

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Corso di Laurea in Fisica

Tesi di Laurea

Il principio di Landauer e la termodinamica dell'informazione

Relatore

Prof. Amos Maritan

Laureanda

Elena Poli

Anno Accademico 2017/2018

Indice

1	Il principio di Landauer	3
2	L'entropia nella teoria dell'informazione	7
3	Il modello unidimensionale per la minima generazione di calore	11
3.1	L'equazione di Langevin	11
3.2	L'equazione di Kramers	12
3.3	La minima generazione di calore	17
4	Applicazione del principio di Landauer: il diavoleto di Maxwell	20
4.1	Il diavoleto di Maxwell	20
4.2	La macchina di Szilard	21
4.3	Risoluzione del paradosso	22
5	Verifiche sperimentali	23
5.1	L'esperimento di Raizen	23
5.2	Verifica della macchina di Szilard	24
5.3	Verifica del principio di Landauer	25
6	Conclusioni	28

Introduzione

Ogni volta che in un processo computazionale viene distrutta dell'informazione e quindi due situazioni precedentemente distinte diventano indistinguibili è necessaria una dissipazione di energia sotto forma di calore. È questo il centro delle argomentazioni proposte da Rolf Landauer nel 1961 volte allo studio dei limiti dell'efficienza di un computer imposti dalle leggi fisiche. Il suo principale risultato afferma che questa produzione di calore, dovuta al semplice fatto che l'informazione viene elaborata da porte logiche irreversibili, ha un limite inferiore. Nella realtà quotidiana la quantità di energia rilasciata da una macchina per un qualsiasi processo di elaborazione dati è decisamente maggiore di molti ordini di grandezza rispetto al minimo stimato da Landauer. Tuttavia, il fatto che la computazione obbedisca a dei limiti energetici imposti dalla fisica ed in particolare dalle leggi della termodinamica stessa rende tale studio interessante e applicabile a diversi ambiti fisici.

In particolare, questo ragionamento è tutto tranne che sterile se relazionato all'esperimento mentale di Maxwell, conosciuto appunto come "diavoletto di Maxwell": sarà presentato, infatti, il ragionamento di Bennett basato sul principio di Landauer che permette di distruggere il paradosso maxwelliano che metteva in crisi il secondo principio della termodinamica.

Perciò, da considerazioni energetiche mirate all'ottimizzazione di processi computazionali, tra l'altro verificate sperimentalmente, è stato possibile confermare l'universalità del principio termodinamico legato all'entropia ed evidenziare quanto il concetto di informazione sia fisico.

Questa tesi propone nella prima parte le argomentazioni di Landauer riguardanti il principio che porta il suo nome, le quali saranno poi dimostrate analiticamente in dettaglio attraverso un modello unidimensionale di dispositivo dotato di memoria suggerito da Shizume. Nella parte finale verranno riportate le applicazioni di tale principio e la sua verifica sperimentale.

Capitolo 1

Il principio di Landauer

Si consideri un supporto di memoria che può contenere un certo numero di bit, ciascuno dei quali può assumere il valore 0 o 1. Tale dispositivo può essere modellizzato con un sistema fisico dotato di due stati di stabilità: in particolare, possiamo immaginare una particella in una doppia buca di potenziale simmetrica rispetto all'asse delle ordinate. Si consideri ora l'operazione di cancellazione, che è proprio quella su cui si vuole focalizzare l'attenzione: essa computazionalmente equivale all'operazione di porre uguale ad 1 ogni locazione di memoria indipendentemente dal suo stato precedente. Quest'ultima affermazione è particolarmente importante perchè se si potessero utilizzare differenti routine di cancellazione a seconda dello stato precedente il sistema fisico considerato non opererebbe in modo analogo ad un computer e questa modellizzazione risulterebbe insensata. Infatti, la capacità di un computer è quella di agire allo stesso modo a prescindere dallo stato del sistema precedente: esso porta un sistema ad uno stato finale desiderato indipendentemente dallo stato iniziale dei dati che sta gestendo. In più, per determinare il futuro modus operandi sarebbe necessaria la memorizzazione dello stato precedente, proprio quello che si vuole cancellare. Ebbene, in figura 1.1 è possibile vedere il sistema fisico da analizzare, in cui la doppia buca analiticamente può essere descritta da un generico polinomio di quarto grado simmetrico e la pallina rappresenta il bit che può stare stabilmente in uno dei due stati.

Con questo modello, l'operazione di cancellazione può avvenire cambiando la forma del

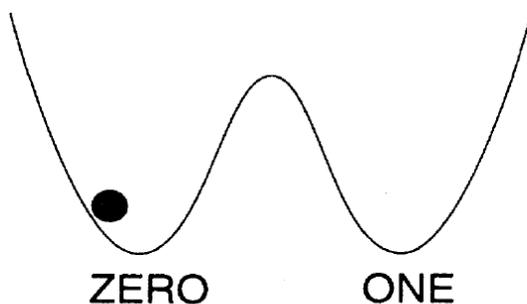


Figura 1.1: Modello del supporto di memoria

potenziale nel tempo: in tal modo la particella potrebbe trovarsi allo stato finale 1 indipendentemente da quello iniziale. La domanda a cui vogliamo rispondere è: che attributi qualitativi dovrebbe possedere una forza $F(t)$ agente sul sistema tale da operare una cancellazione della memoria come precedentemente descritto? Innanzitutto, si noti che le leggi della meccanica ad essa collegate non possono essere quelle relative ad un processo deterministico, dato che nel caso analizzato si vuole arrivare ad un unico stato finale fissato a partire da due stati iniziali differenti. Infatti, se si invertisse la coordinata temporale in un moto deterministico, sarebbe possibile ricondursi univocamente allo stato iniziale a partire da quello finale. Ma tutto ciò in questo modello non avviene: dato lo stato finale di un bit pari ad 1 ottenuto mediante un'operazione di cancellazione, lo stato iniziale del bit considerato potrebbe essere stato sia 0 che 1. Ne segue, dunque, che la forza $F(t)$ dovrà indurre un processo stocastico (probabilistico) nel sistema e dar luogo a fenomeni di dissipazione.

Rolf Landauer riuscì a modellizzare la cancellazione di informazione considerando, quindi, l'azione di una forza dissipativa che porta la pallina da sinistra verso destra. La forza agisce in maniera lenta per cercare di minimizzare l'attrito contro le pareti della buca (che consideriamo essere viscoso, quindi proporzionale alla velocità), ma essa deve essere inevitabilmente dissipativa, con conseguente perdita di una certa quantità di energia. La dissipazione che è un fenomeno fisicamente irreversibile viene strettamente connessa alla cancellazione di informazione che è un fenomeno logicamente irreversibile, ovvero dato l'output è impossibile determinare unicamente l'input.

È facile capire come la cancellazione di informazione avvenga costantemente anche considerando semplici funzioni logiche alla base di ogni processo computazionale ¹, basta che esse siano logicamente irreversibili. Ad esempio, consideriamo la tabella di verità dell'operatore *AND*:

A	B	$A \wedge B$
1	1	1
0	1	0
1	0	0
0	0	0

Si nota che quando l'uscita della porta è 0 si perde informazione perchè non siamo più in grado di risalire allo stato iniziale (una visione schematica della perdita di informazione è visibile in figura 1.2).

Secondo Landauer quindi ogni volta che viene utilizzata una porta logica irreversibile viene dissipata dell'energia nell'ambiente circostante. Il ragionamento esposto fino ad ora tiene in considerazione un solo bit ma può essere generalizzato per un array di bit gestiti dal computer durante la manipolazione dell'informazione. Un array di n bit può assumere 2^n configurazioni differenti: in altre parole, l'array di cui si vuole cancellare l'informazione sarà in uno stato corrispondente a uno dei 2^n punti nello spazio delle fasi, ovvero lo spazio in cui ogni punto corrisponde ad un possibile stato del sistema. Ancora una volta è necessario tener conto di tutti gli stati iniziali possibili perché l'azione che si vuole descrivere deve essere indipendente dallo stato iniziale.

