Università degli Studi di Pisa
Dipartimento di Matematica
Corso di Laurea Magistrale in Matematica

Note Sintetiche di Teoria Ergodica

ispirate dalle lezioni del *Professor Stefano Galatolo*

Marco Tarsia
mtarsia@student.dm.unipi.it

Novembre 2015

Introduzione

Lo scopo principale di questo lavoro è quello di arrivare gradualmente a compiere un primo possibile passo nella direzione dello studio di "eventi rari", come ad esempio potrebbe esser lo studio di un sistema aperto, cioè avente un buco, dal quale sia "davvero improbabile" uscire. Dunque saranno generalmente presenti delle discontinuità nel sistema, e per ciò potrebbe risultare perdente uno studio mediante la teoria classica della perturbazione come quella della linear response.

Quello che invece proponiamo è un teorema astratto per così dire a sé stante - dovuto a G. Keller e C. Liverani, ed anche piuttosto recente - il quale ha indubbiamente elevata potenzialità d'utilizzo, ma che noi sfruttiamo più che altro per ricavare una sorta di *legge di occorrenza di eventi rari* per particolari sistemi "poco caotici".

Il nostro atteggiamento nel corso di tale applicazione comunque astratta sarà in verità piuttosto superficiale rispetto al teorema, nel senso che daremo volutamente per buona la complicata discussione sulla relativa verifica delle numerose ipotesi del teorema stesso. Tuttavia questo resta coerente col fatto di voler compiere solo un primo passo nella teoria, oltre che col carattere appunto sintetico dell'intero elaborato: definizioni ed enunciati ben ordinati e precisi, ma senza dimostrazioni.

Esporremo il tutto in tre sezioni, delle quali le prime due pensate appositamente per arrivare il meglio possibile alla terza, e ciò con l'usuale linguaggio della teoria ergodica, rispettando quasi pienamente la nomenclatura tradizionale di definizioni e risultati consolidati, e dove pure le notazioni simboliche coincideranno all'incirca con quelle comunemente adottate.

I riferimenti bibliografici relativi alle prime due sezioni sono [3] e [4] rispettivamente, mentre quelli relativi alla terza sezione sono [1] e [2].

1 Teoria Ergodica Classica

1.1 Primi Elementi

Sia X un insieme, sia $\mathcal{B}$ una σ-algebra su X, sia $T : X \to X$ $\mathcal{B}$-misurabile e sia μ una misura di probabilità su $\mathcal{B}$. La situazione che consideriamo è quella di un sistema dinamico discreto (X, T)

che sia al tempo stesso uno spazio di probabilità $(X, \mathcal{B}, \mu)$, situazione che particolareggiamo subito mediante un concetto di compatibilità naturale fra i due.

Definizione 1 (*Probabilità Invariante*)**.** Diciamo che la probabilità μ è *T-invariante* se $\mu(T^{-1}(E)) = \mu(E)$ per ogni $E \in \mathcal{B}$ (ovvero se $\mu(T^{-n}(E)) = \mu(E)$ per ogni $n \in \mathbb{N}$ e $E \in \mathcal{B}$). In tal caso, possiamo anche dire che viceversa T è una *trasformazione che preserva la probabilità* μ, per cui possiamo scrivere sinteticamente che il sistema (T, μ) *è p.p.t.* (acronimo appunto di *probability preserving transformation*).

Osservazione 1. Indichiamo con μ_T la probabilità immagine di μ mediante T, ossia la probabilità su $\mathcal{B}$ definita ponendo $\mu_T(E) := \mu(T^{-1}(E))$ per ogni $E \in \mathcal{B}$. Per definizione, quindi, la probabilità μ è T-invariante se e solo se $\mu_T = \mu$.

Nota. Possiamo ovviamente immaginare che sia sempre $\mathcal{B} \neq \{\emptyset, X\}$ (per cui $\#X \geq 2$).

Esercizio 1. Dimostrare che $\int_X \varphi d\mu_T = \int_X \varphi(T) d\mu$ per ogni $\varphi \in L^\infty(\mu)$ e di conseguenza che, se (T, μ) è p.p.t., allora $\int_X f(T) d\mu = \int_X f d\mu$ per ogni $f \in L^1(\mu)$.

Osservazione 2. Ricordando che una combinazione convessa di mappe reali è una loro combinazione lineare finita a coefficienti reali non negativi di somma totale unitaria, risulta evidente che una combinazione convessa di probabilità su $\mathcal{B}$ resta una probabilità su $\mathcal{B}$, ed in particolare che una combinazione convessa di probabilità su $\mathcal{B}$ T-invarianti resta una probabilità su $\mathcal{B}$ T-invariante.

Osservazione 3. Se (T, μ) è p.p.t., allora μ è concentrata su $T^n(X)$ per ogni $n \in \mathbb{N}$, cioè $\mu(E) = 0$ per ogni $E \in \mathcal{B}$ tale che $E \subseteq X \setminus T^n(X)$ quale che sia $n \in \mathbb{N}$ (o equivalentemente $n \in \mathbb{N} \setminus \{0\}$).

Notazioni. In seguito useremo i due simboli $A^c := X \setminus A$, per $A \subseteq X$, e $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$.

Osservazione 4. Esempio speciale di sistema dinamico probabilizzabile $(X, \mathcal{B}, T)$ è dato da X spazio topologico e $\mathcal{B}$ la σ-algebra di Borel su X, e quindi magari con T continua. Per alcuni sistemi di questo genere, infatti, risulta automatica l'esistenza di una probabilità su $\mathcal{B}$ che sia T-invariante (mentre l'unicità, che in generale non vale, va naturalmente studiata di caso in caso).

Teorema 1 (Krylov-Bogolyubov)**.** *Sia X uno spazio metrico compatto, sia $\mathcal{B}$ la σ-algebra di Borel su X e sia $T : X \to X$ continua: allora esiste una probabilità μ su $\mathcal{B}$ che sia T-invariante.*

Teorema 2 (di Ricorrenza di Poincaré)**.** *Se (T, μ) è p.p.t. e se $E \in \mathcal{B}$, allora per μ-q.o. $x \in E$ esiste $N \in \mathbb{N}^*$ tale che $T^N(x) \in E$.*

Corollario 1 (del teorema di ricorrenza di Poincaré)**.** *Se (T, μ) è p.p.t. e se $E \in \mathcal{B}$, allora*

$$\sum_{n=1}^{\infty} \mu(E \cap T^{-1}(E^c) \cap \ldots \cap T^{-(n-1)}(E^c) \cap T^{-n}(E)) = \mu(E).$$

1.2 Ergodicità e Mixing

Definizioni 1 (*Osservabile. Osservabile Invariante. Evento Invariante*)**.** Chiamiamo una funzione $f : X \to \mathbb{R}$ misurabile come *osservabile di X*. Diciamo che un osservabile f di X è un *osservabile T-invariante* se $f(T) = f$ μ-q.c.. In particolare, per ogni $E \in \mathcal{B}$, $\mathbb{1}_E$ è T-invariante se e solo se $\mu(T^{-1}(E) \triangle E) = 0$. Nel caso sia addirittura $T^{-1}(E) = E$, diciamo piuttosto che E è un *evento T-invariante*. Ciò equivale chiaramente ad avere che E^c sia un evento T-invariante.

