

UNIVERSITÀ DEGLI STUDI DI PADOVA  
DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"

Corso di Laurea in Fisica

Tesi di laurea

# Meccanica statistica di non equilibrio e teoria delle grandi deviazioni

**Relatori**

prof. Marco Baiesi  
prof. Attilio Stella

**Candidato**

Francesco Coghi  
matricola: 1059749

Luglio 2015



# Indice

<b>1</b>	<b>Introduzione generale</b>	1
<b>2</b>	<b>Meccanica statistica di non equilibrio</b>	3
2.1	Generalità . . . . .	3
2.2	Reversibilità e Irreversibilità . . . . .	3
2.3	Processi stocastici . . . . .	6
2.3.1	Legame tra processi stocastici e fisica . . . . .	7
2.4	Processi di Markov . . . . .	8
2.5	La Master equation . . . . .	11
2.6	Il bilancio dettagliato . . . . .	13
<b>3</b>	<b>La teoria delle grandi deviazioni</b>	15
3.1	Il principio di grandi deviazioni . . . . .	15
3.1.1	Esempio <i>random bits</i> . . . . .	17
3.2	Il teorema di Gärtner-Ellis . . . . .	17
3.2.1	Teorema di Cramér . . . . .	18
3.3	Proprietà della $\lambda(k)$ e della rate $I(a)$ . . . . .	19
3.4	Grandi deviazioni nel non equilibrio . . . . .	20
3.4.1	Esempio <i>Markov bits</i> . . . . .	22
<b>4</b>	<b>Applicazione ad un modello stocastico di conduzione</b>	23
4.1	Generalità dello studio . . . . .	23
4.2	Il modello . . . . .	24
4.3	Applicazione delle grandi deviazioni . . . . .	27
4.4	Analisi delle rate functions $I(v)$ . . . . .	32
<b>5</b>	<b>Conclusioni</b>	33
	<b>Bibliografia</b>	35



# Capitolo 1

## Introduzione generale

Il titolo concesso a questa tesi comprenderebbe una tale vastità di argomenti che non possono essere sicuramente riassunti in una manciata di pagine di una tesi triennale. Quello che si vuole umilmente fare è presentare in modo sintetico, seguendo un certo rigore concettuale, il modo in cui la *teoria delle grandi deviazioni* entri nel campo della meccanica statistica, ed in particolare della *meccanica statistica di non-equilibrio*.

La teoria delle grandi deviazioni pur essendo un'utile e potente formalizzazione della meccanica statistica di equilibrio, per il non-equilibrio è qualcosa di più. In effetti in quest'ultimo caso diventa, come mostreremo successivamente, una tra le poche strade da seguire nella ricerca. Parallelamente a quanto piaceva pensare a *Mark Twain*, il quale denigrava l'uso scorretto della statistica da parte della maggior parte della gente, definendola come il lampione al quale l'ubriaco si aggrappa più per sostegno che per illuminazione, così potremmo vedere, in questo caso però non in senso negativo, le grandi deviazioni applicate allo studio del non-equilibrio. Tutto ciò per far capire quanto sia veramente essenziale.

La tesi è stata suddivisa in tre macroaree.

Il *primo capitolo* ha lo scopo di presentare i punti salienti del non-equilibrio, si parte cercando di capire come si possa distinguere uno stato di equilibrio da uno di non-equilibrio soffermandoci sulla differenza tra due concetti fondamentali: la reversibilità e l'irreversibilità. Una volta chiara la differenza si comincerà ad analizzare il formalismo matematico necessario nella descrizione dei processi di non-equilibrio costituito dai processi stocastici. A questo punto, in vista dell'applicazione analitico-numerica della tesi, ci soffermeremo ad analizzare quei processi stocastici che hanno una dipendenza temporale limitata allo stadio evolutivo precedente all'ultimo: i processi di Markov. Da qui si passerà agevolmente a derivare la Master Equation, un'importante equazione stocastica che determina la variazione nel tempo delle probabilità di transizione tra certi stati del sistema in oggetto. Un ultimo sforzo concettuale viene fatto in direzione dell'equazione di bilancio dettagliato. Quest'ultima, valida nell'assunzione di un sistema fisico, chiuso e isolato è un'equazione estremamente importante nello studio degli stati di equilibrio. Com'è ben noto i processi di non-equilibrio, per definizione, non attraversano stati di equilibrio, tuttavia è possibile per tali sistemi trovarsi in stati stazionari (*steady state*) e in tal caso si deriva l'equazione di bilancio dettagliato generalizzata, di estrema importanza, come lo è stato anche per noi, per lo studio di evoluzione fuori equilibrio.

Il *secondo capitolo* è interamente dedicato alla teoria delle grandi deviazioni. Si apre mostrando il principio di grandi deviazioni (LDP) e la sua formulazione esponenziale convenzionale che mette in gioco un'importante funzione, la rate function. Quest'ultima raccoglie in se molta informazione riguardo il comportamento asintotico delle fluttuazioni dalla media di una determinata variabile casuale. Una volta intuita l'importanza del LDP e della rate function nascerà il problema pratico della verifica del principio e del calcolo della rate e si vedrà che il calcolo diretto non è sempre così agevole. Entrerà quindi in gioco il teorema di Gärtner-Ellis, che passando per la cosiddetta Scaled Cumulant Generating Function (SCGF) convalida un potente metodo indiretto per compiere questi due compiti. Conclusa questa parte generale di introduzione della teoria, si passa a vedere come questa si applica allo studio dei processi di Markov che descrivono stocasticamente l'evoluzione temporale di sistemi fuori equilibrio.

Si giunge così al *terzo ed ultimo capitolo*, in cui per mettere in atto ciò che si è imparato, si studia in modo analitico e numerico un sistema in uno stato stazionario fuori equilibrio. Il sistema è costituito da un canale con dentro delle trappole. Un flusso di particelle si muove all'interno sotto l'azione di un campo e si può modellizzare il processo tramite una catena di Markov ergodica. Passando poi per il teorema di Gärtner-Ellis si giunge alla SCGF e da questa al calcolo della bramata rate function del processo. Da notare che questo non è un mero esercizio di calcolo, infatti vi è un vivo interesse nei confronti di tale struttura in quanto rende evidente una proprietà significativa, la mobilità differenziale negativa, che consiste in una risposta inaspettata del sistema all'aumento del campo. Ovvero le particelle invece di aumentare la loro velocità media, dopo un certo valore critico la diminuiscono, facendola tendere a 0 per campi molto grandi. Si è voluto quindi indagare come per campi differenti si comporti la rate function associata alla velocità, cercando di capire se anche le fluttuazioni dalla media sono soggette a variazioni del campo.

## Capitolo 2

# Meccanica statistica di non equilibrio

### 2.1 Generalità

Il fine ultimo della meccanica statistica di *equilibrio* è spiegare l'emergenza di uno stato di equilibrio di un sistema macroscopico, composto cioè da un numero elevato di particelle ( $\sim 10^{23}$ ), a partire da uno studio di natura probabilistica sui microstati occupati nello spazio delle fasi, potremmo definirlo uno studio d'*ensemble*.

Caratteristica fondamentale di questo approccio è il piccolo numero di macrostati, in rapporto all'elevato numero di microstati utilizzati per descrivere il sistema a molte particelle da un punto di vista microscopico. Il motivo di ciò è che possiamo descrivere (macroscopicamente) lo stato di equilibrio tramite poche variabili. Gli stati di equilibrio che vengono assunti da tali sistemi sono stati stazionari nel tempo e stabili a piccole perturbazioni.

Lo studio dei sistemi *non in equilibrio* risulta concettualmente più difficile, in quanto oltre allo studio delle fluttuazioni stazionarie, si è interessati alla dinamica del sistema che evolve nel tempo.

Cosa distingue formalmente un sistema in *equilibrio* da un sistema *fuori equilibrio*? Innanzitutto le osservabili dei due sistemi sono descritte da enti matematici differenti, nel primo caso è sufficiente una variabile casuale per descriverne la distribuzione, nel secondo caso invece, a causa dell'instabilità del sistema nel tempo, si fa uso di processi stocastici (si veda il 2.3). Ma non è tutto...

### 2.2 Reversibilità e Irreversibilità

Per chiarire le idee riguardo la differenza tra *equilibrio* e *non equilibrio*, trattiamo il problema matematicamente, senza necessità di rigore per il momento. Definiamo  $\omega_i$  il microstato occupato nello spazio delle fasi dal sistema al tempo  $i$ , quindi  $\omega = (\omega_i)_{i=1}^n$  consiste in una sequenza di  $n$  microstati attraversati dal sistema.

A questo punto si vuole studiare la distribuzione di probabilità  $P(\omega) = P(\omega_1, \dots, \omega_n)$  della traiettoria completa nello spazio delle fasi, se questa distribuzione di probabilità è

equivalente alla distribuzione ottenuta invertendo la sequenza dei microstati, ovvero se  $P(\omega) = P(\omega^R)$ , ove  $\omega^R = (\omega_n, \dots, \omega_1)$ , allora si ha in studio un sistema all'*equilibrio*. Se questa condizione non è soddisfatta, si tratta di un sistema *non in equilibrio* ([9]).

Matematicamente questa condizione è equivalente alla nozione di reversibilità o di bilancio dettagliato (si veda 2.6), in questo caso riferita alle traiettorie piuttosto che alle più usuali rates di transizione.

Risulta interessante approfondire il concetto di *reversibilità*. Il mondo di tutti i giorni è chiaramente irreversibile, come ben descritto dalla seconda legge della termodinamica. I sistemi fisici tendono ad un graduale e macroscopico stato di disordine, e ciò avviene seguendo il verso convenzionalmente positivo della linea temporale. Tuttavia questo sorprende il punto di vista fisico, in quanto le equazioni basilari del moto sono invarianti sotto inversione temporale, con alcune eccezioni nel campo delle interazioni deboli che tuttavia non ci interessano visto che i sistemi fisici che consideriamo evolvono grazie ad interazioni elettromagnetiche ([6]).

Entriamo maggiormente in dettaglio e analizziamo più approfonditamente il *paradosso dell'irreversibilità*. Consideriamo un sistema meccanico costituito da molte particelle in un recipiente, ciascuna di queste particelle ha una traiettoria ben definita nel tempo, nel mondo hamiltoniano infatti è sufficiente conoscere ad un certo istante posizione e momento della particella per conoscerne la traiettoria a tutti i tempi. Concentriamoci sulla particella *i-esima*, all'istante  $t = 0$  occuperà la posizione  $\vec{r}(0)$  e avrà momento  $\vec{p}(0)$ , se lasciamo evolvere il sistema per un tempo  $\tau$ , ritroveremo la particella nella posizione  $\vec{r}(\tau)$  con momento  $\vec{p}(\tau)$ . Ora, assumendo che sia possibile invertire i momenti di tutte le particelle, se lasciassimo evolvere il nostro sistema meccanico (conservativo) per un altro lasso di tempo  $\tau$ , troveremmo che la particella *i-esima* ha assunto posizione  $\vec{r}(2\tau) = \vec{r}(0)$  e momento  $\vec{p}(2\tau) = -\vec{p}(0)$ . Quest'operazione è conosciuta come *time reversal*, e per sistemi hamiltoniani la posizione  $\vec{r}$  e il momento  $\vec{p}$  sono, rispettivamente, funzione pari e dispari sotto questo tipo di trasformazione, e ciò comporta che la traiettoria sia la stessa, in un verso o nell'altro del tempo ([6]).

Si nota chiaramente che il paradosso dell'irreversibilità è ben radicato, dall'esempio sopracitato è evidente che le leggi del moto nel mondo microscopico sono invarianti sotto *time reversal*, tuttavia una simile invarianza non si ha nel mondo macroscopico di tutti i giorni.

Il primo fisico ad abbozzare una spiegazione dell'*irreversibilità* fu L. Boltzmann nella seconda metà dell'800, il quale per primo capì che non sarebbe stato possibile "derivarla" da puri ragionamenti meccanici, ma per farlo bisognava invocare argomenti di natura probabilistica. Un classico esempio è quello di un gas contenuto in un recipiente di volume  $V_{in}$  che dopo l'apertura di una valvola espande liberamente sino ad occupare un volume totale  $V_{fin} = 2V_{in}$ , ci si chiede quale sia la probabilità, una volta avvenuta l'espansione, che il gas torni ad occupare il volume  $V_{in}$ . Molto semplicemente, a posteriori, si potrebbe trattare il problema da un punto di vista puramente probabilistico, considerando la probabilità, che ciascuna particella ha di occupare un certo volume, distribuita come una binomiale, quindi per la particella *i-esima* la probabilità che si trovi in  $V_{in}$  è  $P(V_{in}) = \frac{1}{2} = P(V_{fin} - V_{in})$ , uguale alla probabilità che si trovi in  $V_{fin} - V_{in}$ , essendo il volume finale doppio di quello iniziale. A questo punto, considerando le particelle indipendenti le une dalle altre, la



distribuzione di probabilità di  $N$  particelle è quindi il prodotto delle singole binomiali, si deriva che la probabilità di trovare  $N$  particelle confinate in  $V_{in}$  dopo l'espansione è pari a  $(\frac{1}{2})^N$ .