È stato introdotto il concetto di spazio delle fasi e grazie a questo è possibile introdurre una quantità fondamentale per arrivare alla conclusione di Landauer, l'*entropia*. Precisamente,

¹Concretamente, il trasporto di una certa informazione viene gestito dai circuiti logici nel computer che portano la stringa di bit dallo stato iniziale a quello finale

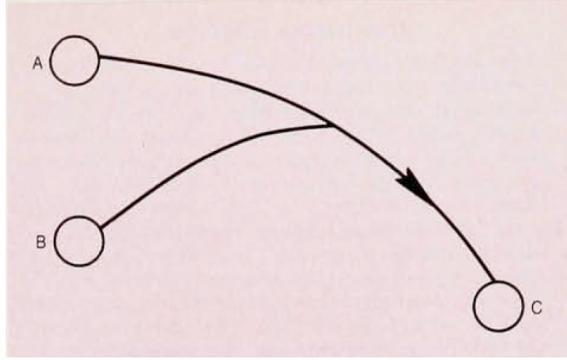


Figura 1.2: Visione simbolica della perdita di informazione nel passaggio dallo stato A o B a C

riprendendo i postulati della meccanica statistica possiamo definire l'entropia in questa maniera:

$$S \equiv k_B \log(W) \quad (1.1)$$

con k_B pari alla costante di Boltzmann e W pari al volume nello spazio delle fasi occupato da tutti gli stati microscopici consistenti con determinati vincoli macroscopici del sistema considerato.²

Per garantire che la cancellazione non dipenda dallo stato iniziale, consideriamo uno stato sconosciuto con un volume nello spazio delle fasi pari a 2^n , ovvero alla totalità dei possibili stati. Infatti ogni bit può assumere 2 configurazioni e, essendoci n bit, si ottiene che le possibili configurazioni iniziali dell'array sono proprio 2^n , ognuna corrispondente ad un punto nello spazio delle fasi. Consideriamo allora la variazione di entropia dovuta al processo di reset: il computer comprime i 2^n stati possibili in uno solo, quello in cui tutti i bit sono 1 se si vuole riprendere l'esempio riportato precedentemente. Si ottiene:

$$\Delta S = k_B \log(2^n) - k_B \log(1) = nk_B \log(2) \quad (1.2)$$

Se supponiamo il sistema totale chiuso e diamo per valido il secondo principio della termodinamica la minima quantità di calore rilasciata nell'ambiente circostante sarà quindi

$$Q = T\Delta S = k_B n T \log(2) \quad (1.3)$$

dove ricordiamo che n è il numero di bit dell'array considerato.

Ecco quindi esplicitato del tutto il principio di Landauer: per ogni bit di informazione cancellato si produce nell'ambiente una quantità di calore $Q \geq k_b T \log(2)$. Per comprimere lo stato logico del calcolatore è necessario comprimere il suo stato fisico con conseguente diminuzione dell'entropia dell'hardware e in accordo al secondo principio della termodinamica questo non può avvenire senza un aumento dell'entropia dell'ambiente maggiore

²Con la funzione $y = \log(x)$ d'ora in poi si intenderà sempre il logaritmo naturale.

o uguale a ΔS calcolato prima. Non è possibile azzerare un registro di memoria o più in generale applicare un'operazione logicamente irreversibile ad un array di bit senza aumentare l'entropia dell'ambiente.

Prima di approcciarsi ad una dimostrazione dettagliata della disuguaglianza enunciata nel principio di Landauer con un modello unidimensionale proposto da Shizume, è importante rivedere il concetto di entropia in termini della cosiddetta teoria dell'informazione.

Capitolo 2

L'entropia nella teoria dell'informazione

Il concetto di entropia nella teoria dell'informazione fu introdotta nel 1948 proprio dal colui che è considerato il padre di questa branca di studio, Claude Shannon. La connessione tra la meccanica statistica e la teoria dell'informazione viene massimamente evidenziata dall'equivalenza dei concetti di entropia utilizzati nelle due branche, resa esplicita nell'utilizzo del principio di massima entropia.

Nella teoria dell'informazione l'entropia di una sorgente di messaggi può essere definita come l'informazione media contenuta in ogni messaggio emesso. In particolare, supponiamo di avere una variabile x che può assumere dei valori discreti (x_1, \dots, x_n) con le corrispondenti probabilità (p_1, \dots, p_n) tali che $\sum_{i=1}^n p_i = 1$. Shannon si chiese se fosse possibile trovare una funzione $H(p_1, \dots, p_n)$ che potesse determinare univocamente la quantità di incertezza rappresentata da questa distribuzione di probabilità e che rispettasse le seguenti tre condizioni:

- $H(p_1, \dots, p_n)$ deve essere una funzione continua nelle variabili p_1, \dots, p_n ;
- Se $p_1, \dots, p_n = \frac{1}{n}$, la quantità $A(n) = H(\frac{1}{n}, \dots, \frac{1}{n})$ è una funzione monotona crescente nella variabile n . In altre parole, al crescere del numero di eventi n , nella condizione di equiprobabilità l'incertezza associata è maggiore;
- $H(p_1, \dots, p_n)$ deve essere additiva: se si raggruppano i primi m eventi con $m \leq n$, si associa loro la probabilità $\omega_1 = p_1 + \dots + p_m$ e si associa ai rimanenti la probabilità $\omega_2 = p_{m+1} + \dots + p_n$, la quantità di incertezza associata a questa unione di eventi sarà $H(\omega_1, \omega_2)$ e $H(p_1, \dots, p_n)$ deve essere tale che

$$H(p_1, \dots, p_n) = H(\omega_1, \omega_2) + \omega_1 H(\frac{p_1}{\omega_1}, \dots, \frac{p_m}{\omega_1}) + \omega_2 H(\frac{p_{m+1}}{\omega_2}, \dots, \frac{p_n}{\omega_2})$$

Ovviamente tale proprietà deve valere anche nel caso di una generalizzazione a k partizioni degli eventi (x_1, \dots, x_n) .

La funzione cercata, a meno di una costante positiva K , è

$$H(p_1, \dots, p_n) = -K \sum_i p_i \log(p_i) \tag{2.1}$$

Questa quantità lega il concetto di probabilità a quello di incertezza: intuitivamente, un'ampia distribuzione di probabilità è connessa ad un'ampia incertezza al contrario di una distribuzione molto piccata. $H(p_1, \dots, p_n)$ è chiamata *entropia della distribuzione di probabilità p_i* .

Per comprendere al meglio la connessione tra incertezza legata ad una stringa ed entropia, consideriamo la trasmissione di un segnale classico, ovvero una stringa di M caratteri, attraverso un canale di comunicazione classico. Ciascun carattere è estratto da un insieme di r lettere diverse in maniera statisticamente indipendente dagli altri con probabilità p_i con $i = 1, \dots, r$ ¹. Chiamiamo ora con m_i il numero di volte che l' i -esimo carattere compare nel messaggio, ma allora il messaggio più probabile sarà quello composto da $m_i = M \cdot p_i$ lettere di tipo i . Passando all'intero messaggio così composto, ricordando che l'evento di ciascuna lettera è statisticamente indipendente da quelli relativi a tutte le altre, la probabilità P con cui il preciso messaggio contenente m_1 lettere di tipo 1, m_2 lettere di tipo 2, etc. ricorrerà è pari a

$$P = p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_r^{m_r} = p_1^{Mp_1} \cdot p_2^{Mp_2} \cdot \dots \cdot p_r^{Mp_r} \quad (2.2)$$

Se tutte le lettere fossero uguali la probabilità di sapere con anticipo il messaggio finale sarebbe $P = 1$, aumentando il numero di lettere possibili invece diminuisce il valore di P e aumenta l'incertezza di ricezione di un certo messaggio. Quest'ultima nozione può essere inclusa nella formula del logaritmo di P , precisamente mediante l'espressione di H precedentemente definita in 2.1:

$$H = \log\left(\frac{1}{P}\right) = -M \sum_i p_i \log(p_i) \quad (2.3)$$

Shannon notò che tale funzione svolge lo stesso ruolo dell'entropia di Gibbs e Boltzmann: massimizzare $\frac{1}{P}$ è del tutto equivalente alla massimizzazione di W in meccanica statistica. Ciò vuol dire che la quantità $\frac{1}{P}$ corrisponde alla degenerazione del messaggio più probabile, in altre parole corrisponde al volume W occupato dal messaggio nello spazio delle fasi formato dal numero totale di messaggi composti da sequenze di lettere con molteplicità $m_i = M \cdot p_i$ e che hanno tutti probabilità P di comparire.

Ora, si verificherà come partendo dall'entropia di Gibbs e Boltzmann sia possibile verificare l'equivalenza con quella dell'informazione e viceversa.

