del livello $H = h$ possono venire agevolmente parametrizzate usando la funzione

$$u \mapsto (u, \pm\sqrt{2}\sqrt{h - V(u)}),$$

dove u varia nelle zone concesse, ovvero negli intervalli in cui $h - V(u)$ è positivo. Per calcolare la forma la cui misura associata è invariante per il flusso Hamiltoniano associato ad H si calcola

$$\frac{\partial(q, p)}{\partial u} = \left(-\frac{1}{\sqrt{2}\sqrt{h - V(u)}} \right),$$

e si ottiene quindi che il determinante Gramiano è la funzione $\sqrt{1 + \frac{V'(u)^2}{2(h - V(u))}}$.

D'altra parte, il gradiente di H è il vettore

$$\nabla H(q(u), p(u)) = \begin{pmatrix} V'(q(u)) \\ p(u) \end{pmatrix} = \begin{pmatrix} V'(u) \\ \pm\sqrt{2}\sqrt{h - V(u)} \end{pmatrix},$$

quindi la sua norma è

$$|\nabla H| = \sqrt{2(h - V(u)) + V'(u)^2}.$$

Segue che la misura invariante è associata alla forma differenziale

$$\tau = \frac{\sqrt{1 + \frac{V'(u)^2}{2(h - V(u))}} du}{\sqrt{2(h - V(u)) + V'(u)^2}} = \frac{\sqrt{\frac{2(h - V(u)) + V'(u)^2}{2(h - V(u))}} du}{\sqrt{2(h - V(u)) + V'(u)^2}} = \frac{du}{\sqrt{2(h - V(u))}}.$$

Questa forma differenziale ha una lunga storia alle spalle. Infatti nei sistemi meccanici ad un grado di libertà vale la relazione $p = \pm\sqrt{2}\sqrt{h - V(q)}$, ma vale anche $\dot{q} = p$. Si ottiene quindi che $\frac{dq}{dt} = \pm\sqrt{2}\sqrt{h - V(q)}$ da cui, per separazione delle variabili, si ottiene

$$\frac{dq}{\pm\sqrt{2(h - V(q))}} = dt$$

Si deduce quindi che la dinamica, per andare dallo stato $(q_0, \sqrt{2}\sqrt{h - V(q_0)})$ al punto $(q_1, \sqrt{2}\sqrt{h - V(q_1)})$, impiega il tempo

$$\Delta t = \int_{t_{q_0}}^{t_{q_1}} dt = \int_{q_0}^{q_1} \frac{dq}{\sqrt{2(h - V(q))}}.$$

(Con ovvia simmetria nel caso in cui si scelga il segno meno.)

Abbiamo quindi scoperto che la misura appena definita associa ad un tratto di orbita il tempo che impiega la dinamica ad andare da un estremo all'altro. Ovviamente il flusso cambierà la lunghezza metrica del tratto di orbita, ma non cambierà il tempo che si impiega a percorrerlo, infatti una delle proprietà del flusso è che vale $\Phi^t \circ \Phi^s = \Phi^{t+s}$, da cui segue facilmente che il tempo di percorrenza è un invariante del moto.

1.4.3 Caso con più integrali primi

Quando il sistema dinamico preserva una misura e le quantità conservate non sono una ma due o più, esempio tipico sono il caso un sistema Hamiltoniano di Hamiltoniana H con un momento conservato J , allora si può definire una misura invariante sulla superficie di livello di tutti gli integrali primi, che è una sottovarietà di codimensione uguale al numero di integrali primi.

Teorema 20 *Sia dato un sistema dinamico $(\mathbb{R}^n, \Phi)$ che preserva la misura di Lebesgue in $\mathbb{R}^n$, e siano $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$, con $i = 1, \dots, k$, integrali del moto. Allora, se $(f_1, \dots, f_k)$ è un valore regolare di $(F_1, \dots, F_k)$, la superficie $M_f = \{F = f\}$ è invariante per la dinamica ed ammette una misura invariante associata alla forma differenziale*

$$\tau = \frac{\text{gram}(\partial_{u^1} x, \dots, \partial_{u^{m-k}} x)}{\text{gram}(\nabla F_i)} du^1 \wedge \dots \wedge du^{m-k}.$$

In altre parole, è conservata la misura $\mu = \frac{1}{\text{gram}(\nabla F_i)} \lambda_{M_f}$, misura assolutamente continua rispetto alla misura di Lebesgue indotta in M_f .

Non daremo dimostrazione di questo fatto, dimostrazione che comunque non è molto diversa da quella presentata per il caso Hamiltoniano con un solo integrale primo. Osserviamo che gli ingredienti sono solo campo vettoriale che conserva la misura di Lebesgue ed ha integrali primi. Questo enunciato include il caso Hamiltoniano come caso particolare, ma non è specifico per i sistemi Hamiltoniani.

1.5 Terzo esempio: la mappa di Poincaré

Finora abbiamo trattato sistemi dinamici continui e non abbiamo ancora visto nessun esempio di sistema dinamico discreto. Esiste una categoria molto vasta ed estremamente importante di sistemi dinamici discreti che vengono definiti e studiati per ottenere informazioni cruciali su sistemi dinamici continui. Vediamo in cosa consistono.

Definizione 21 *Dato un flusso in $\mathbb{R}^n$ associato ad un campo vettoriale X , una ipersuperficie N si dice **sezione di Poincaré** se esiste una mappa continua $T : N \rightarrow \mathbb{R}^+$, $q \mapsto T(q)$ tale che $\Phi_X^{T(q)}(q) \in N$ e $\Phi_X^s(q) \notin N$ per ogni $0 < s < T(q)$. Il reale positivo $T(q)$ si chiama **tempo di primo ritorno**.*

Se inoltre per ogni q in $\mathbb{R}^n$ si ha che $\Phi_X^t(q) \in N$ per qualche $t \in \mathbb{R}$, allora la varietà N si dice **sezione di Poincaré globale**. Data una sezione di Poincaré, rimane definita una mappa Ψ che ad ogni x in N associa $\Psi(q) = \Phi_X^{T(q)}(q)$. Questa mappa si chiama **mappa di Poincaré**. Lo scopo di questa sottosezione è quello di definire una misura Ψ -invariante nella varietà N . Si osservi che in questo modo resta definito un sistema misurabile in cui la dinamica è uno shift e non un flusso.

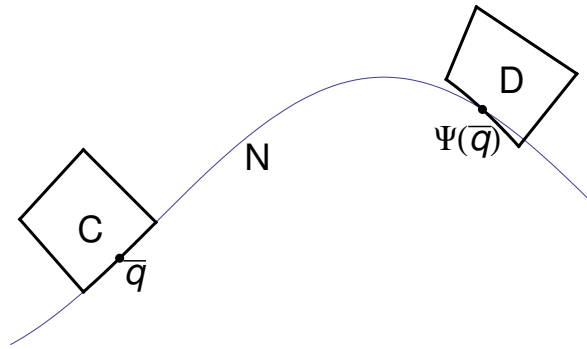
Teorema 22 *Sia X un campo vettoriale in $\mathbb{R}^n$ che preserva la misura di Lebesgue λ , sia N una sezione di Poincaré per il flusso di X , e sia $X_\perp$ la proiezione di X sulla direzione ortogonale ad N . Allora la misura $\mu = |X_\perp| \lambda_N$ definita su N è invariante sotto l'azione della mappa di Poincaré Ψ .*

Dimostrazione. Anche in questo caso provvediamo prima una dimostrazione euristica e poi una dimostrazione formale. Per la dimostrazione euristica fissiamo un punto $\bar{q}$ di N , e scegliamo un elemento infinitesimo di volume A contenente $\bar{q}$. Consideriamo il cilindro

$$C = \{\Phi_X^t(q) \mid q \in A, t \in [0, \varepsilon]\}.$$

Preliminarmente, osserviamo che dato un dominio A contenuto in $\mathbb{R}^{n-1} \subset \mathbb{R}^n$ ed un vettore v in $\mathbb{R}^n$ non appartenente ad $\mathbb{R}^{n-1}$, l'area del cilindro $\{a + \ell v \mid a \in A, \ell \in [0, \varepsilon]\}$ è $\varepsilon \lambda_{\mathbb{R}^{n-1}}(A) |v_\perp|$, dove $v_\perp = (v \cdot e_n) e_n$ è la componente di v ortogonale al tangente ad N . L'estensione di questo fatto al caso non-lineare indica che la misura di Lebesgue di C è, a meno di infinitesimi di ordine ε^2 , il reale positivo $\lambda_N(A) |X_\perp(\bar{q})|$.

Siccome il flusso di X preserva la misura di Lebesgue, il volume di C non varia con il flusso del campo vettoriale X . Quindi, se $T = T(\bar{q})$ è il tempo di primo ritorno per il punto $\bar{q}$, l'insieme $D = \Phi_X^T(C)$ ha stesso volume di C . Ma a meno di infinitesimi di ordine ε^2 , l'insieme D ha volume uguale a quello del cilindro di base $\Psi(A)$ ottenuto fluendo con Φ_X per il tempo ε . Per motivi già detti, tale cilindro ha volume $\lambda_N(\Psi(A)) |X_\perp(\Psi(\bar{q}))| + O(\varepsilon^2)$. Derivando rispetto a ε e calcolando in $\varepsilon = 0$ si dimostra la tesi.



Per una dimostrazione formale, scegliamo delle coordinate locali $u^1, \dots, u^{n-1}$ attorno al generico punto $\bar{q}$ di N , parametrizziamo quindi un intorno di $\bar{q}$ in $\mathbb{R}^n$ con le coordinate in $\mathbb{R}_u^{n-1} \times \mathbb{R}_v \rightarrow \mathbb{R}^n$ date da $(u, v) \mapsto \Phi_X^v(u)$.

La Jacobiana di questo cambio di coordinate è la matrice $\begin{pmatrix} \partial_u x \\ X(\Phi_X^v(u)) \end{pmatrix}$, quindi il gramiano è la matrice $\begin{pmatrix} \text{Gram}(\partial_u x) & * \\ * & X \cdot X \end{pmatrix}$. Usando il fatto che la matrice Gramiana è simmetrica, una operazione su righe e colonne permette di mostrare che il suo determinante è $\text{gram}(\partial_u x) |X_\perp|$. Si ottiene quindi che la forma di Lebesgue $dx^1 \wedge \dots \wedge dx^n$ si riscrive nelle nuove coordinate come $\text{gram}(u^1, \dots, u^{n-1}) |X_\perp| du^1 \wedge \dots \wedge du^{n-1} \wedge dv$.

Il flusso Φ_X^t di X , per ipotesi, conserva il volume Euclideo. Se T è il tempo di primo ritorno per il punto $\bar{q}$, si ha che $\Phi_X^T(u, v) = (\Psi(u), v + T - T(u))$ con Ψ la mappa di Poincaré (per definizione). Si ha quindi che

$$\text{gram}(\partial_u x) |X_\perp| du \wedge dv = dx = (\Phi_X^T)^*(dx) = (\Phi_X^T)^*(\text{gram}(\partial_u x) |X_\perp| du \wedge dv) = \Psi^* \text{gram}(u) |X_\perp| du \wedge dv.$$

Si conclude quindi che $\text{gram}(\partial_u x(u)) |X_\perp(x(u))| du$ è invariante sotto il flusso della mappa di Poincaré. $\square$

Esempio 23 Sia dato un oscillatore armonico forzato, governato dall'equazione del secondo ordine

$$\ddot{q} = -q + \sin t.$$

Studiare la mappa di Poincaré e l'invarianza della misura definita al teorema precedente.

Dimostrazione. Questo sistema non è autonomo, e quindi non c'è unicità delle soluzioni nello spazio $\mathbb{R}_{q,p}^2$. C'è invece unicità delle soluzioni nello spazio $\mathbb{R}_{q,p}^2 \times \mathbb{R}_t$ per il sistema di equazioni

$$\dot{q} = p, \quad \dot{p} = -q + \sin t, \quad \dot{t} = 1.$$

In altre parole si può considerare il campo vettoriale $X(q, p, t) = \begin{pmatrix} p \\ -q + \sin t \\ 1 \end{pmatrix}$. Visto che la forzante $\sin t$ è periodica, si può pensare che il sistema sia ambientato in uno spazio delle fasi $\mathbb{R}_{q,p}^2 \times S_t^1$. Possiamo quindi usare lo spazio $t = 0$ come sezione di Poincaré globale.

La componente normale $X_\perp$ allo spazio $t = 0$ è il vettore $(0, 0, 1)^T$, ed ha sempre norma 1, e la misura di Lebesgue indotta su $\mathbb{R}_{q,p}^2 \subset \mathbb{R}_{q,p}^2 \times S_t^1$ è la misura di Lebesgue di $\mathbb{R}_{q,p}^2$. Quindi, la misura invariante μ è proprio

la misura di Lebesgue λ , associata alla 2-forma $dq \wedge dp$. Per studiare l'invarianza della misura bisogna calcolare il flusso del sistema dinamico, che necessita l'uso del principio di variazione delle costanti. Se la forzante fosse nulla, allora il sistema si direbbe omogeneo e la sua soluzione si scriverebbe usando la matrice esponenziale (detta anche risolvente).

$$\begin{pmatrix} q_t \\ p_t \end{pmatrix} = e^{t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}} \begin{pmatrix} q_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} q_0 \\ p_0 \end{pmatrix}$$

Nel problema in questione cerchiamo le soluzioni al sistema non omogeneo

$$\begin{pmatrix} q \\ p \end{pmatrix}^\bullet = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} + \begin{pmatrix} 0 \\ \sin t \end{pmatrix},$$

che si ottengono sommando alle soluzioni del problema omogeneo una soluzione particolare. La soluzione particolare si calcola assumendo che le costanti di integrazione q_0, p_0 siano funzioni di t . Si cercano quindi funzioni $q_0(t), p_0(t)$ in modo che la funzione

$$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} q_0(t) \\ p_0(t) \end{pmatrix}$$

risolva il problema non-omogeneo.

Chiamiamo A_t la matrice con i seni e coseni, da bravo risolvente $\dot{A}_t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A_t$. Deriviamo.

$$\dot{A}_t \begin{pmatrix} q_0(t) \\ p_0(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A_t \begin{pmatrix} q_0(t) \\ p_0(t) \end{pmatrix} + A_t \begin{pmatrix} \dot{q}_0(t) \\ \dot{p}_0(t) \end{pmatrix}.$$

Si devono quindi determinare delle funzioni $q_0(t), p_0(t)$ tali che $A_t \begin{pmatrix} \dot{q}_0(t) \\ \dot{p}_0(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \sin t \end{pmatrix}$, ovvero si deve risolvere l'equazione

$$\begin{pmatrix} \dot{q}_0(t) \\ \dot{p}_0(t) \end{pmatrix} = A_{-t} \begin{pmatrix} 0 \\ \sin t \end{pmatrix} = \begin{pmatrix} -\sin^2 t \\ \sin t \cos t \end{pmatrix}.$$

Questo sistema si integra e porge

$$q_0(t) = -\frac{t}{2} + \frac{\sin t \cos t}{2} \quad p_0(t) = \frac{1}{2} - \frac{1}{2} \cos^2 t$$

(si sono scelte le costanti di integrazione così che $q_0(0) = p_0(0) = 0$)

Le soluzioni generali sono

$$\begin{pmatrix} q_t \\ p_t \end{pmatrix} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} q_0 \\ p_0 \end{pmatrix} + \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} -\frac{t}{2} + \frac{\sin t \cos t}{2} \\ \frac{1}{2} - \frac{1}{2} \cos^2 t \end{pmatrix}$$

Si ottiene quindi che la mappa di Poincarè è

$$\Phi(q_0, p_0) = \begin{pmatrix} q_{2\pi} \\ p_{2\pi} \end{pmatrix} = \begin{pmatrix} q_0 - \pi \\ p_0 \end{pmatrix}.$$

Questa mappa preserva la misura di Lebesgue. □

1.6 Quarto esempio: I biliardi

I biliardi sono una famiglia di sistemi molto studiata, sorgente di esempi non banali che illustrano la teoria e di complicati problemi aperti. Noi ci interesseremo di una classe particolare di biliardi, i **biliardi convessi**. Un biliardo convesso consiste dei seguenti dati: una immersione liscia γ di un cerchio S^1 in $\mathbb{R}^2$ tale che la componente limitata di $\mathbb{R}^2 \setminus \gamma(S^1)$ sia un insieme convesso, ed un punto materiale P di massa m che si muove come una particella libera (moto rettilineo uniforme) all'interno del dominio convesso e rimbalza in modo totalmente elastico quando impatta il bordo del dominio (l'angolo di incidenza è uguale all'angolo di riflessione).

