

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

Scuola di Scienze  
Dipartimento di Fisica e Astronomia  
Corso di Laurea in Fisica

# Entropia come Missing Information

Relatore:  
Prof.ssa Elisa Ercolessi

Presentata da:  
Alessandro d'Agostino

Anno Accademico 2017/2018



# Indice

<b>1</b>	<b>Elementi di teoria dell'informazione</b>	<b>6</b>
1.1	La teoria dell'informazione di Shannon . . . . .	6
1.1.1	Mutual Information . . . . .	11
1.2	Entropia come Missing Information . . . . .	13
<b>2</b>	<b>Transizione tra MI generale e MI termodinamica</b>	<b>14</b>
2.1	Da una sola a tante particelle . . . . .	15
2.2	Dalle posizioni discrete ad uno spazio continuo . . . . .	18
2.3	La MI dell'impulso . . . . .	19
2.4	La relazione tra le due MI . . . . .	20
2.5	Confronto con le altre definizioni di Entropia . . . . .	21
<b>3</b>	<b>Termodinamica</b>	<b>23</b>
3.1	Termodinamica di un gas perfetto . . . . .	23
3.2	Il principio di massimizzazione dell'entropia . . . . .	25
3.2.1	Indipendenza dall'energia . . . . .	25
3.3	I postulati della termodinamica . . . . .	26
3.3.1	Derivazione delle leggi della termodinamica . . . . .	28
<b>4</b>	<b>Temperature Negative</b>	<b>30</b>
4.1	Sistemi ad energia vincolata . . . . .	30
	<b>Bibliografia</b>	<b>37</b>

# Introduzione

Lo studio della termodinamica vide i suoi primi sviluppi formali all'inizio del XIX secolo per cercare di sistematizzare il funzionamento teorico dei macchinari a vapore.

Sadi Carnot, nel 1824, fu il primo a dimostrare che si potesse ottenere lavoro dallo scambio di calore tra due sorgenti a temperature differenti e fu lui ad introdurre il concetto di *rendimento termodinamico* di un ciclo ideale di trasformazioni, noto come ciclo di Carnot nel suo omonimo teorema.

Solo successivamente Lord Kelvin introdusse il concetto di temperatura assoluta di un sistema, riuscendo così a dare una formulazione della seconda legge della termodinamica nel 1848, e J. P. Joule dimostrò l'uguaglianza delle due forme di energia note come calore e lavoro, abbattendo il concetto di *fluido calorico* (1850).

Furono le considerazioni riguardo all'irreversibilità di un processo termico che portarono allo sviluppo dell'idea di entropia. Il calore venne prima descritto come una forma di energia *secondaria*, cioè che non può essere trasformato completamente in un'altra forma di energia, come il lavoro, senza una qualche dissipazione. Solo successivamente, con i lavori di Clausius (1854), si trovò un'espressione matematica dell'asimmetria tra calore e lavoro che riuscisse a spiegare come individuare processi reversibili o irreversibili tramite la relazione:

$$\int \frac{dQ}{dT} \geq 0. \quad (1)$$

Anche Clausius riconobbe solo successivamente (1865) in questa disuguaglianza una differenza di una certa funzione di stato tra gli stati finale ed iniziale del sistema attraverso la trasformazione termica reversibile, per cui:

$$S - S_0 = \int \frac{dQ}{dT} \geq 0. \quad (2)$$

Clausius propose il nome *entropia* per questa grandezza in maniera poco convinta nelle ultime pagine del suo lavoro coniato un nuovo termine di etimologia greca da *en*, "dentro", e *tropé*, "trasformazione". Aveva trovato un modo di spiegare l'irreversibilità di un processo ma non si era mai interrogato riguardo la natura di questa irreversibilità.

L'interpretazione del significato fisico dell'entropia rimase sempre avvolto da un velo di incomprensione trovando per molto tempo una connessione con la *misura* del

*disordine* di un sistema termodinamico. Il disordine di un sistema però non risulta essere una quantità oggettiva e misurabile in maniera coerente. Questa associazione non aiutò quindi la comprensione approfondita dell'entropia.

Fu Boltzmann, a seguito dell'affermarsi dell'interpretazione atomistica della materia, a trovare una connessione tra la dinamica delle particelle che compongono un sistema e l'entropia di tale sistema, riuscendo a dare una spiegazione dell'irreversibilità di un processo. Altro contributo fondamentale di Boltzmann all'istituzione dell'entropia fu la famosa formula

$$S = k_B \ln W, \quad (3)$$

che trovò nei successivi lavori di Planck sul corpo nero la sua prima vera e propria applicazione, affermandone l'importanza.

Uno degli obiettivi di questa tesi è vedere come questo concetto di entropia sia stato rivoluzionato dallo sviluppo della teoria dell'informazione svolto a metà del XX secolo da Shannon e Von Neumann permettendo a questa grandezza di godere di una ben maggiore comprensione e applicabilità. Le somiglianze tra entropia e un certo tipo di misura della quantità di informazione di un sistema, come vedremo, sono tante eppure non è stato così immediato tracciare un'equivalenza tra questi due concetti, che effettivamente hanno origini molto distanti.