Nel primo caso, partiamo dalla meccanica statistica considerando un gas composto da un numero grande N di particelle in 3 dimensioni. Consideriamo lo spazio a 6 gradi di libertà in cui ogni punto corrisponde ad un possibile stato di una particella di gas determinato da posizione e momento. A questo punto, possiamo dividere questo spazio in s cellette: ogni stato di particella di gas sarà rappresentato da un punto all'interno di una delle s cellette e sarà possibile assegnare a ciascuna di esse il proprio numero di occupazione n_i , ovvero

¹L'intera discussione ha alla base l'interpretazione frequentista della probabilità, ovvero si suppone che le probabilità associate alle lettere siano il rapporto tra numero di volte in cui ciascuna di esse è uscita nel messaggio ed il numero fissato M di prove totali effettuate. Quindi si assume che sia possibile ottenere dal dispositivo un gran numero di messaggi in modo da associare ad ogni uscita a ciascuna lettera la sua frequenza $f_i = \frac{m_i}{M}$. Sempre secondo questa interpretazione frequentista, per un numero n di lanci molto alto questa frequenza coincide con la probabilità associata alla singola lettera ovvero $\lim_{n \rightarrow \infty} f_i = p_i$ e perciò nei passaggi successivi del ragionamento considereremo valida la relazione $p_i = \frac{m_i}{M}$.

il numero di stati nella celletta i , con $i = 1, \dots, s$ e $\sum_i n_i = N$. Fissata la configurazione macroscopica del gas, esistono diverse configurazioni microscopiche corrispondenti ad essa cioè esistono diversi modi di "riempimento" di queste cellette con gli stati delle particelle singole tali che lo stato microscopico sia compatibile con quello macroscopico imposto, ovvero tali che i numeri di occupazione di ogni singola celletta siano preservati. In altre parole, passando allo spazio delle fasi, dove ogni punto corrisponde ad una configurazione dell'intero sistema, il volume totale W degli stati microscopici nello spazio delle fasi è il numero di modi possibili di distribuzione di N elementi nelle cellette disponibili nei limiti della compatibilità macroscopica. Ovvero:

$$W = \frac{N!}{n_1! \dots n_s!} \quad (2.4)$$

Ma allora, sapendo che l'entropia di Gibbs e Boltzmann è proporzionale al logaritmo di W , otteniamo

$$\begin{aligned} S = \log(W) &= \log(N!) - \sum_i \log(n_i!) \approx \\ &\approx N \log(N) - N - \sum_i n_i \log(n_i) + \sum_i n_i = \\ &= - \left[\sum_i n_i \log(n_i) - M \log(N) \right] = \\ &= - \left[\sum_i n_i \log(n_i) - \sum_i n_i \log(N) \right] = \\ &= -N \sum_i \frac{n_i}{N} \log\left(\frac{n_i}{N}\right) \\ &= -N \sum_i p_i \log(p_i) \end{aligned} \quad (2.5)$$

dove l'approssimazione fatta all'inizio è quella di Stirling e nell'ultimo passaggio si è usata l'ipotesi di N grande. Partendo dalla meccanica statistica otteniamo nuovamente la formula 2.3 confermando il fatto che S coincide con quella che in teoria dell'informazione si è chiamata H .

Verifichiamo ora che se si affronta un problema di massimo vincolato risolvibile con i moltiplicatori di Lagrange, la risoluzione attraverso la massimizzazione di H coincide con il principio della massima entropia. Quindi, massimizzando l'entropia definita nella teoria dell'informazione si arriva alle stesse conclusioni che si otterrebbero massimizzando quella della meccanica statistica.

Supponiamo di avere una distribuzione di probabilità associata a dei dati. I due vincoli da imporre sono $\sum_i p_i = 1$ e $\langle E \rangle = \sum_i p_i E_i$ dove $\langle E \rangle$ è il valor medio della grandezza E . L'imposizione di solo questi due vincoli permette di affrontare il problema anche dal punto di vista della meccanica statistica: in particolare, possiamo considerare un ensemble canonico ed imporre il vincolo che fissa il valor medio dell'energia E del sistema all'equilibrio termodinamico. Le distribuzioni di probabilità consistenti con i dati sono differenti

ma quella più probabile si ottiene trovando la distribuzione che massimizza l'entropia e soddisfa i vincoli imposti. Impostando questo problema di massimo vincolato, con α e β moltiplicatori di Lagrange, si ottiene la grandezza

$$F(p_i) = - \sum_i p_i \log(p_i) - \alpha(\sum_i p_i - 1) - \beta(\sum_i p_i E_i - \langle E \rangle) \quad (2.6)$$

Derivando quest'ultima espressione rispetto a p_i e ponendo la derivata uguale a zero si ottiene:

$$\begin{aligned} \frac{\partial F(p_i)}{\partial p_i} &= -1 - \log(p_i) - \alpha - \beta E_i = 0 \\ p_i &= e^{-1-\alpha-\beta E_i} = \frac{e^{-\beta E_i}}{Z} \end{aligned}$$

con Z pari alla funzione di partizione dell'ensemble canonico.

La distribuzione più probabile ottenuta massimizzando l'entropia di Shannon coincide con la distribuzione di Maxwell Boltzmann ottenuta in meccanica statistica dalla definizione di entropia vista al 2.3.

Ancora una volta, quindi, si trova riscontro dell'equivalenza delle due scritte: nei capitoli successivi perciò si è legittimati ad utilizzare la forma analitica di H qualora fosse necessario definire ed studiare l'entropia.

Capitolo 3

Il modello unidimensionale per la minima generazione di calore

Ritornando al principio di Landauer, si è già visto come sia possibile modellizzare un dispositivo dotato di memoria con un modello unidimensionale a doppia buca di potenziale come rappresentato in figura 1.1. Come già detto, l'operazione di cancellazione di informazione equivale a portare il sistema fisico allo stato stabile finale nella buca denotata con il numero 1, indipendentemente da quello iniziale: in questo modo si sarà persa l'informazione proprio nel senso che risulta impossibile dire quale fosse lo stato precedente del sistema senza averlo salvato. Si dimostrerà analiticamente e senza l'assunzione a priori della validità del secondo principio della termodinamica che, indipendentemente dalla variazione nel tempo della buca di potenziale che permette di effettuare l'operazione voluta (si veda la figura 3.1), la quantità di calore totale generata dal procedimento per ogni bit sarà maggiore di $k_B T \ln(2)$.

L'intero ragionamento sarà impostato sull'utilizzo dell'equazione di Kramers che descrive l'evoluzione temporale della funzione densità di probabilità della posizione di una particella sottoposta a forze che possono essere d'attrito o semplici fluttuazioni casuali. In particolare, per tener conto di queste ultime si considera la presenza un rumore gaussiano bianco, ovvero un rumore di fondo che come densità di probabilità ha una distribuzione gaussiana e con spettro costante in tutta la banda di frequenze di interesse. Questa assunzione ci permette di includere la condizione di un rumore di fondo estremamente variabile da un istante all'altro e minimamente correlato agli istanti precedenti.

Per non appesantire la scrittura nei calcoli seguenti la costante di Boltzmann sarà posta $k_B = 1$.

3.1 L'equazione di Langevin

Nel modello unidimensionale ciò che si sposta da uno stato stabile qualsiasi allo stato stabile 1 durante l'operazione di cancellazione è una particella di massa m . Per descrivere il moto della particella impostiamo l'equazione di Langevin corrispondente, ovvero un'equazione differenziale stocastica ottimale per l'analisi di un sistema fisico soggetto a fluttuazioni:

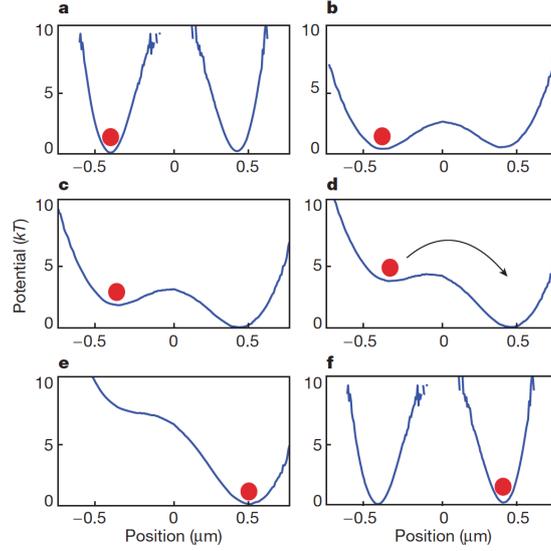


Figura 3.1: Transizione da uno stato iniziale generico (in questo caso 0) allo stato finale 1. Questo modello rappresenta l'operazione di cancellazione di informazione.

$$m \frac{\partial^2 x}{\partial t^2} = -m\gamma \frac{\partial x}{\partial t} - \frac{\partial V(x, t)}{\partial x} + F_r(t) \quad (3.1)$$

dove si vede che l'accelerazione della particella è influenzata da:

- la forza d'attrito viscoso proporzionale alla velocità e con coefficiente γ costante;
- la derivata spaziale del potenziale V che, tuttavia, è variabile anche nel tempo;
- $F_r(t)$, la forza che tiene conto del contributo di rumore bianco gaussiano e che soddisfa la condizione:

$$\langle F_r(t_1) F_r(t_2) \rangle = 2m\gamma T \delta(t_1 - t_2) \quad (3.2)$$

In pratica, questa condizione esprime il fatto che le fluttuazioni termiche dipendono dalla temperatura dell'ambiente, dal coefficiente di attrito, dalla massa della particella e cambiano in continuazione nel tempo. Il coefficiente per cui è moltiplicata la δ di Dirac garantisce che l'equazione 3.1 descrive un sistema che asintoticamente raggiunge l'equilibrio termodinamico ed è caratterizzato da una distribuzione di Maxwell-Boltzmann.