Definizione 2 (*Sistema Ergodico*)**.** Il sistema (T, μ) è *ergodico* se, per ogni $E \in \mathcal{B}$ T-invariante, è o $\mu(E) = 0$ o $\mu(E^c) = 0$ (ovvero $\mu(E) \in \{0, 1\}$). A parole, un sistema è ergodico se la dinamica coinvolge veramente quasi tutto lo spazio.

Proposizione 1. *Se (T, μ) è p.p.t., allora le seguenti tre condizioni sono equivalenti.*

(a) *(T, μ) è ergodico.*

(b) *Per ogni $E \in \mathcal{B}$ con $\mu(T^{-1}(E) \triangle E) = 0$, è $\mu(E) \in \{0, 1\}$.*

(c) *Ogni osservabile T-invariante è μ-q.c. costante.*

Nota. Passare da eventi di $\mathcal{B}$ ad osservabili di X non è semplicemente una generalizzazione fine a sé, bensì contribuirà al passaggio dall'approccio classico alla teoria ergodica a quello analitico-funzionale.

Corollario 2 (del teorema di ricorrenza di Poincaré)**.** *Se (T, μ) è p.p.t. ed ergodico, e se $E \in \mathcal{B}$ ha $\mu(E) > 0$, allora per μ-q.o. $x \in X$ esiste $N \in \mathbb{N}^*$ tale che $T^N(x) \in E$.*

Osservazione 5. Se (T, μ) è p.p.t. e se $E \in \mathcal{B}$, allora definiamo $\tau_E : X \to \mathbb{N}^* \cup \{\infty\}$ ponendo $\tau_E(x) \coloneqq \inf\{n \in \mathbb{N}^* | T^n(x) \in E\}$ per ogni $x \in X$, intendendo al solito che $\inf \emptyset = \infty$: così τ_E è un osservabile di X e possiamo chiamarlo come *il primo istante di scontro su E* (tramite T).
Grazie al teorema 2, è sicuramente $\tau_{E|_E} < \infty$ μ-q.c., mentre se (T, μ) fosse pure ergodico e se $\mu(E) > 0$, allora globalmente $\tau_E < \infty$ μ-q.c. grazie invece al corollario 2.

Teorema 3 (Kac)**.** *Se (T, μ) è p.p.t. ed ergodico, e se $E \in \mathcal{B}$ ha $\mu(E) > 0$, allora è $\int_E \tau_E d\mu = 1$. In altri termini, se indichiamo con μ_E la probabilità indotta da μ sulla traccia $\mathcal{B}_E$ di $\mathcal{B}$ su E, ossia se $\mu_E(A) \coloneqq \frac{\mu(A)}{\mu(E)}$ per ogni $A \in \mathcal{B}$ tale che $A \subseteq E$, allora $\int_E \tau_E d\mu_E = \frac{1}{\mu(E)}$.*

Definizione 3 (*Sistema Mixing*)**.** Il sistema (T, μ) è *mixing* se $\mu(A \cap T^{-n}(B)) \xrightarrow[n\to\infty]{} \mu(A)\mu(B)$ per ogni $A, B \in \mathcal{B}$. Dunque, un sistema mixing è provvisto di un concetto debole d'indipendenza asintotica fra eventi.

Osservazione 6. Un sistema (T, μ) mixing è anche ergodico, e di conseguenza la precedente proposizione 1 descrive parzialmente i sistemi che siano p.p.t. e mixing. Per una caratterizzazione di tali sistemi possiamo di nuovo passare dagli osservabili di X, ma a patto di tirare in gioco lo spazio di Hilbert reale $(L^2(\mu), \|\cdot\|_{L^2(\mu)})$.

Proposizione 2. *Se (T, μ) è p.p.t., allora $\|f(T)\|_{L^2(\mu)} = \|f\|_{L^2(\mu)}$ per ogni $f \in L^2(\mu)$.*

Definizione 4 (*Covarianza*)**.** Definiamo la funzione *covarianza* $\mathbf{Cov} \coloneqq \mathbf{Cov}^\mu : L^2(\mu) \times L^2(\mu) \to \mathbb{R}$ ponendo $\mathbf{Cov}[f, g] \coloneqq \int_X fg d\mu - \int_X f d\mu \int_X g d\mu$ per ogni $f, g \in L^2(\mu)$.

Proprietà basilari (della covarianza)**.**

(1) $\mathbf{Cov}[f, g] = \int_X (f - \int_X f d\mu)(g - \int_X g d\mu) d\mu$ per ogni $f, g \in L^2(\mu)$.

(2) $\mathbf{Cov}$ è un'applicazione bilineare (e simmetrica).

(3) $\mathbf{Cov}[f, c] = 0$ per ogni $f \in L^2(\mu)$ e $c \in \mathbb{R}$.

(4) $\mathbf{Cov}[f, g] = \langle f, g \rangle_{L^2(\mu)}$ per ogni $f, g \in L^2(\mu)$ con $\int_X g d\mu = 0$.

(5) $\mathbf{Cov}[\mathbb{1}_A, \mathbb{1}_B] = \mu(A \cap B) - \mu(A)\mu(B)$ per ogni $A, B \in \mathcal{B}$.

(6) $|\mathbf{Cov}[f,g]| \leq 4 \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$ per ogni $f, g \in L^2(\mu)$.

Proposizione 3. *Se (T,μ) è p.p.t., allora le seguenti due condizioni sono equivalenti.*

(a) *(T,μ) è mixing.*

(b) *Per ogni $f, g \in L^2(\mu)$, $\mathbf{Cov}[f, g(T^n)] \xrightarrow[n\to\infty]{} 0$.*

Nota. Grazie alla precedente caratterizzazione mediante osservabili di X, più che alla definizione di per sé, possiamo pensare un sistema p.p.t. e mixing come ad un sistema più o meno mixing a seconda appunto della velocità ad annullarsi delle covarianze, anche solo su una classe più ristretta di osservabili. In questo modo, possiamo definire lo sfuggevole concetto di *caos* definendone il concetto opposto: proprio quello di *altamente mixing*.

Esercizio 2. Per ogni $f \in L^1(\mu)$ e $\varphi \in L^\infty(\mu)$, definiamo $\mathbf{Cov}[f,\varphi] \coloneqq \int_X f\varphi d\mu - \int_X f d\mu \int_X \varphi d\mu$ ed osserviamo in particolare che $|\mathbf{Cov}[f,\varphi]| \leq 4 \|f\|_{L^1(\mu)} \|\varphi\|_{L^\infty(\mu)}$. Dimostrare che, se (T,μ) è p.p.t., allora (T,μ) è mixing se e solo se $\mathbf{Cov}[f, \varphi(T^n)] \xrightarrow[n\to\infty]{} 0$ per ogni $f \in L^1(\mu)$ e $\varphi \in L^\infty(\mu)$.