Probabilmente il fatto che il concetto di entropia sia stato definito prima dell'affermarsi della teoria atomistica ha reso più accidentato questo percorso. Dai lavori di Clausius infatti vediamo che l'unità di misura da attribuire all'entropia sia quella dei  $\left[\frac{J}{K}\right]$  e invece, dalla teoria cinetica dei gas, riconosciamo la temperatura come una quantità di energia moltiplicata per la costante di Boltzmann  $k_B$ . Se ipotizzassimo di misurare la temperatura non più in kelvin  $[K]$  ma in quantità di energia  $[J]$ , come siamo in grado di giustificare a seguito della teoria cinetica dei gas, l'unità di misura da attribuire all'entropia non sarebbe più  $\left[\frac{J}{K}\right]$  ma essa diventerebbe invece una quantità adimensionale, così come la quantità di informazione. Questa osservazione, seppur di carattere storico, ci aiuta molto ad avvicinare la teoria dell'informazione allo studio dell'entropia.

L'adimensionalità di questa grandezza permette di riconoscere un'entropia in un qualsiasi sistema descritto tramite una distribuzione di probabilità, come vedremo nel primo capitolo. Tramite il principio di massimizzazione dell'entropia vedremo come sfruttare questa grandezza per esplorare e compiere previsioni sul comportamento di questi sistemi.

Nel primo capitolo di questa tesi ripercorreremo la fondazione della teoria dell'informazione e prenderemo familiarità con i concetti necessari per tracciare questa corrispondenza tra entropia e quantità di informazione, in particolare di informazione mancante di un sistema.

Nel secondo capitolo invece vedremo quali sono i passaggi necessari per descrivere compiutamente l'entropia di un sistema termodinamico semplice, come il gas perfetto, partendo da ipotesi esclusivamente di natura informazionale.

Nel terzo capitolo invece vedremo come, tramite il principio di massimizzazione dell'entropia calato a *postulato* della termodinamica, sia possibile costruire completamente una teoria della termodinamica del tutto compatibile con i risultati classici.

Infine nel quarto capitolo vedremo più nel dettaglio un'applicazione di questa nuova formulazione di entropia, più ricca della precedente. Affronteremo lo studio di un semplice sistema che permette di essere descritto tramite l'utilizzo di temperature negative.

# Capitolo 1

## Elementi di teoria dell'informazione

La teoria dell'informazione fu formulata da Shannon per primo nel suo lavoro [10] in cui affronta il problema dell'ottimizzazione di una rete di comunicazione. Per riuscire nel suo intento Shannon razionalizza la definizione di informazione e ne definisce una misura, che risulta essere simile ad una grandezza di natura completamente diversa: l'entropia. In questo capitolo cerchiamo di dare un'introduzione a questo argomento seguendo il filo logico suggerito nel [1, Cap. 3].

### 1.1 La teoria dell'informazione di Shannon

Il concetto di informazione, non essendo di origine prettamente scientifica, si presta a diverse interpretazioni più o meno soggettive. Per poter sviluppare una teoria rigorosa che utilizzi tale grandezza quindi è necessario chiarire che cosa si intenda per informazione. L'accezione con cui viene utilizzata è quella di *quantità di informazione* senza mai andare ad indagare il significato veicolato da tale informazione. La teoria dell'informazione infatti nasce dagli studi di Von Neumann e di Shannon [10] a metà del 20-esimo secolo per cercare di studiare e di ottimizzare le reti di comunicazione in quanto mezzi di trasporto dell'informazione. Per compiere questo tipo di analisi è necessario riconoscere che il significato di una certa informazione risiede sia nel mittente, che nel messaggio ed infine nel ricevente e non è strettamente legato alla forma o alla lunghezza del messaggio: un messaggio scritto in una lingua a noi sconosciuta, seppur redatta con un senso compiuto, può risultare assolutamente priva di significato.

È necessario focalizzarsi solo sulla lunghezza del messaggio trasferito senza alcun riferimento al significato di tale messaggio. È conveniente pensare e misurare questa lunghezza in bit che, allora come adesso, sono l'unità base di qualsiasi processo e sistema d'informazione. Un messaggio scritto in una qualsiasi lingua o sistema di codificazione può essere inteso come una stringa ordinata di diversi caratteri. La disposizione e la frequenza di questi caratteri però spesso non risulta casuale ed equiprobabile. Di solito

troviamo dei percorsi preferenziali tra le diverse lettere e alcuni percorsi del tutto interdetti. Nella lingua inglese per esempio si può vedere come la lettera  $E$  appaia molto più spesso della lettera  $Q$  e che alcune brevi sequenze, come  $TH$  appaiano molto più spesso di altre, come ad esempio  $XP$ . L'analisi dei percorsi più probabili permette di ottimizzare la trasmissione dei messaggi. Una maniera che si può utilizzare per comprimere la quantità di informazione da trasportare può essere quella di indicare le parole e le sequenze più probabili tramite simboli più brevi e invece di rilegare i simboli più lunghi a sequenze meno probabili. In questa maniera il messaggio riformulato risulta mediamente contratto. Questo procedimento è possibile solo se possiamo sfruttare gli squilibri tra le frequenze dei diversi caratteri o sequenze che vanno a comporre un linguaggio. Nel caso il linguaggio in questione fosse invece formato da sequenze del tutto equiprobabili riassegnare taglie diverse a diverse sequenze non darebbe garanzia di contrarre il messaggio. L'ottimizzazione di un messaggio permette quindi in sostanza di poter ridurre la quantità media di bit necessaria per rappresentarlo, sfruttando la distribuzione di probabilità delle diverse sequenze del linguaggio.