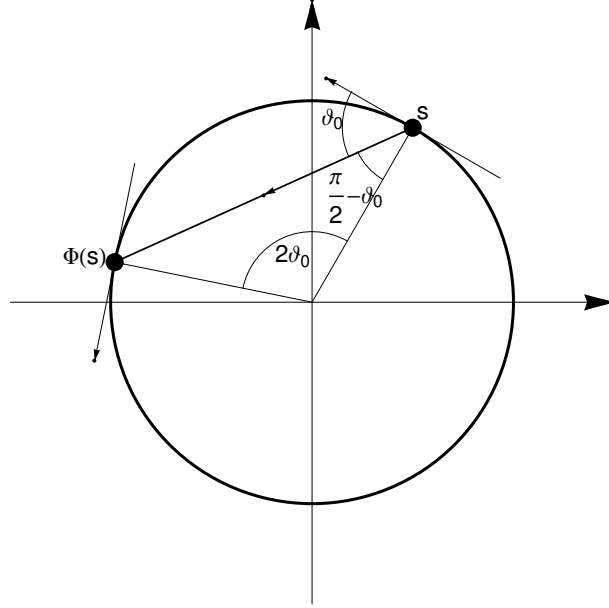
La **mappa del biliardo** è una sorta di riassunto esaustivo della corsa del punto P . È infatti la mappa Φ che ad un punto di impatto ed un angolo di incidenza associa il punto di impatto e l'angolo di incidenza dell'impatto a venire. Scelto un parametro s per la curva $s \mapsto \gamma(s)$, la dinamica del biliardo avviene in $\Phi: S_s^1 \times (0, \pi)_\vartheta \rightarrow S_s^1 \times (0, \pi)_\vartheta$, $(s, \vartheta) \mapsto \Phi(s, \vartheta) = (s', \vartheta')$.

1.6.1 Biliardi a bordo circolare

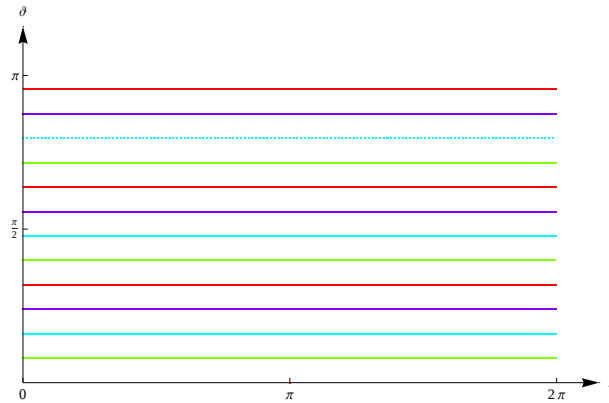
Investighiamo il caso più semplice di biliardo convesso, quello in cui il bordo è un cerchio. Possiamo con evidenza riscalarle le coordinate cartesiane così da assumere che il raggio del cerchio sia unitario.

Una scelta di coordinate s che parametrizzi il cerchio è data da $s \mapsto (\cos s, \sin s)$. Mettiamo in moto il punto materiale P da un punto s_0 della circonferenza con una velocità iniziale che forma un angolo ϑ_0 con il tangente orientato alla circonferenza in s_0 , con semplici considerazioni geometriche (vedi figura) si dimostra che la mappa del biliardo è la mappa

$$\Phi : S^1 \times (0, \pi) \rightarrow S^1 \times (0, \pi), \quad \Phi(s_0, \vartheta_0) = (s_0 + 2\vartheta_0, \vartheta_0)$$



È immediato osservare l'esistenza di una misura invariante: la misura di Lebesgue in $\mathbb{R} \times S^1$, associata alla 2-forma $ds \wedge d\vartheta$. Esiste inoltre una quantità conservata, infatti la mappa $I : (s, \vartheta) \mapsto \vartheta$ è un integrale del moto, e se $i \in (0, \pi)$ allora la curva di livello $I = i$ è topologicamente un cerchio. Infine, il ritratto in fase del sistema dinamico discreto può essere disegnato nel seguente modo.



che è la mappa standard imperturbata. Una caratteristica notevole di questo ritratto in fase è che esiste una famiglia densa (in biiezione con $i\pi\mathbb{Q}$ in $(0, \pi)$) di valori di I per la quale la dinamica nei cerchi $I = i$ è periodica, ma anche esiste una famiglia altrettanto densa per la quale la dinamica nei cerchi $I = i$ è densa. Le perturbazioni di questa mappa producono dinamiche estremamente interessanti.

È facile osservare inoltre che l'insieme delle traiettorie che soddisfano a $I = i$ altro non sono che l'involuppo di un cerchio di raggio $R \cos \bar{i}$, per qualche fissato $\bar{i}$, noto con il nome di **caustica**.

1.6.2 Biliardi a bordo qualsiasi

Per complicare un po' le cose, si consideri ora un biliardo a bordo non circolare. La dinamica è in questo caso notevolmente più complicata. Già per biliardi a bordo ellittico il calcolo della lunghezza della curva porta

agli *integrali ellittici*, per i quali è necessaria una complicata ed interessante teoria e l'introduzione di funzioni trascendenti non studiate nei corsi di base. A vista si notano due traiettorie periodiche di periodo 2, il cui supporto sono gli assi maggiore e minore.

L'automorfismo Φ di $S_s^1 \times (0, \pi)_\vartheta$ può essere riscritto anche in altre coordinate interessanti

$$\Psi : S_s^1 \times (-1, 1)_c \rightarrow S^1 \times (-1, 1)$$

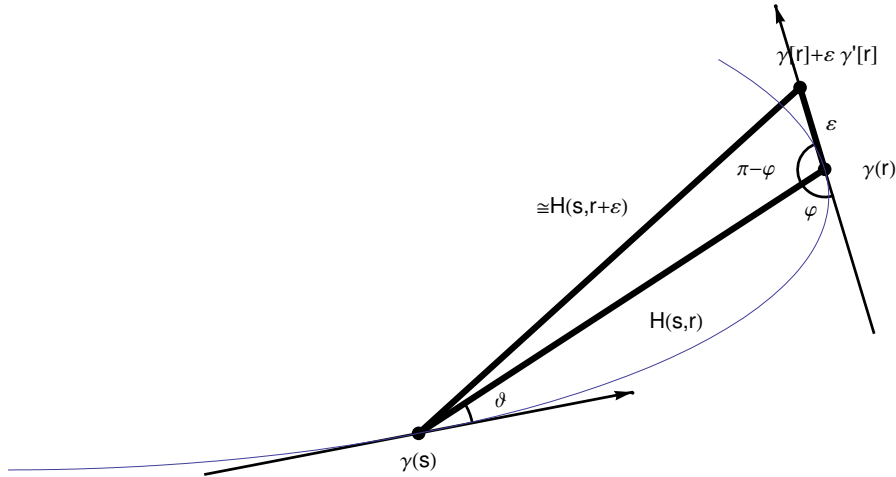
dove c è il coseno dell'angolo di incidenza, oppure nelle coordinate

$$\chi : S_s^1 \times S_r^1 \rightarrow S_s^1 \times S_r^1$$

dove r è la coordinata del punto di prossimo contatto. Naturalmente la funzione $\chi(s, r) = (\chi_1(s, r) = r, \chi_2(s, r))$. In tutte queste mappe, la scelta "corretta" per la parametrizzazione del bordo è il parametro d'arco s , per cui assumiamo che s si muova nella circonferenza di raggio $L/2\pi$, dove L è la lunghezza della curva. La lunghezza L interviene solo in via teorica, non dovremo quindi calcolare esplicitamente integrali ellittici.

Teorema 24 *La mappa del biliardo Ψ conserva la misura μ associata alla 2-forma $ds \wedge dc$.*

Dimostrazione. Usiamo le coordinate r, s nelle quali è definita la funzione χ . Sia H la funzione di $S_s^1 \times S_r^1$ in $\mathbb{R}$ data da $H(s, r) = |\gamma(s) - \gamma(r)|$. Intendiamo studiare la dipendenza di H da s, r , per far questo disegniamo la traiettoria del biliardo tra il punto di impatto di coordinata d'arco s e quello successivo di coordinata d'arco r , supponiamo quindi di aumentare r di poco, ovvero di tramutarlo in $r + \varepsilon$. La formula di Taylor dice che $\gamma(r + \varepsilon) = \gamma(r) + \varepsilon\gamma'(r) + O(\varepsilon^2)$. Consideriamo quindi il triangolo $\gamma(s), \gamma(r), \gamma(r) + \varepsilon\gamma'(r)$



Ricordiamo la legge dei coseni, la quale dice che dato un triangolo di lati A, B, C ed angoli opposti ai lati α, β, γ , allora vale $A^2 = B^2 + C^2 - 2BC \cos \alpha$. Se s, r sono punti di impatto successivi, e ϑ, φ sono angoli di incidenza ad impatti successivi, si ha che

$$H(s, r + \varepsilon) = \sqrt{H^2 + \varepsilon^2 - 2H\varepsilon \cos(\pi - \varphi)} + O(\varepsilon^2),$$

dove si è usato che $|\gamma'(r)| = 1$. Segue quindi che $\partial_r H(s, r) = \partial_\varepsilon H(s, r + \varepsilon)|_{\varepsilon=0} = \cos \varphi$. Per motivi analoghi si ha che $\partial_s H(s, r) = -\cos \vartheta$.

Nelle coordinate (s, c) la mappa Ψ si scrive $\Psi(s, c) = (f(s, c), g(s, c))$. Chiamiamo $s' = f(s, c)$ e $c' = g(s, c)$, e scriviamo la funzione H nelle coordinate s, c , ovvero consideriamo la funzione $\tilde{H}(s, c) = H(s, f(s, c)) = H(s, s')$. Per la regola della catena e per il conto appena fatto si ha che

$$\begin{aligned} \frac{\partial \tilde{H}}{\partial s}(s, c) &= \frac{\partial H}{\partial s}(s, f(s, c)) + \frac{\partial H}{\partial r}(s, f(s, c)) \frac{\partial f}{\partial s}(s, c) = \\ &= \frac{\partial H}{\partial s}(s, s') + \frac{\partial H}{\partial r}(s, s') \frac{\partial f}{\partial s}(s, c) = \\ &= -c + c' \frac{\partial f}{\partial s}(s, c) = -c + g(s, c) \frac{\partial f}{\partial s}(s, c). \end{aligned}$$

D'altro canto, si ha anche che

$$\begin{aligned}\frac{\partial \tilde{H}}{\partial c}(s, c) &= \frac{\partial H}{\partial r}(s, f(s, c)) \frac{\partial f}{\partial c}(s, c) = \\ &= \frac{\partial H}{\partial r}(s, s') \frac{\partial f}{\partial c}(s, c) = \\ &= c' \frac{\partial f}{\partial c}(s, c) = g(s, c) \frac{\partial f}{\partial c}(s, c).\end{aligned}$$

Andiamo ora a calcolare $\partial_c \partial_s \tilde{H}(s, c)$ ed identifichiamolo con $\partial_s \partial_c \tilde{H}(s, c)$. Si ottiene che

$$\begin{aligned}\frac{\partial \left(-c + g(s, c) \frac{\partial f}{\partial c}(s, c) \right)}{\partial c} &= -1 + \frac{\partial g}{\partial c}(s, c) \frac{\partial f}{\partial c}(s, c) + g(s, c) \frac{\partial^2 f}{\partial c \partial s}(s, c) = \\ &= \frac{\partial \left(g(s, c) \frac{\partial f}{\partial c}(s, c) \right)}{\partial s} = \frac{\partial g}{\partial s}(s, c) \frac{\partial f}{\partial c}(s, c) + g(s, c) \frac{\partial^2 f}{\partial c \partial s}(s, c).\end{aligned}$$

da questa equazione segue che

$$1 = \frac{\partial g}{\partial c}(s, c) \frac{\partial f}{\partial c}(s, c) - \frac{\partial g}{\partial s}(s, c) \frac{\partial f}{\partial c}(s, c),$$

il che implica che la Jacobiana della funzione Ψ ha determinante 1. Segue quindi che Ψ preserva il volume Euclideo. $\square$

La mappa H si chiama **funzione generatrice** e gioca un ruolo importante in molte dimostrazioni.

Osservazione 25 *Il diametro di un biliardo è il diametro del minimo cerchio circoscritto. La larghezza di un biliardo è il diametro del massimo cerchio inscritto. I punti del biliardo B che appartengono all'intersezione del bordo del biliardo con il cerchio che realizza il diametro, rispettivamente la larghezza, sono gli estremi di due traiettorie di periodo 2. Segue quindi che esistono sempre almeno due orbite di periodo 2 per la mappa del biliardo.*

Dimostrazione. Fissiamo s un punto di impatto. La funzione $F_s(t) : [0, L] \rightarrow \mathbb{R}$ che, fissato un punto s associa a t l'angolo ad s della retta congiungente $\gamma(s)$ con $\gamma(s+t)$, manda 0 in 0 ed L in π ed è monotona crescente. Invece, la funzione $G_s(t) : [0, L] \rightarrow \mathbb{R}$ che associa a t l'angolo ad $s+t$ della retta congiungente $\gamma(s+t)$ con $\gamma(s)$, manda 0 in π ed L in 0 ed è monotona decrescente.

Sia quindi una funzione $S^1 \rightarrow S^1$, $s \mapsto t(s)$, dove $t(s)$ è l'unico valore per cui vale $F_s(t(s)) = G_s(t(s))$. Per ogni s , il valore $t(s)$ è necessariamente incluso nell'intervallo aperto $(0, L)$. Segue che la funzione $h(s) = H(s, s+t(s)) : S^1 \rightarrow \mathbb{R}$ non è mai nulla ed è quindi ovunque differenziabile in s . Inoltre, la funzione è limitata dal diametro del biliardo.

Si ha perciò che la funzione h ha necessariamente un massimo ed un minimo (potrebbe ovviamente avere più punti critici). Nei punti critici si ha che

$$\begin{aligned}0 = \frac{\partial h}{\partial s} &= \frac{\partial H}{\partial s}(s, s+t(s)) + \frac{\partial H}{\partial r}(s, s+t(s)) \left(1 + \frac{\partial t}{\partial s} \right) = \\ &= \cos \left(F_s(t(s)) \right) - \cos \left(G_s(t(s)) \right) \left(1 + \frac{\partial t}{\partial s} \right) = \\ &= \cos \left(F_s(t(s)) \right) \left(2 + \frac{\partial t}{\partial s} \right).\end{aligned}$$

(l'ultima identità segue dal fatto che, per definizione di $t(s)$, $F_s(t(s)) = G_s(t(s))$). È inoltre evidente che la funzione $s \mapsto s+t(s)$ è crescente, e quindi $2+t'(s) > 0$. Segue che $\cos(F_s(t(s))) = 0$, ovvero che $F_s(t(s)) = \pi/2$. Si sono quindi determinate due orbite di periodo 2 i cui punti di impatto sono s ed $s+t(s)$. $\square$

1.7 Schemi di Bernoulli

Esempio 26 *Chiamiamo vettore di probabilità una lista $p = (p_1, \dots, p_n)$ di numeri reali positivi, tutti minori di 1 e tali che $p_1 + \dots + p_n = 1$ (permettiamo anche che n sia il simbolo ∞ , ovvero p è una successione la cui*

serie associata ha somma 1). Lo spazio delle fasi di uno schema di Bernoulli è lo spazio $B(p)$ i cui elementi sono serie bi-infinite $\sigma = (\dots, \sigma_{-1}, \sigma_0, \sigma_1, \dots)$ con σ_i simboli corrispondenti agli indici dei numeri p_i .

Lo spazio $B(p)$ si dota in modo naturale di uno shift definito come $\Phi(\sigma)_i = \sigma_{i+1}$, tale shift è detto lo **shift a sinistra**, si dota inoltre di una misura Φ -invariante: chiamiamo **cilindri elementari** i sottoinsiemi di $B(p)$ definiti come $C_{\ell_1, \dots, \ell_k}^{j_1, \dots, j_k} = \{\sigma | \sigma_{j_1} = \ell_1, \dots, \sigma_{j_k} = \ell_k\}$, decidiamo che la misura di $C_{\ell_1, \dots, \ell_k}^{j_1, \dots, j_k}$ è il numero reale $\mu(C_{\ell_1, \dots, \ell_k}^{j_1, \dots, j_k}) = p_{\ell_1} \cdots p_{\ell_k}$. Tale definizione si estende in modo unico alla σ -algebra generata dai cilindri elementari, e rende quindi lo spazio $B(p)$ uno spazio con misura e lo shift a sinistra un automorfismo che preserva la misura.

Svolgimento. Controlliamo ora che le definizioni siano ben poste. Per ciò che riguarda la misura, osserviamo che i cilindri elementari sono sempre disgiunti. L'unica cosa da controllare è se, qualora l'unione di cilindri elementari sia un cilindro elementare, allora la misura coincide. Visto che $\sqcup_{\ell=1}^n C_{\ell_1, \dots, \ell_k}^{j_1, \dots, j_k} = C_{\ell_1, \dots, \ell_{k-1}, \ell_{k+1}, \dots, \ell_h}^{j_1, \dots, j_{k-1}, j_{k+1}, \dots, j_h}$, la compatibilità delle misure diventa l'equazione $\sum_{\ell=1}^n p_{\ell_1} \cdots p_{\ell_k} = p_{\ell_1} \cdots p_{\ell_{k-1}} p_{\ell_{k+1}} \cdots p_{\ell_h}$, l'identità segue facilmente dal fatto che $p_1 + \cdots + p_n = 1$. È inoltre ovvio che la misura di M è normalizzata, visto che $\mu(B(p)) = \mu(\sqcup_{\ell=1}^n C_{\ell}^k) = \sum p_{\ell} = 1$.

È infine immediato controllare che lo shift Φ preserva la misura, infatti la preserva trivialmente sui cilindri elementari ed, essendo reversibile, commuta con unioni e complemento. □

2 Sistemi Ergodici

Un primo modo per studiare i sistemi dinamici misurabili è quello di dare delle caratterizzazioni che permettano di distinguere un sistema da un'altro. Una caratterizzazione importante è quella che va sotto il nome di **ergodicità**, essa è una proprietà che i sistemi dinamici possono o no verificare. Inoltre, i sistemi dinamici che sono ergodici permettono di studiare un problema semplificato che ha relazioni con il problema iniziale. La proprietà di ergodicità permette di sostituire un integrale temporale (che dipende dalla dinamica) con un integrale spaziale (che dipende solo dalla misura).