Per ragioni didattiche si potrebbe affrontare il problema con i metodi propri della meccanica statistica, utilizzando, in questo caso a seconda di una personale preferenza nei vincoli da imporre, il formalismo del microcanonico, piuttosto del canonico o ancora del grancanonico, ottenendo ogni volta lo stesso risultato. Seguendo la linea di pensiero di Boltzmann, microcanonico dunque, si suppone che il gas abbia un'energia interna  $E$  che non vari a seguito dell'espansione, del tutto confermato dalle previsioni della termodinamica classica per un'espansione adiabatica, e si vuole calcolare il numero o volume di microstati  $\Omega(E)$  nello spazio delle fasi nello stato iniziale e in quello finale tramite la formula:

$$\Omega(E) = \frac{1}{N!h^N} \int_E^{E+\Delta E} \prod_{i=1}^N d^3r_i d^3p_i \quad (2.1)$$

Si deduce nell'immediato, visto che l'energia si conserva nella trasformazione e le particelle di gas sono indipendenti, che in generale il volume dei momenti nello spazio delle fasi si conserverà, e ciò che varierà sarà solo il volume effettivamente occupato dal gas, quindi:

$$\frac{\Omega_{fin}}{\Omega_{in}} \simeq \left( \frac{V_{fin}}{V_{in}} \right)^N = 2^N \quad (2.2)$$

Da qui, la probabilità di trovare tutte le particella in  $V_{in}$  è  $\sim 2^{-N}$  ([6]). Questo è ciò che Boltzmann dedusse e per l'epoca fu straordinariamente innovativo producendo non poco scalpore nel mondo scientifico. Diverse furono le critiche mosse da diversi illustri scienziati al fine di mostrare l'erroneità nell'argomento statistico. Le principale obiezioni arrivarono da Zermelo e Loschmidt. Zermelo, forte del fenomeno della *ricorrenza di Poincaré*, sosteneva che un sistema Hamiltoniano sarebbe tornato ad occupare lo stesso volume dei microstati nello spazio delle fasi se si fosse stati disposti ad aspettare sufficientemente a lungo, dell'ordine di  $\sim e^N$ , ma effettivamente già per poche particelle si dovrebbero aspettare tempi dell'ordine dell'età dell'Universo. Loschmidt invece attaccava Boltzmann partendo dall'invarianza per *time reversal*, secondo la quale, pur avendo un sistema meccanico arbitrariamente grande, sarebbe stato sufficiente invertire tutte le velocità delle particelle in un determinato momento per mostrare che il sistema sarebbe tornato nel suo stato iniziale, tuttavia si nota che basta un piccolissimo errore nell'invertire le velocità che questo fenomeno non accade più. Questa è una conseguenza della caoticità nella dinamica del sistema, che comporta un'estrema sensibilità alle condizioni iniziali ([6]).

In realtà, analizzando le critiche, si vede come queste non influenzano minimamente il ragionamento di Boltzmann. In altre parole Zermelo e Loschmidt trattano la questione parallelamente a Boltzmann, solo che quest'ultimo sembra non sia impaurito dall'ingresso della statistica nella trattazione dei problemi fisici.

Il ragionamento di Boltzmann mostra, grazie a dei ragionamenti probabilistici (il calcolo del volume dei microstati nello spazio delle fasi), come l'irreversibilità emerga nel caso di sistemi ad elevato numero di gradi di libertà ( $\sim 10^{23}$ ). Da notare inoltre come per un sistema di queste dimensioni le condizioni iniziali occupano un piccolo volume nello spazio

delle fasi, sintomo del fatto che la probabilità del macrostato iniziale diventa velocemente trascurabile una volta che il tutto viene lasciato evolvere. Infine, l'irreversibilità è anche legata al carattere caotico della dinamica del sistema, ove piccole variazioni sulle condizioni iniziali possono portare due traiettorie a divergere esponenzialmente ([6]).

## 2.3 Processi stocastici

Come anticipato poc'anzi, i *processi stocastici* rivestono un ruolo fondamentale nello studio dei sistemi non in equilibrio.

Un processo stocastico è semplicemente una funzione di due variabili, ove una variabile è il tempo  $t$  (continuo o discreto) e l'altra è una variabile casuale  $X$  con spazio degli stati  $\Omega_X = \mathbb{R}$  (rimaniamo generali per il momento):

$$Y_X(t) = f(X, t). \quad (2.3)$$

Per il momento non siamo interessati a definire precisamente lo spazio degli eventi  $\Omega_Y$ . Sostituendo alla variabile casuale uno dei suoi possibili valori, si ottiene una possibile realizzazione del processo  $Y_x(t) = f(x, t)$ . In termini fisici si studia un processo stocastico come un *ensemble* delle sue possibili realizzazioni. Si può facilmente calcolare la media di un determinato processo come:

$$\langle Y(t) \rangle = \int Y_x(t) P_X(x) dx, \quad (2.4)$$

e analogamente il momento  $n$  –esimo:

$$\langle Y(t_1)Y(t_2)\dots Y(t_n) \rangle = \int Y_x(t_1)Y_x(t_2)\dots Y_x(t_n) P_X(x) dx. \quad (2.5)$$

Di particolare interesse risulta la funzione di *auto-correlazione*:

$$k(t_1, t_2) \equiv \langle \langle Y(t_1)Y(t_2) \rangle \rangle = \langle Y(t_1)Y(t_2) \rangle - \langle Y(t_1) \rangle \langle Y(t_2) \rangle \quad (2.6)$$

ove si nota che nel caso in cui  $t_1 = t_2$  si ottiene proprio una varianza dipendente dal tempo. Un'altra proprietà interessante dei processi stocastici è la *stazionarietà*:

$$\langle Y(t_1 + \tau)Y(t_2 + \tau)\dots Y(t_n + \tau) \rangle = \langle Y(t_1)Y(t_2)\dots Y(t_n) \rangle \quad (2.7)$$

in particolare risulta che  $\langle Y \rangle$  è indipendente dal tempo. Ne risulta che la funzione di auto-correlazione di un processo stocastico stazionario dipende da  $|t_1 - t_2|$  solamente; si definisce una costante  $\tau_c$ , chiamata tempo di auto-correlazione, secondo la quale se  $|t_1 - t_2| > \tau_c$  allora  $k(t_1, t_2) \simeq 0$ .

Esiste un modo parallelo a quello appena visto per trattare i processi stocastici che fa uso della *gerarchia delle funzioni di distribuzione*. Si parte semplicemente chiedendosi quale sia la densità di probabilità per  $Y_X(t)$  di assumere il valore  $y$  al tempo  $t$ , come evidente dalla (2.4), risulta:

$$P(y, t) = \int \delta(y - Y_x(t)) P_X(x) dx. \quad (2.8)$$

In modo analogo la densità di probabilità congiunta (*joint probability density function*), che  $Y$  assuma valore  $y_1$  al tempo  $t_1$ ,  $y_2$  al tempo  $t_2$  e così via fino a  $y_n$  al tempo  $t_n$ , è:

$$P(y_1, t_1; y_2, t_2; \dots; y_n, t_n) = \int \delta(y_1 - Y_x(t_1))\delta(y_2 - Y_x(t_2))\dots\delta(y_n - Y_x(t_n))P_X(x)dx, \quad (2.9)$$

da cui discende una formula equivalente alla (2.5) per il momento n-esimo:

$$\langle Y(t_1)Y(t_2)\dots Y(t_n) \rangle = \int y_1 y_2 \dots y_n P(y_1, t_1; y_2, t_2; \dots; y_n, t_n) dy_1 dy_2 \dots dy_n. \quad (2.10)$$

La gerarchia delle funzioni di distribuzione  $P$  obbedisce alle seguenti quattro condizioni di consistenza, che secondo una dimostrazione di Kolmogorov la rendono equivalente ad un processo stocastico:

- i.  $P \geq 0$ ;
- ii.  $P$  non cambia scambiando tra loro  $(y_k, t_k)$  e  $(y_l, t_l)$ ;
- iii.  $\int P(y_1, t_1; \dots; y_{n-1}, t_{n-1}; y_n, t_n) dy_n = P(y_1, t_1; \dots; y_{n-1}, t_{n-1})$ ;
- iv.  $\int P(y_1, t_1) dy_1 = 1$ .

Abbiamo riportato una breve introduzione ai processi stocastici unidimensionali ([5]), tuttavia tutta la trattazione sin qui fatta si potrebbe estendere al caso multidimensionale, ma al fine ultimo della tesi non interessa.

### 2.3.1 Legame tra processi stocastici e fisica

Ciò che invece si ritiene importante è capire più approfonditamente il ruolo di questi processi nella fisica.

Innanzitutto, perchè i processi stocastici entrano nel mondo della fisica? Le ragioni dovrebbero essere abbastanza evidenti, prima fra tutte il fatto che ci sono processi che evolvono nel tempo in modo estremamente complicato, dipendenti da moltissime variabili, per i quali le uniche informazioni che si possono ottenere derivano da medie di qualche osservabile. Quindi, spesso, si argomenta la necessità di utilizzare questi tools probabilistici a causa della nostra ignoranza sul preciso microstato occupato dal sistema, tuttavia ciò è vero solo in parte, in quanto non spiega perchè certe variabili possano essere mediate e altre no (energia totale e composizione di un gas ad esempio). Ciò che sicuramente è vero è che, nonostante la nostra ignoranza sul comportamento di molte delle variabili microscopiche, si evidenziano delle regolarità nel comportamento macroscopico. Quindi risulta evidente che il preciso valore di tali variabili microscopiche non è importante, e si potrebbe tranquillamente mediare su di esse.

Entrando più nel dettaglio, cerchiamo di capire come un processo stocastico entra nella descrizione di un sistema fisico. Ciò, in parte, per le variabili casuali, dovrebbe essere già evidente dallo studio della meccanica statistica di equilibrio. Consideriamo un sistema di  $N$  particelle in una scatola, con la condizione di pareti riflettenti, i cui microstati sono

descritti da  $6N$  coordinate ( $3N$  momenti e  $3N$  posizioni), questo evolverà deterministicamente, una volta fissato lo stato iniziale (chiamiamolo  $x_0$ ), nello spazio delle fasi secondo le equazioni di Hamilton. Nel tempo verranno quindi occupati svariati microstati differenti, corrispondenti tutti allo stesso macrostato. Potremmo essere interessati nel conoscere una quantità fisica  $Y$ , relativa al sistema e funzione delle  $6N$  variabili, due esempi sono: il numero di particelle  $n$  in un definito volume o la pressione esercitata su un pistone. L'idea base della meccanica statistica è quella di considerare tutti i possibili microstati del sistema, associati ad un determinato macrostato, come un *ensemble* con una funzione di densità  $\rho(x)$ , tale che  $\rho(x)dx$  è il numero di microstati nell'elemento di volume  $dx$ . Introducendo questa funzione di densità si è trasformata la variabile  $x$  che identifica il microstato in una variabile casuale  $X$ , in questo modo si viene a definire una funzione di densità di probabilità che a meno di normalizzazione è equivalente proprio alla densità dei microstati:

$$P_X(x) = \frac{\rho(x)}{\int \rho(x')dx'}. \quad (2.11)$$

A questo punto non basta far altro che calcolare la media dell'osservabile  $Y$  tramite la funzione  $P_X(x)$ .

Estremamente importante notare come l'introduzione dell'ensemble ha l'effetto di trasformare ciascuna quantità fisica  $Y(t)$  in un processo stocastico, ad esempio, la pressione sul pistone si calcolerà sfruttando la (2.4), come media d'ensemble delle forze esercitate dalle molecole, piuttosto che una media temporale.

Quindi, riassumendo, l'idea fondamentale della meccanica statistica è quella di sfruttare le medie d'ensemble al posto delle medie temporali.

Tuttavia questo non è sufficiente, è importante fare almeno un'altra assunzione, denominata *caos molecolare*, la quale consiste nell'effettuare medie ripetute sulle variabili irrilevanti, con l'effetto di eliminarne le rapide variazioni, ciò porta alla creazione di equazioni differenziali solamente per le variabili *lente*.

Le equazioni differenziali che si ottengono sono le *equazioni macroscopiche* (ad esempio la legge di Malthus nella crescita di una popolazione). Queste equazioni, viste le assunzioni fatte, sono chiaramente delle approssimazioni, e in effetti sono evidenti delle piccole deviazioni, le cosiddette *fluttuazioni*, che non si possono trovare esattamente, tuttavia grazie alle loro proprietà stocastiche, obbediscono anche loro a delle semplici leggi.

Questo approccio alla teoria delle fluttuazioni è definito *mesoscopico* ([5]), e risulta più dettagliato di quello *macroscopico*, visto che considera le *fluttuazioni*.

## 2.4 Processi di Markov

Con il passaggio dallo studio delle variabili casuali ai processi stocastici, introducendo una dipendenza dal tempo, si allarga lo studio a tutti quei processi che esprimono nella loro evoluzione una certa dipendenza dal passato. Una tra le più importanti sottoclassi dei processi stocastici è quella dei processi di Markov.

Nell'ambito matematico-probabilistico parlare di dipendenza di uno stato rispetto ad un

altro introduce il concetto di probabilità condizionata che per un processo stocastico qualsiasi possiamo riassumere così:  $P(\text{stato}_1, t_1 | \text{stato}_0, t_0) = \frac{P(\text{stato}_1, t_1) \cap P(\text{stato}_0, t_0)}{P(\text{stato}_0, t_0)}$ . Nell’ottica di ciò che andremo a sviluppare nella parte applicativa della tesi, in questa sezione tratteremo i processi di Markov omogenei a stati discreti e a tempo discreto, l’estensione al tempo continuo è abbastanza immediata. Riprendiamo quindi il processo stocastico  $Y(t)$ , ove lo spazio degli stati  $\Omega_Y$  e il tempo  $t$  assumono valori discreti (si trascura la dipendenza dalla variabile casuale  $X$ ), questo sarà definito processo di Markov o catena di Markov se:

$$P(y_n, t_n | y_1, t_1; \dots; y_{n-1}, t_{n-1}) = P(y_n, t_n | y_{n-1}, t_{n-1}), \quad (2.12)$$

com’è evidente dalla definizione, la dipendenza dal passato è limitata al solo stato assunto al tempo  $n - 1$ . Inoltre tale catena di Markov è definita *omogenea* se la parte destra della 2.12 non mostra dipendenza da  $n$ . Abbrevieremo l’espressione catena di Markov omogenea con l’acronimo HMC (*Homogeneous Markov Chain*). É possibile definire una matrice  $\Pi = (\pi_{ij})_{i,j \in \Omega_Y}$  associata ad una HMC, ove:

$$\pi_{ij} = P(j, t_{n+1} | i, t_n), \quad (2.13)$$

chiamata *matrice di transizione*. Dal fatto che le sue entrate sono probabilità, e dal fatto che tramite una transizione a partire da un qualsiasi stato  $i$  si raggiunge un dato stato  $j$ , segue che  $\pi_{i,j} \geq 0$  e che  $\sum_{k \in \Omega_Y} \pi_{ik} = 1$  per tutti gli stati  $i, j$ . Talvolta una matrice di transizione  $\Pi$  è rappresentata da un *grafo di transizione*, un grafo avente per nodi gli stati di  $\Omega_Y$ . Ciascun nodo è poi collegato da delle frecce curve orientate che indicano le transizioni tra un nodo e l’altro, su ciascuna freccia è poi indicata la probabilità di una determinata transizione (per un esempio si veda 3.4.1).