Quello che propone Shannon in *The Mathematical Theory of Communication* è un sistema di misura dell'informazione. I suoi studi in questo lavoro proseguono sull'analisi dell'influenza del rumore su una rete di comunicazione, ma quello a cui siamo interessati è questa parte preliminare. Il sistema per quantificare l'informazione che Shannon individua risulterà formalmente identico, per forma e proprietà, ad una grandezza ben nota in termodinamica: l'entropia. La scelta del nome per tale grandezza non fu scontata, ed è interessante cogliere un estratto di una conversazione tra Shannon e Von Neumann [11] in cui si parla proprio del nome da attribuire a tale grandezza:

*« La mia più grande preoccupazione era come chiamarla. Pensavo di chiamarla informazione, ma la parola era fin troppo usata, così decisi di chiamarla incertezza. Quando discussi della cosa con John Von Neumann, lui ebbe un'idea migliore. Mi disse che avrei dovuto chiamarla entropia, per due motivi: "Innanzitutto, la tua funzione d'incertezza è già nota nella meccanica statistica con quel nome. In secondo luogo, e più significativamente, nessuno sa cosa sia con certezza l'entropia, così in una discussione sarai sempre in vantaggio." »*

Se chiamiamo  $H$  questa funzione entropia possiamo ora andare ad analizzare le proprietà fondamentali che la definiscono e vedere la sua forma. Gli unici fattori da cui  $H(\dots)$  può dipendere sono le probabilità dei diversi eventi possibili. In riferimento ad un linguaggio sarebbero quindi le probabilità dei diversi percorsi tra le lettere o tra le parole. La definizione originale viene data come  $H(p_1, p_2, \dots, p_n)$ , per evidenziare la grande generalità di questo tipo di ragionamento, in funzione di una generica variabile aleatoria discreta  $X$  che si può realizzare in  $n$  diversi eventi possibili con le rispettive probabilità  $p_i$ , richiedendo solo che  $\sum_{i=1}^n p_i = 1$ . L'entropia del sistema che andiamo a valutare infatti dipende esclusivamente dalla distribuzione delle probabilità ed è completamente scollegato dal significato e dal risultato che gli eventi determinino. L'entropia ci



permette di misurare quanto sia vicina o distante la nostra distribuzione di probabilità dal caso uniforme. In altre parole se un sistema ha un'alta entropia significa che il suo stato è avvolto dalla maggior *incertezza* possibile, mentre quando un sistema ha poca entropia è come se avesse poca libertà di cambiamento ma soprattutto delle configurazioni preferenziali tra le sue opzioni, il che lo rendono meno *ignoto*. L'incertezza sullo stato del sistema viene interpretata come una grandezza del sistema stesso e non tanto come un limite conoscitivo dell'osservatore. L'entropia di Shannon ci permette di quantificarla in una maniera altrettanto oggettiva. Shannon propone di valutare il numero medio di domande binarie necessarie per riuscire a determinare lo stato della variabile. Valutiamo il numero medio in quanto abbiamo a che fare con una distribuzione di probabilità, quindi parlare di valori assoluti sarebbe limitante e poco generale e compiamo domande binarie perchè vogliamo che questa quantità di informazione venga tradotta in linguaggio binario. È anche possibile vedere come la domanda binaria sia il modo più efficace di ottenere informazione rispetto allo stato di una variabile. Nella referenza [1, Cap. 3.5] è presente una spiegazione più approfondita: le domande binarie non sono l'unico modo di ottenere l'informazione, ma sono il più veloce e quello che minimizzano il numero medio di domande richieste.

Le proprietà dell'entropia che vengono postulate nel lavoro di Shannon sono tre:

1. *Vogliamo che  $H(\dots)$  sia una funzione continua di ciascuna delle probabilità  $p_i$ .*  
Richiediamo infatti che una lieve modifica delle probabilità degli eventi si ripercuota in una piccola variazione dell'entropia.
2. *Se le probabilità  $p_i$  sono tutte uguali ( $p_i = \frac{1}{n}$ )  $H$  assume il suo valore massimo, poiché rappresenta la condizione di massima incertezza sul risultato della variabile  $X$ . Tale valore massimo deve essere una funzione monotona crescente del numero  $n$  di stati possibili.*  
Risulta chiaro che nel caso in cui tutti gli eventi siano equiprobabili ci si trovi in una situazione di massima "incertezza", che richiede quindi un maggior numero medio di domande per poter trovare lo stato occupato. Nel caso avessimo invece delle probabilità diverse di occupazione, per esempio tre eventi con rispettive probabilità  $\frac{2}{3}$ ,  $\frac{1}{6}$  e  $\frac{1}{6}$ , investigare per primo il primo evento risulterà  $\frac{2}{3}$  delle volte direttamente una scelta vincente, abbassando il numero medio di domande necessarie. Analogamente è altrettanto ragionevole pensare che all'aumentare del numero dei possibili eventi  $n$ , sia maggiore l'informazione mancante per la determinazione dello stato del sistema, cioè che sia necessario mediamente un maggior numero di domande.
3. *Se i possibili eventi della variabile  $X$  vengono separati in un numero  $n$  sottoinsiemi disgiunti l'entropia totale  $H$  deve essere una composizione pesata delle entropie  $H_i$  dei singoli sottosistemi. Questa proprietà viene detta di consistenza.*

Per chiarire meglio quest'ultima proprietà valuteremo un esempio nel prossimo

paragrafo, si può intanto dire che questa proprietà rappresenta l'indipendenza dell'entropia riguardo il raggruppamento degli eventi.