3.2 L'equazione di Kramers

La posizione della particella nel sistema fisico appena riportato è descritta da una distribuzione di probabilità che in generale dipende dalla posizione x , dalla velocità u e dal tempo t . Questa funzione, che sarà denotata con $f(x, u, t)$, può essere ricavata dalla risoluzione dell'equazione di Langevin con l'utilizzo di metodi numerici, ad esempio con metodi Monte

Carlo, e soddisfa l'equazione di Kramers deterministica per le probabilità. Quest'ultima permette di descrivere l'evoluzione di un sistema soggetto ad interazioni deterministiche a cui si aggiunge l'effetto di termini fluttuanti con proprietà statistiche note. In questo caso la fluttuazione di cui si conoscono le proprietà statistiche è il rumore di fondo $F_r(t)$ gaussiano presentato precedentemente. Esplicitamente, l'equazione di Kramers che $f(x, u, t)$ soddisfa è

$$\frac{\partial f(x, u, t)}{\partial t} = \left[-\frac{\partial u}{\partial x} + \frac{\partial}{\partial u} \left(\gamma u + \frac{1}{m} \frac{\partial V(x, t)}{\partial x} \right) + \frac{\gamma T}{m} \frac{\partial^2}{\partial u^2} \right] f(x, u, t) \quad (3.3)$$

Ora, sia \dot{Q} la media di ensemble del calore assorbito dall'ambiente circostante per unità di tempo e sia S l'entropia associata al sistema,

$$S \equiv - \int_{-\infty}^{+\infty} f \log(f) dx du \quad (3.4)$$

Si noti che la definizione di entropia corrisponde all'entropia dell'informazione di Shannon H come esplicitato al capitolo 2 ma questa volta al posto della sommatoria, come visto nella formula 2.1, si ha un integrale esteso allo spazio delle fasi poichè le variabili di integrazione hanno un dominio continuo e la probabilità discreta p_i è sostituita dalla funzione di distribuzione di probabilità $f(x, u, t)$.

A questo punto, si vuole verificare la validità della disuguaglianza del secondo principio della termodinamica

$$\dot{Q} \leq T \frac{dS}{dt} \quad (3.5)$$

Shizume infatti non considera tale principio vero a priori dato che il potenziale non è statico.

Consideriamo l'energia interna totale della particella E data dalla somma di un termine cinetico ed uno potenziale, ovvero

$$E \equiv \frac{mu^2}{2} + V(x, t) \quad (3.6)$$

Derivandola rispetto al tempo si ottiene:

$$\begin{aligned} \frac{\partial E}{\partial t} &= mu(t) \dot{u}(t) + u(t) \frac{\partial V}{\partial x} + \frac{\partial V}{\partial t} = \\ &= u(t) \left[mu \dot{u} + \frac{\partial V}{\partial x} \right] + \frac{\partial V}{\partial t} = \\ &= \frac{\partial V}{\partial t} = W \end{aligned} \quad (3.7)$$

dato che il termine tra parentesi quadre è nullo per il secondo principio della dinamica. Il risultato finale identificato con W corrisponde al lavoro fatto dall'esterno per unità di tempo sulla particella.

Se consideriamo, come detto precedentemente, il sistema immerso in un bagno termico è necessario tener conto anche del calore scambiato con l'esterno che contribuisce all'energia totale. Ma allora il primo membro della disequazione 3.5 è possibile riscriverlo nella seguente maniera:

$$\dot{Q} = \frac{d\langle E \rangle}{dt} - \dot{W} \quad (3.8)$$

che corrisponde al primo principio della termodinamica valutato per unità di tempo e dove tutti i termini sono mediati sullo spazio delle fasi. Si ricordi che per definizione la media di ensemble di una grandezza $A(x, u, t)$ è

$$\langle E \rangle \equiv \int_{-\infty}^{+\infty} f(x, u, t) A(x, u, t) dx du \quad (3.9)$$

In particolare, poichè per quanto detto prima $\frac{\partial V(x, t)}{\partial t}$ è il lavoro fatto dall'esterno sulla particella per unità di tempo, \dot{W} è

$$\dot{W} = \left\langle \frac{\partial V(x, t)}{\partial t} \right\rangle \equiv \int_{-\infty}^{+\infty} f(x, u, t) \frac{\partial V(x, t)}{\partial t} dx du \quad (3.10)$$

A questo punto si può riscrivere l'equazione 3.8 in questo modo

$$\begin{aligned} \dot{Q} &= \frac{d\langle E \rangle}{dt} - \dot{W} = \frac{d}{dt} \left[\int_{-\infty}^{+\infty} f(x, u, t) E dx du \right] - \int_{-\infty}^{+\infty} f(x, u, t) \frac{\partial V(x, t)}{\partial t} dx du = \\ \dot{Q} &= \int_{-\infty}^{+\infty} \frac{\partial f(x, u, t)}{\partial t} E dx du + \int_{-\infty}^{+\infty} f(x, u, t) \frac{\partial V(x, t)}{\partial t} dx du - \int_{-\infty}^{+\infty} f(x, u, t) \frac{\partial V(x, t)}{\partial t} dx du = \\ \dot{Q} &= \int_{-\infty}^{+\infty} \frac{\partial f(x, u, t)}{\partial t} \left(\frac{mu^2}{2} + V(x, t) \right) dx du \end{aligned} \quad (3.11)$$

dove nel penultimo passaggio è stata sfruttata la condizione che l'unico termine dell'energia totale dipendente dal tempo è il potenziale V .

Ora, sostituendo alla derivata temporale di $f(x, u, t)$ l'espressione data dall'equazione di Kramers si ottiene per ciascuno dei tre termini del membro di destra dell'equazione 3.3:

1.

$$\begin{aligned} & \int_{-\infty}^{+\infty} -u \frac{\partial f(x, u, t)}{\partial x} \left(\frac{mu^2}{2} + V(x, t) \right) dx du \\ &= \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} f(x, u, t) \frac{\partial}{\partial x} \left(\frac{mu^3}{2} + uV(x, t) \right) dx = \\ &= \left\langle u \frac{\partial V(x, t)}{\partial x} \right\rangle \end{aligned} \quad (3.12)$$

dove si è integrato per parti in dx e si è sfruttato il fatto che a $\pm\infty$ la funzione di distribuzione di probabilità $f(x, u, t)$ è nulla. Anche negli integrali successivi saranno sfruttate queste proprietà.

2.

$$\begin{aligned}
& \int_{-\infty}^{+\infty} \frac{\partial}{\partial u} \left[\left(\gamma u + \frac{1}{m} \frac{\partial V(x, t)}{\partial x} \right) f(x, u, t) \right] \left(\frac{mu^2}{2} + V(x, t) \right) dx du = \\
& = - \int_{-\infty}^{+\infty} f(x, u, t) \left(\gamma u + \frac{1}{m} \frac{\partial V(x, t)}{\partial x} \right) \frac{\partial}{\partial u} \left(\frac{mu^2}{2} + V(x, t) \right) dx du = \\
& = - \int_{-\infty}^{+\infty} f(x, u, t) \left(\gamma u + \frac{1}{m} \frac{\partial V(x, t)}{\partial x} \right) (mu) dx du = \\
& = - \left\langle \gamma mu^2 + u \frac{\partial V(x, t)}{\partial x} \right\rangle
\end{aligned} \tag{3.13}$$

3.

$$\begin{aligned}
& \gamma \frac{T}{m} \int_{-\infty}^{+\infty} \frac{\partial^2 f(x, u, t)}{\partial u^2} \left(\frac{mu^2}{2} + V(x, t) \right) dx du = \\
& = -\gamma \frac{T}{m} \int_{-\infty}^{+\infty} \frac{\partial f(x, u, t)}{\partial u} \frac{\partial}{\partial u} \left(\frac{mu^2}{2} + V(x, t) \right) dx du = \\
& = -\gamma \frac{T}{m} \int_{-\infty}^{+\infty} \frac{\partial f(x, u, t)}{\partial u} mu dx du = \\
& = \gamma \frac{T}{m} \int_{-\infty}^{+\infty} f(x, u, t) \frac{\partial(mu)}{\partial u} dx du = \\
& = \gamma T \int_{-\infty}^{+\infty} f(x, u, t) dx du = \\
& = \gamma T
\end{aligned} \tag{3.14}$$

