## Università degli Studi di Padova



Dipartimento di Fisica e Astronomia "Galileo Galilei"

Corso di Laurea Magistrale in Fisica

# APPUNTI DEL CORSO DI "FISICA DEI SISTEMI COMPLESSI"

DI

LEONARDO PACCIANI MORI

Questo materiale è rilasciato sotto la licenza Creative Commons Attribuzione - Condividi allo stesso modo 4.0 Internazionale



Ciò significa che questo materiale può essere liberamente modificato e ridistribuito, a patto di citare la fonte e rilasciarlo sempre sotto questa licenza.

# Su questo documento

Questo documento è la trascrizione dei miei appunti del corso di *Fisica dei Sistemi Complessi*, tenuto dal prof. Attilio Stella per il secondo anno del Corso di Laurea Magistrale in Fisica presso il dipartimento di Fisica e Astronomia "Galileo Galilei" dell'Università degli Studi di Padova nel primo semestre dell'Anno Accademico 2016-2017.

Ho deciso di rilasciare questi appunti insieme al loro codice sorgente LATEX, di modo che chiunque possa eventualmente modificarli a seconda delle proprie esigenze. Tutto il materiale (compreso quello relativo anche ad altri corsi) si trova all'indirizzo web leonardo.pm/teaching.

La licenza sotto la quale questo documento è rilasciato è la Creative Commons *Attribuzione - Condividi allo stesso modo 4.0 Internazionale*. In sintesi, ciò significa che questo documento (incluso il suo codice sorgente) può essere modificato e ridistribuito liberamente, a condizione che sia sempre citata la fonte del documento originale e che sia rilasciato sempre sotto questa licenza.

Non è escluso che ci possano essere degli errori, qua e là, anche se ho cercato di essere il più meticoloso possibile nel scovarli. In caso ne siano sopravvissuti, mi scuso anticipatamente.

Padova, Gennaio 2017 Leonardo Pacciani Mori

# **Indice**

In	trodu	zione: cos'è la complessità?	3
1	Il m	oto browniano	8
	1.1	Proprietà generali del moto browniano	. 9
	1.2	Il moto browniano come paradigma	
		1.2.1 Teoria della risposta	
		1.2.2 Processi stocastici	
	1.3	L'equazione di diffusione	
	1.4	Processi stocastici	
		1.4.1 Processi stocastici markoviani continui	
		1.4.2 Il processo di Ornstein-Uhlenbeck	
2	Lar	percolazione	25
	2.1	Introduzione: esempi e applicazioni	
	2.2	Formulazione generale della percolazione	
	2.3	Percolazione di sito	
	2.0	2.3.1 Percolazione di sito in una dimensione	
		2.3.2 Percolazione di sito su reticolo di Bethe	
	2.4	Percolazione e geometria frattale	
	2.5	Percolazione e leggi di scala	
	2.6	Percolazione e modello di Potts	
3	Med	ccanica statistica di non equilibrio	46
J	3.1	Introduzione	
	5.1	3.1.1 La legge bilancio dettagliato	
	3.2	Reversibilità microscopica e irreversibilità macroscopica	
	0.2	3.2.1 Entropia e probabilità condizionata	
		3.2.2 L'ipotesi di regressione di Onsager	
		3.2.3 La matrice dei coefficienti cinetici	. 51
	3.3	Applicazioni	
4	Too	ria della risposta lineare	56
*	4.1	Introduzione	
	4.1	4.1.1 L'oscillatore armonico smorzato	
		4.1.1 L'oscinatore armonico sinorzato	
	4.2	Il teorema di fluttuazione-dissipazione	
	4.4	4.2.1 Il teorema fluttuazione-risposta	
		•	
		4.2.2 Funzioni di risposta e correlatori di equilibrio	
		4 / 5 H regretina di filiffilazione-dissipazione	กว

5	Pro	cessi stocastici	67
	5.1	Processi stazionari	67
		5.1.1 Teorema di Wiener-Khinchin	67
	5.2	Processi markoviani	69
		5.2.1 Equazione di Chapman-Kolmogoroff	<b>7</b> 0
		5.2.2 Esempi e considerazioni	<b>7</b> 0
	5.3		<b>7</b> 3
		5.3.1 Modello di Caldeira-Leggett	<b>7</b> 3
	5.4	L'equazione master	77
		5.4.1 Introduzione	77
		5.4.2 L'equazione master	<b>7</b> 8
		5.4.3 L'equazione master per processi discreti	<b>7</b> 9
6	Terr	nodinamica di sistemi microscopici	82
	6.1	Introduzione	82
		6.1.1 Argomento di plausibilità per $W_{\text{cycle}} < 0$	84
	6.2		86
		6.2.1 Uguaglianza di Jarzynski	87
			88
	6.3		91
7	Teo	ria delle grandi deviazioni	93
	7.1		95
	7.2		97
	7.3		98
		-	01
			02
		00 0	04
			11
			 13

# Introduzione: che cos'è la complessità?

Cosa intendiamo quando in Fisica si parla di "sistema complesso"? Più in generale: cos'è la "complessità" in Fisica?

In termini semplici ed essenziali, un *sistema complesso* è un sistema fisico il cui "comportamento" non è direttamente riconducibile alle proprietà dell'interazione fra le particelle microscopiche che lo costituiscono: in esso si manifestano *effetti cooperativi* che coinvolgono l'insieme di *tutte* le particelle. Un classico esempio di questo tipo di fenomeni lo si riscontra in vicinanza dei punti critici di un sistema termodinamico: in questo caso come noto le correlazioni fra le particelle che compongono un dato sistema si estendono su tutto il sistema stesso<sup>1</sup> (la lunghezza di correlazione diverge), e non possiamo più usare il linguaggio "locale" delle teorie di campo medio per descrivere sistemi del genere ma dobbiamo affrontare il problema in modo "globale"; in particolare nel caso di fenomeni critici questo approccio è dato dalla procedura di coarse graining del Gruppo di Rinormalizzazione.

In realtà come vedremo la complessità è un concetto molto più ampio, che racchiude molto di più dei "semplici" fenomeni critici.

Lo studio di sistemi fisici di questo tipo ha introdotto una "linea di pensiero" nuova rispetto a quella che è stata dominante per molti decenni (e che in alcuni settori continua ad esserlo); vediamo di cosa si tratta.

Sappiamo che in Fisica tradizionalmente si va alla ricerca di leggi fondamentali della Natura, ossia leggi che valgono a tutte le scale, e molti fisici ritenevano (e alcuni tutt'ora ritengono) che la Fisica fondamentale debba limitarsi all'individuazione di nuove leggi di questo tipo; siccome le leggi conosciute sono valide su tutti i range di scale che possiamo testare, cercare nuove leggi fondamentali significa cercare di spingerci a scale molto piccole o grandi, oltre i limiti di ciò che già conosciamo. Sembrerebbe quindi che il ruolo dei fisici fondamentali spetti solo ai particellari e ai cosmologi.

Questa "linea di pensiero" però non tiene conto del fatto che per capire la fisica dei sistemi che ci circondano (sitemi macroscopici ordinari, sistemi biologici e così via) le leggi fondamentali non sono sufficienti: *il riduzionismo non implica il costruttivismo*! In altre parole, dalla conoscenza delle leggi con cui interagiscono i componenti elementari della materia non possiamo ricostruire tutto quello che osserviamo (parafrasando una frase estratta da [6], non possiamo capire come funziona un bulldozer partendo dallo studio dei quarks che lo costituiscono!).

Un chiaro esempio del fatto che le leggi fondamentali della natura sono insufficienti, da sole, a spiegare ciò che ci circonda è il fenomeno della rottura spontanea di simmetria: si

<sup>&</sup>lt;sup>1</sup>Decadono con la distanza come una legge di potenza, e non come un'esponenziale decrescente come avviene in condizioni non critiche.

tratta di un fenomeno (pensiamo a sistemi come il modello di Ising) in completo contrasto con le leggi fondamentali che determinano la dinamica microscopica, ed *emerge* dal fatto che il sistema è composto da un enorme numero di costituenti elementari (idealmente un numero infinito). Possiamo quindi già capire come la Fisica della Complessità permetta di trovare strade per spiegare fenomeni che non sono in contraddizione con le leggi fondamentali della natura, ma che per essere spiegati richiedono qualcos'altro.

Un'altra caratteristica importante dei sistemi complessi che stiamo considerando è la loro *universalità*, ossia il fatto che la loro descrizione può prescindere dal particolare "substrato" considerato; come esempio possiamo considerare la teoria BCS della superconduttività, le cui leggi fondamentali sottostanti erano già note una trentina d'anni prima che venisse formulata. Il meccanismo fondamentale di questa teoria è il seguente: gli elettroni di un metallo possono essere considerati come un mare di Fermi; vicino alla superficie della sfera di Fermi di questo sistema si trova che c'è un'interazione attrattiva fra elettroni di spin opposto mediata dai fononi. Questo porta alla formazione delle cosiddette "coppie di Cooper", le quali sono bosoni e in quanto tali in determinate condizioni (bassa temperatura, ad esempio) condensano: è la condensazione delle coppie di Cooper che rende il metallo superconduttivo.

Ciò che si è visto con gli anni, però, è che questo stesso meccanismo può essere applicato anche a sistemi completamente diversi, come l'<sup>3</sup>He o le stelle di neutroni.

Il fatto che la spiegazione dei fenomeni complessi prescinda dal particolare substrato consierato permette spesso di usare modelli molto semplificati che non ambiscono a partire da princìpi primi ma che comunque permettono di descrivere benissimo il sistema alla scala alla quale lo si osserva.

Altro notevolissimo esempio di sistema complesso (che però non vedremo nel corso) è quello della vortisocità e turbolenza nei fluidi. Tutta la fenomenologia osservata in fluidodinamica è in linea di principio ottenibile dall'equazione fondamentale della fluidodinamica, l'equazione di Navier-Stokes; si tratta un'equazione molto difficile da risolvere, e quindi in generale un approccio di forza bruta non funziona (si possono determinare soluzioni esatte solo in casi particolari, e spesso le soluzioni trovate sono molto instabili). La turbolenza può però essere studiata attraverso gli *automi cellulari*: questi sono dei "toy models" in cui si semplifica moltissimo il problema della descrizione del fluido, ma nonostante questo si riesce (e lo si verifica a posteriori) che la fenomenologia del sistema è effettivamente riprodotta.

Ad esempio, in [6] Goldenfeld e Kadanoff mostrano un semplice modello col quale si possono riprodurre le caratteristiche di un fluido in una simulazione. Mettendoci in d=2 dimensioni per semplicità, supponiamo che le particelle si muovano su un reticolo triangolare; l'automa cellulare viene allora definito dando le regole secondo le quali le particelle possono muoversi sul reticolo (le velocità di una particella possono essere orientate secondo una delle possibili direzioni date dal reticolo, e sono uguali fra loro in modulo):

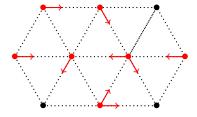


Figura 1: Automa cellulare considerato

I requisiti che richiediamo a questo automa cellulare sono la località (le particelle interagiscono e quindi influenzano la velocità delle altre solo quando vengono a contatto fra loro), la conservazione della massa e del momento, e l'isotropia del fluido. A questo punto le "regole" della dinamica vengono definite nel seguente modo:

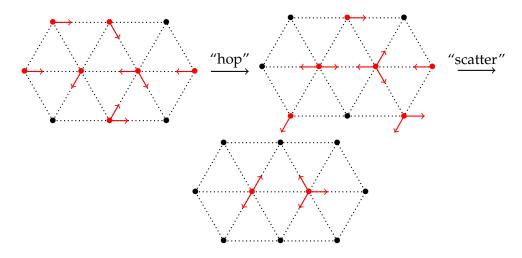


Figura 2: Evoluzione dell'automa cellulare

ove con l'"hop" le particelle si spostano verso il sito su cui punta la loro velocità, mentre nello "scatter" si ha che tutte le volte che in un sito si ha un bilancio totale nullo di momento, le velocità vengono ruotate di  $60^{\circ}$  in senso antiorario. Eseguendo simulazioni di sistemi di questo tipo per tempi lunghi e reticoli grandi risulta proprio che la dinamica riesce a riprodurre quella complessa osservata in fluidi reali.

In generale questi comportamenti complessi sono tipicamente manifestati in sistemi dove non si è all'equilibrio. Lo stato di non equilibrio è qualcosa che favorisce grandemente la manifestazione della complessità, e come vedremo un elemento importantissimo di questi fenomeni è la presenza di fluttuazioni nel sistema (perché molto di quello che potremo studiare di questi fenomeni riguarda proprio le fluttuazioni).

Fin'ora nel nostro percorso formativo non abbiamo mai visto nulla che riguardi la fisica statistica di non equilibrio, ma le attuali frontiere della complessità coinvolgono proprio situazioni fortemente fuori dall'equilibrio. Quello che vogliamo fare in questo corso è anche costruire una sorta di "ponte" fra la nostra attuale "cultura" di equilibrio e una di non equilibrio, anche molto spinto; in particolare vogliamo imparare a descrivere il non equilibrio e capire quali sono le sue leggi fondamentali.

Lo studio della fisica di sistemi non all'equilibrio è nato negli anni '30 ad opera di On-

sager, che ha cercato di capire cosa accade se si sposta *leggermente* un sistema da una situazione di equilibrio.

Riconsideriamo ora le fluttuazioni. Sappiamo che normalmente in un sistema termodinamico le fluttuazioni relative sono di ordine  $1/\sqrt{N}$ , e quindi non facilmente osservabili in sistemi macroscopici: solamente nei pressi di un punto critico le fluttuazioni diventano rilevanti, e questo lo si può vedere ad esempio nella divergenza della *suscettività* di un sistema (pensiamo ad esempio a un modello di Ising).

Analizziamo questa situazione dal punto di vista della teoria di *finite size scaling*. Consideriamo quindi un modello di Ising "standard" in un box finito di lato L in d dimensioni; in assenza di campo esterno l'hamiltoniana del sistema è, come noto:

$$\mathcal{H} = K \sum_{\langle ij \rangle} S_i S_j$$

Secondo la teoria del finite size scaling se f è l'energia libera per spin del sistema, allora effettuando un coarse graining riscalando il sistema di un fattore  $\ell$  si avrà:

$$f(K - K_c, h, L) \simeq \frac{1}{\ell^d} f\left(\ell^{Y_T}(K - K_c), \ell^{Y_H} h, \frac{L}{\ell}\right)$$

Derivando due volte rispetto ad h e ponendo h=0, poiché per definizione  $\partial^2 f/\partial h^2=\chi$  ove  $\chi$  è la suscettività del sistema, allora:

$$\chi(K - K_c, 0, L) \sim \ell^{2Y_H - d} \chi\left(K - K_c, 0, \frac{L}{\ell}\right)$$

Ponendoci al punto critico ( $K=K_c$ ) e ponendo  $\ell=L$  (è un numero arbitrario) si ha quindi:

$$\chi(0,0,L) \sim L^{2Y_H-d}\chi(0,0,1)$$

Ora, per definizione di suscettività si ha:

$$\chi(0,0,L) = \frac{1}{L^d} \left( \langle M^2 \rangle - \langle M \rangle^2 \right)$$

ove M è la magnetizzazione del sistema:

$$M = \sum_{i=1}^{L^d} S_i$$

Sostituendo, e tenendo conto che  $\chi(0,0,1) = \cos t$ .:

$$\frac{1}{L^d} \left( \left\langle M^2 \right\rangle - \left\langle M \right\rangle^2 \right) \sim L^{2Y_H - d} \qquad \Rightarrow \qquad \left\langle M^2 \right\rangle - \left\langle M \right\rangle^2 \sim L^{2Y_H}$$

Come noto, Onsager ha determinato che in d = 2 si ha  $Y_H = 7/4$  e quindi:

$$\left(\left\langle M^2\right\rangle - \left\langle M\right\rangle^2\right)_{d-2} \sim L^{7/2}$$

Ora però, nel sistema bidimensionale ci sono in tutto  $L^2$  spin; se le fluttuazioni del nostro sistema fossero "normali", ossia se potessimo considerare il nostro sistema come composto da spin indipendenti fra loro (o comunque con una eventuale piccola correlazione a

range molto corto), ci aspetteremmo che  $\langle M^2 \rangle - \langle M \rangle^2$  cresca come  $L^2$  per via del teorema del limite centrale². Al punto critico però questo non accade, e le fluttuazioni crescono molto più velocemente: in corrispondenza di esso si ha una forte deviazione dal teorema del limite centrale.

Il punto critico, però, è un punto molto "particolare" del sistema, bisogna "andare a cercarlo" impostando la temperatura (o più in generale di un dato parametro) a un preciso valore.

Einstein fu il primo a capire il ruolo fondamentale delle fluttuazioni in un sistema diverso da quello di un punto critico: i suoi studi sul *moto browniano* mostrarono come in fenomeni di questo tipo si può avere una nozione *diretta* delle fluttuazioni, oltre che a mostrare sperimentalmente per la prima volta la natura *discreta* della materia (la comunità scientifica di allora riteneva infatti che la materia fosse un continuo, rifiutando l'ipotesi atomistica).

 $<sup>^2</sup>$ Il sistema è infatti composto da  $L^2$  spin e per il teorema del limite centrale  $\sigma_M^2 \coloneqq \left\langle M^2 \right\rangle - \left\langle M \right\rangle^2$  cresce come il numero di componenti del sistema, ossia proprio  $L^2$ .

# Capitolo 1

## Il moto browniano

Il termine *moto browniano* prende il nome dal botanico britannico Robert Brown, che agli inizi dell'Ottocento osservò al microscopio il moto di alcuni granuli di polline sospesi in un fluido, ad esempio acqua. In generale una qualunque particella mesoscopica<sup>1</sup> immersa in un fluido risulta muoversi in modo molto caotico, cambiando continuamente direzione e senza risentire visibilmente della forza peso.

Inizialmente fenomeni del genere vennero addirittura attribuiti a una sorta di *vis vitalis* delle particelle in esame (ossia si riteneva che in qualche modo fossero "vive"), teoria poi scartata una volta verificato che lo stesso moto caratterizza anche particelle che sicuramente vive non sono (gli esperimenti vennero ripetuti ad esempio con polvere molto fine di materiale inorganico). È solo grazie ad Einstein che nel 1905 si riuscì a dare una spiegazione teorica al fenomeno: in particolare quest'ultimo, assumendo che la materia fosse costituita da particelle (e non fosse invece un continuo,

Figura 1.1: Moto browniano di microsfere sospese in acqua

come abbiamo detto ritenesse la comunità scientifica dell'epoca²), riuscì a spiegare il moto browniano come conseguenza dei continui urti a cui i granuli (che d'ora in poi chiameremo semplicemente particelle) andavano incontro con le molecole che costituiscono il fluido. Difatti, dato che queste particelle immerse sono mesoscopiche urteranno generalmente con un numero modesto di molecole del fluido, e quindi è perlomeno concepibile che si possano creare degli "squilibri" temporanei di momento su scale temporali piccole (se ad esempio alcune molecole urtano su un lato del granulo, ma un numero insufficiente "controbilancia" questi urti dalla parte opposta). Possiamo quindi pensare che la nostra particella sia soggetta a una "forza fluttuante" che modella l'azione dei continui urti con le molecole. Se chiamiamo  $\vec{f}(t)$  questa forza, l'equazione del moto della particella mesoscopica sarà:

$$\frac{d\vec{v}}{dt} = -\frac{\lambda}{M}\vec{v} + \frac{\vec{f}(t)}{M} \tag{1.1}$$

ove M è la massa della particella, mentre  $\lambda$  è la costante di proporzionalità fra la velocità

<sup>&</sup>lt;sup>1</sup>I granuli di polline sono adatti proprio perché molto più grandi delle molecole d'acqua ma comunque molto piccoli rispetto a un sistema propriamente macroscopico.

<sup>&</sup>lt;sup>2</sup>Il primo modello atomico di Thomson fu proposto appena tre anni prima.

e la forza di attrito viscoso<sup>3</sup> (dalla legge di Stokes, ad esempio, per una sfera di raggio R si ha  $\lambda = 6\pi R\eta$ , con  $\eta$  viscosità del fluido).

Cerchiamo di capire per ora il senso fisico della (1.1), riservandoci di tornare sul suo senso matematico più avanti.

Per ora, poi, non consideriamo l'azione della forza peso perché, come detto prima, sperimentalmente le particelle non ne risentono; vedremo poi più avanti cosa accade se includiamo forze esterne, come appunto la forza peso.

#### 1.1 Proprietà generali del moto browniano

Vediamo alcune proprietà fondamentali dei sistemi che vogliamo studiare.

Innanzitutto, se supponiamo che la particella abbia una velocità iniziale  $\vec{v}_0 \neq \vec{0}$ , ciò che ci aspettiamo è che questa si attenui con un tempo caratteristico  $\tau_{\rm rel} = M/\lambda$  (difatti, se  $\vec{f} = 0$  la soluzione della (1.1) è un semplice esponenziale decrescente), e che  $\langle \vec{v} \rangle$  tenda a zero. Tuttavia, dopo un tempo sufficientemente lungo la particella continua comunque a muoversi (e vedremo infatti che  $\langle \vec{v}^2 \rangle$  non è nullo), e per tempi lunghi è più che ragionevole aspettarsi che questa si equilibri col fluido.

Vediamo poi quali proprietà deve avere la  $\vec{f}(t)$ , visto che per ora è completamente indeterminata. Alcune di queste proprietà sono intuitive ed immediate, ma per le altre una spiegazione esaustiva sarà data solo in 1.4.

L'idea fondamentale è che le sue componenti  $f_i(t)$  siano delle variabili stocastiche dipendenti dal tempo (concetto sul quale comunque torneremo più in dettaglio in seguito); pertanto anche la velocità  $\vec{v}$  della particella, in virtù della (1.1), sarà una variabile aleatoria. Dobbiamo quindi considerare la statistica delle possibili evoluzioni della velocità, in un ensemble opportuno: infatti, per ogni possibile evoluzione della  $\vec{f}(t)$  il sistema avrà una "storia" diversa, e quello che dovremo fare per studiare un sistema di questo tipo sarà andare a considerare l'ensemble di tutte le possibili "storie" della particella (e usarlo per effettuare valori medi ecc.).

Poi, dato che la  $\vec{f}(t)$  descrive le "deviazioni" del moto della particella rispetto al suo "comportamento medio", sicuramente bisognerà che  $\langle f_i(t) \rangle = 0 \quad \forall i, \ \forall t$ . Se però fissiamo un istante t, com'è distribuita  $f_i(t)$ ? In altre parole: qual è la densità di probabilità della  $f_i(t)$  a t fissato?

Quello che facciamo è supporre che la distribuzione di questa variabile aleatoria sia gaussiana; questo perché pensiamo che il numero di molecole che urtano con la particella è comunque grande (anche se "non troppo"), di modo tale che per il teorema del limite centrale la risultante della forza esercitata sulla particella per via degli urti è distribuita normalmente.

L'ultima proprietà della  $\vec{f}(t)$  che dobbiamo analizzare è la correlazione dei suoi valori; ci possiamo infatti chiedere se i valori di questa funzione possano essere in qualche modo correlati fra loro. Quello che assumiamo è che si abbia:

$$\langle f_i(t)f_j(t')\rangle = \delta_{ij}\Lambda\Delta(|t-t'|)$$

Ciò significa che le varie componenti della  $\vec{f}$  sono scorrelate fra loro (per via della presenza della  $\delta_{ij}$ ), e che l'andamento della correlazione nel tempo è regolato dalla funzione  $\Delta$  (ed è ovviamente pari nella differenza dei tempi per via della presenza del modulo);

<sup>&</sup>lt;sup>3</sup>Se la particella ha una velocità iniziale non nulla, la forza che si sviluppa dall'urto con le molecole del fluido è proprio quella responsabile dell'attrito viscoso.

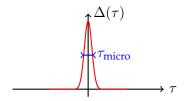


Figura 1.2: Andamento della funzione  $\Delta(\tau)$ 

poiché le scale delle molecole del fluido e della particella browniana sono molto diverse fra loro<sup>4</sup>, ci aspettiamo che la funzione  $\Delta$  sia sensibilmente diversa da zero solo su scale temporali piccole agli effetti macroscopici, o in altre parole che le correlazioni ci siano solo su tempi brevi. Si tratta quindi di una funzione molto piccata su un intervallo temporale, che chiamiamo  $\tau_{\rm micro}$ , molto minore delle scale temporali macroscopiche, rappresentate ad esempio da  $\tau_{\rm rel} = M/\lambda$ . Si ha insomma una situazione del tipo illustrato in figura 1.2. Richiedendo poi che la  $\Delta$  sia normalizzata, ossia che:

$$\int_{-\infty}^{+\infty} \Delta(\tau) d\tau = 1$$

allora agli effetti pratici possiamo considerarla una  $\delta$  di Dirac. Infine, la costante  $\Lambda$  quantifica l'"intensità" di questa correlazione.

Ora, la soluzione formale della (1.1) è:

$$v_i(t) = e^{-\frac{\lambda}{M}t} \left[ \int_0^t dt' e^{\frac{\lambda}{M}t'} \frac{f_i(t')}{M} + v_{0i} \right]$$
 (1.2)

ove  $\vec{v}_0$  è la velocità iniziale. Da questa possiamo ricavare un po' di informazioni utili sul nostro sistema, come ad esempio l'evoluzione nel tempo del valor medio di  $v_i$ :

$$\langle v_i(t) \rangle = e^{-\frac{\lambda}{M}t} \left[ \int_0^t dt' e^{\frac{\lambda}{M}t'} \frac{\langle f_i(t') \rangle}{M} + v_{0i} \right] = e^{-\frac{\lambda}{M}t} v_{0i} \stackrel{t \to \infty}{\longrightarrow} 0$$

(ove abbiamo sfruttato le proprietà della  $\vec{f}$ , in particolare il fatto che  $\langle f_i(t) \rangle = 0$ ). Pertanto, come avevamo intuito poco fa, se la particella ha una velocità iniziale non nulla il valor medio della sua velocità tende a zero in un tempo dell'ordine di  $\tau_{\rm rel} = M/\lambda$ . Ciò però non vuol dire che la particella è ferma tout court: la forza stocastica  $\vec{f}$  continua infatti ad agire, e la particella continua a muoversi (anche se i suoi spostamenti hanno media nulla). Possiamo quindi ricavare ulteriori informazioni sul moto della particella calcolando  $\langle v_i^2(t) \rangle$ :

$$\langle v_{i}^{2}(t) \rangle = e^{-2\frac{\lambda}{M}t} \left[ \int_{0}^{t} dt' \int_{0}^{t'} dt'' e^{\frac{\lambda}{M}(t'+t'')} \frac{\langle f_{i}(t')f_{i}(t'') \rangle}{M^{2}} + v_{0i}^{2} \right] =$$

$$= e^{-2\frac{\lambda}{M}t} \left[ \int_{0}^{t} dt' e^{2\frac{\lambda}{M}t'} \frac{\Lambda}{M^{2}} + v_{0i}^{2} \right] = e^{-2\frac{\lambda}{M}t} \left[ v_{0i}^{2} + \frac{\Lambda}{2\lambda M} \left( e^{2\frac{\lambda}{M}t} - 1 \right) \right]$$
(1.3)

<sup>&</sup>lt;sup>4</sup>Questa differenza di scale è una caratteristica tipica dei fenomeni di non equilibrio. In situazioni del genere c'è infatti sempre una discrepanza fra la scala microscopica e quella alla quale si osserva il sistema.

ove nel primo passaggio abbiamo subito tenuto conto del fatto che nel calcolare  $v_i(t')v_i(t'')$  ci sono alcuni termini lineari in  $f_i(t)$  che hanno media nulla (e pertanto sopravvivono solo i termini quadratici); nel seguito si è poi tenuto conto che, per le proprietà della  $\vec{f}$  di cui abbiamo parlato prima, si ha  $\langle f_i(t')f_i(t'')\rangle = \Lambda\delta(t-t')\delta_{ii} = \Lambda\delta(t-t')$ . Proseguendo:

$$\left\langle v_i^2(t)\right\rangle = e^{-2\frac{\lambda}{M}t}\left(v_{0i}^2 - \frac{\Lambda}{2\lambda M}\right) + \frac{\Lambda}{2\lambda M} \overset{t\to\infty}{\longrightarrow} \frac{\Lambda}{2\lambda M} \coloneqq \left\langle v_i^2(\infty)\right\rangle$$

e quindi il modulo della velocità della particella non si annulla a tempi lunghi.

Si ha poi che correlazioni del tipo  $\langle v_i(t)v_i(0)\rangle$  si annullano per tempi lunghi come un esponenziale della forma  $e^{-t/\tau_{\rm rel.}}$ : possiamo pertanto dire che dopo un tempo  $\tau_{\rm rel}$  la particella "perde memoria" della sua velocità iniziale, e a tempi  $t>\tau_{\rm rel}$  possiamo pensare che lo spostamento della particella sia indipendente dalla storia precedente della particella stessa. È questo ciò che ci "autorizza" a studiare il moto browniano in termini di cammini aleatori.

Ora, la nostra particella è immersa in un fluido che fa anche da bagno termico per il sistema; dopo un tempo sufficientemente lungo a partire dall'istante iniziale, quindi, la nostra particella sarà in equilibrio col bagno termico che la circonda. Supponendo allora che per la particella valga il teorema di equipartizione dell'energia, si avrà:

$$\frac{M}{2} \left\langle v_i^2(\infty) \right\rangle = \frac{1}{2} k_B T \qquad \Rightarrow \qquad \frac{M}{2} \cdot \frac{\Lambda}{2\lambda M} = \frac{1}{2} k_B T$$

Abbiamo quindi trovato la cosiddetta relazione di Einstein:

$$\Lambda = 2\lambda k_B T$$

che permette di esprimere  $\Lambda$  in termini di grandezze misurabili.

Storicamente, però, il valore di  $k_B$  non era noto al tempo in cui fu studiato il moto browniano, o meglio non era ancora noto con precisione il numero di Avogadro<sup>5</sup>. Come vedremo adesso, però, è possibile determinarlo "sperimentalmente" dalla misura dei coefficienti di diffusione.

Per farlo è necessario determinare lo spostamento medio della particella. Si ha innanzitutto:

$$r_i(t) = \int_0^t dt' v_i(t') = \int_0^t dt' e^{-\frac{\lambda}{M}t'} \left[ \int_0^{t'} dt'' e^{\frac{\lambda}{M}t''} \frac{f_i(t'')}{M} + v_{0i} \right]$$

e quindi:

$$\langle r_i(t) \rangle = \int_0^t dt' e^{-\frac{\lambda}{M}t'} v_{0i} = v_{0i} \frac{M}{\lambda} \left( 1 - e^{-\frac{\lambda}{M}t} \right) \stackrel{t \to \infty}{\longrightarrow} v_{0i} \frac{M}{\lambda} = v_{0i} \tau_{\text{rel}}$$

Calcoliamo ora lo scarto quadratico medio dello spostamento, ossia  $\langle \Delta r_i^2(t) \rangle = \langle r_i^2(t) \rangle - \langle r_i(t) \rangle^2$ ; dobbiamo prima di tutto determinare:

$$\langle r_i^2(t) \rangle = \left\langle \int_0^t dt_1 e^{-\frac{\lambda}{M}t_1} \left[ \int_0^{t_1} dt' e^{\frac{\lambda}{M}t'} \frac{f_i(t')}{M} + v_{0i} \right] \cdot \int_0^t dt_2 e^{-\frac{\lambda}{M}t_2} \left[ \int_0^{t_2} dt'' e^{\frac{\lambda}{M}t''} \frac{f_i(t'')}{M} + v_{0i} \right] \right\rangle =$$

$$= \int_0^t dt_1 \int_0^t dt_2 e^{-\frac{\lambda}{M}(t_1 + t_2)} \left[ \frac{1}{M^2} \int_0^{t_1} dt' \int_0^{t_2} dt'' e^{\frac{\lambda}{M}(t' + t'')} \Lambda \delta(t' - t'') + v_{0i}^2 \right]$$

<sup>&</sup>lt;sup>5</sup>Ricordare che  $k_B = R/N_A$ , con R costante dei gas (facilmente misurabile) e  $N_A$  numero di Avogadro.

Notiamo che il termine proporzionale a  $v_{0i}^2$  non è altro che  $\langle r_i(t) \rangle^2$ :

$$\int_0^t dt_1 \int_0^t dt_2 e^{-\frac{\lambda}{M}(t_1 + t_2)} v_{0i}^2 = \left( \int_0^t dt_1 e^{-\frac{\lambda}{M}t_1} v_{0i} \right) \left( \int_0^t dt_2 e^{-\frac{\lambda}{M}t_2} v_{0i} \right) = \langle r_i(t) \rangle^2$$

Quindi:

$$\begin{split} \left< \Delta r_i^2(t) \right> &= \int_0^t dt_1 \int_0^t dt_2 e^{-\frac{\lambda}{M}(t_1 + t_2)} \left[ \frac{1}{M^2} \int_0^{t_1} dt' \int_0^{t_2} dt'' e^{\frac{\lambda}{M}(t' + t'')} \Lambda \delta(t' - t'') \right] = \\ &= \int_0^t dt_1 \int_0^t dt_2 e^{-\frac{\lambda}{M}(t_1 + t_2)} \frac{\Lambda}{M^2} \int_0^{\min(t_1, t_2)} dt' e^{2\frac{\lambda}{M}t'} \end{split}$$

ove si è usato il fatto che:

$$\int_0^{t_1} dt' \int_0^{t_2} dt'' e^{\frac{\lambda}{M}(t'+t'')} \delta(t'-t'') = \int_0^{\min(t_1,t_2)} dt' e^{2\frac{\lambda}{M}t'}$$

Ora,  $t_1$  e  $t_2$  variano da zero a t, e quindi possiamo dividere il dominio d'integrazione in due zone:  $t_1 > t_2$  e  $t_2 > t_1$ . Se  $t_2 < t_1$  possiamo limitare l'integrale in  $dt_2$  a  $t_1$ , e viceversa se  $t_1 < t_2$  limitiamo l'integrale in  $dt_1$  a  $t_2$ : il risultato è lo stesso perché i termini sono simmetrici sotto lo scambio  $t_1 \longleftrightarrow t_2$ ; poiché questi due termini sono uguali, alla fine basta moltiplicare uno dei due contributi per 2. Scegliendo  $t_2 < t_1$  si ha:

$$\langle \Delta r_i^2(t) \rangle = 2 \int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\frac{\lambda}{M}(t_1 + t_2)} \frac{\Lambda}{M^2} \int_0^{t_2} dt' e^{2\frac{\lambda}{M}t'}$$

e alla fine:

$$\left\langle \Delta r_i^2(t) \right\rangle = \frac{\Lambda}{\lambda^2} t + 2 \frac{\Lambda M}{\lambda^3} \left[ e^{-\frac{\lambda}{M}t} - \frac{1}{4} e^{-2\frac{\lambda}{M}t} - \frac{3}{4} \right]$$

Dalla relazione di Einstein, poi,  $\Lambda=2\lambda k_BT$ , e quindi per tempi lunghi si ha:

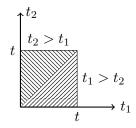
$$\left\langle \Delta r_i^2(t) \right\rangle \stackrel{t \to \infty}{\longrightarrow} 2 \frac{k_B}{\lambda} Tt \equiv 2Dt$$

ove D è il cosiddetto "coefficiente di diffusione",  $D = k_B T/\lambda$ .

Si tratta di una quantità misurabile (Perrin vinse il Nobel nel 1926 per essere riuscito a misurarlo per la prima volta), come T e  $\lambda$ :  $k_B$  può quindi essere determinato "sperimentalmente".

### 1.2 Il moto browniano come paradigma

Vediamo ora come lo studio del moto browniano ci permetta di "intravedere" problemi e tematiche che affronteremo più avanti.



<sup>&</sup>lt;sup>6</sup>Questo è un quadrato:

#### 1.2.1 Teoria della risposta

Abbiamo visto che  $\langle \Delta r_i^2(t) \rangle \sim 2Dt$  per  $t \to \infty$ ; allora si avrà  $\langle \Delta \vec{r}^{\,2}(t) \rangle \sim 6Dt$  sempre nello stesso limite. Possiamo allora pensare di calcolare D come:

$$D = \frac{1}{6} \lim_{t \to \infty} \frac{d}{dt} \left\langle \Delta \vec{r}^{\,2}(t) \right\rangle$$

Tuttavia,  $\langle \Delta \vec{r}^{\,2}(t) \rangle$  si può riscrivere ("per definizione") come:

$$\langle \Delta \vec{r}^{\,2}(t) \rangle = \int_0^t dt_1 \int_0^t dt_2 \langle \vec{v}(t_1) \cdot \vec{v}(t_2) \rangle$$

e quindi:

$$\frac{d}{dt} \left\langle \Delta \vec{r}^{2}(t) \right\rangle = 2 \int_{0}^{t} dt_{1} \left\langle \vec{v}(t_{1}) \cdot \vec{v}(t) \right\rangle$$

In conclusione:

$$D = \frac{1}{3} \lim_{t \to \infty} \int_0^t dt_1 \left\langle \vec{v}(t_1) \cdot \vec{v}(t) \right\rangle$$

Per via della presenza del limite per  $t \to \infty$ , l'integrale che definisce D è sempre più esteso nella "zona di equilibrio", ossia comprende un intervallo di tempo sempre più ampio dove il sistema è all'equilibrio; pertanto l'integrale è largamente dominato dal contributo che proviene dalla cosiddetta "correlazione di equilibrio". Abbiamo quindi espresso il coefficiente di diffusione in termini della correlazione all'equilibrio fra i valori della velocità della particella; la relazione che abbiamo appena trovato può poi essere collegata alle cosiddette "funzioni di risposta" di un sistema.

Per determinare la risposta del nostro sistema, supponiamo di esercitare su di esso una forza esterna  $\vec{F}_{\text{ext}}$ ; chiaramente, la particella si muoverà di conseguenza e il valor medio della velocità della particella cambierà:

$$\langle \vec{v}_{\infty} \rangle = \frac{1}{\lambda} \vec{F}_{\rm ext}$$

In altre parole, la particella "risponde" alla forza esterna acquisendo questa velocità limite. Sapendo però che  $D = k_B T/\lambda$ , possiamo scrivere:

$$\langle \vec{v}_{\infty} \rangle = \frac{D}{k_B T} \vec{F}_{\rm ext}$$

Il coefficiente che lega una forza ad una velocità, come in questo caso, prende il nome di *mobilità*, che in genere si indica con la lettera  $\sigma$ . Si ha pertanto che:

$$\sigma = \frac{D}{k_B T}$$

Come diremo anche più avanti, le "formule di Kubo" permettono di mettere in relazione fra loro funzioni di risposta, come la mobilità  $\sigma$ , con le correlazioni di equilibrio.

#### 1.2.2 Processi stocastici

Abbiamo parlato prima in termini molto generici della forza  $\vec{f}(t)$ . Vogliamo ora andare più in profondità: innanzitutto dobbiamo capire cos'è *davvero* la  $\vec{f}$ , e come usarla. Dopodiché dobbiamo anche capire come trattare da un punto di vista matematicamente

corretto l'equazione (1.1); come vedremo, infatti, analizzandola più a fondo sorgono problemi che noi abbiamo sorvolato con la disinvoltura tipica dei fisici<sup>7</sup>.

Come vedremo, la chiave di tutto starà nell'esprimere la (1.1) non come equazione differenziale, ma in forma *incrementale*; scopriremo poi che le proprietà che la  $\vec{f}$  deve avere per il moto browniano sono quelle che genericamente devono avere termini aleatori di questo tipo per *processi stocastici markoviani continui* (ove per "markoviano" si intende "privo di memoria", nel senso che abbiamo illustrato poco fa).

Cominciamo parlando di variabili aleatorie. Se X è una variabile aleatoria che può assumere valori  $x \in \mathbb{R}$ , questa sarà "qualificata" da una densità di probabilità  $P_X(x)$ , che come noto rappresenta la probabilità che X assuma valori fra x e x + dx. Ad esempio, per una variabile stocastica gaussiana si ha:

$$P_X(x) = \frac{e^{-x^2}}{\sqrt{2\pi}}$$

Le variabili aleatorie che vogliamo considerare, però, *evolvono nel tempo* (come appunto la  $\vec{f}(t)$  e la  $\vec{v}(t)$ ). Oggetti di questo tipo si chiamano *processi stocastici*: come possiamo trattarli? Innanzitutto, se X è una variabile aleatoria e Y = f(X) è una sua funzione, le loro densità di probabilità sono legate dalla relazione:

$$P_Y(y) = \int dx P_X(x) \delta(y - f(x))$$

Volendo, se  $Y = f(X_1, X_2)$  è funzione di due variabili aleatorie  $X_1$  e  $X_2$  che supponiamo siano indipendenti e distribuite identicamente, allora:

$$P_Y(y) = \int dx_1 dx_2 P_{X_1}(x_1) P_{X_2}(x_2) \delta(y - f(x_1, x_2))$$

In questi termini un processo stocastico può essere definito come una funzione di una variabile aleatoria e del tempo:  $Y_t = f(X,t)$ . In questo caso si avrà:

$$P_{Y_t}(y,t) = \int dx P_X(x) \delta(y - f(x,t))$$

Generalmente quando trattiamo le variabili e i processi stocastici "da fisici" non ci preoccupiamo di specificare cos'è la variabile *X*, anche se questa esiste *sempre*.

Fissato quindi  $t = t_1$ , il valor medio del processo  $Y_t$  a  $t_1$  sarà:

$$\langle Y_{t_1} \rangle = \int dx P_X(x) f(x, t_1)$$

e in generale:

$$\langle Y_{t_1} \cdots Y_{t_n} \rangle = \int dx P_X(x) f(x, t_1) \cdots f(x, t_n)$$

Essendo la f(x,t) funzione di una variabile aleatoria e del tempo, fissato  $x=\tilde{x}$  allora  $f(\tilde{x},t)$  rappresenterà una possibile evoluzione del sistema (detta anche "realizzazione"

<sup>&</sup>lt;sup>7</sup>In realtà quello che abbiamo fatto è legittimo, ma dobbiamo vedere meglio quali aspetti non abbiamo considerato.

del processo); il processo stocastico può dunque andare incontro a più possibili evoluzioni, ognuna "pesata" dal valore x della variabile stocastica X.

Ci possiamo anche chiedere se il modo che abbiamo appena visto per costruire processi stocastici sia l'unico possibile. La risposta è negativa, e contenuta in un importante risultato di Kolmogoroff, che illustriamo senza dimostrare.

Consideriamo un insieme di densità di probabilità congiunte per i valori di una data Y (il nostro processo) a tempi fissati:

$$P_n(y_1, t_1; y_2, t_2; \dots; y_n, t_n) = \int dx P_X(x) \delta(y_1 - f(x, t_1)) \cdots \delta(y_n - f(x, t_n))$$

Le  $P_n$  così definite sono simmetriche rispetto alla permutazione dei loro argomenti. Soddisfano inoltre un'altra proprietà:

$$\int dy_n P_n(y_1, t_1; \dots; y_n, t_n) = P_{n-1}(y_1, t_1; \dots; y_{n-1}, t_{n-1})$$

ossia formano una "gerarchia". La normalizzazione di tutte le  $P_n$  è quindi garantita se la  $P_1$  è normalizzata, ossia:

$$\int dy P_1(y,t) = 1$$

Kolmogoroff ha mostrato che se esistono delle  $P_n$  con queste proprietà allora la variabile X che "genera" il processo stocastico esiste (anche se il teorema non dà indicazioni su come determinarla).

Pertanto, un processo stocastico può anche essere determinato assegnando una "gerarchia", nel senso appena specificato, di densità di probabilità congiunte.

Vediamo ora un esempio rilevante di una variabile aleatoria X che entra in gioco nei processi stocastici.