Date queste tre proprietà Shannon verifica che esista solo una la funzione adeguata e che debba avere la seguente forma:

$$H(p_1, p_2, \dots, p_n) = -K \sum_{i=1}^n p_i \log p_i, \quad (1.1)$$

dove  $K$  è una costante arbitraria che possiamo considerare pari ad 1 per ora. Si può notare come questa espressione sia effettivamente identica all'espressione di una delle possibili definizioni di entropia per un sistema termodinamico. La dimostrazione del fatto che questa sia l'unica forma possibile si può trovare in [1, App. F].

Ora che abbiamo la forma analitica della funzione  $H$  possiamo valutare una verifica più rigorosa delle proprietà richieste.

È interessante la verifica del fatto che  $H$  assumi il suo valore massimo proprio in corrispondenza di una distribuzione uniforme delle probabilità. Consideriamo una distribuzione iniziale arbitraria  $p_1, p_2, \dots, p_n$ . Se prendiamo due qualsiasi di queste probabilità,  $p_1$  e  $p_2$ , tali che  $p_1 \neq p_2$  e operiamo un piccolo cambiamento  $dp$  uguale e opposto su  $p_1$  e  $p_2$ , tenendo tutte le altre probabilità  $p_3, \dots, p_n$  inalterate. La variazione risultante su  $H$  è data da:

$$dH = \frac{\partial H}{\partial p_1} dp_1 + \frac{\partial H}{\partial p_2} dp_2 = (-\log p_1 - 1) dp_1 + (-\log p_2 - 1) dp_2, \quad (1.2)$$

Poichè  $dp_1 = dp = -dp_2$  abbiamo che  $dH = -\log\left(\frac{p_1}{p_2}\right) dp$ .

Se nella situazione iniziale abbiamo che  $p_1 > p_2$  e le modifichiamo diminuendo  $p_1$ , quindi aumentando  $p_2$ , abbiamo che  $dp > 0$  e così anche  $dH > 0$ . Analogamente se iniziamo con  $p_2 > p_1$  e le modifichiamo in modo da avvicinare i loro valori abbiamo che  $dp < 0$ , ma anche in questo caso il segno del logaritmo comporta che  $dH > 0$ . Vediamo così che ogni tentativo di livellare le probabilità, verso una distribuzione uniforme comporta un aumento dell'entropia  $H$ . Questa osservazione è coerente con il fatto che la distribuzione uniforme di probabilità sia quella che massimizza l'entropia di una variabile.

Vediamo ora la formulazione della proprietà di *consistenza*. Sia dato un insieme di  $n$  eventi  $\{A_1, A_2, \dots, A_n\}$  della variabile  $X$  e un suo diverso raggruppamento  $\{A'_1, A'_2, \dots, A'_r\}$  con  $r < n$ . Le probabilità  $p'_k$  di questi nuovi eventi saranno semplicemente la somma delle probabilità degli eventi che compongono il sottoinsieme. Denotiamo inoltre con  $m_k$  il numero di eventi di ciascun sottoinsieme, per cui  $\sum_{k=1}^r m_k = n$ . Denotiamo con  $H$  l'entropia costruita sull'insieme iniziale di  $n$  elementi e con  $H_k$  la  $k$ -esima entropia creata sul  $k$ -esimo sottoinsieme. La richiesta di consistenza si può formulare nella seguente maniera:

$$H(X) = H(p_1, p_2, \dots, p_n) = H(p'_1, p'_2, \dots, p'_r) + \sum_{k=1}^r p'_k H_k\left(\frac{p_{l_{k-1}+1}}{p'_k}, \dots, \frac{p_{l_k}}{p'_k}\right). \quad (1.3)$$

Per cui l'incertezza sul sistema originale è uguale all'incertezza sul sistema con gli eventi raggruppati più una sorta di incertezza media di ciascuno dei gruppi in cui è stato diviso l'insieme degli eventi. Si vede che questa espressione è indipendente dal tipo di raggruppamento e permette di formulare l'espressione dell'entropia anche in caso di raggruppamenti successivi e annidati.

Ora che abbiamo presentato le principali proprietà di questa grandezza vale la pena di considerare un esempio su una generica distribuzione di probabilità su  $n$  eventi possibili. Consideriamo quindi la generica variabile  $X$  e la sua distribuzione di probabilità data da  $p_i$ . Come precedentemente visto possiamo scrivere l'entropia  $H$  come

$$H = - \sum_{i=1}^n p_i \log p_i. \quad (1.4)$$

Ricordiamo che l'unico vincolo che si pone alla distribuzione di probabilità è la condizione di normalizzazione

$$\sum_{i=1}^n p_i = 1. \quad (1.5)$$

La più importante proprietà di  $H$  è che abbia un massimo quando tutte le  $p_i$  sono uguali. Per verificare analiticamente questa relazione viene usato il metodo dei moltiplicatori di Lagrange, in quanto può essere vista come una funzione vincolata dalla proprietà di normalizzazione. Definiamo quindi una funzione ausiliara data dalla composizione della nostra funzione e dalla funzione caratteristica del vincolo:

$$F = H(p_1, \dots, p_n) + \lambda \sum_i p_i. \quad (1.6)$$

Operando la derivata parziale rispetto ciascuna delle probabilità  $p_i$  e imponendo che si annullino

$$\left( \frac{\partial F}{\partial p_i} \right)_{p_1, p_{i-1}, \dots, p_{i+1}, p_n} = -\log p_i + \lambda = 0, \quad (1.7)$$

otteniamo  $n$  equazioni indipendenti della forma

$$p_i = \exp(\lambda - 1). \quad (1.8)$$

Se risostituiamo questa forma nella condizione di normalizzazione troviamo che

$$1 = \sum_{i=1}^n p_i = \exp(\lambda - 1) \sum_{i=1}^n 1 = n \exp(\lambda - 1). \quad (1.9)$$

da cui otteniamo il risultato aspettato:

$$p_i = \frac{1}{n}. \quad (1.10)$$

Confermando nuovamente che il valore massimo dell'entropia venga assunto in corrispondenza di una distribuzione uniforme delle probabilità. Possiamo inoltre valutare quale sia il valore  $H_{max}$ :

$$H_{max} = - \sum_{i=1}^n p_i \log p_i = - \sum_{i=1}^n \frac{1}{n} \log \left( \frac{1}{n} \right) = \log n. \quad (1.11)$$

Tutti i logaritmi che abbiamo utilizzato fino ad ora possono essere intesi in base 2, il che riconduce direttamente il valore dell'entropia al numero di *bit* medio necessari a rappresentare lo stato del sistema.

### 1.1.1 Mutual Information

Nei sistemi termodinamici che vogliamo andare a studiare è sempre presente un numero molto grande di particelle, è doveroso quindi compiere qualche considerazione sulle conseguenze che può avere un numero di particelle  $N > 1$ . È possibile infatti che le particelle interagiscano in qualche maniera o che comunque si influenzino a vicenda. Vedremo che la presenza di un qualsiasi tipo di relazione tra le particelle inevitabilmente ci conferisce informazione, diminuendo quindi l'entropia. Consideriamo per prima cosa due particelle  $X$  e  $Y$  con le rispettive distribuzioni di probabilità  $p_X(i)$  e  $p_Y(j)$  riferite ai rispettivi  $n$  ed  $m$  stati disponibili, che rispettino le condizioni di normalizzazione. Chiamiamo  $p(i, j)$  la probabilità congiunta che la prima particella occupi il suo  $i$ -esimo stato e che la seconda il suo  $j$ -esimo stato. La funzione  $H(X, Y)$  definita su tale distribuzione di probabilità sarà:

$$H(X; Y) = - \sum_{i,j} p(i, j) \log p(i, j), \quad (1.12)$$

Le distribuzioni di probabilità delle singole particelle e le corrispondenti entropie saranno:

$$p_i = \sum_{j=1}^m p(i, j) = p_X(i), \quad (1.13)$$

$$H(X) = - \sum_{i=1}^n p_X(i) \log p_X(i) = - \sum_{i=1}^n p(i, j) \log \sum_{j=1}^m p(i, j), \quad (1.14)$$

$$q_j = \sum_{i=1}^n p(i, j) = p_Y(j), \quad (1.15)$$

$$H(Y) = - \sum_{j=1}^m p_Y(j) \log p_Y(j) = - \sum_{j=1}^m p(i, j) \log \sum_{i=1}^n p(i, j). \quad (1.16)$$

Possiamo ora andare a vedere un'importante relazione tra le entropie che abbiamo definito:

$$\begin{aligned}
& H(X) + H(Y) \\
&= - \sum_i^n p_x(i) \log p_x(i) - \sum_j^m p_y(j) \log p_y(j) \\
&= - \sum_{i,j} p(i,j) \log p_x(i) - \sum_{i,j} p_y(i) \log p_y(j) \\
&= - \sum_{i,j} p(i,j) \log \sum_{j=1}^m p(i,j) - \sum_{i,j} p(i,j) \log \sum_{i=1}^n p(i,j).
\end{aligned} \tag{1.17}$$

Possiamo maggiorare quest'ultimo membro sfruttando alcune proprietà della funzioni concave (si veda in particolare la referenza [1, App. H]) per ottenere che:

$$H(X) + H(Y) \geq - \sum_{i,j} p(i,j) \log p(i,j) = H(X, Y). \tag{1.18}$$

Dove l'uguaglianza viene assunta solo nel caso in cui le due variabili siano completamente indipendenti:  $p(i, j) = p_X(i)p_Y(j)$  per cui:

$$H(X, Y) = H(X) + H(Y), \tag{1.19}$$

ciò dimostra che il risultato dell'evento *composto* possiede una quantità di entropia sempre minore o uguale all'entropia dei due esperimenti separati. Come detto precedentemente quindi una qualsiasi relazione tra le variabili studiate comporta una diminuzione dell'entropia del sistema.