A questo punto della trattazione risulta utile introdurre due definizioni che saranno indispensabili nel seguito. Una catena di Markov si dice *irriducibile* se partendo da ogni stato  $i$  c’è una probabilità maggiore di zero di raggiungere ogni altro stato  $j$ , formalmente:

$$\forall i, j \in \Omega_Y, \forall m \in \mathbb{N}, \exists n \in \mathbb{N} : P(j, t_{m+n} | i, t_m) > 0 \quad (2.14)$$

Una catena di Markov si dice *aperiodica* se ogni stato  $i$  è aperiodico, ovvero se vi è una probabilità diversa da 0 per ogni stato di tornare su se stesso in un solo step temporale. Catene di Markov che sono sia irriducibili che aperiodiche si definiscono *ergodiche*.

Particolarmente interessante è lo studio della distribuzione di una HMC. Definiamo la distribuzione di probabilità  $\nu_0(i) = P(i, t_0)$  come la distribuzione iniziale. Dalla legge sequenziale di Bayes deriva che

$$P(i_0, t_0; i_1, t_1; \dots; i_k, t_k) = P(i_0, t_0)P(i_1, t_1 | i_0, t_0) \dots P(i_k, t_k | i_{k-1}, t_{k-1}; \dots; i_0, t_0) \quad (2.15)$$

e da ciò, in vista della proprietà delle HMC e della definizione di matrice di transizione:

$$P(i_0, t_0; i_1, t_1; \dots; i_k, t_k) = \nu(i_0)\pi_{i_0 i_1} \dots \pi_{i_{k-1} i_k}. \quad (2.16)$$

L’equazione 2.16 per tutti i  $k \geq 0$  e tutti gli stati  $i_0, i_1, \dots, i_k$  costituisce la distribuzione della HMC. Questo risultato importante quindi dice che una distribuzione di una HMC a tempo discreto è definita unicamente dalla distribuzione iniziale e dalla sua matrice di transizione. Si può dire anche di più, consideriamo la distribuzione al tempo  $t_n$ ,  $\nu_n(i) = P(i, t_n)$ ,

dalla legge di Bayes e dalla proprietà di Markov discende che  $\nu_{n+1}(j) = \sum_{i \in \Omega_Y} \nu_n(i) \pi_{ij}$ , che in forma matriciale corrisponde a  $\nu_{n+1}^T = \nu_n^T \Pi$ , e iterando si ottiene la forma più utile:

$$\nu_{n+1}^T = \nu_0^T \Pi^n. \quad (2.17)$$

Prima di passare ad un esempio dell'utilizzo delle HMC riporto un altro teorema particolarmente interessante al caso nostro. ([1])

**Teorema 2.4.1** (Teorema della convergenza di una HMC). *Sia definita una HMC a stati discreti  $Y(t)$ , con matrice di transizione associata  $\Pi = \pi_{ij}$  e spazio degli stati  $\Omega_Y$ . Se la catena di Markov è ergodica, la distribuzione di probabilità  $\nu_n$  al tempo  $t_n$ , converge alla distribuzione stazionaria  $\nu$  per ogni distribuzione iniziale di probabilità  $\nu_0$ . Si ha cioè*

$$\forall \nu_0, \forall i, j \in \Omega_Y, \lim_{n \rightarrow \infty} \sum_{i \in \Omega_Y} (\nu_0)_i (\pi)_{ij}^n = \pi_j$$

Un esempio storico ed estremamente interessante per capire bene come si lavora con le HMC discrete a tempo discreto, riguarda il *moto Browniano* ([9]). Per semplificare, senza ledere in generalità, consideriamo una particella che si muove in un reticolo unidimensionale di lunghezza  $L$ , ove per il momento consideriamo condizioni periodiche al contorno. Consideriamo la posizione assunta dalla particella  $Y(t)$  il nostro processo di Markov, i microstati di questo modello consistono nelle posizioni sul reticolo della particella, ovvero lo spazio degli stati è:  $\Omega_Y = \{1, 2, \dots, i, \dots, j, \dots, L\}$ . A tempo discreto questo processo è regolato da una matrice di transizione:

$$\Pi = \begin{pmatrix} 1-p-q & p & 0 & \dots & q \\ q & 1-p-q & p & \dots & 0 \\ 0 & q & 1-p-q & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p & 0 & 0 & \dots & 1-p-q \end{pmatrix}$$

Ove  $0 \leq p \leq 1$  e  $0 \leq q \leq 1-p$  sono le probabilità della particella di muoversi rispettivamente verso destra e verso sinistra per singolo time-step. Da notare che se  $p+q < 1$  la particella ha anche una probabilità finita per time-step di rimanere nella sua posizione. Inoltre risulta facile verificare un risultato, ovvio intuitivamente, che la matrice di transizione appena descritta ha una distribuzione stazionaria di probabilità  $1/L$  per la particella di trovarsi su ciascun sito.

Potremmo essere interessati nel calcolare la corrente media di particelle  $J_n$ , definita nel nostro caso come  $J_1$  la probabilità che la particella attraversi una sezione unitaria nel verso di destra per singolo time-step, associando  $+1$  all'evento *movimento verso destra*,  $-1$  all'evento *movimento verso sinistra*,  $0$  all'evento *nessun movimento*, si ottiene per la singola particella:

$$\langle J_1 \rangle = +1p + (-1)q + 0(1-p-q) = p - q$$

Da notare che l'osservabile  $J_n$  è a sua volta un processo stocastico dipendendo dalla variabile stocastica markoviana *posizione*, in generale:

$$\langle J_n(i) \rangle = \frac{1}{n} \sum_{i=1}^{n-1} f(i, i+1) \quad (2.18)$$

Ove nel nostro caso  $f(i, i+1) = \{+1, 0, -1\}$  a seconda del movimento della particella. Si nota anche che la  $\langle J_n \rangle$  è calcolata come media d'ensemble secondo la 2.4.

È importante spendere ancora due parole sul moto Browniano, ove la posizione della particella (e così anche la velocità) è solo approssimativamente un processo stocastico markoviano. Infatti, se lo spostamento precedente  $Y(t_n) - Y(t_{n-1})$  è grande, è più probabile che nel momento dell'osservazione  $Y(t_n)$  la particella abbia una grande velocità. Questa velocità sopravviverà per un tempo dell'ordine di autocorrelazione della velocità e favorirà un grande spostamento  $Y(t_{n+1}) - Y(t_n)$ . Quindi il fatto che il tempo di autocorrelazione per la velocità non è nullo porta ad una correlazione tra due successivi spostamenti, è quindi evidente che la posizione  $Y(t_{n+1})$  dipenda anche da  $Y(t_{n-1})$  oltre che da  $Y(t_n)$ . Tuttavia questo effetto è tanto più piccolo quanto più il tempo trascorso tra due osservazioni è maggiore del tempo di autocorrelazione della velocità ([5]).

## 2.5 La Master equation

La *Master equation* è una formulazione maneggevole dell'equazione di Chapman-Kolmogorov (C-K) seguente:

$$P(y_3, t_3 | y_1, t_1) = \sum_{y_2 \in \Omega_Y} P(y_3, t_3 | y_2, t_2) P(y_2, t_2 | y_1, t_1) \quad (2.19)$$

che si deriva semplicemente restringendo la 2.15 a  $k = 3$ , sommando su  $y_2$  e dividendo entrambi i lati per  $P(y_1, t_1)$ . Nel ricavare la Master equation passeremo ad una descrizione del processo stocastico markoviano in tempo continuo, accanto alle *probabilità di transizione* di cui abbiamo già parlato estesamente poc'anzi appariranno quindi le *transition rates* o più semplicemente *rates*  $W$ , così definite:

$$\lim_{\tau \rightarrow 0} \frac{P(y', t + \tau | y, t)}{\tau} = W(y \rightarrow y') = W(y' | y). \quad (2.20)$$

La rate è quindi definita come la probabilità per unità di tempo di passare da uno stato  $y$  ad uno stato  $y'$  entrambi nello spazio degli stati  $\Omega_Y$ .

Nel procedere semplificherò la notazione per l'indicazione del tempo, indicherò a pedice il tempo che passa tra l'osservazione di due stati, nel modo seguente:  $P(y', t + \tau | y, t) = P_\tau(y' | y)$ .

Riscriviamo in un modo utile la probabilità di transizione tra due stati, questa consisterà nel togliere a 1 la probabilità che da un definito stato iniziale si passi a tutti i possibili stati finali, tranne quello d'interesse:

$$P_{\tau'}(y_2 | y_1) = \delta_{y_1 y_2} - \sum_{y_2 \in \Omega_Y} W(y_2 | y_1) \tau' + \tau' W(y_2 | y_1) + o(\tau') \quad (2.21)$$

Il fatto che si tengano al massimo termini lineari in  $\tau'$  è giustificato dal fatto che  $\tau'$  è una quantità molto piccola, al limite infinitesima. Sostituendo questa nuova formulazione nell'equazione C-K del 2.19, si ottiene:

$$P_{\tau+\tau'}(y_3|y_1) = \left[ 1 - \sum_{y_2 \in \Omega_Y} W(y_2|y_3)\tau' \right] P_{\tau}(y_3|y_1) + \tau' \sum_{y_2 \in \Omega_Y} W(y_3|y_2)P_{\tau}(y_2|y_1)$$

Da qui, dividendo per  $\tau'$  ed eseguendo  $\lim_{\tau' \rightarrow 0}$ , nonchè notando che  $P_{\tau}(y_2|y_1)$  è identico alla distribuzione  $P(y_2)$  e sopprimendo gli indici superflui, si ricava la *Master equation*:

$$\frac{dP(y, t)}{dt} = \sum_{y' \in \Omega_Y} \{W(y|y')P(y', t) - W(y'|y)P(y, t)\}. \quad (2.22)$$

Come si nota chiaramente la Master equation è un'equazione di tipo guadagno-perdita per lo stato  $y$ , ove si considerano positive le transizioni che "entrano" nello stato d'interesse e negative quelle che "escono" ([5]).

Potremmo porre la 2.22 in una forma più compatta, con le dovute assunzioni:

$$\dot{p}(t) = \mathbb{W}p(t) \quad (2.23)$$

ove la matrice  $\mathbb{W}$  è definita nel modo seguente:

$$\mathbb{W} = \begin{cases} \mathbb{W}_{yy'} = W(y|y') & \text{per } y \neq y', \\ \mathbb{W}_{yy} = -\sum_{y'(\neq y)} W(y'|y). \end{cases} \quad (2.24)$$

Una fondamentale proprietà della Master equation è che per  $t \rightarrow \infty$  tutte le soluzioni tendono alla soluzione stazionaria, o nel caso di una matrice  $\mathbb{W}$  riducibile e decomponibile (vedi irriducibilità al 2.14) ad una delle sue soluzioni stazionarie. Tutto ciò è quanto accade anche per la matrice di transizione  $\Pi$ , si veda il teorema 2.4.1. Quindi, com'era intuibile già dall'inizio, c'è un forte parallelismo tra la trattazione del processo stocastico a tempo discreto con la matrice di transizione  $\Pi$  e la trattazione a tempo continuo con la matrice  $\mathbb{W}$ . Le entrate della matrice  $\Pi$  sono delle probabilità, mentre le entrate della matrice  $\mathbb{W}$  sono delle rates di transizione. Cerchiamo di visualizzare meglio ciò che si intende. Sappiamo qual'è il legame che corre tra rates e probabilità di transizione dalla 2.20, per  $\tau$  sufficientemente piccoli quindi si ha:  $P_{\tau}(y'|y) = \tau W(y'|y)$  e discretizzando anche la Master equation 2.22, si ottiene:

$$P_{t+\tau}(y) = P_t(y) + \sum_{y' \in \Omega_Y} \{P(y|y')P(y', t)\} - \sum_{y' \in \Omega_Y} \{P(y'|y)P(y, t)\}, \quad (2.25)$$

dove possiamo scrivere il secondo termine come  $\sum_{y' \in \Omega_Y} \{P(y'|y)P(y, t)\} = P(y, t)\tau\lambda(y)$ . Sapendo dunque che valgono la 2.17 e la 2.23, possiamo mettere in evidenza il parallelismo:

$$\Pi = \begin{pmatrix} 1 - \tau\lambda(y) & \tau W(y|y') & \tau W(y|y'') & \dots \\ \tau W(y'|y) & 1 - \tau\lambda(y') & \tau W(y'|y''') & \dots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \rightarrow \mathbb{W} = \begin{pmatrix} -\lambda(y) & W(y|y') & W(y|y'') & \dots \\ W(y'|y) & -\lambda(y') & W(y'|y''') & \dots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

Si nota quindi che  $\Pi = \mathbf{1} + \tau\mathbb{W}$ .



## 2.6 Il bilancio dettagliato

In questa sezione riportiamo (senza dimostrazione formale che sarebbe leggermente lunga) l'equazione del *bilancio dettagliato* ([5]). Ad ogni modo è importante concentrarsi sulle ipotesi di partenza che permettono di ricavarla.

Anzitutto ci poniamo in un sistema *fisico* che sia contemporaneamente *chiuso* ed *isolato*. Con sistema *fisico* intendiamo un sistema che sia descrivibile microscopicamente in termini delle equazioni di Hamilton (o per un sistema quantistico dall'equazione di Schrödinger), il termine *chiuso* si riferisce al fatto che non ci possano essere scambi di materia con l'ambiente esterno e *isolato* significa che non ci sono forze dipendenti dal tempo che agiscono sul sistema, quindi l'energia è un integrale primo, e la traiettoria del sistema che evolve nello spazio delle fasi è confinata ad una singola shell energetica. Oltre a ciò è necessario che l'Hamiltoniana del sistema e pure la variabile stocastica  $Y(t)$  siano funzioni dispari dei momenti. Le assunzioni fin qui fatte permettono una dimostrazione generale del bilancio dettagliato.