2.1 Medie temporali e teorema di Birkhoff-Kinchin

Sia dato uno spazio con misura (M, μ) , ipotizziamo che $\mu(M) = 1$, supponiamo data una funzione sommabile $f : M \rightarrow \mathbb{R}$. Si può quindi calcolare l'integrale $\int_M f d\mu$, che in questo contesto denoteremo con il simbolo $\langle f \rangle$, e lo chiameremo **media spaziale** o **media in fase** di f .

Se lo spazio (M, μ) è anche dotato di un sistema dinamico $\Phi : M \rightarrow M$ che preserva la misura allora, data una funzione sommabile, si può anche calcolare la sua **media temporale (nel futuro)**, che è una funzione di M in $\mathbb{R}$ che denoteremo con $\bar{f}$. Dato un punto $q \in M$ ed una funzione sommabile $f : M \rightarrow \mathbb{R}$, la media temporale di f è per definizione la funzione

$$\bar{f}(q) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(\Phi^s(q)) ds$$

ovvero, se il sistema dinamico è discreto

$$\bar{f}(q) = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{s=0}^N f(\Phi^s(q))$$

L'ergodicità di un sistema poggia sulla considerazione che un sistema evolve in modo "casuale" se visita in modo uniforme tutto lo spazio. Quindi se il sistema è ergodico, la funzione $q \mapsto \bar{f}(q)$ dovrebbe essere indipendente da q e dovrebbe sempre dare il valore medio di f in M , che è il numero $\langle f \rangle$.

Prima di poter rendere rigorosa una affermazione del genere, abbiamo l'onere di dimostrare che la definizione di media temporale è una buona definizione. Nello specifico: la funzione $q \mapsto \bar{f}(q)$ è sempre definita? La risposta è no, infatti

Esempio 27 Si consideri lo spazio M delle sequenze bi-infinite $\sigma = (\dots, E_{-1}, E_0, E_1, \dots)$ di lancio di una moneta, i possibili eventi sono quindi equiprobabili e del tipo $E_i = T$, testa o C , croce. La misura in M associa all'insieme in cui solo l'evento i -esimo è determinato il volume $1/2$. Il sistema dinamico reversibile è la traslazione a sinistra $\Phi(\sigma)_i = \sigma_{i+1}$. Si osservi che questo sistema è precisamente lo schema di Bernoulli $B(\frac{1}{2}, \frac{1}{2})$. In questo sistema, esiste una funzione $f : M \rightarrow \mathbb{R}$ ed un elemento $\sigma \in M$ tale che $\bar{f}(\sigma)$ non è definito.

Dimostrazione. Consideriamo la funzione sommabile $f(\sigma) = 1$ se $\sigma_0 = T$ ed $f(\sigma) = 0$ se $\sigma_0 = C$. Sia $\bar{\sigma}$ il punto di M

$$(\cdots, \sigma_0 = T, \sigma_1 = C, \sigma_2 = C, \cdots, \sigma_{2+\dots+2^{2m-1}+1} = T, \cdots, \sigma_{2+\dots+2^{2m}} = T, \sigma_{2+\dots+2^{2m}+1} = C, \cdots, \sigma_{2+\dots+2^{2m+1}} = C, \cdots)$$

Mostriamo che la successione che definisce $\bar{f}(\sigma)$ non ha limite. Infatti la somma parziale $\frac{1}{N} \sum_{s=0}^{1+\dots+2^{2m}} f(\Phi^s(\bar{\sigma}))$ porge il valore

$$\frac{1 + 2^2 + 2^4 + \cdots + 2^{2m-2} + 2^{2m}}{1 + 2 + 2^2 + \cdots + 2^{2m-1} + 2^{2m}} = \frac{1 + 4 + 4^2 + \cdots + 4^{m-1} + 4^m}{1 + 2 + 2^2 + \cdots + 2^{2m-1} + 2^{2m}}.$$

Il numeratore è serie geometrica di ragione 4 mentre il denominatore è serie geometrica di ragione 2, quindi il rapporto è

$$\frac{1 - 4^{m+1}}{1 - 4} / \frac{1 - 2^{2m+1}}{1 - 2} = \frac{1}{3} \frac{1 - 2^{2m+2}}{1 - 2^{2m+1}}.$$

Il termine $\frac{1-2^{2m+2}}{1-2^{2m+1}}$ è sempre maggiore di 2, quindi la somma parziale vale sempre più di 2/3.

D'altro canto, la somma parziale troncata alla fine del blocco successivo di C , $\frac{1}{N} \sum_{s=0}^{1+\dots+2^{2m+1}} f(\Phi^s(\bar{\sigma}))$, ha valore

$$\frac{1 + 2^2 + 2^4 + \cdots + 2^{2m-2} + 2^{2m}}{1 + 2 + 2^2 + \cdots + 2^{2m} + 2^{2m+1}} = \frac{1 + 4 + 4^2 + \cdots + 4^{m-1} + 4^m}{1 + 2 + 2^2 + \cdots + 2^{2m} + 2^{2m+1}}.$$

Il numeratore è una serie geometrica di ragione 4 mentre il denominatore è una serie geometrica di ragione 2, quindi il rapporto è

$$\frac{1 - 4^{m+1}}{1 - 4} / \frac{1 - 2^{2m+2}}{1 - 2} = \frac{1}{3}.$$

Segue che ci sono due troncate, una delle quali ha valore sempre 1/3, l'altra ha invece valore sempre maggiore di 2/3. Si deduce che la serie che definisce $\bar{f}(\bar{\sigma})$ non può convergere, quindi $\bar{f}(\bar{\sigma})$ non è definito. $\square$

Ci sono anche sistemi dinamici misurabili per i quali la media temporale è definita per ogni funzione misurabile f ed in ogni loro punto q .

Esempio 28 Sia dato lo spazio delle fasi S^1 con misura $d\vartheta$ e con sistema dinamico Φ la traslazione $\vartheta \mapsto \vartheta + 2\pi/p$. Data una qualsiasi funzione sommabile, la media temporale è definita per ogni $\vartheta \in S^1$.

Dimostrazione. Il sistema dinamico è dato dalla mappa

$$\Phi : S^1 \rightarrow S^1 \quad \vartheta \mapsto \vartheta + 2\pi \frac{q}{p}.$$

Ogni naturale $N \gg 0$ è del tipo $N = Mp + r$. Data una qualsiasi funzione f sommabile, si ha che

$$\begin{aligned} \bar{f}(\vartheta) &= \lim_{N \rightarrow \infty} \frac{f(\vartheta) + f(\vartheta + 2\pi \frac{1}{p}) + \cdots + f(\vartheta + 2\pi \frac{1}{p} N)}{N + 1} = \\ &= \lim_{n \rightarrow \infty} \frac{M}{Mp + r} \left(f(\vartheta) + \cdots + f(\vartheta + 2\pi \frac{1}{p} (p-1)) \right) + \frac{f(\vartheta) + \cdots + f(\vartheta + 2\pi \frac{1}{p} r)}{Mp + r} \end{aligned}$$

Passando al limite si ottiene che $\bar{f}(\vartheta) = \frac{f(\vartheta) + \cdots + f(\vartheta + 2\pi \frac{1}{p} (p-1))}{p}$. Quindi la funzione $\bar{f}$ ha le stesse caratteristiche di f (è liscia se f è liscia, continua se f è continua, misurabile se f è misurabile). $\square$

Si riesce a dare un senso alla definizione di media spaziale $q \mapsto \bar{f}(q)$? Ci viene in aiuto il teorema ergodico di Birkhoff-Kinchin

Teorema 29 (Birkhoff-Kinchin) Sia (M, μ, Φ) un sistema dinamico misurabile, e sia f una funzione sommabile. Allora per quasi ogni q in M esiste il limite $\bar{f}(q) = \lim_{N \rightarrow +\infty} \frac{1}{N+1} \sum_{s=0}^N f(\Phi^s(q))$ ed inoltre vale

$$\begin{aligned} \bar{f}(q) &= \bar{f}(\Phi(q)) \\ < \bar{f} > &= < f > . \end{aligned}$$

Se il sistema è reversibile, allora anche il limite

$$\tilde{f}(q) = \lim_{N \rightarrow +\infty} \frac{1}{N+1} \sum_{s=0}^N f(\Phi^{-s}(q))$$

esiste per q.o. q e coincide con $\bar{f}(q)$.

Per una dimostrazione di questo teorema si può consultare le note [B] di G. Benettin oppure il libro [G].

2.2 Definizione di sistema ergodico

Cominciamo quindi con il definire il concetto di ergodicità

Definizione 30 Un sistema dinamico misurabile e reversibile (M, μ, Φ) si dice **ergodico** se data una funzione sommabile $f : M \rightarrow \mathbb{R}$, la sua media temporale $q \mapsto \bar{f}(q)$ è costante quasi ovunque ed assume quasi ovunque il valore $\langle f \rangle$.

Per dare dei criteri equivalenti di ergodicità, dobbiamo introdurre il concetto di insieme invariante e di funzione invariante per un sistema dinamico misurabile (M, μ, Φ) .

Definizione 31 Sia dato un sistema dinamico misurabile (M, μ, Φ) . Un insieme $A \subset M$ misurabile si dice **invariante** se $\Phi^{-1}(A) \Delta A$ ha misura nulla² (od equivalentemente, se il sistema è reversibile, quando $\Phi(A) \Delta A$ ha misura nulla). Una funzione $f : M \rightarrow \mathbb{R}$ misurabile si dice **invariante** se $f(\Phi(q)) = f(q)$ per quasi ogni q in M (od equivalentemente, quando il sistema è reversibile, se $f(\Phi^{-1}(q)) = f(q)$ per quasi ogni q in M).

Vi sono altre proprietà caratterizzanti l'ergodicità ed altrettanto utili

Proposizione 32 Sia dato un sistema dinamico misurabile e reversibile (M, μ, Φ) , le seguenti proprietà sono equivalenti:

- (E1) il sistema è ergodico,
- (E2) ogni sottoinsieme A misurabile ed invariante ha misura 0 od 1,
- (E3) ogni funzione f misurabile ed invariante è costante q.o.

Dimostrazione. Cominciamo con il mostrare che E1 implica E2. Sia dato un insieme A misurabile ed invariante, la funzione caratteristica χ_A è sommabile. La condizione (E1) implica che la media temporale di χ_A è quasi ovunque uguale a $\mu(A)$. Essendo A invariante, la media temporale di χ_A è quasi ovunque uguale a χ_A , ne segue che χ_A è quasi ovunque uguale alla funzione costante $\mu(A)$. Questo può succedere solo in due casi: o $\mu(A) = 0$ e $\chi_A = 0$ quasi ovunque, oppure $\mu(A) = 1$ e $\chi_A = 1$ quasi ovunque.

Dimostriamo ora che E2 implica E3. Sia f una funzione misurabile ed invariante, e sia $M_e = \{x | f(x) < e\}$. L'invarianza di f implica che questi insiemi sono insiemi invarianti. Allora, per E2, gli insiemi M_e sono contenuti uno nell'altro, ed hanno misura 0 od 1. Sia $\bar{e} = \sup\{e | \mu(M_e) = 0\} = \inf\{e | \mu(M_e) = 1\}$, si ha che $f = \bar{e}$ quasi ovunque.

Dimostriamo infine che E3 implica E1. Presa una funzione sommabile f , consideriamo la funzione $g : q \mapsto \bar{f}(q) - \langle f \rangle$. Per il teorema di Birkhoff-Kinchin, la funzione g è sommabile ed invariante, quindi per E3 $g(q) = c$ con c costante per quasi ogni q . Sempre per il teorema di Birkhoff-Kinchin si ha che $\langle g \rangle = \langle \bar{f} \rangle - \langle f \rangle = 0$. D'altro canto $\langle g \rangle = c$, da cui segue che $g(q) = 0$ per quasi ogni q . Segue che $\bar{f} = \langle f \rangle$ quasi ovunque. $\square$

Osservazione 33 La definizione di ergodicità potrebbe anche essere data in altri modi,

- (E4) per ogni insieme misurabile A , la media temporale della funzione caratteristica χ_A vale q.o. $\mu(A)$
- (E5) per ogni coppia di insiemi misurabili A, B , vale

$$\lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{s=0}^N \mu(A \cap \Phi^{-s}B) = \mu(A)\mu(B)$$

Dimostrazione. che E1 implichi E4 è ovvio, mentre E4 implica E1 con argomenti di densità. È inoltre evidente che E4 implica E5, infatti quasi ovunque

$$\frac{1}{N+1} \sum_{s=0}^N \mu(A \cap \Phi^{-s}B) = \frac{1}{N+1} \left\langle \sum_{s=0}^N \chi_A \chi_{\Phi^{-s}B} \right\rangle = \int_{q \in M} \left[\chi_A(q) \frac{1}{N+1} \sum_{s=0}^N \chi_B(\Phi^s q) \right] d\mu$$

ma il secondo termine converge alla media temporale di χ_B , che per E4 converge quasi ovunque alla funzione costante $\mu(B)$, quindi il limite della funzione integranda tende quasi ovunque a $\mu(A)\mu(B)$. Viceversa, mostriamo che E5 implica E2. Sia U un insieme misurabile ed invariante, segue che per ogni A misurabile

$$\frac{1}{N+1} \sum_{s=0}^N \mu(A \cap \Phi^{-s}U) \equiv \mu(A \cap U),$$

²In questa formula $A \Delta B = (A - B) \cup (B - A)$.

ma per E5 converge a $\mu(A)\mu(U)$. Scelto $A = U^c$ si ha che $0 = \mu(\emptyset) = \mu(U)\mu(U^c)$ e quindi o $\mu(U) = 0$ o $\mu(U) = 1$. $\square$

2.3 Esempi di sistemi Ergodici e non

Esempio 34 La rotazione di S^1 in sè di angolo α

$$\Phi : S^1 \rightarrow S^1 \quad \vartheta \mapsto \vartheta + \alpha$$

è ergodica se e solo se $\alpha/2\pi$ è irrazionale

Dimostrazione. Se $\alpha/2\pi$ è razionale, allora $\alpha = 2\pi p/q$. Segue che la dinamica è periodica con orbite di periodo q . Sia $0, 2\pi/q, \dots, (q-1)2\pi/q$ l'orbita di 0 (si osservi che quello presentato non è necessariamente l'ordine in cui la dinamica visita i punti dell'orbita), allora l'insieme

$$\bigcup_{k=0}^{q-1} \left[\frac{2\pi}{q}k - \varepsilon, \frac{2\pi}{q}k + \varepsilon \right]$$

con $0 < \varepsilon < \pi/q$ è un insieme invariante, ed ha misura $0 < 2\varepsilon q < 2\pi$ (in questo esercizio, la misura di S^1 non è normalizzata ad 1). Quindi il sistema non è ergodico visto che non vale E2.

Se invece $\alpha/2\pi$ non è razionale, sia A un insieme misurabile ed invariante, e sia $f = \chi_A : M \rightarrow \mathbb{R}$ la funzione caratteristica di A . La funzione f appartiene, in particolare, ad L_2 , il quale è uno spazio di Hilbert con base Hilbertiana data dalle funzioni $e^{ik\vartheta}$, $k \in \mathbb{N}$, inoltre, la funzione f è quasi ovunque uguale a

$$\sum \hat{f}_k e^{ik\vartheta}$$

dove $\hat{f}_k = \int_{S^1} f(\vartheta) e^{-ik\vartheta}$ è il coefficiente di Fourier di indice k . Applicando lo shift si ha che la funzione $f(\Phi(\vartheta))$ è quasi ovunque uguale a

$$\sum \hat{f}_k e^{ik\alpha} e^{ik\vartheta}.$$

Dall'invarianza di A segue che le due funzioni $\vartheta \mapsto f(\vartheta)$ e $\vartheta \mapsto f(\Phi(\vartheta))$ sono uguali quasi ovunque, ma questo vale se e solo se $\hat{f}_k(1 - e^{ik\alpha}) = 0$ per ogni k . Se α è del tipo $2\pi r$ con r irrazionale, il secondo termine non è mai zero ad eccezione del caso $k = 0$. Quindi $\hat{f}_k$ deve essere zero per ogni $k \neq 0$, mentre nessuna condizione si impone a $\hat{f}_0$.

Per essere invariante, deve essere che $f = \hat{f}_0$ quasi ovunque, il che vuol dire che A ha misura 0 od 1. $\square$

Esempio 35 Ogni sistema Hamiltoniano 1D non è ergodico nello spazio $\mathbb{R}^2$ ma è ergodico nelle componenti connesse dei livelli regolari $M_h = \{H = h\}$

Dimostrazione. Il sistema è transitivo (l'orbita è tutto lo spazio) sulle componenti connesse di M_h , le quali sono curve immerse in $\mathbb{R}^2$. Segue che tali componenti connesse sono gli unici insiemi invarianti. $\square$

Esempio 36 Un oscillatore armonico disaccoppiato $H = \sum_i H_i$ con $H_i = \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)$ non è ergodico nello spazio e nemmeno nelle componenti connesse dei livelli $H = c$. Lo è nei tori $H_i = c_1$ a patto che le ω_i siano massimalmente irrazionali

Dimostrazione. Nelle ipotesi, gli insiemi di livello $M_h = \{H_i = h_i\}$ anche in questo caso gli unici insiemi invarianti. $\square$

Esempio 37 La trasformazione di T^2 in T^2 detta "gatto di Arnol'd" è ergodica.