Per quantificare questa correlazione tra le variabili è utile definire ciò che viene chiamata *mutua informazione* (*mutual information*):

$$I(X; Y) \equiv H(X) + H(Y) - H(X, Y). \tag{1.20}$$

Sostituendo le diverse espressioni otteniamo:

$$\begin{aligned}
I(X; Y) &= - \sum_i^n p_x(i) \log p_x(i) - \sum_j^m p_y(j) \log p_y(j) + \sum_{i,j} p(i,j) \log p(i,j) \\
&= \sum_{i,j} p(i,j) \log \left[ \frac{p(i,j)}{p_x(i)p_y(j)} \right] \\
&= \sum_{i,j} p(i,j) \log g(i,j) \geq 0,
\end{aligned} \tag{1.21}$$

dove  $g(i, j)$  è la *correlazione* tra i due eventi  $x_i$  e  $y_j$ . Questo fattore esprime il rapporto tra la probabilità dell'evento congiunto  $p(i, j)$  e il prodotto delle singole probabilità  $p_X(i)$

e  $p_Y(j)$ , se esso assume il valore unitario significa che gli eventi sono indipendenti, ma nel momento in cui il suo valore devia capiamo che l'avvenimento di uno dei due eventi influenza la probabilità di avvenire dell'altro. La definizione di  $I(X; Y)$  quindi misura la media del logaritmo della *correlazione* tra le due variabili  $X$  e  $Y$ . Tutte queste definizioni e considerazioni si possono facilmente generalizzare ad un numero qualsiasi di variabili discrete, proprietà che useremo nella descrizione di un sistema termodinamico composto da molte particelle.

## 1.2 Entropia come Missing Information

Per usufruire di questi studi per la descrizione di un sistema termodinamico si può compiere un passo intermedio verso la descrizione di un sistema discreto. Per esempio la descrizione della posizione di una moneta all'interno di una tra tante scatole disponibili. Tale problema non è poi così diverso dalla descrizione della posizione di una particella di gas all'interno di un contenitore. Compiendo alcuni passi di generalizzazione sarà infatti possibile andare da qui a costruire una definizione consistente di entropia termodinamica per un gas e poter costruire tutta una teoria termodinamica che ne descriva il comportamento.

Facendo riferimento al sistema della moneta nascosta in una delle scatole l'informazione che vogliamo quantificare è chiaramente a riguardo della posizione di tale moneta. Se cambiamo il numero delle scatole o le probabilità con cui la moneta occupi una scatola rispetto alle altre capiamo che questa informazione che vogliamo descrivere viene modificata, aumentando o diminuendo. Più precisamente questa informazione rappresenta il numero medio di domande binarie necessarie per determinare la posizione della moneta in ciascuna diversa configurazione del sistema. È quindi più corretto riferirsi a tale grandezza come informazione mancante *MI* (*missing information*) del sistema. Si noti che nuovamente ci riferiamo a questa informazione come un grandezza propria del sistema. È questo l'unico modo di considerare la *MI* come una grandezza oggettiva e quantificabile in maniera univoca.

L'idea sarà quella di usufruire del concetto di *MI* per andare a descrivere un sistema termodinamico. Uno degli obiettivi di questa tesi è infatti esplorare questa definizione di entropia e valutare lo sviluppo di una termodinamica consistente da questo punto di partenza oltre che di promuovere proprio questa visione dell'entropia in favore ad altre interpretazioni più comuni. Solo recentemente infatti si è cercato di abbandonare l'interpretazione dell'entropia come misura del *disordine* di un sistema. Il disordine non è una grandezza oggettiva e non gode di proprietà matematiche come l'additività per cui è un candidato equivocabile.

## Capitolo 2

# Transizione tra MI generale e MI termodinamica

Per poter applicare il concetto di *MI* ad un sistema termodinamico dobbiamo passare attraverso qualche generalizzazione, come accennato precedentemente. Questi passaggi formali e logici sono sostanzialmente quattro:

1. *Passare da un sistema composto da un piccolo numero di particelle ad un sistema termodinamico composto da un numero enorme di particelle, tipicamente almeno dell'ordine del numero di Avogadro  $10^{23}$ .*
2. *Passare da uno spazio di posizioni discrete ad una descrizione in uno spazio continuo, come lo spazio delle configurazioni.*
3. *Passare da un solo tipo di informazione (solo quella sulla posizione) a due tipi di informazione (su posizione e sulla velocità delle particelle).*
4. *Passare dal considerare tutte le fonti di informazione indipendenti tra loro al riconoscere delle connessioni tra di esse.*

Nel compiere questi passaggi seguiremo la trattazione portata avanti nel capitolo 4 di [1].

## 2.1 Da una sola a tante particelle

Nella descrizione dei sistemi termodinamici abbiamo la necessità di tenere conto di un numero di particelle molto elevato e della loro interazione reciproca, pur semplice che sia. Consideriamo quindi come prima cosa il caso di una sola particella che ha a disposizione un numero di stati equiprobabili  $M$  arbitrariamente grande. La probabilità di trovare la particella nel  $j$ -esimo stato quindi è data semplicemente da  $P_j = \frac{1}{M}$  che porta il sistema ad avere una corrispondente  $MI$  data da  $H = \log M$ . All'aumentare del numero di stati possibili  $M$  vediamo come la corrispondente  $MI$  cresca, in accordo con la seconda proprietà dell'entropia promossa da Shannon.