L'equazione che si ricava (a tempo continuo) è:

$$W(y|y')P^e(y', t) = W(y'|y)P^e(y, t), \quad (2.26)$$

da notare che non stabilisce semplicemente che  $\frac{dP(y,t)}{dt} = 0$ , permettendo di semplificare la 2.22 considerando che all'equilibrio la somma di tutte le transizioni per unità di tempo che portano in un determinato stato  $y$  è bilanciata dalla somma delle transizioni che escono da  $y$  e portano in un qualunque altro stato  $y'$ , ma la 2.26 è un'asserzione ben più forte restringendosi a ciascun paio di  $y$  e  $y'$  separatamente.

Concludiamo ricollegandoci ai discorsi fatti all'inizio per distinguere processi all'*equilibrio* da processi *fuori equilibrio* (si veda 2.2), in effetti il *bilancio dettagliato* è equivalente alla nozione di *reversibilità*, quindi se è garantito ci troviamo in un sistema all'equilibrio. Tuttavia se ne può trovare una formulazione analoga, definita *bilancio dettagliato generalizzato* valida per stati stazionari di sistemi fuori equilibrio (si veda il capitolo finale 4). L'espressione d'uso comune per il bilancio dettagliato generalizzato è [7]:

$$W(y|y')P^s(y', t) = W(y'|y)P^s(y, t)e^{\gamma_0 f(y,y')}, \quad (2.27)$$

ove per  $\gamma_0 = 0$  si ritorna all'usuale bilancio dettagliato 2.26, mentre per  $\gamma_0 \neq 0$  si ha che  $s$  ad apice indica che siamo interessati allo stato stazionario, e  $f(y, y')$  è la quantità di cui varia per singolo time-step l'osservabile alla quale siamo interessati (ad esempio nel caso in cui l'osservabile sia la corrente di particelle 2.18,  $f(y, y') = \{0, \pm 1\}$ ).



## Capitolo 3

# La teoria delle grandi deviazioni

"[...] we shall argue, by accumulating several correspondences between statistical mechanics and large deviation theory, that the mathematics of statistical mechanics, as a whole, is the theory of large deviations, in the same way that differential geometry, say, is the mathematics of general relativity." [8].

La teoria delle grandi deviazioni ebbe inizio grazie a Cramér negli anni '30 del secolo scorso, tuttavia progredì poco e in solitaria senza incontrare la fisica fino agli inizi degli anni '70, quando si sviluppò una teoria unificata grazie a Donsker e Varadhan e da qui grazie al lavoro di fisici-matematici come Gärtner e Ellis cominciarono ad essere pubblicati nel mondo scientifico degli articoli in cui si formula la meccanica statistica in termini della teoria delle grandi deviazioni ([8]). Al giorno d'oggi questa teoria è uno dei principali strumenti in meccanica statistica ed in particolare nello studio del non equilibrio. A posteriori si può riconoscere come già a partire dalla famosa equazione di Boltzmann  $S = k \ln W$ , scritta in questa forma da Planck, fosse stata interpretata da Einstein come una formula con cui esprimere la probabilità di deviazioni dallo stato di equilibrio. [2][4]

In questo capitolo presenteremo i risultati fondamentali di questa teoria attraverso una formulazione matematica non strettamente rigorosa. Inoltre limiteremo lo studio della teoria a variabili casuali reali, discrete o continue. Le variabili che utilizzeremo nel seguito saranno sempre tali da *convergere debolmente*, ovvero data la probabilità  $P(A_n \in da)$  che  $A_n$  assuma valori nell'intervallo  $[a, a + da]$ , sarà possibile passare al limite continuo  $P(A_n \in da) \rightarrow p(A_n \in da)da$  per  $n \rightarrow \infty$ , con  $n$  parametro naturale da cui dipende la variabile  $A_n$ .

Nel corso della spiegazione si farà uso di alcuni esempi che aiuteranno a inquadrare intuitivamente i concetti fondamentali di questa teoria.

### 3.1 Il principio di grandi deviazioni

Il nucleo centrale di questa teoria è proprio questo principio ([8]), nel quale si evidenzia come nei processi casuali la probabilità di grandi fluttuazioni da un comportamento medio

decada esponenzialmente (col tempo nel caso di un processo stocastico, o con il numero di osservazioni nel caso di una semplice variabile casuale).

**Teorema 3.1.1** (Large deviation principle LDP). *Sia data  $A_n$  variabile casuale dipendente dal parametro  $n \in \mathbb{N}$ , e sia  $P(A_n \in da)$  la probabilità che  $A_n$  assuma un valore nell'intervallo  $[a, a + da]$  con  $a \in \mathbb{R}^d$ . Se esiste*

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \ln P(A_n \in da) = I(a) \quad (3.1)$$

con  $I(a) : \mathbb{R}^d \supseteq D \rightarrow [0, +\infty[$  funzione continua, allora si dice che  $A_n$  soddisfa il principio di grandi deviazioni con rate function  $I(a)$ .

Il significato di questo limite è proprio che per  $n$  molto grande il comportamento dominante della probabilità  $P_n$  è esponenzialmente decrescente nella prima potenza di  $n$ . Infatti dire che  $A_n$  soddisfa al 3.1.1 equivale a dire che la probabilità  $P_n$  può essere scritta come:

$$P(A_n \in da) = e^{-nI(a) + o_{+\infty}(n)} \quad (3.2)$$

ove  $o_{+\infty}(n)$  sono termini trascurabili rispetto ad  $n$  per  $n \rightarrow \infty$ . Facendo il limite come in 3.1 si ottiene proprio la rate function  $I(a)$ .

Tutto ciò giustifica la notazione più immediata,  $P_n \underset{n \rightarrow +\infty}{\approx} e^{-nI}$  che in un contesto di grandi deviazioni viene tradizionalmente scritta:

$$P_n \underset{n \rightarrow +\infty}{\asymp} e^{-nI}. \quad (3.3)$$

Il limite nella 3.1 sottintende che la rate  $I(a)$  debba essere positiva, in caso contrario si otterrebbe un comportamento divergente della probabilità per  $n \rightarrow \infty$ , inoltre si hanno problemi in cui  $I(a)$  è identicamente 0 o nel caso in cui sia  $+\infty$ . Si assiste allora rispettivamente a comportamenti *sub-esponenziali* o *super-esponenziali*, di non interesse per la nostra trattazione. Ciò che invece è importante è che la  $I(a)$  possa avere un numero finito di punti in cui vale 0, infatti sono questi i punti che  $A_n$  assume con maggiore probabilità per  $n \rightarrow +\infty$ .

Tutto il discorso fatto fino adesso per una variabile casuale discreta, considerando quindi la sua distribuzione di probabilità, si può estendere a variabili casuali continue, considerandone la densità di probabilità. È sufficiente esplicitare la probabilità  $P(A_n \in [a, a + da]) = p_n(a)da$  e riprendere tutto ciò sin qui fatto per mostrare che se  $A_n$  soddisfa il principio di grandi deviazioni allora  $p_n(a) \underset{n \rightarrow +\infty}{\asymp} e^{-nI(a)}$ .

Si può riassumere quanto finora riportato dicendo che tutte le variabili casuali  $A_n$  che soddisfano il LDP 3.1.1 sono accomunate dallo stesso comportamento universale, l'andamento esponenzialmente decrescente della loro probabilità (o della loro pdf) per grandi  $n$ , ciò che invece caratterizza tale andamento per ciascuna  $A_n$  è la sua rate function. Conoscere la rate  $I(a)$  di  $A_n$  equivale a conoscere l'andamento asintotico di  $A_n$ .

Si vogliono trovare quindi dei metodi per capire se una data variabile casuale soddisfa il LDP e successivamente ricavarne la sua rate function, a tal fine si possono seguire due strade ([8]):

- i. **metodo diretto:** si esplicita  $P_n$  o  $p_n$  e si ricava tramite opportune approssimazioni nel limite  $n \rightarrow \infty$  una forma di grandi deviazioni analoga a 3.1;
- ii. **metodo indiretto:** ci si avvale del *Teorema di Gärtner-Ellis*.

### 3.1.1 Esempio *random bits*

Si consideri una variabile casuale discreta  $B$ , il nostro bit, che può assumere valore nello spazio degli stati  $\Omega_B = \{0,1\}$  con la stessa probabilità (ad esempio il lancio di una moneta non truccata), si consideri ora la sequenza  $b = (B_1, B_2, \dots, B_n)$  di  $n$  estrazioni indipendenti di  $B$ . Definiamo la media aritmetica  $R_n = \frac{1}{n} \sum_{i=1}^n B_i$ .

Si vuole conoscere la probabilità  $P(R_n = r)$  che  $R_n$  assuma uno dei valori  $r = \{0, 1/n, 2/n, \dots, 1\}$ . Dal momento che la probabilità di ottenere una definita sequenza  $b$  è  $P(b) = 1/2^n$ , e poiché i lanci sono indipendenti si ha che  $P(R_n = r) = \sum_{b|R_n(b)=r} P(b) = \sum_{b|R_n(b)=r} 1/2^n$ , somma quindi dei contributi di tutte le sequenze che vanno bene per dare come media aritmetica  $r$ .

La probabilità cercata equivale alla probabilità di ottenere  $k = rn$  bit uguali a 1 in  $n$  estrazioni, la quale è espressa dalla binomiale:

$$P(k, n) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{con } p = 1/2$$

Per quanto appena detto si ottiene quindi:

$$P(R_n = r) = \frac{1}{2^n} \frac{n!}{(rn)!(1-r)n!}.$$

Considerando ora un grande numero di lanci  $n \rightarrow +\infty$  si può approssimare l'espressione della probabilità  $P(R_n = r)$  tramite la *formula di Stirling*  $n! \approx n^n e^{-n}$ . Si ricava allora che il contributo dominante della probabilità ha una forma:

$$P(R_n = r) \underset{n \rightarrow +\infty}{\approx} e^{-nI(r)} \quad \text{con } I(r) = \ln 2 + t \ln(r) + (1-r) \ln(1-r) \quad r \in [0,1].$$

Questa approssimazione è un chiaro esempio di grande deviazione. Si osserva che  $I(r)$  è positiva e strettamente convessa per tutti i  $t \in [0,1]$  e che il suo unico minimo nonchè zero si ha a  $r = 1/2$ .

## 3.2 Il teorema di Gärtner-Ellis

Come anticipato, tramite il teorema di Gärtner-Ellis siamo in grado di studiare tramite un metodo indiretto la rate function di una variabile casuale che soddisfi al LDP 3.1.1. Nucleo di questo teorema è la conoscenza di un'importante funzione di  $A_n$ , la *Scaled Cumulant Generating Function (SCGF)*  $\lambda(k)$  di  $A_n$ , definita come:

$$\lambda(k) = \lim_{n \rightarrow +\infty} \ln \langle W_n(k) \rangle \tag{3.4}$$

con

$$W_n(k) = e^{nkA_n} = \int_{\mathbb{R}^d} e^{nka} p_n(a) da \quad \text{ove } k \in D \subseteq \mathbb{R}^d. \quad (3.5)$$

Nella 3.4 appare la  $W_n(k)$ , la *funzione generatrice dei momenti* di  $A_n$ .

**Teorema 3.2.1** (Teorema di Gärtner-Ellis GE). *Sia  $A_n$  una variabile casuale che assume valori in  $\mathbb{R}^d$  dipendente dal parametro  $n \in \mathbb{N}$  e sia data la sua SCGF  $\lambda(k)$ . Se  $\lambda(k)$  esiste ed è differenziabile per ogni  $k \in D \subseteq \mathbb{R}^d$ , allora  $A_n$  soddisfa il LDP 3.1.1 e la sua rate function  $I(a)$  è data dalla  $\lambda^*(k)$ , trasformata di Legendre-Fenchel di  $\lambda(k)$ :*

$$I(a) = \lambda^*(k) = \sup_{k \in D} \{k \cdot a - \lambda(k)\} \quad (3.6)$$

con  $k \cdot a$  prodotto scalare in  $\mathbb{R}^d$  tra  $k$  e  $a$ .

Il grosso vantaggio derivante dal Teorema di Gärtner-Ellis 3.2.1 è quello di riuscire a stabilire il LDP 3.1.1 e la rate function per una data variabile casuale senza conoscerne esplicitamente la densità di probabilità ([8]). Tuttavia, non sempre  $\lambda(k)$  esiste, e inoltre non tutte le diverse tipologie di rate functions possono essere ottenute tramite la trasformata di Legendre-Fenchel di  $\lambda(k)$ . Prima di terminare è importante notare che nel caso in cui  $\lambda(k)$  oltre ad essere differenziale è anche strettamente convessa, la trasformata di Legendre-Fenchel si riduce alla trasformata di Legendre:

$$I(a) = \{a \cdot k(a) - \lambda(k(a))\} \quad (3.7)$$

con  $k(a) \in \mathbb{R}^d$  l'unica radice dell'equazione  $\nabla \lambda(k) = a$ . Circostanza che nelle applicazioni spesso si verifica.

### 3.2.1 Teorema di Cramér

Nel caso in cui la variabile casuale  $A_n$  sia una *media aritmetica* di una sequenza di variabili indipendenti e identicamente distribuite (IID) si ricade nel risultato classico del *teorema di Cramér*.