In meccanica statistica abbiamo visto come trattare un sistema classico all'equilibrio: sappiamo che in questo caso  $(q,p)\in\Gamma$  è il punto rappresentativo del sistema; possiamo allora identificare la X con lo stato microscopico. Difatti, essendo normalizzata, la densità di probabilità  $\rho(x)$  (microcanonica, canonica ecc.) che introduciamo nello spazio  $\Gamma$  può essere considerata come una  $P_X(x)$  ove appunto la X è lo stato microscopico del sistema e x=(q,p) il particolare punto rappresentativo.

Sappiamo poi che in linea di principio il punto rappresentativo evolve secondo le equazioni di Hamilton, e quindi esiste una funzione f(x,t)=x(t)=(q(t),p(t)); se quindi vogliamo studiare la dinamica di una qualunque funzione che descrive una proprietà del sistema, anche all'equilibrio, quello che dobbiamo fare è studiare la dinamica di un processo stocastico Y(X,t).

Quest'ottica ha però un punto debole: in linea di principio è infatti molto "elegante" vedere (q,p) come la variabile aleatoria che genera i processi che stanno alla base delle grandezze osservabili in Fisica statistica, ma il problema è che per conoscere la f(x,t) che fa evolvere il sistema bisogna risolvere le equazioni di Hamilton, cosa che sappiamo non essere possibile in generale.

#### Il teorema del limite centrale

Abbiamo citato qualche volta il teorema del limite centrale; visto che ci servirà anche in seguito, illustriamolo.

Supponiamo di avere n variabili aleatorie  $X_1,\ldots,X_n$  indipendenti e identicamente distribuite, non necessariamente gaussiane, con varianza  $\sigma^2$  e media nulla (in generale se le  $X_i$  non hanno media nulla basta "traslarle"). Il teorema del limite centrale asserisce che allora per  $n\to\infty$  la variabile  $Y=X_1+X_2+\cdots+X_n$  tende a distribuirsi secondo una gaussiana con scarto quadratico medio pari a $^8$   $\sigma\sqrt{n}$ .

Da notare che l'unica informazione sulle distribuzioni delle  $X_i$  che sopravvive nel limite  $n \to \infty$  è la loro varianza; le  $X_i$  potrebbero infatti avere gli altri momenti non banalmente collegabili al secondo<sup>9</sup>: dei tanti parametri che potrebbero caratterizzare queste distribuzioni, però, l'unico rilevante è  $\sigma$  (è una forma molto strana di "universalità").

Vediamo quindi, qualitativamente, come verificare il teorema del limite centrale. Innanzitutto, si ha:

$$P_Y(y) = \int dx_1 \cdots dx_n P_{X_1}(x_1) \cdots P_{X_n}(x_n) \delta(y - (x_1 + \dots + x_n))$$
 (1.4)

Introduciamo poi la funzione caratteristica di una data densità di probabilità  $P_X(x)$ :

$$f(k) := \int dx e^{ikx} P_X(x)$$

che non è altro che la trasformata di Fourier della  $P_X$ . Sviluppando l'esponenziale:

$$f(k) = \int dx \left( P_X(x) + ikx P_X(x) - \frac{k^2}{2} x^2 P_X(x) + \cdots \right)$$

Integrando possiamo esplicitare i vari momenti della  $P_X$ :

$$f(k) = 1 + ik \langle x \rangle_{P_X} - \frac{k^2}{2} \langle x^2 \rangle_{P_X} + \cdots$$

La funzione caratteristica è quindi una *funzione generatrice* dei momenti della  $P_X$ . Nel caso specifico di una gaussiana, ad esempio, completando il quadrato ad esponente si ha:

$$f(k) = \int dx e^{ikx} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma}} = e^{-\frac{k^2}{2}\sigma^2}$$

ossia la funzione caratteristica di una gaussiana è un'altra gaussiana, ma con varianza  $1/\sigma^2$ .

Ritornando alla situazione che stavamo considerando, calcoliamo la funzione caratteristica della  $P_Y$ :

$$F_n(k) = \int dy e^{iky} \int dx_1 \cdots dx_n P_{X_1}(x_1) \cdots P_{X_n}(x_n) \delta \left( y - (x_1 + \dots + x_n) \right) =$$

$$= \int dx_1 \cdots dx_n e^{ik(x_1 + \dots + x_n)} P_{X_1}(x_1) \cdots P_{X_n}(x_n) = f(k)^n$$

Notare che questo significa che  $Y = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)$  è distribuita come una gaussiana di scarto quadratico medio pari a  $\sigma$ .

<sup>&</sup>lt;sup>9</sup>Ad esempio nel caso di una gaussiana i momenti dispari sono tutti nulli e quelli pari sono esprimibili in termini del secondo momento (e quindi banalmente collegabili con esso).

Questo ci suggerisce che se definiamo  $Y = \frac{1}{\sqrt{n}}(X_1 + \cdots X_n)$  la funzione caratteristica della sua densità di probabilità sarà:

$$F_n' = f\left(\frac{k}{\sqrt{n}}\right)^n = \left(1 - \frac{\sigma^2}{2} \cdot \frac{k^2}{n} + O\left(\frac{k^3}{n^{3/2}}\right)\right)^n \xrightarrow{n \to \infty} e^{-\frac{k^2}{2}\sigma^2}$$

che è proprio la funzione caratteristica di una gaussiana! Questo significa quindi che la Y è distribuita normalmente, come volevamo mostrare.

Da notare che assunzioni *fondamentali* per la validità del teorema del limite centrale sono l'indipendenza delle  $X_i$  e il fatto che siano identicamente distribuite, nonché il fatto che bisogna prendere il limite  $n \to \infty$ .

La letteratura su questo teorema e le sue applicazioni è vastissima, e importanti risultati sono stati ad esempio raggiunti grazie alla teoria del Gruppo di Rinormalizzazione (opportunamente "ripulita" e interpretata da un punto di vista probabilistico) per generalizzare il teorema del limite centrale nel caso in cui le  $X_i$  siano debolmente dipendenti. Come vedremo più avanti, il teorema del limite centrale va "completato" con la teoria delle grandi deviazioni (capitolo 7).

#### 1.3 L'equazione di diffusione

Fin'ora abbiamo studiato il moto browniano di una singola particella, andando a studiare come evolvono nel tempo alcune sue proprietà, come la velocità o la posizione. Possiamo però anche studiare il problema da un altro punto di vista, definendo una densità di probabilità  $\rho(\vec{r},t)$  di trovare la particella in  $\vec{r}$  all'istante t e chiedendoci se possiamo determinare l'equazione che ne regola l'evoluzione temporale.

Supponiamo quindi che la nostra particella browniana si trovi in  $\vec{r}$  all'istante t; chiamiamo allora  $\mathcal{P}(\delta \vec{r}, \delta t | \vec{r}, t)$  la densità di probabilità condizionata che partendo da  $\vec{r}$  all'istante t la particella "salti" in  $\vec{r} + \delta \vec{r}$  all'istante  $t + \delta t$ . Si ha quindi:

$$\rho(\vec{r}, t + \delta t) = \int d(\delta \vec{r}) \mathcal{P}(\delta \vec{r}, \delta t | \vec{r} - \delta \vec{r}, t) \rho(\vec{r} - \delta \vec{r}, t)$$
(1.5)

Definito ora  $\vec{r}' := \vec{r} - \delta \vec{r}$ , vogliamo esprimere l'integrando nella (1.5) in funzione di esso ed espanderlo attorno a  $\delta \vec{r} = 0$ , ossia  $\vec{r}' = \vec{r}$ . Si ha:

$$\mathcal{P}(\delta\vec{r},\delta t|\vec{r}',t)\rho(\vec{r}',t) = \mathcal{P}(\delta\vec{r},\delta t|\vec{r},t)\rho(\vec{r},t) - \delta\vec{r} \cdot \left[\vec{\nabla}_{\vec{r}'}\left(\mathcal{P}(\delta\vec{r},\delta t|\vec{r}',t)\rho(\vec{r}',t)\right)\right]_{|\vec{r}'=\vec{r}} + \frac{1}{2}\sum_{i,j=1}^{3} \delta r_{i}\delta r_{j} \left[\frac{\partial}{\partial r'_{i}}\frac{\partial}{\partial r'_{j}}\mathcal{P}(\delta\vec{r},\delta t|\vec{r}',t)\rho(\vec{r}',t)\right]_{|\vec{r}'=\vec{r}} + \cdots$$

e quindi la (1.5) diventa:

$$\rho(\vec{r}, t + \delta t) = \underbrace{\int d(\delta \vec{r}) \mathcal{P}(\delta \vec{r}, \delta t | \vec{r}, t)}_{=1} \rho(\vec{r}, t) - \vec{\nabla}_{\vec{r}'} \left( \langle \delta \vec{r} \rangle_{\vec{r}'}, \rho(\vec{r}', t) \right)_{|\vec{r}' = \vec{r}} + \frac{1}{2} \sum_{i,j=1}^{3} \frac{\partial^{2}}{\partial r'_{i} \partial r'_{j}} \left( \langle \delta r_{i} \delta r_{j} \rangle_{\vec{r}'}, \rho(\vec{r}', t) \right)_{|\vec{r}' = \vec{r}} + \cdots$$

ove abbiamo definito:

$$\langle f(\delta r) \rangle_{\vec{r}'} = \int d(\delta \vec{r}) \mathcal{P}(\delta \vec{r}, \delta t | \vec{r}', t) f(\delta \vec{r})$$

Riscriviamo quindi la (1.5) come:

$$\rho(\vec{r},t+\delta t)-\rho(\vec{r},t) = -\vec{\nabla}_{\vec{r}^{\,\prime}} \left( \langle \delta \vec{r} \rangle_{\vec{r}^{\,\prime}} \, \rho(\vec{r}^{\,\prime},t) \right)_{|\vec{r}^{\,\prime}=\vec{r}} + \frac{1}{2} \sum_{i,j=1}^{3} \frac{\partial^{2}}{\partial r_{i}^{\prime} \partial r_{j}^{\prime}} \left( \langle \delta r_{i} \delta r_{j} \rangle_{\vec{r}^{\,\prime}} \, \rho(\vec{r}^{\,\prime},t) \right)_{|\vec{r}^{\,\prime}=\vec{r}} + \cdots$$

Il primo termine del secondo membro è nullo, perché la densità di probabilità  $\mathcal{P}$  è isotropa in  $\delta \vec{r}$ , e quindi ha primo momento nullo:  $\langle \delta \vec{r} \rangle_{\vec{r}'} = 0$ . Nel secondo termine, poi,  $\langle \delta r_i \delta r_j \rangle_{\vec{r}'}$ , è nullo se  $i \neq j^{10}$ , mentre se i = j allora sappiamo che  $\langle \delta r_i^2 \rangle_{\vec{r}'} = 2D\delta t$ . Pertanto:

$$\rho(\vec{r}, t + \delta t) - \rho(\vec{r}, t) = \frac{1}{2} \sum_{i=1}^{3} \frac{\partial^{2}}{\partial r_{i}^{2}} \left( 2D\delta t \rho(\vec{r}', t) \right)_{|\vec{r}' = \vec{r}} = D\delta t \nabla^{2} \rho(\vec{r}, t)$$

Dividendo ambo i membri per  $\delta t$  e prendendo il limite  $\delta t \to 0$  si determina finalmente l'equazione che regola l'evoluzione temporale della  $\rho$ , che è l'*equazione di diffusione*:

$$\frac{\partial}{\partial t}\rho(\vec{r},t) = D\nabla^2\rho(\vec{r},t)$$

Rispetto alla situazione che abbiamo studiato adesso si possono considerare anche casi leggermente più complicati. Ad esempio possiamo immaginare che la particella sia soggetta ad un potenziale esterno  $U(\vec{r})$  (e quindi ad una forza  $\vec{F} = -\vec{\nabla} U(\vec{r})$ ), e il problema andrà quindi risolto tenendo conto anche delle condizioni fisiche del sistema stesso (in questo caso l'equazione che regola l'evoluzione della  $\rho$  è l'equazione di Fokker-Planck, che è una generalizzazione dell'equazione di diffusione). Ad esempio se la viscosità del fluido è molto elevata il moto della particella avviene in condizione ipersmorzata, e dall'equazione di Fokker-Planck del sistema si ricava che all'equilibrio  $\rho_{\rm eq}(\vec{r}) \propto e^{-U(\vec{r})/k_BT}$  (che è un'espressione che ricorda molto quello che abbiamo visto in meccanica statistica), e questo deriva dal fatto che nel sistema si instaurano due correnti: una dinamica dovuta al potenziale esterno, e una diffusiva dovuta alla normale diffusione delle particelle, e all'equilibrio queste due correnti si devono controbilanciare (questa condizione di reciproca eliminazione di questi due contributi vedremo più avanti corrispondere esattamente alla condizione di *bilancio dettagliato*, tipica delle situazioni di equilibrio).

#### 1.4 Processi stocastici

Cerchiamo ora di trattare in modo matematicamente più corretto i processi stocastici; in questo modo riusciremo anche a spiegare e giustificare meglio l'origine di alcune assunzioni che abbiamo fatto sulla forza fluttuante  $\vec{f}(t)$ .

Cominciamo considerando un processo deterministico: questo è descritto da una funzione X(t) che soddisfa a un'equazione differenziale ordinaria del tipo:

$$\frac{d}{dt}X(t) = A(X(t), t) \tag{1.6}$$

 $<sup>^{10}</sup>$ Essendo le componenti della  $\vec{f}$  scorrelate fra loro, anche le componenti della posizione della particella non sono correlate.

ove A è una qualche funzione regolare dei suoi argomenti. Questa equazione può essere espressa equivalentemente in un forma "meno elegante", ossia in forma incrementale:

$$X(t+dt) = X(t) + A(X(t), t)dt$$
  $dt > 0$  (1.7)

e supporremo sempre dt>0 perché pensiamo a un processo che evolve nel tempo (alias "il tempo scorre in avanti").

Nonostante questa forma sia decisamente esteticamente meno appagante della precedente ha comunque il pregio di essere facilmente interpretabile come *equazione iterativa*, ossia come una sorta di "algoritmo" da usare per determinare l'evoluzione del sistema: conoscendo il valore di X a t (e ovviamente conoscendo anche A) siamo facilmente in grado di determinarne il valore a un istante immediatamente successivo, e ripetendo più volte l'operazione possiamo determinare tutta l'evoluzione del processo.

Il fatto che questa equazione ci permetta di determinare X a un istante immediatamente successivo ci permette comunque di definire il processo come "deterministico", e per il fatto che da questa derivi  $X(t+dt) \to X(t)$  per  $dt \to 0$  possiamo dire che il processo è continuo; il fatto poi che X(t+dt) dipenda esclusivamente dal valore di X in t, e non in istanti precedenti, ci permette di chiamare il processo privo di memoria, o più semplicemente markoviano. Insomma, le due equazioni che abbiamo appena visto descrivono un processo deterministico markoviano continuo, e sono risolvibili nell'àmbito del calcolo differenziale standard.

Ora, affinché l'equazione (1.6) sia sensata, X(t) dev'essere ovviamente differenziabile. Potremmo però pensare che la richiesta di differenziabilità della X sia troppo restrittiva in determinati casi, e quindi la (1.7) risulta decisamente più utile in quanto ha validità generale e non presuppone nulla sulla differenziabilità della X. Potremmo allora pensare che in determinati processi possano valere equazioni diverse ma della stessa forma, ad esempio:

$$X(t+dt) = X(t) + A(X(t),t)(dt)^{1/2}$$
(1.8)

In questo modo la X(t) sicuramente non è differenziabile, pur essendo continua; se infatti riscriviamo quest'ultima equazione come:

$$\frac{X(t+dt) - X(t)}{dt} = \frac{A(X(t),t)}{(dt)^{1/2}}$$

nel limite  $dt \to 0$  il secondo membro diverge.

Indipendentemente da questo fatto, però, la (1.8) non è comunque una possibile valida equazione per la descrizione dell'evoluzione di un processo perché *non* è autoconsistente. Consideriamo infatti l'intervallo [t,t+dt] e dividiamolo in due; ovviamente l'equazione che regola il processo, per essere sensata, dovrà essere tale che il risultato della sua applicazione nel far evolvere il sistema da t a t+dt sia lo stesso che si ha facendolo evolvere prima fra t e t+dt/2 e poi fra t+dt/2 e t+dt. Se però applichiamo due volte la (1.8) allora:

$$X\left(t + \frac{dt}{2}\right) = X(t) + A(X(t), t) \left(\frac{dt}{2}\right)^{1/2}$$

$$X(t + dt) = X\left(t + \frac{dt}{2}\right) + A\left(X\left(t + \frac{dt}{2}\right), t + \frac{dt}{2}\right) \left(\frac{dt}{2}\right)^{1/2}$$

$$\downarrow \downarrow$$

$$X(t+dt) = = X(t) + A(X(t),t) \left(\frac{dt}{2}\right)^{1/2} + A\left(X\left(t + \frac{dt}{2}\right), t + \frac{dt}{2}\right) \left(\frac{dt}{2}\right)^{1/2} =$$

$$= X(t) + 2A(X(t),t) \left(\frac{dt}{2}\right)^{1/2} + \cdots$$

$$\downarrow \qquad \qquad \qquad \downarrow$$

$$X(t+dt) = X(t) + 2^{1/2}A(X(t),t)(dt)^{1/2}$$

che non coincide col risultato della (1.8)!

Un processo markoviano continuo, pertanto, può solo essere descritto da un'equazione della forma (1.7).

#### 1.4.1 Processi stocastici markoviani continui

Quello che ora vogliamo fare è generalizzare queste considerazioni al caso in cui X è una variabile stocastica, ossia X(t) è un processo stocastico; parleremo in questo caso di processi stocastici markoviani continui.

Indichiamo con  $\mathcal{N}(m, \sigma^2)$  una generica variabile aleatoria gaussiana, la cui densità di probabilità ha media m e varianza  $\sigma^2$ ; in altre parole se  $\mathcal{N}$  può assumere valori  $y \in \mathbb{R}$  allora:

$$P(y) = \frac{e^{-\frac{(y-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

Le variabili aleatorie distribuite normalmente soddisfano a due particolari proprietà <sup>11</sup>:

$$\alpha + \beta \mathcal{N}(m, \sigma^2) = \mathcal{N}(\alpha + \beta m, \beta^2 \sigma^2) \qquad \forall \alpha, \beta \in \mathbb{R}$$
 (1.9)

$$\mathcal{N}(m_1, \sigma_1^2) + \mathcal{N}(m_2, \sigma_2^2) = \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$$
(1.10)

In particolare, se  $\overline{N}$  è una variabile distribuita normalmente con media nulla e varianza unitaria, ossia  $\overline{N} = \mathcal{N}(0,1)$ , allora la (1.9) si riduce a:

$$\alpha + \beta \overline{N} = \mathcal{N}(\alpha, \beta^2) \tag{1.11}$$

A questo punto vogliamo individuare qual è la forma più generale possibile dell'incremento condizionato per la X. In altre parole, definiamo:

$$\Xi(dt; x, t) = X(t + dt) - X(t)$$
 con la condizione che  $X(t) = x$ 

e vogliamo capire quale forma deve avere  $\Xi$  per soddisfare alcune proprietà che adesso illustriamo.

Da notare che l'incremento  $\Xi$  dipende solo da x (il valore di X in t), t e dt: la memoria del processo è limitata alla sola x, e non c'è traccia dell'evoluzione del processo a tempi

$$P(z) = \int dy \frac{e^{-\frac{(y-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \delta(z - \alpha - \beta y)$$

Anche la seconda proprietà si verifica nello stesso modo, ma è più laboriosa.

<sup>&</sup>lt;sup>11</sup>Ad esempio la prima si verifica andando a calcolare:

minori di t.

Facciamo quindi alcune ipotesi su  $\Xi$ : dato che vogliamo studiare processi markoviani *continui*, supponiamo che la  $\Xi$  sia una funzione regolare dei suoi argomenti, e che per ogni x e t si abbia  $\Xi \to 0$  per  $dt \to 0$ . Ciò che vogliamo dimostrare è che con queste ipotesi l'incremento deve avere *necessariamente* la forma:

$$\Xi(dt;x,t) = A(x,t)dt + D^{1/2}(x,t)N(t)(dt)^{1/2}$$
(1.12)

ove A è detta funzione di drift e D funzione di diffusione; la N(t) inoltre è una variabile distribuita normalmente ad ogni t, con media nulla e varianza unitaria (in altre parole, ad ogni istante esiste una  $N = \mathcal{N}(0,1)$  tale che N(t) e N(t') sono indipendenti per  $t \neq t'$ ). Mostriamo la validità della (1.12) usando il teorema del limite centrale.

Consideriamo l'intervallo [t,t+dt] e immaginiamo di dividerlo in n sottointervalli ognuno lungo dt/n; chiamiamo allora  $t_0=t$ ,  $t_1=t+dt/n$ , in generale  $t_i=t+i\cdot dt/n$ , fino a  $t_n=t+dt$ . Poiché il processo deve essere consistente, nel senso che abbiamo illustrato precedentemente, allora:

$$\Xi(dt; X(t), t) = \sum_{i=1}^{n} \Xi\left(\frac{dt}{n}; X(t_{i-1}), t_{i-1}\right) = \sum_{i=1}^{n} \Xi_i\left(\frac{dt}{n}; x, t\right)$$

Nell'ultima uguaglianza abbiamo in realtà effettuato implicitamente più passaggi; innanzitutto abbiamo sfruttato la continuità della  $\Xi$  dai suoi argomenti, grazie alla quale possiamo infatti "permetterci" di sostituire i vari  $t_i$  con l'istante iniziale t, e porre quindi anche X(t) uguale al suo valore inziale x (questo può solo far sorgere dei piccoli termini correttivi, che trascuriamo). Abbiamo poi usato la notazione  $\Xi_i$  per indicare che tutti gli incrementi, pur avendo la stessa forma ed essendo tutti condizionati dal (singolo) valore x di X, sono indipendenti fra loro (è come se campionassimo n volte una stessa variabile aleatoria).

Ora, dato che stiamo sommando n (con n arbitrariamente grande) variabili statisticamente indipendenti (le  $\Xi_i$  sono indipendenti fra loro per la markovianità del processo), per il teorema del limite centrale la loro somma  $\Xi$  è distribuita normalmente (e quindi per autoconsistenza lo sono anche le  $\Xi_i$ ). Valgono allora le proprietà che abbiamo mostrato prima relative alla variabili gaussiane:

$$\operatorname{var}\left\{\Xi(dt;x,t)\right\} = \sum_{i} \operatorname{var}\left\{\Xi_{i}\left(\frac{dt}{n};x,t\right)\right\} = n \cdot \operatorname{var}\left\{\Xi\left(\frac{dt}{n};x,t\right)\right\}$$

$$\operatorname{mean}\left\{\Xi(dt;x,t)\right\} = \sum_{i} \operatorname{mean}\left\{\Xi_{i}\left(\frac{dt}{n};x,t\right)\right\} = n \cdot \operatorname{mean}\left\{\Xi\left(\frac{dt}{n};x,t\right)\right\}$$

ove ovviamente per "mean" intendiamo la media e per "var" la varianza della distribuzione di una variabile.

A questo punto, sfruttiamo un teorema di analisi (che non dimostriamo); nella sua formulazione generale asserisce che se h(z) è una qualunque funzione regolare di z tale che h(z) = nh(z/n) per ogni n intero positivo, allora h(z) = Cz con C indipendente da z. Applicato al nostro caso questo teorema ci permette di stabilire che:

$$\operatorname{var}\left\{\Xi(dt;x,t)\right\} = D(x,t)dt$$
 mean  $\left\{\Xi(dt;x,t)\right\} = A(x,t)dt$ 

ove A e D possono dipendere da x e t ma non da dt (e D è sicuramente non negativa). Pertanto, poiché  $\Xi$  è distribuita normalmente possiamo scrivere:

$$\Xi(dt; x, t) = \mathcal{N}(A(x, t)dt, D(x, t)dt)$$

Usando la (1.11) si può quindi scrivere effettivamente:

$$\Xi(dt;x,t) = X(t+dt) - X(t) = A(X(t),t)dt + D^{1/2}(X(t),t)\mathcal{N}(0,1)(dt)^{1/2}$$

e considerando che la variabile normale unitaria va presa istante per istante, chiamandola N(t) possiamo concludere che:

$$X(t+dt) = X(t) + A(X(t),t)dt + D^{1/2}(X(t),t)N(t)(dt)^{1/2}$$

che è proprio la (1.12). Il termine proporzionale a dt è generalmente detto "di drift", mentre quello proporzionale a  $(dt)^{1/2}$  è detto "di diffusione".

Possiamo quindi concludere che questa è l'unica forma che può avere l'equazione che regola un processo stocastico markoviano continuo.

Ora, osservando l'equazione si potrebbe però obiettare che stiamo sommando due termini di ordini diversi, e che in particolar modo per  $dt \to 0$  dòmini quello proporzionale a  $(dt)^{1/2}$ . In realtà questo non è vero, perché il termine in  $(dt)^{1/2}$  è moltiplicato per la variabile aleatoria N(t) che può assumere sia valori positivi che negativi: di fatto la N(t) "attenua" il termine diffusivo, rendendolo dello stesso ordine di quello di drift.

Proviamo poi ad adottare l'ottica del fisico nei confronti dell'equazione (1.12); dividendo ambo i membri per dt si ottiene:

$$\frac{d}{dt}X(t) = A(X(t), t) + D^{1/2}(X(t), t) \frac{N(t)}{(dt)^{1/2}}$$

che, analogamente a quello che abbiamo detto discutendo la validità della (1.8), nel limite  $dt \to 0$  è singolare<sup>12</sup>; ciò lo si può anche capire dal fatto che  $N(t)/(dt)^{1/2} = \mathcal{N}(0,1)/(dt)^{1/2} = \mathcal{N}(0,1/dt)$  nel limite  $dt \to 0$  è una variabile la cui densità di probabilità è una gaussiana di varianza infinita (mandando  $dt \to 0$  la gaussiana si "spancia" su tutto l'asse reale). Ciò che tradizionalmente si fa per dare un senso all'equazione che abbiamo appena scritto, e in particolar modo alla nozione di "derivata" dX(t)/dt, è di definire il  $tumore\ bianco\ gaussiano\ come$ :

$$\Gamma(t) = \lim_{dt \to 0} \mathcal{N}\left(0, \frac{1}{dt}\right)$$

In questo modo l'equazione precedente può essere riscritta come:

$$\frac{d}{dt}X(t) = A(X(t),t) + D^{1/2}(X(t),t)\Gamma(t)$$

detta *equazione di Langevin*, che altro non è che la generalizzazione della (1.6) per processi stocastici.

Le proprietà del rumore bianco gaussiano  $\Gamma(t)$  discendono semplicemente dalla sua definizione:

$$\langle \Gamma(t) \rangle = 0$$
  $\qquad \langle \Gamma(t)\Gamma(t+t') \rangle = \begin{cases} 0 & t' \neq 0 \\ \frac{1}{dt} & t' = 0 \end{cases} = \delta(t')$ 

così come quelle di N(t):

$$\langle N(t) \rangle = 0$$
  $\langle N(t)N(t') \rangle = \begin{cases} 0 & t \neq t' \\ 1 & t = t' \end{cases}$   $\langle X(t')N(t) \rangle = 0 \text{ se } t' \leq t$ 

 $<sup>\</sup>overline{\ \ \ }^{12}$ Questo non accadrebbe se il termine diffusivo D(x,t) fosse nullo, ma in questo caso il processo sarebbe deterministico e non più stocastico quindi la situazione perderebbe di interesse.

ove l'ultima rappresenta proprio la markovianità del processo.

Se ora ripensiamo a quello che abbiamo detto in 1.1 studiando le proprietà generali del moto browniano, possiamo riconoscere in  $\Gamma$  il ruolo della funzione fluttuante  $\vec{f}$  che modellizza gli urti della particella con le molecole del fluido: è quindi chiara l'origine delle proprietà della  $\vec{f}$  che in 1.1 abbiamo solo elencato.

#### 1.4.2 Il processo di Ornstein-Uhlenbeck

Consideriamo ora un particolare tipo di processo stocastico (uno dei pochi esempi completamente caratterizzabili), il cosiddetto *processo di Ornstein-Uhlenbeck*. Si tratta del processo col quale possiamo studiare il moto browniano, e nel linguaggio che abbiamo sviluppato precedentemente l'incremento della X(t) per esso è dato da:

$$X(t+dt) = X(t) - \frac{x}{\tau}dt + c^{1/2}N(t)(dt)^{1/2}$$
(1.13)

ossia  $A(x,t) = -x/\tau$  e D(x,t) = c; si dice anche che il processo di Ornstein-Uhlenbeck in esame ha *tempo di rilassamento*  $\tau$  e *costante di diffusione c*.

Un'importante proprietà di questo processo, innanzitutto, è che per esso la X(t) risulta essere una variabile distribuita normalmente che evolve nel tempo, ossia la densità di probabilità della X è una gaussiana che cambia nel tempo; pertanto note le leggi secondo le quali evolvono sia la sua media che la sua varianza abbiamo completamente determinato il processo. La gaussianità della X(t) discende dalle proprietà delle variabili distribuite normalmente, in particolare la (1.10) (ossia il fatto che la somma di due variabili gaussiane indipendenti è ancora distribuita normalmente). Infatti, supponiamo che all'istante iniziale  $t_0$  si abbia  $X(t_0) = x_0$ ; poiché in quell'istante conosciamo esattamente il valore di  $x_0$ , la densità di probabilità della X può essere considerata una  $\delta$  di Dirac, ossia una gaussiana di varianza nulla. In altre parole,  $X(t_0) = x_0 = \mathcal{N}(x_0,0)$  e allora per la (1.13) si ha:

$$X(t+dt) = x_0 - \frac{x_0}{\tau}dt + c^{1/2}N(t_0)(dt)^{1/2}$$

Quindi, dato che X(t+dt) è una somma di variabili gaussiane indipendenti allora è essa stessa una variabile distribuita normalmente. Analogamente, applicando lo stesso ragionamento si ha che anche X(t+2dt) è gaussiana, e reiterando possiamo concludere che X(t) è gaussiana per ogni  $t>t_0$ .

Si può dimostrare che l'espressione della X(t) come funzione di t per un processo di Ornstein-Uhlenbeck descritto dalla (1.13) è:

$$X(t) = \mathcal{N}\left(x_0 e^{-\frac{t-t_0}{\tau}}, \frac{c\tau}{2}\left(1 - e^{-2\frac{t-t_0}{\tau}}\right)\right)$$

Da notare come  $\tau$  sia proprio la scala di tempo caratteristica nella quale la media e la varianza della distribuzione della X rilassino verso i loro valori asintotici 0 e  $c\tau/2$ . In altre parole si ha:

$$\lim_{t \to \infty} X(t) = \mathcal{N}\left(0, \frac{c\tau}{2}\right)$$

ossia col passare del tempo il processo tende ad essere *stazionario* (vedi 5.1), e la stazionarietà viene raggiunta con un tempo esponenziale caratteristico pari a  $\tau$ . **Processo di Wiener** Un caso interessante si ha nel limite  $\tau \to \infty$ , nel quale la (1.13) diventa:

$$X(t+dt) = X(t) + c^{1/2}N(t)(dt)^{1/2}$$

ossia è presente solo il termine di rumore. In questo caso il processo prende il nome di processo di Wiener, e se il dato iniziale è  $X(0)=x_0$  allora:

$$\langle X(t)\rangle = x_0$$
  $\langle X(t)^2\rangle = c(t - t_0) + x_0^2$ 

Pertanto l'espressione di X(t) è data da:

$$X(t) = \mathcal{N}(x_0, c(t - t_0))$$

## Capitolo 2

# La percolazione

La *percolazione* è un problema "classico" nell'àmbito dei sistemi complessi e disordinati, che ha numerosissime applicazioni. Per comprenderlo più intuitivamente, vediamo prima alcune di queste e andiamo poi ad "astrarre" il problema per dare una formulazione generale della percolazione.

#### 2.1 Introduzione: esempi e applicazioni

In àmbito industriale è possibile creare filtri usando materiali ceramici creando dei "vuoti" all'interno di una matrice, con una densità desiderata (un po' come se fosse una spugna). È quindi fondamentale riuscire a capire innanzitutto quale sia la densità minima di questi vuoti affinché il filtro permetta il passaggio di fluidi, e in secondo luogo come il flusso passante per il filtro vari in funzione di questa densità di vuoti.

Ci sarà infatti una soglia di densità oltre la quale è possibile individuare un "cammino" nella matrice dato da vuoti che si "uniscono" e permettono al fluido di passare (e vedremo fra poco che questo prende il nome di *cluster infinito* o *percolante*).

Consideriamo ora una situazione diversa. Supponiamo di avere una miscela di microsfere di due materiali diversi, alcune che conducono corrente e altre isolanti. Immaginiamo poi di mettere questa miscela di microsfere in un grande contenitore, ad esempio una scatola con le pareti metalliche, e di andare ad applicare una differenza di potenziale fra due lati di questa scatola e misurare la corrente che passa nel sistema. Variando la densità di sfere conduttrici si trova una situazione simile alla precedente: per densità troppo basse il sistema non conduce, poi oltre una certa soglia si creano "cammini" che permettono alla corrente di passare da un lato all'altro del contenitore, e la corrente aumenta all'aumentare della densità di sfere conduttrici.

Situazioni di questo tipo, come vedremo fra poco, sono descrivibili con lo studio del fenomeno della percolazione.

### 2.2 Formulazione generale della percolazione

La prima formulazione "matematica" della percolazione risale agli anni '50 ad opera di Hammersley. In genere si distinguono due diverse possibili formulazioni della percolazione: quella "di sito" e quella "di legame". Noi analizzeremo quella di sito, per poi

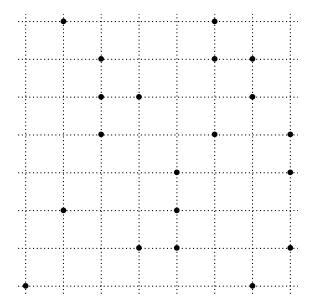


Figura 2.1: Esempio di configurazione di siti occupati

vedere brevemente al momento opportuno le differenze con quella di legame.

Consideriamo quindi un reticolo (per semplicità pensiamolo bidimensionale, ma ovviamente può essere in qualunque dimensione), e supponiamo che ogni sito possa essere occupato indipendentemente dagli altri con una data probabilità p, data a priori; fissata dunque p il reticolo avrà una data configurazione di siti occupati. In figura 2.1 ne è raffigurato un esempio.

Stabiliamo poi che se due siti primi vicini sono occupati li connettiamo insieme; chiamiamo poi *cluster* un insieme di siti occupati che sono primi vicini fra loro (in figura 2.2 i vari cluster sono evidenziati in colori diversi e dai legami in grassetto); supponiamo anche che ogni sito occupato isolato costituisca un cluster. La percolazione allora non è altro che lo studio delle proprietà statistiche di cluster connessi di questo tipo; come vedremo meglio in seguito, nel limite termodinamico (ossia nel limite di un reticolo infinitamente esteso) i cluster si divideranno in due "categorie": *finiti* e *infiniti*. Per *cluster infinito* si intende un cluster¹ che si estende su *tutto* il sistema, raggiungendo anche il suo bordo all'infinito (nel caso degli esempi precedenti, un cluster infinito rende possibile la "comunicazione" fra le due facce opposte del filtro o della scatola metallica); un cluster finito si estende invece solo su una porzione finita di sistema, trascurabile nel limite termodinamico (in figura 2.3 sono rappresentati degli esempi di clusters finiti e infiniti). Lo studio della percolazione è quindi "ridotto" allo studio delle proprietà statistiche e geometriche di questi clusters in funzione della probabilità *p*, che è il parametro del nostro sistema che gioca un ruolo simile a quello della temperatura in modelli come quello di Ising.

La percolazione di legame è molto simile a quella di sito che abbiamo appena visto. In questo caso, però, si suppone che tutti i siti siano occupati, e che p sia la probabilità che si stabilisca un legame fra primi vicini. Rispetto al caso della percolazione di sito, quindi,

<sup>&</sup>lt;sup>1</sup>Si può anche dimostrare, ma non è affatto semplice, che in un sistema percolativo esiste un unico cluster infinito.

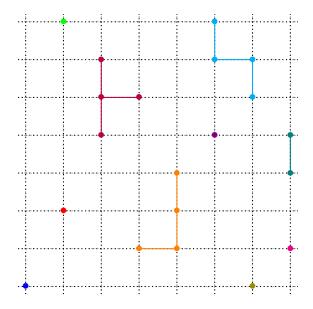


Figura 2.2: Cluster della configurazione di figura 2.1

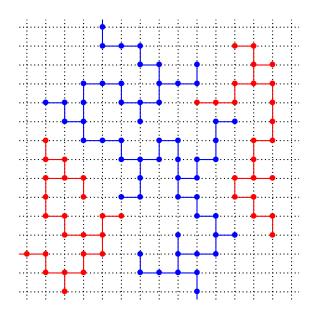


Figura 2.3: Cluster finiti (in rosso) e infinito (in blu)

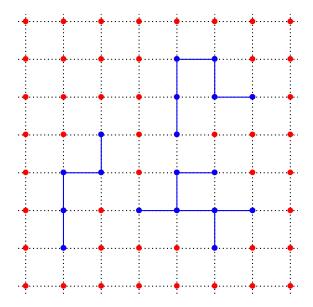


Figura 2.4: Esempio di percolazione di legame

nella percolazione di legame *tutti* i siti fanno parte di un cluster (nel caso precedente sono solo quelli occupati); un esempio di configurazione di un sistema di percolazione di legame è dato in figura 2.4.

Le differenze concettuali fra questi due sistemi sono quindi minime, ma comunque esistono; noi continueremo sempre nel seguito a considerare un problema di percolazione di sito, se non diversamente specificato.

#### 2.3 Percolazione di sito

Consideriamo quindi un reticolo di N siti, e chiamiamo S il numero di siti in un cluster connesso; se poi definiamo  $N_S(p)$  il numero totale di cluster a S siti data la probabilità p di occupazione del generico sito del reticolo, ci aspettiamo allora che  $N_S(p)/N$  tenda ad una quantità ben definita nel limite termodinamico:

$$\frac{N_S(p)}{N} \stackrel{N \to \infty}{\longrightarrow} n_S(p)$$

ove  $n_S(p)$  è la frazione media di cluster finiti a S siti.

Consideriamo ad esempio  $SN_S(p)/N$ : questa è la frazione media di siti che appartengono a un cluster di S siti; pertanto nel limite termodinamico si ha:

$$\frac{SN_S(p)}{N} \stackrel{N \to \infty}{\longrightarrow} sn_s(p)$$

che è la frazione media di siti in un cluster a s siti, ossia la probabilità che un sito appartenga a un cluster finito di s siti.

Chiamiamo poi P(p) la probabilità che un generico sito appartenga a un cluster infinito. Come possiamo collegare P(p) alle quantità appena viste?

Innanzitutto, affinché il sito appartenga a un cluster infinito deve essere perlomeno occupato, poi bisogna vedere se appartiene a un cluster finito o infinito. Insomma, vale la

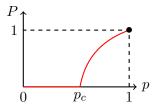


Figura 2.5: Andamento di P in funzione di p

seguente relazione:

$$p = P(p) + \sum_{s=1}^{\infty} s n_s(p)$$
(2.1)

ossia, "a parole", si ha che se un sito è occupato (e lo è con probabilità p) allora o appartiene a un cluster infinito (con probabilità P(p)) oppure appartiene a un *qualunque* cluster finito (con probabiltà data dalla sommatoria).

Come abbiamo visto anche dagli esempi precedenti, quello che ci aspettiamo è che nel limite termodinamico P(p) sia nulla fino ad un valore critico  $p_c$  di p, oltre il quale² cresce fino a che P(p)=1 quando p=1 (come mostrato in figura 2.5). Si tratta insomma di una situazione analoga a quella della magnetizzazione spontanea di un sistema in funzione della sua temperatura.

Data quest'analogia fra sistemi percolativi e magnetici, ci aspettiamo quindi che  $P(p) \sim (p-p_c)^{\beta}$  per  $p \to p_c^+$ .

Quest'analogia va però molto oltre.

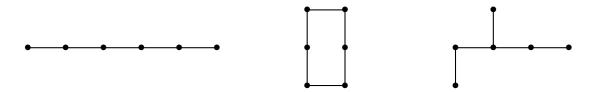
Possiamo infatti caratterizzare un cluster, oltre che dal numero dei siti che lo compongono, dalla sua geometria; cerchiamo quindi di definire le dimensioni medie di un cluster finito.

Fissato il numero S di siti in un cluster, questo può presentarsi in moltissime configurazioni<sup>3</sup>; per ognuna di esse, considerandole come un insieme di punti materiali, possiamo determinare il raggio d'inerzia rispetto al centro di massa. Chiamando  $R_s$  il raggio d'inerzia medio per un cluster di s siti<sup>4</sup>, allora possiamo definire le dimensioni lineari di un cluster finito come:

$$\xi(p) = \frac{\sum_{s} n_s(p) R_s}{\sum_{s} n_s(p)}$$

Se p è piccola, il sistema presenterà pochi e piccoli clusters, e via via che p cresce questi diventano sempre più estesi e numerosi. A  $p=p_c$  come abbiamo detto devono comparire

 $<sup>^3\</sup>mathrm{A}$ puro titolo di esempio, seS=6alcuni dei possibili cluster sono:



 $<sup>^4</sup>$ Da notare che la determinazione di  $R_s$  è anche un problema "enumerativo", perché per calcolarlo si devono determinare tutte le possibili configurazioni del cluster e valutarne il raggio.

<sup>&</sup>lt;sup>2</sup>Se il sistema è finito la transizione a  $p=p_c$  in figura 2.5 non è brusca, ma graduale; diventa "improvvisa" solo nel limite termodinamico.

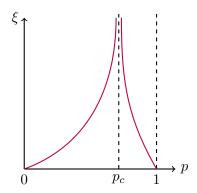


Figura 2.7: Andamento di  $\xi$  in funzione di p

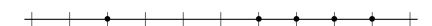


Figura 2.8: Esempio di configurazione per un sistema percolativo unidimensionale

clusters infiniti, e quindi ci aspettiamo che  $\xi(p)$  diverga per  $p \to p_c^-$ . Se invece partiamo da p=1 il nostro sistema è costituito da un unico cluster infinito; facendo diminuire p all'interno di questo cluster si formano delle vacanze, all'interno delle quali si possono a loro volta creare dei cluster finiti. Al diminuire di p verso  $p_c$  le dimensioni di questi cluster aumentano, e il loro comportamento rispecchia esattamente quello che si aveva per  $p < p_c$ .

Pertanto,  $\xi(p)$  diverge sia per  $p \to p_c^-$  che per  $p \to p_c^+$  (come illustrato in figura 2.7):

$$\xi(p) \sim |p - p_c|^{\nu}$$

Pensando all'analogia col caso magnetico, quindi,  $\xi$  è il corrispettivo percolativo della lunghezza di correlazione.

Possiamo poi anche definire, sempre seguendo l'analogia, una funzione di correlazione per il nostro sistema percolativo.

Chiamiamo  $\Gamma(\vec{r})$  la probabilità che un sito a distanza  $\vec{r}$  da un sito occupato (che per semplicità poniamo nell'origine del nostro sistema di riferimento) appartenga al suo stesso cluster *finito*.

È importante richiedere che il cluster considerato sia *finito*, perché se così non fosse allora ci aspettiamo che:

$$\Gamma(\vec{r}) \stackrel{\vec{r} \to \infty}{\longrightarrow} P^2(p)$$

(questa richiesta deriva dalla necessità di avere definizioni non ambigue; nel caso in cui  $\Gamma$  vada come  $P^2$  per  $\vec{r} \to \infty$ , oggetti come  $\sum_{\vec{r}} \Gamma(\vec{r})$  non convergerebbero).

#### 2.3.1 Percolazione di sito in una dimensione

Proviamo ad applicare quello che abbiamo visto allo studio di un sistema percolativo nel caso più semplice possibile, ossia quello di reticolo unidimensionale.