Dimostrazione. La matrice di questa trasformazione è $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Gli autovalori di A sono $(3 \pm \sqrt{5})/2$. Un autovalore è maggiore di 1, mentre l'altro è in modulo minore di 1. I rispettivi autovettori sono $\begin{pmatrix} 1 \mp \sqrt{5} \\ -2 \end{pmatrix}$, che sono rette a pendenza irrazionale.

Queste due rette sono rispettivamente la varietà instabile e quella stabile dell'origine, che è punto fisso del sistema. Queste due varietà riempiono in modo denso lo spazio delle fasi T^2 . Se ne deduce che qualsiasi integrale del moto $f : T^2 \rightarrow \mathbb{R}$ continuo deve essere costantemente uguale a $f(0)$. Si conclude per densità. $\square$

Esempio 38 La trasformazione del panettiere è un sistema dinamico con misura invariante il cui spazio delle fasi è $\mathcal{P} = [0, 1) \times [0, 1)$, la misura è quella di Lebesgue. Questo spazio è dotato di uno shift dato da

$$\Phi: (x, y) \mapsto \begin{cases} (2x, \frac{1}{2}y) & \text{se } x < 1/2, \\ (2x - 1, \frac{1}{2}y + \frac{1}{2}) & \text{se } x \geq 1/2 \end{cases}$$

Questo sistema è ergodico.

Dimostrazione. La trasformazione del panettiere non è una mappa continua, ma è reversibile. In questo caso il punto $(0, 0)$ è punto fisso, e le sue varietà instabile consiste nelle rette a coordinate y razionale con denominatore potenza di 2. Tale reticolo è denso nello spazio, e ciò permette di concludere allo stesso modo che per l'esempio precedente. $\square$

Esempio 39 Gli schemi di Bernoulli sono ergodici.

Dimostreremo più avanti che gli schemi di Bernoulli godono di una proprietà che è più forte dell'ergodicità.

2.4 Sistemi iperbolici ed ergodicità

Supponiamo dato un sistema dinamico liscio (M, Φ) , dove M è una varietà liscia e Φ è un diffeomorfismo di M in sé. La nozione di punto iperbolico si può estendere a sottoinsiemi Λ di M detti **insiemi iperbolici**.

Definizione 40 Un insieme Λ si dice **iperbolico** se è compatto, Φ -invariante e se è tale che in ogni suo punto $q \in \Lambda$ esiste una decomposizione del tangente in sottospazi $E^s(q)$, $E^u(q)$ che variano in modo liscio con q e tali che:

$$(I1) \quad E^s(\Phi^n(q)) \oplus E^u(\Phi^n(q)) = T_{\Phi^n(q)}M,$$

$$(I2) \quad E^s(\Phi^n(q)) = d_q \Phi^n E^s(q), \text{ e così anche per } E^u,$$

$$(I3) \quad \text{esistono un } 0 < \lambda < 1 \text{ ed una costante } c \text{ tali che per ogni } \xi \text{ in } E^s(q) \text{ vale } \|d_q \Phi^n \xi\| \leq c\lambda^n \|\xi\|, \text{ e per ogni } \xi \text{ in } E^u(q) \text{ vale } \|d_q \Phi^{-n} \xi\| \leq c\lambda^n \|\xi\|.$$

Cominciamo con il dare esempi di insiemi iperbolici.

Esempio 41 Dato un equilibrio iperbolico q^* ed una sua intersezione omoclina q_0 , l'insieme $\Lambda = q^* \cup \left(\bigcup_{n \in \mathbb{Z}} \Phi^n(q_0) \right)$ è un insieme iperbolico.

Dimostrazione. Le varietà stabile ed instabile definiscono, con i loro tangenti, gli spazi E^s , E^u con le proprietà richieste. $\square$

Gli insiemi iperbolici contenuti nello spazio delle fasi sono insiemi invarianti che caratterizzano (parte della) dinamica, si pensi alla *horseshoe di Smale*. Vi sono sistemi fatti solo di punti iperbolici, tali sistemi si dicono **sistemi di Anosov**.

Esempio 42 Lo shift del gatto di Arnol'd è un sistema di Anosov. Così pure la mappa del panettiere (la mappa non è però liscia).

Dimostrazione. Dato un sistema di Anosov, per ogni punto q della varietà M passano due varietà immerse, dette varietà stabile $W^s(q)$ e varietà instabile $W^u(q)$. Tali varietà sono invarianti per la dinamica Φ , nel senso che $\Phi(W^*(q)) = W^*(\Phi(q))$ e su di esse la dinamica contrae rispettivamente dilata le distanze, nel senso che se $q' \in W^*(q)$, allora

$$\|\Phi^n(q') - \Phi^n(q)\| \leq C\lambda^n \|q' - q\|$$

$\square$

Proposizione 43 Ogni sistema di Anosov reversibile, con $\dim E^s \neq 0, n$ e con misura invariante è ergodico.

Dimostrazione. (Argomento di Hopf) Per ogni punto $\bar{q}$ di un dato insieme iperbolico, si possono definire varietà $W^s(\bar{q})$ e $W^u(\bar{q})$ passanti per $\bar{q}$ e tangenti a $E^s(\bar{q})$ e $E^u(\bar{q})$ rispettivamente. Tali varietà sono definite come

$$W^s(\bar{q}) = \{q \mid \lim_n \|\Phi^n(q) - \Phi^n(\bar{q})\| = 0\}.$$

Mostriamo come la media temporale $\bar{f}$ di qualsiasi funzione continua f coincida q.o. con la media in fase.

Presi due punti abbastanza vicini q, q' , la varietà stabile di q deve intersecare la varietà instabile di q' in un punto $\tilde{q}$. Quindi, $\Phi^n(\tilde{q})$ tende allo stesso punto a cui tende $\Phi^n(q)$, e quindi la media temporale $\bar{f}(q)$ ha lo stesso valore di $\bar{f}(\tilde{q})$, diciamo c .

Similmente, usando la media nel passato, $c = \bar{f}(\tilde{q}) = \bar{f}(q')$. Con un argomento di aperto/chiuso, segue che dati due punti vicini la media temporale di f coincide. Quindi $\bar{f}(q) = c$ quasi ovunque. Dal momento che $\langle f \rangle = \langle \bar{f} \rangle = c$, si conclude che la media temporale di f coincide con la sua media in fase. $\square$

Abbiamo quindi un altro modo per dimostrare che il sistema dinamico del *gatto di Arnol'd* e la trasformazione del panettiere sono sistemi dinamici ergodici.

2.5 Misure ergodiche

Proviamo ora a rovesciare la domanda. Sia dato un sistema dinamico reversibile (M, Φ) , ci sono delle misure μ che siano Φ -invarianti e che rendono ergodico il sistema misurabile reversibile (M, μ, Φ) ? Come è fatto lo spazio di tali misure?

Facciamo una ipotesi semplificatrice, che ci permette di confrontare le misure: supponiamo la σ -algebra su cui è definita la misura fissata una volta per tutte. Ricordiamo che

Definizione 44 *Data una misura μ , una seconda misura ν definita sulla stessa σ -algebra si dice*

- μ -**regolare** ovvero **assolutamente continua** rispetto a μ se per ogni A misurabile tale che $\mu(A) = 0$, allora $\nu(A) = 0$.
- μ -**singolare** se esistono U, V misurabili tali che

$$\mu(U) = 1, \mu(V) = 0, \nu(U) = 0, \nu(V) = 1.$$

Osservazione 45 *Date due misure μ, ν definite sulla stessa σ -algebra, se ν è μ -regolare rispetto a μ , allora esiste $f: M \rightarrow \mathbb{R}$ sommabile tale che $\nu = f\mu$.*

Questo è il celebrato teorema di Radon-Nycodim, la cui dimostrazione è parte dei corsi di teoria della misura. lo citiamo soprattutto per far osservare che, nel caso di varietà lisce ed immerse dotate di forme volume τ , si è visto come la misura associata a τ sia regolare rispetto alla misura di Lebesgue.

Supponiamo quindi dato un sistema ergodico (M, μ, Φ) . Intendiamo dedicarci alla ricerca delle misure ν invarianti e μ -regolare. La prossima proposizione dimostra che la ricerca non sarà fruttuosa. Infatti

Proposizione 46 *Dato un sistema ergodico (M, μ, Φ) ed una misura ergodica ν tale che $\nu(M) = 1$, allora*

- $\nu = \mu$, oppure
- ν è μ -singolare

Dimostrazione. Supponiamo che $\mu \neq \nu$, allora esiste una funzione sommabile f tale che $\langle f \rangle_\mu \neq \langle f \rangle_\nu$. Siccome la misura μ è ergodica, allora esiste un insieme misurabile ed invariante U di μ -misura piena tale che $\bar{f}|_U \equiv \langle f \rangle_\mu$. D'altro canto anche ν è ergodica, quindi esiste un insieme misurabile ed invariante V di ν -misura piena tale che $\bar{f}|_V \equiv \langle f \rangle_\nu$.

Ma necessariamente U e V sono disgiunti, quindi necessariamente vale $\mu(V) = 0$ e $\nu(U) = 0$. $\square$

Quindi due misure ergodiche diverse sono supportate in insiemi disgiunti. Se μ, ν sono misure ergodiche, allora una loro combinazione convessa $\eta_\lambda = \lambda\mu + (1 - \lambda)\nu$ è ancora una misura invariante e normalizzata, ma non è ergodica. Infatti i due insiemi U, V sono insiemi invarianti per η_λ ed hanno misura non nulla. D'altro canto l'insieme di tutte le misure invarianti è ovviamente convesso.

Un fatto che mostreremo essere ragionevole alla fine delle note, è che le misure ergodiche per un dato sistema dinamico (M, Φ) sono il bordo dell'insieme delle misure invarianti, nel senso che se $\mathcal{M}$ è lo spazio delle misure ergodiche, allora lo spazio delle misure invarianti è l'involuppo convesso di $\mathcal{M}$.

Esempio 47 Sia M lo spazio delle sequenze bi-infinite sui simboli T, C , dotato del sistema dinamico Φ della traslazione a sinistra. Lo spazio M ha una famiglia 1-parametrica di misure ergodiche, determiniamo i loro supporti disgiunti.

Svolgimento. È immediato osservare che lo spazio M è isomorfo agli schemi di Bernoulli $B(p, 1-p)$ per ogni scelta di p . Sappiamo che gli schemi di Bernoulli sono tutti ergodici. La scelta di p definisce una famiglia 1-parametrica di misure ergodiche μ_p in M . Questo vuol dire che esistono insiemi V_p di μ_p -misura piena e di μ_q -misura nulla per ogni $q \neq p$.

Consideriamo la funzione $f = \chi_{C_T^0}$. Per ogni $p \in (0, 1)$ fissato il teorema di Birkhoff-Kinchin dice che $\bar{f}$, la media temporale di f , è costante quasi ovunque e quasi ovunque uguale a $\mu(C_T^0) = p$. Sia quindi $V_p = \{\sigma \in \{0, 1\}^{\mathbb{Z}} | \bar{f}(\sigma) = p\}$, si ha che $\mu_p(V_p) = 1$ e che $V_p \cap V_{p'} = \emptyset$, e quindi $\mu_p(V_{p'}) = 0$ per ogni $p' \neq p$.

Gli insiemi V_p sono quindi gli insiemi che supportano la misura μ_p e dimostrano che le varie misure μ_p sono singolari una rispetto all'altra. $\square$

3 Sistemi Mescolanti

Un'altra proprietà importante nei sistemi dinamici misurabili è quella di essere mescolanti. Come dice la parola, questa caratterizzazione si rifà al comportamento tipico delle soluzioni: supponiamo di avere un bicchiere riempito per 99 parti di acqua e per 1 parte di inchiostro, supponiamo inoltre che, in un istante iniziale, l'inchiostro venga versato nell'acqua, e quindi che ci sia una regione A ben definita del bicchiere (di volume $1/100$) occupata da inchiostro puro.

Lasciamo passare il tempo ed aspettiamo che l'inchiostro si sia diluito. Prendiamo ora in considerazione la parte del liquido che si trova nella regione B del bicchiere, che chiameremo finestra di osservazione. Una domanda ragionevole da fare è, dopo un tempo N abbastanza lungo, quanta parte del liquido nella finestra di osservazione è inchiostro e quanta parte è acqua. Se le molecole di inchiostro ed acqua si muovono in accordo con una legge dinamica Φ , la domanda appena posta è quanto vale $\mu(\Phi^N(A) \cap B)$, od equivalentemente, quanto vale $\mu(A \cap \Phi^{-N}(B))$?

Dire che un sistema è mescolante vuol dire che, una volta passato abbastanza tempo, le molecole di inchiostro si sono uniformemente distribuite, ovvero quando N è abbastanza grande, allora $\mu(A \cap \Phi^{-N}(B)) \simeq \mu(A)\mu(B)$. La formulazione matematica di "tempo abbastanza lungo" richiede un passaggio al limite,

$$\lim_{N \rightarrow \infty} \mu(A \cap \Phi^{-N}(B)) = \mu(A)\mu(B).$$

Definizione 48 Sia (M, μ, Φ) un sistema dinamico misurabile reversibile. Il sistema si dice **mescolante** se per ogni coppia di insiemi misurabili A e B vale

$$(M1) \quad \lim_{N \rightarrow \infty} \mu(A \cap \Phi^{-N}(B)) = \mu(A)\mu(B)$$

Anche in questo caso, la caratterizzazione di sistema mescolante può essere data in più di un modo.

Proposizione 49 Un sistema dinamico misurabile (M, μ, Φ) è mescolante se e solo se per ogni coppia di funzioni sommabili f, g risulta

$$(M2) \quad \lim_N \int f(\Phi^N(q))g(q)d\mu = \langle f \rangle \langle g \rangle$$

Dimostrazione. Mostriamo che se il sistema è mescolante allora soddisfa alla condizione integrale M2.

Certamente, la formula integrale vale se f e g sono funzioni caratteristiche, infatti se $f = \chi_B$, allora $f \circ \Phi^N(q) = \chi_B(\Phi^N(q)) = \chi_{\Phi^{-N}(B)}(q)$, se inoltre $g = \chi_A$ allora $\lim_N \int f(\Phi^N(q))g(q)d\mu = \lim_N \mu(A \cap \Phi^{-N}(B))$, che per M1 è $\mu(A)\mu(B) = \langle f \rangle \langle g \rangle$. Segue immediatamente che l'identità vale per combinazioni finite a coefficienti reali di funzioni caratteristiche e quindi, per densità, vale per funzioni sommabili qualsiasi.

Mostriamo l'implicazione inversa, ovvero che se vale la condizione integrale allora il sistema è mescolante. Siano dati due insiemi misurabili A e B , consideriamo le rispettive funzioni caratteristiche χ_A e χ_B . Queste due funzioni sono sommabili, quindi l'ipotesi che valga l'identità integrale implica che

$$\lim_N \mu(A \cap \Phi^{-N}(B)) = \lim_N \int_M \chi_A(q)\chi_B(\Phi^N(q))d\mu = \langle \chi_A \rangle \langle \chi_B \rangle = \mu(A)\mu(B).$$

Questo conclude la dimostrazione. $\square$

Esiste una relazione tra il fatto che un sistema sia mescolante ed il fatto che un sistema sia ergodico?

Proposizione 50 *Ogni insieme mescolante è ergodico.*

Dimostrazione. Sia (M, μ, Φ) un sistema mescolante, mostriamo che vale la proprietà (E2) dei sistemi ergodici. Sia A misurabile ed invariante per Φ . Scegliendo $B = A$ si ha che

$$\mu(A)^2 = \lim_N \mu(A \cap \Phi^{-N}(A)) = \mu(A),$$

quindi A deve avere misura 1 o 0. $\square$

Diamo ora un po' di esempi di sistemi mescolanti

Esempio 51 *Il gatto di Arnol'd è mescolante.*

Dimostrazione. Dimostriamo che per questo sistema vale M1. Prendiamo due insiemi misurabili A e B , e consideriamo $f = \chi_A$, $g = \chi_B$. Siccome $\mu(M) = 1$ si ha che $f, g \in L_2$. Sviluppando le funzioni in serie di Fourier si ha che

$$f = \sum \hat{f}_k e^{2\pi i k x}, \quad g = \sum \hat{g}_k e^{2\pi i k x},$$

dove ovviamente vale $\hat{f}_0 = \langle f \rangle = \mu(A)$, ed $\hat{g}_0 = \langle g \rangle = \mu(B)$. Chiamando X la matrice della mappa del gatto di Arnol'd si ha che $f \circ \Phi^N(q) = \sum \hat{f}_k e^{2\pi i k X^N q} = \sum \hat{f}_k e^{2\pi i (X^t)^N k \cdot q}$. Per ogni vettore k , il vettore intero $(X^t)^N k$ tende in norma all'infinito. Tronchiamo all'infrarosso le funzioni, ovvero approssimiamo f con $f^{\leq n} = \sum_{|k| \leq n} \hat{f}_k e^{2\pi i k x}$ e g con $g^{\leq n} = \sum_{|k| \leq n} \hat{g}_k e^{2\pi i k x}$. Calcoliamo

$$\int f^{\leq n} \circ \Phi^N(q) g^{\leq n}(q) = \sum_{|k|, |\ell| \leq n} \hat{f}_k \hat{g}_\ell \int e^{2\pi i (X^t)^N k \cdot q} e^{2\pi i \ell \cdot q}.$$

Con l'eccezione del termine $k = \ell = 0$, ogni termine di questa somma è nullo, perchè l'ultimo integrale è il prodotto scalare $(e^{2\pi i (X^t)^N k \cdot q}, e^{2\pi i \ell \cdot q})$ che è nullo per $N \gg 0$. Questo porge l'identità $\langle f \rangle \langle g \rangle = \mu(A) \mu(B)$. Quanto abbiamo trascurato? il resto è infinitesimo, infatti

$$\int f \circ \Phi^N(q) g(q) d\mu = \int f^{\leq n} \circ \Phi^N(q) g^{\leq n}(q) + \int f \circ \Phi^N(q) g^{>n}(q) d\mu + \int f^{>n} \circ \Phi^N(q) g^{>n}(q) d\mu,$$

dove chiaramente $f^{>n} = \sum_{|k| > n} \hat{f}_k e^{2\pi i k \cdot q}$ e similmente per $g^{>n}$. Gli ultimi due termini si stimano facilmente come minore di ε a patto che si scelga $n \gg 0$. Basta infatti usare il fatto che in spazi di Hilbert vale $|(u, v)| \leq \|u\| \|v\|$, e che $g^{>n}$ è la coda di una serie convergente. $\square$

Esempio 52 *Il sistema dinamico del Panettiere è mescolante.*

Svolgimento. Un punto $(x, y) \in [0, 1] \times [0, 1]$ di $\mathcal{P}$ si può scrivere espandendo in forma binaria i due numeri x, y . Siano

$$x = 0, a_1 a_2 \dots \quad y = 0, b_1 b_2 \dots$$

con $a_i, b_i = 0, 1$. Qualsiasi tipo di espansione ha ambiguità di scrittura, in questo caso l'ambiguità consiste nel fatto che $\dots a_i 0 \bar{1} = \dots a_i 1$. Per convenzione, decidiamo che nell'espansione dei numeri x, y non ci siano sequenze infinite di 1.