Ovvero, data  $A_n = \frac{1}{n} \sum_{i=1}^n X_i$  ove le  $X_i$  sono variabili casuali IID, la SCGF si ricava facilmente:

$$\lambda(k) = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \langle e^{nk \frac{1}{n} \sum_{i=1}^n X_i} \rangle \stackrel{IND.}{=} \lim_{n \rightarrow +\infty} \frac{1}{n} \prod_{i=1}^n \langle e^{kX_i} \rangle \stackrel{ID.}{=} \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \langle e^{kX_i} \rangle^n = \ln \langle e^{kX} \rangle$$

ove  $X$  è una qualsiasi delle  $X_i$ . La SCGF così ottenuta è differenziabile e quindi rispetta le ipotesi del teorema GE 3.2.1, per ottenere la rate function basterà quindi calcolare la trasformata di Legendre-Fenchel di  $\lambda(k) = \ln \langle e^{kX} \rangle$ . Possiamo vederne un'immediata applicazione con una pdf gaussiana:

$$\lambda(k) = \ln \langle e^{kX} \rangle = \ln \int_{\mathbb{R}} e^{kx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{x-\mu}{\sqrt{2\sigma}}\right)^2} dx,$$

e svolgendo l'integrale a ultimo membro si ottiene  $\lambda(k) = k\mu + \frac{\sigma^2}{2}k^2$  ove  $k \in \mathbb{R}$ . Come previsto la SCGF è differenziabile e poichè anche strettamente convessa ( $\lambda''(k) = \sigma^2 > 0 \quad \forall k \in \mathbb{R}$ ) il calcolo della trasformata di Legendre-Fenchel si riduce a quello della trasformata di Legendre:

$$I(s) \underset{k \in \mathbb{R}}{=} \{sk(s) - \lambda(k(s))\} = \left(\frac{s - \mu}{\sqrt{2\sigma}}\right)^2 \quad \text{con } k(s) = (\lambda')^{-1}(s) = \frac{s - \mu}{\sigma^2}.$$

### 3.3 Proprietà della $\lambda(k)$ e della rate $I(a)$

Come mostrato precedentemente, il teorema di GE 3.2.1 permette di ricavare la  $I(a)$  per  $A_n$  attraverso la trasformata di Legendre-Fenchel senza conoscere esplicitamente la pdf  $p_n(a)$ , inoltre il calcolo analitico della rate risulta agevole, come evidenziato nell'esempio della pdf gaussiana.

Tuttavia, ci sono casi in cui la  $\lambda(k)$  non esiste, tipicamente sono quelli in cui la  $A_n$  non verifica il LDP, ma è da notare che la non esistenza della SCGF non è comunque una condizione sufficiente a stabilire che la  $A_n$  non soddisfa il LDP. Inoltre può essere che non tutte le rate functions siano calcolabili come trasformate di Legendre-Fenchel delle  $\lambda(k)$ . I casi particolari sono tanti, ma noi ci concentriamo solo su quelli che spiegano le ipotesi del teorema di GE 3.2.1.

La differenziabilità di  $\lambda(k)$  è una condizione sufficiente ma non necessaria per avere che  $I = \lambda^*$ , infatti per funzioni convesse, ma non strettamente, si ha comunque  $I = \lambda^*$ , ma  $\lambda = I^*$  non è differenziabile. L'unica affermazione forte che si può fare che spieghi le ipotesi del teorema di GE è la seguente: *sia  $\lambda(k)$  differenziabile,  $I$  è strettamente convessa se e solo se  $I = \lambda^*$ .* Ad ogni modo l'utilità di questa affermazione è di mettere in luce anche il principale limite del teorema, ovvero che permette di calcolare rates  $I$  che siano solo strettamente convesse.

Concentriamoci ora sulle proprietà ([8]) che derivano dalle trasformate di Legendre-Fenchel (o di Legendre semplicemente), nel caso in cui la  $\lambda(k)$  oltre a essere *differenziabile* è anche *strettamente convessa*. In questa situazione ci si restringe al caso in cui per calcolare la rate  $I$  si utilizza la trasformata di Legendre. Ebbene la  $\lambda'(k)$  è monotona crescente, quindi la funzione  $k(a)$  che soddisfa alla  $\lambda'(k(a)) = a$  può essere invertita al fine di ottenere la  $a(k)$  soddisfacente la  $\lambda'(k) = a(k)$ . Da questa equazione, applicando la trasformata di Legendre si ottiene  $I'(a(k)) = k$  e  $I'(a) = k(a)$ . Quindi, in questo caso, e solamente in questo caso, la pendenza di  $\lambda$  è in relazione biunivoca con la pendenza di  $I$ , questa proprietà prende il nome di *proprietà duale* ed è riassunta graficamente nella figura 3.1.

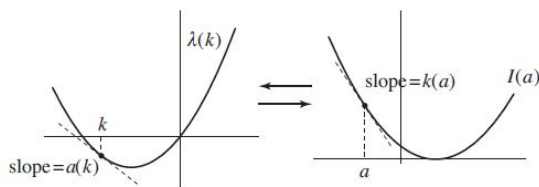


Figura 3.1: Relazione tra pendenze di  $\lambda(k)$  e  $I(a)$

Sempre tramite questa *proprietà duale* si può vedere come i punti critici di  $\lambda(k)$  ove la funzione diverge vengono traslati dalla trasformata di Legendre in code della rate function  $I(a)$  che sono lineari o asintoticamente lineari. I ragionamenti di sfondo sono gli stessi fatti poc'anzi, riportiamo di seguito nella figura 3.2 un riassunto grafico.

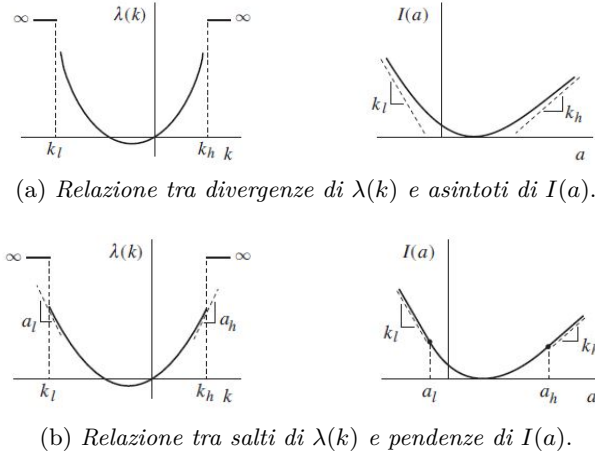


Figura 3.2: Relazione tra divergenze e salti della  $\lambda(k)$  con asintoti e pendenze della  $I(a)$

Per concludere analizziamo due interessanti proprietà puntuali che ci saranno molto utili nel seguito. Dal fatto che la probabilità è normalizzata:

$$\lambda'(0) = \lim_{n \rightarrow +\infty} \left. \frac{\langle A_n e^{nkA_n} \rangle}{\langle e^{nkA_n} \rangle} \right|_{k=0} = \lim_{n \rightarrow +\infty} \langle A_n \rangle_{\text{medie IID}} = \mu, \quad (3.8)$$

$$\lambda''(0) = \lim_{n \rightarrow +\infty} n \left( \langle A_n^2 \rangle - \langle A_n \rangle^2 \right) = \lim_{n \rightarrow +\infty} n \text{var}(A_n)_{\text{medie IID}} = \sigma^2. \quad (3.9)$$

Ovvero, nel caso in cui  $A_n$  è una variabile casuale che descrive la media semplice tra variabili casuali IID,  $\lambda'(0)$  e  $\lambda''(0)$  sono proprio, rispettivamente, la media e la varianza di  $A_n$ . Applicando a ciò la *proprietà duale* introdotta in precedenza (valida per  $\lambda(k)$  strettamente convessa) si deduce che il valore della media  $\mu$  è zero e punto di minimo della rate function  $I(a)$ , mentre la varianza  $\sigma^2$  è legata all'ampiezza della rate function.

### 3.4 Grandi deviazioni nel non equilibrio

Come introdotto nel capitolo precedente sulla meccanica statistica di non equilibrio, sono i processi stocastici, e tra questi i processi di Markov, che descrivono in maniera utile le osservabili di un sistema fisico fuori dall'equilibrio. In questa sezione quindi ci proponiamo di indagare come la teoria delle grandi deviazioni si unisca allo studio dei processi di Markov al fine di ottenere informazioni utili sulle osservabili di nostro interesse. Per quello che abbiamo visto poc'anzi, l'informazione essenziale che ci deriva dallo studio della teoria



delle grandi deviazioni, è contenuta tutta nella rate function, dato che questa funzione ci permette di capire quale sia il comportamento asintotico di un'osservabile. Vediamo quindi come sia possibile unificare il tutto. [8][9] In analogia a quanto fatto in precedenza nella sezione 2.4 trattiamo catene di Markov finite, a stati discreti e a tempo discreto. Consideriamo un processo stocastico  $Y_n = (y_1, y_2, \dots, y_n)$ , ove la  $n$  a pedice indica gli steps temporali e lo spazio degli stati discreto è il solito  $\Omega_Y$ . Siamo interessati allo studio dell'osservabile  $S_n = \frac{1}{n} \sum_{i=1}^n f(y_i)$ , che consiste in una media aritmetica che coinvolge una funzione arbitraria  $f : \Omega_Y \rightarrow \mathbb{R}^d$ ,  $d \geq 1$ .

Assumiamo quindi che  $Y_n$  formi una HMC e quindi obbedisca alla relazione 2.16, ovvero:

$$P(Y_n) = \nu(y_1) \prod_{i=2}^n \pi(y_i | y_{i-1}), \quad (3.10)$$

ove  $\nu(y_1)$  indica la distribuzione iniziale e  $\pi(y_i | y_{i-1})$  la probabilità di transizione condizionata di passare da  $y_{i-1}$  a  $y_i$ . Nel derivare il LDP 3.1.1 per  $S_n$  si procede secondo le linee guida dettate dal teorema di GE 3.2.1, quindi prima di tutto si calcola la  $\lambda(k)$ , e per farlo si parte da:

$$\begin{aligned} \langle W_n(k) \rangle &= \langle e^{nk \cdot S_n} \rangle = \sum_{y_1, y_2, \dots, y_n} \nu(y_1) e^{k \cdot f(y_1)} \pi(y_2 | y_1) e^{k \cdot f(y_2)} \dots \pi(y_n | y_{n-1}) e^{k \cdot f(y_n)} \\ &= \sum_{y_1, y_2, \dots, y_n} \nu_k(y_1) \pi_k(y_2 | y_1) \pi_k(y_n | y_{n-1}), \end{aligned} \quad (3.11)$$

definendo  $\nu_k(y_1) = \nu(y_1) e^{k \cdot f(y_1)}$  e  $\pi_k(y_i | y_{i-1}) = \pi(y_i | y_{i-1}) e^{k \cdot f(y_i)}$ . Si riconosce nella seconda equazione una sequenza di prodotti matriciali che coinvolgono  $\nu_k(y_1)$  e la matrice di transizione  $\pi_k(y_i | y_{i-1})$ . Per una riscrittura più esplicita, ricordando 2.17 e denotando  $(\nu_k)_i = \nu_k(y_1 = i)$  e  $(\Pi_k)_{ji} = \pi_k(j | i)$ , si ottiene:

$$\langle W_n(k) \rangle = \sum_{j \in \Omega_Y} \left( \Pi_k^{n-1} \nu_k \right)_j. \quad (3.12)$$

L'espressione  $\lambda(k)$ , ricordando la 3.4, si ricava da quest'ultima espressione determinando il comportamento asintotico del prodotto  $\left( \Pi_k^{n-1} \nu_k \right)_j$  facendo uso della teoria di Perron-Frobenius applicata alla matrice  $\Pi_k$  ([8]). Per quanto ci riguarderà nel seguito, trattiamo il solo caso in cui la matrice  $\Pi_k$  è ergodica (irriducibile e aperiodica, si veda 2.4), in tal caso vi è un'unica distribuzione di probabilità stazionaria  $\nu^*$ , associata all'autovalore 1 di  $\Pi$ , tale che  $\Pi \nu^* = \nu^*$ . La sola distribuzione di probabilità stazionaria indica per  $\Pi_k$  la presenza di un solo *autovalore dominante* (autovalore di modulo maggiore)  $\zeta(\Pi_k)$  che stabilisce la dinamica del sistema. Questo autovalore dominante della matrice  $\Pi_k$  sarà quello che per  $k = 0$ , caso in cui la matrice  $\Pi_k$  si riduce alla matrice  $\Pi$ , varrà 1. A questo punto segue immediatamente che:

$$\langle e^{nk \cdot S_n} \rangle \asymp \zeta(\Pi_k)^n \quad (3.13)$$

e quindi per 3.4 si ottiene:

$$\lambda(k) = \ln \zeta(\Pi_k). \quad (3.14)$$

Si potrebbe mostrare che avendo assunto  $\Pi_k$  finita,  $\zeta(\Pi_k)$  è una funzione analitica in  $k$ . Possiamo concludere applicando alla SCGF 3.14 il teorema di GE 3.2.1, ottenendo quindi tramite trasformata di Legendre-Fenchel la rate function:

$$I(s) = \sup_k \{k \cdot s - \ln \zeta(\Pi_k)\} \quad (3.15)$$

### 3.4.1 Esempio *Markov bits*

Per ragioni di chiarezza riportiamo nel seguito un semplice esempio di quanto appena discusso. Riprendiamo dall'esempio dei *bits* trattati precedentemente in cui si aveva un processo stocastico del tipo  $b_n = (B_1, B_2, \dots, B_n)$ , assumiamo inoltre il seguente grafo della catena di Markov ( $\alpha \in [0,1]$ ):

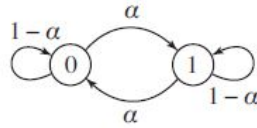


Figura 3.3: Grafo di Markov per random bits

La matrice di transizione associata alla catena di Markov è:

$$\Pi = \begin{pmatrix} \pi(0|0) & \pi(0|1) \\ \pi(1|0) & \pi(1|1) \end{pmatrix} = \begin{pmatrix} 1 - \alpha & \alpha \\ \alpha & 1 - \alpha \end{pmatrix}$$

e si cerca l'autovalore dominante di:

$$\Pi_k = \begin{pmatrix} (1 - \alpha)e^0 & (\alpha)e^0 \\ (\alpha)e^k & (1 - \alpha)e^k \end{pmatrix}.$$

Ove si sono moltiplicati i singoli termini della  $\Pi$  per i fattori  $e^{kf(y_i)}$ , in particolare nel caso in cui ci si trovi nello stato 0 allora  $f(0) = 0$  e nel caso ci si trovi nello stato 1 allora  $f(1) = 1$ .