Consideriamo quindi un reticolo in una dimensione con una data configurazione a p fissato, come mostrato in figura 2.8. Come sappiamo  $sn_s(p)$  è la frazione media di siti occupati che appartengono a un s-cluster (ossia un cluster a s siti), oppure possiamo

interpretarla anche come la probabilità che un generico sito appartenga a un s-cluster. Cominciamo supponendo ad esempio s=1: cos'è allora  $1n_1(p)$ ? Un 1-cluster è un cluster composto da un singolo sito isolato, quindi affinché un sito appartenga a un cluster di questo tipo deve essere occupato e i suoi primi vicini devono essere vacanti. Pertanto:

$$1n_1(p) = p(1-p)^2$$

Se invece consideriamo s=3, esistono tre possibilità affinché un dato sito appartenga a un 3-cluster:

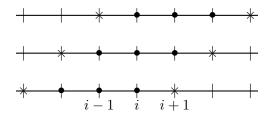


Figura 2.9: Possibili modi in cui un dato sito può appartenere ad un 3-cluster (× indica un sito vacante, • uno occupato)

Dato che il peso di ogni configurazione è  $p^3(1-p)^2$ , si ha:

$$3n_3(p) = 3p^3(1-p)^2$$

ed è pertanto facile intuire che in generale:

$$sn_s(p) = sp^s(1-p)^2$$

Consideriamo ora l'equazione (2.1); sostituendo l'espressione appena vista per  $sn_s(p)$  si ha:

$$p = P(p) + \sum_{s=1}^{\infty} sp^{s}(1-p)^{2}$$

La sommatoria può essere calcolata semplicemente sfruttando la serie geometrica:

$$\sum_{s=1}^{\infty} sp^{s} (1-p)^{2} = (1-p)^{2} \sum_{s=1}^{\infty} sp^{s} = (1-p)^{2} p \frac{\partial}{\partial p} \left( \sum_{s=1}^{\infty} p^{s} \right) =$$

$$= (1-p)^{2} p \frac{\partial}{\partial p} \left( \frac{p}{1-p} \right) = (1-p)^{2} p \frac{1}{(1-p)^{2}} = p$$

Pertanto:

$$p = P(p) + p$$
  $\Rightarrow$   $P(p) = 0$ 

Vediamo quindi come in una dimensione non esista una soglia finita di percolazione (si può avere P(p)=1 solo se p=1); in altre parole, la percolazione ha una dimensione critica infieriore pari ad uno.

Proviamo ora, sempre in una dimensione, a calcolare  $\Gamma(r)$  ( $\vec{r}$  diventa semplicemente un indice, r).

Chiaramente, affinché il sito in r sia nello stesso cluster di quello in r = 0 si deve avere:

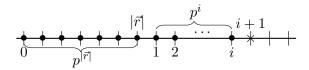


Figura 2.10: Cluster finito unidimensionale (× indica sempre un sito vacante)

Affinché il cluster sia finito i siti oltre r possono essere occupati in modo finito, ossia fino ad un i finito. La probabilità di ognuna di queste configurazioni è<sup>5</sup>:

$$\underbrace{(1-p)\left(\sum_{j=0}^{\infty}p^{j}\right)\cdot p^{|r|}\cdot (1-p)\left(\sum_{i=0}^{\infty}p^{i}\right)}_{\text{contribute di distra}} = p^{|r|}$$

ove abbiamo sfruttato la somma della serie geometrica. Pertanto:

$$\Gamma(r) = p^{|r|} \stackrel{|r| \to \infty}{\longrightarrow} 0$$

che è un'altra evidenza del fatto che in una dimensione non si può avere percolazione (peraltro, anche stavolta il limite può essere non nullo solo se p=1). Ora, per un sistema non critico ci aspettiamo che:

$$\Gamma(r) \stackrel{|r| \to \infty}{\sim} e^{-|r|/\xi(p)}$$

e possiamo scrivere  $\Gamma$  in questa forma anche per il caso di percolazione unidimensionale:

$$\Gamma(r) = p^{|r|} = e^{-|\ln p| \cdot |r|}$$

Possiamo quindi esprimere la dimensione lineare dei clusters  $\xi$  in funzione di p:

$$\xi(p) = \frac{1}{|\ln p|}$$

che ha la stessa forma dell'espressione della lunghezza di correlazione per il modello di Ising in una dimensione.

Infine, per determinare l'esponente critico  $\nu$  possiamo scrivere  $p=1-\varepsilon$ : sostituendo e sviluppando il logaritmo si ha  $\xi \sim 1/\varepsilon$  e quindi  $\nu=1$ .

Nei fenomeni magnetici sappiamo che in generale si può definire la suscettività come:

$$\chi = \sum_{\vec{r}} \Gamma(\vec{r})$$

Nel nostro caso, poiché  $\Gamma(r) = p^{|r|}$ , si ha:

$$\chi = \sum_{r=-\infty}^{+\infty} p^{|r|} = 1 + 2\sum_{r=1}^{+\infty} p^r = 1 + 2\frac{p}{1-p} = \frac{1+p}{1-p}$$

<sup>&</sup>lt;sup>5</sup>Per "contributo di destra" indichiamo il contributo dato dagli eventuali siti occupati alla destra di  $|\vec{r}|$ , mentre per "contributo di sinistra" indichiamo quello relativo a eventuali siti occupati a sinistra di 0 (che non sno rappresentati in figura 2.10).

e la  $\chi$  diverge per  $p \to 1^-$  (è effettivamente l'analogo di una suscettività). Si ha poi:

$$\chi = \frac{1+p}{1-p} = \frac{2}{1-p} - 1 \overset{p \approx p_c}{\sim} \frac{1}{p_c - p} \qquad \Rightarrow \qquad \gamma = -1$$

Possiamo però scrivere la  $\chi$  in un'altra forma: questo parametro infatti misura la grandezza media dei cluster finiti che contengono un dato sito occupato<sup>6</sup>. Questa dimensione media risulta infatti<sup>7</sup>:

 $\frac{\sum_{s} s \cdot s n_{s}(p)}{\sum_{s} s n_{s}(p)} = \dots = \frac{1+p}{1-p} = \chi$ 

#### 2.3.2 Percolazione di sito su reticolo di Bethe

Come abbiamo appena visto, il problema della percolazione in una dimensione è molto semplice da trattare ma ha soluzione banale. Possiamo però definire un problema bidimensionale altrettanto semplice ma che esibisce una soglia di percolazione: è il problema della percolazione di sito definito su un *reticolo di Bethe* (noto anche come *albero di Cayley*). Si tratta di un reticolo *gerarchico* definito recursivamente: si parte da un "sito 0", l'*origine* del nostro reticolo, e a questo accostiamo  $\zeta$  primi vicini; ad ognuno di questi primi vicini si aggiungono poi altri  $\zeta-1$  primi vicini, ai quali si aggiungono altri  $\zeta-1$  primi vicini ciascuno e così via.

D'ora in poi considereremo sempre il caso semplice in cui  $\zeta=3$ , nel quale un reticolo di Bethe ha il seguente aspetto:



Figura 2.11: Reticolo di Bethe

La "superficie" di un sistema di questo tipo, intesa come il numero di siti che si trovano sulla "frontiera" (ossia negli ultimi rami), cresce molto velocemente all'aumentare del numero di ramificazioni.

Come vedremo fra poco, questo tipo di reticolo permette di ricavare soglie di percolazione non banali con non molta più fatica del caso unidimensionale; la proprietà che sfrutteremo per determinarle è il fatto che nel limite termodinamico la struttura di una branca è uguale a quella di ogni sua sottobranca (proprio perché il reticolo è infinito).

Se quindi vogliamo studiare la percolazione di sito su un reticolo del genere, P(p) sarà la probabilità che il sito di origine appartenga a un cluster infinito, ove stavolta per "cluster infinito" si intende un cluster che raggiunge la frontiera del reticolo.

Definiamo quindi una quantità ausiliaria che ci servirà nei conti: chiamiamo Q(p) la probabilità che il sito al vertice di una branca non sia connesso all'infinito; per calcolarla sfruttiamo la struttura gerarchica del reticolo stesso. Considerando il sito evidenziato in

<sup>&</sup>lt;sup>6</sup>In altre parole, dà una misura di quanto sono grandi mediamente i clusters finiti che contengono il sito considerato.

 $<sup>^{7}</sup>$ Per svolgere il calcolo si sfrutta lo stesso truppo che abbiamo visto poco fa per dimostrare che P(p)=0 in una dimensione, ossia si scrivono le sommatorie come derivate (dell'ordine opportuno) della serie geometrica.

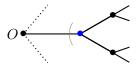


Figura 2.12: Uso della gerarchia del reticolo per il calcolo di Q(p)

blu in figura 2.12 come il vertice della nostra branca, allora si avrà:

$$Q(p) = (1 - p) + pQ^{2}(p)$$

che è facilmente interpretabile: un dato sito non è connesso all'infinito se esso stesso è vacante (e lo è con probabilità 1-p) oppure se il sito è occupato ma entrambe le sue sottobranche sono vacanti (e questo avviene con probabilità  $pQ^2(p)$ ). Risolvendo:

$$Q(p) = \frac{1 \pm \sqrt{1 - 4p(1 - p)}}{2p} = \frac{1 \pm \sqrt{(1 - 2p)^2}}{2p} = \begin{cases} \frac{1 - p}{p} & \text{se } p > 1/2 \\ 1 & \text{altrimenti} \end{cases}$$

Già da ora possiamo quindi capire che effettivamente esiste una soglia di percolazione, in corrispondenza di  $p=p_c=1/2$ : se  $p< p_c$  allora Q(p)=1, e quindi sicuramente nessun sito è connesso all'infinito da un cluster; se invece  $p>p_c$  questa probabilità è minore di uno, e quindi un sito può effettivamente appartenere a un cluster infinito.

Cerchiamo ora di calcolare la probabilità P(p) relativa al sito d'origine; affinché questo appartenga a un cluster infinito bisogna innanzitutto che sia occupato, e poi i suoi tre primi vicini non devono essere tutti sconnessi dalla frontiera del reticolo. Pertanto si ha:

$$P(p) = p(1 - Q(p)^3) = p \left[ 1 - \left(\frac{1-p}{p}\right)^3 \right]$$

ove ovviamente stiamo supponendo  $p>p_c=1/2$ . Vediamo anche qual è l'andamento di P(p) per  $p\to p_c^+$ ; scrivendo  $p=1/2+\varepsilon$ , sostituendo ed espandendo in serie di Taylor si ottiene:

$$P(p) = \left(\frac{1}{2} + \varepsilon\right) \left[1 - \left(\frac{1 - 1/2 - \varepsilon}{1/2 + \varepsilon}\right)^3\right] = \left(\frac{1}{2} + \varepsilon\right) \left[1 - \left(\frac{1/2 - \varepsilon}{1/2 + \varepsilon}\right)^3\right] = \left(\frac{1}{2} + \varepsilon\right) (1 - 1 + 12\varepsilon + O(\varepsilon^2)) = 6\varepsilon + O(\varepsilon^2)$$

(ove abbiamo sfruttato il fatto che  $[(1/2-\varepsilon)/(1/2+\varepsilon)]^3=[(1-2\varepsilon)/(1+2\varepsilon)]^3=(1-4\varepsilon+O(\varepsilon^2))^3=1-12\varepsilon+O(\varepsilon^2)$ ). Pertanto, P(p) cresce linearmente subito dopo la soglia,come mostrato in figura 2.13. Quindi, poiché  $\varepsilon=p-p_c$  allora per  $p\to p_c^+$  si ha  $P(p)\sim (p-p_c)^1$ , e quindi  $\beta=1$ .

Calcoliamo ora l'equivalente della suscettività che abbiamo definito prima, rinominandola S(p):

$$S(p) = \frac{\sum_{s} s^2 n_s(p)}{\sum_{s} s n_s(p)}$$

 $<sup>^8</sup>$ Notare che la probabilità che la sottobranca sia vacante è ancora Q(p), proprio perché come appena detto nel limite termodinamico le strutture di una branca e delle sue sottobranche sono le stesse.

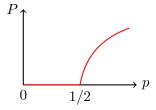


Figura 2.13: Andamento di *P* in funzione di *p* per il reticolo di Bethe

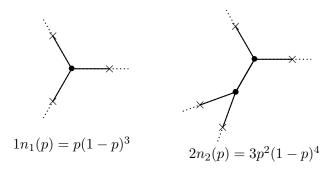


Figura 2.14: Esempi per il calcolo di  $sn_s(p)$  per s=1 e s=2

In questo contesto si presenta però il problema (di natura "enumerativa") di come si possa determinare, ad esempio,  $sn_s(p)$ .

Se partiamo da casi semplici, come s=1 o s=2, le cose non sono troppo complicate come si può vedere dalla figura 2.14. Se però aumentiamo s anche di poco, ad esempio considerando s=5, le cose si complicano un po'; ad esempio, la probabilità di entrambi i clusters in figura 2.15 è proporzionale a  $p^5(1-p)^7$  (ma il coefficiente di proporzionalità non è semplice da determinare, in generale).

In generale, quindi, per un cluster a s siti si ha:

$$sn_s(p) \propto p^s (1-p)^{s+2} \tag{2.2}$$

perché la frontiera del cluster è sempre composta da s+2 siti; lo possiamo mostrare "per induzione".

Innanzitutto, come possiamo vedere in figura 2.14, un cluster composto da pochi siti (s=1 o s=2, nell'esempio), ha una frontiera che è effettivamente formata da s+2 siti. Anche nel caso s=5 dalla figura 2.15 si può vedere come indipendentemente dalla configurazione del cluster la sua frontiera ha s+2=7 siti.

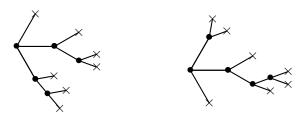


Figura 2.15: Esempi per il calcolo di  $sn_s(p)$  per s=5

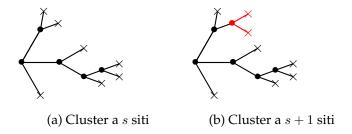


Figura 2.16: Procedura di induzione

Consideriamo ora un generico cluster a s siti, come in figura 2.16a (per semplicità prendiamo il secondo degli esempi usati in figura 2.15), e supponiamo di volerci aggiungere un sito occupato e trasformarlo quindi in un cluster a s+1 siti: per farlo bisogna quindi prendere un sito con numero di coordinazione 1 o 2 ed aggiungere una branca a partire da esso, come mostrato in figura 2.16b (ove i siti modificati sono evidenziati in rosso). Nell'aggiungere questo sito s è aumentato di una unità, così come il numero di siti sulla frontiera. Pertanto la frontiera del cluster continua ad essere composta da s+2 siti, per ogni valore di s.

Il problema è che nella (2.2) non abbiamo determinato il coefficiente numerico, ossia il numero di modi in cui un dato sito può essere compreso in un *s*-cluster, che può essere molto complicato da determinare per *s* non piccoli. Questo problema può però essere risolto sfruttando ancora una volta la struttura gerarchica del reticolo.

Definiamo quindi come T(p) la grandezza media dei clusters che includono il vertice apicale di una branca. Allora (facendo riferimento a una situazione simile a quella rappresentata in figura 2.12):

$$T = p(1 + 2T)$$

perché il sito apicale dev'essere occupato, e anche le sue sottobranche hanno grandezza media T. Risolvendo:

$$T(p) = \frac{p}{1 - 2p}$$

che, comprensibilmente, diverge per  $p \to p_c^-$  (con  $p_c = 1/2$ ). Se S(p) è la grandezza media dei cluster che includono il sito di

Se S(p) è la grandezza media dei cluster che includono il sito di origine, si ha:

$$S(p) = 1 + 3T = 1 + \frac{3p}{1 - 2p} = \frac{1 + p}{1 - 2p} \stackrel{p \approx p_c}{\sim} \frac{1}{p_c - p}$$

(dal quale si ha  $\gamma = 1$ ). Sappiamo poi che la S(p) può anche essere scritta come:

$$S(p) = \sum_{\vec{r}} \Gamma(\vec{r})$$

ove per  $p < p_c$  la  $\Gamma(\vec{r})$  è la probabilità che un sito a distanza  $\vec{r}$  da uno occupato appartenga al suo stesso cluster; verifichiamolo.

La connessione fra l'origine del sistema di riferimento e  $\vec{r}$  è garantita da un unico cammino, e quindi tutti i siti del cammino stesso devono essere sicuramente occupati. Allora  $\Gamma(\vec{r}) = p^{|\vec{r}|}$ , ove  $|\vec{r}|$  è la "distanza" dall'origine calcolata come "numero di passi" del cammino (è detta anche "distanza chimica").

Pertanto:

$$S(p) = \sum_{\vec{r}} p^{|\vec{r}|} = 1 + 3p + 3 \cdot 2p^2 + 3 \cdot 2^2 p^3 + \dots = 1 + 3p[1 + 2p + (2p)^2 + \dots] = 1 + 3p \cdot \frac{1}{1 - 2p} = \frac{1 + p}{1 - 2p}$$

che è esattamente lo stesso risultato di prima.

Vogliamo ora cercare di capire qualcosa di più sull'andamento di  $n_s(p)$  per  $p \sim p_c$ ; per farlo, consideriamo la quantità  $n_s(p)/n_s(p_c)$ :

$$\frac{n_s(p)}{n_s(p_c)} = \left(\frac{p}{p_c}\right)^s \left(\frac{1-p}{1-p_c}\right)^2 \left(\frac{1-p}{1-p_c}\right)^s = \left(\frac{1-p}{1-p_c}\right)^2 \left(\frac{p}{p_c} \cdot \frac{1-p}{1-p_c}\right)^s$$

(ove abbiamo sfruttato la (2.2)). Il primo termine è del tutto innocuo, e non presenta anomalie per  $p \sim p_c$ . Il secondo termine, invece, vogliamo riscriverlo come  $e^{-c(p)s}$ , e pertanto:

$$c(p) = -\ln\left(\frac{p}{p_c} \cdot \frac{1-p}{1-p_c}\right)$$

Sviluppando attorno a  $p=p_c$  risulta:

$$c(p) \propto (p - p_c)^2$$
  $\Rightarrow$   $\frac{n_s(p)}{n_s(p_c)} \sim e^{-(p - p_c)^2 s}$ 

A questo punto facciamo un'"ipotesi educata": essendo  $p_c$  un punto critico del nostro sistema possiamo supporre che in prossimità di esso si abbia<sup>9</sup>:

$$n_s(p) \sim s^{-\tau} e^{-c(p)s}$$

Cerchiamo di determinare  $\tau$ ; per farlo, teniamo conto del fatto che  $S(p) \sim (p_c - p)^{-1}$  e che:

$$S(p) \propto \sum_{s} s^2 n_s(p) \sim \sum_{s=1}^{\infty} s^{2-\tau} e^{-c(p)s}$$

Vogliamo capire come questa somma diverge; per farlo (è un ragionemanto poco rigoroso, ma comunque efficace) possiamo "trasformarla" in un integrale, perché gli eventuali contributi singolari provengono dai termini con s grande, e per s molto grande la somma può essere approssimata con un integrale (la differenza fra i punti su cui si somma è trascurabile). Si ha quindi:

$$\int ds s^{2-\tau} e^{-c(p)s} = c(p)^{-(3-\tau)} \cdot \underbrace{\int dx e^{-x} x^{2-\tau}}_{\text{o un purpose}} \sim |p - p_c|^{-2(3-\tau)}$$

ove al primo passaggio abbiamo effettuato il cambio di variabile x=c(p)s. Poiché sappiamo che S(p) deve andare come  $(p-p_c)^{-1}$ , si ha  $\tau=5/2$ . Insomma, per  $p\sim p_c$  e grandi s si ha:

$$n_s(p) \sim s^{-5/2} e^{-|p-p_c|^2 s}$$

$$n_s(p) \sim s^{-\tau} f\left(|p - p_c|^{1/\sigma} s\right)$$

ove f è una funzione che si annulla per grandi argomenti; nel nostro caso,  $f(x) = e^{-x}$  e  $\sigma = 1/2$ .

<sup>&</sup>lt;sup>9</sup>In generale si può supporre:

Si tratta di una legge di scaling, con la quale (come tutte le leggi di questo tipo) possiamo collegare tutti gli esponenti critici (ossia, esprimere  $\gamma$  e  $\beta$  in funzione di  $\sigma$  e  $\tau$ ). Pertanto, determinati  $\sigma$  e  $\tau$  si può "risolvere" il problema della percolazione su questo reticolo (cosa in genere non fattibile con altri tipi di reticolo).

### 2.4 Percolazione e geometria frattale

Il problema della percolazione è strettamente collegato a problemi di natura geometrica che coinvolgono frattali, e anche storicamente è stato uno dei primi esempi in cui ci si è resi conto che le questioni geometriche entrano in gioco in fenomeni critici. In particolare, come vediamo adesso, esiste un collegamento fra la geometria dei clusters e gli esponenti critici del sistema.

Supponiamo di essere a  $p \sim p_c^-$ ; il sistema presenta allora clusters molto grandi. Dallo studio dei fenomeni critici sappiamo che se "fotografiamo" una particolare configurazione del sistema e la osserviamo a "risoluzioni" diverse l'aspetto del sistema è sempre lo stesso (è l'invarianza di scala tipica dei fenomeni critici). Nel caso della percolazione, questo suggerisce che se il cluster ha una dimensione lineare R, allora ci possiamo aspettare che il numero di siti in esso contenuti aumenti come  $R^{\overline{d}}$ , con  $\overline{d}$  possibilmente diversa dalla dimensione d del reticolo. La  $\overline{d}$  è un esempio di dimensione frattale<sup>10</sup>.

Vediamo un modo molto semplice per determinare la dimensione frattale di un cluster percolativo.

Consideriamo un sistema in d dimensioni e supponiamo che  $p\gtrsim p_c$ . Sappiamo che allora nel sistema c'è un cluster infinito e molti cluster finiti di grandi dimensioni; ci aspettiamo allora che non ci siano grandi differenze fra uno di questi cluster finiti e una porzione di quello infinito.

Per stimare la P(p) possiamo sfruttare la densità del cluster infinito, espressa in termini di rapporto fra siti appartenenti al cluster e siti totali del reticolo. Questa ci aspettiamo che sarà la stessa anche per i clusters finiti; ma la densità dei clusters finiti è facilmente calcolabile usando la loro dimensione lineare  $\xi$ . Si avrà quindi<sup>11</sup>:

$$P(p) \simeq \xi(p)^{\overline{d}-d}$$
  $\Rightarrow$   $|p-p_c|^{\beta} \simeq |p-p_c|^{-\nu(\overline{d}-d)}$ 

Pertanto:

$$\overline{d} = d - \frac{\beta}{\nu}$$

Si può anche mostrare (noi lo diciamo e basta) che  $1/\nu \coloneqq \overline{d}'$  ha un preciso significato. Supponiamo di osservare un cluster connesso in soglia di percolazione; questo può essere collegato ad un altro cluster da una "catena" di siti (nel senso che "tagliando" un legame i due clusters risultano separati), detti in generale  $\mathit{red}$  bonds. Si ha allora che  $1/\nu \coloneqq \overline{d}'$  è la dimensione frattale dei red bonds del sistema.

 $<sup>^{10}</sup>$ Un esempio "classico" di frattale è il *triangolo di Sierpiński*. Questo è invariante per cambio di scala, e per costruzione è tale per cui raddoppiandone le dimensioni la "materia" in esso contenuta (ossia il numero di siti, se lo pensiamo come un reticolo) triplica; la sua dimensione è quindi  $\overline{d} = \ln 3 / \ln 2$ .

<sup>&</sup>lt;sup>11</sup>In altre parole:  $\xi(p)^{\overline{d}}$  è il numero di siti presenti in un cluster percolativo, mentre  $\xi(p)^d$  è il volume del clusters.

### 2.5 Percolazione e leggi di scala

Vediamo come alcuni concetti e idee di base che abbiamo visto studiando i fenomeni critici si possano applicare al contesto della percolazione.

Studiando il modello di Ising abbiamo visto che l'energia libera per sito si può "scindere" in un contributo regolare e uno singolare  $f_s$ , che soddisfa lo scaling di Kadanoff:

$$f_s(K - K_c, h) = \ell^{-d} f_s \left( \ell^{Y_T} (K - K_c), \ell^{Y_H} h \right)$$

Ora, la percolazione non è un problema termico, quindi non sappiamo come definire un'energia libera associata al sistema (a priori non ha neanche senso definirla); possiamo però cercare di definire una quantità che ne sia un analogo. Ad esempio possiamo introdurre un campo ausiliario h (è privo di senso fisico, ma matematicamente utile) e definire la funzione:

$$f(p,h) = \sum_{s} n_s(p)e^{-hs}$$

Da notare che se  $h \neq 0$  l'esponenziale "uccide" il comportamento singolare di  $n_s$  per grandi s: è un comportamento analogo a quello che si ha in modelli come quello di Ising, ove la presenza di un campo esterno  $h \neq 0$  rimuove la singolarità del punto critico. Possiamo notare che:

$$-\frac{\partial f}{\partial h}_{|h=0} = \sum_{s} s n_s(p)$$

che è il "blocco fondamentale" che definisce P(p). Derivando una seconda volta:

$$\frac{\partial^2 f}{\partial h^2}|_{h=0} = \sum_s s^2 n_s(p)$$

che, a meno di un fattore di normalizzazione, è  $\chi$ .

Insomma, la funzione f contiene nelle sue derivate oggetti che ci interessano.

Per capire come possiamo usarla nel contesto della percolazione, dobbiamo comprendere cosa significhi effettuare una trasformazione di rinormalizzazione per un problema del genere.

Nel caso del modello di Ising sappiamo che si arriva allo scaling di Kadanoff a partire da una trasformazione  $K'_{\ell}(K,h)$ ,  $h'_{\ell}(K,h)$  delle costanti di accoppiamento, che altro non è che una procedura di coarse graining su domini di dimensione lineare  $\ell$ ; se chiamiamo  $F_{N,\text{sing}}$  la parte singolare dell'energia libera di un sistema finito a N siti ciò che si assume è che la nostra trasformazione la preservi in forma, ossia che valga:

$$F_{N,\mathrm{sing}}(K,h) = F_{N',\mathrm{sing}}(K'_\ell,h'_\ell)$$

ove  $N' = N/\ell^d$  è il nuovo numero di siti dopo la trasformazione di coarse graining. Dividendo per N ambo i membri e prendendo il limite termodinamico si ha:

$$\frac{1}{N}F_{N,\mathrm{sing}}(K,h) = \frac{1}{N}F_{N',\mathrm{sing}}(K'_{\ell},h'_{\ell}) \qquad \Rightarrow \qquad f_{\mathrm{sing}}(K,h) = \ell^{-d}f_{\mathrm{sing}}(K'_{\ell},h'_{\ell})$$

che è proprio la legge di scaling di Kadanoff.

In aggiunta a questo anche la lunghezza di correlazione  $\xi$  cambia una volta applicata la trasformazione:

$$\xi(K,h) = \ell^d \xi(K'_\ell, h'_\ell)$$

Riconsideriamo ora il nostro problema di percolazione. Ricordando che  $N_S(p)$  è il numero totale di clusters a S siti data la probabilità p che un generico sito sia occupato, in base a quanto abbiamo visto prima possiamo definire l'analogo dell'energia libera singolare come:

$$F_{N,\text{sing}} = \sum_{S} N_S(p) e^{-hS}$$

In analogia a quello che abbiamo detto per il caso del modello di Ising, la trasformazione nel caso della percolazione sarà del tipo  $(p,h) \to (p',h')$ , e dovrà ad esempio essere tale che  $\xi(p',h')=\ell^{-1}\xi(p,h)$ . Se effettuiamo il coarse graining il numero di siti contenuti in un cluster trasformerà come  $S\to S/\ell^{\overline{d}}$ ; ciò che richiediamo stavolta è che nell'effettuare il coarse graining di grandi clusters si abbia:

$$N_S(p) = N_{S/\ell^{\overline{d}}}(p')$$

ossia che  $N_S$  non cambi in forma per coarse graining *identici*, che sono quelli tali per cui clusters precedentemente separati *non* si uniscono fra loro. Con questa richiesta si ha:

$$\sum_{S} N_{S}(p) e^{-hS} = \sum_{S} N_{S/\ell^{\overline{d}}}(p') e^{-hS} = \sum_{S} N_{S/\ell^{\overline{d}}}(p') e^{-h\ell^{\overline{d}} \frac{S}{\ell^{\overline{d}}}} = \sum_{S'} N_{S'}(p') e^{-h\ell^{\overline{d}}S'}$$

ove nell'ultimo passaggio abbiamo definito  $S'=S/\ell^{\overline{d}}.$ 

La nostra "energia libera" è quindi invariante se ridefiniamo  $h' = h\ell^{\overline{d}}$ . Dividendo poi ambo i membri per N, come prima, si trova che nel limite termodinamico:

$$f_{\text{sing}}(p,h) = \ell^{-d} f_{\text{sing}}(p',h')$$

Consideriamo il caso semplice di percolazione in una dimensione, con  $\ell=2$ ; vogliamo quindi trovare una p' tale per cui  $\xi(p')=\xi(p)/2$ . Questo può accadere se  $p'=p^2$ : p' deve essere infatti la probabilità che il nuovo blocco sia occupato, e dato che bisogna garantire che la connettività di un cluster sia mantenuta (altrimenti il coarse graining non è più identico) allora entrambi i siti del bocco devono essere occupati, cosa che avviene appunto con probabilità  $p^2$ .

Ci chiediamo ora: esistono punti fissi della nostra trasformazione, ossia esistono dei  $p^*$  tali che  $p^* = (p^*)^2$ ? Sì, e sono banali:  $p^* = 0$  e  $p^* = 1$ .

Il punto fisso  $p^*=0$  è quello verso cui porta l'iterazione del coarse graining se p<1, perché ad ogni iterazione la  $\xi$  diminuisce fino ad annullarsi. Quello a  $p^*=1$  è invece relativo alla "criticalità limite" che abbiamo visto la percolazione ha per p=1. In corrispondenza di esso si ha<sup>12</sup>:

$$\frac{\partial p'}{\partial p}_{|p^*=1} = 2p^* = 2 = 2^{1/\nu}$$

$$\xi(p' - p_c) = \frac{1}{\ell}\xi(p - p_c)$$

e poiché  $\xi(p-p_c) \sim |p-p_c|^{1/\nu}$  allora:

$$\frac{1}{|p'-p_c|^\nu} = \frac{1}{\ell} \cdot \frac{1}{|p-p_c|^\nu}$$

 $<sup>^{12}</sup>$ L'origine dell'ultimo passaggio, in particolare la presenza dell' $1/\nu$ , deriva dalle proprietà generali dello scaling. Infatti, in generale si ha:

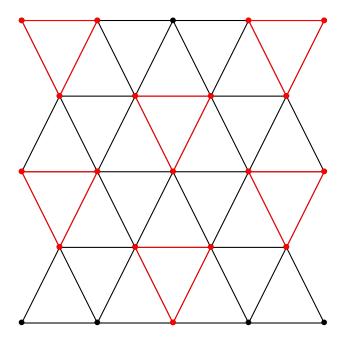


Figura 2.17: Coarse graining su reticolo triangolare

e quindi da questo possiamo dedurre che  $\nu = 1$ .

Vediamo ora una situazione un po' più complessa: consideriamo percolazione di sito su un reticolo triangolare, sul quale effettuiamo una trasformazione di coarse graining a tre siti, come illustrato in figura 2.17. Il nuovo reticolo è ancora triangolare, e si avrà:

$$\xi(p') = \frac{1}{\sqrt{3}}\xi(p)$$

Stavolta i blocchi sono sicuramente occupati se tutti e tre i loro siti lo sono, ma rispetto al caso unidimensionale la connettività può essere comunque garantita anche se i blocchi non sono completamente pieni. Insomma, i possibili casi che si possono presentare, con le rispettive probabilità, sono:

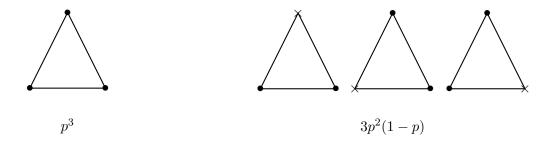
Poiché però in generale si ha:

$$\frac{p' - p_c}{p - p_c} = \frac{\partial p'}{\partial p}_{|p_c}$$

allora:

$$\frac{1}{\left|\frac{\partial p'}{\partial p}_{|p_c}(p-p_c)\right|^{\nu}} = \frac{1}{\ell} \cdot \frac{1}{|p-p_c|^{\nu}} \qquad \Rightarrow \qquad \frac{\partial p'}{\partial p}_{|p_c} = \ell^{1/\nu}$$

(e nel nostro caso particolare  $\ell = 2$ ).



Teoricamente potremmo considerare anche altri casi, ma supponiamo che gli unici possibili siano questi.

La trasformazione sarà allora:

$$p' = p^3 + 3p^2(1-p)$$

e i suoi punti fissi sono  $p^*=0$  e  $p^*=1/2$ ; quest'ultimo è esattamente uguale al risultato noto per il reticolo triangolare (non è un risultato *esatto*, ma comunque ottenibile con argomenti sensati; questa coincidenza è del tutto fortuita). Ciò che non torna rispetto a quello che ci si aspetta è il valore di  $\nu$  (andando a calcolarlo in base a  $\partial p'/\partial p_{|p_c=1/2}=(\sqrt{3})^{1/\nu}$  l'esponente  $\nu$  non concorda con quello osservato nelle simulazioni).

#### 2.6 Percolazione e modello di Potts

Il problema della percolazione in realtà può essere messo in relazione ad altri modelli "classici" della fisica statistica, come il modello di Potts. Quello che vogliamo vedere ora è come questi due problemi apparentemente molto diversi siano collegati.

Come noto, il modello di Potts è un modello di reticolo nel quale ad ogni sito viene associata una variabile  $\sigma=1,2,\ldots,q$ ; considerando un modello di Potts su reticolo quadrato, la sua hamiltoniana ridotta sarà:

$$-\beta \mathcal{H} = K \sum_{\langle ij \rangle} \delta_{\sigma_i \sigma_j} + h \sum_i \delta_{\sigma_i, 1}$$

ove la seconda somma favorisce le configurazioni con  $\sigma=1$  (ma potevamo scegliere qualunque altro valore).

Se q=2 il modello è chiaramente equivalente a un modello di Ising, mentre per q>2 ne costituisce una generalizzazione invariante sotto permutazioni (anziché sotto parità).

Quello che vogliamo mostrare è che considerando la "continuazione analitica" del modello (ossia permettendo a q di avere valori reali) e detta  $Q_N$  la funzione di partizione del sistema, allora:

$$\frac{\partial}{\partial q} \left( \frac{1}{N} \ln Q_N \right)_{|q=1} = f\left( p(h, K), h^{\mathsf{p}}(h, K) \right) \tag{2.3}$$

ove  $h^p$  è il campo ausiliario che abbiamo introdotto per la percolazione, e f è la densità di "energia libera" del sistema; in questo modo possiamo anche capire quale sia la sua origine, e ne possiamo vedere gli esponenti critici come un "caso limite" di quelli del modello di Potts.

Vediamo quindi come mostrare la validità della (2.3).

Sappiamo che nell'espansione in cluster per i fluidi reali si definiscono le *funzioni di Mayer* come:

$$f_{ij} = e^{-\beta v_{ij}} - 1$$

ove  $v_{ij}$  è il potenziale d'interazione fra le particelle i e j, e la funzione di partizione può essere scritta come:

$$Z_N = \frac{1}{\Lambda^{3N}} \int d\vec{r_i} \cdots d\vec{r_N} \prod_{i < j} (1 + f_{ij})$$

e ogni contributo può essere interpretato graficamente con un diagramma. Cose analoghe avvengono anche per i modelli di spin; la funzione di partizione del modello di Potts è infatti:

$$Q_N = \sum_{\{\sigma\}} \prod_{\langle ij \rangle} e^{K\delta_{\sigma_i \sigma_j}} \prod_i e^{h\delta_{\sigma_i, 1}}$$

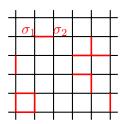
e usiamo il "trucco" di scrivere:

$$e^{K\delta_{\sigma_i\sigma_j}} = \underbrace{(e^K - 1)}_{:=u(K)} \delta_{\sigma_i\sigma_j} + 1$$

(considerando i casi i=j e  $i\neq j$  si verifica subito che entrambi i membri coincidono). Riscriviamo allora  $Q_N$  come:

$$Q_N = \sum_{\{\sigma\}} \prod_{\langle ij \rangle} (u(K)\delta_{\sigma_i \sigma_j} + 1) \prod_i e^{h\delta_{\sigma_i,1}}$$

Sviluppando la prima produttoria si generano termini che diagrammaticamente possono essere interpretati come tutte le possibilità di associare o meno a un legame del reticolo la  $u^{13}$ . Possiamo ad esempio rappresentare la situazione graficamente nel seguente modo:



Ogni possibile configurazione dei legami, quindi, è rappresentata da un diagramma G nello sviluppo della  $Q_N$ . Possiamo allora scrivere:

$$Q_N = \sum_{G} \sum_{\{\sigma\}} u(K)^{B(G)} \delta_{\sigma_{i_1} \sigma_{i_2}} \cdots \delta_{\sigma_{B(G)-1} \sigma_{B(G)}} \prod_{i} e^{h \delta_{\sigma_i, 1}}$$

ove B(G) è il numero dei legami che compaiono in G. Le  $\delta$  poi vincolano i siti legati fra loro ad avere la stessa  $\sigma$ : si tratta di una sorta di cluster percolativo.

 $<sup>^{13}</sup>$ In generale, sviluppando la produttoria otterremo termini del tipo  $1 + u\delta_{\sigma_1\sigma_2} + u\delta_{\sigma_2\sigma_3} + \cdots + u^2\delta_{\sigma_1\sigma_2}\delta_{\sigma_3\sigma_4} + \cdots$  ecc. Il primo termine, ad esempio, corrisponde al non associare nessun termine ai legami del reticolo, mentre il primo corrisponde ad associare un u al legame fra  $\sigma_1$  e  $\sigma_2$ ; analogamente, il termine  $u^2\delta_{\sigma_1\sigma_2}\delta_{\sigma_3\sigma_4}$  corrisponde all'associare un u al legame  $\sigma_1$ - $\sigma_2$  e uno a  $\sigma_3$ - $\sigma_4$ , e così via per tutti i termini successivi.

Una volta scelto G, poi, c'è una restrizione sulle possibili  $\sigma$  dei vari siti. Considerando quelli che nella figura precedente sono  $\sigma_1$  e  $\sigma_2$ , ad esempio, si ha:

$$\sum_{\sigma_1, \sigma_2} e^{h\delta_{\sigma_1, 1}} e^{h\delta_{\sigma_2, 1}} \delta_{\sigma_1, \sigma_2} = e^{2h} + (q - 1)$$

ove il primo termine è relativo al caso  $\sigma_1 = \sigma_2 = 1$ , mentre il secondo a  $\sigma_1 = \sigma_2 \neq 1$  (in questo caso il sommando è pari ad 1, e ci sono q-1 sommandi). Pertanto, detto C un cluster e  $n_C$  il numero dei siti che contiene:

$$Q_N = \sum_G u(K)^{B(G)} \prod_{C \in G} \left( e^{hn_C} + q - 1 \right)$$

Abbiamo quindi trasformato la funzione di partizione in un oggetto "geometrico". Notiamo anche che adesso q è un semplice parametro, che può assumere i valori che vogliamo.

Vogliamo a questo punto mostrare (teorema di Fortuin-Kasteleyn) che l'"energia libera" per la percolazione di legame è data da:

$$f = \frac{1}{N} \frac{\partial}{\partial q} \ln Q_N_{|q=1}$$

Calcoliamo innanzitutto:

$$\frac{1}{N} \frac{\partial}{\partial q} \ln Q_N \Big|_{q=1} = \frac{1}{N} \cdot \left( \frac{1}{Q_N} \frac{\partial Q_N}{\partial q} \right)_{|q=1} = \frac{1}{N} \cdot \frac{\sum_G u(K)^{B(G)} \sum_{C \in G} \prod_{C' \in G, C' \neq C} e^{hn_{C'}}}{\sum_G u(K)^{B(G)} \prod_{C \in G} e^{hn_{C}}}$$

Pertanto, per ogni diagramma G andiamo a considerare tutti i clusters che ci appartengono, e per ognuno di essi andiamo a calcolare la produttoria che compare a numeratore. Ora, in percolazione di legame qualunque sia il diagramma G che scegliamo si ha:

$$\sum_{C \in G} n_C = N$$

ove N è il numero totale di siti del reticolo<sup>14</sup>. Alla luce di questo la produttoria presente a denominatore può essere riscritta come:

$$\prod_{C \in G} e^{hn_C} = e^{h\sum_{C \in G} n_C} = e^{hN}$$

Quest'osservazione ci permette anche di semplificare l'espressione a numeratore:

$$\prod_{\substack{C' \in G \\ C' \neq C}} e^{hn_{C'}} = \exp\left(h \sum_{\substack{C' \in G \\ C' \neq C}} n_{C'} + hn_C - hn_C\right) = e^{hN - hn_C}$$

<sup>&</sup>lt;sup>14</sup>Ricordare che in percolazione di legame *ogni* sito, anche se non collegato ad altri primi vicini, costituisce un cluster.

Se quindi chiamiamo M il numero totale di legami fra primi vicini, moltiplicando a numeratore e denominatore per  $e^{-MK}$  si ha<sup>15</sup>:

$$\begin{split} \frac{1}{N} \cdot \frac{\partial}{\partial q} \ln Q_N &= \frac{1}{N} \cdot \frac{\sum_G u(K)^{B(G)} e^{-MK} \sum_{C \in G} e^{-hn_C}}{\sum_G u(K)^{B(G)} e^{-MK}} = \\ &= \frac{1}{N} \cdot \frac{\sum_G \left(\frac{u(K)}{e^K}\right)^{B(G)} \left(e^{-K}\right)^{M-B(G)} \sum_{C \in G} e^{-hn_C}}{\sum_G \left(\frac{u(K)}{e^K}\right)^{B(G)} \left(e^{-K}\right)^{M-B(G)}} = \\ &= \frac{1}{N} \sum_G p^{B(G)} (1-p)^{M-B(G)} \sum_{C \in G} e^{-hn_C} \end{split}$$

ove nell'ultima riga abbiamo definito:

$$p := \frac{u(K)}{e^K} = \frac{e^K - 1}{e^K} = 1 - e^{-K}$$

(di modo anche che  $e^{-K}=1-p$ ). Fissato ora un diagramma G, ossia una data configurazione dei clusters, possiamo anche scrivere:

$$\sum_{C \in G} e^{-hn_C} = \sum_{s} \sum_{C \in G} e^{-hs} \delta_{s,n_C}$$

ove  $s=1,2,\ldots$  è il numero di siti che possono appartenere a un dato cluster. Pertanto:

$$\frac{1}{N} \frac{\partial}{\partial q} \ln Q_N \Big|_{q=1} = \sum_{s=1}^{\infty} \frac{e^{-hs}}{N} \sum_G p^{B(G)} (1-p)^{M-B(G)} \sum_{C \in G} \delta_{s,n_C}$$

Tuttavia:

$$N_S(p) = \sum_G p^{B(G)} (1-p)^{M-B(G)} \sum_{C \in G} \delta_{s,n_C}$$

(presa la configurazione di un diagramma, andiamo a sommare su ogni cluster il numero di siti in esso contenuto; alla fine facciamo la media pesata di questo numero di siti, pesata con la probabilità del diagramma, e alla fine otteniamo quindi effettivamente il numero medio di cluster a S siti) e poiché  $N_S(p)/N = n_S(p)$ , alla fine si ha proprio:

$$\frac{1}{N} \frac{\partial}{\partial q} \ln Q_N = \sum_s n_s(p) e^{-hs}$$

che è la forma dell'"energia libera" della percolazione di legame.

 $<sup>^{15}</sup>$ Nell'ultima riga il denominatore è uguale ad 1 perché  $p^{B(G)}(1-p)^{M-B(G)}$  è la probabilità di un dato diagramma, e sommando su tutte le possibili configurazioni dei diagrammi si ottiene ovviamente 1.