Lo spazio $\mathcal{P}$ si dimostra così isomorfo allo spazio delle successioni bi-infinite con entrate 0, 1.

$$\sigma = (\dots, b_1, a_1, \dots) \in \{0, 1\}^{\mathbb{Z}}.$$

La trasformazione del panettiere Φ , in questa rappresentazione dello spazio, diventa lo *shift a sinistra*, ovvero la mappa che manda una successione bi-infinita σ nella successione $\Phi(\sigma)$ tale che $\Phi(\sigma)_i = \sigma_{i+1}$.

La misura di Lebesgue si trasporta nello spazio delle successioni bi-infinite. Cerchiamo di descrivere gli insiemi misurabili e la misura in tale spazio. Un misurabile di $[0, 1] \times [0, 1]$ è unione arbitraria di plurirettangoli. Preso un punto $\bar{x}$ in $[0, 1]$ con $\bar{x} = 0.\bar{a}_1 \dots$, l'insieme degli x tali che

$$x = 0.\bar{a}_1 \dots \bar{a}_n * \dots$$

è un intervallo di lunghezza $1/2^{-n}$ che contiene $\bar{x}$. Chiamiamo $R_{\bar{a}_1, \dots, \bar{a}_n}^{1, \dots, n}$ questo intervallo. A partire da questa osservazione, si arriva a capire che $R_{\bar{a}_{i_1}, \dots, \bar{a}_{i_n}}^{j_1, \dots, j_n} = \{x | x = 0 \dots \bar{a}_{i_1} \dots \bar{a}_{i_n} \dots\}$ è una unione di intervalli, tutti di lunghezza $1/2^{-j_n}$, con $j_n \geq n$, uno dei quali contenente $\bar{x}$.

Ricordando ora che la mappa del panettiere è un sistema con due variabili x, y in $[0, 1]$, si definisce il

$$R_{a_{i-m}, \dots, \bar{a}_{i-1}, \bar{a}_0, \bar{a}_{i_1}, \dots, \bar{a}_{i_n}}^{-j-m, \dots, -j-1, j_0, j_1, \dots, j_n} = \{(x, y) | x = 0 \dots \bar{a}_{i_0} \dots \bar{a}_{i_n} \dots, y = 0 \dots \bar{a}_{i-1} \dots \bar{a}_{i-m} \dots\},$$

il plurirettangolo un unione di rettangoli di larghezza $1/2^{n+1}$ e di altezza $1/2^m$.

Abbiamo quindi dimostrato che una famiglia di plurirettangoli che genera la σ -algebra dei Boreliani è in corrispondenza con i cilindri elementari di uno schema di Bernoulli $B(\frac{1}{2}, \frac{1}{2})$. Oltretutto, la misura di Lebesgue di tali plurirettangoli è precisamente la misura associata ai rispettivi cilindri elementari di tale schema di Bernoulli. In realtà, con pochi controlli, si può dimostrare che il sistema dinamico del panettiere è isomorfo a tale schema di Bernoulli.

Dimostriamo ora che il sistema del panettiere è mescolante. Siano $A = C_{\ell_1, \dots, \ell_h}^{j_1, \dots, j_h}$ e $B = C_{\ell'_1, \dots, \ell'_k}^{j'_1, \dots, j'_k}$ dei cilindri elementari. Per definizione di shift,

$$\Phi^{-N}(A) = \Phi^{-N}(C_{\ell_1, \dots, \ell_h}^{j_1, \dots, j_h}) = C_{\ell_1, \dots, \ell_h}^{j_1+N, \dots, j_h+N}.$$

Scegliendo quindi un N tale che $j_1 + N$ sia maggiore di j'_k , si ha che l'intersezione $\Phi^{-N}(A) \cap B$ è proprio il cilindro elementare $C_{\ell'_1, \dots, \ell'_k, \ell_1, \dots, \ell_h}^{j'_1, \dots, j'_k, j_1+N, \dots, j_h+N}$, la cui misura è $\frac{1}{2^{h+k}} = \frac{1}{2^h} \frac{1}{2^k}$, il prodotto delle due misure. Si conclude quindi che per i cilindri elementari non è necessario far tendere n all'infinito. Dati invece due insiemi misurabili A, B , per ogni ε esistono due unioni disgiunte di cilindri elementari C_A contenuta in A ed C_B contenuta in B tali che $\mu(A \setminus C_A) < \varepsilon$, idem per B . La addittività della misura, ed il risultato ottenuto con i cilindri elementari, permette di ottenere il risultato. $\square$

Esempio 53 *Gli schemi di Bernoulli sono mescolanti*

Dimostrazione. L'argomento per dimostrare questo fatto è identico all'argomento usato nell'esempio precedente. $\square$

Corollario immediato di questo esempio è che gli schemi di Bernoulli sono ergodici.

4 Entropia

Abbiamo finora definito e studiato l'ergodicità e la mescolanza, due proprietà che un sistema misurabile può avere o può non avere. Queste proprietà permettono una (rozza) discriminazione tra sistemi dinamici misurabili diversi. In questo Capitolo introduciamo un invariante, l'entropia, che a differenza di ergodicità e mescolanza — le quali hanno valore nello spazio *si o no* — è una quantità che si associa ad un dato sistema dinamico misurabile ed ha valori nei reali non-negativi. È quindi una quantità caratteristica che discrimina in modo più fine i sistemi dinamici misurabili.

4.1 L'entropia delle partizioni

Lo spazio ambiente è, anche in questo caso, un sistema dinamico misurabile reversibile (M, μ, Φ) con misura normalizzata $\mu(M) = 1$.

Definizione 54 Lo spazio delle **partizioni misurabili finite**, $\mathcal{P}(M)$, è un insieme di classi di equivalenza di partizioni. Diciamo che due partizioni $\alpha, \beta \in \mathcal{P}(M)$ sono da considerare uguali se per ogni A in α esiste un B in β tale che $A \Delta B$ è un insieme di misura nulla.

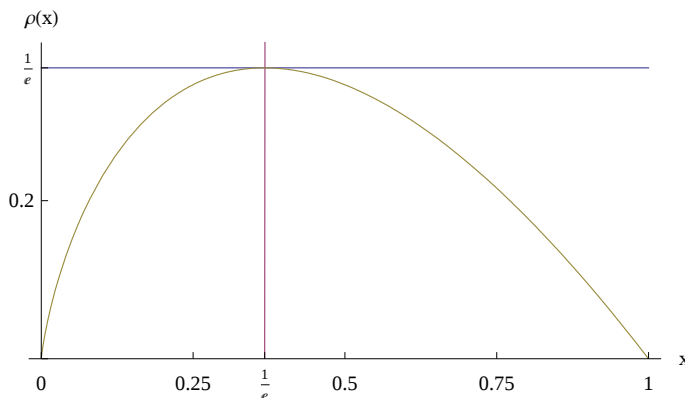
Non appesantiremo le notazioni indicando una partizione misurabile con il simbolo $[\alpha]$, ma raccomandiamo di tener presente che una partizione misurabile è una classe di equivalenza. Data una partizione misurabile α esiste un rappresentante con numero minimo di elementi, che è una qualsiasi partizione della classe di equivalenza α i cui elementi hanno tutti misura non nulla. Indichiamo con $\mathcal{P}_n(M)$ le partizioni con n elementi. Distilliamo ora dalla partizione una informazione numerica.

Definizione 55 Sia $\eta : \mathcal{P}(M) \rightarrow \mathbb{R}^{\geq}$ la funzione che, ad una partizione α associa il reale non-negativo

$$\eta(\alpha) = - \sum_{A \in \alpha} \mu(A) \log(\mu(A)).$$

Il valore $\eta(\alpha)$ si chiama **entropia della partizione** α .

Cerchiamo di capire questa definizione: la funzione $\rho : [0, 1] \rightarrow \mathbb{R}$, $x \mapsto -x \log x$, estesa per continuità in 0 ponendo $\rho(0) = 0$, ha grafico



è positiva, vale 0 agli estremi 0,1 dell'intervallo in cui è definita, ed ha massimo nel punto $1/e$, su cui assume il valore $1/e$. La funzione è strettamente concava³, e questo ha degli effetti sulla funzione $\eta = \sum_{A \in \alpha} \rho(A)$.

È evidente che se due partizioni α ed α' sono equivalenti allora l'entropia ad esse associata è lo stesso numero. Per questo motivo alcune frasi si dicono in modo più conciso se il dominio della funzione η è l'insieme delle classi di equivalenza. Due proprietà importanti della funzione entropia sono le seguenti.

Proposizione 56 Sia α una partizione misurabile, allora

³Una funzione è concava se e solo se l'ipografico è un insieme convesso, ovvero se e solo se vale la disuguaglianza $\rho(\lambda x + (1-\lambda)y) \geq \lambda \rho(x) + (1-\lambda)\rho(y)$ per ogni $\lambda \in [0, 1]$. La stretta concavità porge una disuguaglianza stretta quando $\lambda \in]0, 1[$ a meno che $x = y$.

Segue per induzione che $\rho(\sum_{i=1}^n \lambda_i p_i) \geq \lambda_1 \rho(p_1) + \dots + \lambda_n \rho(p_n)$ per ogni p_i e $\lambda_i \in]0, 1[$ tali che $\lambda_1 + \dots + \lambda_n = 1$ con disuguaglianza stretta a meno che i p_i siano tutti uguali.

($\eta 1$) $\eta(\alpha) = 0$ se e solo se $\alpha = \{M\}$.

($\eta 2$) In $\mathcal{P}_n(M)$, l'entropia è massima sulle (e solo sulle) partizioni misurabili α tali che $\mu(A) = 1/n$ per ogni $A \in \alpha$.

Dimostrazione. ($\eta 1$) Se $\alpha = \{M\}$ allora $\eta(\alpha) = -1 \log 1 = 0$. Se $\alpha = \{A_1, \dots, A_n\}$ con nessuno degli A_i di misura 1, allora esistono almeno due elementi A_1 ed A_2 in α tali che $0 < \mu(A_i) < 1$. Quindi, per questi due elementi vale $\rho(\mu(A_i)) > 0$. Segue che $\eta(\alpha) > 0$.

($\eta 2$) Se una partizione $\bar{\alpha}$ in $\mathcal{P}_n$ è composta da elementi di misura $1/n$, allora la sua entropia vale $\eta(\bar{\alpha}) = n\rho(1/n) = -n(1/n \log(1/n)) = \log n$. Dobbiamo mostrare che se $\alpha \in \mathcal{P}_n$ allora $\eta(\alpha) \leq \log n$, con disuguaglianza stretta a meno che $\mu(A) = 1/n$ per ogni A in α .

Usiamo il fatto che ρ è strettamente concava, e quindi $\rho(\sum_i \lambda_i p_i) \geq \sum_i \lambda_i \rho(p_i)$ per ogni scelta di $\lambda_i \in]0, 1[$ con disuguaglianza stretta a meno che tutti i p_i siano uguali tra loro e quindi uguali ad $1/n$. Applichiamo la disuguaglianza quando ogni $\lambda_i = 1/n$ e $p_i = \mu(A_i)$ al variare degli A_i in α . Si ottiene che

$$n\rho\left(\sum_i p_i/n\right) \geq \sum_i \rho(p_i) = \eta(\alpha),$$

con disuguaglianza stretta a meno che tutti i p_i non siano uguali tra loro, e quindi uguali ad $1/n$. Ma

$$n\rho\left(\sum_{i=1}^n \frac{p_i}{n}\right) = -n \left(\sum_{i=1}^n \frac{p_i}{n}\right) \log \left(\sum_{i=1}^n \frac{p_i}{n}\right) = \log n.$$

Si ottiene quindi la disuguaglianza desiderata, stretta se α non è composta di n atomi di misura $1/n$. □

Definiamo alcune proprietà ed operazioni sulle partizioni.

Definizione 57 La partizione β si dice un **raffinamento** di α , e si scrive $\beta \succcurlyeq \alpha$, se per ogni $B \in \beta$ esiste un $A \in \alpha$ tale che $\mu(B - A) = 0$. Date due partizioni, la partizione $\alpha \vee \beta = \{A \cap B | A \in \alpha, B \in \beta\}$ si dice **partizione generata** da α e β ed è il più piccolo raffinamento sia di α che di β .

Due partizioni α, β si dicono **indipendenti** se $\mu(A \cap B) = \mu(A)\mu(B)$ per ogni $A \in \alpha, B \in \beta$.

L'idea di indipendenza è fondamentale in teoria della probabilità, dove ad essere indipendenti sono delle variabili aleatorie. Una tipica coppia di partizioni indipendenti sono A la partizione di un quadrato in rettangoli verticali e B la partizione di un quadrato in rettangoli orizzontali. L'indipendenza di due partizioni significa che, il sapere che q appartiene ad un $A \in \alpha$ non ha nessuna influenza sulla probabilità che q appartenga ad uno dei $B \in \beta$.

Osservazione. È interessante notare che la mescolanza è una indipendenza asintotica tra i misurabili A ed $\Phi^{-N}(B)$, mentre l'ergodicità è una più debole indipendenza in media tra A e la famiglia $\{\Phi^{-N}(B) | n \in \mathbb{N}\}$.

Vediamo ora come la funzione entropia si comporta rispetto a queste operazioni.

Proposizione 58 Siano α, β due partizioni.

($\eta 3$) È sempre verificata la disuguaglianza $\eta(\alpha \vee \beta) \leq \eta(\alpha) + \eta(\beta)$. Inoltre vale $\eta(\alpha \vee \beta) = \eta(\alpha) + \eta(\beta)$ se e solo se α, β sono indipendenti.

($\eta 4$) Se β è un raffinamento di α allora $\eta(\beta) \geq \eta(\alpha)$. Inoltre, se β è un raffinamento stretto allora $\eta(\beta) > \eta(\alpha)$.

Dimostrazione. Dimostriamo ($\eta 3$) con una sequenza di disequazioni. Per la definizione di entropia

$$\eta(\alpha \vee \beta) = - \sum_{A \in \alpha, B \in \beta} \mu(A \cap B) \log \left(\mu(B) \frac{\mu(A \cap B)}{\mu(B)} \right) = \dots$$

usando le proprietà del logaritmo

$$\dots = - \sum_{A \in \alpha, B \in \beta} \mu(A \cap B) \left[\log(\mu(B)) + \log \left(\frac{\mu(A \cap B)}{\mu(B)} \right) \right]$$

distribuendo ed usando $\sum_{A \in \alpha} \mu(A \cap X) = \mu(X)$

$$\dots = - \sum_{B \in \beta} \mu(B) \log(\mu(B)) - \sum_{A \in \alpha, B \in \beta} \mu(A \cap B) \log \left(\frac{\mu(A \cap B)}{\mu(B)} \right) = \dots$$

tornando ad usare la definizione di entropia

$$\dots = \eta(\beta) - \sum_{A \in \alpha, B \in \beta} \mu(B) \frac{\mu(A \cap B)}{\mu(B)} \log \left(\frac{\mu(A \cap B)}{\mu(B)} \right)$$

Osserviamo che il secondo termine è del tipo $\sum_{A \in \alpha} \left(\sum_{B \in \beta} \mu(B) \rho \left(\frac{\mu(A \cap B)}{\mu(B)} \right) \right)$. Per la stretta concavità di ρ si ha che, fissato A misurabile, ogni somma sui $B \in \beta$ all'interno delle parentesi soddisfa a

$$\sum_{B \in \beta} \mu(B) \rho \left(\frac{\mu(A \cap B)}{\mu(B)} \right) \geq \rho \left(\sum_{B \in \beta} \mu(B) \frac{\mu(A \cap B)}{\mu(B)} \right) = \rho(\mu(A)).$$

Sommando infine su A si ottiene per definizione $\eta(\alpha)$. Oltretutto l'uguaglianza vale se e solo se $\mu(A \cap B)/\mu(B)$ è costante per ogni B , ma $\mu(A \cap B) \equiv c\mu(B)$ per ogni B implica che $\mu(A) = \sum_B \mu(A \cap B) = c \sum \mu(B) = c$, ovvero che $\mu(A \cap B) = \mu(A)\mu(B)$.