1.3 Teoremi Ergodici

Insistendo con l'idea che T sia data mentre μ T-invariante sia misteriosa e da trovare, ammesso che esista, chiediamoci se sia possibile dedurla per via statistica: se sia possibile cioè distribuirla su X in base a come T distribuisca i punti di X stesso (cosa dalla quale si dovrebbe dedurre l'eventuale ergodicità del sistema senza appunto conoscere μ). Le risposte matematiche a domande del genere sono i *teoremi ergodici*, e noi ne enunciamo uno piuttosto generale che però rienunciamo in un caso più particolare, anche perché la relativa dimostrazione si rivelerebbe elementare ma brillante. Sottolineiamo comunque che questa sottosezione sì completa i discorsi precedenti, ma che al tempo stesso non servirà direttamente per il seguito.

Notazione. Per ogni osservabile f di X, poniamo $S_n^f \coloneqq \frac{1}{n+1}\sum_{k=0}^n f(T^k)$ per ogni $n \in \mathbb{N}$.

Osservazione 7. Per ogni osservabile f di X e per ogni $n \in \mathbb{N}$, anche S_n^f è un osservabile di X. Inoltre, per ogni $p \in [1,\infty]$, $S_n^f \in L^p(\mu)$ per ogni $n \in \mathbb{N}$ se e solo se $f \in L^p(\mu)$. In particolare, se (T,μ) è p.p.t. e se $f \in L^1(\mu)$, allora $\int_X S_n^f d\mu = \int_X f d\mu$ per ogni $n \in \mathbb{N}$. Infine f è T-invariante se e solo se $S_n^f = f$ μ-q.c. per ogni $n \in \mathbb{N}$.

Osservazione 8. Per ogni $E \in \mathcal{B}$ e per ogni $n \in \mathbb{N}$, essendo $\sum_{k=0}^n \mathbb{1}_E(T^k) = \sum_{k=0}^n \mathbb{1}_{T^{-k}(E)} = \#\{k = 0,\dots,n | T^k \in E\}$, è $S_n^{\mathbb{1}_E}(x) = \frac{\#\{k=0,\dots,n|T^k(x)\in E\}}{n+1}$ per ogni $x \in X$, e ciò coincide con la frequenza di scontri su E da parte di x (tramite T) degli $n+1$ possibili.

Teorema 4 (Ergodico di Birkhoff). *Se (T,μ) è p.p.t. e se $f \in L^1(\mu)$, allora esiste $\bar{f} \in L^1(\mu)$ che sia T-invariante e tale che $S_n^f \xrightarrow[n\to\infty]{\mu-q.c.} \bar{f}$. In particolare, se (T,μ) fosse pure ergodico, allora sarebbe $\bar{f} = \int_X f d\mu$ μ-q.c..*

Corollario 3 (di Birkhoff). *Se (T,μ) è p.p.t. ed ergodico e se $E \in \mathcal{B}$, allora $\frac{\#\{k=0,\dots,n|T^k(x)\in E\}}{n+1} \xrightarrow[n\to\infty]{} \mu(E)$ per μ-q.o. $x \in X$.*

Teorema 5 (Ergodico di Von-Neumann)**.** *Se (T, μ) è p.p.t. e se $f \in L^2(\mu)$, allora esiste $\bar{f} \in L^2(\mu)$ che sia T-invariante e tale che $S_n^f \xrightarrow[n\to\infty]{L^2(\mu)} \bar{f}$. In particolare, se (T, μ) fosse pure ergodico, allora sarebbe $\bar{f} = \int_X f d\mu$ μ-q.c..*

Corollario 4 (di Von-Neumann)**.** *Se (T, μ) è p.p.t. ed ergodico, allora è anche debolmente mixing nel senso che $\frac{1}{n+1} \sum_{k=0}^{n} \mu(A \cap T^{-k}(B)) \xrightarrow[n\to\infty]{} \mu(A)\mu(B)$ per ogni $A, B \in \mathcal{B}$.*

Nota. Riconosciamo immediatamente che vale pure il viceversa del corollario 4: precisamente, se (T, μ) è debolmente mixing, allora è ergodico.

Esercizio 3. Dimostrare che un sistema mixing è anche debolmente mixing.

2 Approccio Analitico-funzionale alla Teoria Ergodica

2.1 Operatore di Trasferimento

Insistiamo giustamente con l'approccio statistico piuttosto che puntuale, chiedendoci precisamente come la dinamica vada trasformando una certa densità di massa fissata.

Sia quindi X uno spazio metrico, sia $\mathcal{B}$ la σ-algebra di Borel su X e sia $T : X \to X$ $\mathcal{B}$-misurabile. Indicando con $SM(X)$ lo spazio vettoriale reale delle misure reali finite con segno "su X" (intendendo su $\mathcal{B}$), e continuando ad indicare per ogni $\mu \in SM(X)$ con μ_T la misura immagine di μ mediante T su X definita ponendo $\mu_T(E) := \mu(T^{-1}(E))$ per ogni $E \in \mathcal{B}$, abbiamo allora che $\mu_T \in SM(X)$ e che la mappa $SM(X) \to SM(X)$, $\mu \mapsto \mu_T$, è un'applicazione lineare e cioè un operatore di $SM(X)$.

Sia adesso $m \in SM(X) \setminus \{0\}$ con $m(X) = 1$ che per noi sia come di riferimento, nel senso che per noi, fra tutte le misure $\mu \in SM(X)$, siano interessanti tutte e sole quelle assolutamente continue rispetto a m stessa (ovvero tali che, per ogni $E \in \mathcal{B}$, se $m(E) = 0$ allora pure $\mu(E) = 0$): in simboli, le $\mu << m$.

Supponiamo dunque che T sia *non singolare* rispetto a m nel senso che, per ogni $E \in \mathcal{B}$, $m(T^{-1}(E)) = 0$ se e solo se $m(E) = 0$ (nonostante che in generale m non sia T-invariante): in questo modo abbiamo sicuramente che, per ogni $\mu \in SM(X)$, se $\mu << m$ allora $\mu_T << m$.

Ora, grazie ad un teorema di Radon-Nikodym, possiamo dire di esser interessati a tutte e sole le $\mu \in SM(X)$ per le quali esista $f \in L^1(m)$ tale che $\mu(E) = \int_E f dm$ per ogni $E \in \mathcal{B}$ (e così f è la densità di μ rispetto a m): in simboli, tale che $\mu = f dm$.

Volendo quindi considerare, più che l'intero operatore di $SM(X)$ sopra definito, la sua restrizione alle $\mu \in SM(X) << m$, possiamo equivalentemente considerare l'operatore $L : L^1(m) \to L^1(m)$ di $L^1(m)$ definito ponendo $L(f) := \frac{(f dm)_T}{dm}$ per ogni $f \in L^1(m)$ e chiamato *operatore di trasferimento* (o *di Perron-Frobenius*).

Per definizione, L semplicemente associa alla densità di ogni $\mu \in SM(X) << m$ quella di μ_T (rispetto a m per entrambe): in particolare, vale evidentemente che $\int_E L(f) dm = \int_{T^{-1}(E)} f dm$ per ogni $f \in L^1(m)$ e $E \in \mathcal{B}$ (da cui giustamente $\int_X L(f) dm = \int_X f dm$).