## Capitolo 3

# Meccanica statistica di non equilibrio

#### 3.1 Introduzione

La meccanica statistica, come sappiamo, si basa sull'assegnazione di una densità di probabilità  $\rho(q,p)$  nello spazio delle fasi  $\Gamma$  del sistema, attraverso la quale determinarne le proprietà che ci interessano. Se ad esempio siamo interessati a studiare un'osservabile descritta dalla funzione A(q,p), l'evoluzione del sistema è data da:

$$\frac{d}{dt}A(q,p) = \{A, H\}_{\rho} := \mathcal{L}A(q,p)$$

ove H è l'hamiltoniana del sistema, e  $\mathcal{L} = \{\bullet, H\}_{\rho}$  è l'operatore di Liouville. La soluzione formale di quest'equazione è:

$$A(t) = A(q_t, p_t) = e^{\mathcal{L}t} A(q, p)$$

Il valor medio di questa funzione sarà dato da:

$$\langle A(t) \rangle = \int dq dp \rho(q, p) A(q_t, p_t)$$

Per determinare quindi il valor medio di una funzione nel tempo si adotta un'ottica simile a quella della visuale di Heisenberg della meccanica quantistica: lo stato del sistema (la  $\rho$ ) resta inalterato mentre la funzione A evolve.

Quest'impostazione è assolutamente generale: a seconda della forma della  $\rho$ , infatti, il nostro sistema potrà essere all'equilibrio o fuori da esso (e sappiamo dalla meccanica statistica standard quali devono essere le forme della  $\rho$  affinché il sistema sia all'equilibrio).

Possiamo però pensare di invertire questo punto di vista e adottarne uno più simile alla visuale di Schrödinger della meccanica quantistica, ossia anziché tenere fissa  $\rho$  al suo valore iniziale possiamo farla evolvere nel tempo e tenere fisse le osservabili. In particolare, se chiamiamo  $\rho(q,p) := \rho(q,p;0)$  la densità di probabilità nello spazio delle fasi all'istante iniziale, richiediamo che valga  $\rho(q,p;0) = \rho(q_t,p_t;t)$  (il valore della  $\rho$  in  $(q_t,p_t)$  all'istante t è lo stesso della  $\rho$  in (q,p) se  $(q_t,p_t)$  è l'evoluto all'istante t di (q,p)). In questo modo, ad esempio, la normalizzazione della  $\rho$  è garantita ad ogni istante; se infatti vale:

$$\int \rho(q, p; 0) dq dp = 1$$

allora sfruttando la richiesta che facciamo e il teorema di Liouville ( $dqdp = dq_tdp_t$ ) si ha anche:

$$\int \rho(q_t, p_t; t) dq_t dp_t = 1$$

In questo caso, quindi, l'ensemble del sistema evolve nel tempo; l'equazione che regola quest'evoluzione, ossia l'equazione che regola l'andamento di  $\rho$  nel tempo, è l'equazione di Liouville. Per determinarla deriviamo rispetto a t l'ipotesi  $\rho(q, p; 0) = \rho(q_t, p_t; t)$ :

$$0 = \frac{\partial \rho}{\partial t} + \{\rho, H\} \qquad \Rightarrow \qquad \frac{\partial \rho}{\partial t} = -\mathcal{L}\rho$$

In questo modo possiamo scrivere il valor medio della generica funzione *A* nella "visuale di Schrödinger":

$$\langle A(t) \rangle = \int dq dp \rho(q, p; 0) A(q_t, p_t) = \int dq_t dp_t \rho(q_t, p_t; t) A(q_t, p_t) = \int dq dp \rho(q, p; t) A(q, p)$$

ove nell'ultimo passaggio abbiamo rinominato le variabili d'integrazione.

Sempre in quest'ottica, poi, la densità di probabilità  $\rho$  sarà d'equilibrio se non dipende dal tempo, ossia:

$$\frac{\partial}{\partial t}\rho_{\rm eq}(q,p;t) = 0$$

Ma questo implica che  $\{\rho, H\} = 0$ , ossia che  $\rho$  dev'essere una funzione dell'hamiltoniana<sup>1</sup>.

#### 3.1.1 La legge bilancio dettagliato

All'equilibrio, poi, vale anche la *legge di bilancio dettagliato*. Questa asserisce che, se Y(q,p) è un'osservabile che ci interessa, allora la densità di probabilità congiunta  $P^{eq}(Y=y_1,t;Y=y_0,0)$  che Y assuma il valore  $y_0$  al tempo t=0 e il valore  $y_1$  all'istante t è tale per cui:

$$P^{\text{eq}}(y_1, t; y_0, 0) = P^{\text{eq}}(y_0, t; y_1, 0)$$
(3.1)

Si tratta di una proprietà *fondamentale* in processi di equilibrio, e come vedremo discende direttamente dalla reversibilità microscopica. È anche uno degli strumenti fondamentali per capire cosa accade in sistemi che si trovano leggermente fuori dall'equilibrio.

Per mostrare la validità della (3.1) consideriamo un sistema la cui hamiltoniana è invariante per inversione dei momenti, ossia H(q,p) = H(q,-p) (è un'ipotesi lecitissima, visto che generalmente i momenti compaiono solo nel termine cinetico, che è quadratico nelle p), e supponiamo che lo sia anche l'osservabile Y che ci interessa: in questo caso allora il sistema è invariante per inversione temporale<sup>2</sup>. Possiamo allora scrivere la densità di probabilità congiunta che ci interessa, in una situazione d'equilibrio:

$$P^{\text{eq}}(y_1, t; y_0, 0) = \int dq dp \rho_{\text{eq}}(q, p) \delta\left(Y(q, p) - y_0\right) \delta\left(Y(q_t, p_t) - y_1\right)$$

Sfruttando il teorema di Liouville e il fatto che  $\rho_{\rm eq}$  è una densità d'equilibrio possiamo scrivere  $dqdp\rho_{\rm eq}(q,p)=dq_tdp_t\rho_{\rm eq}(q_t,p_t)$ ; definendo poi  $p_t'=-p_t$ , sfruttando l'invarianza

<sup>&</sup>lt;sup>1</sup>Notare che tutte le densità di probabilità di equilibrio, dalla microcanonica alla grancanonica, sono effettivamente funzioni (eventualmente singolari, come la  $\delta$  di Dirac) dell'hamiltoniana.

<sup>&</sup>lt;sup>2</sup>Partendo infatti da (q,p) a t=0, dopo un tempo  $\tau$  il sistema si troverà in  $(q_{\tau},p_{\tau})$ ; se definiamo allora  $q'_t=q_{\tau-t}$  e  $p'_t=-p_{\tau-t}$  queste sono ancora una soluzione delle equazioni di Hamilton, e sono tali che dal tempo t=0 all'istante t evolvono da  $(q_{\tau},p'_{\tau})=(q_{\tau},-p_{\tau})$  a (q,-p).

per riflessione nei momenti della  $\rho_{eq}^3$  e della Y, questa densità di probabilità diventa uguale a:

$$\int dq_t dp_t' \rho_{\text{eq}}(q_t, p_t') \delta\left(Y(q, -p) - y_0\right) \delta\left(Y(q_t, p_t') - y_1\right)$$

Sfruttando ora l'invarianza per inversione temporale, il punto (q, -p) non è altro che l'evoluto al tempo t del nuovo punto iniziale  $(q_t, p_t')$ : pertanto quella che abbiamo appena scritto è proprio la probabilità congiunta di trovare  $Y = y_0$  al tempo t avendo  $Y = y_1$  al tempo t = 0, ossia:

$$\int dq_t dp'_t \rho_{\text{eq}}(q_t, p'_t) \delta(Y(q, -p) - y_0) \delta(Y(q_t, p'_t) - y_1) = P^{\text{eq}}(y_0, t; y_1, 0)$$

Vale quindi proprio la (3.1).

### 3.2 Reversibilità microscopica e irreversibilità macroscopica

#### 3.2.1 Entropia e probabilità condizionata

Cosa significa quindi osservare un sistema macroscopico in un dato stato (q,p)? In meccanica statistica, quando si studia un sistema macroscopico, siamo generalmente interessati a determinate osservabili estensive, ossia funzioni  $X_1(q,p),\ldots,X_{\nu}(q,p)$  che descrivono le proprietà che ci interessano.

All'infuori dei fenomeni critici, però, fin'ora non abbiamo mai visto le fluttuazioni di queste funzioni assumere un ruolo rilevante; Einstein fu il primo a introdurre un modo completamente originale per trattarle, che ora vediamo.

Supponiamo di avere un sistema termodinamico isolato, quindi descritto da una densità di probabilità microcanonica nello spazio delle fasi. La sua entropia è allora:

$$S(E) = k_B \ln \Gamma(E) = k_B \ln \left( \frac{1}{h^{3N} N!} \int_{E \le H(q, p) \le E + \Delta} dq dp \right)$$
(3.2)

Se però consideriamo un sistema descritto dalle variabili  $X_{\alpha}(q,p)$ , alle quali permettiamo di fluttuare attorno ai loro valori medi, possiamo dare un'espressione dell'entropia legata ai valori di queste variabili:

$$S(x_1, \dots, x_{\nu}) = k_B \ln \left[ \frac{1}{h^{3N} N!} \int_{E \le H(q, p) \le E + \Delta} dq dp \prod_{\alpha = 1}^{\nu} \delta \left( X_{\alpha}(q, p) - x_{\alpha} \right) \right]$$
(3.3)

ossia, all'interno della shell di energia nello spazio delle fasi "selezioniamo" solo i (q, p) per i quali le  $X_{\alpha}$  hanno i valori desiderati.

Qual è la relazione fra queste due entropie?

Prendendo il limite termodinamico (ossia considerando un sistema sempre più grande), la (3.3) tenderà ad avere un massimo sempre più pronunciato in corrispondenza di valori di equilibrio  $\vec{x}^{\text{eq}} = (x_1^{\text{eq}}, \dots, x_{\nu}^{\text{eq}})$  tali per cui la (3.3) calcolata in essi coincide con la (3.2).

<sup>&</sup>lt;sup>3</sup>Notare che essendo una densità di probabilità d'equilibrio è una funzione dell'hamiltoniana, che supponiamo essere invariante.

La (3.3), quindi, a differenza della (3.2), permette alle  $X_{\alpha}$  di fluttuare attorno ai loro valori di equilibrio  $\vec{x}^{\text{ eq}}$ ; poiché la probabilità che le variabili  $X_1, \ldots, X_{\nu}$  assumano i valori  $x_1, \ldots, x_{\nu}$  all'equilibrio è<sup>4</sup>:

$$P^{\text{eq}}(x_1, \dots, x_{\nu}) \propto e^{\frac{1}{k_B}S(x_1, \dots, x_{\nu})}$$

possiamo sviluppare S attorno a  $\vec{x}$  eq e studiare le proprietà del sistema quando questo è leggermente fuori dall'equilibrio. Scriviamo quindi:

$$S(x_1, \dots, x_{\nu}) = S(x_1^{\text{eq}}, \dots, x_{\nu}^{\text{eq}}) + \frac{1}{2} \sum_{\alpha, \beta} \frac{\partial^2 S}{\partial x_{\alpha} \partial x_{\beta}} \Big|_{\vec{x}^{\text{eq}}} (x_{\alpha} - x_{\alpha}^{\text{eq}}) (x_{\beta} - x_{\beta}^{\text{eq}}) + \dots$$

Se quindi trascuriamo gli ordini successivi, chiamiamo N' la costante di normalizzazione di  $P^{eq}$  e definiamo:

$$S_{\alpha\beta} := -\frac{1}{k_B} \cdot \frac{\partial^2 S}{\partial x_\alpha \partial x_\beta}_{|\vec{x}|^{\text{eq}}}$$

allora otteniamo:

$$P^{\text{eq}}(x_1,\ldots,x_{\nu}) = N'e^{-\frac{1}{2}\sum_{\alpha,\beta}S_{\alpha\beta}(x_{\alpha}-x_{\alpha}^{\text{eq}})(x_{\beta}-x_{\beta}^{\text{eq}})}$$

Se infine definiamo anche  $y_{\alpha} := x_{\alpha} - x_{\alpha}^{\text{eq}}$ :

$$P^{\text{eq}}(y_1, \dots, y_{\nu}) = N' e^{-\frac{1}{2} \sum_{\alpha, \beta} S_{\alpha\beta} y_{\alpha} y_{\beta}}$$

Cerchiamo ora di determinare il fattore di normalizzazione N'.

Innanzitutto, S è una matrice simmetrica e definita positiva e quindi è sicuramente diagonalizzabile; detta A la matrice che la diagonalizza e posto  $y'_{\alpha} = \sum_{\beta} A_{\alpha\beta} y_{\beta}$  si ha:

$$P^{\text{eq}}(y'_1, \dots, y'_{\nu}) = N' e^{-\frac{1}{2} \sum_{\alpha} S'_{\alpha\alpha} y'_{\alpha}^2}$$

ove S' è la diagonalizzazione di S. Pertanto, integrando su tutte le  $y'_{\alpha}$ :

$$1 = \int dy'_1 \cdots dy'_{\nu} P^{\text{eq}}(y'_1, \dots, y'_{\nu}) = \int dy'_1 \cdots dy'_{\nu} N' e^{-\frac{1}{2} \sum_{\alpha} S'_{\alpha\alpha} y'_{\alpha}^2} = N' \sqrt{\frac{(2\pi)^{\nu}}{\prod_{\alpha} S'_{\alpha\alpha}}}$$

Dato che S' è una matrice diagonale  $\prod_{\alpha} S'_{\alpha\alpha} = \det S' = \det S$ , e quindi  $N' = \sqrt{\det S}/(2\pi)^{\nu/2}$ . Pertanto, riesprimendo  $P^{\text{eq}}$  in funzione delle  $y_{\alpha}$ , si ha:

$$P^{\text{eq}}(y_1, \dots, y_{\nu}) = \frac{\sqrt{\det S}}{(2\pi)^{\nu/2}} e^{-\frac{1}{2} \sum_{\alpha, \beta} S_{\alpha\beta} y_{\alpha} y_{\beta}}$$

Quello che adesso vogliamo fare è usare la probabilità che abbiamo appena determinato per calcolare correlazioni di equilibrio, ossia valori medi del tipo  $\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\rm eq}$ . In particolare possiamo scrivere:

$$\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} = \int d^{\nu}y d^{\nu}y_0 P^{\text{eq}}(\vec{y}_0)y_{0\alpha}y_{\beta}P(\vec{y},t|\vec{y}_0,0)$$
(3.4)

 $<sup>^4</sup>$ Infatti, in generale il volume nello spazio delle fasi di un sistema con entropia S, come sappiamo, è  $e^{S/k_B}$ . La costante di proporzionalità è quella che serve a normalizzare la probabilità, che dopo determineremo.

Il problema di quest'espressione è che la probabilità condizionata  $P(\vec{y},t|\vec{y_0},0)$  non è nota, e soprattutto per determinarla bisognerebbe essere in grado di risolvere le equazioni del moto del sistema. Infatti:

$$P(\vec{y}, t | \vec{y_0}, 0) = \frac{P(\vec{y}, t; \vec{y_0}, 0)}{P^{\text{eq}}(\vec{y_0})} = \frac{\int_{E \le H \le E + \Delta} dq dp \prod_{\beta} \delta\left(Y_{\beta}(q_t, p_t) - y_{\beta}\right) \prod_{\gamma} \delta\left(Y_{\gamma}(q, p) - y_{0\gamma}\right)}{\int_{E \le H \le E + \Delta} dq dp \prod_{\alpha} \delta\left(Y_{\alpha}(q, p) - y_{0\alpha}\right)}$$
(3.5)

ma in generale non è possibile determinare  $Y_{\beta}(q_t, p_t)$ . A denominatore, ad esempio, l'integrale in q e p è ristretto alle  $y_0$ , ossia nella shell di energia  $E \leq H \leq E + \Delta$  selezioniamo l'insieme ulteriormente specificato dai valori delle  $y_0^5$ .

Notare che stiamo approcciando il problema da un punto di vista prettamente statistico: le  $Y_{\alpha}$  sono in generale variabili che si annullano all'equilibrio, ma che comunque fluttuano nel tempo. In termodinamica, invece (in particolare in quella dei processi reversibili), abbiamo a che fare con uno scenario più semplice: il sistema è descritto da variabili macroscopiche che evolvono nel tempo rilassando verso i propri valori di equilibrio, senza che entri mai in gioco il concetto di fluttuazione.

#### 3.2.2 L'ipotesi di regressione di Onsager

La (3.5) è una quantità che abbiamo determinato supponendo che il nostro sistema sia all'equilibrio; possiamo immaginare però che possa riferirsi anche a situazioni diverse, ad esempio situazioni fuori equilibrio. In questo caso, supponendo che a t=0 il sistema sia fuori equilibrio e che la densità di probabilità iniziale nello spazio delle fasi (che in generale non conosciamo) sia  $\rho(q,p;0)$ , la probabilità condizionata sui valori di  $\vec{y}$  diventa:

$$P^{\text{n.e.}}(\vec{y}, t | \vec{y_0}, 0) = \int_{E \le H \le E + \Delta} dq dp \prod_{\beta} \delta(Y_{\beta}(q_t, p_t) - y_{\beta}) \, \rho(q, p; 0)$$
 (3.6)

(ove "n.e." sta per "non equilibrio").

Pertanto, a seconda se inizialmente il sistema è o meno all'equilibrio l'espressione della probabilità condizionata può essere data dalla (3.5) o dalla (3.6). Tuttavia, possiamo immaginare che a t=0 il sistema sia in  $\vec{y}_0$  non per via di semplici fluttuazioni statistiche della Y, ma perché il sistema è stato perturbato e portato fuori dall'equilibrio. In questo secondo caso, ovviamente, bisogna considerare una scelta diversa della  $\rho$ .

Ora, in quali casi posso pensare che la  $\rho$  che scelgo sia effettivamente equivalente a quella che uso per determinare la (3.5)? L'idea è che se il sistema evolve dinamicamente, e in tempi  $\mathit{molto}$   $\mathit{più}$   $\mathit{rapidi}$  rispetto a quelli su cui variano le Y, il sistema va verso un  $\mathit{equilibrio}$   $\mathit{incompleto}$ , ove con quest'ultima espressione si intende che nella porzione della shell energetica  $E \leq H \leq E + \Delta$  compatibile col valore  $y_0$  di Y assumiamo ci sia microcanonicità; in altre parole, per "equilibrio incompleto" intendiamo che si ha "microcanonicità ristretta" nella porzione di spazio delle fasi disponibile dove Y assume il valore  $y_0$ .

Insomma, l'ipotesi di Onsager è che la (3.5) si possa identificare con la (3.6); con quest'ipotesi, quindi, non possiamo distinguere fra l'aver "trovato" il sistema in  $y_0$  per via di fluttuazioni statistiche o l'avercelo "trovato" perché è stato perturbato dall'esterno; questi due aspetti si conciliano appunto supponendo l'esistenza di questo "equilibrio incompleto".

 $<sup>^5</sup>$ In altre parole ancora, stiamo considerando il sottoinsieme dell'ensemble microcanonico del sistema dove le  $Y_{\alpha}$  hanno dei valori assegnati.

#### 3.2.3 La matrice dei coefficienti cinetici

A questo punto però la situazione è ancora non del tutto definita, perché non possiamo in generale deterinare  $Y_{\beta}(q_t, p_t)$ .

Riscriviamo la (3.4) come:

$$\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} = \int d^{\nu}y_{0}P^{\text{eq}}(\vec{y_{0}})y_{0\alpha}\underbrace{\int d^{\nu}yy_{\beta}P(\vec{y},t|\vec{y_{0}},0)}_{:=\langle y_{\beta}\rangle_{\vec{y_{0}}}}$$

ove abbiamo definito il valor medio condizionato  $\langle y_\beta \rangle_{\vec{y_0}}$ ; data l'identificazione che abbiamo appena analizzato, ci aspettiamo che questi valori medi condizionati "si comportino" come variabili termodinamiche, ossia ci aspettiamo che per essi valgano le stesse leggi fenomenolgiche termodinamiche che valgono per le  $Y_\alpha$ . Queste leggi sono in generale della forma<sup>6</sup>:

$$\frac{d}{dt} \langle y_{\beta}(t) \rangle_{\vec{y}_0} = -\sum_{\gamma} G_{\beta\gamma} \langle y_{\gamma}(t) \rangle_{\vec{y}_0}$$
(3.7)

ove G è una matrice (positiva) che, appunto, racchiude tutti gli aspetti fenomenologici sull'evoluzione del sistema.

La soluzione di quest'equazione è  $\langle \vec{y}(t) \rangle_{\vec{y_0}} = e^{-Gt} \, \langle \vec{y}(0) \rangle_{\vec{y_0}} = e^{-Gt} \, \vec{y_0}$ , e pertanto:

$$\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} = -\int d^{\nu}y_{0}y_{0\alpha}P^{\text{eq}}(\vec{y}_{0})\left(e^{-Gt}\vec{y}_{0}\right)_{\beta}$$

Considerando il limite  $t \to 0^+$  si ha:

$$\begin{aligned} \langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} &= \int d^{\nu}y_{0}y_{0\alpha}P^{\text{eq}}(\vec{y}_{0})\left[(\mathbb{1} - tG)\vec{y}_{0}\right]_{\beta} = \\ &= \int d^{\nu}y_{0}P^{\text{eq}}(\vec{y}_{0})y_{0\alpha}y_{0\beta} - t\int d^{\nu}y_{0}P^{\text{eq}}(\vec{y}_{0})y_{0\alpha}G_{\beta\gamma}y_{0\gamma} \end{aligned}$$

Anche se non è evidente, il primo integrale è uguale a  $S_{\alpha\beta}^{-1}$  (più avanti mostriamo che è così, e come in questo àmbito può essere usato il teorema di Wick), pertanto complessivamente:

$$\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} = S_{\alpha\beta}^{-1} - t\sum_{\gamma} G_{\beta\gamma}S_{\alpha\gamma}^{-1}$$
 (3.8)

Ora, in termodinamica si definiscono le "forze generalizzate" come:

$$F_{\alpha} = \frac{\partial S}{\partial y_{\alpha}|\vec{y}} = -\sum_{\beta} S_{\alpha\beta} y_{\beta}$$

e quindi  $\vec{y} = -S^{-1}\vec{F}$ ; possiamo allora riscrivere la (3.7) come:

$$\frac{d}{dt} \langle \vec{y}(t) \rangle_{\vec{y}_0} = GS^{-1} \left\langle \vec{F}(t) \right\rangle_{\vec{y}_0} \tag{3.9}$$

la matrice  $GS^{-1}$  è comunemente detta matrice dei coefficienti cinetici:

$$L := GS^{-1}$$

<sup>&</sup>lt;sup>6</sup>Da notare che è la presenza del segno negativo nell'equazione che permette al sistema di rilassare verso uno stato d'equilibrio.

Si tratta quindi della matrice che lega le derivate nel tempo dei valori medi condizionati alle forze generalizzate termodinamiche.

Poiché S è simmetrica lo è anche  $S^{-1}$ , e quindi la (3.8) si può riscrivere come:

$$\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} = S_{\alpha\beta}^{-1} - tL_{\beta\alpha}$$
 (3.10)

Se avessimo calcolato  $\langle y_\alpha(t)y_\beta(0)\rangle_{\rm eq}$  avremmo trovato lo stesso risultato, con  $\alpha$  e  $\beta$  invertiti:

$$\langle y_{\alpha}(t)y_{\beta}(0)\rangle_{\text{eq}} = S_{\beta\alpha}^{-1} - tL_{\alpha\beta}$$
 (3.11)

Sappiamo però anche che per la legge di bilancio dettagliato deve valere:

$$\langle y_{\alpha}(0)y_{\beta}(t)\rangle_{\text{eq}} = \langle y_{\beta}(0)y_{\alpha}(t)\rangle_{\text{eq}}$$

e quindi la (3.10) e la (3.11) devono coincidere. Pertanto:

$$L_{\alpha\beta} = L_{\beta\alpha}$$

ossia *la matrice dei coefficienti cinetici è simmetrica*. La reversibilità microscopica ("incarnata" dalla legge di bilancio dettagliato) interviene quindi nella descrizione di quantità che "regolano" l'irreversibilità macroscopica.

**Teorema di Wick** Vogliamo ora mostrare che in generale vale quello che abbiamo anticipato, ossia:

$$\langle y_{0\alpha}y_{0\beta}\rangle_{P^{\text{eq}}(\vec{y_0})} = \int d^{\nu}y_0 P^{\text{eq}}(\vec{y_0})y_{0\alpha}y_{0\beta} = S_{\alpha\beta}^{-1}$$

Consideriamo delle variabili  $\varphi_i$ , e A una matrice invertibile positiva; consideriamo allora l'integrale:

$$\int d\varphi_1 \cdots d\varphi_N e^{-\frac{1}{2} \sum_{ij} \varphi_i A_{ij} \varphi_j}$$

Con questo possiamo calcolare valori medi, in particolare quello di nostro interesse:

$$\langle \varphi_k \varphi_\ell \rangle = \frac{\int d\varphi_1 \cdots d\varphi_N e^{-\frac{1}{2} \sum_{ij} \varphi_i A_{ij} \varphi_j} \varphi_k \varphi_\ell}{\int d\varphi_1 \cdots d\varphi_N e^{-\frac{1}{2} \sum_{ij} \varphi_i A_{ij} \varphi_j}}$$

Introducendo  $Z(\vec{j})$ :

$$Z(\vec{j}) = \int d\varphi_1 \cdots d\varphi_N e^{\sum_i j_i \varphi_i} e^{-\frac{1}{2} \sum_{k\ell} \varphi_k A_{k\ell} \varphi_\ell}$$

il denominatore di  $\langle \varphi_k \varphi_\ell \rangle$  è Z(0), e più in generale possiamo scrivere:

$$\langle \varphi_k \varphi_\ell \rangle = \frac{1}{Z(0)} \frac{\partial^2 Z(\vec{j})}{\partial j_k \partial j_\ell}_{|\vec{j}=0}$$

(la Z è una sorta di funzione generatrice per i correlatori). Si può mostrare, poi (usando le proprietà dell'integrazione gaussiana), che:

$$Z(\vec{j}) = Z(0)e^{-\frac{1}{2}\sum_{\ell m} j_{\ell}(A^{-1})_{\ell m} j_{m}}$$

e quindi:

$$\langle \varphi_k \varphi_s \rangle = \frac{\partial^2}{\partial j_k \partial j_s} \left( e^{-\frac{1}{2} \sum_{\ell m} j_\ell (A^{-1})_{\ell m} j_m} \right)_{|\vec{j}=0} = A_{ks}^{-1}$$

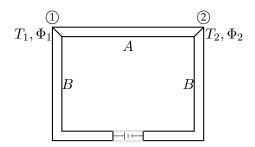


Figura 3.1: Schema dell'apparato sperimentale per lo studio degli effetti termoelettrici

Il caso precedentemente considerato lo si ritrova attraverso le seguenti identificazioni:

$$A \longleftrightarrow S \qquad \varphi \longleftrightarrow y_0 \qquad P^{\text{eq}}(\vec{y}_0) \longleftrightarrow \frac{e^{-\frac{1}{2}\sum_{ij}\varphi_i A_{ij}\varphi_j}}{\int d\varphi_1 \cdots d\varphi_N e^{-\frac{1}{2}\sum_{ij}\varphi_i A_{ij}\varphi_j}}$$

Se consideriamo il valor medio  $\langle \varphi_k \varphi_\ell \rangle$  come "contrazione" fra due variabili, possiamo poi usare una "generalizzazione" (o meglio, una reinterpretazione alla luce del senso che abbiamo appena dato alla contrazione fra due variabili) del teorema di Wick per calcolare oggetti più complicati come  $\langle \varphi_k \varphi_\ell \varphi_m \varphi_n \rangle$ .

### 3.3 Applicazioni

Vediamo ora delle applicazioni di quello che abbiamo appena visto.

L'esempio storico più importante di questo tipo di applicazioni è dato dai cosiddetti *effetti termoelettrici*.

Per capire di che cosa si tratta, consideriamo un sistema composto da una doppia giunzione metallica come in figura 3.1, i cui due estremi hanno due temperature diverse  $T_1$  e  $T_2$ , e sono sottoposte a due diversi potenziali elettrici  $\Phi_1$  e  $\Phi_2$ . Regolando questi parametri, oppure imponendo il passaggio di corrente attraverso la giunzione, si sono osservati sperimentalmente moltissimi effetti; noi ne consideriamo due:

**Effetto Peltier:** si collega la giunzione a un generatore di corrente, e imponiamo  $T_1 = T_2$ . Ciò che risulta è che una delle due giunzioni cede calore, mentre l'altra lo assorbe: si instaura quindi una corrente di calore (oltre a quella elettrica) fra le due giunzioni

**Effetto Seebeck:** mantenendo una differenza di temperatura ai due estremi della giunzione (tenendo il circuito aperto) si genera una differenza di potenziale

A metà Ottocento Thomson (colui che poi diventerà Lord Kelvin) scoprì una relazione a prima vista difficilmente spiegabile: determinò che esiste una relazione (la *relazione di Thomson*, che vedremo fra poco) fra il coefficiente che regola l'effetto Seebeck (che lega  $\Delta \Phi$  a  $\Delta T$ ) e quello che regola l'effetto Peltier (che lega  $\Delta T$  alla corrente I). Formulò anche una teoria per cercare di spiegarla, ma si rivelò poi falsa: fu Onsager il primo che riuscì a darne una spiegazione, usando le idee che abbiamo appena visto. Vediamo come.

Consieriamo il nostro sistema composto da una doppia giunzione metallica, come illustrato in figura 3.1. Il sistema è inizialmente isolato, quindi la sua energia interna U sarà la somma delle energie interne delle due giunzioni:

$$U = U_1 + U_2$$

relazioni analoghe varranno anche per la carica Q e l'entropia S:

$$Q = Q_1 + Q_2 \qquad \qquad S = S_1 + S_2$$

Per il principio di conservazione di carica ed energia:

$$dU = 0$$
  $\Rightarrow$   $dU_1 = -dU_2$   $dQ = 0$   $\Rightarrow$   $dQ_1 = -dQ_2$ 

Dalla termodinamica sappiamo che:

$$dU_1 = T_1 dS_1 + \Phi_1 dQ_1$$

e analogamente per  $dU_2$  ("a parole": l'energia interna di una giunzione può variare per una variazione di entropia o per effetto del lavoro elettrico esercitato su essa). Considerando l'entropia, quindi, si ha:

$$dS = dS_1 + dS_2 = \frac{dU_1}{T_1} - \frac{\Phi_1}{T_1}dQ_1 - \frac{dU_1}{T_2} + \frac{\Phi_2}{T_2}dQ_1$$

ove abbiamo sfruttato tutto quello che abbiamo appena visto. Pertanto:

$$dS = \left(\frac{1}{T_1} - \frac{1}{T_2}\right) dU_1 + \left(\frac{\Phi_2}{T_2} - \frac{\Phi_1}{T_1}\right) dQ_1$$

ove i coefficienti fra parentesi sono, per definizione, le componenti  $F_1$  e  $F_2$  della forza termodinamica generalizzata di cui abbiamo parlato precedentemente.

Vogliamo ora scrivere le equazioni del moto per l'energia e la carica, ad esempio relative alla prima giunzione; in altre parole, vogliamo determinare  $dU_1/dt$  e  $dQ_1/dt$ . Dalla (3.9) quindi, tenendo conto che il ruolo delle y è adesso rappresentato da  $U_1$  e  $Q_1$  si ha:

$$\frac{dU_1}{dt} = L_{11} \left( \frac{1}{T_1} - \frac{1}{T_2} \right) + L_{12} \left( \frac{\Phi_2}{T_2} - \frac{\Phi_1}{T_1} \right)$$

$$\frac{dQ_1}{dt} = L_{21} \left( \frac{1}{T_1} - \frac{1}{T_2} \right) + L_{22} \left( \frac{\Phi_2}{T_2} - \frac{\Phi_1}{T_1} \right)$$

Consideriamo ora l'effetto Seebeck; ponendo  $T_1 \sim T_2 \coloneqq T$  e  $\Delta T \coloneqq T_2 - T_1$  e analogamente per  $\Phi$  si ha:

$$\frac{1}{T_1} - \frac{1}{T_2} = \frac{\Delta T}{T^2}$$
  $\frac{\Phi_2}{T_2} - \frac{\Phi_1}{T_1} = \frac{\Phi_2 T_1 - \Phi_1 T_2}{T_1 T_2} \simeq \frac{\Delta \Phi}{T}$ 

e quindi:

$$\frac{dQ_1}{dt} = L_{21}\frac{\Delta T}{T^2} + L_{22}\frac{\Delta \Phi}{T} \qquad \qquad \frac{dU_1}{dt} = L_{11}\frac{\Delta T}{T^2} + L_{12}\frac{\Delta \Phi}{T}$$

Nell'esperimento di Seebeck, però, il circuito resta aperto e quindi dobbiamo porre  $dQ_1/dt = 0$ . Pertanto:

$$\frac{\Delta\Phi}{T} = -\frac{L_{21}}{L_{22}}\frac{\Delta T}{T^2} \qquad \Rightarrow \qquad \Delta\Phi = -\frac{L_{21}}{L_{22}}\frac{\Delta T}{T}$$

Generalmente questa relazione si scrive come  $\Delta\Phi=a\Delta T$ , e quindi il coefficiente che regola l'effetto Seebeck è  $a=-(L_{21}/L_{22})/T$ .

Passiamo ora all'effetto Peltier; in questo caso imponiamo  $\Delta T = 0$ , ottenendo:

$$\frac{dQ_1}{dt} = L_{22} \frac{\Delta \Phi}{T} \qquad \qquad \frac{dU_1}{dt} = L_{12} \frac{\Delta \Phi}{T}$$

Il coefficiente che regola l'effetto Peltier è quindi il rapporto fra queste due correnti:

$$b = \frac{\dot{U}_1}{\dot{Q}_1} = \frac{L_{12}}{L_{22}}$$

Notiamo quindi che poiché la matrice L è simmetrica (ed è qui che entra pesantemente in gioco il ragionamento precedente di Onsager) allora esiste una relazione fra i due coefficienti:

 $a = -\frac{b}{T}$ 

Questa è la *relazione di Thomson* precedentemente citata. È solo grazie a ragionamenti di questo tipo, che coinvolgono lo studio di sistemi leggermente fuori dall'equilibrio, che è stato possibile dedurla correttamente.

## Capitolo 4

# Teoria della risposta lineare

#### 4.1 Introduzione

La *teoria della risposta lineare* cerca di studiare a livello statistico dei fenomeni nei quali un sistema "risponde" a una qualche sollecitazione esterna. Si tratta di una teoria macroscopica, di cui però la meccanica statistica permette di dare un'interpretazione microscopica.

Come esempio di sistema che possiamo trattare con la teoria della risposta lineare, consideriamo un oscillatore armonico unidimensionale sottoposto ad una forza dissipativa (ad esempio immerso in un fluido viscoso); se x è la coordinata che ne descrive lo spostamento dalla posizione di equilibrio, la sua equazione del moto sarà:

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x = f(t)$$

ove  $\omega_0 = k/m$  e f(t) = F(t)/m, con k costante elastica della molla e F(t) forza esterna. Si tratta di un sistema macroscopico molto semplice, per il quale però *non* vale l'invarianza per inversione temporale; se infatti cerchiamo una soluzione tale per cui  $x'(t) = x(\tau - t)$ , si ha:

$$\ddot{x}' - \gamma \dot{x}' + \omega_0^2 x' = f(\tau - t)$$

che non è la stessa equazione di prima<sup>1</sup>!

In situazioni di questo tipo è interessante andare a studiare come il sistema risponde alla forzante esterna; ora, in situazioni come quella che stiamo considerando (ossia sistemi regolati da equazioni differenziali ordinarie lineari) vale il principio di sovrapposizione, e quindi la "relazione" fra x e f è di tipo lineare². Questo ci suggerisce che la x(t) si possa scrivere come:

$$x(t) = \int dt' \chi(t - t') f(t')$$
(4.1)

(ossia è un funzionale di f(t), e tale per cui "traslando" nel tempo la forzante "trasla" anche la soluzione), ove la funzione  $\chi$  è la "suscettività" del sistema, in generale detta funzione di risposta.

Per il principio di causalità, poi, la  $\chi$  non può essere diversa da zero se t' > t: x(t) può dipendere dai valori della f solo per istanti precedenti a t. In altre parole, supponiamo

<sup>&</sup>lt;sup>1</sup>Da notare, però, che in assenza del termine dissipativo l'invarianza per inversione temporale torna a valere.

<sup>&</sup>lt;sup>2</sup>In altre parole, se  $x_1$  è la soluzione con forzante  $f_1$  e  $x_2$  con forzante  $f_2$ ,  $x_1 + x_2$  è la soluzione con forzante  $f_1 + f_2$ .

che  $\chi(\tau) = 0$  se  $\tau < 0$ .

Usiamo ora anche la trasformata di Fourier; si avrà ad esempio:

$$\tilde{x}(\omega) = \int e^{i\omega t} x(t) dt = \tilde{\chi}(\omega) \tilde{f}(\omega)$$

perché la x(t) è la convoluzione di  $\chi$  e f.

Se continuiamo la  $\tilde{\chi}$  nel piano complesso, ossia la scriviamo come  $\tilde{\chi}(\omega+i\eta)$ , il principio di causalità impone che  $\tilde{\chi}$  abbia proprietà di analiticità che pongono vincoli molto forti sulla sua forma, e in particolare legano la sue parti reali e immaginaria. Si può poi anche mostrare che  $\operatorname{Im} \tilde{\chi}$  è legata alla dissipazione del sistema.

Ciò che in questi casi la meccanica statistica riesce a fare è mostrare che, se non siamo troppo distanti dall'equilibrio, queste funzioni di risposta si possono esprimere in termini di correlazioni di equilibrio dipendenti dal tempo di determinate grandezze.

#### 4.1.1 L'oscillatore armonico smorzato

Continuando a ragionare sull'esempio dell'oscillatore armonico smorzato, in termini di trasformate di Fourier l'equazione del moto diventa:

$$-\omega^2 \tilde{x}(\omega) - i\gamma \omega \tilde{x}(\omega) + \omega_0^2 \tilde{x}(\omega) = \tilde{f}(\omega)$$

Quindi, poiché  $\tilde{\chi} = \tilde{x}/\tilde{f}$ :

$$\tilde{\chi}(\omega) = -\frac{1}{\omega^2 - \omega_0^2 + i\gamma\omega}$$

Scriviamone esplicitamente le parti reale e immaginaria, che chiamiamo rispettivamente  $\tilde{\chi}'$  e  $\tilde{\chi}''$ :

$$\tilde{\chi}(\omega) = -\frac{1}{\omega^2 - \omega_0^2 + i\gamma\omega} \frac{(\omega^2 - \omega_0^2) - i\gamma\omega}{(\omega^2 - \omega_0^2) - i\gamma\omega} = \frac{\omega_0^2 - \omega^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2} + i\frac{\gamma\omega}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2}$$

$$\Rightarrow \qquad \tilde{\chi}'(\omega) = \frac{\omega_0^2 - \omega^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2} \qquad \tilde{\chi}''(\omega) = \frac{\gamma\omega}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2}$$

"Promuoviamo" adesso la  $\tilde{\chi}$  a funzione nel piano complesso, usando  $z=\omega+i\eta$  come suo argomento.

La prima proprietà importante della  $\tilde{\chi}(z)$  è che è analitica nel semipiano complesso superiore  ${\rm Im}\,z>0.$  Si ha infatti:

$$\tilde{\chi}(z) = -\frac{1}{z^2 - \omega_0^2 + i\gamma z}$$

e le radici del denominatore sono:

$$z_{\pm} = \frac{-i\gamma \pm \sqrt{-\gamma^2 + 4\omega_0^2}}{2}$$

Se  $\gamma^2 > 4\omega_0^2$  le soluzioni sono puramente immaginarie ma con  ${\rm Im}\,z_\pm < 0$ , mentre in caso contrario sono reali ma pur sempre con parte immaginaria negativa. Pertanto, eventuali poli di  $\tilde{\chi}(z)$  si trovano tutti nel semipiano inferiore.

Questa analiticità nel semipiano superiore è in qualche modo anche una conseguenza del principio di causalità; infatti, la trasformata di Fourier della  $\chi(z)$  si può scrivere come:

$$\tilde{\chi}(z) = \int_0^{+\infty} e^{izt} \chi(t) dt$$

ove l'estremo inferiore di integrazione è zero perché come abbiamo detto prima  $\chi(t)=0$  per t<0. A questo punto con argomenti di analisi matematica si può effettivamente dimostrare che così definita la  $\tilde{\chi}$  è analitica nel semipiano superiore<sup>3</sup>.

Vediamo ora come la  $\tilde{\chi}''$  sia legata alla dissipazione. Supponiamo di applicare una forzante oscillante al sistema, del tipo:

$$f(t) = f_0 \cos(\omega t)$$

Proviamo a calcolare la potenza dissipata dal sistema, che sappiamo essere:

$$\frac{d}{dt}W(t) = \dot{x}(t)f(t)$$

Riscriviamo innanzitutto la (4.1), effettuando un cambio di variabili, come:

$$x(t) = \int_0^{+\infty} \chi(t') f(t - t') dt'$$

Abbiamo allora, usando l'espressione di f(t) scelta:

$$x(t) = \int_0^{+\infty} dt' \chi(t') f_0 \cos\left(\omega(t - t')\right) = \int_0^{+\infty} dt' \chi(t') \operatorname{Re}\left[f_0 e^{-i\omega(t - t')}\right]$$

ove nell'ultimo passaggio abbiamo semplicemente riscritto il tutto in una forma più conveniente. Si ha allora:

$$\frac{d}{dt}x(t) = \frac{d}{dt}\operatorname{Re}\left[f_0e^{-i\omega t}\int_0^{+\infty} dt'\chi(t')e^{i\omega t'}\right] = 
= f_0\frac{d}{dt}\operatorname{Re}\left[\left(\cos(\omega t) - i\sin(\omega t)\right)\left(\tilde{\chi}'(\omega) + i\tilde{\chi}''(\omega)\right)\right] = 
= f_0\operatorname{Re}\left[\left(-\omega\sin(\omega t) - i\omega\cos(\omega t)\right)\left(\tilde{\chi}'(\omega) + i\tilde{\chi}''(\omega)\right)\right] = 
= f_0\left(-\omega\sin(\omega t)\tilde{\chi}''(\omega) + \omega\cos(\omega t)\tilde{\chi}''(\omega)\right)$$

e quindi:

$$\frac{dW}{dt} = f_0^2 \cos(\omega t) \left( -\omega \sin(\omega t) \tilde{\chi}'(\omega) + \omega \cos(\omega t) \tilde{\chi}''(\omega) \right)$$

La vera quantità interessante è la media di questa potenza dissipata su un periodo; poiché si ha  $\langle \cos(\omega t) \sin(\omega t) \rangle = 0$  e  $\langle \cos^2(\omega t) \rangle = 1/2$ , allora:

$$\left\langle \frac{dW}{dt} \right\rangle = \frac{f_0^2 \omega}{2} \tilde{\chi}''(\omega) \tag{4.2}$$

$$\tilde{\chi}_n(z) = \int_0^n e^{izt} \chi(t) dt$$

che sono tutte analitiche (per l'argomento che abbiamo mostrato precedentemente) per  ${\rm Im}\,z>0$ ; in questo caso quindi anche il limite della successione, cioè la  $\tilde\chi$ , è analitica nel semipiano superiore.

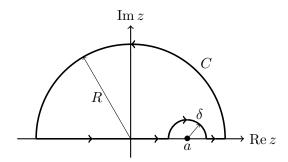
<sup>&</sup>lt;sup>3</sup>Per farlo si usa un teorema che, in linea generale, asserisce quanto segue: se abbiamo una successione di funzioni che convergono uniformemente a una data funzione, e sono tutte analitiche in un certo dominio del piano complesso, allora anche la funzione limite della successione è analitica nello stesso dominio. In questo caso la successione è data da:

che, notiamo, è *sempre* positiva. Vediamo quindi come effettivamente la potenza media dissipata in un periodo della forzante è direttamente proporzionale alla parte immaginaria della suscettività.