La proprietà ($\eta 4$) si dimostra raggruppando i termini ed usando la stretta concavità di ρ assieme al fatto che $\rho(0) = 0$. Infatti vale

$$\eta(\beta) = - \sum_{B \subset \beta} \mu(B) \log(\mu(B)) = - \sum_{A \in \alpha} \sum_{B \subset A} \mu(B) \log(\mu(B)) = \sum_{A \in \alpha} \sum_{B \subset A} \rho(\mu(B)).$$

Se mostriamo che $\sum_{B \subset A} \rho(\mu(B)) \geq \rho(A)$ abbiamo finito. Sia $p = \mu(A)$, allora per ogni $B \subset A$ si ha che $\mu(B) = \lambda_B p$ con $\lambda_B \in]0, 1[$. Per la stretta convessità si ha che

$$\rho(\lambda_B p) = \rho((1 - \lambda_B)0 + \lambda_B p) \leq (1 - \lambda_B)\rho(0) + \lambda_B \rho(p) = \lambda_B \rho(p).$$

Sommando sui $B \supset A$ si conclude. □

Le quattro proprietà $\eta 1, \dots, \eta 4$ dimostrate sopra caratterizzano la funzione entropia, infatti

Proposizione 59 *Una funzione di $\mathcal{P}(M)$ a valori reali che soddisfi ad $\eta 1$ – $\eta 4$ è un multiplo di η .*

Non diamo una dimostrazione di questo fatto. Osserviamo solo che la costante moltiplicativa corrisponde all'arbitraria scelta della base del logaritmo. Scelte tipiche sono la base 2 quando si lavora in teoria dell'informazione (in linea con il concetto di *bite*) e la base e quando si vuole evitare il tipico apparire di noiose costanti in alcune formule.

L'entropia associata ad una partizione è un modo per formalizzare la quantità di informazione legata ad una partizione. Si supponga che gli elementi A di α siano **finestre di osservazione**, e che una **misurazione** sia la registrazione di quale insieme $A \in \alpha$ contenga un evento $q \in M$. Se la partizione $\alpha = \{A_1, A_2\}$ con $\mu(A_i) = 1/2$, allora una misurazione porta una quantità di informazione $\log 2$ (che è 1 se si sceglie di usare il logaritmo in base 2). Due misurazioni corrispondono ad una coppia $(q_1, q_2) \in M \times M$ ricoperta dalla partizione $\alpha^2 = \{A_j \times A_k\}$, la cui entropia è $\log 2^2 = 2$, e così via. Se invece $\alpha = \{A_1, A_2\}$ con $\mu(A_1) \cong 1$, allora l'informazione che ci si può aspettare a priori da una misurazione è circa nulla.

4.2 L'entropia di Kolmogorov-Sinai

Per il momento, la definizione di entropia ha a che fare con lo spazio con misura (M, μ) , ma non ha nulla a che fare con una dinamica. Supponiamo ora sia dato un sistema dinamico misurabile (M, μ, Φ) , ed adoperiamo la funzione di entropia appena introdotta per le partizioni al fine di ottenere informazioni sul sistema dinamico misurabile in questione.

Definizione 60 *Data una partizione α ed un numero naturale N , la partizione*

$$\alpha^{(N)} = \alpha \vee \Phi^{-1}(\alpha) \vee \dots \vee \Phi^{-N}(\alpha)$$

*si dice la **partizione generata da α e Φ al passo N** .*

La partizione $\alpha^{(N)}$ si arricchisce sempre più in elementi in base a come l'isomorfismo Φ rimescola le parti di M .

Definizione 61 L'entropia del sistema dinamico Φ associato alla partizione α è il numero reale

$$h(\Phi, \alpha) = \lim_{N \rightarrow \infty} \frac{\eta(\alpha^{(N)})}{N} := \lim_{N \rightarrow \infty} h_N(\Phi, \alpha)$$

La prima domanda da porsi è se la definizione appena data sia sempre sensata, ovvero se ci sono delle situazioni in cui il limite che definisce $h(\Phi, \alpha)$ non esiste.

Proposizione 62 Sia dato (M, μ, Φ) sistema dinamico misurabile. Il limite usato nella definizione di $h(\Phi, \alpha)$ esiste per ogni scelta di α .

Dimostrazione. Dalla definizione di $\alpha^{(N)}$, e dalla proprietà $\eta 3$, si ha che $\eta(\alpha^{(N+M)}) = \eta(\alpha^{(N)} \vee \alpha^{(M)}) \leq \eta(\alpha^{(N)}) + \eta(\Phi^{-N} \alpha^{(M)}) = \eta(\alpha^{(N)}) + \eta(\alpha^{(M)})$. Segue quindi che

$$h_{N+M}(\Phi, \alpha) = \frac{\eta(\alpha^{(N+M)})}{N+M} \leq \frac{N}{N+M} \frac{\eta(\alpha^{(N)})}{N} + \frac{M}{N+M} \frac{\eta(\alpha^{(M)})}{M} = \frac{N}{N+M} h_N(\Phi, \alpha) + \frac{M}{N+M} h_M(\Phi, \alpha).$$

Quando una successione soddisfa una disuguaglianza del genere, si ha che il limite $h_N(\Phi, \alpha)$ esiste e coincide con l'estremo inferiore $h = \inf_N h_N(\Phi, \alpha)$. Per dimostrare questo fatto si deve dimostrare che per ogni N abbastanza grande $h_N(\Phi, \alpha) - h$ è arbitrariamente piccolo, ovvero che $\forall \varepsilon > 0$ esiste un N_ε tale che se $N > N_\varepsilon$ allora $h_N(\Phi, \alpha) - h \leq \varepsilon$.

Per definizione di $\inf$, per ogni ε esiste un N_ε tale che $h_{N_\varepsilon}(\Phi, \alpha) \leq h + \varepsilon$. Dalla disuguaglianza dimostrata sopra segue che, per ogni intero n si ha che $h_{nN_\varepsilon}(\Phi, \alpha) \leq h + \varepsilon$ (una facile dimostrazione per induzione). Prendiamo quindi un N maggiore di N_ε , necessariamente $N = nN_\varepsilon + r$ per qualche n, r interi con $r < N_\varepsilon$. Segue quindi che

$$h_N(\Phi, \alpha) \leq \frac{nN_\varepsilon}{N} h_{nN_\varepsilon}(\Phi, \alpha) + \frac{r}{N} h_r(\Phi, \alpha) \leq h + \varepsilon + \frac{r}{N} h_r(\Phi, \alpha) \leq h + \varepsilon + \frac{1}{n} h_r(\Phi, \alpha).$$

Se N è abbastanza grande, allora n è abbastanza grande da assicurare che $h_r(\Phi, \alpha)/n \leq \varepsilon$ qualsiasi sia $r = 0, \dots, N_\varepsilon - 1$. Segue che per tali N , tutti gli $h_N(\Phi, \alpha)$ sono stimati da sopra da $h + 2\varepsilon$. $\square$

Una osservazione, per prendere confidenza con le definizioni, è che il caso peggiore di crescita di $h_N(\Phi, \alpha)$ è quello in cui $|\alpha| = n$, gli elementi di α hanno tutti misura $1/n$, e la dinamica mescola gli elementi di α in modo così uniforme da far sì che $|\alpha^{(N)}| = n^{N+1}$ e che tutti gli elementi di $\alpha^{(N)}$ abbiano misura $1/n^{N+1}$. Allora

$$\eta(\alpha^{(N)}) = - \sum_{A \in \alpha^{(N)}} \log(1/n^{N+1})/n^{N+1} = (N+1) \log n,$$

da cui segue che $h(\Phi, \alpha) = \log n$.

Definizione 63 Dato un sistema dinamico discreto (M, μ, Φ) , la sua **entropia di Kolmogorov-Sinai** (detta anche **entropia metrica**) è il numero in $[0, +\infty]$ definito come

$$h(\Phi) = \sup_{\alpha \in \mathcal{P}} h(\Phi, \alpha)$$

Un primo fatto importante è che.

Proposizione 64 Sia (M, μ, Φ) un sistema dinamico misurabile reversibile, vale $h(\Phi^{-1}) = h(\Phi)$.

Dimostrazione. Ogni troncata N -esima $h_N(\Phi, \alpha) = \eta(\alpha \vee \dots \vee \Phi^{-N}(\alpha))/N$ coincide con $h_N(\Phi^{-1}, \alpha) = \eta(\Phi^N(\alpha) \vee \dots \vee \alpha)/N$. Basta infatti applicare Φ^N alla partizione $\alpha \vee \dots \vee \Phi^{-N}(\alpha)$. L'uguaglianza discende passando al limite per N che tende a ∞ e poi passando al sup sulla partizione α . $\square$

Osservazione. Potevamo equivalentemente definire l'entropia di un sistema dinamico misurabile reversibile come

$$\lim_{N \rightarrow +\infty} \frac{\eta(\alpha^{[N]})}{2N}$$

dove $\alpha^{[N]} = \Phi^{-N}(\alpha) \vee \dots \vee \alpha \vee \dots \vee \Phi^N(\alpha)$.

4.2.1 L'entropia dei sistemi continui

Finora abbiamo definito e studiato l'entropia di sistemi dinamici nei quali la mappa Φ è uno shift. Qualora il sistema dinamico sia a tempo continuo, ovvero $\Phi = \Phi^t$ sia un flusso, allora viene spontaneo definire

Definizione 65 Dato un sistema dinamico continuo (M, μ, Φ) , l'entropia di Kolmogorov-Sinai del sistema è l'entropia di Kolmogorov-Sinai del sistema dinamico discreto (M, μ, Φ^1) , ovvero $h(\Phi) := h(\Phi^1)$.

La domanda d'obbligo a cui bisogna rispondere è se questa definizione è indipendente dalla scelta della scala temporale. Ovvero, l'entropia $h(\Phi^T)$ del sistema discreto Φ^T con $T \neq 1$ ha qualche relazione con l'entropia $h(\Phi)$ definita sopra? In realtà vale il fatto seguente.

Proposizione 66 Sia dato un sistema dinamico misurabile continuo (M, μ, Φ) , e sia T un reale. L'entropia associata al sistema dinamico discreto di mappa Φ^T è $h(\Phi^T) = |T|h(\Phi)$.

Dimostrazione. Supponiamo $T = p/q$ con p, q coprimi. Non è restrittivo supporre che $p < q$, perchè se $p > q$ allora basta definire $\Psi^s := \Phi^{\frac{p}{q}s}$, osservare che per questo nuovo flusso $\Psi^1 = \Phi^{\frac{p}{q}}$, $\Psi^{\frac{q}{p}} = \Phi^1$, e dimostrare la formula per Ψ .

Consideriamo la partizione $\tilde{\alpha} = \alpha \vee \Phi^{-\frac{1}{q}}(\alpha) \vee \Phi^{-\frac{2}{q}}(\alpha) \cdots \vee \Phi^{-1}(\alpha)$. Si ha che

$$\begin{aligned} h_{qN}(\Phi^T, \tilde{\alpha}) &= \frac{1}{qN} \eta(\alpha \vee \Phi^{-\frac{1}{q}}(\alpha) \vee \cdots \vee \Phi^{-Np-1}(\alpha)) \\ &= \frac{pN}{qN} \frac{1}{pN} \eta(\alpha \vee \Phi^{-\frac{1}{q}}(\alpha) \vee \cdots \vee \Phi^{-Np-1}(\alpha)) = \\ &= \frac{p}{q} h_{pN}(\Phi^1, \tilde{\alpha}). \end{aligned}$$

Facendo tendere N all'infinito si ha che $h(\Phi^T, \tilde{\alpha}) = \frac{p}{q} h(\Phi^1, \tilde{\alpha})$. Passando poi al sup al variare di $\tilde{\alpha}$ ed osservando che le entropie relative a $\tilde{\alpha}$ maggiorano quelle relative ad α , si ottiene il risultato.

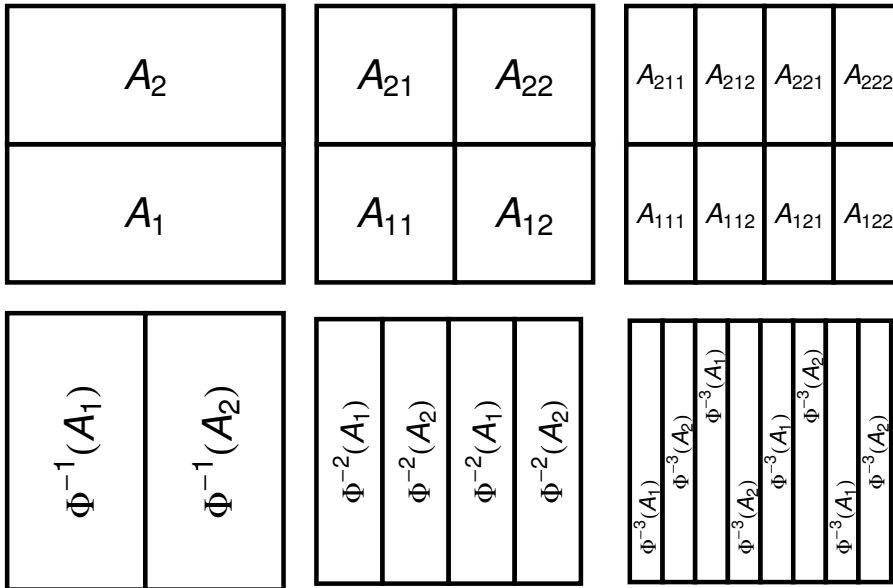
Ottenere il risultato per T irrazionale è più laborioso perchè, dato α , non ci si può limitare alla scelta una volta per tutte di un suo raffinamento $\tilde{\alpha}$. Non scriviamo la dimostrazione. $\square$

4.2.2 Prime stime di Entropia

La definizione di entropia di Kolmogorov-Sinai rende facile stimare da sotto le entropie dei sistemi dinamici misurabili: si sceglie un opportuno ricoprimento α e si calcola l'entropia del sistema dinamico Φ associata a quel ricoprimento. Il valore di $h(\Phi, \alpha)$ è una stima da sotto. Un paio di esempi.

Esempio 67 L'entropia della trasformazione del panettiere è stimata da sotto da $\log 2$.

Dimostrazione. Consideriamo la partizione α formata da due rettangoli $A_1 = [0, 1) \times [0, 1/2)$ ed $A_2 = [0, 1) \times [1/2, 1)$. La partizione $\alpha^{(1)}$ è formata da 4 aperti di misura $1/4$. La partizione $\alpha^{(2)}$ è formata da 2^3 aperti di misura 2^{-3} . In generale, la partizione $\alpha^{(N)}$ è formata da 2^{N+1} aperti di misura 2^{-N-1} .



Ne segue che l'entropia di $\alpha^{(N)}$ è il numero $\eta(\alpha^{(N)}) = -2^{N+1}2^{-N-1} \log 2^{-N-1} = (N+1) \log 2$. Quindi $h(\Phi, \alpha) = \log 2$, che fornisce una stima da sotto dell'entropia del sistema. $\square$

Esempio 68 L'entropia di uno schema di Bernoulli $B(p_1, \dots, p_n)$ è stimata da sotto da $-\sum_{\ell=1}^n p_\ell \log p_\ell$ (anche quando $n = +\infty$).

Dimostrazione. Consideriamo la partizione α formata dagli n cilindri elementari C_ℓ^0 . La partizione $\alpha^{(1)}$ è formata dagli n^2 cilindri elementari $C_{\ell_0, \ell_1}^{0,1}$. Continuando allo stesso modo si ottiene che la partizione $\alpha^{(N)}$ è formata da n^{N+1} cilindri elementari $C_{\ell_0, \dots, \ell_N}^{0, \dots, N}$.

Segue che l'entropia di $\alpha^{(N)}$ è il numero $\eta(\alpha^{(N)}) = \sum p_{\ell_0} \cdots p_{\ell_N} \log(p_{\ell_0} \cdots p_{\ell_N})$. Riorganizzando i termini, la somma si riscrive come

$$\begin{aligned} & \sum_{\ell_0, \dots, \ell_{N-1}} \sum_{\ell=1}^n p_{\ell_0} \cdots p_{\ell_{N-1}} p_\ell [\log(p_{\ell_0} \cdots p_{\ell_{N-1}}) + \log p_\ell] = \\ & \sum_{\ell_0, \dots, \ell_{N-1}} p_{\ell_0} \cdots p_{\ell_{N-1}} \log(p_{\ell_0} \cdots p_{\ell_{N-1}}) + \sum_{\ell_0, \dots, \ell_{N-1}} p_{\ell_0} \cdots p_{\ell_{N-1}} \left[\sum_{\ell=1}^n p_\ell \log p_\ell \right] = \dots \\ & (N+1) \sum_{\ell=1}^n p_\ell \log p_\ell. \end{aligned}$$

Dividendo per N e passando al limite si ottiene che $h(\Phi, \alpha) = \sum_{\ell=1}^n p_\ell \log p_\ell$, che fornisce una stima da sotto dell'entropia di uno schema di Bernoulli. $\square$

L'esempio precedente, con un vettore di probabilità infinito, porge un caso di sistema dinamico la cui entropia associata è infinita.