Proprietà basilari (di L)**.**

(1) L è un operatore positivo: per ogni $f \in L^1(m)$, se $f \geq 0$ m-q.c., allora pure $L(f) \geq 0$ m-q.c..

(2) Per ogni $f \in L^1(m)$, $|L(f)| \leq L(|f|)$ m-q.c..

(3) L è un operatore continuo e di norma unitaria: per ogni $f \in L^1(m)$, $\|L(f)\|_{L^1(m)} \leq \|f\|_{L^1(m)}$ e $\|L\|_{\mathcal{L}(L^1(m))} = 1$.

(4) Per ogni $f \in L^1(m)$, $L(f)$ è l'unico elemento di $L^1(m)$ tale che sia $\int_X \varphi L(f) dm = \int_X \varphi(T) f dm$ per ogni $\varphi \in L^\infty(m)$.

(5) Per ogni $f \in L^1(m)$ e $g \in L^\infty(m)$, è $L(g(T)f) = gL(f)$ m-q.c..

(6) Se (T,m) è p.p.t., allora $L(f) \circ T = \mathbf{E}^m[f|\sigma(T)]$ m-q.c. per ogni $f \in L^1(m)$.

(7) Se (T,m) è p.p.t., allora $|\mathbf{Cov}^m[f, \varphi(T^n)]| \leq \|\varphi\|_{L^\infty(m)} \left\|L^n(f) - \int_X f dm\right\|_{L^1(m)}$ per ogni $n \in \mathbb{N}$, $f \in L^1(m)$ e $\varphi \in L^\infty(m)$.

Proposizione 4.

(1) *Se esistono* $f \in L^1(m) \setminus \{0\} \geq 0$ *m-q.c. e* $h \in L^1(m)$ *tali che* $L^n(f) \xrightarrow[n\to\infty]{L^1(m)} h \int_X f dm$, *allora* $h \geq 0$ *m-q.c. con* $\int_X h dm = 1$ *ed è tale che* $L(h) = h$ *m-q.c., ossia* (T, hdm) *è p.p.t..*

(2) *Se* (T, hdm) *è p.p.t. e se* $L^n(f_0) \xrightarrow[n\to\infty]{L^1(m)} 0$ *per ogni* $f_0 \in L^1(m)$ *con* $\int_X f_0 dm = 0$, *allora* (T, hdm) *è mixing.*

(3) *Se* $h \in L^1(m)$ *è tale che* $L^n(f) \xrightarrow[n\to\infty]{L^1(m)} h \int_X f dm$ *per ogni* $f \in L^1(m)$, *allora* (T, hm) *è p.p.t. e mixing.*

Nota. Per ogni $(f_n)_{n\in\mathbb{N}}$, f in $L^1(m)$, $f_n \xrightarrow[n\to\infty]{L^1(m)} f$ se e solo se $\int_X \varphi f_n dm \xrightarrow[n\to\infty]{} \int_X \varphi f dm$ per ogni $\varphi \in L^\infty(m)$.

2.2 Quasi-compattezza e Spectral Gap

Per studiare il comportamento asintotico di $L^n(f)$ (in n), almeno per una classe più ristretta di $f \in L^1(m)$, risulta indubbiamente potente il metodo che trattiamo in seguito.

Notazioni. Sia $(V, \|\cdot\|_V)$ uno spazio di Banach reale e sia $P \in (\mathcal{L}(V), \|\cdot\|_{\mathcal{L}(V)}) \setminus \{0\}$, e poniamo $\|\cdot\| := \|\cdot\|_{\mathcal{L}(V)}$. Ricordiamo che un autovalore di P è un $\lambda \in \mathbb{R}$ in corrispondenza del quale esiste un $\varphi \in V \setminus \{0\}$ tale che $P(\varphi) = \lambda\varphi$; viceversa, ogni $\varphi \in V \setminus \{0\}$ avente tale proprietà è un λ-autovettore di P. Il sottospazio di V formato dallo 0 di V e dagli λ-autovettori di P, e cioè $Ker(\lambda I_V - P)$, è il λ-autospazio di P. Un autovalore λ di P è semplice per P se $dim(Ker(\lambda I_V - P)) = 1$. Notiamo in particolare che un 1-autovettore di P è semplicemente un punto fisso di P. Indicando ora con $\sigma(P) := \{z \in \mathbb{R} | zI_V - P \text{ non ha inversa limitata}\}$ lo spettro di P - per cui ogni autovalore di P sta in $\sigma(P)$ - ricordiamo che è $\sigma(P) \subseteq [-\|P\|, \|P\|]$ e che $\sigma(P)$ è chiuso (ovvero compatto). Indicando infine con $\rho(P) := \sup_{z\in\sigma(P)} |z|$ il raggio spettrale di P - per cui appunto $\rho(P) \leq \|P\|$ - ricordiamo la fondamentale doppia formula di Gelfand $\rho(P) = \lim_{n\to\infty} \sqrt[n]{\|P^n\|} = \inf_{n\in\mathbb{N}^*} \sqrt[n]{\|P^n\|}$, dalla quale deduciamo in particolare che, per ogni $\varepsilon > 0$, esistono $C := C_\varepsilon \in (0, \infty)$ e $n_\varepsilon \in \mathbb{N}^*$ tale che, per ogni $n \in \mathbb{N}^* \geq n_\varepsilon$ e $\psi \in V \setminus \{0\}$, sia $\frac{\|P^n(\psi)\|_V}{\|\psi\|_V} \leq Ce^{n\varepsilon}\rho(P)^n$.

Esercizio 4 (sull'operatore di trasferimento L)**.** Sia (T, hdm) p.p.t., e consideriamo le seguenti cinque condizioni.

(a) h è > 0 m-q.c. e tale che $h, \frac{1}{h} \in L^\infty(m)$ e (T, hdm) sia mixing.

(b) $L^n(f_0) \xrightarrow[n\to\infty]{L^1(m)} 0$ per ogni $f_0 \in L^1(m)$ con $\int_X f_0 dm = 0$.

(c) $L^n(f) \xrightarrow[n\to\infty]{L^1(m)} h \int_X f dm$ per ogni $f \in L^1(m)$.

(d) Esiste uno ed un solo autovalore λ di L avente modulo 1 ed è proprio $\lambda = 1$, per cui $\rho(L) = 1$, ed inoltre tale autovalore è semplice per L.

(e) h è l'unica densità rispetto a m che abbia $\int_X hdm = 1$ e tale che hdm sia T-invariante.

Dimostrare che $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e)$.

Definizione 5 (*Spectral Gap*)**.** P ha *spectral gap* se esistono $\lambda \in \mathbb{R}$ e $R, Q \in \mathcal{L}(V)$ tali che

[a] $P = \lambda R + Q$;

[b] $R^2 = R$ e $dim(Im(R)) = 1$ (ovvero R è una proiezione di V su una retta di V stesso);

[c] $\rho(Q) < |\lambda|$;

[d] $Q(R) = R(Q) = 0$.

Proprietà basilari (di P avente spectral gap)**.**

(1) λ è un autovalore di P ed ogni elemento di $Im(R) \setminus \{0\}$ è un λ-autovettore di P.

(2) Per ogni $n \in \mathbb{N}^*$, $P^n = \lambda^n R + Q^n$.