#### 4.1.2 Le relazioni di Kramers-Kronig

Come abbiamo appena visto, la parte immaginaria della funzione di risposta è direttamente coinvolta nella dissipazione del sistema. Tuttavia sarebbe auspicabile riuscire ad ottenere una conoscenza completa della  $\tilde{\chi}$ , compresa la sua parte reale; a questo proposito sono estremamente utili le *relazioni di Kramers-Kronig*, che come vediamo adesso legano fra loro le parti reale e immaginaria della  $\tilde{\chi}$ .

Ora, consideriamo una generica funzione  $\chi(z)$  e supponiamo che sia analitica per  ${\rm Im}\, z>0$ ; supponiamo poi anche che per  $|z|\to\infty$  (e sempre  ${\rm Im}\, z>0$ )  $|\chi(z)|$  decresca almeno come cost./|z|. Possiamo allora applicare il teorema di Cauchy alla funzione  $\chi(z)/(z-a)$  con a un qualche punto sull'asse reale; per applicarlo consideriamo un contorno d'integrazione di questo tipo:



Poiché la  $\chi$  è analitica nel semipiano superiore, l'integrale su questo cammino è nullo e quindi (evidenziamo i vari contributi dell'integrale):

$$\underbrace{\int_{-R}^{a-\delta} \frac{\chi(x)}{x-a} dx + \int_{a+\delta}^{R} \frac{\chi(x)}{x-a} dx}_{(1)} + \underbrace{\int_{0}^{\pi} iRe^{i\theta} \frac{\chi(Re^{i\theta})}{Re^{i\theta}-a} d\theta}_{(2)} + \underbrace{\int_{\pi}^{0} i\delta e^{i\theta} \frac{\chi(a+\delta e^{i\theta})}{\delta e^{i\theta}} d\theta}_{(3)} = 0 \quad (4.3)$$

Si può dimostrare che il modulo del contributo 2 è sempre maggiorato da  $R|\chi(Re^{i\theta})|/|R-a|$ , che si annulla nel limite  $R\to\infty$ .

Nello stesso limite, il contributo (1) tende alla parte principale dell'integrale:

mentre per il contributo (3) si ha, per continuità della  $\chi$ :

$$(3) \xrightarrow{\delta \to 0} -i\pi \chi(a)$$

Pertanto, la (4.3) nel limite  $R \to \infty$  e  $\delta \to 0$  diventa:

$$\mathcal{P}\left[\int_{-\infty}^{+\infty} \frac{\chi(x)}{x-a} dx\right] = i\pi\chi(a)$$

Se quindi inseriamo la  $\tilde{\chi}$  al posto della  $\chi$ , uguagliando parti reale e immaginaria otteniamo le *relazioni di Kramers-Kronig*:

$$\mathcal{P}\left[\int_{-\infty}^{+\infty} \frac{\tilde{\chi}'(x)}{x-a} dx\right] = -\pi \tilde{\chi}''(a)$$

$$\mathcal{P}\left[\int_{-\infty}^{+\infty} \frac{\tilde{\chi}''(x)}{x-a} dx\right] = \pi \tilde{\chi}'(a)$$

che, come vediamo, permettono di esprimere la parte reale o immaginaria della  $\tilde{\chi}$  a partire dalla conoscenza dell'altra.

### 4.2 Il teorema di fluttuazione-dissipazione

Fin'ora la nostra trattazione non ha niente di statistico: dobbiamo ancora vedere cosa ha da dire la meccanica statistica sulle funzioni di risposta, ed è quello che faremo adesso.

#### 4.2.1 Il teorema fluttuazione-risposta

Consideriamo quindi un sistema che all'equilibrio è descritto da una densità di probabilità  $\rho$  canonica nello spazio delle fasi, con hamiltoniana H(q,p) e a temperatura T. La sua funzione di partizione è allora:

$$Z(H) = \frac{1}{h^{3N}N!} \int dq dp e^{-\beta H}$$

Supponiamo ora di aggiungere un piccolo termine perturbativo all'hamiltoniana:

$$H(q,p) \longrightarrow H(q,p) + V(q,p)$$

e scriviamo V nella forma V(q,p) = -fA(q,p); f è un coefficiente, che regola l'intensità della perturbazione, mentre A è la funzione che ne regola la "forma".

Quello che ci interessa sapere è come possiamo approssimare al primo ordine in f i valori medi di date quantità; ad esempio, se ci interessa la quantità descritta dalla funzione B(q,p):

$$\langle B \rangle = \frac{\int dq dp e^{-\beta H} B}{\int dq dp e^{-\beta H}}$$

Detto  $\overline{B}$  il valor medio di B in presenza della perturbazione, si ha:

$$\overline{B} = \frac{\int dq dp e^{-\beta(H+V)} B}{\int dq dp e^{-\beta(H+V)}}$$

e, espandendo in V:

$$\overline{B} = \frac{Z(H)}{Z(H)} \cdot \frac{\int dq dp e^{-\beta H} (1 - \beta V) B}{\int dq dp e^{-\beta H} (1 - \beta V)} = \frac{\langle B \rangle - \beta \langle BV \rangle}{1 - \beta \langle V \rangle} \simeq \langle B \rangle - \beta (\langle BV \rangle - \langle B \rangle \langle V \rangle)$$

Pertanto, definendo il correlatore connesso come:

$$\langle BV \rangle_c := \langle BV \rangle - \langle B \rangle \langle V \rangle$$

si ha:

$$\overline{B} = \langle B \rangle - \beta \langle BV \rangle_c$$

Questo risultato è noto col nome di teorema fluttuazione-risposta.

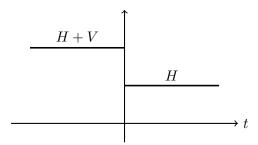


Figura 4.1: Evoluzione considerata dell'hamiltoniana

#### 4.2.2 Funzioni di risposta e correlatori di equilibrio

Per poter studiare effettivamente le funzioni di risposta, però, abbiamo bisogno di vedere cosa accade se f, invece di essere costante, dipende dal tempo.

Supponiamo quindi che la perturbazione V agisca da  $t=-\infty$  a t=0, e che in quest'ultimo istante si "spenga" (e pertanto resta solo H nell'hamiltoniana), come illustrato in figura 4.1. Vogliamo vedere come il sistema reagisce a questo "sbalzo".

Per t < 0 il sistema è ben equilibrato,  $\overline{B}$  è indipendente dal tempo e soddisfa il teorema di fluttuazione-risposta che abbiamo appena visto; per t > 0 invece,  $\overline{B}$  diventa una funzione del tempo  $\overline{B}(t)$ , che vogliamo determinare.

Fino a t=0 la densità canonica nello spazio delle fasi ha H+V come hamiltoniana; sfruttando la "visuale di Heisenberg", ossia tenendo fissa la  $\rho$  e facendo evolvere B, possiamo dire che per t>0 si ha:

$$\overline{B}(t) = \frac{\int dq dp e^{-\beta(H+V)} B(t)}{\int dq dp e^{-\beta(H+V)}}$$

e B(t) è l'evoluto di B secondo H. Quindi:

$$\overline{B}(t < 0) = \langle B \rangle - \beta \langle BV \rangle_c \qquad \overline{B}(t > 0) = \langle B \rangle - \beta \langle B(t)V(0) \rangle_c$$

Ricordando poi che V = -fA:

$$\overline{B}(t>0) = \langle B \rangle + \beta f \langle B(t)A(0) \rangle_c$$

che sono tutte quantità pertinenti all'equilibrio con hamiltoniana H. Vediamo dunque che:

$$\delta \overline{B} = \overline{B}(t) - \langle B \rangle = \beta f \langle B(t)A(0) \rangle_c$$

Scrivendo la f come funzione del tempo nella forma  $f(t) = f\Theta(-t)$ , la funzione di risposta  $\chi_{BA}$  (ove il pedice sta ad indicare che la funzione rappresenta una risposta di B dovuta ad una sollecitazione di A) dovrà essere tale che:

$$\delta \overline{B}(t) = \int_{-\infty}^{t} dt' \chi_{BA}(t - t') f(t') = f \int_{-\infty}^{0} dt' \chi_{BA}(t - t') = f \int_{t}^{+\infty} dt'' \chi_{BA}(t'')$$

ove nell'ultimo passaggio abbiamo effettuato il cambio di variabile t'' = t - t'. Si ha allora:

$$\frac{d}{dt}\delta \overline{B}(t) = \begin{cases} -f\chi_{BA}(t) & t > 0\\ 0 & t < 0 \end{cases}$$

Usando invece quanto visto precedentemente, si ha:

$$\frac{d}{dt}\delta \overline{B}(t) = \begin{cases} \beta f \left\langle \dot{B}(t)A(0) \right\rangle_c & t > 0\\ 0 & t < 0 \end{cases}$$

e quindi:

$$\chi_{BA}(t) = -\beta \Theta(t) \left\langle \dot{B}(t) A(0) \right\rangle_{c} \tag{4.4}$$

che era il risultato che volevamo mostrare, ossia che le funzioni di risposta si possono esprimere tramite correlatori di equilibrio dipendenti dal tempo.

Analizziamo ora qualche caso particolare che ci sarà utile in seguito. Se B=A, allora:

$$\chi_{AA}(t) = -\beta\Theta(t)\frac{d}{dt}\langle A(t)A(0)\rangle_c$$

ossia, se il sistema in esame viene portato fuori dall'equilibrio questo rilassa verso l'equilibrio stesso con la dinamica delle fluttuazioni all'equilibrio; è un'altra forma dell'ipotesi di regressione di Onsager.

Sempre in questo caso, secondo quanto visto prima si ha:

$$\chi_{AA}(t) = -\beta\Theta(t) \left\langle \dot{A}(t)A(0) \right\rangle_{c}$$

In generale si definisce la funzione di autocorrelazione (o autocorrelatore) come:

$$C_{AA}(t) = \langle A(t)A(0)\rangle_c$$

che normalmente ha proprietà "standard", per esempio è tale che  $C_{AA}(t) = C_{AA}(-t)$ ; infatti:

$$C_{AA}(t) = \langle A(t)A(0)\rangle = \langle A(0)A(-t)\rangle = \langle A(-t)A(0)\rangle = C_{AA}(-t)$$

ove abbiamo sfruttato la stazionarietà dei fenomeni di equilibrio, ossia l'invarianza di A all'equilibrio sotto traslazioni temporali.

Consideriamo poi anche  $\langle (A(t)-A(0))^2 \rangle$ , quantità sicuramente positiva; si ha:

$$\langle A(t)^2 \rangle + \langle A(0)^2 \rangle \ge 2 \langle A(t)A(0) \rangle$$

Sottraendo membro a membro  $2\langle A\rangle^2$  e considerando che  $\langle A(t)^2\rangle - \langle A\rangle^2 = \langle A(0)^2\rangle - \langle A\rangle^2 = C_{AA}(0)$ , si ha:

$$2C_{AA}(0) \ge 2C_{AA}(t)$$
  $\Rightarrow$   $C_{AA}(0) \ge C_{AA}(t)$ 

e quindi la funzione di autocorrelazione decresce. Si può dimostrare anche che si ha  $C_{AA}(t) \ge -C_{AA}(0)$ , e generalmente la  $C_{AA}$  decade esponenzialmente per grandi tempi:

$$C_{AA}(t) \stackrel{t \to \infty}{\sim} e^{-|t|/\tau}$$

Altra proprietà della funzione di autocorrelazione che ci servirà è che  $\dot{C}_{AA}(t)=-\dot{C}_{AA}(-t).$ 

#### 4.2.3 Il teorema di fluttuazione-dissipazione

Riconsideriamo l'esempio precedente dell'oscillatore armonico smorzato. La parte immaginaria della trasformata di Fourier della funzione di risposta è:

$$\tilde{\chi}''(\omega) = \operatorname{Im} \int dt e^{i\omega t} \chi(t) = -\beta \int_0^{+\infty} dt \sin(\omega t) \dot{C}(t) = -\frac{\beta}{2} \int_{-\infty}^{+\infty} dt \sin(\omega t) \dot{C}(t) =$$

$$= -\frac{\beta}{2i} \int_{-\infty}^{+\infty} dt e^{i\omega t} \dot{C}(t)$$

(ove nell'ultimo passaggio si è aggiunto all'integrando un termine dispari,  $\cos(\omega t)\dot{C}(t)$ , il cui integrale è nullo) e poiché:

$$\tilde{\dot{C}}(\omega) = -i\omega\tilde{C}(\omega)$$

allora:

$$\tilde{\chi}''(\omega) = \frac{\beta\omega}{2}\tilde{C}(\omega)$$

Usando questo risultato nella (4.2) si ottiene:

$$\left\langle \frac{dW}{dt} \right\rangle = \beta \frac{f_0^2 \omega^2}{4} \tilde{C}(\omega)$$

ossia stiamo esprimendo la potenza media dissipata in termini della funzione di autocorrelazione. Questo risultato è noto col nome di *teorema di fluttuazione-dissipazione*.

Abbiamo visto con l'equazione (4.4) che la funzione di risposta di un sistema si può esprimere in termini del correlatore connesso  $\langle \dot{B}(t)A(0)\rangle_c$ . Dato che è spesso utile, vogliamo mostrare che questo correlatore può essere riscritto come:

$$\left\langle \dot{B}(t)A(0)\right\rangle_{c} = -\left\langle B(t)\dot{A}(0)\right\rangle_{c}$$
 (4.5)

Innanzitutto:

$$\langle B(t)A(0)\rangle_c = \langle B(0)A(-t)\rangle_c$$

ove abbiamo effettuato una traslazione temporale. Derivando ambo i membri rispetto al tempo, allora:

$$\left\langle \dot{B}(t)A(0)\right\rangle_{c} = -\left\langle B(0)\dot{A}(-t)\right\rangle_{c} = -\left\langle B(t)\dot{A}(0)\right\rangle_{c}$$

ove nell'ultimo passaggio abbiamo sfruttato l'invarianza per traslazione temporale.

Esempio: relazione di Green-Kubo Quello che abbiamo appena visto risulta utile ad esempio nel caso che vogliamo discutere adesso, che ha a che fare con la conduzione di cariche in un mezzo. Consideriamo un sistema di volume  $\Omega$  composto di cariche mobili (ad esempio un metallo) di intensità q, come quello in figura 4.2, e supponiamo di esercitare un campo elettrico su di esso. Come risposta a questa sollecitazione si instaurerà una corrente nel sistema.

La perturbazione V all'hamiltoniana la possiamo scrivere come:

$$V = -E(t)q\sum_{i} x_{i}$$

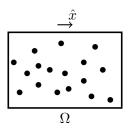


Figura 4.2: Sistema considerato

ove, rispetto alla notazione generale che abbiamo adottato, E(t) ha il ruolo della funzione f(t), mentre  $q \sum_i x_i$  è la funzione A (che in questo caso dipende solo dalle posizioni delle particelle, e non dai loro momenti).

Come abbiamo detto la quantità che ci interessa studiare è la corrente, ossia:

$$B = q \sum_{i} \dot{x}_{i}$$

Quindi  $B = \dot{A}$ , e possiamo anche scriverla come  $B = \dot{A} = j\Omega$ , ove j è la densità di corrente. Pertanto:

$$\delta \overline{B}(\omega) = \Omega j(\omega) = \beta q^2 E(\omega) \int_0^{+\infty} dt e^{i\omega t} \sum_{i,j} \langle \dot{x}_i(t) \dot{x}_j(0) \rangle_{c,E=0}$$
 (4.6)

ove nella prima uguaglianza abbiamo tenuto conto del fatto che a campo nullo non c'è corrente, ossia j=0; anche se non è evidente, poi, abbiamo usato la (4.5). Il pedice E=0 indica esplicitamente che il correlatore va calcolato all'equilibrio (in questo caso, in assenza di campo esterno).

Vediamo meglio come si giunge a questo risultato: innanzitutto, per quanto visto in 4.1 si ha  $\delta \overline{B}(\omega) = \chi(\omega) f(\omega)$  (per semplicità non indichiamo le trasformate di Fourier con la tilde, ma solo esplicitando il loro argomento), ove  $f(\omega) = E(\omega)$ . Dalla (4.4), poi, sappiamo che  $\chi(t) = -\beta \Theta(t) \left\langle \dot{B}(t) A(0) \right\rangle_c$ ; pertanto:

$$\chi(\omega) = -\beta \int_0^{+\infty} dt e^{i\omega t} \left\langle \dot{B}(t) A(0) \right\rangle_c = \beta \int_0^{+\infty} dt e^{i\omega t} \left\langle B(t) \dot{A}(0) \right\rangle_c$$

ove nell'ultimo passaggio abbiamo esplicitamente usato la (4.5). A questo punto, tenendo conto che:

$$B(t) = \dot{A}(t) = q \sum_{i} \dot{x}_{i}(t)$$

si trova proprio la (4.6).

Volendo,  $\delta \overline{B}(\omega)$  si può scrivere in modo leggermente diverso:

$$\delta \overline{B}(\omega) = \beta \Omega^2 E(\omega) \int_0^{+\infty} dt e^{i\omega t} \langle j(t)j(0) \rangle_{c,E=0} = \Omega j(\omega)$$

Una quantità interessante del sistema è la sua conducibilità  $\sigma$ . Poiché in generale questa è definita come la quantità tale che  $j(\omega)=\sigma(\omega)E(\omega)$ , usando l'espressione di  $j(\omega)$ 

ottenibile dall'equazione precedente si ha:

$$\sigma(\omega) = \beta \Omega \int_0^{+\infty} dt \, \langle j(t)j(0) \rangle_{c,E=0} \, e^{i\omega t}$$

Nel limite statico  $\omega \to 0$  si ha:

$$\sigma(0) = \beta \Omega \int_0^{+\infty} dt \, \langle j(t)j(0) \rangle_{c,E=0}$$

e quindi la conducibilità è legata alle fluttuazioni di corrente nel metallo *a campo esterno nullo*. Quella che abbiamo appena visto è una delle *relazioni di Green-Kubo*, che in generale legano correlatori a coefficienti di trasporto (in questo caso la conducibilità, appunto).

**Esempio:** il moto browniano Un altro esempio che è utile vedere in questo àmbito è il moto browniano: in generale infatti le relazioni che abbiamo appena visto coinvolgono correlatori molto difficili da calcolare. Il moto browniano ci permette però di vedere all'opera tutto quello che abbiamo appena visto senza troppa difficoltà.

Avevamo visto nel capitolo 1 che (ponendoci in una dimensione):

$$\frac{dv}{dt} = -\gamma v + \sqrt{\frac{2\gamma k_B T}{m}} f(t) + \frac{1}{m} F(t)$$
(4.7)

ove f è la forza fluttuante e F la forzante esterna; con questa scelta dei coefficienti, poi, si ha  $\langle f(t_1)f(t_2)\rangle=\delta(t_2-t_1)$ .

Sappiamo che dopo un transiente iniziale, in assenza della forzante F, la particella ha  $\langle v \rangle_{\rm eq} = 0$ ; pertanto la risposta del sistema alla F sarà una variazione di velocità  $\delta \overline{v}$ . Per determinarla facciamo la media (perturbata) della (4.7) (tenendo conto che, poiché è fluttuante,  $\overline{f}(t) = 0$ ):

$$\frac{d\overline{v}}{dt} = -\gamma \overline{v} + \sqrt{\frac{2\gamma k_B T}{m}} \overline{f}(t) + \frac{1}{m} F(t) = -\gamma \overline{v} + \frac{1}{m} F(t)$$

Facendo la trasformata di Fourier ad ambo i membri:

$$-i\omega \overline{v}(\omega) = -\gamma \overline{v}(\omega) + \frac{1}{m} F(\omega) \qquad \Rightarrow \qquad \overline{v}(\omega) = \frac{1}{m(\gamma - i\omega)} F(\omega)$$

e quindi possiamo subito dire che:

$$\tilde{\chi}(\omega) = \frac{1}{m(\gamma - i\omega)}$$

Usiamo ora quello che abbiamo visto con la teoria della risposta lineare per calcolare  $\tilde{\chi}$ . La perturbazione esterna al sistema è data da V=-F(t)x con x posizione della particella, e quindi si ha A=x e  $B=\dot{A}$  (la risposta è data dalla velocità della particella). Usando i risultati precedenti, quindi, si ha:

$$\tilde{\chi}(\omega) = \beta \int_{0}^{+\infty} dt e^{i\omega t} \langle v(t)v(0)\rangle_{c}$$

e usando il risultato valido per il moto browniano<sup>4</sup>:

$$\langle v(t)v(0)\rangle_c = \frac{1}{\beta m}e^{-\gamma|t|}$$

<sup>&</sup>lt;sup>4</sup>Discende dal valore del correlatore  $\langle X(t)X(0)\rangle$ .

allora:

$$\tilde{\chi}(\omega) = \beta \int_0^{+\infty} dt e^{i\omega t} \frac{1}{\beta m} e^{-\gamma|t|} = \frac{1}{m} \int_0^{+\infty} dt e^{-\gamma t} e^{i\omega t} = \frac{1}{m} \cdot \frac{1}{\gamma - i\omega}$$

Il risultato è quindi (giustamente) sempre lo stesso.

**Teoria della risposta lineare quantistica** In tutto lo studio della teoria della risposta abbiamo sempre considerato per semplicità un sistema classico, e mai quantistico. Nel caso in cui si voglia formulare una teoria della risposta quantistica non ci sono grandi differenze concettuali rispetto a quello che abbiamo visto, ma ci sono alcune problemi "tecnici": in particolare in generale H e V non commutano, e quindi  $e^{H+V} \neq e^H e^V$  (bisogna usare identità come quella di Baker-Campbell-Hausdorff).

## Capitolo 5

## Processi stocastici

In questo capitolo torniamo a studiare più in dettaglio i processi stocastici, concentrandoci in particolar modo su quelli markoviani; allo stesso tempo rivisiteremo alcuni aspetti del moto browniano.

Senza ripetere tutto quello che abbiamo visto in 1.2.2, sappiamo che i processi markoviani sono di grande interesse in Fisica perché permettono di descrivere molti fenomeni interessanti; sappiamo anche che un processo stocastico può essere formulato fornendo una "gerarchia" di densità di probabilità congiunte  $P_n(y_1, t_1; ...; y_n, t_n)$ .

#### 5.1 Processi stazionari

Fra i vari possibili processi stocatstici hanno molto interesse in fisica i cosiddetti *processi stazionari*, ossia quelli tali che (sfruttando sempre la stessa notazione usata in 1.2.2):

$$P_n(y_1, t_1 + \tau; y_2, t_2 + \tau; \dots; y_n, t_n + \tau) = P_n(y_1, t_1; y_2, t_2; \dots; y_n, t_n)$$

Se quindi Y è un processo stocastico stazionario il suo valor medio non dipenderà dal tempo:  $\langle Y(t) \rangle = \langle Y \rangle = \text{cost.}$ ; inoltre si ha:

$$\langle Y(t_1)Y(t_2)\rangle - \langle Y\rangle^2 = k(t_2 - t_1)$$

ove k è una funzione pari come conseguenza dell'invarianza per traslazioni temporali, e generalmente  $k(\tau) \sim e^{-|\tau|/\tau_0}$ .

#### 5.1.1 Teorema di Wiener-Khinchin

Un importante risultato che riguarda i processi stocastici stazionari è il *teorema di Wiener-Khinchin*, che vogliamo mostrare.

Consideriamo un processo stazionario a media nulla ( $\langle Y \rangle = 0$ ); per esso anche  $\langle Y^2 \rangle$  è indipendente dal tempo. L'idea fondamentale alla base del teorema è di definire uno "spettro" delle fluttuazioni e analizzare i contributi di ogni frequenza.

Supponiamo quindi che il nostro processo  $Y_x(t)$  sia definito per un intervallo temporale molto grande,  $t \in [0,T]$ . Fissato il valore x della variabile stocastica X "soggiacente" al processo, quindi, possiamo pensare di sviluppare in serie di Fourier ogni realizzazione del processo, ossia definire:

$$A_{n,x} = \frac{2}{T} \int_0^T \sin\left(\frac{n\pi}{T}t\right) Y_x(t) dt$$
 (5.1)

di modo che:

$$Y_x(t) = \sum_{n=1}^{\infty} A_{n,x} \sin\left(\frac{n\pi}{T}t\right)$$

L'identità di Parseval per quest'espressione di  $Y_x(t)$  è:

$$\int |Y_x(t)|^2 dt = \sum_{n,n'} \int_0^T A_{n,x} A_{n',x} \sin\left(\frac{n\pi}{T}t\right) \sin\left(\frac{n'\pi}{T}t\right) dt = \frac{T}{2} \sum_n A_{n,x}^2$$

ove nell'ultimo passaggio abbiamo sfruttato il fatto che  $\int_0^T \sin(n\pi t/T)\sin(n'\pi t/T)dt = \delta_{n,n'}T/2$ . Pertanto:

$$\frac{1}{T} \int Y_x^2(t) dt = \frac{1}{2} \sum_n A_{n,x}^2$$

e mediando su x ad ambo i membri:

$$\frac{1}{T} \int \left\langle Y^2 \right\rangle dt = \frac{1}{2} \sum_n \left\langle A_{n,x}^2 \right\rangle \qquad \Rightarrow \qquad \left\langle Y^2 \right\rangle = \frac{1}{2} \sum_n \left\langle A_{n,x}^2 \right\rangle$$

Vediamo quindi come alle fluttuazioni di Y contribuiscano più frequenze. Posto ovviamente che le frequenze sono nella forma  $\omega=n\pi/T$ , definiamo lo *spettro delle fluttuazioni* come la funzione  $S(\omega)$  tale che:

$$\frac{1}{2} \sum_{\omega \le n\pi/T \le \omega + \Delta\omega} \left\langle A_{n,x}^2 \right\rangle = S(\omega) \Delta\omega \tag{5.2}$$

ove la somma è stata ristretta al range degli n tali per cui  $n\pi/T$  si trova in un piccolo intervallo di frequenze<sup>1</sup>, fra  $\omega$  e  $\omega + \Delta \omega$ .

Abbiamo pertanto che  $S(\omega)$  è il contributo alla fluttuazione totale  $\langle Y^2 \rangle$  dato dalla frequenza  $\omega$ .

Il teorema di Wiener-Khinchin asserisce che lo spettro delle fluttuazioni ha la seguente espressione:

$$S(\omega) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \cos(\omega \tau) k(\tau) d\tau$$

ove  $k(\tau)$  è la funzione di autocorrelazione che abbiamo visto poco sopra. In altre parole, il teorema di Wiener-Khinchin asserisce che lo spettro delle fluttuazioni è la cosiddetta "trasformata del coseno" della funzione di autocorrelazione.

Per dimostrare il teorema, iniziamo notando che dalla (5.1) si ha:

$$\langle A_{n,x}^2 \rangle = \frac{4}{T^2} \int_0^T \sin\left(\frac{n\pi}{T}t\right) \sin\left(\frac{n\pi}{T}t'\right) \langle Y_x(t)Y_x(t')\rangle dtdt'$$

e si ha  $\langle Y_x(t)Y_x(t')\rangle = k(t-t')$ . Postuliamo che la funzione k sia sensibilmente diversa da zero in un intervallo molto piccolo rispetto a quello di integrazione; effettuando poi il cambio di variabili  $\tau = t' - t$  si ha:

$$\langle A_{n,x}^2 \rangle = \frac{4}{T^2} \int_0^T dt \sin\left(\frac{n\pi}{T}t\right) \int_{-t}^{T-t} d\tau \sin\left(\frac{n\pi}{T}(t+\tau)\right) k(\tau)$$

<sup>&</sup>lt;sup>1</sup>Notare che poiché T è molto grande,  $\omega$  ha valori "quasi continui".

e l'ultimo seno, usando le formule trigonometriche, è uguale a:

$$\sin\left(\frac{n\pi}{T}(t+\tau)\right) = \underbrace{\sin\left(\frac{n\pi}{T}t\right)\cos\left(\frac{n\pi}{T}\tau\right)}_{\boxed{1}} + \underbrace{\cos\left(\frac{n\pi}{T}t\right)\sin\left(\frac{n\pi}{T}\tau\right)}_{\boxed{2}}$$

Per l'ipotesi che abbiamo fatto su k l'integrazione in  $d\tau$  è di fatto estesa a un intervallo molto piccolo (quello su cui k è diversa da zero), quindi possiamo fare l'integrale in  $\tau$  su tutto l'asse reale, trascurando gli effetti di bordo. Il contributo di 2 è allora nullo (complessivamente c'è un integrale in dt di  $\sin(n\pi t/T)\cos(n\pi t/T)$ , che è nullo) , mentre non lo è quello di 1:

$$\left\langle A_{n,x}^2\right\rangle = \frac{4}{T^2} \int_0^T dt \sin^2\left(\frac{n\pi}{T}t\right) \int_{-\infty}^{+\infty} d\tau \cos\left(\frac{n\pi}{T}\tau\right) k(\tau) = \frac{4}{T^2} \cdot \frac{T}{2} \int_{-\infty}^{+\infty} \cos\left(\frac{n\pi}{T}\tau\right) k(\tau) d\tau$$

Riconsideriamo ora la (5.2): per un dato  $\Delta\omega$  i termini che contribuiscono saranno ovviamente quelli tali per cui  $\Delta\omega = \Delta n \cdot \pi/T$ . Pertanto (sommando su questo intervallo di valori per n):

$$S(\omega)\Delta\omega = \frac{1}{2} \cdot \frac{T}{\pi} \Delta\omega \frac{2}{T} \int_{-\infty}^{+\infty} \cos(\omega\tau) k(\tau) d\tau$$

ossia:

$$S(\omega) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \cos(\omega \tau) k(\tau) d\tau$$

che era proprio quello che volevamo dimostrare.

#### 5.2 Processi markoviani

Abbiamo già visto in 1.2.2 che, dato un processo stocastico generico, possiamo definire in generale una probabilità condizionata<sup>2</sup>:

$$P_{\ell|k}(y_{k+1}, t_{k+1}; \dots; y_{k+\ell}, t_{k+\ell}|y_1, t_1; \dots; y_k, t_k) = \frac{P_{k+\ell}(y_1, t_1; \dots; y_k, t_k; y_{k+1}, t_{k+1}; \dots; y_{k+\ell}, t_{k+\ell})}{P_k(y_1, t_1; \dots; y_k, t_k)}$$

Se ad esempio consideriamo  $\ell=k=1$  si avrà:

$$P_{1|1}(y_2, t_2|y_1, t_1) = \frac{P_2(y_1, t_1; y_2, t_2)}{P_1(y_1, t_1)}$$

e questa è la probabilità che Y assuma valore  $y_2$  a  $t_2$  sapendo che valeva  $y_1$  a  $t_1$ . In altre parole: selezionando fra tutte le realizzazioni di  $Y_x(t)$  quelle che passano per  $y_1$  a  $t_1$ , la frazione di questo "sottoensemble" che passa fra  $y_2$  e  $y_2 + dy_2$  all'istante  $t_2$  è data da  $P_{1|1}(y_2,t_2|y_1,t_1)dy_2$ .

Un processo stocastico, a questo punto, è markoviano se tutte le  $P_{\ell|k}$  coincidono con una  $P_{\ell|1}$ , nel senso che la conoscenza dell'ultimo valore  $y_k$  del processo a  $t_k$  determina completamente le probabilità dell'evento successivo. Più formalmente, per un processo markoviano si ha:

$$P_{\ell|k}(y_{k+1}, t_{k+1}; \dots; y_{k+\ell}, t_{k+\ell}|y_1, t_1; \dots; y_k, t_k) = P_{\ell|1}(y_{k+1}, t_{k+1}; \dots; y_{k+\ell}, t_{k+\ell}|y_k, t_k)$$

Un processo markoviano può quindi essere completamente ricostruito a partire dalla conoscenza di  $P_1$  e  $P_{1|1}$  (quest'ultima è anche detta *probabilità di transizione*).

 $<sup>^{2}</sup>$ Usiamo una notazione in cui a pedice indichiamo il numero di punti (y, t) coinvolti.

### 5.2.1 Equazione di Chapman-Kolmogoroff

Vogliamo ora arrivare determinare e studiare una delle equazioni fondamentali dei processi markoviani, l'equazione di Chapman-Kolmogoroff. Come vedremo, quest'equazione fornisce le condizioni sotto le quali una data  $P_1$  e  $P_{1|1}$  possono descrivere un processo markoviano.

Consideriamo un processo stocastico a tre step,  $t_1 < t_2 < t_3$ . La probabilità congiunta  $P_3(y_1, t_1; y_2, t_2; y_3, t_3)$  sarà allora data da:

$$P_3(y_1, t_1; y_2, t_2; y_3, t_3) = P_2(y_1, t_1; y_2, t_2) P_{1|1}(y_3, t_3|y_2, t_2) =$$

$$= P_1(y_1, t_1) P_{1|1}(y_2, t_2|y_1, t_1) P_{1|1}(y_3, t_3|y_2, t_2)$$

Integrando in  $y_2$  ad ambo i membri:

$$\int dy_2 P_3(y_1, t_1; y_2, t_2; y_3, t_3) = P_2(y_1, t_1; y_3, t_3) = P_1(y_1, t_1) P_{1|1}(y_3, t_3|y_1, t_1) =$$

$$= P_1(y_1, t_1) \int P_{1|1}(y_2, t_2|y_1, t_1) P_{1|1}(y_3, t_3|y_2, t_2) dy_2$$

e quindi:

$$P_{1|1}(y_3, t_3|y_1, t_1) = \int P_{1|1}(y_3, t_3|y_2, t_2) P_{1|1}(y_2, t_2|y_1, t_1) dy_2$$
(5.3)

che è l'equazione di Chapman-Kolmogoroff.

Affinché la  $P_1$  e la  $P_{1|1}$  descrivano effettivamente un processo markoviano, però, oltre a soddisfare la (5.3) devono anche soddisfare necessariamente la relazione:

$$P_1(y_2, t_2) = \int P_{1|1}(y_2, t_2|y_1, t_1) P_1(y_1, t_1) dy_1$$

#### 5.2.2 Esempi e considerazioni

Vediamo ora qualche esempio e qualche considerazione sui processi markoviani.

**Processo di Wiener** Innanzitutto, considerando un processo di Wiener (che abbiamo incontrato in 1.4.2), per esso si avrà:

$$P_{1|1}(y_{2}, t_{2}|y_{1}, t_{1}) = \frac{1}{\sqrt{2\pi(t_{2} - t_{1})}} \exp\left(-\frac{(y_{2} - y_{1})^{2}}{2(t_{2} - t_{1})}\right)$$

$$P_{1|1}(y_{2}, t_{2}|y_{1}, t_{1}) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y_{2} - y_{1})^{2}}{2(t_{2} - t_{1})}\right)$$

Non è un processo propriamente stazionario (la  $P_1$  dipende dal tempo) ma è temporalmente omogeneo, ossia la  $P_{1|1}$  dipende solo dalla differenza dei tempi.

**Processo di Ornstein-Uhlenbeck** Il processo di Ornstein-Uhlenbeck è invece un esempio di processo markoviano stazionario. In realtà così come lo abbiamo visto in 1.4.2 non è propriamente stazionario (anche se rilassa verso la stazionarietà per tempi lunghi), ma il processo di Ornstein-Uhlenbeck propriamente detto è stazionario, e definito da:

$$P_1(y,t) = \frac{e^{-y^2/2}}{\sqrt{2\pi}} \qquad P_{1|1}(y_2, t_2|y_1, t_1) = T_{(t_2-t_1)}(y_2|y_1)$$

ove la seconda è una notazione generale che si usa per processi temporalmente omogenei, e in particolare per questo processo definendo  $\tau=t_2-t_1$  si ha:

$$T_{\tau}(y_2|y_1) = \frac{1}{\sqrt{2\pi(1 - e^{-2\tau})}} \exp\left(-\frac{(y_2 - y_1 e^{-\tau})^2}{2(1 - e^{-2\tau})}\right)$$

A questo punto si ha, ad esempio:

$$\int dy_2 T_{\tau}(y_3|y_2) T_{\tau'}(y_2|y_1) = T_{\tau+\tau'}(y_3|y_1) \qquad \qquad \int dy_0 T_{\tau}(y|y_0) \frac{e^{-y_0^2/2}}{\sqrt{2\pi}} = \frac{e^{-y^2/2}}{\sqrt{2\pi}}$$

Esiste poi un teorema, il *teorema di Doob*, che asserisce che questo è l'*unico* processo markoviano, gaussiano e stazionario.

**Moto browniano** Come abbiamo visto, la velocità di una particella browniana è ben descritta da un processo stocastico markoviano: la probabilità di variazione della velocità della particella in un certo istante dipende solo dalla velocità in quello stesso istante, o in altre parole non è necessario conoscere la storia della velocità per determinarne la sua evoluzione futura.

Strettamente questa è un'approssimazione, anche se buonissima: possiamo infatti pensare che in realtà la storia precedente della particella possa aver in qualche modo influenzato il fluido che la circonda (ad esempio instaurando dei moti vorticosi), e quindi la sua evoluzione futura può essere "un po" influenzata dalla sua storia. Ovviamente la markovianità diventa un'ottima approssimazione nel momento in cui queste correlazioni si estendono su scale temporali piccolissime.

Il moto browniano, poi, può essere anche studiato in termini dello spostamento della particella; in questo caso allora la markovianità è un'ottima approssimazione se il tempo caratteristico dello spostamento della particella è molto maggiore di quello di autocorrelazione della velocità.

Consideriamo una particella che passa per diverse posizioni  $(\dots, x_{k-1}, x_k, x_{k+1}, \dots)$ , e supponiamo che i "salti" da una posizione all'altra siano indipendenti fra loro. Supponiamo anche di sapere che il salto  $x_{k-1} \to x_k$  sia grande: allora è ragionevole ritenere che in  $x_k$  la velocità della particella sia molto grande. Questo però rende molto più probabile che anche il salto  $x_k \to x_{k+1}$  sia molto grande, e quindi in realtà si potrebbe in qualche modo instaurare una correlazione fra gli spostamenti. Questa correlazione però diventa assolutamente trascurabile se il tempo di autocorrelazione della velocità della particella è molto minore dei suoi tempi caratteristici di spostamento.

**Processi deterministici** I processi deterministici possono essere descritti come processi stocastici markoviani "singolari". Se infatti il processo è descritto da un'equazione differenziale assolutamente generale del tipo  $\dot{x}=f(x)$ , la cui soluzione è  $x(t)=\varphi(t,x_0)$  (ove  $(x_0,t_0)$  è il punto iniziale), allora:

$$P_{1|1}(x,t|x_0,t_0) = \delta(x - \varphi(t,x_0))$$

"Perdita" di markovianità In generale, se "riduciamo" il numero di variabili coinvolte in un processo markoviano (ad esempio integrando su quel grado di libertà, oppure agendo con determinati proiettori) non è detto che il processo risultante sia ancora markoviano.

Un caso dove questo *non* avviene è la velocità  $\vec{v} = (v_x, v_y, v_z)$  di una particella browniana (pensata come processo stocastico markoviano): se consideriamo una sola delle sue componenti il processo resta ancora markoviano.

Se invece ad esempio consideriamo un sistema composto da più reagenti chimici e due di questi, A e B, possono associarsi e dissociarsi ( $AB \longleftrightarrow A+B$ ), possiamo supporre che la probabilità di associazione/dissociazione dipenda dalle concentrazioni istantanee dei componenti, e quindi che il processo di associazione/dissociazione sia markoviano; in realtà se consideriamo solo i componenti A e B (perdendo l'informazione su tutti gli altri) questo processo non è markoviano.

In generale, però, la markovianità di un processo può essere ripristinata includendo delle nuove variabili.

Consideriamo ad esempio una particella browniana soggetta a un campo di forza esterno (che ovviamente dipenderà dalla posizione), e studiamo il processo dato dai suoi spostamenti; in questo caso allora lo spostamento della particella non dipenderà più solo dalla sua velocità istantanea, ma anche dalla sua posizione. Per conoscere la posizione, però, bisogna conoscere la storia della velocità della particella, e quindi il processo non è markoviano. La markovianità è però recuperata se consideriamo come variabili del problema spostamento e velocità.

In generale, quindi, includendo le giuste variabili si può recuperare la markovianità di un processo.

Catene di Markov e matrici stocastiche Spesso (come in parte abbiamo visto all'inizio di 5.2) per un dato processo stocastico stazionario è opportuno considerarne un "sottoensemble"; in pratica quello che si fa è una "selezione" delle possibili realizzazioni del processo che soddisfano a date condizioni, ad esempio il fatto che a t=0 si abbia  $y=y_0$ , con  $y_0$  dato. Passando a questo "sottoensemble" stiamo quindi modificando la  $P_1$  e la  $P_{1|1}$ :

$$P_1(y) \longrightarrow P_1^*(y,t) = T_t(y|y_0) \qquad \qquad P_{1|1}(y_2|y_1) \longrightarrow P_{1|1}^*(y_2,t_2|y_1,t_1) = T_{t_2-t_1}(y_2|y_1)$$

Ci aspetteremmo quindi che  $P_1^*(y,t) \stackrel{t \to \infty}{\longrightarrow} P_1(y)$ , ossia che il processo del "sottoensemble" rilassi verso il processo stazionario (è una situazione analoga a quella vista in 1.4.2 per il processo di Ornstein-Uhlenbeck). Molto spesso è cruciale (soprattutto per le applicazioni) riuscire a capire se un processo tenda alla stazionarietà come abbiamo appena visto; diventa più semplice trattare problemi del genere se i processi che consideriamo vengono descritti in termini di *catene di Markov*. Una *catena di Markov* è fondamentalmente un processo stocastico markoviano come quelli che abbiamo considerato, ma nel quale sia Y che il tempo assumono valori discreti (si tratta di una sorta di "discretizzazione" di un processo continuo).

Se la catena è N-dimensionale (ossia y può assumere N diversi possibili valori) allora  $T_{\tau}(y'|y)$  sarà una matrice  $N\times N$ , e per l'equazione di Chapman-Kolmogoroff si avrà ad esempio che  $T_{\tau}=(T_1)^{\tau}$ . Le matrici che descrivono catene di Markov sono dette *matrici* 

stocastiche, le quali soddisfano a due importanti proprietà:

$$T_{mn} \ge 0$$
 
$$\sum_{m} T_{mn} = 1 \quad \forall n$$

ossia sono semidefinite positive e la somma dei termini appartenenti ad ogni colonna dà 1.

Notare che per via della seconda proprietà si ha sicuramente:

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} T = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

ossia  $(1,\ldots,1)^T$  è sicuramente autovalore sinistro di T con autovalore 1 (in generale la T non è simmetrica). Dovrà allora esistere un autovalore destro  $p_s$  di T tale che  $Tp_s=p_s$ . Un teorema di Perron e Frobenius garantisce che questo autovalore esiste, e che questo sia lo stato stazionario del sistema per  $t\to\infty$ .

## 5.3 Equazione di Langevin

Quello che vogliamo cercare di fare adesso è capire come l'equazione di Langevin "emerga" in quest'impostazione, in particolare usando gli strumenti della teoria della risposta.

## 5.3.1 Modello di Caldeira-Leggett

Sono stati formulati molti modelli per giustificare l'equazione di Langevin nel passaggio dal microscopico al macroscopico; un modello molto utile fra questi è stato quello di Caldeira e Leggett.

Ne consideriamo una formulazione semplice e in un caso anch'esso semplice, ossia in una dimensione. Supponiamo quindi di avere una particella di massa M, posizione X e impulso P che interagisce con un bagno termico di molecole di massa  $m_{\lambda}$ , posizione  $x_{\lambda}$  e impulso  $p_{\lambda}$ . Prendiamo come hamiltoniana del sistema:

$$H = \frac{P^2}{2M} + \sum_{\lambda} \frac{p_{\lambda}^2}{2m_{\lambda}} + \frac{1}{2} \sum_{\lambda} m_{\lambda} \omega_{\lambda}^2 (X - x_{\lambda})^2$$

ove quindi supponiamo che l'interazione fra la particella e le molecole sia di tipo armonico; H è chiaramente invariante per traslazioni. Svolgendo il quadrato:

$$H = \frac{P^2}{2M} + \frac{\kappa}{2}X^2 + \sum_{\lambda} \left(\frac{p_{\lambda}^2}{2m_{\lambda}} + \frac{m_{\lambda}}{2}\omega_{\lambda}^2 x_{\lambda}^2\right) - X\sum_{\lambda} m_{\lambda}\omega_{\lambda}^2 x_{\lambda}$$
 (5.4)

ove  $\kappa = \sum_{\lambda} m_{\lambda} \omega_{\lambda}^2$ ; questa riscrittura pur non essendo palesemente invariante per traslazioni è utile perché evidenzia il carattere "armonico" sia della particella che delle molecole del bagno termico.