Tutto al fine di ottenere la rate function che riportiamo accanto alla  $\lambda(k)$  nella figura 3.4 sottostante:

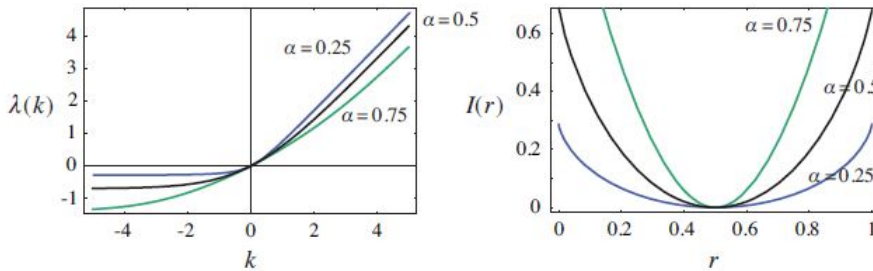


Figura 3.4:  $\lambda(k)$  e rate  $I(r)$  per Markov bits

## Capitolo 4

# Applicazione ad un modello stocastico di conduzione

### 4.1 Generalità dello studio

Per mostrare come la teoria delle grandi deviazioni sia un'utile strumento nello studio di sistemi fuori equilibrio, abbiamo deciso di applicarla allo studio di un modello stocastico di conduzione.

Immaginiamo quindi di avere un canale, il cui interno non è però cavo e liscio, ma ci sono delle pareti sporgenti che creano delle trappole. All'interno del tubo si posiziona una particella e vi si applica un campo di forza esterno, la particella è quindi condizionata nel suo movimento, il quale avviene preferenzialmente lungo la direzione del campo e non si ha il caso già discusso di moto browniano (per il quale appunto il campo esterno è nullo). Per un tale sistema, non in equilibrio a causa del campo esterno, si è studiato come varia la velocità della particella in funzione del campo esterno applicato, intuitivamente dovrebbe risultare che la velocità media della particella aumenta all'aumentare del campo esterno. Tuttavia non è così e si ottiene l'andamento mostrato nella figura 4.1 sottostante:

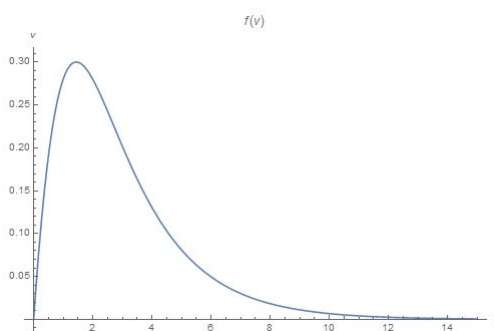


Figura 4.1: Andamento della velocità in funzione del campo

Essenzialmente si nota che solo per un certo range ristretto di forze la velocità aumenta con il campo, poi la presenza delle trappole modifica l'andamento e fa' si che da una certa

forza in poi la velocità diminuisca invece di crescere, si tratta del fenomeno della *mobilità differenziale negativa*.

Ciò che si è fatto è stato di studiare tale sistema, che come vedremo poi può essere rappresentato da una catena di Markov omogenea (si veda il 2.4), tramite la teoria delle grandi deviazioni, come presentato nel 3.4. Essenzialmente, si sono considerati tre valori di campo peculiari, un valore nel range di diretta proporzionalità, il valore che identifica il massimo della curva, e un valore nel range di mobilità differenziale negativa. Per ciascuno di questi valori si è studiato poi l'andamento della velocità, velocità media e fluttuazioni dalla media, tramite la rate function.

Il fine ultimo è stato quindi quello di ottenere delle rate functions e di studiarle al variare di determinati parametri (presentati nel seguito) per ottenere informazioni sulla conduzione nel canale, attraverso lo studio della velocità della particella.

Prima di cominciare è doveroso sottolineare che, pur essendo fuori equilibrio, gli stati che andremo a studiare sono stati stazionari. Le velocità delle particelle sono quindi pertinenti a questi stati, la cui presenza è facilitata dall'instaurare condizioni periodiche al contorno. L'assunzione che si fa per restare in uno stato stazionario è che i tempi d'interazione delle particelle in movimento all'interno del canale, con il bagno termico, siano di ordini di grandezza superiori rispetto ai tempi d'interazione della particelle con il campo. Questo ci permette di considerare valida l'equazione di bilancio dettagliato generalizzato 2.27 nella trattazione seguente.

## 4.2 Il modello

Come anticipato il sistema oggetto di studio è costituito da un canale al cui interno, ad intervalli regolari sono presenti delle trappole. All'interno del canale le particelle si muovono soggette ad un campo di forze esterno ( $\vec{f}$ ). Uno schema visivo è riportato nella figura 4.2 sottostante.

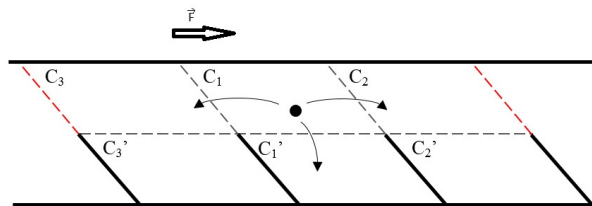


Figura 4.2: Schema del sistema

Nello schema sono visualizzate sei differenti posizioni in cui la particella può trovarsi, le distinguiamo con una lettera, un numero e la presenza o meno di un apice. In questo caso particolare la particella è stata posta nello stato  $C_1$ , com'è evidente dalle frecce la particella può muoversi lungo due assi, l'asse verticale o quello orizzontale, movimenti in diagonale non sono permessi. Sebbene esplicativa, questa rappresentazione non è quella minimale, infatti si nota che la struttura più piccola che si ripete è costituita da due soli

stati, uno di conduzione nella parte superiore del canale e una trappole sottostante, tuttavia, come sarà chiaro a posteriori la presenza dei sei stati è concettualmente importante per la costruzione della matrice  $\Pi_k$ . Inoltre consideriamo il canale di lunghezza infinita, quindi poniamo delle condizioni periodiche al contorno (segnalate da trattini rossi nella figura 4.2), in altre parole se la particella si dovesse trovare in  $C_2$  allora potrebbe: muoversi a sinistra e tornare in  $C_1$ , in basso e finire in  $C'_2$  oppure anche a destra e spostarsi in  $C_3$ . Ragionamento analogo vale se la particella si dovesse trovare in  $C_3$ .

Cominciamo a definire quantitativamente il nostro sistema associando alle possibili transizioni che la particella può fare delle probabilità che esse avvengano. Supponiamo per il momento di trovarci nella banda di conduzione del canale, qui sarà ovviamente più probabile che la particella si muova in direzione concorde al campo piuttosto che contro il campo, associamo quindi alla probabilità per unità di tempo di muoversi verso destra il termine di rate  $W_r$  e alla probabilità per unità di tempo di muoversi verso sinistra il termine di rate  $W_l$ . Inoltre una particella nella banda di conduzione ha anche la possibilità di muoversi verso il basso e di entrare quindi in una trappola. Potrebbe sembrare che il movimento lungo l'asse perpendicolare alla conduzione non sia dipendente dal campo, ma la dipendenza è invece esplicita avendo disegnato le trappole inclinate (si veda 4.2), in tal caso sarà più probabile entrare in una trappola che uscirne. Definiamo  $W_d$  la rate di transizione per una particella che si muove verso il basso e  $W_u$  per una particella che si muove verso l'alto. In maniera abbastanza arbitraria si decidono dei valori per le rates, al fine di una maggior chiarezza riporto i valori scelti nella tabella sottostante, le colonne indicano a quale rate ci stiamo riferendo, le righe indicano lo stato in cui si può trovare la particella:

	$W_r$	$W_l$	$W_u$	$W_d$
$C_1$	$e^{\frac{f}{2}}$	$e^{-\frac{f}{2}}$	0	$\epsilon e^{2(1-a)\frac{f}{2}}$
$C_2$	$e^{\frac{f}{2}}$	$e^{-\frac{f}{2}}$	0	$\epsilon e^{2(1-a)\frac{f}{2}}$
$C_3$	$e^{\frac{f}{2}}$	$e^{-\frac{f}{2}}$	0	$\epsilon e^{2(1-a)\frac{f}{2}}$
$C'_1$	0	0	$\epsilon e^{-2a\frac{f}{2}}$	0
$C'_2$	0	0	$\epsilon e^{-2a\frac{f}{2}}$	0
$C'_3$	0	0	$\epsilon e^{-2a\frac{f}{2}}$	0

Tabella 4.1: Rates del modello

Come anticipato le rates si possono scegliere in maniera abbastanza arbitraria, con l'unico vincolo di rispettare la condizione locale di bilancio dettagliato generalizzato. Dalla 2.27 si ricava che ([7]):

$$\frac{P^s(y, t)}{P^s(y', t + \delta t)} = \frac{W(y|y')}{W(y'|y)} = e^f \quad (4.1)$$

ove  $f$  è proporzionale al campo applicato. Quindi si dovrà avere:  $\frac{W_r}{W_l} = \frac{W_d}{W_u} = e^f$ .

Discutiamo ora brevemente il significato delle rates. Ciascuna rate identifica la probabilità per unità di tempo di effettuare un certo spostamento, ogni rate è dipendente dal campo e lo spostamento concorde al campo è rappresentato da una rate di modulo maggiore rispetto allo spostamento nella stessa direzione ma opposto al campo. Inoltre, nel caso degli spostamenti in ingresso e in uscita dalle trappole abbiamo introdotto due parametri:  $a$  e  $\epsilon$ . La loro presenza modula la probabilità di entrare ed uscire dalle trappole. Facendo un'analogia possiamo immaginare che il passaggio da uno stato di conduzione alla trappola sottostante sia rappresentato da una strada a due corsie, il termine  $\epsilon$  si occupa di stringere ed allargare l'intera strada, lasciando invariate le dimensioni relative delle due corsie. Quindi, contemporaneamente, aumenta, all'aumentare di  $\epsilon$  o diminuisce, al diminuire di  $\epsilon$ , le probabilità di entrare e di uscire da una trappola. Leggermente differente è il significato di  $a$ , il quale ha la funzione di allargare o stringere sempre entrambe le corsie, ma in funzione di quanto sia forte il campo  $f$ . In parole povere  $a$  e  $\epsilon$  agiscono entrambi allargando o stringendo la strada e lasciando invariato il rapporto tra le due corsie, tuttavia a differenza di  $\epsilon$  che è costante, il contributo di  $a$  dipende da quanto intenso è il campo. Operando a stati discreti e a tempo discreto e notando che la probabilità che avvenga una determinata transizione dipende solo dallo stato precedentemente occupato possiamo modellizzare il nostro sistema tramite una catena di Markov  $Y_n$  (2.4) e in particolare tramite un processo di Markov omogeneo. Di seguito ne riportiamo il grafo:

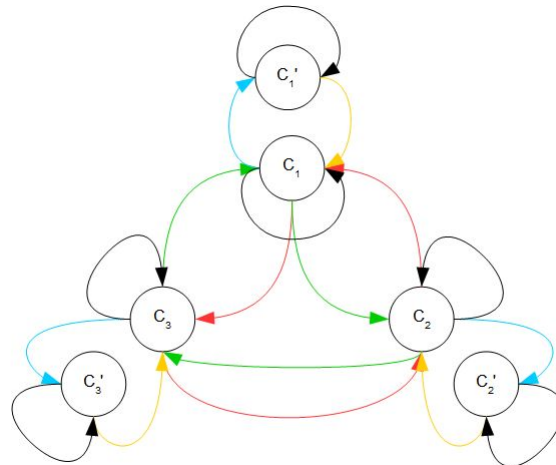


Figura 4.3: Grafo del processo di Markov di conduzione stocastica

Il grafo mostra gli stati e tramite delle frecce le possibili transizioni che possono avvenire. Per chiarezza i colori delle frecce sono gli stessi utilizzati nella tabella 4.1. In azzurro sono mostrate le frecce che da uno stato di conduzione portano in una trappola, in arancione quelle che da una trappola riportano in uno stato di conduzione, in nero le transizioni che da uno stato portano nello stesso stato, in verde le transizioni nella banda di conduzione che sono concordi al campo e in rosso quelle discordi.

Al fine di ottenere una matrice di transizione del tipo 2.13 dobbiamo sostituire le rates con delle probabilità secondo la 2.20 (come alla fine del 2.5). Per ottenere delle probabilità

è quindi sufficiente moltiplicare le rates per un  $\delta t$  sufficientemente piccolo. Ne risulta la matrice di transizione seguente:

$$\Pi = \begin{pmatrix} \pi(C_1|C_1) & \pi(C_1|C_2) & \pi(C_1|C_3) & \pi(C_1|C'_1) & \pi(C_1|C'_2) & \pi(C_1|C'_3) \\ \pi(C_2|C_1) & \pi(C_2|C_2) & \pi(C_2|C_3) & \pi(C_2|C'_1) & \pi(C_2|C'_2) & \pi(C_2|C'_3) \\ \pi(C_3|C_1) & \pi(C_3|C_2) & \pi(C_3|C_3) & \pi(C_3|C'_1) & \pi(C_3|C'_2) & \pi(C_3|C'_3) \\ \pi(C'_1|C_1) & \pi(C'_1|C_2) & \pi(C'_1|C_3) & \pi(C'_1|C'_1) & \pi(C'_1|C'_2) & \pi(C'_1|C'_3) \\ \pi(C'_2|C_1) & \pi(C'_2|C_2) & \pi(C'_2|C_3) & \pi(C'_2|C'_1) & \pi(C'_2|C'_2) & \pi(C'_2|C'_3) \\ \pi(C'_3|C_1) & \pi(C'_3|C_2) & \pi(C'_3|C_3) & \pi(C'_3|C'_1) & \pi(C'_3|C'_2) & \pi(C'_3|C'_3) \end{pmatrix}$$

$$= \begin{pmatrix} 1-\delta t(z^{-1}+z+\epsilon z^{2(1-a)}) & \delta t z^{-1} & \delta t z & \delta t \epsilon z^{-2a} & 0 & 0 \\ \delta t z & 1-\delta t(z^{-1}+z+\epsilon z^{2(1-a)}) & \delta t z^{-1} & 0 & \delta t \epsilon z^{-2a} & 0 \\ \delta t z^{-1} & \delta t z & 1-\delta t(z^{-1}+z+\epsilon z^{2(1-a)}) & 0 & 0 & \delta t \epsilon z^{-2a} \\ \delta t \epsilon z^{2(1-a)} & 0 & 0 & 1-\delta t \epsilon z^{-2a} & 0 & 0 \\ 0 & \delta t \epsilon z^{2(1-a)} & 0 & 0 & 1-\delta t \epsilon z^{-2a} & 0 \\ 0 & 0 & \delta t \epsilon z^{2(1-a)} & 0 & 0 & 1-\delta t \epsilon z^{-2a} \end{pmatrix}$$

ove  $z = e^{\frac{f}{2}}$ .