In definitiva, sommando i tre termini si ottiene perciò che

$$\begin{aligned}
\dot{Q} &= \left\langle u \frac{\partial V(x, t)}{\partial x} - \gamma mu^2 - u \frac{\partial V(x, t)}{\partial x} \right\rangle + \gamma T = \\
&= \gamma (T - \langle mu^2 \rangle)
\end{aligned} \tag{3.15}$$

Per quanto riguarda invece il secondo membro della disequazione 3.5, differenziamo la 3.4 rispetto al tempo e sfruttiamo nuovamente la 3.3. Nei passaggi successivi per non appesantire la scrittura considereremo $f(x, u, t) \equiv f$ e $V(x, t) \equiv V$ pur sempre ricordando da che variabili dipendono. Si ottiene:

$$\begin{aligned}
\frac{\partial S}{\partial t} &= -\frac{\partial}{\partial t} \int_{-\infty}^{+\infty} f \log(f) dx du \\
&= - \int_{-\infty}^{+\infty} \frac{\partial f}{\partial t} \log(f) dx du - \int_{-\infty}^{+\infty} \frac{\partial f}{\partial t} dx du = \\
&= - \int_{-\infty}^{+\infty} \frac{\partial f}{\partial t} \log(f) dx du - \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} f dx du = \\
&= - \int_{-\infty}^{+\infty} \frac{\partial f}{\partial t} \log(f) dx du = \\
&= - \int_{-\infty}^{+\infty} \left(-u \frac{\partial f}{\partial x} + \frac{\partial}{\partial u} \left(\gamma u + \frac{1}{m} \frac{\partial V}{\partial x} \right) f + \gamma \frac{T}{m} \frac{\partial^2 f}{\partial u^2} \right) \log(f) dx du
\end{aligned} \tag{3.16}$$

Si noti che nel terzo passaggio il secondo termine è nullo perchè la densità di probabilità, essendo normalizzata, se integrata su tutto lo spazio dà 1 e perciò la derivata di tale integrale è nulla. In più, nell'ultimo passaggio è stata effettuata la solita sostituzione con la formula di Kramers 3.3. Come effettuato precedentemente, calcoliamo singolarmente i tre termini dell'integrale:

1.

$$\begin{aligned}
& \int_{-\infty}^{+\infty} -u \frac{\partial f}{\partial x} \log(f) dx du \\
&= - \int_{-\infty}^{+\infty} u du \int_{-\infty}^{+\infty} \frac{\partial f}{\partial x} \log(f) dx = \\
&= \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} f \frac{1}{f} \frac{\partial f}{\partial x} dx = \\
&= 0
\end{aligned} \tag{3.17}$$

dato che ancora una volta si è integrato per parti in dx e si è sfruttato il fatto che a $\pm\infty$ la funzione di distribuzione di probabilità $f(x, u, t)$ è nulla. Anche negli integrali successivi saranno sfruttate queste condizioni.

2.

$$\begin{aligned}
& \int_{-\infty}^{+\infty} \log(f) \frac{\partial}{\partial u} \left[\left(\gamma u + \frac{1}{m} \frac{\partial V}{\partial x} \right) f \right] dx du \\
&= - \int_{-\infty}^{+\infty} f \frac{1}{f} \frac{\partial f}{\partial u} \left(\gamma u + \frac{1}{m} \frac{\partial V}{\partial x} \right) dx du = \\
&= \int_{-\infty}^{+\infty} f \gamma dx du + \int_{-\infty}^{+\infty} f \frac{1}{m} \frac{\partial^2 V}{\partial x \partial u} dx du = \\
&= \gamma
\end{aligned} \tag{3.18}$$

dove nell'ultimo passaggio si è sfruttato il fatto che la derivata di V rispetto a u è nulla.

3.

$$\begin{aligned}
& \gamma \frac{T}{m} \int_{-\infty}^{+\infty} \frac{\partial^2 f}{\partial u^2} \log(f) dx du = \\
&= -\gamma \frac{T}{m} \int_{-\infty}^{+\infty} \frac{1}{f} \frac{\partial f}{\partial u} \frac{\partial f}{\partial u} dx du = \\
&= -\gamma \frac{T}{m} \int_{-\infty}^{+\infty} f \left(\frac{\frac{\partial f}{\partial u}}{f} \right)^2 dx du = \\
&= -\gamma \frac{T}{m} \left\langle \left(\frac{\partial \log(f)}{\partial u} \right)^2 \right\rangle
\end{aligned} \tag{3.19}$$

Riunendo i risultati ottenuti e ricordandosi che davanti all'integrale della 3.16 c'era un segno $-$ si ottiene:

$$\frac{\partial S}{\partial t} = \gamma \left[\frac{T}{m} \left\langle \left(\frac{\partial \log(f)}{\partial u} \right)^2 \right\rangle - 1 \right] \tag{3.20}$$

A questo punto, utilizzando tutti i risultati ottenuti e riconducendoci alla 3.5:

$$\begin{aligned}
\dot{Q} - T \frac{\partial S}{\partial t} &= \gamma(T - \langle mu^2 \rangle) - T\gamma \left[\frac{T}{m} \left\langle \left(\frac{\partial \log(f)}{\partial u} \right)^2 \right\rangle - 1 \right] = \\
&= -\gamma \langle mu^2 \rangle - \gamma \frac{T^2}{m} \left\langle \left(\frac{\partial \log(f)}{\partial u} \right)^2 \right\rangle + 2\gamma T = \\
&= -\frac{\gamma}{m} \left\langle \left[mu + T \frac{\partial \log(f)}{\partial u} \right]^2 \right\rangle + 2\gamma T + 2\gamma T \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} u \frac{\partial f}{\partial u} f \frac{1}{f} du = \\
&= -\frac{\gamma}{m} \left\langle \left[mu + T \frac{\partial \log(f)}{\partial u} \right]^2 \right\rangle + 2\gamma T - 2\gamma T \int_{-\infty}^{+\infty} f dx du = \\
&= -\frac{\gamma}{m} \left\langle \left[mu + T \frac{\partial \log(f)}{\partial u} \right]^2 \right\rangle \leq 0
\end{aligned} \tag{3.21}$$

Questa disuguaglianza finale permette di verificare completamente la validità del secondo principio della termodinamica in questo tipo di problema dove $V(x, t)$ è dipendente dal tempo.

Si può notare che l'uguaglianza vale se e solo se

$$mu = -T \frac{\partial \log(f)}{\partial u} \tag{3.22}$$

ovvero ciò accade se f è tale che

$$f(x, u, t) = k(x, t) e^{-\frac{mu^2}{2T}} \tag{3.23}$$

con k costante di normalizzazione. In altre parole, l'uguaglianza della 3.21 vale se la funzione di distribuzione di probabilità delle velocità coincide con quella che si ha all'equilibrio alla temperatura T fissata.

3.3 La minima generazione di calore

Riassumendo quanto detto nei paragrafi precedenti, ipotizzando che il moto della particella nella doppia buca di potenziale dipendente nel tempo sia descrivibile mediante un'equazione del moto di Langevin è possibile ricavare la forma dell'equazione di Kramers per descrivere la variazione temporale della funzione di distribuzione di probabilità associata alla posizione della particella e da queste semplici assunzioni si ricava che anche se il potenziale non è statico vale comunque il secondo principio della termodinamica. Ritornando al principio di Landauer, la doppia buca di potenziale è il modello di un dispositivo di memoria che può assumere due stati stabili. La dipendenza del potenziale dal tempo ci assicura la possibilità di passare ad uno stato finale certo (ad esempio lo stato 1) in modo tale da poter riprodurre l'operazione di cancellazione alla base del principio di Landauer. Chiamata $\Delta Q_{out}(t_i; t_f)$ l'energia sotto forma di calore dissipata nell'ambiente nell'intervallo di tempo $\Delta t = t_f - t_i$, poichè il sistema di memoria e l'ambiente esterno formano un sistema chiuso vale l'uguaglianza $\Delta Q_{out}(t_i; t_f) = -\Delta Q_{in}(t_i; t_f)$ e per la disuguaglianza 3.21 si può dire che

$$\Delta Q_{out}(t_i; t_f) = \int_{t_i}^{t_f} (-\dot{Q}_{in}) dt \geq T [S(t_i) - S(t_f)] \quad (3.24)$$

e perciò

$$\Delta Q_{out,min}(t_i; t_f) = T [S(t_i) - S(t_f)] \quad (3.25)$$

Ora, consideriamo un singolo dispositivo di memoria modellizzato da una doppia buca di potenziale, siano $f_0(x, u)$ e $f_1(x, u)$ le funzioni di distribuzione della particella quando si trova rispettivamente nello stato 0 o 1 e siano p_0 e p_1 le probabilità di trovarsi in ciascuno dei due stati. L'entropia dell'informazione iniziale risulta:

$$S_i = - \int_{-\infty}^{\infty} (p_0 f_0 + p_1 f_1) \log(p_0 f_0 + p_1 f_1) dx du \quad (3.26)$$