Notare che in questa forma l'hamiltoniana può essere vista come composta da un termine fondamentale e uno perturbativo (l'ultima sommatoria).

Il nostro obiettivo è determinare le equazioni che regolano il moto della particella nell'ipotesi che questa sia in equilibrio termico col bagno a una data temperatura. Le equazioni di Hamilton del sistema, molto semplicemente, sono:

$$\dot{P} = -\kappa X + \sum_{\lambda} m_{\lambda} \omega_{\lambda}^{2} x_{\lambda} \qquad \dot{X} = \frac{P}{M}$$
 (5.5)

$$\dot{p}_{\lambda} = -m_{\lambda}\omega_{\lambda}^{2}(x_{\lambda} - X) \qquad \dot{x}_{\lambda} = \frac{p_{\lambda}}{m_{\lambda}}$$
 (5.6)

Consideriamo per ora solo le versioni "imperturbate" di queste equazioni, ossia quelle che derivano dall'hamiltoniana *H* nella (5.4) senza l'ultima sommatoria:

$$X(t) \sum_{\lambda} m_{\lambda} \omega_{\lambda}^2 x_{\lambda} := X(t) F(t)$$

Considerando infatti questo termine come una piccola perturbazione, sfruttando la notazione che abbiamo usato in 4.2.1 possiamo considerare la X(t) come l'analogo della f(t) e la F(t) come l'equivalente della  $A(q_t, p_t)$ .

Supponendo che il bagno termico sia all'equilibrio e imperturbato, chiaramente la F(t) potrà fluttuare ma avrà necessariamente media nulla (il sistema è composto da più oscillatori armonici indipendenti, e ogni  $x_{\lambda}$  ha media nulla); la presenza della perturbazione, però, farà sì che il valor medio della F venga "spostato" di una quantità  $\overline{\delta F}(t)$ , data dalla teoria della risposta. In altre parole, rispetto al caso imperturbato l'unica (importante) differenza è che la F(t) non ha più media nulla, ma data dalla teoria della risposta. In particolare, usando gli strumenti fin'ora sviluppati, possiamo dire che:

$$\overline{\delta F}(t) = \int_0^t dt' \chi(t - t') X(t')$$

ove abbiamo supposto che la "perturbazione" X(t) sia diversa da zero per t>0, ossia in altre parole poniamo la particella nel bagno a t=0.

Sappiamo però anche che la funzione di risposta si può scrivere come:

$$\chi(t - t') = -\beta \frac{d}{dt} \left\langle F(t)F(t') \right\rangle$$

o più semplicemente, definendo t'' = t - t':

$$\overline{\delta F}(t) = \int_0^t dt'' \chi(t'') X(t - t'') \qquad \qquad \chi(t'') = -\beta \dot{c}(t'') = -\beta \frac{d}{dt''} \left\langle F(t'') F(0) \right\rangle$$

ove c è la funzione di autocorrelazione.

Si ha pertanto:

$$\overline{\delta F}(t) = -\beta \int_0^t \dot{c}(t'') X(t - t'') dt''$$
(5.7)

ossia, la teoria della risposta permette di esprimere lo shift  $\overline{\delta F}$  in termini della funzione di autocorrelazione relavita però al sistema *imperturbato* all'equilibrio termico.

Possiamo quindi scrivere la (5.5) come:

$$\dot{P} = -\kappa X + F(t) + \overline{\delta F}(t)$$

(ove la F è relativa al sistema imperturbato). Esplicitiamo quindi la (5.7) integrando per parti:

$$\begin{split} \overline{\delta F}(t) &= -\beta c(t'') X(t-t'') \big|_{t''=0}^{t''=t} -\beta \int_0^t dt'' c(t'') \dot{X}(t-t'') = \\ &= \beta c(0) X(t) - \beta c(t) X(0) -\beta \int_0^t dt'' c(t'') \dot{X}(t-t'') \end{split}$$

Vediamo ora come quest'espressione si semplifica valutandola per tempi grandi, e qual è il significato dei vari termini.

Innanzitutto, sappiamo che  $c(t) = \langle F(t)F(0)\rangle$ ; le soluzioni dell'equazione del moto sono tali per cui:

$$\ddot{x}_{\lambda} = -m_{\lambda}\omega_{\lambda}^2 x_{\lambda}(t) \qquad \Rightarrow \qquad m_{\lambda}\omega_{\lambda}^2 x_{\lambda}(t) = m_{\lambda}\omega_{\lambda}^2 x_{\lambda}(0)\cos(\omega t) + \omega_{\lambda}p_{\lambda}(0)\sin(\omega t)$$

e quindi:

$$c(t) = \langle F(t)F(0) \rangle = \left\langle \sum_{\lambda,\lambda'} m_{\lambda} \omega_{\lambda}^2 x_{\lambda}(t) \cdot m_{\lambda'} \omega_{\lambda'}^2 x_{\lambda'}(0) \right\rangle =$$

$$= \left\langle \sum_{\lambda,\lambda'} \left[ m_{\lambda} \omega_{\lambda}^2 x_{\lambda}(0) \cos(\omega_{\lambda} t) + \omega_{\lambda} p_{\lambda}(0) \sin(\omega_{\lambda} t) \right] m_{\lambda'} \omega_{\lambda'}^2 x_{\lambda'}(0) \right\rangle$$

Essendo in equilibrio termico, per gli oscillatori vale il teorema di equipartizione dell'energia e quindi:

$$m_{\lambda}\omega_{\lambda}^{2}\langle x_{\lambda}(0)x_{\lambda'}(0)\rangle = k_{B}T\delta_{\lambda,\lambda'} \quad \forall \lambda,\lambda' \qquad \langle x_{\lambda}(0)p_{\lambda'}(0)\rangle = 0 \quad \forall \lambda,\lambda'$$

Pertanto:

$$c(t) = k_B T \sum_{\lambda} m_{\lambda} \omega_{\lambda}^2 \cos(\omega_{\lambda} t)$$
 (5.8)

e in particolare, ricordando la definizione di  $\kappa$ :

$$c(0) = \frac{\kappa}{\beta}$$

Il primo termine della  $\overline{\delta F}$ , quindi, cancella il primo termine della  $\dot{P}$ .

Passiamo ai termini successivi, in particolare consideriamo  $\beta c(t)X(0)$ : questo termine per grandi tempi è trascurabile, perché normalmente (come abbiamo già avuto modo di specificare) ci si aspetta che la funzione di autocorrelazione c(t) decada velocemente nel tempo.

Possiamo quindi concludere che l'evoluzione della *P*, a tempi grandi, è data da:

$$\dot{P} = F(t) - \beta \int_0^t dt' c(t') \dot{X}(t - t')$$

Si vede chiaramente che in questo caso il sistema *non* presenta markovianità, perché la  $\dot{P}$  dipende dalla storia passata della  $\dot{X}$  (ci sono effetti di memoria<sup>3</sup>).

Quello che vogliamo fare adesso è vedere come si può arrivare da quest'equazione a quella di Langevin, e all'introduzione del rumore bianco gaussiano.

Consideriamo il termine integrale che contribuisce alla  $\dot{P}$ : supponendo che i tempi caratteristici del moto della particella siano molto maggiori del range temporale in cui la c(t) è diversa da zero, allora questo lo si può approssimare con:

$$-\beta \dot{X}(t) \int_{0}^{+\infty} c(t')dt'$$

 $<sup>^{3}</sup>$ Questi effetti di memoria emergono praticamente sempre quando si effettuano procedure simili al coarse graining, ossia quando si integra su variabili microscopiche.

ove abbiamo portato  $\dot{X}$  fuori dall'integrale perché t' varia pochissimo rispetto a t, e l'integrale può essere esteso da t=0 a  $t=+\infty$  perché in questo modo si introduce comunque un errore trascurabile (sempre perché l'intervallo su cui c(t) è diversa da zero è molto piccolo). In questo modo non ci sono più effetti di memoria, perché così  $\dot{P}(t)$  dipende solo dalla  $\emph{velocità}$  della particella nello stesso istante. Pertanto, definendo:

$$-\frac{\beta}{M} \int_0^{+\infty} c(t)dt = -\gamma$$

si ha:

$$\dot{P} = F(t) - \gamma P$$

Per come l'abbiamo scritto, poi, F(t) è una somma di variabili gaussiane, e quindi F stessa è distribuita normalmente. Il fatto poi che il suo autocorrelatore sia diverso da zero in un range temporale molto piccolo ha come conseguenza che questa "forza gaussiana" si possa effettivamente considerare come un rumore bianco<sup>4</sup>.

Abbiamo quindi effettivamente trovato un'equazione del tipo Langevin, in cui la derivata temporale della velocità della particella (o equivalentemente del momento) è data da una forza fluttuante che ha le caratteristiche di un rumore bianco gaussiano e da un termine viscoso, proporzionale alla velocità stessa.

Il caso delle frequenze continue Nella derivazione che abbiamo appena visto c'è però una piccola difficoltà, ossia il fatto che per avere una F(t) che abbia veramente le proprietà desiderate bisogna passare da un modello a frequenze discrete come il nostro a uno a frequenze continue. Come si può passare da una situazione all'altra?

Definiamo la seguente funzione:

$$J(\omega) = \frac{\pi}{2} \sum_{\lambda} m_{\lambda} \omega_{\lambda}^{3} \delta(\omega - \omega_{\lambda})$$

Allora si ha:

$$\frac{2}{\pi} \int_0^{+\infty} \frac{d\omega}{\omega} J(\omega) = \sum_{\lambda} m_{\lambda} \omega_{\lambda}^2$$

Supponiamo poi di voler generalizzare l'espressione di  $J(\omega)$ , continuandola ad esempio per  $\omega$  negative supponendo che  $J(-\omega)=-J(\omega)$ . Allora dalla (5.8) si ha:

$$\frac{1}{2}\beta c(t) = \frac{1}{\pi} \int_0^{+\infty} \frac{J(\omega)}{\omega} \cos(\omega t) d\omega = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi\omega} J(\omega) e^{i\omega t}$$

ove abbiamo sfruttato nell'ultimo passaggio la parità di  $J(\omega)/\omega$ . La  $J(\omega)$  è quindi una sorta di trasformata di Fourier della c(t).

Per poter fare un esempio nel quale svolgere conti semplicemente, supponiamo che la  $J(\omega)$  abbia una determinata forma, e in particolare prendiamo:

$$J(\omega) = M\gamma\omega \frac{\omega_c^2}{\omega^2 + \omega_c^2}$$

<sup>&</sup>lt;sup>4</sup>Ricordare che fra le ipotesi fatte in 1.1 sulla forza fluttuante  $\vec{f}$  c'è il fatto che il suo autocorrelatore sia regolato da una funzione molto piccata su un piccolissimo intervallo temporale (agli effetti pratici poi l'abbiamo considerata una  $\delta$  di Dirac).

(ossia  $J/\omega$  è una distribuzione lorentziana), e ci aspettiamo che  $\omega_c$  sia molto grande perché  $\tau_c=1/\omega_c$  è il range temporale nel quale la c(t) è non nulla. Si ha allora:

$$\frac{1}{2}\beta c(t) = M\gamma \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\omega_c^2}{(\omega + i\omega_c)(\omega - i\omega_c)} e^{i\omega t}$$

Per calcolarlo usiamo il teorema di Cauchy, e per t>0 prendiamo un cammino d'integrazione chiuso nel semipiano complesso superiore; l'integrale è quindi uguale al residuo dell'integrando nel polo  $\omega=+i\omega_c$ , e quindi:

$$\frac{1}{2}\beta c(t) = \frac{M\gamma}{2\pi} 2\pi i \frac{\omega_c^2}{2i\omega_c} e^{-\omega_c t} = \frac{1}{2} M\gamma \omega_c e^{-\omega_c t}$$

e analogamente se t<0. Se quindi  $\omega_c^{-1}=\tau_c$  è molto minore dei tempi caratteristici del moto della particella è giustificata l'assunzione che il rumore gaussiano sia bianco.

## 5.4 L'equazione master

#### 5.4.1 Introduzione

In 5.2 abbiamo parlato di processi markoviani, e abbiamo visto come un generico processo markoviano è completamente determinato una volta date due probabilità, che abbiamo chiamato  $P_1(y,t)$  e  $P_{1|1}(y_2,t_2|y_1,t_1)$  grazie all'equazione di Chapman-Kolmogoroff. Abbiamo poi anche parlato di catene di Markov, ossia processi nei quali le y e il tempo t assumono valori discreti.

Quello che vogliamo fare adesso è fornire una caratterizzazione continua dei processi markoviani equivalente a quella data dall'equazione di Chapman-Kolmogoroff detta master equation, che altro non è che una "traduzione" integro-differenziale della prima. Il vantaggio, come vedremo, è che la master equation è meno "astratta" di quella di Chapman-Kolmogoroff, e più semplicemente collegabile alla fisica.

**Esempio preliminare** Prima di procedere alla determinazione e allo studio della master equation conviene mostrare un esempio nel quale usiamo le manipolazioni che ci serviranno in seguito.

Consideriamo il processo dato dal decadimento di N atomi radioattivi. Supponiamo che  $N(t=0)=n_0$ , e chiamiamo  $P_1(n_1,t)$  la probabilità che al tempo t ci siano ancora  $n_1$  atomi non decaduti, e  $P_{1|1}(n_2,t_2|n_1,t_1)$  la probabilità condizionata che ci siano  $n_2$  atomi non decaduti a  $t_2$  sapendo che a  $t_1$  ce n'erano  $n_1$  (notare che stiamo supponendo il tempo continuo).

Ovviamente ogni atomo ha una probabilità w(t) di "sopravvivere" fino al tempo t; come varia nel tempo questa funzione? Supponendo che la probabilità che un atomo decada in un intervallo temporale infinitesimo dt sia proporzionale ad esso, in particolare uguale a  $\gamma dt$  con  $\gamma$  costante, allora si avrà  $w(t+dt)=w(t)(1-\gamma dt)$ . Pertanto  $w(t)=w(0)e^{-\gamma t}=e^{-\gamma t}$ , perché w(0)=1. Allora:

$$P_1(n_1, t_1) = \binom{n_0}{n_1} w(t)^{n_1} (1 - w(t))^{n_0 - n_1}$$

e inoltre:

$$P_{1|1}(n_2, t_2|n_1, t_1) = \binom{n_1}{n_2} w(t_2 - t_1)^{n_2} (1 - w(t_2 - t_1))^{n_1 - n_2}$$

Supponiamo di voler calcolare il numero medio di atomi non decaduti all'istante  $t_1$ ; si ha quindi:

$$\langle N(t_1) \rangle = \sum_{n_1} n_1 \binom{n_0}{n_1} w(t)^{n_1} (1 - w(t))^{n_0 - n_1} = \frac{\partial}{\partial w} \left[ (w + v)^{n_0} \right]_{|v = 1 - w} = w(t) n_0$$

Ci chiediamo ora come vari la forma della probabilità di transizione quando la si valuta su intervalli temporali sempre più piccoli. Tenendo conto che  $w(t)=e^{-\gamma t}$  ed espandendo gli esponenziali, definendo  $\tau=t_2-t_1$  si ha:

$$P_{1|1}(n_2, t_2|n_1, t_1) = \binom{n_1}{n_2} (1 - \gamma \tau + \cdots)^{n_2} (1 - 1 + \gamma \tau + \cdots)^{n_1 - n_2}$$

Supponendo poi ad esempio  $n_1 = n_2$  (ossia nessun atomo è decaduto), allora:

$$P_{1|1}(n_2, t_2|n_1, t_1) = 1 - n_1 \gamma \tau + \cdots$$

mentre se  $n_2 = n_1 - 1$  (ossia è decaduto un solo atomo):

$$P_{1|1}(n_2, t_2|n_1, t_1) = n_1 \gamma \tau + \cdots$$

Pertanto, all'ordine  $\tau$  si può scrivere:

$$P_{1|1}(n_2, t_2|n_1, t_1) = \delta_{n_1, n_2}(1 - n_1\gamma\tau) + \delta_{n_1 - 1, n_2}n_1\gamma\tau$$

Questo è un caso particolare di un risultato generale che adesso discuteremo.

## 5.4.2 L'equazione master

Per un processo stocastico markoviano continuo sappiamo che vale l'equazione di Chapman-Kolmogoroff (5.3), e se il processo è temporalmente omogeneo allora vale anche:

$$T_{\tau+\tau'}(y_3|y_1) = \int dy_2 T_{\tau'}(y_3|y_2) T_{\tau}(y_2|y_1)$$

Quello che vogliamo fare adesso è vedere quale forma assume l'equazione nel limite  $\tau' \to 0$ .

"Ispirati" dal caso particolare che abbiamo appena visto, a destra possiamo sostituire  $T_{\tau'}(y_3|y_2)$  con:

$$T_{\tau'}(y_3|y_2) = (1 - \tau' a_0(y_2))\delta(y_3 - y_2) + \tau' W(y_3|y_2) + \cdots$$

ove<sup>5</sup>:

$$a_0(y_2) = \int dy_3 W(y_3|y_2)$$

e W è una probabilità di transizione per unità di tempo:

$$W(y_3|y_2) = \lim_{\tau' \to 0} \frac{T_{\tau'}(y_3|y_2)}{\tau'}$$

e quindi effettivamente  $a_0(y_2) = \int dy_3 W(y_3|y_2)$ .

<sup>&</sup>lt;sup>5</sup>Questo risultato si determina dalla condizione di normalizzazione di  $T_{\tau'}(y_3|y_2)$ :

Si ha allora:

$$T_{\tau+\tau'}(y_3|y_1) = \int dy_2 \left\{ [1 - \tau' a_0(y_2)] \delta(y_3 - y_2) T_{\tau}(y_2|y_1) + \tau' W(y_3|y_2) T_{\tau}(y_2|y_1) \right\} =$$

$$= T_{\tau}(y_3|y_1) - \tau' a_0(y_3) T_{\tau}(y_3|y_1) + \tau' \int W(y_3|y_2) T_{\tau}(y_2|y_1) dy_2$$

ove  $a_0(y_3)$  lo possiamo scrivere come:

$$a_0(y_3) = \int dy_2 W(y_2|y_3)$$

Portando  $T_{\tau}(y_3|y_1)$  a sinistra e dividendo ad ambo i membri per  $\tau'$ :

$$\frac{T_{\tau+\tau'}(y_3|y_1) - T_{\tau}(y_3|y_1)}{\tau'} = \int dy_2 \left[ W(y_3|y_2) T_{\tau}(y_2|y_1) - W(y_2|y_3) T_{\tau}(y_3|y_1) \right]$$

e nel limite  $\tau' \to 0$  a primo membro compare una derivata:

$$\frac{\partial}{\partial \tau} T_{\tau}(y_3|y_1) = \int dy_2 \left[ W(y_3|y_2) T_{\tau}(y_2|y_1) - W(y_2|y_3) T_{\tau}(y_3|y_1) \right]$$
 (5.9)

Questa è la master equation.

Rispetto all'equazione di Chapman-Kolmogoroff, quest'equazione ha il vantaggio di mettere in evidenza la probabilità di transizione W, la cui stima può derivare anche dallo studio microscopico del sistema. Ovviamente la Chapman-Kolmogoroff e la master equation sono equivalenti, ma per quanto abbiamo appena detto quest'ultima è più usata in Fisica.

Spesso la master equation è espressa in termini di probabilità che a un dato tempo t i parametri del sistema assumano determinati valori; in altre parole, invece che in termini di  $T_{\tau}$  l'equazione è espressa in termini di  $P_1$ . In realtà, però, quelle che vanno considerate non sono le  $P_1$  ma le  $P_1^*$  (ossia le probabilità relative al sottoensemble); insomma, alla fine si ha:

$$\frac{\partial}{\partial t} P_1^*(y_3, t) = \int dy_2 \left[ W(y_3|y_2) P_1^*(y_2, t) - W(y_2|y_3) P_1^*(y_3, t) \right]$$

(e in genere l'1 e il \* sono omessi).

### 5.4.3 L'equazione master per processi discreti

Le applicazioni della master equation sono innumerevoli; per ora limitiamoci a studiarla per un processo a valori discreti (il tempo continuiamo a supporlo continuo, quindi non stiamo studiando catene di Markov in realtà).

In questo caso il sistema è caratterizzato da un "vettore" di probabilità  $p_n(t)$  (eventualmente finito), e la master equation assume la forma:

$$\dot{p}_n(t) = \sum_{n'} \left[ W_{nn'} p_n(t) - W_{n'n} p_{n'}(t) \right] \tag{5.10}$$

Nel caso dell'esempio precedente sul decadimento radioattivo, quest'equazione diventa:

$$\dot{p}_n(t) = \gamma(n+1)p_{n+1}(t) - \gamma n p_n(t)$$

Se invece consideriamo un random walk su tutta la retta reale, allora:

$$\dot{p}_n(t) = p_{n+1}(t) + p_{n-1}(t) - 2p_n(t)$$

Quest'ultimo processo in realtà non ha uno stato stazionario, perché lasciandolo evolvere la densità di probabilità di trovare la particella si "sparpaglia" su tutta la retta; se invece ponessimo dei vincoli, ad esempio  $\dot{p}_N(t)=p_{N-1}(t)-p_N(t)$  (il sistema si "blocca" in n=N), allora il numero di stati possibili del sistema è finito, e la densità di probabilità tende ad uniformarsi su tutto l'intervallo (e quindi il sistema ha uno stato stazionario). Vediamo quindi anche l'importanza che può avere il tipo di dominio su cui è definita la master equation (e di conseguenza le sue soluzioni).

Si può anche intuire che considerando catene di Markov non sorgono problemi fintanto che le matrici che entrano in gioco nella descrizione di un processo sono finite, mentre se diventano infinite (o peggio ancora la variabile del sistema assume valori continui) possono sorgere molti problemi.

Introduciamo ora una nuova notazione, definendo:

$$\mathbb{W} = \begin{cases} W_{nn'} & \text{se } n \neq n' \\ -\sum_{n \neq n'} W_{n'n} & \text{se } n = n' \end{cases}$$

In questo modo la (5.10) si può riscrivere come:

$$\dot{p}_n(t) = \sum_{n'} \mathbb{W}_{nn'} p_{n'}(t)$$

Le W sono matrici con proprietà molto particolari e importanti: dalla definizione che abbiamo dato vediamo che gli elementi fuori dalla diagonale, ad esempio, sono tutti non negativi, mentre quelli sulla diagonale sono negativi; inoltre:

$$\sum_{n'} \mathbb{W}_{n'n} = 0 \qquad \forall n$$

ossia la somma dei termini presenti in ogni colonna dà sempre zero. Poiché poi non sono simmetriche, in genere i loro autovettori e autovalori non sono rilevanti.

Queste proprietà non si conservano sotto trasformazioni di similarità, ma permamgono se si permutano contemporaneamente le loro righe e colonne (in pratica si fa un "relabeling" degli stati); dopo una manipolazione del genere si può addirittura avere che  $\mathbb W$  sia scrivibile nella forma:

$$\mathbb{W} = \begin{pmatrix} \mathbb{A} & \mathbb{0} \\ \mathbb{0} & \mathbb{B} \end{pmatrix}$$

In questo caso allora il processo è in realtà costituito da due processi indipendenti fra loro.

Dalle proprietà delle W si ha poi:

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \mathbb{W} = 0$$

e quindi  $(1,1,\ldots,1)^T$  è autovettore sinistro di  $\mathbb{W}$  con autovalore nullo; dovrà quindi esistere un autovettore destro  $\varphi_s$  tale che  $\mathbb{W}\varphi_s=0$ . Questo autovettore (ma non abbiamo la

garanzia che in generale le sue componenti siano non negative e normalizzate) dovrebbe essere uno stato stazionario del sistema.

In realtà si può dimostrare che se le  $\mathbb{W}$  hanno le proprietà che abbiamo illustrato e hanno dimensione finita, allora questo stato stazionario esiste effettivamente e il sistema ci tende per tempi lunghi.

Ci possono poi essere anche matrici  $\mathbb{W}$  con la seguente struttura:

$$\mathbb{W} = \begin{pmatrix} \mathbb{A} & \mathbb{D} \\ \mathbb{0} & \mathbb{B} \end{pmatrix}$$

In questo caso si scopre allora che il sistema è soggetto a uno "svuotamento" delle probabilità degli stati contenuti in  $\mathbb{B}$ , e sopravvivono solo quelli contenuti in  $\mathbb{A}$ ; in altre parole, gli stati di  $\mathbb{B}$  sono transienti e si "svuotano" a favore di quelli di  $\mathbb{A}$ .

La dimostrazione (che non vediamo) parte dal considerare l'insieme degli autovalori  $\varphi_i(t)$ ; in un generico istante alcune di questi saranno positivi, altri nulli e altri ancora negativi. Chiamiamo  $\varphi_u(t)$  quelli positivi,  $\varphi_v(t)$  quelli negativi e  $\varphi_w(t)$  quelli nulli. A questo punto si studia l'evoluzione nel tempo della somma:

$$U(t) = \sum_{u} \varphi_u(t)$$

che può variare perché le  $\varphi_u$  evolvono nel tempo, e in particolare il loro numero cambia a tempi discreti. Quando il loro numero varia (perché una delle  $\varphi_i$  è diventata negativa, oppure da negativa è diventata positiva), in quell'istante la variazione di U è continua, mentre negli istanti intermedi U può solo decrescere (e questo discende dalle proprietà della  $\mathbb{W}$ ). A partire da questo si può dimostrare l'esistenza del limite per  $t \to \infty$  dell'unico stato stazionario del sistema, e che l'evoluzione tende ad esso nello stesso limite.

# Capitolo 6

# Termodinamica di sistemi microscopici

In questo capitolo vogliamo andare a studiare la termodinamica di sistemi microscopici, e in particolare ci interessa capire come in questo contesto si caratterizzi il secondo principio della termodinamica.

Andando a scale microscopiche, infatti, le fluttuazioni statistiche assumono una grandissima rilevanza nell'evoluzione delle grandezze termodinamiche, e dobbiamo quindi capire come tutto quello che conosciamo per sistemi macroscopici vada riformulato.

#### 6.1 Introduzione

Di recente sono state sviluppate tecniche sperimentali che permettono di andare a studiare direttamente la termodinamica di sistemi microscopici.

Un importante esempio di questo tipo di strumenti sono le pinze ottiche.

Consideriamo delle sfere di polistirene in sospensione; con particolari tecniche è possibile "agganciare" ad esse singoli filamenti di RNA. Le sfere di polistirene hanno particolari proprietà ottiche che permettono di controllarle e spostarle usando fasci laser<sup>1</sup> (generalmente una delle due sfere è tenuta ferma da una pipetta, mentre l'altra è mossa dal laser); in particolare il laser fornisce alla sfera una buca di potenziale all'interno della quale collocarsi (vedi figura 6.1). Possiamo quindi esercitare lavoro sul filamento di RNA spostando le sfere con questo dispositivo; poiché la molecola non è un sistema macroscopico in esperimenti del genere possiamo vedere direttamente le caratteristiche termodinamiche di un sistema fortemente fluttuante<sup>2</sup>.

bagno termico. <sup>2</sup>Da un punto di vista "storico", prima dello sviluppo di queste tecniche sperimentali i problemi della termodinamica fluttuante non erano ancora stati esplorati da un punto di vista teorico.

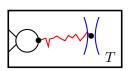


Figura 6.1: Setup sperimentale per il "tiraggio" di filamenti di RNA con trappole ottiche

 $<sup>^{1}</sup>$ L'analogo macroscopico di quest'esperimento è l'allungamento o accorciamento di un elastico in un

Quello che vedremo è che in questi àmbiti il secondo principio della termodinamica assume una formulazione statistica, all'interno della quale però le fluttuazioni saranno rilevantissime (d'altra parte, i valori medi delle grandezze termodinamiche continueranno a soddisfare le leggi note).

Vediamo brevemente come in quest'àmbito si debba riconsiderare il secondo principio della termodinamica.

Sappiamo che il secondo principio asserisce che, se *A* e *B* sono stati d'equilibrio:

$$\int_{A}^{B} \frac{\delta Q}{T} \le S_B - S_A$$

Se poi la trasformazione avviene a temperatura costante,  $Q/T \le \Delta S$ . Dal primo principio, poi, sappiamo che:

$$\Delta U = W + Q$$

ove abbiamo usato la convenzione di considerare W positivo se è eseguito sul sistema; sappiamo allora che detta F=U-TS l'energia libera si avrà:

$$W > \Delta F = \Delta U - T\Delta S$$

Nel caso particolare di un ciclo termodinamico, ossia una trasformazione dove B=A, allora:

$$W_{\text{cycle}} \ge 0$$

e eseguendo sempre lo stesso ciclo con un sistema termodinamico macroscopico otteniamo sempre la stessa quantità di lavoro.

Nel caso di sistemi microscopici come il filamento di RNA che stiamo considerando, però, ciò che risulta sperimentalmente è che non solo il valore di  $W_{\rm cycle} \geq 0$  cambia sensibilmente da ciclo a ciclo, ma addirittura in alcuni casi si ottiene  $W_{\rm cycle} < 0$ ! Come abbiamo detto, sono le fluttuazioni che in sistemi di scale così ridotte assumono un ruolo rilevante; ci aspettiamo che comunque si abbia  $\langle W_{\rm cycle} \rangle \geq 0$  anche in questo caso.

In seguito considereremo sempre un sistema termodinamico microscopico come quello dell'esempio precedente, il quale passa da uno stato A a uno B con una generica legge oraria  $\lambda(t)$  con  $0 \le t \le \tau$  ( $\lambda$  è un generico parametro che descrive una caratteristica del sistema, nel caso del filamento di RNA ad esempio può essere l'allungamento del filamento stesso). Una volta che un generico sistema è giunto in B non si trova strettamente parlando all'equilibrio con l'ambiente circostante, ma rilassa fino a raggiungerlo.

Se eseguiamo più volte un esperimento come quello che abbiamo citato, quindi, possiamo determimare una densità di probabilità  $\rho(W)$  del lavoro del processo termodinamico.

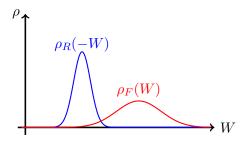
Se consideriamo però un ciclo invece di una semplice trasformazione (ossia un processo del tipo  $A \longrightarrow B \longrightarrow A$ ) allora è utile dividere il processo in una componente "forward"  $(A \longrightarrow B)$ , che avrà una sua legge oraria  $\lambda_F(t)$ , e in una "reverse"  $(B \longrightarrow A)$  che segue la legge  $\lambda_R(t) = \lambda_F(\tau - t)$ . Eseguendo un ciclo del genere, considerando le singole componenti forward e reverse in assenza di fluttuazioni si ha  $W_F \ge \Delta F$  e  $-W_R \le \Delta F$ , ossia:

$$-W_R \le \Delta F \le W_F$$

Considerando invece complessivamente tutto il ciclo:

$$W_R + W_F \ge \Delta F$$

Mantenendo sempre la distinzione fra le due componenti del processo possiamo allora determinare sia una  $\rho_F(W)$  che una  $\rho_R(W)$ ; particolarmente interessante è il confronto fra  $\rho_F(W)$  e<sup>3</sup>  $\rho_R(-W)$ . In generale infatti il grafico delle due densità di probabilità è:



Come evidenzieremo anche in seguito, il punto di sovrapposizione fra le due densità di probabilità si ha per  $W=\Delta F$ ; la differenza fra le medie delle due  $\rho$ , poi, è il lavoro medio eseguito per ciclo, e come si vede è comunque positivo.

## 6.1.1 Argomento di plausibilità per $W_{\text{cycle}} < 0$

Quello che vogliamo fare adesso è capire come sia possibile che si possa avere  $W_{\rm cycle} < 0$  in un ciclo isotermo, e darne un argomento di plausibilità. Pensiamo sempre a un sistema come quello del filamento di RNA; il sistema complessivo formato dalla molecola e dal bagno è complessivamente isolato, e quindi per esso possiamo dare una descrizione microcanonica.

Quello che immaginiamo (e in parte speriamo) possa avvenire è che in sistemi del genere grandezze come  $W_F$  e  $W_R$  diventino variabili stocastiche, che possono ovviamente fluttuare ma tali che le leggi termodinamiche continuino a valere in termini di valori medi; in altre parole, anche se in una particolare realizzazione di un processo si può avere  $W_F + W_R < 0$ , normalmente si ha comunque  $\langle W_F + W_R \rangle \geq 0$ .

In generale, quindi, consideriamo un processo ciclico  $A \longrightarrow A$ , chiamiamo  $H(\lambda,x)$  l'hamiltoniana del sistema (bagno termico incluso) con x=(q,p) punto rappresentativo nello spazio  $\Gamma$  e  $\lambda$  parametro che descrive il processo che vogliamo studiare; se  $x_0$  è il punto iniziale del sistema nello spazio  $\Gamma$ , allora il lavoro speso per compiere un qualunque ciclo risulta essere:

$$W(x_0) = H(\lambda(\tau), x_{\tau}(x_0)) - H(\lambda(0), x_0)$$

ove  $x_{\tau}$  è l'evoluto di  $x_0$  al tempo  $\tau$ . Il lavoro  $W(x_0)$  è quindi una variabile stocastica. Supponiamo che in generale si abbia  $H(\lambda(0),x_0)\leq U_0$ , con  $U_0$  generica costante, quando  $x_0\in C$ , con C un sottoinsieme dello spazio  $\Gamma$ ; facendo evolvere il sistema, i vari  $x_0\in C$  si muoveranno lungo le rispettive traiettorie nello spazio delle fasi, e complessivamente tutto l'insieme C evolverà in un altro insieme  $C_{\tau}$  (che ovviamente per il teorema di Liouville avrà lo stesso volume di C); in particolare si possono presentare due situazioni:

$$-\langle W_R \rangle = -\int_{-\infty}^{+\infty} W \rho_R(W) dW = \int_{-\infty}^{+\infty} \rho_R(-W) W dW$$

ove nell'ultimo passaggio abbiamo prima effettuato il cambio di variabile W'=-W, e poi rinominato W' con W.

<sup>&</sup>lt;sup>3</sup>Usiamo  $\rho_R(-W)$  e non  $\rho_R(W)$  perché ci interessa determinare  $-\langle W_R \rangle$  (entra in gioco nella versione "mediata" di  $-W_R \leq \Delta F \leq W_F$ , che comunque deve continuare a valere anche per un sistema microscopico), e si ha:

 $C_{\tau} = C$ : avviene quando il processo è talmente lento da essere reversibile

 $C_{\tau} \neq C$ : è la situazione che ci aspettiamo avvenga in generale. In questo caso allora esisteranno dei punti  $\overline{x} \in C$  ma  $\overline{x} \notin C_{\tau}$ .

Consideriamo uno di questi  $\overline{x}$ : necessariamente anch'esso è l'evoluto di un qualche  $\overline{x}_0 \notin C$ , e quindi la realizzazione del processo ha portato l'energia dall'avere un valore maggiore di  $U_0$  (perché appunto  $\overline{x}_0 \notin C$ ) a uno minore di  $U_0$  (perché  $\overline{x} \in C$ ). Per questa particolare traiettoria, quindi, il sistema ci sta effettivamente dando energia, e quindi il lavoro compiuto in essa è negativo.

Possono quindi esistere delle traiettorie nello spazio delle fasi per le quali il lavoro svolto sul sistema durante il processo è negativo.

Quello che ora vogliamo poter mostrare è che comunque il valor medio del lavoro svolto durante un ciclo è non negativo.

In generale, se un processo porta il parametro di lavoro dal valore  $\lambda(0) := A$  al valore  $\lambda(\tau) := B$ , il valor medio del lavoro svolto in questa traiettoria è:

$$\langle W \rangle = \int dx_0 f_A(x_0) (H_B(x_\tau) - H_A(x_0)) \ge \Delta F$$

ove  $f_A$  è la distribuzione di equilibrio delle condizioni iniziali quando  $\lambda = \lambda(0) = A$ . Per un ciclo, in particolare, si dovrà avere:

$$\langle W_{\text{cycle}} \rangle = \int dx_0 f_A(x_0) (H_A(x_\tau) - H_A(x_0)) \ge 0$$

Consideriamo il seguente valor medio:

$$\left\langle e^{-\beta W_{\text{cycle}}} \right\rangle = \int dx_0 f_A(x_0) e^{-\beta (H_A(x_\tau) - H_A(x_0))}$$

Ora, teoricamente dovremmo prendere una  $f_A(x_0)$  microcanonica, ma è più conveniente da un punto di vista dei conti prenderla canonica (tanto sappiamo che canonico e microcanonico sono equivalenti)<sup>4</sup>; allora:

$$\left\langle e^{-\beta W_{\text{cycle}}} \right\rangle = \int dx_0 \frac{e^{-\beta H_A(x_0)}}{Z_A} e^{-\beta (H_A(x_\tau) - H_A(x_0))} = \frac{1}{Z_A} \int dx_0 e^{-\beta H_A(x_\tau(x_0))}$$

Per il teorema di Liouville, ora, lo jacobiano  $|\partial x_0/\partial x_\tau|$  è pari ad 1; possiamo quindi cambiare variabile d'integrazione da  $x_0$  a  $x_\tau$ , ottenendo:

$$\left\langle e^{-\beta W_{\text{cycle}}} \right\rangle = \frac{1}{Z_A} \int dx_{\tau} e^{-\beta H_A(x_{\tau})} = \frac{Z_A}{Z_A} = 1$$

e quindi:

$$\left\langle e^{-\beta W_{\text{cycle}}} \right\rangle = 1$$
 (6.1)

 $<sup>^4</sup>$ In realtà facciamo questa scelta anche perché se prendessimo la  $f_A$  microcanonica si possono trovare controesempi "patologici" per i quali  $\langle W_{\rm cycle} \rangle < 0$ . In realtà questo avviene se usiamo il "metodo" di Boltzmann per definire il microcanonico (ossia considerando solo gli stati in una shell sottile attorno a una data energia); se invece usiamo il "metodo" di Gibbs (consideriamo tutti gli stati fino a una data energia), non sorgono problemi. Questo fatto si è inserito nel dibattito che ultimamente sta cercando di determinare quale di questi due "metodi" sia effettivamente il più consistente con la termodinamica.

L'esponenziale è una funzione convessa, e per le funzioni convesse di variabili aleatorie vale la *disuguaglianza di Jensen*, che illustriamo e dimostriamo in seguito; nel nostro caso questa asserisce che:

$$\left\langle e^{-\beta W_{\text{cycle}}} \right\rangle \ge e^{-\beta \left\langle W_{\text{cycle}} \right\rangle}$$

Usando questo risultato nella (6.1) si ha quindi:

$$e^{-\beta \langle W_{\text{cycle}} \rangle} < 1$$

ossia:

$$\langle W_{\rm cycle} \rangle \ge 0$$

come auspicavamo.

Per poter completare la "dimostrazione" bisogna quindi mostrare la validità della disuguaglianza di Jensen.

**Disuguaglianza di Jensen** Data una qualunque funzione convessa f(x), per definizione si ha  $f(x) \geq f(a) + (x-a)f'(a)$  ove a è un qualunque punto del dominio dove f è definita; nelle nostre idee x è una variabile random sulla quale poi vogliamo fare valori medi. La relazione che abbiamo appena visto vale per qualunque x, e pertanto al posto di x possiamo usare una funzione g(x) di x stesso; in questo modo si ha:

$$f(g(x)) \ge f(a) + (g(x) - a)f'(a)$$

Questa disequazione vale però anche per qualunque a, e quindi possiamo prendere  $a = \langle g(x) \rangle$ :

$$f(g(x)) \ge f(\langle g(x) \rangle) + (g(x) - \langle g(x) \rangle)f'(\langle g(x) \rangle)$$

e prendendo il valore medio ad ambo i membri:

$$\langle f(g(x)) \rangle \ge f(\langle g(x) \rangle)$$

(nel caso di nostro interesse,  $f(x) = e^{-x}$  e  $g(x) = \beta x$ ).

### 6.2 Il teorema di Crooks

In questa sezione vogliamo arrivare a dimostrare il *teorema di Crooks*, il quale asserisce che:

$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{\beta(W - \Delta F)}$$

Non si tratta di un risultato facilissimo da ottenere perché bisogna tenere conto dell'effetto del bagno termico sul nostro sistema; come vedremo possiamo affrontare la questione in più modi (in particolare due, uno che fa uso della meccanica hamiltoniana e uno che sfrutta la modellizzazione di processi markoviani).

Per poter dimostrare questo teorema è prima necessario dimostrare l'uguaglianza di Jarzynski.

### 6.2.1 Uguaglianza di Jarzynski

L'uguaglianza di Jarzynski afferma che per la componente forward di un generico processo si ha:

 $\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$ 

Supponiamo che all'istante iniziale del processo si abbia  $H(\lambda(0), x_0) = H_A(x_0)$ , e che il sistema sia in  $x_0$  in uno stato di equilibrio (canonico); si avrà allora:

$$f_A(x_0) = \frac{e^{-\beta H_A(x_0)}}{Z_A}$$

Il lavoro compiuto durante il processo è poi:

$$W(x_0) = H_B(x_\tau(x_0)) - H_A(x_0)$$

Si ha allora:

$$\left\langle e^{-\beta W} \right\rangle = \int dx_0 f_A(x_0) e^{-\beta [H_B(x_\tau(x_0)) - H_A(x_0)]} = \int dx_0 \frac{e^{-\beta H_A(x_0)}}{Z_A} e^{-\beta [H_B(x_\tau(x_0)) - H_A(x_0)]} = \int dx_0 \frac{e^{-\beta H_B(x_\tau(x_0)) - H_A(x_0)}}{Z_A} = \int dx_0 \frac{e^{-\beta H_B(x_\tau(x_0))}}{Z_A}$$

Ancora una volta, per il teorema di Liouville  $dx_0 = dx_\tau$  e quindi cambiando variabile si ha:

$$\left\langle e^{-\beta W} \right\rangle = \int dx_{\tau} \frac{e^{-\beta H_B(x_{\tau})}}{Z_A} = \frac{Z_B}{Z_A} = e^{-\beta(F_B - F_A)} = e^{-\beta\Delta F}$$

e quindi effettivamente:

$$\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$$

Con questa disuguaglianza si possono dimostrare risultati interessanti.

Ad esempio, sappiamo che in generale si dovrà avere  $\langle W \rangle \geq \Delta F$ , ma possiamo chiederci se ci sono casi in cui  $W \leq \Delta F - nk_BT$  con n intero, e con quali probabilità avvengono. Per determinarlo, tenendo conto che:

$$\left\langle e^{-\beta W} \right\rangle = \int dW \rho(W) e^{-\beta W}$$

allora:

$$\Pr(W \le \Delta F - nk_B T) = \int_{-\infty}^{F - nk_B T} \rho(W) dW$$

Se poi nell'integrale inseriamo un integrando maggiore di uno,  $e^{-\beta(W-\Delta F+nk_BT)}$ , allora:

$$\Pr(W \le \Delta F - nk_B T) \le \int_{-\infty}^{F - nk_B T} \rho(W) e^{-\beta(W - \Delta F + nk_B T)} dW$$

e possiamo "rafforzare" la disuguaglianza estendendo l'integrazione a tutto l'asse reale:

$$\Pr(W \le \Delta F - nk_B T) \le \int_{-\infty}^{+\infty} \rho(W) e^{-\beta(W - \Delta F + nk_B T)} dW =$$

$$= e^{\beta \Delta F - n} \int_{-\infty}^{+\infty} e^{-\beta W} \rho(W) dW = e^{\beta \Delta F - n} \left\langle e^{-\beta W} \right\rangle = e^{-n}$$

Pertanto possono effettivamente esserci violazioni al fatto che  $W \ge \Delta F$ , ma queste sono estremamente improbabili già per n piccoli.