Esempio 69 Scelta a piacere una successione $n_1^{-1}, \dots, n_k^{-1}, \dots$ la cui somma converge ad 1 ed in cui gli $n_i \in \mathbb{N}^>$. Lo schema di Bernoulli infinito il cui vettore di probabilità è dato dalla successione $n_1^{-1}2^{-n_1}$ ripetuto 2^{n_1} , seguito da $n_2^{-1}2^{-n_2}$ ripetuto 2^{n_2} e così via ha entropia $+\infty$.

Dimostrazione. Applicando il risultato della proposizione precedente, si ottiene che l'entropia dello schema di Bernoulli appena definito è stimata da sotto dal limite della serie

$$\begin{aligned} & -2^{n_1} \frac{1}{n_1 2^{n_1}} \log \left(\frac{1}{n_1 2^{n_1}} \right) - \dots - 2^{n_k} \frac{1}{n_k 2^{n_k}} \log \left(\frac{1}{n_k 2^{n_k}} \right) - \dots = \\ & = \frac{1}{n_1} (\log n_1 + n_1 \log 2) + \dots + \frac{1}{n_k} (\log n_k + n_k \log 2) + \dots \end{aligned}$$

Si vede che la serie è la somma di una serie a termini positivi che potrebbe anche convergere ed una serie a termini costanti positivi che sicuramente diverge a $+\infty$. $\square$

4.3 Partizioni generanti e calcolo dell'entropia

Finora abbiamo dato la definizione di entropia di un sistema dinamico misurabile, ed abbiamo stimato da sotto l'entropia di un po' di esempi. Bisogna ora capire come calcolare l'entropia in modo esatto. La definizione richiede di calcolare il sup tra le entropie del sistema dinamico associate a tutte le partizioni misurabili di M . Ovviamente questo sup non verrà operativamente ottenuto calcolando $h(\Phi, \alpha)$ per tutte le partizioni α . Si dimostrerà invece che è sufficiente calcolare $h(\Phi, \bar{\alpha})$ per **una** partizione $\bar{\alpha}$ ben scelta, che realizza il sup.

Definizione 70 Sia dato un sistema dinamico misurabile (M, μ, Φ) una partizione α si dice **Φ -generante** (la σ -algebra) se gli atomi della partizione assieme alle loro Φ -iterate sono generatori della σ -algebra. Ovvero se la σ -algebra è generata dal sottoinsieme delle parti di M i cui elementi sono $\{\Phi^j(A) \mid A \in \alpha, j \in \mathbb{Z}\}$.

Vale il seguente, cruciale teorema.

Teorema 71 Se α è partizione generante, allora $h(\Phi, \alpha) = h(\Phi)$.

Postponiamo alla Sezione 4.7.1 la discussione sulle partizioni generanti e la dimostrazione del prossimo teorema, la quale richiede alcune definizioni e risultati che richiedono un po' di lavoro. Passiamo invece alla applicazione del teorema ad alcuni sistemi dinamici misurabili.

Osservazione 72 La partizione dello spazio delle fasi del sistema dinamico del panettiere data nella sezione precedente è una partizione generante, segue che l'entropia del sistema dinamico del panettiere è $\log 2$.

La partizione in cilindri elementari $\alpha = \{C_\ell^0 \mid \ell = 1, \dots, k\}$ di uno schema di Bernoulli $B(p_1, \dots, p_n)$ sono generanti, segue che l'entropia di uno schema di Bernoulli è $-\sum_{j=1}^n p_j \log p_j$.

Dimostrazione. Il fatto che per la mappa del panettiere la partizione è generante è evidente, infatti le iterate dei due rettangoli genera una quadrettatura infinitesima del quadrato. Per gli schemi di Bernoulli, è facile osservare che iterando i cilindri C_ℓ^0 si ottiene la famiglia di cilindri elementari $\{C_\ell^j \mid j \in \mathbb{Z}, \ell = 1, \dots, n\}$, e che questi, per definizione, generano la σ -algebra. $\square$

4.4 Entropia come invariante dinamico

Definizione 73 Due sistemi dinamici misurabili (M, μ, Φ) e (N, ν, Ψ) sono **isomorfi** se esiste A sottoinsieme di M di misura 1 e B sottoinsieme di N di misura 1 ed una biiezione misurabile $f : A \rightarrow B$ con inversa misurabile $f^{-1}B \rightarrow A$ che preserva la misura, ovvero tale che $\nu(S) = \mu(f^{-1}(S))$ per ogni $S \in Y$ misurabile e che coniuga i sistemi dinamici, ovvero tale che $f(\Phi(x)) = \Psi(f(x))$.

Proposizione 74 Sistemi dinamici misurabili isomorfi hanno la stessa entropia

Dimostrazione. Tutto si corrisponde!!! $\square$

4.5 gli schemi di Bernoulli e l'entropia

Abbiamo visto come l'entropia sia un numero reale positivo che permette di distinguere sistemi dinamici misurabili non isomorfi. L'invariante ideale è una quantità che si usa per dimostrare non solo quando due sistemi dinamici misurabili sono diversi, ma anche quando questi sono uguali. Un invariante di tal fatta permetterebbe una classificazione completa all'interno della categoria dei sistemi dinamici misurabili. Ovviamente, un tale tipo di invariante è il fine di qualsiasi teoria di classificazione, ed è difficile da ottenere a meno di circoscrivere la categoria in cui si lavora. Infatti l'entropia non permette una classificazione dei sistemi dinamici misurabili, ma se ristretta alla categoria degli shift di Bernoulli permette di classificarli. La dimostrazione dei due risultati elencati sotto è molto laboriosa, li citiamo per dare un'idea di parte della ricerca nel settore.

Teorema 75 Due schemi di Bernoulli $B(p_1, \dots, p_n)$ e $B(q_1, \dots, q_m)$ sono isomorfi se e solo se hanno la stessa entropia.

La dimostrazione di questo fatto, addirittura con una condizione stringente sul tipo di isomorfismo, si trova in M. Keane e M. Smorodisky, *Bernoulli schemes of the same entropy are finitarily isomorphic*, Annals of Mathematics, **109**/2 397–406 (1979).

Ma la vita non è poi così facile.

Esempio 76 Esistono sistemi dinamici non isomorfi ma con stessa entropia.

Anche questi risultati sono difficili da dimostrare. Per una costruzione di due sistemi dinamici non isomorfi ma con stessa entropia si veda D. Rudolph, *Two nonisomorphic K -automorphisms with isomorphic squares*, Israel Journal of Mathematics, **23**/3-4 274–287 (1976).

4.6 Entropia e sistemi classici

Abbiamo visto come una parte importante di sistemi dinamici misurabili provengano da un campo vettoriale definito su di una varietà liscia. In tali varietà la tentazione è di calcolare l'entropia su oggetti che siano lisci, od almeno lisci a tratti. Chiamiamo **sistema dinamico classico** un sistema dinamico misurabile (M, μ, Φ) in cui M è varietà liscia, che non è restrittivo supporre immersa in qualche spazio Euclideo $\mathbb{R}^N$, Φ è il flusso di un campo vettoriale liscio oppure un diffeomorfismo liscio, e μ è una misura assolutamente continua rispetto alla misura di Lebesgue, ovvero è una misura associata ad una forma volume. Chiameremo tali forme differenziali **regolari**. Nella categoria dei sistemi dinamici misurabili lisci definiamo

Definizione 77 Una partizione α di M si dice **partizione classica** se ciascun atomo della partizione è unione finita di poliedri (o complessi) semiaperti. Ovvero di insiemi A semplicemente connessi, tali che $A \subset (A^\circ)^-$ e con bordo ∂A regolare a tratti.

Questo tipo di partizioni tornano utili perchè

Lemma 78 *Data una varietà M compatta ed una misura μ regolare, allora le partizioni classiche sono dense, nello spazio $\mathcal{P}$ delle partizioni misurabili dotato della metrica dist .*

Dimostrazione. Dobbiamo mostrare che per qualunque partizione misurabile α e per ogni $\varepsilon > 0$ esiste una partizione classica $\tilde{\alpha}$ tale che $\text{dist}(\alpha, \tilde{\alpha}) < \varepsilon$.

Sia A un insieme misurabile, non è restrittivo supporre che A sia contenuto in un aperto di carta. Siccome i misurabili in una varietà coincidono con i boreliani di una carta, per ogni $\varepsilon > 0$ esistono una famiglia finita di parallelepipedi tali che la loro unione $\tilde{A}^\varepsilon$ è contenuta in A e la differenza $A - \tilde{A}^\varepsilon$ ha misura minore di ε .

Sia quindi data una partizione misurabile $\alpha = \{A_1, \dots, A_n\}$ e sia preso ad arbitrio $\varepsilon > 0$. Siano infine $\tilde{A}_i^{\varepsilon/n}$ insiemi classici con le proprietà sopra. La partizione

$$\tilde{\alpha} = \{B_1 = \tilde{A}_1^{\varepsilon/n}, \dots, B_n = \tilde{A}_n^{\varepsilon/n}, B_0 = M - \cup_i \tilde{A}_i^{\varepsilon/n}\}$$

ha le proprietà richieste. Infatti $|\mu(B_j|A_i) - \delta_{i,j}| \leq |\mu(A_i|B_j) - \delta_{i,j}| \leq \varepsilon$ per ogni i, j , quindi applicare a queste misure relative la funzione ρ , che è continua e vale 0 in 0,1, e poi applicare una somma finita, porge il fatto che $\eta(\alpha|\tilde{\alpha})$ e $\eta(\tilde{\alpha}|\alpha)$ sono vicine a piacere a zero. $\square$

Definizione 79 *Dato un sistema dinamico classico, il coefficiente di dilatazione lineare di Φ è il più piccolo reale positivo λ tale che $\|d_q\Phi(v)\| \leq \lambda\|v\|$ per ogni $q \in M$ e $v \in T_qM$.*

Teorema 80 (Teorema di Kouchnirenko) *L'entropia nei sistemi classici è finita ed ammette la stima*

$$h(\Phi) \leq (\dim M) \log \lambda.$$

dove λ è il coefficiente di dilatazione lineare associato alla dinamica Φ .

Dimostrazione. Visto il risultato di densità appena dimostrato ed il fatto che la funzione h è continua, possiamo limitarci a calcolare h nelle partizioni classiche.

Assumiamo, e non è restrittivo visto un teorema di geometria Riemanniana che afferma l'esistenza di immersioni metriche, che la misura in M provenga da una metrica indotta dalla immersione di M in uno spazio Euclideo, e sia λ il coefficiente di dilatazione lineare di Φ . Ogni atomo di una partizione classica ha bordo liscio ed $n - 1$ dimensionale. La struttura Euclidea permette di calcolare la superficie del bordo, denotiamo con $\mathcal{S}(A)$ la misura della superficie del bordo di A . La conservazione della misura implica che $\mathcal{S}(\Phi^{-1}(A)) \leq \lambda \mathcal{S}(A)$.

Data una partizione classica, indichiamo con $\mathcal{S}(\alpha)$ la somma delle aree dei suoi atomi. È immediato osservare che $\mathcal{S}(\alpha \vee \beta) \leq \mathcal{S}(\alpha) + \mathcal{S}(\beta)$. Iterando si ha che

$$\mathcal{S}(\alpha \vee \dots \vee \Phi^{-N}(\alpha)) \leq (1 + \lambda + \dots + \lambda^N) \mathcal{S}(\alpha) = \frac{1 - \lambda^{N+1}}{1 - \lambda} \mathcal{S}(\alpha)$$

Usiamo ora un'altro fatto di geometria Riemanniana noto come disuguaglianza isoperimetrica, che afferma l'esistenza di costanti C, D tali che se $\mu(A) \leq D$ allora $\mu(A) \leq C \mathcal{S}(A)^{\frac{m}{m-1}}$, dove $m = \dim M$. Supponiamo ora di avere una partizione classica α i cui elementi hanno tutti misura limitata da D . Si ha che

$$\begin{aligned} \eta(\alpha) &= - \sum \mu(A) \log(\mu(A)) = m \sum \mu(A) \log(\mu(A)^{-\frac{1}{m}}) \\ &\leq m \log\left(\sum \mu(A) \mu(A)^{-\frac{1}{m}}\right) = m \log\left(\sum \mu(A)^{\frac{m-1}{m}}\right) \\ &\leq m \log\left(C^{\frac{m-1}{m}} \sum \mathcal{S}(A)\right) = m \log(\mathcal{S}(\alpha)) + (m-1) \log C \end{aligned}$$

Da questa disuguaglianza segue che

$$\begin{aligned} \eta(\alpha \vee \dots \vee \Phi^{-N}(\alpha)) &\leq m \log(\mathcal{S}(\alpha \vee \dots \vee \Phi^{-N}(\alpha))) + (m-1) \log C \\ &\leq m \log\left(\frac{1 - \lambda^{N+1}}{1 - \lambda}\right) + m \log(\mathcal{S}(\alpha)) + (m-1) \log C. \end{aligned}$$

Dividendo per N e passando al limite per N che tende all'infinito si ottiene una maggiorazione indipendente da α da cui segue che

$$h(\Phi) \leq m \log \lambda.$$

$\square$

Nella prossima sezione enunceremo un risultato che porge la disuguaglianza di Kouchnirenko. È comunque notevole avere una stima da sopra dell'entropia di un sistema, visto che le stime da sotto sono facili da ottenere.

Corollario 81 *L'entropia della traslazione in T^n è nulla.*

Dimostrazione. Tutti i sistemi dinamici che si possono scrivere come una traslazione in qualche sistema di coordinate hanno coefficiente di dilatazione 1, da cui si stima che $h(\Phi) \leq 0$. $\square$

4.7 Entropia ed esponenti di Liapunov

In questa Sezione enunceremo un raffineremo della disuguaglianza di Kouchnirenko in una identità nota come **formula di Pesin**. Questa formula non si limita a stimare l'entropia usando un coefficiente di espansione massima, ma la determina a partire dall'intero spettro di possibili dilatazioni, espressi sotto forma di esponenti di Lyapunov.

Teorema 82 *Dato un sistema dinamico classico misurabile (M, μ, Φ) , si ha che*

$$h(\Phi) = \int_M \sum_{\lambda \in Liap^+} \lambda(q) d\mu,$$

dove $Liap^+$ sono gli esponenti di Liapunov positivi.

Non dimostriamo questo teorema, il cui enunciato è comunque così conciso da non aver bisogno della dimostrazione per essere capito ed usato.

Corollario 83 *Se il sistema dinamico è ergodico allora tutte le λ sono costanti e quindi $h = \sum_{\lambda \in Liap^+} \lambda$*

Corollario 84 *L'entropia della traslazione del toro è 0. L'entropia del gatto di Arnol'd è $\log(3 + \sqrt{5})/2$*

Dimostrazione. Gli esponenti di Lyapunov delle traslazioni del toro sono tutti 0. Gli esponenti di Lyapunov del gatto di Arnol'd sono $\log(3 \pm \sqrt{5})/2 = \pm \log(3 + \sqrt{5})/2$. L'unico positivo è $\log((3 + \sqrt{5})/2)$. $\square$

4.8 Entropia condizionata e distanza sullo spazio delle partizioni misurabili

La funzione $h(\Phi, \alpha)$ che ad una partizione associa un numero reale positivo è quindi una funzione dallo spazio delle partizioni misurabili a valori nei reali positivi. Attraverso questa funzione abbiamo definito l'entropia, ma non ci siamo soffermati abbastanza su alcune caratteristiche di h che torneranno utili (come ad esempio il Teorema 71), ed in ogni caso il calcolo esplicito dell'entropia passa attraverso una adeguata comprensione di h . In questa sezione studieremo la funzione h , ma ha anche ne approfitteremo per introdurre dei concetti e sviluppare delle tecniche che appartengono alla teoria della probabilità.

Cominciamo introducendo il concetto di misura condizionata e di entropia condizionata.

Definizione 85 *Dati due insiemi misurabili A, B , la misura di A condizionata a B è il numero reale in $[0, 1]$ dato da $\mu(A|B) = \mu(A \cap B)/\mu(B)$.*

Data una partizione α ed un insieme misurabile B , l'entropia di α condizionata a B è la quantità

$$\eta(\alpha|B) = \sum_{A \in \alpha} \rho(\mu(A|B)).$$

Date due partizioni α, β l'entropia di α relativa a β è la quantità

$$\eta(\alpha|\beta) = \sum_{B \in \beta} \mu(B) \eta(\alpha|B).$$

Con queste definizioni, la proprietà $\eta 3$, dimostrata nella Proposizione 56 si può affinare nel seguente modo

Lemma 86 *Date due partizioni α, β , vale sempre che $\eta(\alpha|\beta) \leq \eta(\alpha)$, ed inoltre vale*

$$\eta(\alpha|\beta) = \eta(\alpha \vee \beta) - \eta(\beta).$$

Dimostrazione. Usando la definizione,

$$\eta(\alpha|\beta) = - \sum_{B \in \beta} \left(\sum_{A \in \alpha} \mu(A \cap B) \left(\log(\mu(A \cap B)) - \log(\mu(B)) \right) \right) = \dots$$

distribuendo, si ha che

$$\dots = - \sum_{B \in \beta} \left(\sum_{A \in \alpha} \mu(A \cap B) \log(\mu(A \cap B)) \right) - \sum_{B \in \beta} \left(\sum_{A \in \alpha} -\mu(A \cap B) \log(\mu(B)) \right) = \dots$$

il primo termine è proprio $\eta(\alpha \vee \beta)$, mentre sommando su A si ha che il secondo termine è $-\eta(\beta)$.