(3) Per ogni $\psi \in V$, $\|Q^n(\psi)\|_V = o_{n\to\infty}(|\lambda^n|)$.

(4) Per ogni $\psi \in V$, $(\lambda^{-1}P)^n(\psi) \xrightarrow[n\to\infty]{V} R(\psi)$ con velocità esponenziale.

(5) Vale il viceversa della proprietà (1), ovvero che ogni λ-autovettore di P appartiene a $Im(R) \setminus \{0\}$. Riassumendo, $Ker(\lambda I_V - P) = Im(R)$ e per ciò λ è semplice per P.

(6) Esiste $\gamma_0 \in (0, \infty)$ tale che $\sigma(P) \subseteq \{\lambda\} \cup \{z \in \mathbb{R} | \, |z| \leq e^{-\gamma_0} |\lambda|\}$, e di conseguenza $\rho(P) = |\lambda|$.

Definizione 6 (*Quasi-compattezza*)**.** P è *quasi-compatto* se $\rho(P) > 0$ e se esistono $F, H < V$ ed un $\rho \in (0, \rho(P))$ tali che

[a] $V = F \oplus H$;

[b] F, H sono chiusi e P-invarianti (nel senso che $P(F) < F$, $P(H) < H$);

[c] $dim(F) < \infty$ ed ogni autovalore di $P_{|_F} : F \to F$ ha modulo $> \rho$ (per cui $\rho(P_{|_F}) > \rho$);

[d] $\rho(P_{|_H}) < \rho$.

Proposizione 5. *Se P è quasi-compatto, se esiste uno ed un solo autovalore $\bar{\lambda}$ di P avente $|\bar{\lambda}| = \rho(P)$, e se tale $\bar{\lambda}$ è semplice per P, allora P ha spectral gap (con $\lambda = \bar{\lambda}$).*

Dunque la quasi-compattezza risulta una fra le condizioni sufficienti per avere lo spectral gap. Enunciamo quindi un importante teorema avanzato che dìa criteri su come ottenere proprio la quasi-compattezza, la quale può esser vista così come "ponte" tra le assunzioni di tale teorema e la condizione di spectral gap.

Teorema 6 (Hennion). *Se esiste una seminorma $\|\cdot\|' : V \to \mathbb{R}$ su V tale che*

(a) $\|\cdot\|'$ *è continua su* V*;*

(b) *per ogni* $(f_n)_{n\in\mathbb{N}}$ *in* V *con* $\sup_{n\in\mathbb{N}} \|f_n\|_V < \infty$*, esiste una sottosuccessione* $n_k \xrightarrow[k\to\infty]{} \infty$ *ed esiste* $f \in V$ *tali che* $\|P(f_{n_k}) - f\|' \xrightarrow[k\to\infty]{} 0$*;*

(c) *esiste* $M \in (0,\infty)$ *tale che* $\|P(f)\|' \leq M \|f\|'$ *per ogni* $f \in V$*;*

(d) *esistono* $N \in \mathbb{N}^*$*,* $r \in (0, \rho(P))$ *e* $D \in (0,\infty)$ *tali che* $\left\|P^N(f)\right\|_V \leq r^N \|f\|_V + D\|f\|'$ *per ogni* $f \in V$*;*

allora P *è quasi-compatto.*

Nota. La disuguaglianza in (d), conosciuta come *disuguaglianza di Lasota-Yorke* (o *di Doeblin-Fortet*), resterebbe vera nella stessa forma per ogni $n \in \mathbb{N}^* \geq N$, e con $r \in (0,1)$, se solo fosse $\|P\| \leq 1$.

Corollario 5 (della proposizione 5). *Nel caso speciale che sia* $V < L^1(m)$ *chiuso contenente le costanti con precisamente* $\|\cdot\|_{L^1(m)} \leq cost \|\cdot\|_V$ *su* V *e tale che* $L(V) < V$*, consideriamo* $P := L_{|V}$*. Se* (T, hdm) *è p.p.t. e se*

[a] $h \in V$ *(oppure* $h\mathbb{1}_E \in V$ *per ogni* $E \in \mathcal{B}$*);*

[b] $L^n(f_0) \xrightarrow[n\to\infty]{L^1(m)} 0$ *per ogni* $f_0 \in V$ *con* $\int_X f_0 dm = 0$*;*

[c] P *è quasi compatto;*

allora P *ha spectral gap con* $\lambda = 1$ *e* $R(f) = h\int_X f dm = hm(f)$ *per ogni* $f \in V$*, quindi* $P = hm + Q$*. Di conseguenza, segue che* $L^n(1) \xrightarrow[n\to\infty]{V} h$ *con velocità esponenziale.*

Nota. Possiamo sostituire l'ipotesi [c] di quasi-compattezza di P con opportune condizioni sufficienti per ottenere [c] stessa, naturalmente più esplicite e prese dal teorema 6. Per questo possiamo prendere $\|\cdot\|' := \|\cdot\|_{L^1(m)}$ su V e supporre ad esempio che P sia compatto e che soddisfi una disuguaglianza di Lasota-Yorke.

Applicazione. Se $X := [0,1]$, se m è la misura di Lebesgue sui boreliani di $[0,1]$, e se $T : [0,1] \to [0,1]$ è monotòna $\mathcal{C}^1$ a tratti ed espandente, cioè con $|T'| > 1$ "sui tratti", allora T è non singolare rispetto a m e soprattutto, se $V := BV([0,1])$ è lo spazio delle funzioni $f : [0,1] \to \mathbb{R}$ a variazione limitata su $[0,1]$, allora $L(V) < V$ e P è quasi-compatto grazie ad un teorema di Rychlik. Più precisamente, in questo caso vale una Lasota-Yorke con $\|\cdot\|' := \|\cdot\|_{L^1(m)}$ su V e vera per ogni $N \in \mathbb{N}^*$.

3 Un Risultato Astratto di Perturbazione

3.1 Il Teorema di Keller-Liverani

Sia V uno spazio vettoriale reale normato dalle due norme $\|\cdot\|_V \leq \|\cdot\|_V^*$, e sia $(V', \|\cdot\|_{V'})$ il suo spazio duale. Se indichiamo con $(\mathcal{L}(V), \|\cdot\|) \coloneqq (\mathcal{L}((V, \|\cdot\|_V), (V, \|\cdot\|_V)), \|\cdot\|_{\mathcal{L}((V,\|\cdot\|_V),(V,\|\cdot\|_V))})$ e con $(\mathcal{L}(\|\cdot\|_V^*, \|\cdot\|_V), \|\cdot\|^*) \coloneqq (\mathcal{L}((V, \|\cdot\|_V^*), (V, \|\cdot\|_V)), \|\cdot\|_{\mathcal{L}((V,\|\cdot\|_V^*),(V,\|\cdot\|_V))})$, allora chiaramente $\|\cdot\|^* \leq \|\cdot\|$ e quindi $\mathcal{L}(V) < \mathcal{L}(\|\cdot\|_V^*, \|\cdot\|_V)$. A proposito, indichiamo fin da subito $I \coloneqq I_V$.