Se aumentiamo il numero di particelle del nostro sistema, per esempio a due, dobbiamo riformulare il calcolo delle possibili configurazioni del sistema. Consideriamo dapprima il caso di due particelle *diverse*, che etichettiamo con gli indici 1 e 2 in un numero di stati equiprobabili  $M$  arbitrariamente grande. Se presupponiamo che non ci siano interazioni tra le particelle e che ciascuno stato possa accogliere un qualsiasi numero di particelle la probabilità di trovare ciascuna configurazione è data da:

$$p_{i,j}(1, 2) = \frac{1}{M^2} = \frac{1}{M} \frac{1}{M} = p_i(1)p_j(2). \quad (2.1)$$

È chiaro come in questo caso la  $MI$  sia data da:

$$H(1, 2) = \log M^2 = \log M + \log M = H(1) + H(2). \quad (2.2)$$

L'entropia globale risulta semplicemente dalla somma delle entropie delle due singole particelle, questo è dovuto alla loro completa indipendenza.

Nel caso invece in cui le due particelle non possano condividere lo stesso stato il ragionamento cambia leggermente: il numero di configurazioni accessibili non è più  $M^2$  ma è dato dal numero di disposizioni possibili  $M(M - 1)$ . Per cui:

$$p_{i,j}(1, 2) = \frac{1}{M(M - 1)} = p_i(1)p_j(2/1), \quad (2.3)$$

dove  $p_j(2/1)$  è la probabilità condizionale di trovare la particella 2 nel  $j$ -esimo stato dato che la particella 1 occupi il  $i$ -esimo, con  $i \neq j$ . Possiamo esprimere in modo efficace questa correlazione introducendo la funzione di correlazione:

$$g_{i,j}(1, 2) = \frac{p_{i,j}(1, 2)}{p_i(1)p_j(2/1)} = \frac{M^2}{M(M - 1)} = \frac{1}{1 - \frac{1}{M}}. \quad (2.4)$$

Per cui possiamo esprimere:

$$p_{i,j}(1, 2) = p_i(1)p_j(2/1)g_{i,j}(1, 2) = \frac{1}{M} \frac{1}{M} \frac{1}{1 - \frac{1}{M}}. \quad (2.5)$$



Questa correlazione tra le particelle è dovuta alla limitatezza del numero di siti disponibili. La funzione di correlazione  $g_{i,j}(1, 2)$  infatti tende all'unità al crescere del numero di stati  $M$ .

È interessante vedere come ciò influisca sull'entropia del sistema valutando la *mutual information* tra le particelle:

$$\begin{aligned} I(1, 2) &= H(1) + H(2) - H(1, 2) \\ &= \log M + \log M - \log M(M - 1) \\ &= \log \frac{M^2}{M(M - 1)} = \log[g_{i,j}(1, 2)] = -\log\left(1 - \frac{1}{M}\right). \end{aligned} \quad (2.6)$$

Questa grandezza risulta essere sempre positiva e va quindi a ridurre l'entropia del sistema.

Se consideriamo ora il caso di due particelle *identiche* invece dobbiamo riformulare il numero di configurazioni possibili tenendo conto della ridondanza che porterebbe il considerare le diverse permutazioni di una stessa situazione. Si dovrebbero affrontare gli stessi passaggi dell'esempio precedente, ma quello che ci interessa è il caso in cui il numero di stati  $M$  è molto maggiore del numero di particelle  $N = 2$  che compongono il sistema, per cui la probabilità di una mutua occupazione del singolo stato è trascurabile ( $M \gg N > 1$ ). In questo caso il numero  $W$  di possibili configurazioni risulta essere:

$$W = \frac{M(M - 1)}{2} \xrightarrow{M \rightarrow \infty} \frac{M^2}{2}. \quad (2.7)$$

La differenza fondamentale rispetto al caso precedente è che adesso le particelle sono *indistinguibili* e non è più sensato etichettarle in una maniera diversa. Ha senso valutare quindi la probabilità che una qualsiasi particella occupi un certo stato energetico. Riusciamo in questo caso a definire la correlazione tra le particelle:

$$p_i(1) = \frac{1}{M}, \quad p_i(2) = \frac{1}{M}, \quad p_{i,j}(1, 2) = \frac{2}{M^2}. \quad (2.8)$$

Da cui:

$$g_{i,j}(1, 2) = \frac{p_{i,j}(1, 2)}{p_i(1)p_j(2/1)} = 2. \quad (2.9)$$

Questa correlazione può essere espressa anche in termini di entropia:

$$\begin{aligned} I(1, 2) &= H(1) + H(2) - H(1, 2) \\ &= \log M + \log M - \log \frac{M^2}{2} = \log 2 > 0. \end{aligned} \quad (2.10)$$

Possiamo osservare come l'entropia del sistema diminuisca solo per il fatto di aver considerato le particelle *indistinguibili*. Al contrario di quanto potrebbe sembrare, la perdita

delle etichette che individuano le singole particelle non aumenta l'informazione mancante, bensì la diminuisce. Se non c'è più modo di distinguere le particelle il sistema perde tutta quella quantità di informazione necessaria per individuarle e il numero di configurazioni del sistema diminuisce portando ad una diminuzione della  $MI$ .