La seconda affermazione è una facile applicazione della proprietà $\eta 3$ e di quanto appena mostrato. Infatti $\eta(\alpha|\beta) = \eta(\alpha \vee \beta) - \eta(\beta)$ ed $\eta(\alpha \vee \beta) \leq \eta(\alpha) + \eta(\beta)$ con uguaglianza se e solo se α è indipendente da β . $\square$

L'entropia condizionata gode di svariate proprietà, che ci torneranno utili.

Lemma 87 *Siano α, β, γ tre partizioni misurabili, allora*

1. $\beta \succcurlyeq \alpha$ se e solo se $\eta(\alpha|\beta) = 0$,
2. Se $\beta \succcurlyeq \alpha$ allora per ogni partizione misurabile γ vale $\eta(\beta|\gamma) \geq \eta(\alpha|\gamma)$ e $\eta(\gamma|\beta) \leq \eta(\gamma|\alpha)$,
3. $\eta(\alpha \vee \beta|\gamma) \leq \eta(\alpha|\gamma) + \eta(\beta|\gamma)$.

Dimostrazione. (1) $\beta \succcurlyeq \alpha$ se e solo se per ogni $A \in \alpha$ e $B \in \beta$ vale $\mu(A|B) = 0, 1$, perchè $\mu(A \cap B) = 0, \mu(B)$, se e solo se per ogni $A \in \alpha$ e $B \in \beta$ vale $\rho(\mu(A|B)) = 0$, se e solo se $\eta(\alpha|\beta) = 0$.

(2) Siano date due partizioni $\beta \succcurlyeq \alpha$, per il Lemma 86 vale

$$\eta(\beta|\gamma) = \eta(\beta \vee \gamma) - \eta(\gamma).$$

Siccome $\beta \vee \gamma \succcurlyeq \alpha \vee \gamma$ si ha che

$$\eta(\beta \vee \gamma) - \eta(\gamma) \geq \eta(\alpha \vee \gamma) - \eta(\gamma) = \eta(\alpha|\gamma).$$

Per l'altra disuguaglianza, utilizziamo la concavità di ρ . Fissiamo un $C \in \gamma$ ed un $A \in \alpha$ ed osserviamo che

$$\begin{aligned} \sum_{B \subset A} \frac{\mu(B)}{\mu(A)} \rho(\mu(C|B)) &\leq \rho \left(\sum_{B \subset A} \frac{\mu(B)}{\mu(A)} \mu(C|B) \right) \\ &= - \left(\sum_{B \subset A} \frac{\mu(C \cap B)}{\mu(A)} \right) \log \left(\sum_{B \subset A} \frac{\mu(C \cap B)}{\mu(A)} \right) \\ &= - \frac{\mu(C \cap A)}{\mu(A)} \log \left(\frac{\mu(C \cap A)}{\mu(A)} \right) = \rho(\mu(C|A)). \end{aligned}$$

Quindi $\sum_{B \subset A} \mu(B) \rho(\mu(C|B)) \leq \mu(A) \rho(\mu(C|A))$. Segue che

$$\begin{aligned} \eta(\gamma|\beta) &= \sum_{B \in \beta} \mu(B) \left(\sum_{C \in \gamma} \rho(\mu(C|B)) \right) \\ &= \sum_{A \in \alpha} \left(\sum_{B \subset A} \mu(B) \left(\sum_{C \in \gamma} \rho(\mu(C|B)) \right) \right) \\ &\leq \sum_{A \in \alpha} \left(\sum_{C \in \gamma} \mu(A) \rho(\mu(C|A)) \right) \\ &= \eta(\gamma|\alpha) \end{aligned}$$

(3) si ha che

$$\begin{aligned} \eta(\alpha \vee \beta|\gamma) &= \eta(\alpha \vee \beta \vee \gamma) - \eta(\gamma) \\ &= \eta(\alpha \vee \beta \vee \gamma) \pm \eta(\beta \vee \gamma) - \eta(\gamma) \\ &= \eta(\alpha|\beta \vee \gamma) + \eta(\beta|\gamma) \\ &\leq \eta(\alpha|\gamma) + \eta(\beta|\gamma) \end{aligned}$$

□

Possiamo ora dotare lo spazio delle partizioni misurabili di una funzione distanza

Definizione 88 *Date due partizioni misurabili α, β , la distanza tra le partizioni è il reale positivo*

$$\text{dist}(\alpha, \beta) = \eta(\alpha|\beta) + \eta(\beta|\alpha)$$

Mostriamo quindi che la funzione *dist* appena definita è realmente una distanza

Proposizione 89 *La funzione *dist* è una distanza.*

Dimostrazione. Dobbiamo dimostrare le tre proprietà della distanza, ovvero che $\text{dist}(\alpha, \beta) \geq 0$, con uguaglianza se e solo se $\alpha = \beta$, la simmetria e la disuguaglianza triangolare. La simmetria è evidente dalla definizione, la positività segue dal fatto che le entropie relative sono positive, se $\text{dist}(\alpha, \beta) = 0$ allora $\eta(\alpha|\beta) = 0$ e $\eta(\beta|\alpha) = 0$, da cui segue che $\beta \succcurlyeq \alpha$ ed $\alpha \succcurlyeq \beta$, e quindi $\alpha = \beta$.

Per la disuguaglianza triangolare,

$$\begin{aligned} \text{dist}(\alpha, \gamma) &= \eta(\alpha|\gamma) + \eta(\gamma|\alpha) \\ &= \eta(\alpha \vee \gamma) - \eta(\gamma) + \eta(\gamma \vee \alpha) - \eta(\alpha) \\ &\leq \eta(\alpha \vee \gamma \vee \beta) \pm \eta(\beta \vee \gamma) - \eta(\gamma) + \eta(\gamma \vee \alpha \vee \beta) \pm \eta(\beta \vee \alpha) - \eta(\alpha) \\ &= \eta(\alpha|\gamma \vee \beta) + \eta(\beta|\gamma) + \eta(\gamma|\alpha \vee \beta) + \eta(\beta|\alpha) \\ &\leq \eta(\alpha|\beta) + \eta(\beta|\gamma) + \eta(\gamma|\beta) + \eta(\beta|\alpha) \\ &= \text{dist}(\alpha, \beta) + \text{dist}(\beta, \gamma) \end{aligned}$$

□

La funzione distanza gioca un ruolo importante nella teoria dell'entropia.

Proposizione 90 *L'entropia è funzione Lipschitziana con costante di Lipschitz 1 dallo spazio metrico $(\mathcal{P}, \text{dist})$ in $\mathbb{R}$. In formula, $|h(\Phi, \alpha) - h(\Phi, \beta)| \leq \text{dist}(\alpha, \beta)$.*

Dimostrazione. Basta mostrare che per ogni N naturale, si ha che $|h_N(\Phi, \alpha) - h_N(\Phi, \beta)| \leq \text{dist}(\alpha, \beta)$. Non è restrittivo supporre che $h(\Phi, \alpha) > h(\Phi, \beta)$, il che permette di non usare il modulo per $N \gg 0$. Si ha che

$$\begin{aligned} \eta(\alpha \vee \dots \vee \Phi^{-N}(\alpha)) - \eta(\beta \vee \dots \vee \Phi^{-N}(\beta)) &\leq \\ &\leq \eta(\alpha \vee \dots \vee \Phi^{-N}(\alpha) \vee \beta \vee \dots \vee \Phi^{-N}(\beta)) - \eta(\beta \vee \dots \vee \Phi^{-N}(\beta)) \\ &= \eta(\alpha \vee \dots \vee \Phi^{-N}(\alpha) | \beta \vee \dots \vee \Phi^{-N}(\beta)) \\ &\leq \sum_{s=0}^N \eta(\Phi^{-s}(\alpha) | \beta \vee \dots \vee \Phi^{-N}(\beta)) \\ &\leq \sum_{s=0}^N \eta(\Phi^{-s}(\alpha) | \Phi^{-s}(\beta)) \\ &= \sum_{s=0}^N \eta(\alpha|\beta) = (N+1)\eta(\alpha|\beta). \end{aligned}$$

Da cui si ricava che

$$h_N(\Phi, \alpha) - h_N(\Phi, \beta) \leq \frac{N+1}{N} \eta(\alpha|\beta) \leq \frac{N+1}{N} \text{dist}(\alpha, \beta)$$

(visto che ovviamente $\eta(\alpha|\beta) \leq \text{dist}(\alpha, \beta)$). Si ottiene quindi che

$$h_N(\Phi, \alpha) - h_N(\Phi, \beta) \leq \frac{N+1}{N} \text{dist}(\alpha, \beta)$$

per tutti gli $N \gg 0$. Passando al limite per N che tende ad infinito, questa disuguaglianza porge la disuguaglianza che volevamo dimostrare. □

4.8.1 Dimostrazione del teorema sulle partizioni generanti

Ricordiamo che un sottoinsieme γ delle parti di M genera la σ -algebra se ogni misurabile U si scrive come unione numerabile di intersezioni numerabili di elementi di γ o di loro complementari, ovvero se $U = \cup_i (\cap_j C_{i,j}^*)$, dove $C_{i,j} \in \gamma$ e la stellina indica l'identità od il complemento algebrico. Quindi una partizione α è Φ -generante se ogni misurabile si scrive come $\cup_i (\cap_j \Phi^{n_{i,j}}(A_{i,j})^*)$ dove $A_{i,j} \in \alpha$ ed $n_{i,j} \in \mathbb{Z}$. Nel caso di partizioni Φ -generanti, la caratterizzazione degli elementi della σ -algebra generata è più semplice, infatti

Lemma 91 *Una partizione α è Φ -generante se il sottoinsieme $\alpha^{[\infty]}$ delle parti di M ottenuto unendo le partizioni $\alpha^{[N]} = \Phi^{-N}(\alpha) \vee \dots \vee \alpha \vee \dots \vee \Phi^N(\alpha)$ genera la σ -algebra.*

Se α è Φ -generante, allora ogni insieme U misurabile si scrive come $\sqcup_{i=1}^\infty C_i$ dove ogni C_i appartiene a $\alpha^{[N_i]}$ per qualche $N_i \in \mathbb{N}$.

Dimostrazione. Dimostriamo la prima affermazione. Ogni $\Phi^n(A)$ con $n \in \mathbb{Z}$ ed $A \in \alpha$ è unione di elementi di $\alpha^{[n]}$, e quindi appartiene alla σ -algebra generata da $\alpha^{[\infty]}$. Viceversa, ogni elemento di $\alpha^{[\infty]}$ è intersezione di elementi del tipo $\Phi^n(A)$, e quindi appartiene alla σ -algebra generata da tali elementi.

In base a quanto appena dimostato, sappiamo che la partizione $\alpha^{[\infty]}$ genera la σ -algebra, quindi ogni insieme misurabile U è unione numerabile di intersezioni di elementi di $\alpha^{[\infty]}$ o di loro complementari, ovvero $U = \cup(\cap C_{i,j}^*)$ dove $C_{i,j} \in \alpha^{[n_{i,j}]}$ e la stellina sta eventualmente per il complemento algebrico. Se la stellina rappresenta l'identità allora $C_{i,j}$ è l'intersezione $\cap_{\ell=-n_{i,j}}^{n_{i,j}} \Phi^\ell(A_\ell)$, oppure è unione di elementi così fatti, quindi $U = \cup_i (\cap_j (\cup_\ell C_{i,j,\ell}))$, con $C_{i,j,\ell}$ elemento di $\alpha^{[n_{i,j}]}$.

Basta ora far commutare l'intersezione con l'unione per osservare che U è del tipo $\cup_i (\cap_j C_{i,j})$. $\square$

Abbiamo quindi caratterizzato una famiglia (si spera non vuota) di partizioni che hanno caratteristiche di *esautività* relativamente alla misura. Dimostriamo ora che le partizioni generanti sono partizioni che realizzano il sup tra i valori di $h(\Phi, \alpha)$. Cominciamo con il dimostrare che, trattandosi di partizioni, gli insiemi misurabili si generano in modo più agile.

Lemma 92 *Se la partizione α è generante, allora per ogni $\varepsilon > 0$ e per ogni insieme misurabile A esiste $N \gg 0$ e $C_1, \dots, C_m$ in $\alpha^{[N]}$ e contenuti in A tali che $\mu(A - \sqcup C_i) < \varepsilon$.*

Dimostrazione. È conseguenza immediata di partizione generante. La famiglia delle partizioni $\Phi^j(\alpha)$ genera la σ -algebra se ogni insieme misurabile $A = \sqcup C_i$ con $C_i \in \alpha^{[N_i]}$. Segue che la serie $\sum \mu(C_i)$ converge a $\mu(A)$. Questo implica che per ogni ε esiste una troncata $\sqcup_{s=1}^N C_s$ che approssima da dentro A in misura, ovvero $\sqcup_{s=1}^N C_s \subset A$ e $\mu(A - \sqcup_{s=1}^N C_s) < C\varepsilon$. $\square$

Utilizziamo il lemma per dare un risultato di densità.

Lemma 93 *Sia α una partizione generante, allora l'insieme delle partizioni $\Gamma = \{\gamma \in \mathcal{P}(M) \mid \gamma \preceq \alpha^{[N]}, \exists N\}$ è denso nell'insieme delle parti finite di M per la distanza dist .*

Dimostrazione. Bisogna mostrare che comunque presa una partizione $\beta = \{B_1, \dots, B_n\}$ e comunque scelto un ε , esiste un γ in Γ tale che $\text{dist}(\beta, \gamma) < \varepsilon$.

Per il lemma appena dimostrato, per ogni atomo B_i esistono un N_i e C_i tali che C_i è unione di atomi in $\alpha^{[N_i]}$ e $\mu(B_i - C_i) < \varepsilon$. Segue quindi che la partizione $\gamma = \{C_1, C_2, \dots, C_n, C_{n+1} = M - \cup_{i=1}^n C_i\}$ è elemento di Γ (usando come N il sup degli N_i). Applicando le definizioni si ha che

$$\begin{aligned} \text{dist}(\beta, \gamma) &= \eta(\beta|\gamma) + \eta(\gamma|\beta) = 2\eta(\beta \vee \gamma) - \eta(\beta) - \eta(\gamma) = \\ &= 2 \sum_{i=1}^n \rho(\mu(C_i)) + 2 \sum_{i=1}^n \rho(\mu(B_i \cap C_{n+1})) - \sum_{i=1}^{n+1} \rho(\mu(C_i)) - \sum_{i=1}^n \rho(\mu(B_i)) \leq \\ &\leq (2n+1)\rho(\mu(C_{n+1})) + \sum_{i=1}^n \left(\rho(\mu(C_i)) - \rho(\mu(B_i)) \right) \end{aligned}$$

Usando con cura la continuità di ρ , si conclude che $\text{dist}(\gamma, \beta) < C\varepsilon$ con C indipendente da γ . $\square$

Siamo infine pronti a dimostrare il teorema sull'entropia e le partizioni generanti

Teorema 94 *Se α è partizione generante, allora $h(\Phi, \alpha) = h(\Phi)$.*

Dimostrazione. Dimostriamo che, presa α una partizione generante e β una partizione qualsiasi, vale la disuguaglianza $h(\Phi, \alpha) \geq h(\Phi, \beta)$.

Se β appartiene a Γ , allora la disuguaglianza vale perchè

$$\begin{aligned} h(\Phi, \beta) &\leq h(\Phi, \alpha^{[N]}) = \lim_{N' \rightarrow \infty} \frac{1}{N'} \eta(\alpha^{[N]} \vee \dots \Phi^{-N'}(\alpha^{[N]})) \\ &= \lim_{N' \rightarrow \infty} \frac{1}{N'} \eta(\Phi^{-N-N'}(\alpha) \vee \dots \Phi^N(\alpha)) \\ &= \lim_{N' \rightarrow \infty} \frac{2N + N'}{N'} \frac{1}{2N + N'} \eta(\Phi^{-2N-N'}(\alpha) \vee \dots \alpha) \\ &= h(\Phi, \alpha). \end{aligned}$$

Per densità, la disuguaglianza rimane vera anche se β è una qualsiasi partizione finita. $\square$

Bibliografia

Una introduzione alle forme differenziali con impostazione geometrica si trova in [BT] R. Bott e L.W. Tu, *Differential Forms in Algebraic Topology*, Springer-Verlag (1991).

Come la metrica di $\mathbb{R}^n$ permetta di definire integrali m -dimensionali, e varie ed eventuali di teoria dell'integrazione sono spiegate in [DM] G. De Marco, *Analisi Due/2*, Zanichelli (1993).

Gli esempi del capitolo 1, ed i capitoli 2,3 e 4, sono una rielaborazione delle note [B] G. Benettin, *Capitolo I, "Introduzione alla teoria Ergodica"* e si trovano in rete all'indirizzo

www.math.unipd.it/~benettin/.

Il capitolo 5 è invece una rielaborazione del libro [G] G. Gallavotti, *Aspetti della teoria ergodica, qualitativa e statistica del moto*, Quaderni dell'Unione Matematica Italiana, Pitagora Editrice (1981). Una versione tradotta in Inglese e scritta da G.G. in collaborazione con F. Bonetto e G. Gentile si intitola *Aspects of the ergodic, qualitative and statistical properties of motion* e si trova in rete all'indirizzo

<http://141.108.20.11/~giovanni/glib.html#E>.

Infine, un libro interessante da leggere è [HK] B. Hasselblatt & A. Katok, *A first course in dynamics, with a panorama of recent developments*, Cambridge University Press (2003).