### 4.3 Applicazione delle grandi deviazioni

Il fine ultimo è quello di derivare delle rate functions dal nostro modello markoviano di conduzione stocastica. Per farlo si applicherà passo passo quanto visto nel 3.4. L'osservabile che si vuole studiare è la velocità  $v$  della particella nel canale, le rates che si ricaveranno saranno del tipo  $I(v)$ . Tramite comparazione delle rates si vogliono ottenere informazioni sulle fluttuazioni dalla media delle velocità e sulla possibile dinamica delle particelle nelle seguenti diverse situazioni:

- i. Al variare del campo  $f$ , tenendo costante  $a$  e  $\epsilon$ . Sono stati scelti tre valori arbitrari (si confronti con 4.1)  $f_i = 0.22$  nella fase di crescita proporzionale,  $f_m = \log(2 + \sqrt{5})$  il massimo della curva e  $f_f = 4.56$  nella parte di mobilità differenziale negativa;
- ii. Al variare del parametro  $\epsilon$ , tenendo costante  $f$  e  $a$ . Si sono scelti tre parametri che rappresentassero bene i dati,  $\epsilon = \{0.1, 1, 10\}$ ;
- iii. Al variare del parametro  $a$ , tenendo costante  $f$  e  $\epsilon$ . Si sono scelti tre parametri che rappresentassero bene i dati,  $a = \{0.25, 0.5, 0.75\}$ .

Ricordiamo che la posizione è modellizzata da una catena di Markov  $Y_n$  e che, come anticipato, l'osservabile in considerazione è la velocità media  $\langle v \rangle$  della particella:

$$\langle v \rangle = \frac{1}{n} \sum_{i=1}^{n-1} f(y_i, y_{i+1}), \quad \text{si veda 2.18}$$

ove lo spazio degli stati del processo markoviano della posizione  $Y_n$  è  $\Omega_Y = \{y_1, y_2, \dots, y_6\} = \{C_1, C_2, C_3, C'_1, C'_2, C'_3\}$ , e  $f$  è funzione dello stato  $i$  e di quello precedente. Nel nostro caso

$f(y_i, y_{i-1}) = \{+1, 0, -1\}$  a seconda, rispettivamente, se il movimento è verso destra (anche nel caso di ingresso in trappole), se la particella rimane ferma, o se si muove verso sinistra (anche nel caso di uscita da trappole).

A questo punto si calcola la matrice  $\Pi_k$  e per farlo è sufficiente moltiplicare i termini della matrice  $\Pi$  per i fattori  $e^{kf(y_i, y_{i+1})}$ . Risulta:

$$\Pi_k = \begin{pmatrix} 1 - \delta t(z^{-1} + z + \epsilon z^{2(1-a)}) & \delta t z^{-1} e^{-k} & \delta t z e^k & \delta t \epsilon z^{-2a} e^{-k} & 0 & 0 \\ \delta t z e^k & 1 - \delta t(z^{-1} + z + \epsilon z^{2(1-a)}) & \delta t z^{-1} e^{-k} & 0 & \delta t \epsilon z^{-2a} e^{-k} & 0 \\ \delta t z^{-1} e^{-k} & \delta t z e^k & 1 - \delta t(z^{-1} + z + \epsilon z^{2(1-a)}) & 0 & 0 & \delta t \epsilon z^{-2a} e^{-k} \\ \delta t \epsilon z^{2(1-a)} e^k & 0 & 0 & 1 - \delta t \epsilon z^{-2a} & 0 & 0 \\ 0 & \delta t \epsilon z^{2(1-a)} e^k & 0 & 0 & 1 - \delta t \epsilon z^{-2a} & 0 \\ 0 & 0 & \delta t \epsilon z^{2(1-a)} e^k & 0 & 0 & 1 - \delta t \epsilon z^{-2a} \end{pmatrix}$$

Visto che la matrice è ergodica (si veda 2.4), facendo uso della teoria di Perron-Frobenius si dovrebbero calcolare gli autovalori per identificare quello dominante. Tutto ciò è analiticamente complesso e si hanno problemi anche tramite l'utilizzo del software *Mathematica*, si è quindi deciso di passare ad una matrice 2x2 che rappresenta una proiezione di questa 6x6 su solamente 2 stati, uno di conduzione  $C$  e una trappola  $C'$ .

Ci si potrebbe chiedere perchè non si sia considerato fin dall'inizio un processo markoviano a due stati, il motivo è che in tal caso non si sarebbe potuto calcolare agilmente la matrice  $\Pi_k$  non avendo avuto bene idea di cosa volesse dire restare fermi nello stato di conduzione (infatti a quest'evenienza sarebbe corrisposta l'effettiva immobilità della particella, ma anche i movimenti orizzontali).

Operando come si è fatto è invece chiaro il significato di ciascun termine, si ottiene:

$$\Pi_{k_p} = \begin{pmatrix} 1 + \delta t \left( z^{-1} e^{-k} + e^k z + \epsilon z^{-2a} e^{-k} - z^{-1} - z - \epsilon z^{2(1-a)} \right) & \delta t \epsilon z^{-2a} \\ \delta t \epsilon z^{2(1-a)} & 1 - \delta t \epsilon z^{-2a} \end{pmatrix}$$

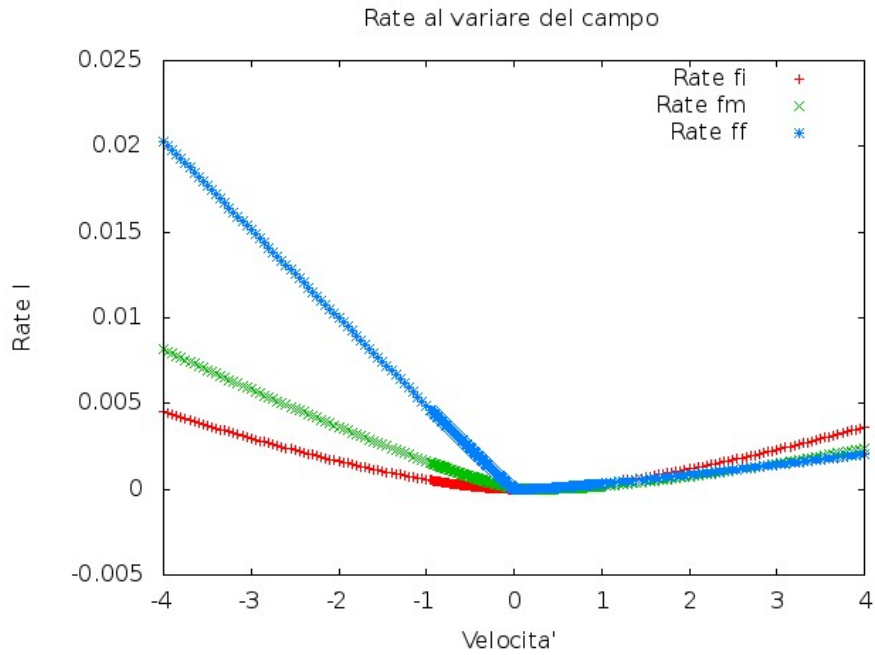
Infatti, ad esempio nel termine  $\pi_{k_p}(C|C)$  che consiste nella probabilità di restare in  $C$  non è solo presente il termine di effettiva immobilità, ma anche quelli di spostamento verso destra (derivante da  $\pi_k(C_1|C_3)$ ) e spostamento verso sinistra (derivante da  $\pi_k(C_1|C_2)$ ).

La matrice  $\Pi_{k_p}$  è chiaramente ancora ergodica e per ciascun caso riportato nell'elenco 4.3 risulta agevole con *Mathematica* il calcolo degli autovalori  $\lambda_{1,2}(k)$  e l'individuazione di quello dominante  $\zeta(\Pi_{k_p})$ .

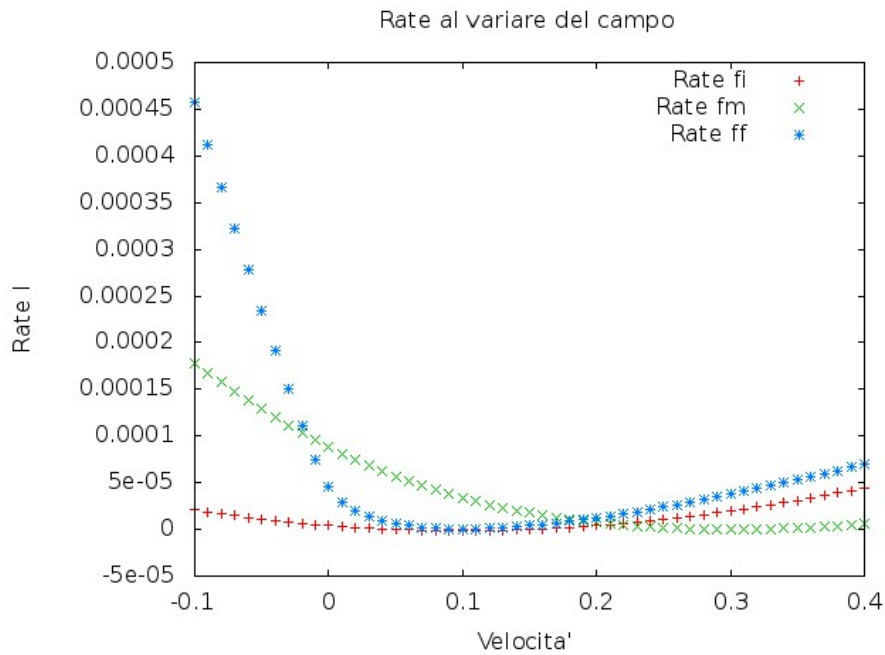
A questo punto, per ciascun caso, è sufficiente calcolare la rate  $I(v)$  tramite la 3.15. Questa operazione risultava complessa analiticamente, si è deciso quindi di scrivere un programma in  $C++$ , il quale, essenzialmente, una volta fissato un valore di  $v$  cerca il  $k$  che massimizza  $kv - \ln \zeta(\Pi_{k_p})$  e ripete quest'operazione per tutti i  $v$  necessari a capire l'andamento della rate function.

Di seguito si riportano, per ciascun caso, i grafici che visualizzano l'andamento generale delle rate functions, fiancheggiati dai grafici che ingrandiscono la regione dei minimi per verificare se effettivamente i minimi e zeri corrispondono alle velocità medie previste (secondo quanto abbiamo visto dalle proprietà duali nel 3.3), in base al valore del campo, dal grafico 4.1.



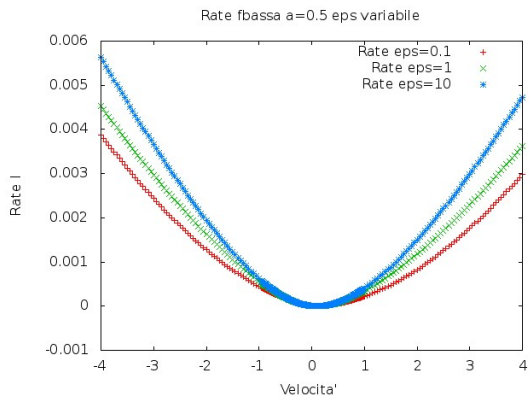


(a) *Andamento generale.*

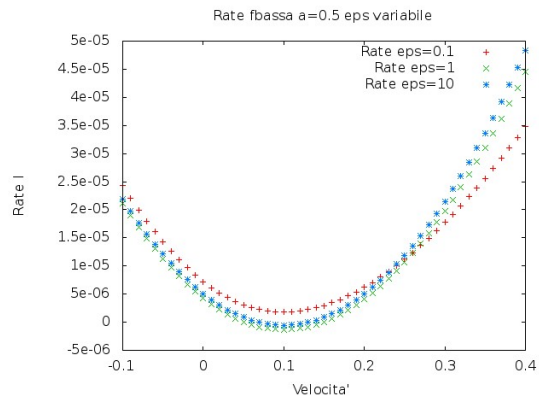


(b) *Studio dei minimi.*

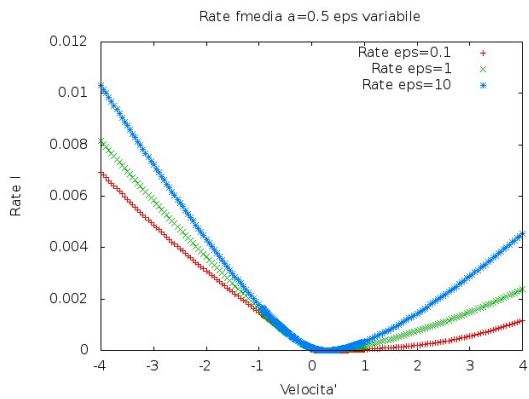
Figura 4.4: Rate al variare del campo  $f_i < f_m < f_f$



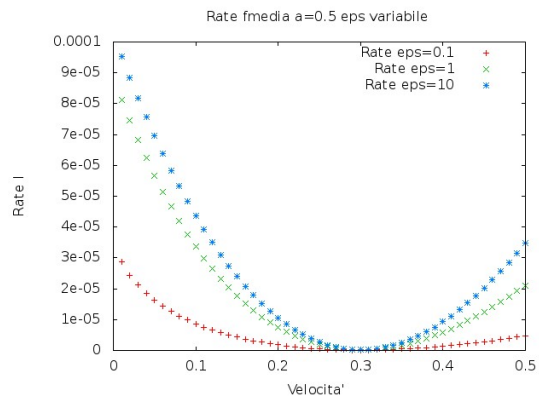
(a) Andamento generale per  $f_i$  ad  $\epsilon$  variabile.