Supponiamo che le due funzioni di distribuzione abbiano una sovrapposizione trascurabile¹ e che quindi dell'integrale 3.26 rimangano solo i contributi che non contengono il prodotto tra $f_0(x, u)$ e $f_1(x, u)$, ovvero:

$$\begin{aligned} S_i &\approx - \int_{-\infty}^{\infty} p_0 f_0 \log(p_0 f_0) + p_1 f_1 \log(p_1 f_1) dx du = \\ &= p_0 S[f_0] + p_1 S[f_1] - p_0 \log(p_0) - p_1 \log(p_1) \end{aligned} \quad (3.27)$$

dove è stato sfruttato il fatto che l'integrale della f_i su tutto lo spazio delle fasi è pari all'unità perché è una funzione di distribuzione di probabilità normalizzata. Se ora applichiamo l'operazione di cancellazione di informazione all'istante finale la probabilità che la particella sia nello stato 1 è $p_1 = 1$ e quindi $p_0 = 0$. Allora l'entropia all'istante t_f sarà:

$$S_f = - \int_{-\infty}^{\infty} f_1 \log(f_1) dx du = S[f_1] \quad (3.28)$$

Inserendo le equazioni 3.27 e 3.28 in 3.25 si ricava che l'operazione di cancellazione di informazione libera nell'ambiente libera come minimo la seguente quantità di calore:

$$\Delta Q_{out,min}(t_i; t_f) = T (p_0 S[f_0] + p_1 S[f_1] - p_0 \log(p_0) - p_1 \log(p_1) - S[f_1]) \quad (3.29)$$

Assumiamo che $S[f_0] = S[f_1]$, in quanto nessuno dei due stati è privilegiato rispetto all'altro essendo del tutto simmetrici e si ottiene:

¹Questa approssimazione ci assicura che sia sempre possibile misurare lo stato della particella e quindi il valore del bit della memoria, ovvero che non esistano stati misti intermedi.

$$\begin{aligned}\Delta Q_{out,min}(t_i; t_f) &= T((p_0 + p_1)S[f_1] - p_0 \log(p_0) - p_1 \log(p_1) - S[f_1]) = \\ &= T(-p_0 \log(p_0) - p_1 \log(p_1))\end{aligned}\quad (3.30)$$

Se $p_0 = p_1 = \frac{1}{2}$, come accade nel caso di un singolo dispositivo di memoria che può assumere solo i valori 0 o 1 con uguale probabilità, si ricava proprio il limite inferiore dato dal principio di Landauer ovvero:

$$\Delta Q_{out,min}(t_i; t_f) = T \log(2) \quad (3.31)$$

Questo procedimento può essere generalizzato ad un array di N bit, ciascuno dei quali contribuisce all'aumento di calore nell'ambiente esterno con un ΔQ dato dalla 3.31 perciò quello totale sarà semplicemente $\Delta Q_{tot} = Nk_B T \log(2)$, dove è stata esplicitata la costante di Boltzmann per evidenziare la corretta unità di misura della quantità finale ricavata². Tale modello unidimensionale di Shizume permette di verificare partendo da poche ipotesi il principio di Landauer senza assumere a priori la validità del secondo principio della termodinamica, ma dimostrandone la correttezza anche nel caso di un potenziale dipendente dal tempo. Si può perciò affermare definitivamente che la cancellazione di un bit di informazione è *sempre* accompagnata dalla generazione di calore nell'ambiente circostante che come minimo vale $k_B T \log(2)$.

²Infatti l'entropia dell'informazione e l'entropia della meccanica statistica differiscono per una costante moltiplicativa in quanto la prima è adimensionale e la seconda ha unità di misura $\frac{J}{K}$. La costante moltiplicativa è proprio la costante di Boltzmann k_B .

Capitolo 4

Applicazione del principio di Landauer: il diavoleto di Maxwell

È interessante chiedersi concretamente quali siano le conseguenze di aver esteso la seconda legge della termodinamica alla teoria dell'informazione e se sia possibile riscontrare la "fisicità" del concetto di informazione connesso al principio di Landauer. Il primo indizio di una connessione tra informazione e termodinamica è il problema proposto nel 1867 da Clerk Maxwell conosciuto sotto il nome di "diavoleto di Maxwell" che trova soluzione proprio nell'applicazione del principio di Landauer.

Prima di affrontare il problema, tuttavia, è bene ricordare che l'operazione di cancellazione di informazione è associata ad una dissipazione di energia perchè è un'operazione *irreversibile*: una volta effettuata non è possibile ricondursi allo stato iniziale in quanto tutti i bit dello stato finale, convenzionalmente posti pari ad 1, non sono connessi biunivocamente allo stato iniziale del bit corrispondente. Sembrerebbe efficace effettuare delle operazioni anche irreversibili salvando lo stato precedente: in questo modo potremmo ricondurci allo stato precedente e quindi rendere l'operazione reversibile. Tuttavia, per salvare la cronologia degli stati precedenti è necessario un array di bit apposito e si sta solo postponendo il problema di cancellazione alla fine del procedimento: una volta arrivati allo stato finale a cui si è interessati, la cronologia salvata sarà cancellata irreversibilmente in modo da poter essere usata nelle operazioni successive, a meno che non si abbia un dispositivo con memoria infinita che dal punto di vista pratico è impossibile.

Nel problema del diavoleto di Maxwell è fondamentale il concetto di irreversibilità: è proprio questo che permetterà la risoluzione di questo paradosso che per moltissimi anni ha messo in crisi la validità del secondo principio della termodinamica.

4.1 Il diavoleto di Maxwell

Il diavoleto di Maxwell è una creatura intelligente capace di monitorare delle singole molecole di gas contenute in una scatola divisa in due sottocamere adiacenti, come mostrato in figura 4.1. Le due camere all'inizio hanno uguale temperatura che, dalla teoria cinetica dei gas, si sa essere connessa alla velocità quadratica media delle molecole dalla formula

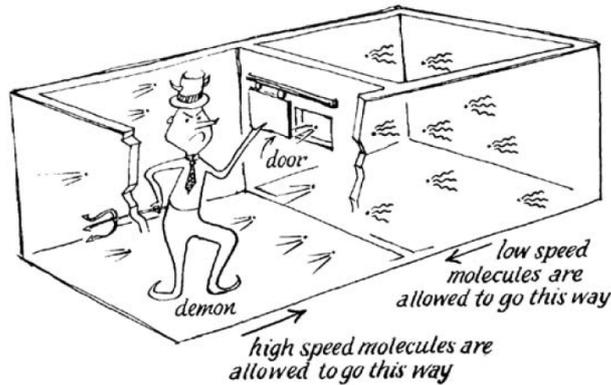


Figura 4.1: Diavoletto di Maxwell (1867)

$$T = \frac{mN_a \langle v \rangle^2}{3R} \quad (4.1)$$

dove m è la massa delle molecole, N_a il numero di Avogadro, R la costante del gas perfetto. Essendo fissa la velocità quadratica media, è del tutto lecito affermare che alcune molecole si muovano più velocemente delle altre: il compito del diavoletto è quello di regolare il passaggio di molecole attraverso la porta separatoria delle due sottocamere facendo passare da una camera all'altra in una direzione solo le molecole più veloci e nell'altra direzione solo quelle più lente. Alla fine del processo le due camere avranno due gas distinti con due differenti temperature e questo può essere sfruttato per compiere un lavoro sull'ambiente esterno. Il secondo principio della termodinamica vieta l'esistenza di un apparato che sia capace di effettuare tale processo: il diavoletto, ammettendo che la porta di passaggio si muova senza attrito, crea un flusso di calore da una sezione all'altra che continua anche dopo che una delle due ha assunto una temperatura più alta dell'altra. Tutto questo provoca una diminuzione di entropia della zona con molecole a velocità bassa maggiore dell'aumento di entropia nella zona con molecole a velocità alta. Perciò, convertendo l'informazione acquisita (riguardante posizione e velocità di ogni particella) in energia, il diavoletto sembra riuscire a far diminuire l'entropia del sistema senza compiere un lavoro in contraddizione con il secondo principio della termodinamica.