Noi siamo interessati a degli $P \in \mathcal{L}(V) \setminus \{0\}$ per i quali esistano $\lambda \in \mathbb{R} \setminus \{0\}$, $\varphi \in V \setminus \{0\}$, $\nu \in V' \setminus \{0\}$ e $Q \in \mathcal{L}(V)$ tali che P si scriva nella forma $\lambda^{-1}P = \varphi\nu + Q$ e tali che sussistano fra loro almeno le seguenti quattro relazioni:

[a] $P(\varphi) = \lambda\varphi$;

[b] $Q(\varphi) = 0$;

[c] $\nu(P) = \lambda\nu$;

[d] $\nu(Q) = 0$.

Proprietà basilari (in tale situazione)**.**

[1] $\nu(\varphi) = 1$.

[2] $\nu = \nu(\lambda^{-1}P) = \nu(\varphi\nu)$.

[3] $Q(\varphi\nu) = 0$.

[4] $(\lambda^{-1}P)(\varphi\nu) = (\varphi\nu)(\lambda^{-1}P) = \varphi\nu$.

[5] Per ogni $n \in \mathbb{N}^*$, $(\lambda^{-1}P)(Q^n) = (Q^n)(\lambda^{-1}P) = Q^{n+1}$.

[6] Per ogni $n \in \mathbb{N}^*$, $(\lambda^{-1}P)^n = \varphi\nu + Q^n$.

[7] Per ogni $n \in \mathbb{N}^*$, $(I - \lambda^{-1}P)(\lambda^{-1}P)^n = (I - \lambda^{-1}P)(Q^n) = Q^n - Q^{n+1}$.

Sia dunque $E \subseteq \mathbb{R}$ un chiuso di cardinalità infinita avente 0 come punto di accumulazione, affinché naturalmente abbia senso l'operazione $\lim_{\varepsilon\to 0} \coloneqq \lim_{\varepsilon\in E,\, \varepsilon\to 0}$. Ad esempio, possiamo immaginare che sia $E = [0, \varepsilon_1]$ con $0 < \varepsilon_1 << 1$.

Siano allora $(P_\varepsilon)_{\varepsilon\in E}$ in $\mathcal{L}(V) \setminus \{0\}$, $(\lambda_\varepsilon)_{\varepsilon\in E}$ in $\mathbb{R} \setminus \{0\}$, $(\varphi_\varepsilon)_{\varepsilon\in E}$ in $V \setminus \{0\}$, $(\nu_\varepsilon)_{\varepsilon\in E}$ in $V' \setminus \{0\}$ e $(Q_\varepsilon)_{\varepsilon\in E}$ in $\mathcal{L}(V)$ in corrispondenza dei quali poniamo $\eta_\varepsilon \coloneqq \|\nu_0(P_0 - P_\varepsilon)\|_{V'}$ e $\Delta_\varepsilon \coloneqq \nu_0((P_0 - P_\varepsilon)(\varphi_0))$ per ogni $\varepsilon \in E$, e consideriamo le seguenti condizioni.

(A1) Per ogni $\varepsilon \in E$, $\lambda_\varepsilon^{-1}P_\varepsilon = \varphi_\varepsilon\nu_\varepsilon + Q_\varepsilon$; inoltre $C_0 \coloneqq \sup_{\varepsilon\in E} \|P_\varepsilon\| < \infty$.

(A2) Per ogni $\varepsilon \in E$, $P_\varepsilon(\varphi_\varepsilon) = \lambda_\varepsilon\varphi_\varepsilon$, $Q_\varepsilon(\varphi_\varepsilon) = 0$, $\nu_\varepsilon(P_\varepsilon) = \lambda_\varepsilon\nu_\varepsilon$, $\nu_\varepsilon(Q_\varepsilon) = 0$.

(A3) $C_1 \coloneqq \sum_{n=0}^{\infty} \sup_{\varepsilon\in E} \|Q_\varepsilon^n\| < \infty$.

(A3*) $C_1^* \coloneqq \sum_{n=0}^{\infty} \sup_{\varepsilon\in E} \|Q_\varepsilon^n\|^* < \infty$.

(A4) Per ogni $\varepsilon \in E$, $\nu_0(\varphi_\varepsilon) = 1$; inoltre $C_2 \coloneqq \sup_{\varepsilon\in E} \|\varphi_\varepsilon\|_V < \infty$.

(A4*) Per ogni $\varepsilon \in E$, $\nu_0(\varphi_\varepsilon) = 1$; inoltre $C_2^* := \sup_{\varepsilon\in E} \|\varphi_\varepsilon\|_V^* < \infty$.

(A5) $\eta_\varepsilon \xrightarrow[\varepsilon\to 0]{} 0$.

(A6) Esiste $C_3 \in (0,\infty)$ tale che $\eta_\varepsilon \|(P_0 - P_\varepsilon)(\varphi_0)\|_V \le C_3 |\Delta_\varepsilon|$ per ogni $\varepsilon \in E$.

(A6*) Esiste $C_3^* \in (0,\infty)$ tale che $\eta_\varepsilon \|(P_0 - P_\varepsilon)(\varphi_0)\|_V^* \le C_3^* |\Delta_\varepsilon|$ per ogni $\varepsilon \in E$.

Nota. Chiaramente (A3)$\Rightarrow$(A3*), mentre (A4*)$\Rightarrow$(A4) e (A6*)$\Rightarrow$(A6).

Notazione. Dati $(a_\varepsilon)_{\varepsilon\in E}$ in $\mathbb{R}$ e $(b_\varepsilon)_{\varepsilon\in E}$ in $(0,\infty)$, possibilmente con $b_\varepsilon \xrightarrow[\varepsilon\to 0]{} 0$, scriviamo $a_\varepsilon = \mathcal{O}(b_\varepsilon)$ intendendo che esiste $C \in (0,\infty)$ tale che $|a_\varepsilon| \le C b_\varepsilon$ per ogni $\varepsilon \in E$.

Osservazione 9. Assumendo (A1), (A2), (A4) o (A4*), e (A5), vale che $\Delta_\varepsilon \xrightarrow[\varepsilon\to 0]{} 0$ e $\lambda_\varepsilon \xrightarrow[\varepsilon\to 0]{} \lambda_0$, ed anzi che $\Delta_\varepsilon = \mathcal{O}(\eta_\varepsilon)$ e $\lambda_\varepsilon - \lambda_0 = \mathcal{O}(\eta_\varepsilon)$.

Teorema 7 (Keller-Liverani (2009))**.** *Si consideri il contesto fin ora descritto e si assumino* (A1), (A2), (A3), (A4), (A5) e (A6)*, o in alternativa* (A1), (A2), (A3*), (A4*), (A5) e (A6*).