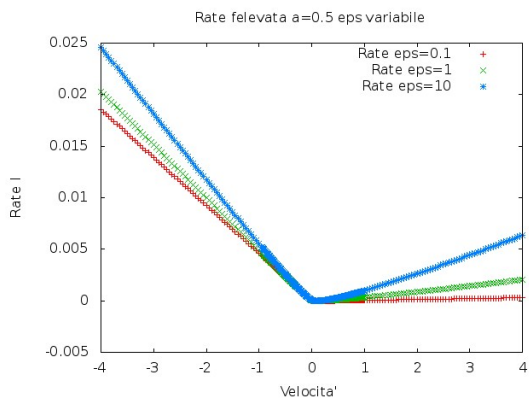
(b) Studio dei minimi per  $f_i$  ad  $\epsilon$  variabile.



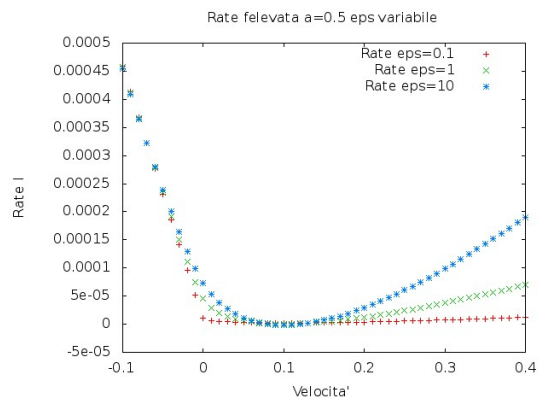
(c) Andamento generale per  $f_m$  ad  $\epsilon$  variabile.



(d) Studio dei minimi per  $f_m$  ad  $\epsilon$  variabile.

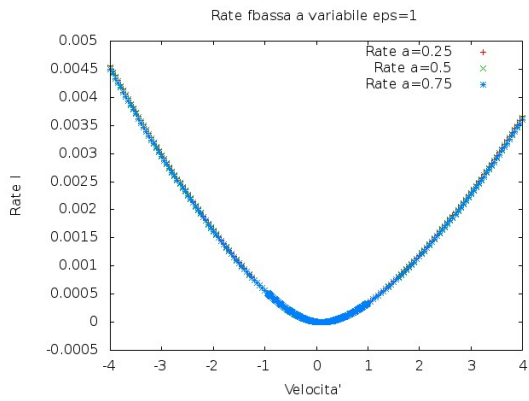


(e) Andamento generale per  $f_f$  ad  $\epsilon$  variabile.

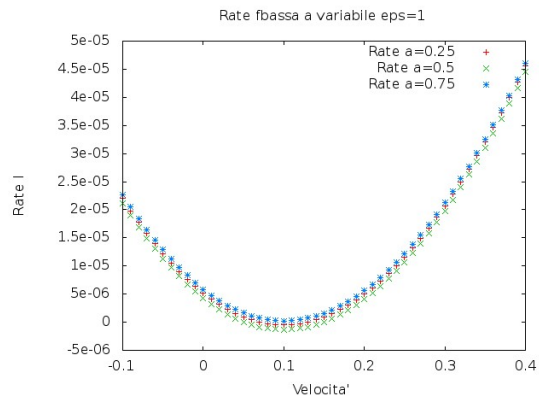


(f) Studio dei minimi per  $f_f$  ad  $\epsilon$  variabile.

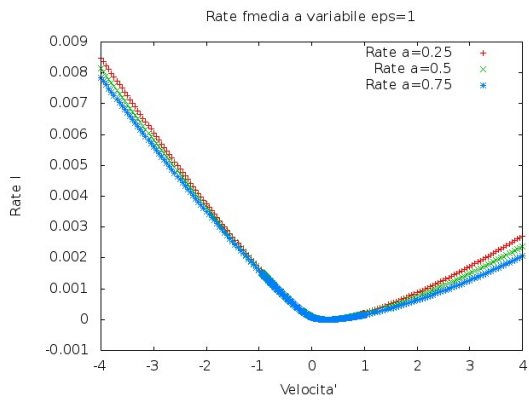
Figura 4.5: Rates al variare di  $f$  e di  $\epsilon$



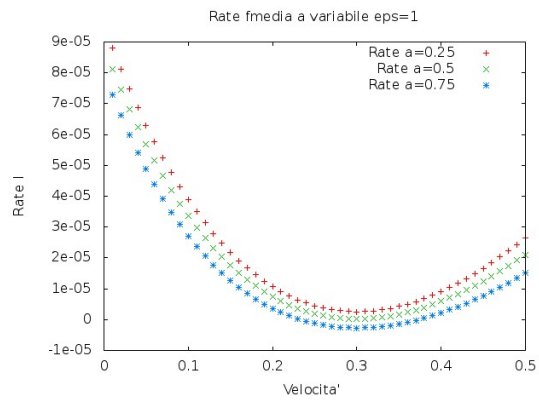
(a) Andamento generale per  $f_i$  con  $a$  variabile.



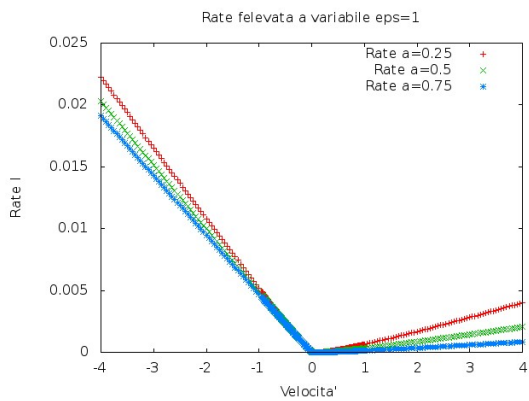
(b) Studio dei minimi per  $f_i$  con  $a$  variabile.



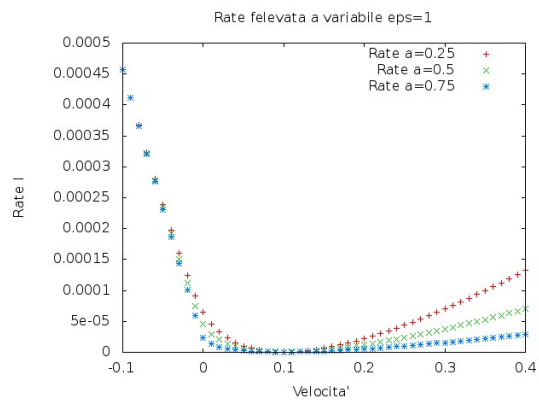
(c) Andamento generale per  $f_m$  con  $a$  variabile.



(d) Studio dei minimi per  $f_m$  con  $a$  variabile.



(e) Andamento generale per  $f_f$  con  $a$  variabile.



(f) Studio dei minimi per  $f_f$  con  $a$  variabile.

Figura 4.6: Rates al variare di  $f$  e di  $a$

#### 4.4 Analisi delle rate functions $I(v)$

Cominciamo analizzando il caso in cui si sono considerate rates functions associate a campi diversi (figura 4.4), dallo zoom di destra si verifica che le velocità medie corrispondono proprio ai minimi e zeri delle rate functions (corrispondenza suggellata dalla figura 4.1). Nell'andamento generale di sinistra è evidente invece il fatto che all'aumentare del campo la rate  $I(v)$  perde simmetria e tende ad inclinarsi sempre più verso destra, ciò sta a significare, secondo la 3.3, che all'aumentare del campo si registrano più fluttuazioni dalla media per velocità maggiori della velocità media piuttosto che per velocità minori.

Dal punto di vista dinamico questo testimonia che all'aumentare del campo per le particelle risulta più facile infiltrarsi in una trappola (la trappola inclinata infatti contribuisce positivamente alla velocità) e sempre più difficile uscirvi, quindi le particelle che entreranno nelle trappole ad  $f$  elevati vi passeranno molto più tempo all'interno che ad  $f$  bassi. Infatti per una rate simmetrica  $I(v)$ , visto che le fluttuazioni sono simmetriche si ha una stessa probabilità di entrare e di uscire da una trappola (è il caso del moto browniano).

Concentriamoci ora sui grafici di figura 4.5. Per campi poco intensi è evidente una dipendenza dal parametro  $\epsilon$  interessante, qualitativamente le rates hanno tutte lo stesso grado di asimmetria, ciò che varia abbastanza uniformemente sia per velocità maggiori della velocità media che per velocità minori è l'ampiezza. All'aumentare di  $\epsilon$  (per  $f_i$  costante) la rate si chiude sempre più, mostrando quindi che favorendo al contempo l'entrata e l'uscita dalle trappole le fluttuazioni dalla media diminuiscono. Le particelle così come entrano nelle trappole vi escono, invece per  $\epsilon$  bassi le fluttuazioni sono decisamente maggiori, indice che le particelle tenderanno a passare più tempo nelle trappole, tuttavia vista la forma abbastanza simmetrica della rate, le stesse particelle che vi entrano, poi escono. Cosa che invece non accade all'aumentare dell'intensità del campo, è vero che la tendenza all'aumentare di  $\epsilon$  è sempre quella di diminuire l'ampiezza della rate, tuttavia visto che  $f$  aumenta, al contempo si registra anche un aumento di asimmetria, una sempre maggior inclinazione verso fluttuazioni per velocità maggiori della velocità media. In tali casi infatti è vero che all'aumentare di  $\epsilon$  si favorisce contemporaneamente l'entrata e l'uscita dalle trappole, diminuendo in senso relativo le fluttuazioni, ma la presenza di un campo maggiore rende sempre più difficile uscire dalle trappole una volta entrati, portando, in senso assoluto al variare del campo  $f$ , un aumento delle fluttuazioni per velocità maggiori della velocità media. In sintesi, aumentando  $f$  e  $\epsilon$  si registrano meno fluttuazioni, ma quelle che si registrano sono dovute al fatto che è più facile che una particella entri in una trappola piuttosto che vi esca, molte particelle rimarranno quindi intrappolate per lunghi tempi senza possibilità di uscire.

Dall'analisi dei grafici in 4.6, per campo poco intenso  $f_i$  non si registra una dipendenza significativa dal parametro  $a$ , le rates sono tutte sovrapposte, e abbastanza simmetriche, ma non completamente, altrimenti vorrebbe dire che il campo è nullo e siamo nel caso browniano. L'aumentare o il diminuire del parametro  $a$ , non influenza le fluttuazioni per campi poco intensi. Tuttavia all'aumentare di  $f$  si registrano dipendenze sempre più significative dal parametro  $a$ , laddove si hanno  $f$  più intensi, l'aumentare di  $a$  tende ad allargare la rate e ad inclinarla maggiormente verso velocità positive. Ciò indica una maggior propensione delle particelle ad entrare nelle trappole piuttosto che uscirne, ed un maggior tempo di permanenza all'interno.

## Capitolo 5

# Conclusioni

Nell'elaborato esposto abbiamo cercato di mostrare come la teoria delle grandi deviazioni abbia una vasta applicazione nella ricerca del non-equilibrio, estendendo quindi il suo significato di mera formalizzazione se applicata allo studio della meccanica statistica di equilibrio.

I processi stocastici, in particolare quelli di Markov incontrano gli strumenti importanti delle grandi deviazioni, quali il principio delle grandi deviazioni, il teorema di Gärtner-Ellis e il teorema di Cramer al fine di sviluppare un metodo da seguire nello studio di evoluzione di sistemi fisici fuori equilibrio. In particolare si è cercato di mostrare quanto tutto il succo dell'informazione sia raccolto in una semplice funzione, la rate function. Tale funzione è in grado di spiegare come, per un processo stocastico che evolva nel tempo, nello studio di una determinata osservabile, si possano avere delle fluttuazioni dalla sua media, fornendo un utile mezzo di interpretazione anche per la dinamica che segue il sistema. Nell'equilibrio invece la rate function ricopriva un ruolo analogo a quello dell'entropia o dell'energia libera in termodinamica.

Si è cercato di applicare tutto ciò che abbiamo imparato dallo studio di questa teoria per il non-equilibrio, ad un sistema di conduzione di particelle all'interno di un canale con delle trappole. L'interesse per lo studio di tale sistema consiste nell'emergenza di una proprietà interessante, chiamata mobilità differenziale negativa. Ovvero, all'aumentare del campo, superato un certo valore massimo, la velocità media delle particelle che corrono nel canale, invece di aumentare, diminuisce. Lo scopo è stato quindi quello di valutare, tramite lo studio delle rate functions, se ci fossero delle variazioni nelle fluttuazioni dalla media della velocità delle particelle per campi differenti. Così è stato, le rate functions si sono mostrate, al variare di diversi parametri, piuttosto differenti, permettendoci di dare possibili interpretazioni alla dinamica interna al canale.

Concludiamo quindi rimarcando il fatto che la teoria delle grandi deviazioni si mostra essere un ottimo strumento di lettura e analisi del non-equilibrio.



# Bibliografia

- [1] P. Brémaud, *Markov chains*, 9, United States of America, Springer 1999.
- [2] A. Einstein, *The theory of the Opalescence of Homogeneous Fluids and Liquid Mixtures near the Critical State*, in *The collected papers of Albert Einstein*, vol. 3 *The Swiss years, writings 1909-1911*: 231-249, 1994.
- [3] K. Huang, *Meccanica statistica*, Bologna, Zanichelli Editore 1997.
- [4] G. Jona-Lasinio, *Large deviations and the Boltzmann entropy formula*, *Braz. J. Probab. Stat.*, vol. 29(2): 494-501, 2015.
- [5] N.G. van Kampen, *Stochastic processes in physics and chemistry*, prima edizione, Netherlands, North-Holland 1981.
- [6] M. Le Bellac, *Non equilibrium statistical mechanics in Lectures given at Les Houches Predoctoral School*, Istitut Non Linéaire de Nice, 2007, 1-65.
- [7] K. Mallick, *Some recent developments in non-equilibrium statistical physics*, *Pramana - J. Phys.*, vol. 73(3): 417-451, 2009.
- [8] H. Touchette, *The large deviation approach to statistical mechanics*, *Phys.Rep.*, vol. 478(1-3): 1-69, 2009.
- [9] H. Touchette e R.J. Harris, *Large deviation approach to nonequilibrium systems*, Preprint arxiv.org, 1-24, 2012.