4.2 La macchina di Szilard

Si può vedere in maniera più semplice questo paradosso facendo riferimento ad un particolare modello di diavoletto ideato nel 1929 da Szilard, in cui è esplicito il lavoro meccanico che si può trarre dalla situazione sopra descritta. Il diavoletto di Szilard (si veda in figura 4.2) considera un gas formato da una sola molecola e il piano separatorio delle due camere è sostituito da un pistone. Il diavoletto ha il compito di identificare la posizione della particella che, come in un array di due bit, occupa ciascuna sezione con probabilità $p = \frac{1}{2}$. Se la molecola si trova nella camera di sinistra il diavoletto attacca un peso a sinistra del pistone: capiterà che la molecola abbia degli urti contro il pistone mobile, il quale quindi tenderà a spostarsi verso destra fino ad arrivare alla parete finale e questo movimento

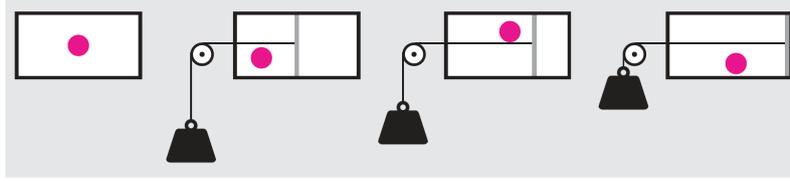


Figura 4.2: Diavoletto di Szilard (1929)

comporta il sollevamento del peso attaccato senza lo svolgimento di un lavoro perchè il pistone si muove nel vuoto (si ricordi che nella seconda regione della camera non è presente gas). Se il diavoletto individua la molecola nella parte destra avviene lo stesso processo in maniera speculare.

4.3 Risoluzione del paradosso

Il principio di Landauer è fondamentale per la risoluzione di questo paradosso. Il primo ad associare la teoria dell'informazione al problema del diavoletto fu Bennett nel 1982, notando che uno dei punti centrali della situazione sopra descritta è l'acquisizione e la memorizzazione dell'informazione effettuata dal diavoletto: esso deve immagazzinare ciò che ha acquisito durante la misurazione per poterne trarre vantaggio e ciò lo si vede bene nell'apparato di Szilard dove è necessario ricordare in quale delle due parti si trova la molecola di gas per aggiungere il peso dalla parte corretta. L'atto irreversibile è costituito dalla cancellazione dell'informazione acquisita alla fine di un ciclo per riportare la memoria ad uno stato standard in preparazione alla prossima misura, invece la misurazione in sé non genera dissipazione. La raccolta di informazione è assolutamente necessaria per l'operazione di diminuzione dell'entropia del sistema ed è la sua operazione di reset alla fine del ciclo che comporta un aumento di entropia dell'ambiente tale da compensare la diminuzione di quella del sistema, ovvero tale che $\Delta S_{tot} = \Delta S_{sist} + \Delta S_{ext} \geq 0$.

Il diavoletto di Maxwell è quindi un'importante applicazione del principio di Landauer: portare la macchina di Szilard allo stato iniziale corrisponde ad una compressione nello spazio delle fasi dai due stati possibili ad uno solo coincidente con quello standard iniziale. Questa compressione porta ad una dissipazione di energia sotto forma di calore nell'ambiente e, se si immagina il diavoletto come un dispositivo tecnologico dotato di una memoria di n bit, il calore rilasciato è $Q \geq nk_b T \log(2)$.

Ma è necessario cancellare la memoria? E se il diavoletto avesse memoria molto grande e non fosse necessario cancellarla ad ogni ciclo? In questo caso si sta solo rimandando il problema della cancellazione che prima o poi deve avvenire e in questo caso il ciclo non sarebbe più un ciclo perchè la memoria acquisirebbe ogni volta un bit di informazione in aggiunta a quelli precedenti. È come se il dispositivo per non far aumentare l'entropia dell'ambiente facesse aumentare la propria e questa cosa, concretamente, danneggia il processo analizzato. In questo caso, il possesso da parte del diavoletto di un'informazione ormai inutile per il processo diventa uno svantaggio: al contrario, il costo del reset eguaglia energeticamente il beneficio che il diavoletto ne ha tratto memorizzandola.

Capitolo 5

Verifiche sperimentali

Il diavoletto di Maxwell nacque come esperimento mentale e rimase tale per quasi un secolo. Anche una sua versione più meccanica, come la macchina ideata da Szilard, funziona con un gas formato da una sola molecola e nella pratica era impossibile realizzare questo apparato: ciò che si poteva gestire con l'esperimento era un gas dal punto di vista macroscopico, non microscopico. Tuttavia, con lo sviluppo tecnologico degli ultimi decenni è stato possibile realizzare esperimenti con singoli atomi e piccole particelle, rendendo possibile la realizzazione di macchine che hanno potuto verificare sperimentalmente il principio di Landauer.

5.1 L'esperimento di Raizen

Uno dei primi esperimenti fu quello realizzato da Mark Raizen e i suoi collaboratori nell'Università del Texas, è possibile vederlo schematizzato nella figura 5.1. Alcuni atomi sono confinati in una trappola magnetica ovvero un dispositivo sperimentale capace di raffreddare atomi a temperature dell'ordine dei μK attraverso delle combinazioni di fasci laser e un quadrupolo magnetico. Inizialmente gli atomi occupano tutti lo stesso stato cioè gli elettroni che li compongono hanno tutti una stessa configurazione energetica. Due fasci laser accostati costituiscono la barriera ottica: un fascio promuove gli atomi in uno stato eccitato (nel disegno è quello rosso) e l'altro (nel disegno è in arancione) è regolato in modo tale da avere effetto solamente per gli atomi eccitati, che vengono respinti, ma non agisce su quelli allo stato fondamentale.

Un atomo dopo aver interagito con il laser rosso si trova nello stato eccitato (il laser lo eccita per assorbimento di un singolo fotone) e interagendo con il fascio arancione rimane intrappolato nella seconda metà della camera. Un atomo che invece incontra prima il fascio arancione non interagisce con esso e continua il suo moto. Alla fine capiterà che tutti gli atomi incontreranno per primo il laser rosso e poi quello arancione e per questo si concentreranno tutti sulla seconda metà della trappola. La loro temperatura è uguale a quella iniziale ma occupano la metà del volume. Se si spengono i laser, gli atomi tornano ad occupare il volume iniziale e nell'espansione il gas si raffredda. In più, gli atomi dopo un certo tempo torneranno allo stato fondamentale con l'emissione di un fotone in direzione casuale.

Questo dispositivo, di fatto, si comporta come un diavoletto di Maxwell: la porta unidirezionale è costituita dai laser. L'informazione prima acquisita e poi cancellata che permette

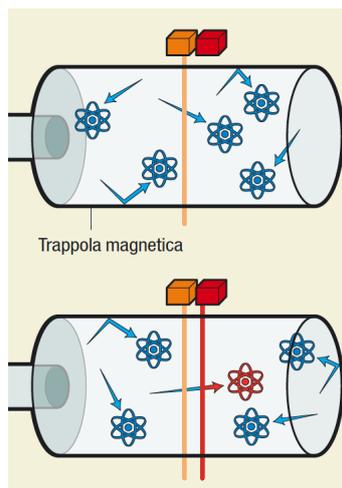


Figura 5.1: Modello dell'esperimento di Raizen

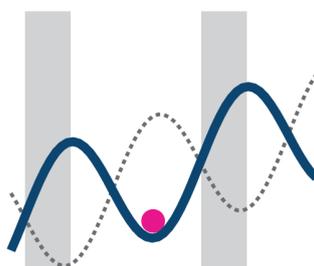


Figura 5.2: Potenziale per la macchina di Szilard

di selezionare gli atomi in questo caso, fisicamente, corrisponde al fotone assorbito dall'atomo di gas e poi rilasciato in direzione casuale. Ogni volta che un atomo diffonde un fotone, ritornando allo stato fondamentale, l'entropia del sistema aumenta in quanto il fotone che precedentemente si stava propagando coerentemente al fascio laser ora si propaga in direzione casuale. Il gruppo di lavoro di Raizen nel 2006 riuscì a dimostrare che l'aumento di entropia nell'ambiente è sempre maggiore della riduzione di entropia prodotta dal processo di raffreddamento atomico. Il diavoletto esiste e funziona: l'informazione acquisita e poi rilasciata permette di confermare la validità del secondo principio della termodinamica per questo tipo di macchine.

5.2 Verifica della macchina di Szilard

Nel 2010 il gruppo sperimentale di Masaki Sano all'Università di Tokyo realizzò una macchina di Szilard con una singola particella colloidale con dimensioni dell'ordine dei μm immersa in un bagno fluido di temperatura uniforme. La particella, intrappolata su un potenziale ottico periodico schematizzato in figura 5.2, si muove per fluttuazioni energetiche. Monitorando la particella con un dispositivo CCD è possibile distinguere i movimenti in avanti della particella da quelli all'indietro: in particolare, inserendo una barriera ottica dietro la particella ogni volta che essa avanza si è sicuri di farla avanzare solo in avanti in ogni ciclo. In questo modo la particella compie un lavoro contro la forza esercitata dalla barriera di potenziale: questo sistema usa l'acquisizione dell'informazione riguardante il

verso del moto casuale della singola particella per estrarre energia da un sistema immerso in un ambiente termico uniforme.