[I] *Se $\Delta_\varepsilon = 0$ per infiniti $\varepsilon \in E$ arbitrariamente vicini a 0, allora esiste $\varepsilon_0 > 0$ tale che $\Delta_\varepsilon = 0$ e $\lambda_\varepsilon = \lambda_0$ per infiniti $\varepsilon \in E$ con $|\varepsilon| \le \varepsilon_0$.*

[II] *Se $\Delta_\varepsilon \neq 0$ definitivamente per $\varepsilon \to 0$, allora poniamo $q_{k,\varepsilon} := \frac{1}{\Delta_\varepsilon}\nu_0((P_0 - P_\varepsilon)(P_\varepsilon^k)(P_0 - P_\varepsilon)(\varphi_0))$ per ogni $k \in \mathbb{N}$ ed in corrispondenza dei medesimi $\varepsilon \in E$. Sia quindi assunta anche la seguente condizione.*

(A7) *Per ogni $k \in \mathbb{N}$, esiste finito il limite $q_k := \lim_{\varepsilon\to 0} q_{k,\varepsilon}$.*

Allora esiste finito il limite

$$\lim_{\varepsilon\to 0} \frac{\lambda_0 - \lambda_\varepsilon}{\Delta_\varepsilon} = 1 - \sum_{k=0}^{\infty} \lambda_0^{-(k+1)} q_k.$$

Di conseguenza, posto $\vartheta := 1 - \sum_{k=0}^{\infty} \lambda_0^{-(k+1)} q_k$, è

$$\lambda_\varepsilon \approx_{\varepsilon\to 0} \lambda_0 \exp(-\Delta_\varepsilon(\frac{\vartheta}{\lambda_0} + o_{\varepsilon\to 0}(1))).$$

Nota. Si sappia che uno dei primi calcoli della dimostrazione del teorema 7 porterebbe a scoprire che $1 - \nu_\varepsilon(\varphi_0) = \mathcal{O}(\eta_\varepsilon)$.

3.2 Un'Applicazione Astratta

Sia X uno spazio metrico, sia $\mathcal{B}$ la σ-algebra di Borel su X, sia m una probabilità su X e sia $T : X \to X$ $\mathcal{B}$-misurabile non singolare rispetto a m, e sia quindi $L : L^1(m) \to L^1(m)$ l'operatore di trasferimento.

L'idea ora sarebbe quella di trovare come V un $V < L^1(m)$ chiuso contenente le costanti con precisamente $\|\cdot\|_{L^1(m)} \le cost \|\cdot\|_V$ su V e tale che $L(V) < V$, più magari che $f\mathbb{1}_A \in V$ per ogni $f \in V$ e $A \in \mathcal{B}$, e tale che a $P := L_{|V}$ corrispondano $\lambda = 1$, $\varphi \in V \setminus \{0\} \ge 0$ m-q.c., $\nu = m$ e Q avente $\rho(Q) < 1$, per cui $P = \varphi m + Q$ avrebbe spectral gap.

A questo punto, posto $E := [0, \varepsilon_1]$ con $0 < \varepsilon_1 << 1$, vorremmo trovare $(A_\varepsilon)_{\varepsilon \in E}$ in $\mathcal{B}$ regolari abbastanza affinché si possa prendere $P_\varepsilon(f) := P(f \mathbb{1}_{A_\varepsilon^c})$ per ogni $\varepsilon \in E$ e $f \in V$, nel senso cioè che $(P_\varepsilon)_{\varepsilon \in E}$ soddisfi le ipotesi iniziali del teorema 7.

Osserviamo in particolare che, se fosse $m(A_0) = 0$, allora sarebbe $P_0 = P$. Comunque, più in generale, è intuitivo che debba essere $m(A_\varepsilon) << 1$ per ogni $\varepsilon \in E$, per cui possiamo pensare agli A_ε come ad *eventi rari* rispetto ad ogni $\mu \in SM(X) << m$.

Immaginiamo direttamente che sia $P_0 = P$, da cui $\nu_0 = m$, e che $A_\varepsilon \downarrow A_0$ aventi $m(A_\varepsilon) > 0$ per ogni $\varepsilon \in E \setminus \{0\}$. Notiamo come $P_0 = P$ dìa subito che $(P_0 - P_\varepsilon)(f) = P(f \mathbb{1}_{A_\varepsilon})$ per ogni $\varepsilon \in E$ e $f \in V$, mentre dalla sola $A_\varepsilon \downarrow A_0$ avremmo avuto $(P_0 - P_\varepsilon)(f) = P(f \mathbb{1}_{A_\varepsilon \setminus A_0})$.

Proprietà basilari (in tale situazione).

(1) $\int_X P_\varepsilon^n(f) d\nu_0 = \int_X f \prod_{i=0}^{n-1} \mathbb{1}_{A_\varepsilon^c}(T^i) d\nu_0$ per ogni $\varepsilon \in E$, $n \in \mathbb{N}^*$ e $f \in V$.

(2) $\int_X (P_0 - P_\varepsilon)(P_\varepsilon^k)(P_0 - P_\varepsilon)(f) d\nu_0 = \int_X f \mathbb{1}_{A_\varepsilon} \prod_{j=1}^{k} \mathbb{1}_{A_\varepsilon^c}(T^j) \mathbb{1}_{A_\varepsilon}(T^{k+1}) d\nu_0$ per ogni $\varepsilon \in E$, $k \in \mathbb{N}$ e $f \in V$.

Seguirebbe in particolare che $\Delta_\varepsilon = \nu_0(\varphi_0 \mathbb{1}_{A_\varepsilon}) = \int_{A_\varepsilon} \varphi_0 d\nu_0$ per ogni $\varepsilon \in E$, cioè che $\Delta_\varepsilon = \mu_0(A_\varepsilon)$ se $\mu_0 := \varphi_0 d\nu_0$. Assumendo così che $\mu_0(A_\varepsilon) > 0$ per ogni $\varepsilon \in E \setminus \{0\}$, mentre per forza $\mu_0(A_0) = 0$, otterremmo che $q_{k,\varepsilon} = \frac{1}{\mu_0(A_\varepsilon)} \mu_0(A_\varepsilon \cap T^{-1}(A_\varepsilon^c) \cap \cdots \cap T^{-k}(A_\varepsilon^c) \cap T^{-(k+1)}(A_\varepsilon))$ (il quale $\in [0,1]$) per ogni $k \in \mathbb{N}$ e $\varepsilon \in E \setminus \{0\}$. Ma ora, grazie al corollario 1, è $\sum_{k=0}^{\infty} q_{k,\varepsilon} = 1$ per ogni $\varepsilon \in E \setminus \{0\}$: supponendo pertanto che esista finito $q_k = \lim_{\varepsilon \downarrow 0} q_{k,\varepsilon}$ per ogni $k \in \mathbb{N}$, avremmo in definitiva che $\vartheta = 1 - \sum_{k=0}^{\infty} q_k \in [0,1]$ e che $\lim_{\varepsilon \downarrow 0} \frac{1 - \lambda_\varepsilon}{\mu_0(A_\varepsilon)} = \vartheta$, o scritto altrimenti che

$$\lambda_\varepsilon \approx_{\varepsilon \downarrow 0} \exp(-\mu_0(A_\varepsilon) \vartheta (1 + o_{\varepsilon \downarrow 0}(1))).$$