#### 6.2.2 Dimostrazioni del teorema di Crooks

Passiamo ora a dimostrare il teorema di Crooks, che ricordiamo asserisce:

$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{\beta(W - \Delta F)}$$

Notare che è da questo risultato che troviamo il fatto che il crossing fra le densità di probabilità  $\rho_F$  e  $\rho_R$  avviene a  $W=\Delta F$  (in corrispondenza di questo valore il rapporto vale uno).

Vediamo la dimostrazione in due modi: la prima fa uso della meccanica hamiltoniana, mentre la seconda è particolarizzata al caso di processi markoviani.

#### Dimostrazione con la meccanica hamiltoniana

Nello stato iniziale, il sistema (che include il bagno termico) ha energia E; la densità di probabilità che il lavoro svolto durante un processo  $x_0 \longrightarrow x_f(x_0)$  che porta l'hamiltoniana da  $H_i(x_0)$  ad  $H_f(x_f(x_0))$  sia pari a W è:

$$\rho_F^E(W) = \frac{\int dx_0 \delta(H_i(x_0) - E) \delta(H_f(x_f) - H_i(x_0) - W)}{\int dx_0 \delta(H_i(x_0) - E)}$$

ove chiamiamo il denominatore, per brevità,  $\Omega_i(E)$ . Nella componente forward del processo, quindi, partendo da una certa energia E siamo in grado di raggiungere la shell di energia E+W nello spazio delle fasi, con W dipendente dalla singola traiettoria.

Nella componente reverse, invece, l'energia complessiva iniziale è E+W e il lavoro svolto -W, quindi:

$$\rho_R^{E+W}(-W) = \frac{\int d\overline{x}_f \delta(H_f(\overline{x}_f) - E - W) \delta(H_f(\overline{x}_f) - H_i(\overline{x}_i) - W)}{\int dx_f \delta(H_f(x_f) - E - W)}$$

ove anche stavolta per brevità chiamiamo  $\Omega_f(E+W)$  il denominatore; inoltre a numeratore  $\overline{x}$  è la traiettoria x percorsa a tempo invertito (abbiamo applicato un'inversione temporale).

Sfruttando la seconda delta a numeratore possiamo scrivere  $H_f(\overline{x}_f) = H_i(\overline{x}_i) + W$ , e tenendo conto che  $d\overline{x}_f = d\overline{x}_i$ :

$$\rho_R^{E+W}(-W) = \frac{\int d\overline{x}_i \delta(H_i(\overline{x}_i) - E) \delta(H_f(\overline{x}_f) - H_i(\overline{x}_i) - W)}{\Omega_f(E+W)}$$

In questo modo i numeratori di  $\rho_F^E$  e  $\rho_R^{E+W}$  sono uguali; allora:

$$\frac{\rho_F^E(W)}{\rho_R^{E+W}(-W)} = \frac{\Omega_f(E+W)}{\Omega_i(E)}$$

che ovviamente si può riscrivere come:

$$\frac{\rho_F^E(W)}{\rho_R^{E+W}(-W)} = e^{\frac{1}{k_B}(S_f(E+W) - S_i(E))}$$

Supponendo che le temperature iniziali e finali coincidano, sfruttando la definizione di energia libera (F=U-TS) allora si ha effettivamente:

$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{\beta(W - \Delta F)}$$

(ci aspettiamo anche che la dipendenza da E nelle densità di probabilità svanisca nel limite<sup>5</sup>  $E \to \infty$ ).

L'uguaglianza di Jarzynski è poi "contenuta" nel teorema di Crooks. Infatti:

$$\left\langle e^{-\beta W} \right\rangle = \int dW e^{-\beta W} \rho_F(W) = \int dW \rho_R(-W) e^{-\beta \Delta F} =$$

$$= e^{-\beta \Delta F} \underbrace{\int dW \rho_R(-W)}_{-1} = e^{-\beta \Delta F}$$

#### Dimostrazione per processi markoviani

Vediamo ora un'altra dimostrazione del teorema di Crooks.

Consideriamo un processo stocastico discreto, nel quale il sistema percorre la "traiettoria":

$$X := x_0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow \cdots \longrightarrow x_{\tau}$$

Supponendo il processo markoviano, si avrà che la probabilità di una traiettoria X è:

$$P(X) = P^{eq}(x_0)P_{\lambda_1}(x_0 \longrightarrow x_1)P_{\lambda_2}(x_1 \longrightarrow x_2)\cdots P_{\lambda_{\tau}}(x_{\tau-1} \longrightarrow x_{\tau})$$

ove il parametro di lavoro  $\lambda$  entra in gioco nelle varie probabilità perché varia nel tempo. In questo caso l'effetto del bagno termico entra in gioco perché supponiamo che le probabilità di transizione siano relative ad un sistema che per un tempo limitato è all'equilibrio con un bagno termico; insomma, supponiamo che una volta determinato un nuovo  $\lambda$  il sistema riesca ad equilibrarsi col bagno prima di fare il "salto" successivo.

Stiamo quindi supponendo che tutte le probabilità di transizione obbediscano alla legge di bilancio dettagliato, tutte con la stessa temperatura ma con  $\lambda$  diversi:

$$e^{-\beta H(x,\lambda)}P_{\lambda}(x\longrightarrow x') = e^{-\beta H(x',\lambda)}P_{\lambda}(x'\longrightarrow x)$$
 (6.2)

Ovviamente, anche in questo caso se il processo è un ciclo distinguiamo fra le sue componenti forward e reverse; si avrà  $\lambda_R(t) = \lambda_F(\tau - t)$ , e chiamiamo  $P_F(X)$  la probabilità di una traiettoria e  $P_R(X^\dagger)$  quella della sua inversa (indichiamo con  $X^\dagger$  la traiettoria X percorsa al contrario). Per ogni singola traiettoria, poi, dobbiamo definire in modo consistente il calore scambiato e il lavoro eseguito sul sistema.

Il calore Q(X) scambiato durante una traiettoria X è definito come l'insieme delle variazioni dell'hamiltoniana H conseguenti ai "salti" del sistema a parità del parametro di lavoro  $\lambda$ :

$$Q(X) = \sum_{t=1}^{\tau} \left[ H(x_t, \lambda_t) - H(x_{t-1}, \lambda_t) \right]$$
 (6.3)

Il lavoro, invece, è l'insieme delle variazioni di H dovute alla variazione di  $\lambda$ :

$$W(X) = \sum_{t=1}^{\tau} \left[ H(x_{t-1}, \lambda_t) - H(x_{t-1}, \lambda_{t-1}) \right]$$
 (6.4)

<sup>&</sup>lt;sup>5</sup>Stiamo "rimuovendo" gli effetti del bagno, ma non lo abbiamo fatto brutalmente supponendo che non ci sia interazione fra bagno e sistema.

Notiamo che da entrambe le definizioni segue che  $Q_F(X) = -Q_R(X^{\dagger})$  e  $W_F(X) = -W_R(X^{\dagger})$ .

Per la conservazione dell'energia, poi, deve valere:

$$Q(X) + W(X) = \Delta U(X) = H(x_{\tau}, \lambda_{\tau}) - H(x_{0}, \lambda_{0})$$
(6.5)

e la verifica della validità di quest'equazione segue dalla semplice somma delle equazioni (6.3) e (6.4); sommando Q(X) e W(X) si cancellano infatti tutti i termini tranne quelli che compaiono effettivamente nella (6.5):

$$Q(X) + W(X) = H(x_1, \lambda_1) - H(x_0, \lambda_1) + H(x_2, \lambda_2) - H(x_1, \lambda_2) + \dots + H(x_{\tau}, \lambda_{\tau}) - H(x_{\tau-1}, \lambda_{\tau}) + \dots + H(x_0, \lambda_1) - H(x_0, \lambda_0) + \dots + H(x_1, \lambda_2) - H(x_1, \lambda_1) + \dots + H(x_{\tau-1}, \lambda_{\tau}) - H(x_{\tau-1}, \lambda_{\tau-1})$$

e quindi effettivamente vale la (6.5):

$$Q(X) + W(X) = H(x_{\tau}, \lambda_{\tau}) - H(x_0, \lambda_0)$$

Pertanto, usando la stessa notazione sfruttata in 6.2.1 per i punti iniziale e finale di una traiettoria e della sua inversa, per una traiettoria che descrive una componente forward di un processo stocastico si ha:

$$P_F(X) = f_A^{\text{eq}}(x_0) P_{\lambda_1^F}(x_0 \longrightarrow x_1) \cdots P_{\lambda_{\tau}^F}(x_{\tau-1} \longrightarrow x_{\tau})$$

e analogamente per la componente reverse. Il rapporto delle due probabilità è:

$$\frac{P_F(X)}{P_R(X^{\dagger})} = \frac{f_A^{\text{eq}}(x_0) P_{\lambda_1^F}(x_0 \longrightarrow x_1) \cdots P_{\lambda_{\tau}^F}(x_{\tau-1} \longrightarrow x_{\tau})}{f_B^{\text{eq}}(x_{\tau}) P_{\lambda_0^R}(x_{\tau} \longrightarrow x_{\tau-1}) \cdots P_{\lambda_{\tau-1}^R}(x_1 \longrightarrow x_0)}$$

A questo punto possiamo sfruttare il bilancio dettagliato (6.2) per esprimere i vari rapporti fra le singole P (ad esempio  $P_{\lambda_1^F}/P_{\lambda_{\tau-1}^R}$ ); poiché nella (6.2) il parametro  $\lambda$  è sempre lo stesso, ognuno di questi rapporti contribuisce a una componente di Q(X). Complessivamente, quindi, si ha:

$$\frac{P_F(X)}{P_R(X^{\dagger})} = \frac{f_A^{\text{eq}}(x_0)}{f_B^{\text{eq}}(x_\tau)} e^{-\beta Q_F(X)}$$

e tenendo conto che:

$$f_A^{\text{eq}}(x_0) = \frac{e^{-\beta H(x_0, \lambda_0)}}{Z_A}$$

(e analogamente per  $f_B^{\mathrm{eq}}$ ), allora:

$$\frac{P_F(X)}{P_R(X^{\dagger})} = e^{-\beta H(x_0,\lambda_0) + \beta H(x_{\tau},\lambda_{\tau})} \cdot e^{-\beta(F_B - F_A)} \cdot e^{-\beta Q_F(X)} =$$

$$= e^{-\beta \Delta F - \beta Q_F(X) + \beta \Delta U(X)} = e^{\beta(W_F(X) - \Delta F)}$$

ove nell'ultimo passaggio abbiamo usato la (6.5).

Il risultato che abbiamo trovato dipende dalla traiettoria particolare considerata; quello che vorremmo riuscire a fare è "svincolarlo" e formularlo in generale. Innanzitutto data una  $P_F(X)$ , la  $\rho_F(W)$  sarà:

$$\rho_F(W) = \int dX P_F(X) \delta(W - W_F(X))$$

ove dX è una notazione per  $dx_0 dx_1 \cdots dx_{\tau}$ . Quindi:

$$\rho_F(W) = \int dX P_R(X^{\dagger}) e^{\beta(W_F(X) - \Delta F)} \delta(W - W_F(X))$$

A questo punto cambiamo variabile d'integrazione da X a  $X^{\dagger}$  (tenendo conto che  $dX = dX^{\dagger}$ ), sfruttiamo la  $\delta$  per sostituire  $W_F$  con W (e portiamo l'esponenziale fuori dall'integrale), e infine usiamo il fatto che  $W_F(X) = -W_R(X^{\dagger})$ :

$$\rho_F(W) = e^{\beta(W - \Delta F)} \int dX^{\dagger} P_R(X^{\dagger}) \delta(W + W_R(X^{\dagger})) = e^{\beta(W - \Delta F)} \rho_R(-W)$$

Pertanto:

$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{\beta(W - \Delta F)}$$

e quindi il teorema di Crooks effettivamente vale.

In questa dimostrazione, però, rispetto alle precedenti è molto più chiaro quale sia il ruolo del bagno termico.

### 6.3 Il "confine" fra la reversibilità e l'irreversibilità

Concludiamo questo capitolo con una considerazione sul "passaggio" dalla reversibilità microscopia all'irreversibilità macroscopica.

Un esempio molto usato quando si parla di reversibilità e irreversibilità è il seguente: se ci viene presentato un filmato (ad esempio) di un pendolo privo di *qualunque* tipo di dissipazione che oscilla, non siamo in grado di stabilire se il filmato stia procedendo in avanti o al contrario, mentre se il filmato rappresenta un altro fenomeno come la rottura di un bicchiere siamo in grado di capire in che verso è proiettato.

"Trasportiamo" quest'esempio al caso dei sistemi termodinamici fluttuanti che stiamo studiando, come l'allungamento e l'accorciamento del filamento di RNA: come possiamo capire se un dato "filmato" di un processo è un forward o un reverse proiettato al contrario? Fino a che punto riusciamo a distinguere fra le due cose? In parole un po' più generali, ci stiamo chiedendo come fare a distinguere il verso del tempo in sistemi termodinamici fluttuanti.

Possiamo provare a rispondere determinando la plausibilità L(F|X) che un dato processo sia un forward; poiché sicuramente  $L(F|X) \propto P_F(X)$ , e inoltre:

$$L(F|X) + L(R|X^{\dagger}) = 1$$

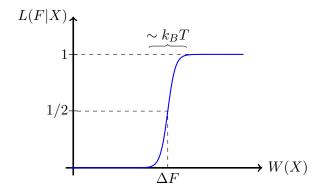
allora:

$$L(F|X) = \frac{P_F(X)}{P_F(X) + P_R(X^{\dagger})}$$

(questa è l'unica stima ragionevole che possiamo dare di L(F|X), non avendo altre informazioni su cui basarci). Per il teorema di Crooks, allora, si avrà (raccogliendo  $P_F(X)$ ):

$$L(F|X) = \frac{1}{1 + e^{-\beta(W(X) - \Delta F)}}$$

Al variare di W(X) il valore di questa funzione passa rapidamente da 0 a 1:



Pertanto, per "capire" quale sia il verso del tempo basta confrontare il lavoro eseguito con la variazione di energia libera: se  $W(X) \ll \Delta F$  il processo è sicuramente un reverse proiettato al contrario, mentre per  $W(X) \gg \Delta F$  è un forward.

Da notare che nella regione  $W \approx \Delta F$  si "perde" la cognizione del tempo: in questo caso il verso del tempo non è inequivocabile.

# Capitolo 7

# Teoria delle grandi deviazioni

La *teoria delle grandi deviazioni*, come suggerisce il nome, è una teoria che cerca di studiare fluttuazioni di grande entità in variabili stocastiche.

Una sua prima formulazione "inconscia" è contenuta nella formula di Einstein per le fluttuazioni; in 3.2.1 abbiamo visto che nel caso in cui un sistema sia descrivibile tramite  $\nu$  variabili estensive, la probabilità che esse assumano i valori  $x_1, \ldots, x_{\nu}$  all'equilibrio è:

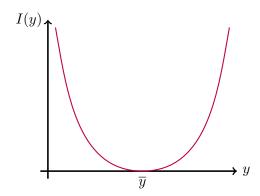
$$P^{\text{eq}}(x_1, \dots, x_{\nu}) \sim e^{\frac{1}{k_B}S(x_1, \dots, x_{\nu})}$$

Come vedremo, questa può essere vista come una prima formulazione della teoria delle grandi deviazioni.

Consideriamo una generica variabile stocastica  $Y_n$  che dipende da un parametro estensivo n (che poi manderemo all'infinito, pensando a un limite termodinamico); in generale la  $Y_n$  non sarà estensiva, e si dice che essa "soddisfa un principio di grandi deviazioni" se:

$$P(Y_n = y) \stackrel{n \to \infty}{\sim} e^{-nI(y)}$$

ove I(y) è detta "rate function". Generalmente, questo tipo di funzioni hanno la seguente forma (non necessariamente parabolica):



Per  $n \to \infty$ , allora, la  $Y_n$  tende a "concentrarsi" su  $\overline{y}$  (o meglio,  $\overline{y}$  diventa il valore di  $Y_n$  più probabile).

Questo "meccanismo" in realtà lo abbiamo già incontrato più volte, ed è fondamental-

mente quello che serve a giustificare le forme degli ensembles statistici<sup>1</sup>; è anche lo stesso ragionamento che abbiamo usato per ricavare l'espressione di Einstein<sup>2</sup> di  $P^{eq}$ .

Storicamente, la teoria delle grandi deviazioni si è sviluppata in àmbito matematico attorno agli anni '30 ed è stata sfruttata molto alla fine del secolo scorso per studiare proprietà di equilibrio; un nuovo impulso l'ha poi ricevuto negli ultimi decenni, con lo studio della termodinamica fuori equilibrio: in questo caso infatti i principî di grandi deviazioni vengono sfruttati per sopperire alla mancanza di potenziali termodinamici con delle funzioni che svolgano un ruolo analogo (prima di allora la teoria delle grandi deviazioni era semplicemente un modo alternativo per spiegare fatti che si conoscevano già).

**Teoria delle grandi deviazioni e teorema del limite centrale** All'interno dell'approccio allo studio delle grandi deviazioni è contenuto il teorema del limite centrale, o meglio il principio di grandi deviazioni può essere visto come un'"estensione" del teorema del limite centrale; vediamo brevemente come.

Data una variabile aleatoria  $Y_n$ , sviluppando I(y) attorno a  $\overline{y}$  si ha:

$$P(Y_n = y) \sim \exp \left[ -n \frac{I''(\overline{y})}{2} (y - \overline{y})^2 + n \frac{I'''(\overline{y})}{3!} (y - \overline{y})^3 + \cdots \right]$$

Definendo allora la variabile  $z := \sqrt{n} (y - \overline{y})$  (e l'idea è appunto che questa sia una somma di n variabili aleatorie  $Y_n$  indipendenti e identicamente distribuite, vedi anche 1.2.2), si ha:

$$P(z) \sim e^{-\frac{I''(\bar{y})}{2}z^2 + O(z^3/\sqrt{n})}$$

che è una gaussiana; abbiamo proprio ritrovato il teorema del limite centrale (fintantoché z non è troppo grande).

**Esempio** Consideriamo la variabile:

$$S_n = \sum_{i=1}^n \frac{\varepsilon_i}{n}$$

ove  $\varepsilon_i=\pm 1$  (è una variabile come lo spin di un modello di Ising). Vogliamo determinare la probabilità che  $S_n$  sia uguale a un dato valore +r; poiché ciascuna delle configurazioni possibili ha probabilità  $1/2^n$ , e poiché fissato r ci sono tante possibili configurazioni quanti sono i modi di ordinare n elementi in gruppi di n(1+r)/2 (è il numero delle variabili  $\varepsilon_i=+1$ ), allora:

$$P(S_n = r) = \frac{1}{2^n} \binom{n}{\frac{1}{2}n(1+r)}$$

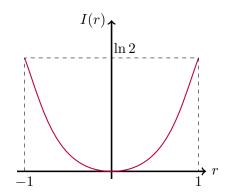
 $<sup>^1</sup>$ Tutta la meccanica statistica classica in realtà si può ottenere dalla teoria delle grandi deviazioni. Sappiamo infatti che in meccanica statistica viene assegnata una probabilità a priori agli stati microscopici nello spazio  $\Gamma$ ; si scelgono poi delle grandezze fisiche che obbediscono a princip $\hat{\Gamma}$  di grandi deviazioni nel limite termodinamico, di modo tale che nella quasi totalità dei punti disponibili nello spazio  $\Gamma$  la grandezza assume il valore che annulla la rate function I (esiste insomma un valore tipico della grandezza in esame). È questa in realtà la vera giustificazione della media di ensemble in meccanica statistica (non è sufficiente dire che una media esista, ma affinché sia significativa bisogna che sia simile al valore "tipico" della grandezza fisica).

 $<sup>^2</sup>$ Da notare, però, che mentre Einstein si è limitato a studiare piccole deviazioni (ha effettuato uno sviluppo attorno a un punto di equilibrio fino al second'ordine), l'ambizione della teoria delle grandi deviazioni è proprio quella di studiare P(y) per valori arbitrari di y.

Usando l'approssimazione di Stirling per sviluppare quest'espressione, alla fine si trova che:

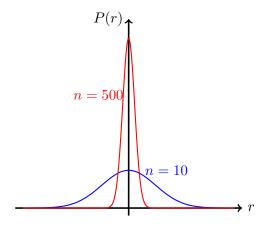
$$I(r) = \ln 2 + \frac{1+r}{2} \ln \frac{1+r}{2} + \frac{1-r}{2} \ln \frac{1-r}{2}$$

che ha un andamento di questo tipo:



Per  $n \to \infty$ , quindi, si avrà r = 0.

La distribuzione di probabilità per n crescenti ma finiti sarà dunque la seguente:



Questo è un esempio "fortunato" nel quale è possibile ricavare esplicitamente la rate function in modo semplice.

# 7.1 Teoria delle grandi deviazioni e termodinamica

Vediamo ora un altro esempio, nel quale ricaveremo la rate function, che mostra bene il collegamento fra teoria delle grandi deviazioni e termodinamica.

Consideriamo un gas di N particelle a temperatura T in un volume V (vedi figura 7.1), e consideriamo in esso un volumetto v molto più piccolo di V ma comunque macroscopico; il numero n di particelle in esso contenuto non sarà fisso, ma in generale potrà fluttuare. Vogliamo studiare  $\rho=n/v$  (ossia la densità locale nel volumetto), e in particolare determinare la probabilità  $P(n/v=\rho)$  che questa sia uguale a un dato  $\rho$ , sapendo che  $\rho_0=N/V$ .

Supponendo il gas non ideale (altrimenti il problema sarebbe banale), ma comunque con

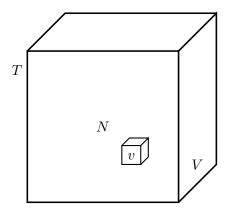


Figura 7.1: Sistema termodinamico considerato

le "giuste" proprietà (ad esempio il fatto che l'interazione fra le particelle è a corto raggio, di modo tale che il limite termodinamico esista), possiamo scrivere questa proprietà come un rapporto fra funzioni di partizione:

$$P\left(\frac{n}{v} = \rho\right) = \frac{Z_n(v, T)Z_{N-n}(V - v, T)}{Z_N(V, T)}$$

(ove stiamo implicitamente assumendo che l'interazione fra il volumetto e il resto del sistema sia trascurabile).

Sappiamo che  $Z_N(V,T) = \exp[-\beta V f(\rho_0,T)]$ , ove f è l'energia libera per unità di volume (e analogamente per  $Z_n(v,T)$ ); pertanto:

$$P\left(\frac{n}{v} = \rho\right) = \frac{e^{-\beta v f(\rho, T)} e^{-\beta(V - v) f\left(\frac{N - n}{V - v}, T\right)}}{e^{-\beta V f(\rho_0, T)}} \tag{7.1}$$

Ora, tenendo conto che:

$$f\left(\frac{N-n}{V-v},T\right) = f(\rho_0,T) + f_{\rho}(\rho_0,T)\left(\frac{N-n}{V-v} - \rho_0\right) + \cdots$$

allora prendendo il logaritmo della (7.1) ed espandendo, si ha:

$$\ln P\left(\frac{n}{v} = \rho\right) = -\beta v f(\rho, T) - \beta (V - v) f(\rho_0, T) -$$

$$-\beta (V - v) f_\rho(\rho_0, T) \left(\frac{N - n}{V - v} - \rho_0\right) + \beta V f(\rho_0, T) =$$

$$= -\beta v [f(\rho, T) - f(\rho_0, T)] - \beta (N - n) f_\rho(\rho_0, T) + \beta \rho_0 (V - v) f_\rho(\rho_0, T)$$

Tenendo infine conto che  $N=\rho_0 V$  e  $n=\rho v$  si ottiene:

$$\ln P\left(\frac{n}{v} = \rho\right) = -v\beta[f(\rho, T) - f(\rho_0, T)] + v\beta(\rho - \rho_0)f_\rho(\rho_0, T)$$

che è effettivamente nella forma  $-v\beta I(\rho)$ , con:

$$I(\rho) = f(\rho, T) - f(\rho_0, T) - (\rho_0 - \rho)f_{\rho}(\rho_0, T)$$

Come ci aspettiamo (e com'è ragionevole che sia), questa rate function si annulla per  $\rho = \rho_0$ . In questo caso, poi, possiamo vedere esplicitamente il collegamento che sussiste

fra la rate function e i potenziali termodinamici all'equilibrio.

Nel caso di sistemi fuori dall'equilibrio il principio di grandi deviazioni può essere usato generalizzando quello che abbiamo appena visto definendo un *rate functional*  $I[\rho(\vec{r})]$ ; anche in questo caso è comunque necessario conoscere esplicitamente la forma dell'energia libera (cosa spesso assolutamente non semplice).

#### 7.2 Teorema di Gärtner-Ellis

In questa sezione vogliamo arrivare a illustrare (senza dimostrare) il *teorema di Gärtner-Ellis*, che permette di stabilire se una data variabile soddisfa il principio di grandi deviazioni e come costruirne la rate function; inoltre, stabilisce un collegamento strettissimo fra teoria delle grandi deviazioni e meccanica statistica, spiegando anche perché per passare da un ensemble all'altro bisogna usare le trasformate di Legendre.

Cominciamo con un esempio che ci servirà per comprendere meglio il teorema, e che riprenderemo dopo averlo illustrato.

Consideriamo la variabile:

$$S_n = \frac{1}{n} \sum_{i=1}^n X_i$$

ove le  $X_i$  sono variabili gaussiane indipendenti ed identicamente distribuite:

$$P(X_i = x) = \frac{\exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right]}{\sqrt{2\pi\sigma^2}}$$

Poiché  $S_n$  è una somma di variabili indipendenti, la densità di probabilità che sia uguale a un dato s sarà data da:

$$P(S_n = s) = \int dx_1 \cdots dx_n \delta\left(S_n - \frac{1}{n} \sum_{i=1}^n x_i\right) P(x_1) \cdots P(x_n)$$

Ora, tenendo conto che sommando variabili gaussiane indipendenti si ottiene una variabile che ha media pari alla somma delle medie, e varianza uguale alla somma delle varianze, per il teorema del limite centrale si ha:

$$P(nS_n = s) = \frac{\exp\left[-\frac{(s - n\mu)^2}{2n\sigma^2}\right]}{\sqrt{2\pi n\sigma^2}}$$

e quindi:

$$P(S_n = s) = \sqrt{\frac{n}{2\pi\sigma^2}} e^{-\frac{n(s-\mu)^2}{2\sigma^2}}$$

In questo caso, allora, la rate function del sistema è:

$$I(s) = \frac{(s-\mu)^2}{2\sigma^2}$$
 (7.2)

Mostriamo a questo punto il teorema di Gärtner-Ellis.

Consideriamo una variabile stocastica  $Y_n$ , e definiamo la sua "scaled cumulant generating function" come:

$$\lambda(k) := \lim_{n \to \infty} \frac{1}{n} \ln \left\langle e^{nkY_n} \right\rangle$$

ove k è una variabile reale. Se possiamo stabilire che  $\lambda(k)$  è definita su tutto l'asse reale ed è differenziabile ovunque, allora  $Y_n$  soddisfa il principio di grandi deviazioni, e si ha:

$$I(y) = \sup_{k \in \mathbb{R}} (ky - \lambda(k))$$

Se poi  $\lambda(k)$  è anche strettamente convessa (ossia  $\lambda''(k) > 0$  ovunque), quest'ultimo risultato è riscrivibile come:

$$I(y) = k(y)y - \lambda(k(y))$$

ove k(y) è l'espressione di k in funzione di y che si trova dall'inversione di  $y = \lambda'(k)$ ; si tratta proprio di una trasformata di Legendre.

Applichiamo ora questo teorema all'esempio precedente; ci aspettiamo ovviamente di ritrovare la (7.2).

Dobbiamo quindi calcolare:

$$\lambda(k) = \lim_{n \to \infty} \frac{1}{n} \ln \left\langle e^{k \sum_{i=1}^{n} X_i} \right\rangle = \lim_{n \to \infty} \frac{1}{n} \ln \left( \left\langle e^{kX} \right\rangle^n \right) = \ln \left\langle e^{kX} \right\rangle$$

Ora, si ha:

$$\left\langle e^{kX} \right\rangle = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} dx e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{kx}$$

Cambiando variabile in  $x' = x - \mu$  e completando poi il quadrato a esponente, si ottiene:

$$\left\langle e^{kX} \right\rangle = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} dx' e^{-\frac{x'^2}{2\sigma^2}} e^{kx'} e^{k\mu} = e^{k\mu} e^{\frac{k^2\sigma^2}{2}}$$

e pertanto:

$$\lambda(k) = \ln\left(e^{k\mu}e^{\frac{k^2\sigma^2}{2}}\right) = \mu k + \frac{k^2\sigma^2}{2}$$

che è effettivamente definita su tutto  $\mathbb{R}$  e differenziabile ovunque. Essendo anche strettamente convessa, la rate function sarà  $I(s) = k(s)s - \lambda(k(s))$ , ove k(s) è la soluzione di  $\lambda'(k) = s$ :

$$\mu + k\sigma^2 = s$$
  $\Rightarrow$   $k(s) = \frac{s - \mu}{\sigma^2}$ 

Pertanto:

$$I(s) = s\frac{s-\mu}{\sigma^2} - \mu\frac{s-\mu}{\sigma^2} + \frac{\sigma^2}{2}\left(\frac{s-\mu}{\sigma^2}\right)^2 = \frac{(s-\mu)^2}{2\sigma^2}$$

che è esattamente il risultato della (7.2) (come ci aspettavamo).

# 7.3 Teoria delle grandi deviazioni fuori dall'equilibrio

Fin'ora abbiamo visto esempi e applicazioni della teoria delle grandi deviazioni a sistemi che si trovano all'equilibrio; come sappiamo, però, questa teoria è stata fortemente sfruttata per studiare sistemi fuori dall'equilibrio. Cerchiamo quindi di capire come questa entri in gioco in sistemi del genere.

In generale, se all'equilibrio  $x \in \Gamma$  e  $M_n$  è una generica quantità fisica, la probabilità che questa sia uguale a un dato valore m è:

$$P(M_n = m) = \int dx \delta(M_n(x) - m)P(x)$$

e per  $n \to \infty$  questa si "concentra" su un certo valore, nel senso e nel modo che abbiamo già visto.

Fuori dall'equilibrio, invece di pensare ad una probabilità di stati microscopici conviene considerare probabilità di traiettorie (del tipo  $X := x_1 \longrightarrow x_2 \longrightarrow \cdots \longrightarrow x_\tau$ , come abbiamo già visto) nello spazio delle fasi; con queste possiamo costruire delle osservabili dinamiche del tipo:

$$M_n(t) = \frac{1}{n} \sum_{i=1}^n f(x_i)$$

ove f è una generica funzione.

Osservabili molto interessanti sono quelle la cui estensività è data dal tempo stesso, ossia che sono nella forma:

$$F_{\tau}(X) = \frac{1}{\tau} \sum_{i=1}^{\tau} f(x_i)$$

(e ci aspettiamo che per grandi  $\tau$  questa soddisfi il principio di grandi deviazioni). Un tipico esempio di grandezze di questo tipo sono le correnti (ove però compariranno termini del tipo  $f(x_i,x_{i-1})$ ), che sono grandezze che compaiono tipicamente in stati stazionari fuori equilibrio.

Cerchiamo quindi di studiare le proprietà di questi stati stazionari fuori equilibrio.

Usiamo un formalismo in cui il tempo è continuo, e le configurazioni C del sistema supponiamo che siano numberabili; una possibile evoluzione del sistema è allora una "traiettoria" del tipo  $C_0 \xrightarrow{t_1} C_1 \xrightarrow{t_2} C_2 \xrightarrow{t_3} \cdots \xrightarrow{t_{\tau}} C_{\tau}$ .

Detta<sup>3</sup>  $P_t(C)$  la probabilità di una di queste "traiettorie" (e  $P_t(C^{\dagger})$  quella della sua inversa), dalla master equation si ha:

$$\frac{d}{dt}P_t(C) = MP_t(C)$$

(ove M è la matrice che in 5.4.2 avevamo chiamato W).

La condizione minimale per avere equilibrio termico, a questo punto, è:

$$\sum_{C' \neq C} M(C, C') e^{-\frac{E(C')}{k_B T}} = e^{-\frac{E(C)}{k_B T}} \sum_{C' \neq C} M(C', C)$$

(ove  $e^{-\beta E(C)}/Z=P_{\rm eq}(C)$ ) che è meno restrittiva della legge di bilancio dettagliato, perché non abbiamo supposto che M sia simmetrica, conseguenza della reversibilità microscopica. Se supponiamo che questa reversibilità ci sia, allora possiamo riscrivere l'equazione come:

$$M(C, C')P_{eq}(C') = M(C', C)P_{eq}(C)$$
 (7.3)

e dividendo ambo i membri per  $P_{\text{eq}}^{1/2}(C)P_{\text{eq}}^{1/2}(C')$ :

$$P_{\text{eq}}^{-1/2}(C)M(C,C')P_{\text{eq}}^{1/2}(C') = P_{\text{eq}}^{-1/2}(C')M(C',C)P_{\text{eq}}^{1/2}(C)$$
(7.4)

Se a questo punto definiamo le matrici:

$$Q = \begin{pmatrix} P_{\text{eq}}^{1/2}(C_1) & 0 & \cdots & 0 \\ 0 & P_{\text{eq}}^{1/2}(C_2) & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & \cdots & P_{\text{eq}}^{1/2}(C_{\tau}) \end{pmatrix}$$

 $<sup>^3</sup>$ Il t a pedice indica semplicemente che P dipende dal tempo, oltre che da C.

e:

$$\mathcal{W} = Q^{-1}MQ$$

la (7.4) si può riscrivere come:

$$\mathcal{W}^\dagger = \mathcal{W}$$

ossia, W è hermitiana; la reversibilità microscopica dà quindi la possibilità di definire una matrice più facile da trattare, cosa non possibile in generale.

Ora, si può scrivere:

$$\frac{d}{dt}P_t(C) = \sum_{C'} \left[ M(C, C')P_t(C') - M(C', C)P_t(C) \right] = \sum_{C'} J_t(C, C') \tag{7.5}$$

ove abbiamo definito J, il flusso netto di probabilità da C' a C; la variazione di P(C(t)) è quindi data dalla somma di tutti questi possibili flussi. Ora, per definizione si ha:

$$J_t(C,C') = -J_t(C',C)$$

e dato che la probabilità totale si conserva, vale anche:

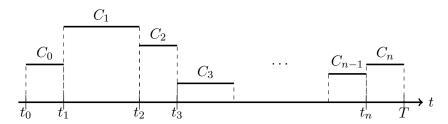
$$\sum_{C'} J_t(C, C') = 0 (7.6)$$

Per uno stato stazionario, poi, si dovrà avere:

$$\sum_{C'} J_{\infty}(C, C') = 0$$

È utile precisare cos'è in questo contesto la  $P_t(C)$ .

"Graficamente", l'evoluzione della traiettoria è rappresentabile nel seguente modo:



Per determinare la probabilità  $P_t(C)$  della traiettoria dobbiamo innanzitutto determinare la probabilità che il sistema permanga in  $C_0$  fino a  $t_1$ ; in generale, la probabilità che la traiettoria non cambi in un tempo dt infinitesimo è  $1 + M(C_0, C_0)dt$  (nota: M(C, C) < 0). Sfruttando poi la markovianità del processo, la probabilità che il sistema permanga in  $C_0$  fino a un tempo  $t_1$  (non infinitesimo) è:

$$[1 + M(C_0, C_0)dt]^{\frac{t_1}{dt}} \xrightarrow{dt \to 0} e^{M(C_0, C_0)t_1}$$

Al tempo  $t_1$  il sistema "salta" nella configurazione  $C_1$ , e questo avviene con una probabilità  $M(C_1, C_0)dt_1$ . Proseguendo quindi la traiettoria, alla fine si avrà:

$$P_t(C) = P_{eq}(C_0)e^{M(C_0,C_0)t_1}M(C_1,C_0)dt_1e^{M(C_1,C_1)(t_2-t_1)}M(C_2,C_1)dt_2\cdots$$

$$\cdots dt_nM(C_n,C_{n-1})e^{M(C_n,C_n)(T-t_n)}$$

<sup>&</sup>lt;sup>4</sup>Nota: non ci preoccupiamo della normalizzazione di queste probabilità perché alla fine ci interesseranno solo *rapporti* fra probabilità, quindi gli eventuali fattori di normalizzazione si semplificano comunque.

Considerando la traiettoria inversa  $C^{\dagger}(t)$  (per la quale vale  $C^{\dagger}(T-t) = C(t)$ ) si ha invece:

$$P_t(C^{\dagger}) = P_{\text{eq}}(C_n)e^{M(C_n, C_n)(T - t_n)}dt_n M(C_{n-1}, C_n)e^{M(C_{n-1}, C_{n-1})(t_n - t_{n-1})} \cdots \cdots dt_1 M(C_0, C_1)e^{M(C_0, C_0)(t_1 - t_0)}$$

Consideriamo il rapporto fra queste due probabilità; tutti i  $dt_i$  e gli esponenziali si semplificano (perché compaiono in entrambe), e quindi "sopravvivono" solo i rapporti fra le  $P_{\rm ed}$  e le M:

$$\frac{P_t(C)}{P_t(C^{\dagger})} = \frac{P_{\text{eq}}(C_0)}{P_{\text{eq}}(C_n)} \cdot \frac{M(C_1, C_0)}{M(C_0, C_1)} \cdots \frac{M(C_n, C_{n-1})}{M(C_{n-1}, C_n)}$$

Se però sfruttiamo la (7.3) (che ovviamnte vale solo all'equilibrio) allora nel rapporto compaiono delle  $P_{eq}$  che complessivamente si semplificano:

$$\frac{P_t(C)}{P_t(C^{\dagger})} = \frac{P_{\text{eq}}(C_0)}{P_{\text{eq}}(C_n)} \cdot \frac{P_{\text{eq}}(C_1)}{P_{\text{eq}}(C_0)} \cdot \frac{P_{\text{eq}}(C_2)}{P_{\text{eq}}(C_1)} \cdot \cdot \cdot \frac{P_{\text{eq}}(C_n)}{P_{\text{eq}}(C_{n-1})} = 1$$
 (7.7)

## 7.3.1 Entropia in processi markoviani

In questa sezione vogliamo cercare di definire e studiare l'entropia di un processo stocastico, e come possiamo determinarla in termini di quantità che conosciamo.

In termodinamica la variazione di entropia di un sistema è, in termini generali, composta da due contributi, uno generato internamente dal sistema e uno causato dall'interazione con l'ambiente esterno:

$$dS = d_i S + d_e S$$

(ove "i" ed "e" stanno proprio per "interno" ed "esterno"). Dal secondo principio della termodinamica sappiamo che  $d_iS \geq 0$  (e può essere uguale a zero solo per processi reversibili), mentre non ci sono vincoli su  $d_eS$  (può essere sia positivo che negativo, o anche nullo).

Se un sistema è strettamente e adiabaticamente isolato, allora:

$$dS = d_i S > 0 \qquad \Rightarrow \qquad dS > 0$$

mentre se il sistema è accoppiato a un serbatoio termico a temperatura T:

$$dS = d_i S + \frac{\delta Q}{T} \ge \frac{\delta Q}{T} \qquad \Rightarrow \qquad dS \ge \frac{\delta Q}{T}$$

Ora, come possiamo identificare l'entropia e i suoi due contributi in un processo stocastico markoviano?

Dato che il sistema è descritto dalla probabilità  $P_t(C)$ , possiamo definire l'entropia del sistema come:

$$S(t) = -\sum_{C} P_t(C) \ln P_t(C)$$

che è un'entropia adimensionale dipendente dal tempo (è l'entropia di Shannon). Considerando l'equazione master (7.5):

$$\frac{dP_t(C)}{dt} = \sum_{C'} J_t(C, C')$$

possiamo scrivere la derivata temporale di S(t) come:

$$\frac{dS(t)}{dt} = -\sum_{C} \frac{dP_t(C)}{dt} \left[ \ln P_t(C) + 1 \right] = -\sum_{C,C'} J_t(C,C') \left[ \ln P_t(C) + 1 \right] = -\sum_{C,C'} J_t(C,C') \ln P_t(C)$$

$$= -\sum_{C,C'} J_t(C,C') \ln P_t(C)$$

ove nell'ultimo passaggio abbiamo sfruttato la conservazione della probabilità (7.6). Ora, poiché  $J_t(C, C') = -J_t(C', C)$  possiamo anche scrivere:

$$\frac{dS(t)}{dt} = \sum_{C,C'} J_t(C',C) \ln P_t(C)$$

e rinominando gli indici sommati (che sono muti):

$$\frac{dS(t)}{dt} = \sum_{C,C'} J_t(C,C') \ln P_t(C')$$

Abbiamo quindi due espressioni equivalenti per dS/dt; varrà allora:

$$\frac{dS(t)}{dt} = \frac{1}{2} \sum_{C,C'} J_t(C,C') \left( \ln P_t(C') - \ln P_t(C) \right) = \frac{1}{2} \sum_{C,C'} J_t(C,C') \ln \left( \frac{P_t(C')}{P_t(C)} \right) = 
= \frac{1}{2} \sum_{C,C'} J_t(C,C') \ln \left( \frac{P_t(C')M(C,C')}{P_t(C)M(C',C)} \right) - \frac{1}{2} \sum_{C,C'} J_t(C,C') \ln \left( \frac{M(C,C')}{M(C',C)} \right)$$

ove nell'ultimo passaggio abbiamo moltiplicato e diviso per la stessa quantità (ossia M(C,C')/M(C',C)).

Notiamo quindi come in generale dS/dt è effettivamente scrivibile come somma di due termini, e le loro proprietà corrispondono a quelle che ci aspettiamo da  $d_iS$  (il primo termine) e  $d_eS$  (il secondo).

Consideriamo ad esempio il primo termine: se vale la legge di bilancio dettagliato (7.3) (e quindi vale la reversibilità) è sicuramente nullo, e si può anche dimostrare (ma noi non lo facciamo) che è sempre non negativo.

Sul secondo, invece, non esistono vincoli, e più avanti vedremo come lo si può interpretare proprio come lo scambio di entropia con l'ambiente esterno.

#### 7.3.2 La legge generalizzata di bilancio dettagliato

Vogliamo ora capire come si può generalizzare la legge di bilancio dettagliato in situazioni fuori dall'equilibrio.

Prima di tutto rivediamo brevemente cosa accade all'equilibrio, di modo da introdurre una notazione che ci servirà in seguito. Consideriamo quindi un sistema (descritto da un processo stocastico) in contatto con un bagno termico a temperatura T. In generale le configurazioni di un processo stocastico C hanno un'energia E(C) associata, e nella situazione che stiamo considerando (in cui il nostro sistema è descritto da un ensemble canonico) si ha:

$$P_{\text{eq}}(C) = \frac{e^{-\frac{E(C)}{k_B T}}}{Z}$$

e la legge di bilancio dettagliato asserisce che:

$$M(C', C)e^{-\frac{E(C)}{k_BT}} = M(C, C')e^{-\frac{E(C')}{k_BT}}$$

ora, nella transizione  $C \to C'$  l'energia del sistema varia da E(C) a E(C'), e possiamo quindi definire la quantità:

$$\varepsilon \coloneqq E(C') - E(C)$$

che non è altro che l'energia che il sistema scambia col serbatoio nella transizione. Si ha allora:

$$M(C', C) = e^{-\frac{E(C') - E(C)}{k_B T}} M(C, C') = e^{-\varepsilon/k_B T} M(C, C')$$

e riscriviamo quest'equazione con questa nuova notazione (che vedremo sarà utile in seguito):

$$M_{\varepsilon}(C',C) = e^{-\varepsilon/k_B T} M_{-\varepsilon}(C,C') \tag{7.8}$$

Consideriamo ora invece una situazione in cui il nostro sistema è in uno stato stazionario ma fuori equilibrio, in particolare supponiamo che sia in contatto con *due* bagni termici, alle temperature  $T_a$  e  $T_b$ .