5.3 Verifica del principio di Landauer

Nel 2011 un gruppo sperimentale alla scuola Normale superiore di Lione realizzò l'esperimento mentale di Landauer utilizzando sempre una particella colloidale intrappolata in una doppia buca di potenziale prodotta e modulata da due fasci laser. Questo sistema, già visto in figura 3.1, ha due stati stabili: quello di destra e quello di sinistra. Ciò che venne ricreato fu perciò proprio il modello unidimensionale di dispositivo di memoria che può essere usato per immagazzinare un bit di informazione: questo può essere anche cancellato portando la particella ad uno stato finale fissato. Inizialmente la barriera tra le due buche può essere abbassata gestendo l'intensità dei laser, poi una leggera forza indotta da un motore piezoelettrico porta la particella verso destra e infine il potenziale ritorna alla forma iniziale. Come voluto, la particella si trova allo stato finale prestabilito a prescindere dal suo stato iniziale. Per un ciclo completo, il calore dissipato nell'ambiente è uguale al lavoro impiegato per modulare il potenziale. E secondo le leggi della termodinamica il calore medio dissipato ottenuto mediando le varie traiettorie della particella deve essere sempre maggiore di $\langle Q \rangle \geq -T\Delta S = k_b T \log(2) = \langle Q \rangle_{Landauer}$. Ovviamente vengono mediate tutte le traiettorie possibili quindi sia quelle in cui la particella parte dalla posizione 0 sia dalla posizione 1. Tuttavia, poichè era inconveniente selezionare casualmente la configurazione iniziale della particella, i due casi vennero trattati separatamente. Il gruppo sperimentale riuscì a valutare la quantità di calore dissipata nel seguente modo: la modulazione dell'altezza della barriera venne fatta in maniera molto lenta (o comunque in un tempo molto più lento del tempo di reazione della particella), per questo motivo è un'azione quasi-reversibile e non contribuisce alla dissipazione di calore; ciò di cui si terrà conto sarà solo il contributo derivante dalla forza esterna che agisce indipendentemente dal fatto che la pallina sia in posizione 1 o 0. Visualizzando in un grafico la probabilità di successo nel reset (ovvero la probabilità che la particella alla fine del processo si trovi nella posizione 1) in funzione del modulo della forza esterna, si nota che tale percentuale satura attorno al 95% (figura 5.3 (a)). Regolato opportunamente il modulo della forza, tracciando il grafico della quantità di calore dissipato per la transizione $0 \rightarrow 1$ si verifica che il valor medio è sempre maggiore o uguale al limite imposto da Landauer: ad esempio, in figura 5.3 (b) la linea tratteggiata corrisponde al limite di Landauer, quella rossa al valor medio della distribuzione di dati considerati. Venne replicato lo stesso processo per l'analisi della transizione $1 \rightarrow 1$ e si provò anche a svolgere i processi in tempi più lunghi modificando il modulo della forza e quindi la probabilità di successo dell'operazione. Quest'ultima analisi è visibile nella figura 5.3 (c), in cui è rappresentata la quantità di calore dissipato in funzione della durata τ del protocollo e la linea tratteggiata è il limite di Landauer: l'andamento atteso è verificato perchè aumentando τ ci si avvicina ad avere sempre più un processo quasi-reversibile e quindi ci si avvicina sempre più al limite di Landauer. In particolare i dati possono essere interpolati con la funzione

$$\langle Q \rangle = \langle Q \rangle_{Landauer} + \left[A e^{-\frac{t}{\tau_k}} + 1 \right] \frac{B}{\tau} \quad (5.1)$$

dove τ_k è il tempo di rilassamento di Cramer.

I risultati dell'esperimento indicano che si tende al limite termodinamico connesso alla

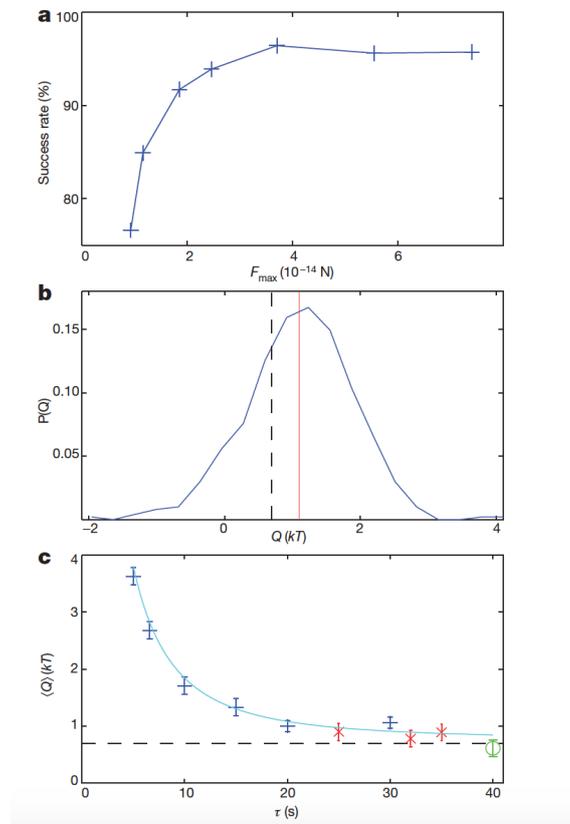


Figura 5.3: Grafici dell'esperimento per la verifica del principio di Landauer

cancellazione di informazione, il limite di Landauer, se si verificano le condizioni di regime quasi statico. Questo esperimento conferma definitivamente, indipendentemente dal dispositivo che si sta utilizzando e dall'operazione che si sta compiendo, i limiti fisici della computazione irreversibile.

Capitolo 6

Conclusioni

Dopo aver introdotto in generale il principio di Landauer nella teoria dell'informazione e aver presentato il concetto di entropia in questo ambito, ne è stata riconosciuta l'equivalenza con l'entropia termodinamica. È stato possibile dimostrare rigorosamente la validità del secondo principio della termodinamica nel modello unidimensionale di Shizume, ovvero nel caso di una particella in un potenziale dipendente dal tempo in presenza di viscosità e rumore bianco gaussiano. Con quest'ultimo, facendo poche assunzioni iniziali, si è potuta verificare la disuguaglianza alla base del principio di Landauer confermando analiticamente che la quantità di calore dissipato nei processi di cancellazione di un bit di informazione è $Q \geq k_b T \log(2)$. Infine, è stato possibile riconoscere l'importanza delle applicazioni di questo principio e con lo sviluppo tecnologico degli ultimi anni gli esperimenti mentali del diavoleto di Maxwell e la macchina di Szilard sono diventati realtà.

Grazie al principio di Landauer è stato possibile estendere il secondo principio della termodinamica ampliandone, così, il dominio di applicabilità al mondo dell'informazione.

Come si è visto, l'informazione è un'entità fisica: la computazione ha a che fare con i gradi di libertà di un sistema ed obbedisce a leggi fisiche tramite dispositivi e circuiti che ne implementano la realizzazione. Inoltre, si è visto che l'entropia di informazione è matematicamente equivalente all'entropia termodinamica ma, soprattutto, si è dimostrato che l'informazione, attraverso un processo di cancellazione, può essere convertita in energia.

Bibliografia

- [1] Landauer, *Information is Physical*, Physics Today 44, 5, 23 (1991).
- [2] Landauer, *Irreversibility and Heat Generation in the Computing Process*, IBM Journal of Research and development 5, 3, July (1961)
- [3] Bennett, *The Thermodynamics of Computation*, International Journal of Theoretical Physics 21, 12 (1982)
- [4] Shizume, *Heat generation required by information erasure*, Physical Review E 52, 4 (1995)
- [5] Jaynes, *Information Theory and Statistical Mechanics*, The Physical Review, Vol. 106, 4, May 15 (1957)
- [6] Pressè, Gosh, Lee, Dill, *Principles of Maximum entropy and maximum caliber in statistical physics*, Review of Modern Physics, Vol 85, July-September (2013)
- [7] Bérut, Arakelyan, Petrosyan, Ciliberto, Dillenschneider, Lutz, *Experimental verification of Landauer's principle linking information and thermodynamics*, Nature, 483, 187-189, March (2012)
- [8] *Maxwell's Demon 2 - Entropy, Classical and Quantum Information, Computing*, Institute of Physics Publishing, Bristol and Philadelphia (2003)
- [9] Lutz, Ciliberto *Information: from Maxwell's Demon to Landauer's eraser*, Physics today, 68, 9, 30 (2015)
- [10] Raizen *Diavoli, entropia e la corsa verso lo zero assoluto*, Le Scienze, 513, 58-63 Maggio (2011)
- [11] Bennet, Landauer *I limiti fisici fondamentali del calcolo*, Le Scienze, 205, 50-58 Settembre (1985)