Applicazione. Immaginiamo che $A_0 = \{z\}$, $z \in X$ scelto, e che ad esempio T sia continua su X. Se z non fosse periodico per T, allora $\lim_{\varepsilon \downarrow 0} q_{k,\varepsilon} = 0$ per ogni $k \in \mathbb{N}$, cioè $q_k = 0$ per ogni $k \in \mathbb{N}$ e $\vartheta = 1$. Se invece z fosse p-periodico rispetto a T, $p \in \mathbb{N}^*$, allora $\lim_{\varepsilon \downarrow 0} q_{k,\varepsilon} = 0$ per ogni $k \in \mathbb{N} \setminus \{p-1\}$ mentre $\lim_{\varepsilon \downarrow 0} q_{p-1,\varepsilon} = \lim_{\varepsilon \downarrow 0} \frac{\mu_0(A_\varepsilon \cap T^{-p}(A_\varepsilon))}{\mu_0(A_\varepsilon)}$, da cui $\vartheta = 1 - \lim_{\varepsilon \downarrow 0} \frac{\mu_0(A_\varepsilon \cap T^{-p}(A_\varepsilon))}{\mu_0(A_\varepsilon)}$. Pertanto, $\vartheta < 1$ sarebbe possibile solo se z fosse un punto periodico rispetto a T.
In aggiunta a tutto questo, assumendo pure che (T, μ_0) sia mixing, avremmo grazie all'osservazione 5 e al teorema 3 che $\tau_{A_\varepsilon} < \infty$ μ_0-q.c. e che $\int_{A_\varepsilon} \tau_{A_\varepsilon} d(\mu_0)_{A_\varepsilon} = \frac{1}{\mu_0(A_\varepsilon)}$ per ogni $\varepsilon \in E \setminus \{0\}$. Ma questo "tempo tipico" di ritorno in A_ε è legato al tempo tipico di scontro su A_ε, e cioè col tempo tipico del verificarsi dell'evento raro A_ε: cerchiamo di stimare per ciò la distribuzione di τ_{A_ε} rispetto a μ_0.
In generale, per ogni $\varepsilon \in E$, $n \in \mathbb{N}^*$ e $f \in V$, è $\int_{A_\varepsilon^c \cap \{\tau_{A_\varepsilon} \geq n\}} f d\nu_0 = \int_X f \mathbb{1}_{A_\varepsilon^c \cap \{\tau_{A_\varepsilon} \geq n\}} d\nu_0 = \int_X f \prod_{i=0}^{n-1} \mathbb{1}_{A_\varepsilon^c}(T^i) d\nu_0 = \int_X P_\varepsilon^n(f) d\nu_0 = \lambda_\varepsilon^n \nu_\varepsilon(f) \nu_0(\varphi_\varepsilon) + \lambda_\varepsilon^n \int_X Q_\varepsilon^n(f) d\nu_0$, da cui sicuramente $\int_{A_\varepsilon^c \cap \{\tau_{A_\varepsilon} \geq n\}} f d\nu_0 = \lambda_\varepsilon^n \nu_\varepsilon(f) + \mathcal{O}(|\lambda_\varepsilon^n| \, \|Q_\varepsilon^n\| \, \|f\|_V)$. In particolare, la scelta $f = \varphi_0$ dà

$$\mu_0(A_\varepsilon^c \cap \{\tau_{A_\varepsilon} \geq n\}) \approx_{\varepsilon \downarrow 0,\, n \to \infty} \lambda_\varepsilon^n \approx_{\varepsilon \downarrow 0,\, n \to \infty} \exp(-n \mu_0(A_\varepsilon) \vartheta (1 + o_{\varepsilon \downarrow 0}(1))).$$

Di conseguenza, se comunque presi $t \in (0, \infty)$ e $(\varepsilon_n)_{n \in \mathbb{N}^*}$ in $E \setminus \{0\}$ tali che $\varepsilon_n \downarrow 0$ per $n \to \infty$ tali per cui $\lim_{n \to \infty} n \mu_0(A_{\varepsilon_n}) = t$, come se fosse $\mu_0(A_{\varepsilon_n}) = \frac{t}{n} + o_{n \to \infty}(\frac{1}{n})$ per $n >> t$, allora appunto

$$\mu_0(A_{\varepsilon_n}^c \cap \{\tau_{A_{\varepsilon_n}} \geq n\}) \approx_{n \to \infty} \exp(-t \vartheta (1 + o_{n \to \infty}(1))).$$

Osserviamo che a sinistra compare la μ_0-probabilità che l'evento raro A_{ε_n} impieghi tempi $\geq n = t(\frac{n}{t}) \approx_{n \to \infty} t \frac{1}{\mu_0(A_{\varepsilon_n})} =$ "t volte il tempo tipico" per verificarsi.

3.2.1 Un Calcolo Piu' Esplicito

Consideriamo $X = [0,1]$, m di Lebesgue e $T : [0,1] \to [0,1]$ la mappa di raddoppiamento dell'angolo, ossia $T(x) := 2x \; mod \; 1$ per ogni $x \in [0,1]$. Allora (T, m) stesso è p.p.t. e mixing. Prendiamo quindi $V := BV([0,1])$ e $A_\varepsilon \downarrow \{z\}$ dove $m(A_\varepsilon) = \varepsilon$ e $z \in [0,1]$ è p-periodico rispetto a T, $p \in \mathbb{N}^*$: allora $\lim_{\varepsilon\downarrow 0} \frac{\mu_0(A_\varepsilon \cap T^{-p}(A_\varepsilon))}{\mu_0(A_\varepsilon)} = 2^{-p}$ e quindi $\lim_{\varepsilon\downarrow 0} \frac{1-\lambda_\varepsilon}{m(A_\varepsilon)} = 1 - 2^{-p}$, o anche $\lambda_\varepsilon = 1 - \varepsilon(1 - 2^{-p}) + o_{\varepsilon\downarrow 0}(\varepsilon)$. In particolare, la funzione $\varepsilon \mapsto \lambda_\varepsilon$ è perlomeno differenziabile in $\varepsilon = 0$.

La vera difficoltà della situazione è riuscire a garantire tutte le ipotesi per le P_ε. Osserviamo comunque, sempre in una scelta particolare di (T, m) "simile alla precedente", che almeno da $\eta_\varepsilon \leq \nu_0(A_\varepsilon) = \varepsilon$ e da $\|(P_0 - P_\varepsilon)(\varphi_0)\|_V \leq cost \, \|\varphi_0 \mathbb{1}_{A_\varepsilon}\|_V$ avremmo che per (A6) basterà avere ad esempio $\|\varphi_0 \mathbb{1}_{A_\varepsilon}\|_V \leq cost \frac{1}{\nu_0(A_\varepsilon)} \int_{A_\varepsilon} \varphi_0 d\nu_0$, la quale è certamente vera se $\inf \varphi_{0|_{A_{\varepsilon_1}}} > 0$.

Riferimenti bibliografici

[1] Keller, G. - Liverani, C. (2009), *Rare Events, Escape Rates and Quasistationarity: Some Exact Formulae*, Journal of Statistical Physics.

[2] Keller, G. (2011), *Rare Events, Exponential Hitting Times and Extremal Indices via Spectral Perturbation*, Journal of Statistical Physics.

[3] Sarig, O. (2009), *Lecture Notes on Ergodic Theory*, The Weizmann Institute of Science (Israel).

[4] Sarig, O. (2012), *Introduction to the Transfer Operator Method*, The Weizmann Institute of Science (Israel).