In questo caso ci sarà chiaramete un'equazione del tipo master che regola il processo che descrive il nostro sistema, ma quello che non sappiamo e vogliamo capire è se per essa esiste un vincolo come la (7.8), ed eventualmente che forma ha.

Ora, il sistema complessivo composto dai due serbatoi e dal sistema propriamente detto è complessivamente isolato, e tende spontaneamente ad equilibrarsi (col passare del tempo il calore passa, attraverso il sistema, da un serbatoio all'altro fino a che entrambi non hanno la stessa temperatura<sup>5</sup>); essendo isolato, però, per esso all'equilibrio varrà la legge di bilancio dettagliato.

Se quindi chiamiamo  $P(\{C, E_a, E_b\} \to \{C', E'_a, E'_b\})$  la probabilità che il sistema passi dalla configurazione C a C' e che le energie dei due serbatoi varino, rispettivamente, da  $E_a$  ad  $E'_a$  e da  $E_b$  a  $E'_b$ , allora si avrà<sup>6</sup>:

$$e^{-S_a(E_a)/k_B}e^{-S_b(E_b)/k_B}P(\{C, E_a, E_b\} \to \{C', E_a, E_b\}) =$$

$$= e^{-S_a(E'_a)/k_B}e^{-S_b(E'_b)/k_B}P(\{C', E'_a, E'_b\} \to \{C, E'_a, E'_b\})$$

ove, ad ambo i membri, il prodotto degli esponenziali è il numero degli stati dei due serbatoi con le energie di partenza, mentre le probabilità di transizione sono a energie dei serbatoi fissate.

A questo punto, portiamo tutti gli esponenziali al secondo membro; per il serbatoio con energia  $E_a$  si ha:

$$e^{-\frac{1}{k_B}(S_a(E_a') - S_a(E_a))} \sim e^{-\frac{1}{k_BT_a}(E_a' - E_a)} = e^{-\varepsilon_a/k_BT_a}$$

Ricordare poi che, poiché il sistema è complesivamente isolato, si avrà  $E(C) + E_a + E_b = E(C') + E'_a + E'_b$ .

<sup>&</sup>lt;sup>5</sup>Ovviamente, dato che i serbatoi termici vengono come al solito considerati come molto più grandi del sistema considerato, o comunque con un'elevatissima capacità termica, questo processo di equilibrazione è *lunghissimo*, e all'atto pratico se si osserva il sistema per periodi di tempo ragionevoli questo "appiattimento" della differenza di temperatura fra i due serbatoi è impossibile da misurare.

<sup>&</sup>lt;sup>6</sup>Notare che in realtà  $\{C, E_a, E_b\} \rightarrow \{C', E'_a, E'_b\}$  non è una singola transizione, ma un insieme di transizioni: stiamo infatti specificando solo l'energia dei serbatoi termici, non il loro stato, e possono ovviamente esserci più stati ai quali corrisponde la stessa energia.

ove nel primo passaggio abbiamo espanso l'entropia attorno al suo massimo, e sfruttato il fatto che i serbatoi hanno una capacità termica divergente nel limite termodinamico per trascurare gli ordini superiori; una relazione analoga varrà anche per l'altro serbatoio. Pertanto, rinominando le probabilità di transizione si ha<sup>7</sup>:

$$M_{\varepsilon_a,\varepsilon_b}(C',C) = e^{-\varepsilon_a/k_B T_a} e^{-\varepsilon_b/k_B T_b} M_{-\varepsilon_a,-\varepsilon_b}(C,C')$$
(7.9)

che è la legge di bilancio dettagliato generalizzata.

Una relazione assolutamente analoga la si può ottenere supponendo che il nostro sistema sia a contatto con dei "serbatoi di particelle" a potenziali chimici diversi; in altre parole possiamo supporre che il nostro sistema scambi particelle con questi serbatoi, i quali hanno due diverse fugacità  $z_a$  e  $z_b$ . La differenza fra i potenziali chimici dei serbatoi innesca quindi una corrente di materia; ragionando esattamente come abbiamo appena fatto, e chiamando  $q_a$  e  $q_b$  le quantità di particelle scambiate fra il sistema e i due serbatoi nella transizione  $C \to C'$  del sistema, si ottiene:

$$M_{q_a,q_b}(C',C) = z_a^{q_a} z_b^{q_b} M_{-q_a,-q_b}(C,C')$$
(7.10)

Ora, dato che sistema e serbatoi si scambiano particelle potremmo immaginare che in realtà si scambino anche energia (perché comunque entrano in contatto termico); in generale questo avviene, ma come vedremo fra poco questo risultato si applica in modo molto naturale anche a modelli dove l'energia di un particolare stato non è neanche definita (mentre l'entropia sì, e dipende solo dalla configurazione del sistema).

Concludiamo con una piccola nota.

Abbiamo detto alla fine di 7.3.1 che il contributo esterno alla variazione di entropia di un processo markoviano nel tempo è:

$$\frac{d_e S(t)}{dt} = -\frac{1}{2} \sum_{C,C'} J_t(C,C') \ln \left( \frac{M(C,C')}{M(C',C)} \right)$$

Dalla (7.9) si ha che il logaritmo contenuto in quest'espressione è uguale a:

$$\ln\left(\frac{M(C,C')}{M(C',C)}\right) = \frac{\varepsilon_a}{k_B T_a} + \frac{\varepsilon_b}{k_B T_b}$$

che, a meno della costante di Boltzmann, è proprio l'entropia scambiata dal sistema coi due serbatoi. Vediamo quindi come effettivamente la (7.9) sia in accordo con l'interpretazione che abbiamo dato a questo contributo di dS/dt.

#### 7.3.3 Modelli SSEP, ASEP e TASEP

Ogni àmbito della Fisica Statistica si è sviluppato a partire dallo studio di uno o più modelli adeguati, più o meno trattabili matematicamente e soprattutto tali da poter far luce sul fenomeno studiato. Nel caso delle transizioni di fase, ad esempio, questo modello fondamentale è stato il modello di Ising.

<sup>&</sup>lt;sup>7</sup>La notazione che abbiamo introdotto nella (7.8) è molto utile in questo caso, perché permette proprio di mettere in evidenza le energie scambiate dal sistema coi singoli serbatoi.

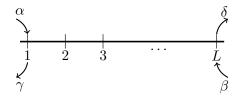


Figura 7.2: Impostazione del modello SSEP

Nell'àmbito dello studio dei processi stazionari fuori dall'equilibrio i modelli che hanno permesso di studiarne le caratteristiche sono il SSEP ("simple symmetric exclusion process"), l'ASEP ("asymmetric simple exclusion process") e il TASEP ("totally antisymmetric simple exclusion process"); questi fanno parte di una famiglia più ampia di modelli, detti generalmente "driven lattice gases".

Questa classe di modelli sono stati inizialmente introdotti per studiare problemi di tipo biofisico, ma si sono rivelati utilissimi per lo studio di problemi fondamentali di meccanica statistica come le proprietà di stati stazionari fuori equilibrio (intuiamo anche come abbiano un'enorme potenzialità applicativa).

Vediamo come sono definiti.

#### Il modello SSEP

Consideriamo (per semplicità) un reticolo unidimensionale di L siti, nei quali in un dato istante può essere presente non più di una particella (da qui il termine "exclusion" nel nome del modello); supponiamo poi che i siti al bordo del reticolo (il primo e l'L-esimo, in questo caso) possano scambiare particelle con un serbatoio (in generale ognuno col proprio rate di assorbimento o emissione, vedi figura 7.2). Quante sono allora le possibili configurazioni del sistema?

In generale una configurazione C è data da  $C=(\tau_1,\tau_2,\ldots,\tau_L)$ , ove i  $\tau_i=\{0,1\}$  sono i numeri di occupazione dei vari siti; poiché il numero di particelle contenute nel reticolo è variabile (da nessuna a L, ovviamente), le possibili configurazioni del sistema sono  $2^L$ . Supponiamo poi che in un intervallo di tempo infinitesimo dt ogni particella possa "saltare" in un sito primo vicino (ovviamente a meno che non sia già occupato) con probabilità  $p=q=\frac{1}{2}dt$ , sia a destra che a sinistra (da qui il termine "symmetric" nel nome del modello); se però una particella si trova al bordo del reticolo allora potrà essere assorbita con una probabilità che dipende dal rate di assorbimento del serbatoio stesso ( $\delta dt$  per il sito L-esimo nel caso rappresentato in figura 7.2), mentre viceversa se un sito al bordo del reticolo è vuoto questo potrà essere occupato con una probabilità che dipende dal rate di assorbimento del serbatoio ( $\alpha dt$  per il primo sito sempre nel caso della figura 7.2).

Ciò che si fa con questo modello è studiare le correnti e i profili di densità delle particelle al variare dei rate di assorbimento/emissione dei serbatoi (ovviamente, sempre nel caso di figura 7.2, se  $\alpha = \delta$  e  $\beta = \gamma$  allora non ci sono correnti nette e il sistema è all'equilibrio).

Usando varie tecniche si è riusciti a studiare e risolvere la dinamica di stati stazionari (ma ovviamente fuori equilibrio per via della presenza delle correnti) per questo tipo di modelli; è stato addirittura visto che possono verificarsi "transizioni di fase dinamiche", ossia passaggi fra (ad esempio) regimi di alta o bassa densità che presentano non analiti-

cità tipiche delle transizioni di fase.

Sistemi di questo tipo soddisfano alla legge generalizzata di bilancio dettagliato (7.10); notare anche come nella discussione di questo modello non è mai entrata in gioco l'eventuale energia delle particelle o delle configurazioni del reticolo.

#### I modelli ASEP e TASEP

Il modello ASEP ha lo stesso setup, ma le probabilità pdt e qdt che in un intervallo dt la particella "salti" in un sito primo vicino (ovviamente libero) rispettivamente a destra oppure a sinistra non coincidono più<sup>8</sup>.

In questo caso allora si instaura *spontaneamente* una corrente di particelle nel sistema (è come se sottoponesimo le particelle del modello precedente a un potenziale esterno che favorisce il loro spostamento in una direzione anziché nell'altra).

In questo caso è allora più comodo "rimuovere" i serbatoi dal sistema, e porre condizioni periodiche al contorno sul reticolo; come noto, il sistema può allora essere pensato come definito su un reticolo disposto come una circonferenza. In questo caso però il numero di particelle N è costante, e quindi il numero di possibili configurazioni del sistema sarà dato da  $\binom{L}{N}$ .

Si può anche poi dimostrare che nello stato stazionario del sistema si ha che tutte le configurazioni sono equiprobabili.

Consideriamo infatti una data configurazione di un modello TASEP (per semplicità, i ragionamenti per l'ASEP sono analoghi); in essa ci saranno varie "isole" di siti occupati. Possiamo allora determinare il "rate di uscita" da una data configurazione: se supponiamo che le particelle possano muoversi solo in senso antiorario, considerata un'isola di siti occupati il sistema potrà passare ad un'altra configurazione se il sito "più" antiorario di tutti effettua uno step (separandosi dall'isola). Questo ragionamento vale per tutte le isole, quindi il rate totale di uscita da una data configurazione sarà la probabilità p di effettuare un salto per il numero delle isole di siti occupati $^9$ . D'altra parte però il "rate di ingresso" nella nostra configurazione iniziale è uguale, in quanto considerata una singola isola, questa proverrà da una configurazione dove il sito "meno" antiorario di tutti si trova separato dall'isola ed effettua un salto per unircisi. Insomma, poiché c'è questo "equilibrio" fra i rates di "ingresso" e "uscita" in una data configurazione, tutte le configurazioni sono equiprobabili.

Nel caso del modello TASEP, invece, i "salti" sono permessi solo in una direzione (e quindi la probabilità di spostarsi nell'altra è nulla).

**Applicazioni biofisiche** Prima di passare al prossimo argomento vediamo un attimo perché sono stati introdotti questi modelli.

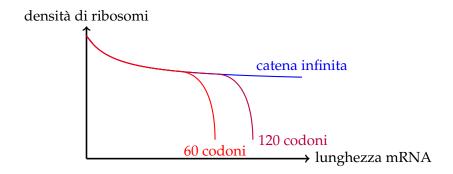
In biologia cellulare, l'mRNA è una molecola che codifica l'informazione del DNA attraverso una sequenza di unità, dette "codoni", che servono per sintetizzare proteine; il ribosoma è la molecola che "agganciandosi" all'mRNA elabora l'informazione dei codoni per capire qual è l'amminoacido da usare per sintetizzare una proteina; una volta

<sup>&</sup>lt;sup>8</sup>Nota: dato che le probabilità vere e proprie di "saltare" in una delle due possibili direzioni sono pdt e qdt, p e q non sono singolarmente normalizzate. In altre parole, non è necessariamente detto che p+q=1 (anche se generalmente per comodità vengono rinormalizzate di modo tale che questa relazione valga).

<sup>&</sup>lt;sup>9</sup>Nota: anche un singolo sito isolato costituisce un'isola a sé stante.

che l'amminoacido selezionato è stato usato, il ribosoma passa al codone successivo (non può tornare indietro e non può essere "scavalcato" da altri ribosomi). Alla fine, raggiunto l'ultimo codone, il ribosoma viene eiettato e la proteina è sintetizzata.

Il TASEP è stato introdotto proprio per cercare di modellizzare sistemi di questo tipo. Sperimentalmente, poi, è possibile misurare la densità di ribosomi in funzione della posizione lungo l'mRNA (per mRNA di diverse lunghezze), e i risultati sperimentali sono in ottimo accordo con le previsioni del TASEP:



#### Proprietà degli stati stazionari

Cerchiamo quindi di capire quali sono le proprietà dello stato stazionario di un modello come quelli che abbiamo appena visto. In particolare vogliamo studiare queste proprietà e le loro fluttuazioni anche per capire se in sistemi del genere vale il principio di grandi deviazioni.

Vediamo innanzitutto quali sono le equazioni alle quali soddisfano i vari  $\langle \tau_i \rangle$ . Se prendiamo ad esempio i=1, poiché:

$$\langle \tau_1 \rangle = 1 \cdot P_1(1) + 0 \cdot P_1(0) = P_1(1)$$

(ove con  $P_i(j)$  indichiamo la probabilità che il sito i abbia numero di occupazione  $j = \{0, 1\}$ ) allora:

$$\frac{d}{dt} \left\langle \tau_1 \right\rangle = \frac{d}{dt} P_1(1)$$

Ora, per determinare la variazione nel tempo di  $P_1(1)$  dobbiamo tener conto di tutte le possibili evoluzioni del sistema nei pressi del primo sito; se quindi indichiamo con  $P_{1,2}(\tau_1,\tau_2) \coloneqq P(\tau_1,\tau_2)$  la densità di probabilità congiunta delle variabili relative ai primi due siti, allora:

$$\frac{d}{dt}P_1(1) = \alpha(P(0,0) + P(0,1)) - \gamma(P(1,1) + P(1,0)) - P(1,0) + P(0,1)$$

ove il primo termine è relativo al caso in cui il primo sito sia inizialmente vuoto e venga occupato da una particella assorbita dal serbatoio, il secondo al caso in cui invece il sito sia inizialmente occupato e la particella venga assorbita dal serbatoio, il terzo al "salto" nel secondo sito della particella che inizialmente sta nel primo sito e l'ultimo termine alla situazione opposta.

Ora,  $P(0,0)+P(0,1)=P_1(0)$ , e in base alla definizione di  $\langle \tau_1 \rangle$  si ha  $P_1(0)=1-\langle \tau_1 \rangle$ ; inoltre,  $P(1,1)+P(1,0)=P_1(1)=\langle \tau_1 \rangle$ . Aggiungendo e sottraendo P(1,1) all'espressione di  $dP_1(1)/dt$ , tenendo conto che:

$$-P(1,0) - P(1,1) = -\langle \tau_1 \rangle$$
  $P(0,1) + P(1,1) = \langle \tau_2 \rangle$ 

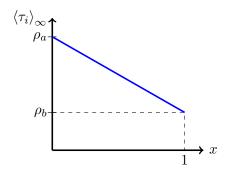


Figura 7.3: Andamento di  $\langle \tau_i \rangle_{\infty}$ 

si ottiene:

$$\frac{d\langle \tau_1 \rangle}{dt} = \alpha - (\alpha + \gamma + 1)\langle \tau_1 \rangle + \langle \tau_2 \rangle$$

Una relazione analoga vale per  $\langle \tau_L \rangle$ :

$$\frac{d\langle \tau_L \rangle}{dt} = \delta - (1 + \beta + \delta) \langle \tau_L \rangle + \langle \tau_{L-1} \rangle$$

Nella parte centrale del sistema, invece, bisogna tener conto che non entrano direttamente in gioco i serbatoi ma bisogna considerare oltre al generico sito i-esimo quelli (i-1)-esimo e (i+1)-esimo; alla fine si trova:

$$\frac{d\langle \tau_i \rangle}{dt} = \langle \tau_{i+1} \rangle - 2\langle \tau_i \rangle + \langle \tau_{i-1} \rangle$$

Adesso, per determinare il "profilo di densità" nello stato stazionario basta imporre che tutte le derivate temporali siano nulle; senza entrare nei dettagli, risulta che:

$$\langle \tau_i \rangle_{\infty} = \frac{\rho_a \left( L + \frac{1}{\beta + \delta} - i \right) + \rho_b \left( i - 1 + \frac{1}{\alpha + \gamma} \right)}{L + \frac{1}{\alpha + \gamma} + \frac{1}{\beta + \delta} - 1}$$

ove:

$$\rho_a = \frac{\alpha}{\alpha + \gamma} \qquad \qquad \rho_b = \frac{\delta}{\delta + \beta}$$

In termini più concreti, definita x = i/L (che è una variabile quasi continua definita fra 0 e 1) allora:

$$\langle \tau \rangle = (1 - x)\rho_a + x\rho_b$$

come ci potevamo ragionevolmente aspettare; la situazione è rappresentata in figura 7.3.

Consideriamo ora la corrente che fluisce nel sistema. Stabilendo per convenzione che questa è positiva se scorre nella direzione degli i crescenti (ossia verso destra), questa sarà data dal numero di particelle che passano dal sito i al sito i+1 nell'unità di tempo, ossia:

$$J = P_{i,i+1}(1,0) - P_{i,i+1}(0,1)$$

Aggiungendo e sottraendo  $P_{i,i+1}(1,1)$  si ottiene:

$$J = \langle \tau_i \rangle - \langle \tau_{i+1} \rangle$$

che è anche convieniente riscrivere come:

$$J = \langle \tau_i \rangle \left( 1 - \langle \tau_{i+1} \rangle \right) - \langle \tau_{i+1} \rangle \left( 1 - \langle \tau_i \rangle \right)$$

Nello stato stazionario del sistema, la corrente è quindi pari a:

$$J = \frac{\rho_a - \rho_b}{L + \frac{1}{\alpha + \gamma} + \frac{1}{\beta + \delta} - 1}$$

Se L è sufficientemente grande possiamo trascurare gli altri termini del denominatore, e l'espressione di J diventa fondamentalmente la legge di Fick (ossia, la corrente è proporzionale al gradiente del profilo di densità).

Consideriamo ora un ASEP periodico; come sappiamo allora i serbatoi non ci sono più (ma questo non significa che non saremo in grado di scrivere la legge di bilancio dettagliato generalizzata), il numero di particelle N è fisso e il numero di configurazioni possibili è  $\binom{L}{N}$ . Chiamiamo poi p la probabilità che una particella "salti" in verso antiorario e q in senso orario.

A questo punto, allora, detta C(t) la configurazione del sistema in funzione del tempo e Y(C(t)) il numero di "salti" effettuati dalle particelle al tempo t, se in totale ci sono state n cambi di configurazione allora  $Y(C(t)) = y_1 + y_2 + \cdots + y_n$ , con  $y_i$  il numero di salti all'i-esima "transizione". Allora l'idea è che per tempi lunghi questa quantità sia tale che  $\langle Y(C(t))/t \rangle \to J$ .

Quello che vogliamo fare è particolarizzare l'equazione master a un sistema che può fare "salti" come quello che stiamo considerando. Detta allora  $P_t(C,Y)$  la probabilità che a un certo istante t il sistema sia nella configurazione C e che Y abbia un dato valore, e introducendo la notazione  $\mathbb{W}_y$  per le matrici stocastiche (quelle che abbiamo trovato in 5.4.3) possiamo scrivere:

$$\frac{d}{dt}P_t(C,Y) = \sum_{y} \sum_{C'} \left[ \mathbb{W}_y(C,C')P_t(C',Y-y) + \delta_{C,C'} \mathbb{W}_y(C,C)P_t(C,Y) \right]$$
(7.11)

ove  $\mathbb{W}_y(C,C) = \sum_{C' \neq C} \mathbb{W}_y(C',C)$ .

Introduciamo ora la seguente funzione:

$$F_t(C,\lambda) = \sum_{Y} e^{\lambda Y} P_t(C,Y)$$

In questo modo, se moltiplichiamo ambo i membri della (7.11) per  $e^{\lambda Y}$  e sommiamo su tutti i valori di Y, a primo membro si ha dF/dt, mentre a secondo membro si ottiene (escludendo la somma su C'):

$$\sum_{y,Y} e^{\lambda Y} P_t(C', Y - y) = \sum_y e^{\lambda y} \sum_{Y'} e^{\lambda Y'} P_t(C', Y') = \sum_y e^{\lambda y} F_t(C', \lambda)$$

ove nel primo passaggio abbiamo definito Y' = Y - y. Complessivamente si ha:

$$\frac{d}{dt}F_t(C,\lambda) = \sum_{C'} \underbrace{\sum_{y} \left[ e^{\lambda y} \mathbb{W}_y(C,C') + \delta_{C,C'} \mathbb{W}_y(C,C) \right]}_{:= \mathcal{W}_{\lambda}(C,C')} F_t(C',\lambda)$$

ove abbiamo definito l'operatore lineare  $W_{\lambda}(C,C')$ . Poiché da quest'equazione la derivata temporale di  $F_t(C,\lambda)$  è data fondamentalmente dall'applicazione della "matrice"  $W_{\lambda}$  al "vettore"  $F_t(C,\lambda)$ , possiamo dire che la sua soluzione per grandi tempi è:

$$F_t(C,\lambda) = e^{\mathcal{W}_{\lambda}t} F_0(C_0,\lambda_0)$$

Possiamo notare adesso che se vogliamo determinare  $\langle e^{\lambda Y_t} \rangle$ , che è una scaled cumulant generating function, si ha:

$$\left\langle e^{\lambda Y_t} \right\rangle = \sum_{C,Y} P_t(C,Y) e^{\lambda Y} = \sum_C F_t(C,\lambda)$$

Quello che normalmente si fa, quindi, è diagonalizzare<sup>10</sup>  $W_{\lambda}$  e determinarne l'autovettore  $v(C, \lambda)$  relativo all'autovalore massimo  $g(\lambda)$ ; dato che questo sarà l'autovalore che dominerà la dinamica del sistema per tempi lunghi, allora:

$$\left\langle e^{\lambda Y_t} \right\rangle \stackrel{t \to \infty}{\sim} e^{g(\lambda)t}$$

Per il teorema di Gärtner-Ellis allora la grandezza  $j := Y_t/t$  soddisferà al principio di grandi deviazioni se:

$$\frac{1}{t} \ln \left\langle e^{\lambda Y_t} \right\rangle \stackrel{t \to \infty}{\longrightarrow} \varepsilon(\lambda)$$

con  $\varepsilon(\lambda)$  definita e differenziabile ovunque. È dunque chiaro che se riusciamo a risolvere il problema agli autovalori di  $\mathcal{W}_{\lambda}$  si può determinare la scaled cumulant generating function, e sfruttando il teorema di Gärtner-Ellis anche la rate function del sistema.

Vogliamo ora verificare che anche nel caso di un modello ASEP periodico si può determinare una legge generalizzata di bilancio dettagliato, nonostante l'assenza dei serbatoi. "Ispirati" da quello che abbiamo visto in 7.3.2, possiamo dire che vale:

$$\mathbb{W}_{y}(C',C)P_{\text{staz.}}(C) = \mathbb{W}_{-y}(C,C')P_{\text{staz.}}(C')e^{\gamma_{0}y}$$
(7.12)

Si tratta quindi di capire per quale valore di  $\gamma_0$  questa relazione è soddisfatta.

Nel nostro caso, ovviamente,  $\mathbb{W}_y(C',C)=p$  (perché corrisponde a un "salto" a destra effettuato da una particella) e  $\mathbb{W}_{-y}(C,C')=q$  (che corrisponde a un "salto" a sinistra); inoltre  $P_{\text{staz.}}(C)=P_{\text{staz.}}(C')=1/N_{\text{conf.}}$  con  $N_{\text{conf.}}$  numero delle configurazioni. Pertanto, la (7.12) è soddisfatta se:

$$\gamma_0 = \ln\left(\frac{q}{p}\right)$$

Nel caso di un TASEP periodico si può determinare che la corrente che si instaura nello stato stazionario del sistema è:

$$J = \lim_{t \to \infty} \left\langle \frac{Y_t}{t} \right\rangle = \frac{N(L - N)}{L(L - 1)}$$

 $<sup>^{10}</sup>$ Dato che la diagonalizzazione di  $W_{\lambda}$  nel caso di uno spazio delle configurazioni infinito-dimensionale è tutt'altro che semplice, in genere si cerca (se possibile) di "ridurre" lo spazio delle configurazioni ad uno finito-dimensionale.

che è di fatto uguale a P(1,0); infatti:

$$P(1,0) = P_1(1)P_2(0)_{|1,1} = \frac{N}{L} \cdot \frac{L-N}{L-1}$$

ove con  $P_2(0)_{|1,1}$  abbiamo indicato la probabilità condizionata che il secondo sito sia vuoto sapendo che il primo è occupato. Si ha  $P_1(1)=N/L$  perché possiamo occupare N degli L siti disponibili; poi, poiché tutte le configurazioni sono equiprobabili,  $P_2(0)_{|1,1}$  è la probabilità che gli L-N siti rimanenti siano vuoti.

Più complicato è calcolare in questo caso la fluttuazione della corrente attorno al suo valore medio; quello che risulta è che:

$$D := \lim_{t \to \infty} \frac{\left\langle Y_t^2 \right\rangle - \left\langle Y_t \right\rangle^2}{t} = \frac{LN(L-N)}{(L-1)(2L-1)} \cdot \frac{\binom{2L}{2N}}{\binom{L}{N}^2}$$

A questo punto J e D entrano in gioco una volta che andiamo a determinare il limite di  $\ln \langle e^{\lambda Y_t} \rangle / t$  per grandi tempi, poiché risulta:

$$\frac{1}{t}\ln\left\langle e^{\lambda Y_t}\right\rangle \stackrel{t\to\infty}{\sim} J\lambda + \frac{D}{2}\lambda^2 + \cdots$$

In altre parole, J e D sono i primi coefficienti dello sviluppo di  $\varepsilon(\lambda)$  (o, in altre parole ancora,  $\varepsilon(\lambda)$  è la funzione generatrice della corrente e delle sue fluttuazioni).

Il caso dell'ASEP è completamente noto, ad esempio è possibile calcolare esattamente  $\varepsilon(\lambda)$ ; questo implica che  $j=Y_t/t$  soddisfa il principio di grandi deviazioni:

$$P\left(\frac{Y_t}{t} = j\right) \overset{t \to \infty}{\sim} e^{-t\varphi(j)}$$

ove  $\varphi(i)$  è la rate function.

Sfruttando questo principio di grandi deviazioni si ha:

$$\left\langle e^{\lambda Y_t} \right\rangle = \int dY_t P(Y_t) e^{\lambda Y_t} = t \int dj e^{-t\varphi(j) + \lambda tj}$$

e detto:

$$\varepsilon(\lambda) \coloneqq \max_{j} (\lambda j - \varphi(j))$$

allora:

$$\left\langle e^{\lambda Y_t} \right\rangle \sim e^{\varepsilon(\lambda)t}$$
 (7.13)

Insomma, la scaled cumulant generating function (come sappiamo) è la trasformata di Legendre della rate function. Equivalentemente, infatti, si può ottenere  $\varphi(j)$  come:

$$\varphi(j) = \max_{\lambda} (\lambda j - \varepsilon(\lambda)) \tag{7.14}$$

#### 7.3.4 Teorema di Gallavotti-Cohen

Esploriamo ora alcune conseguenze di quello che abbiamo visto.

Abbiamo visto (equazione (7.7)) che se C è una traiettoria e la sua inversa è  $C^{\dagger}$ , nel caso in cui valga la legge di bilancio dettagliato il rapporto delle loro probabilità vale 1. Nel nostro caso, invece, dato che vale la legge generalizzata di bilancio dettagliato, si avrà:

$$\frac{P(C(t))}{P(C^{\dagger}(t))} = e^{\gamma_0 Y(C(t))} \tag{7.15}$$

Ora, in generale si ha

$$\sum_{C(t)} P(C(t))e^{(\gamma-\gamma_0)Y(C(t))} = \sum_{C(t)} P(C^{\dagger}(t))e^{\gamma Y(C(t))}$$

ove abbiamo usato la (7.15).

Ad ogni traiettoria C(t), poi, corrisponde una e una sola  $C^{\dagger}(t)$ , e pertanto  $\sum_{C(t)} = \sum_{C^{\dagger}(t)}$ . Quindi, sfruttando anche la disparità della Y sotto inversione temporale, ossia  $Y(C^{\dagger}(t)) = -Y(C(t))$  (Y è somma di quantità dispari per inversione temporale), si ha:

$$\sum_{C(t)} P(C(t)) e^{(\gamma - \gamma_0) Y(C(t))} = \sum_{C^{\dagger}(t)} P(C^{\dagger}(t)) e^{-\gamma Y(C^{\dagger}(t))}$$

o, più semplicemente:

$$\left\langle e^{(\gamma - \gamma_0)Y_t} \right\rangle = \left\langle e^{-\gamma Y_t} \right\rangle$$

Supponendo che esista una scaled cumulant generating function  $\varepsilon(\gamma)$ , poiché  $\left\langle e^{-\gamma Y_t}\right\rangle \stackrel{t\to\infty}{\sim} e^{\varepsilon(-\gamma)t}$ , allora:

$$\varepsilon(\gamma - \gamma_0) = \varepsilon(-\gamma) \tag{7.16}$$

Vediamo quali sono le conseguenze di questa relazione (una sorta di "parità generalizzata") sulla rate function  $\varphi(j)$ . Sostituendo  $\gamma' = -\lambda$  nella (7.14) si ha:

$$\varphi(j) = \max_{\gamma'} (-\gamma' j - \varepsilon(-\gamma'))$$

Rinominando  $\gamma'$  con  $\gamma$  e sfruttando la (7.16):

$$\varphi(j) = \max_{\gamma} (-\gamma j - \varepsilon(\gamma - \gamma_0))^{\gamma' := \frac{\gamma}{\gamma} - \gamma_0} \max_{\gamma'} (-\gamma' j - \gamma_0 j - \varepsilon(\gamma')) = -\gamma_0 j + \underbrace{\max_{\gamma'} (-\gamma' j - \varepsilon(\gamma'))}_{\varphi(-j)}$$

Pertanto:

$$\varphi(i) = \varphi(-i) - \gamma_0 i$$

Questo risultato è noto col nome di teorema di Gallavotti-Cohen.

Come conseguenza di questo teorema si ha:

$$\frac{P\left(\frac{Y_t}{t} = j\right)}{P\left(\frac{Y_t}{t} = -j\right)} \stackrel{t \to \infty}{\sim} e^{\gamma_0 jt}$$

che è un'espressione la cui struttura ricorda molto quella del teorema di Crooks.

Originariamente il teorema di Gallavotti-Cohen è stato formulato pensando a situazioni diverse da quelle che abbiamo considerato noi; vediamo quali.

Consideriamo una particella con velocità  $\vec{v}$  immersa in un fluido viscoso (che fa anche da serbatoio termico) in uno stato stazionario; qual è allora l'entropia trasferita al serbatoio nel tempo t in questo stato stazionario?

Possiamo definire questa entropia trasferita come:

$$\sigma = \frac{Q_t}{Tt} = \frac{d_e S}{dt}$$

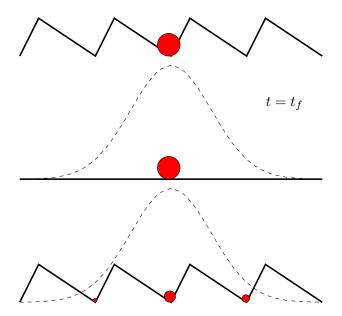


Figura 7.4: Schematizzazione dei motori molecolari. Nell'ultima figura la dimensione del cerchio rappresenta la probabilità di trovare la molecola in una posizione (più grande il cerchio, maggiore la probabilità)

dove  $Q_t$  è il calore scambiato col fluido nel tempo t; in questo caso allora il teorema di Gallavotti-Cohen stabilisce che:

$$\frac{P\left(\frac{d_e S}{dt} = \sigma\right)}{P\left(\frac{d_e S}{dt} = -\sigma\right)} = e^{\sigma t}$$

Quest'equazione è molto importante perché mostra che esiste sempre una possibilità (anche se sempre più piccola al passare del tempo) che l'entropia venga assorbita dal serbatoio, invece che ceduta ad esso. Se quindi consideriamo un sistema non macroscopico ci possono essere comportamenti "anomali" rispetto a quelli che ci aspetteremmo per un sistema macroscopico.

Un sistema biofisico dove il teorema di Gallavotti-Cohen trova applicazione sono i *motori* molecolari.

#### 7.3.5 Motori molecolari

I *motori molecolari* sono essenzialmente delle "macchine" molecolari che realizzano una "rettificazione" di moti browniani, supportate ovviamente da un input esterno di energia.

A livello cellulare i motori molecolari operano in un notevole numero di processi fondamentali per la vita (come il trasporto di materiale all'interno di una cellula).

I motori molecolari possono essere modellizzati nel seguente modo: consideriamo un tubulo, un filamento, una "rotaia dentata" sulla quale una molecola (il nostro motore) può spostarsi nel tempo. Queste strutture su cui i motori si muovono sono tipicamente caratterizzate da una periodicità spaziale dell'ordine della decina di nanometri; l'energia necessaria alla molecola per muoversi viene fornita tramite l'idrolisi dell'ATP (ATP  $\leftrightarrow$  ADP + P + energia).

Schematicamente, possiamo immaginare che al tempo iniziale t=0 la "rotaia" crei un

potenziale a dente di sega asimmetrico, e il motore sarà una "particella" localizzata in uno dei minimi di questo potenziale (vedi figura<sup>11</sup> 7.4). L'energia fornita dall'idrolisi dell'ATP "spiana" il potenziale fino al tempo  $t=t_f$ , e in questo intervallo di tempo la particella può diffondere. Finita l'energia dell'ATP il potenziale torna allo stato precedente; a seconda di dove si trova la particella in  $t_f^{12}$  questa potrà ricadere sullo stesso minimo di prima o in uno di quelli adiacenti. La probabilità di spostarsi in un minimo adiacente è però asimmetrica per via dell'asimmetria del potenziale a dente di sega, e favorisce lo spostamento in una direzione (negli esempi di figura 7.4 è verso destra).

Quali sono i parametri che entrano in gioco in questa modellizzazione?

Sicuramente la velocità di migrazione della molecola (che ovviamente nello stato stazionario del sistema è costante); dobbiamo poi anche considerare il rate di consumo di ATP: la reazione ATP  $\leftrightarrow$  ADP + P + energia deve infatti essere "sbilanciata" verso destra. Il rate di consumo di ATP è dato da  $\Delta\mu=\mu_{\rm ATP}-\mu_{\rm ADP}-\mu_{\rm P}$ . Le due "forzanti" esterne che agiscono sul sistema sono quindi  $\Delta\mu$  e la forza esterna  $F_{\rm ext}$  esercitata dal carico.

Per comodità chiamiamo a la periodicità spaziale del potenziale, e la usiamo per adimensionalizzare tutte le grandezze in gioco ridefinendo (detta X la posizione del motore):

$$v \longrightarrow rac{1}{a} rac{dX}{dt} \qquad \Delta \mu \longrightarrow rac{\Delta \mu}{k_B T} \qquad F_{\mathrm{ext}} \longrightarrow rac{F_{\mathrm{ext}} a}{k_B T} \coloneqq f_{\mathrm{ext}}$$

Quello che il modello deve essere quindi in grado di fornirci è allora la velocità  $v(f_{\rm ext},\Delta\mu)$  del motore e il rate  $r(f_{\rm ext},\Delta\mu)$  di consumo di ATP, e in particolare i loro valori stazionari.

Ora, chiaramente se  $f_{\rm ext}=\Delta\mu=0$  si ha v=r=0 e la molecola diffonde solamente (si osservano solo fluttuazioni, e ci aspettiamo che  $(\langle X_t^2\rangle - \langle X_t\rangle^2)/t \stackrel{t\to\infty}{\longrightarrow} D_1$  e  $(\langle Q_t^2\rangle - \langle Q_t\rangle^2)/t \stackrel{t\to\infty}{\longrightarrow} D_2$ , con Q energia dell'ATP consumata).

Nel caso del moto browniano abbiamo visto che possiamo legare fra loro il coefficiente di diffusione e la mobilità, che è la funzione di risposta del sistema; in particolare, se  $f_{\rm ext}$  è la forza esterna agente su una particella browniana, la sua "risposta" è l'acquisto della velocità  $v=\sigma f_{\rm ext}$  con  $\sigma=D/(k_BT)$ . Ci aspettiamo che anche in questo caso esista una relazione che leghi i "campi esterni"  $f_{\rm ext}$  e  $\Delta\mu$  ai parametri di risposta v=r del sistema:

$$v = L_{11} f_{\text{ext}} + L_{12} \Delta \mu$$
  $r = L_{21} f_{\text{ext}} + L_{22} \Delta \mu$  (7.17)

almeno per piccoli  $f_{\text{ext}}$  e  $\Delta \mu$ , e con  $L_{12} = L_{22}$ .

Vediamo come in effetti questo risultato sia una conseguenza del teorema di Gallavotti-Cohen.

Nel nostro caso la (7.13) diventa:

$$\left\langle e^{\gamma X_t + \lambda Q_t} \right\rangle \stackrel{t \to \infty}{\sim} e^{\varepsilon(\gamma, \lambda)t}$$
 (7.18)

ove comprensibilmente entrano in gioco due parametri,  $\lambda$  e  $\gamma$ . A questo punto, poiché vale una legge di bilancio dettagliato generalizzata, la  $\varepsilon$  soddisfa una relazione del tipo Gallavotti-Cohen:

$$\varepsilon(\gamma - f_{\text{ext}}, \lambda - \Delta\mu) = \varepsilon(-\gamma, -\lambda)$$

<sup>&</sup>lt;sup>11</sup>In realtà qui non stiamo tenendo conto del fatto che il carico trasportato dal motore esercita una forza costante su di esso. Il potenziale effettivo, quindi, andrebbe inclinato in avanti.

<sup>&</sup>lt;sup>12</sup>A priori non possiamo predire esattamente dove si troverà; sappiamo solo che, poiché ha diffuso per un poò di tempo, la densità di probabilità di trovarla in una data posizione è gaussiana.

che possiamo riscrivere come:

$$\varepsilon(-\gamma - f_{\text{ext}}, -\lambda - \Delta\mu) = \varepsilon(\gamma, \lambda) \tag{7.19}$$

In che modo questo teorema e la conoscenza di  $\varepsilon$  permette di trarre informazioni sul sistema?

Sappiamo che la  $\varepsilon$  è tale che<sup>13</sup>:

$$\frac{\partial \varepsilon}{\partial \gamma_{\mid (0,0)}} = \frac{\langle X_t \rangle}{t} = v \qquad \qquad \frac{\partial \varepsilon}{\partial \lambda_{\mid (0,0)}} = \frac{\langle Q_t \rangle}{t} = r$$

e inoltre si ha:

$$\frac{\partial^{2} \varepsilon}{\partial \gamma^{2}|_{(0,0)}} = \frac{\left\langle X_{t}^{2} \right\rangle - \left\langle X_{t} \right\rangle^{2}}{t} = D_{1} \qquad \qquad \frac{\partial^{2} \varepsilon}{\partial \lambda^{2}|_{(0,0)}} = \frac{\left\langle Q_{t}^{2} \right\rangle - \left\langle Q_{t} \right\rangle^{2}}{t} = D_{2}$$

Ora, derivando la (7.19) rispetto a  $\gamma$  e  $\lambda$  e usando i risultati appena visti otteniamo:

$$\frac{\partial \varepsilon}{\partial \gamma_{\parallel (0,0)}} = v = -\frac{\partial \varepsilon}{\partial \gamma} (-f_{\rm ext}, -\Delta \mu) \qquad \qquad \frac{\partial \varepsilon}{\partial \lambda_{\parallel (0,0)}} = r = -\frac{\partial \varepsilon}{\partial \lambda} (-f_{\rm ext}, -\Delta \mu)$$

Vediamo ora cosa accade in situazioni prossime all'equilibrio; espandendo queste ultime equazioni per piccoli  $f_{\text{ext}}$  e  $\Delta\mu$  otteniamo:

$$v = \frac{\partial^2 \varepsilon}{\partial \gamma^2}_{|(0,0)} f_{\text{ext}} + \frac{\partial^2 \varepsilon}{\partial \lambda \partial \gamma}_{|(0,0)} \Delta \mu + \cdots \qquad \qquad r = -\frac{\partial^2 \varepsilon}{\partial \gamma \partial \lambda}_{|(0,0)} f_{\text{ext}} + \frac{\partial^2 \varepsilon}{\partial \lambda^2}_{|(0,0)} \Delta \mu + \cdots$$

Se quindi chiamiamo:

$$\frac{\partial^2 \varepsilon}{\partial \gamma^2} = D_1 = L_{11} \qquad \quad \frac{\partial^2 \varepsilon}{\partial \lambda^2} = D_2 = L_{22} \qquad \quad L_{12} = \frac{\partial^2 \varepsilon}{\partial \lambda \partial \gamma} = \frac{\partial^2 \varepsilon}{\partial \gamma \partial \lambda} = L_{21}$$

ritroviamo proprio la (7.17), compreso il risultato (di Onsager, come sappiamo)  $L_{12} = L_{21}$ .

 $<sup>^{13}</sup>$ Dalla (7.18) si ha che per grandi tempi  $\varepsilon(\gamma,\lambda)=\gamma X_t+\lambda Q_t.$ 

# Bibliografia

- [1] P. W. Anderson, *Physics: The opening to complexity*, Proc. Natl. Acad. Sci. USA, **92**., pp. 6653-6654, 1995.
- [2] E. Boksenbojm et al., *Nonequilibrium thermodynamics at the microscale: work relations and the second law*, arXiv:1002.1230v1, 2010.
- [3] C. Bustamante et al., *The Nonequilibrium Thermodynamics of Small Systems*, Physics Today, July 2005.
- [4] B. Derrida, Non equilibrium steady states: fluctuations and large deviations of the density and of the current, arXiv:0703762v1, 2008.
- [5] D. T. Gillespie, *The mathematics of Brownian motion and Johnson noise*, American Journal of Physics, **64** (3), 1996.
- [6] N. Goldenfeld and L. P. Kadanoff, *Simple Lessons from Complexity*, Science **284** (87), 1999.
- [7] M. Le Bellac, Non equilibrium statistical mechanics, lectures given at Les Houches postdoctoral school, 2007.
- [8] K. Mallick, *Some recent developments in non-equilibrium statistical physics*, Pramana, **73** (3), pp. 417-451, 2009.
- [9] L. Peliti, Appunti di meccanica statistica, Bollati Boringhieri, 2003.
- [10] M. Stoppato, Appunti del corso: Fisica dei Sistemi Complessi, 2016.
- [11] H. Touchette, *The large deviation approach to statistical mechanics*, arXiv:0804.0327v2, 2009
- [12] N. G. Van Kampen, *Stochastic processes in Physics and Chemistry*, North Holland, terza edizione, 2007.