L'ultimo passo è quello di generalizzare il ragionamento per un numero qualsiasi di particelle. Consideriamo un caso con  $N$  particelle ed  $M$  siti a disposizione con  $M > N$  e quindi la probabilità di occupazione multipla di un singolo sito trascurabile. Se abbiamo  $N$  particelle *distinguibili* il numero delle configurazioni possibili sarà:

$$W^D = M(M-1)(M-2)\dots(M-(N-1)) = \frac{M!}{(M-N)!}. \quad (2.11)$$

Se invece consideriamo  $N$  particelle *indistinguibili* dobbiamo eliminare tutte le configurazioni che risultano una permutazione delle disposizioni precedenti. Otteniamo quindi:

$$W^{ID} = \frac{M!}{(M-N)!N!} = \binom{M}{N}. \quad (2.12)$$

La probabilità di occupazione di una singola configurazione sarà  $\binom{M}{N}^{-1}$  e la funzione di correlazione corrispondente sarà:

$$g(1, 2, \dots, N) = \frac{p_{i_1, i_2, \dots, i_N}(1, 2, \dots, N)}{[p_1(1)]^N} = \frac{\binom{M}{N}^{-1}}{\left(\frac{1}{M}\right)^N} \xrightarrow{M \rightarrow \infty} N!. \quad (2.13)$$

Vediamo come in questo caso esista una correlazione tra le particelle, dovuta alla loro indistinguibilità, che risulta essere la naturale generalizzazione del caso semplice di due sole particelle. La variazione di entropia risultante da questa correlazione può essere quantificata:

$$\begin{aligned} \Delta H &= H^{ID}(1, 2, \dots, N) - H^D(1, 2, \dots, N) \\ &= \log \binom{M}{N} - \log \frac{M!}{(M-N)!} \xrightarrow{M \gg N} \log \frac{1}{N!} = -\log N!. \end{aligned} \quad (2.14)$$

Vediamo quindi come la rimozione delle etichette sulle particelle, cioè il passaggio da particelle distinguibili ad indistinguibili porti ad una variazione di entropia macroscopica non trascurabile. Infatti, a differenza della correlazione riguardante la posizione, che all'aumentare di  $M$  diventava trascurabile, questa non scompare secondo tale limite. È una correlazione sempre positiva e crescente con  $N$ .

Sarà fondamentale tenere conto di questo risultato nella formulazione dell'entropia di alcuni sistemi termodinamici anche semplici, come il gas perfetto. In questo paragrafo abbiamo utilizzato il concetto di particelle *indistinguibili*, che è un concetto che esula

dalla trattazione classica della meccanica. Nella meccanica classica sappiamo che tutte le particelle, anche se identiche tra loro, godono della distinguibilità mentre nella meccanica quantistica particelle identiche sono indistinguibili l'una dalle altre. Questo non è l'unico strumento di origine quantistica che utilizzeremo per formulare l'entropia di un sistema termodinamico, dovremo infatti introdurre anche il concetto di quantizzazione dello spazio delle fasi.

## 2.2 Dalle posizioni discrete ad uno spazio continuo

Per rappresentare efficacemente il comportamento di un sistema termodinamico abbiamo bisogno di descrivere le posizioni delle particelle in maniera più raffinata. L'informazione sulla posizione di una particella deve essere una variabile continua in uno spazio delle configurazioni continuo e non più una variabile discreta come fino ad ora. Se consideriamo che la posizione di una particella sia uniformemente distribuita in un segmento dello spazio lungo  $L$  possiamo dire che la densità di probabilità sarà  $p(x) = \frac{1}{L}$ . La generalizzazione più semplice, partendo dalla definizione che abbiamo dato di entropia, è quella di trasformare la sommatoria su tutti i possibili eventi in un integrale di una distribuzione continua sul segmento, per cui:

$$H = - \int_a^b p(x) \log p(x) dx. \quad (2.15)$$

Le proprietà dell'entropia in seguito a questo passaggio non rimangono del tutto inalterate. In linea di principio non è più possibile valutare il valore particolare di entropia di un sistema ma è solo possibile valutare differenze di tale grandezza<sup>1</sup>.

Nella pratica, come Shannon propose, si continua ad usare la definizione di  $MI$  come nella (3.15) per cui  $H = \log L$ . Se quindi consideriamo una particella dentro un volume cubico, per semplicità, con ogni spigolo lungo  $L$  associamo ad ognuna delle tre dimensioni una variabile aleatoria  $X$ ,  $Y$  e  $Z$  appartenenti ad un intervallo lungo  $L$  per determinare la sua posizione  $\mathbf{R}$ , per cui:

$$H(\mathbf{R}) = H(X) + H(Y) + H(Z) = 3 \log L = \log V. \quad (2.16)$$

Possiamo ora scrivere l'entropia di  $N$  particelle *distinguibili* o *indistinguibili* in un volume  $V$ :

---

<sup>1</sup>Nel passaggio tra discreto e continuo compare nell'espressione di  $H$  un termine che dipende dalla finezza con cui vogliamo individuare la posizione della particelle nel sistema, che nel limite di una precisione arbitraria diverge, rendendo di scarso utilizzo il valore esatto di  $H$ . Tale termine problematico scompare nel momento in cui si compie una differenza tra l'entropia di due diverse configurazioni del sistema. Poichè però la precisione ottenuta in un qualsiasi esperimento è ben lontana dall'essere arbitraria ci teniamo distanti dalla divergenza di tale fattore. Per una trattazione più completa indichiamo l'appendice I di [1].

$$H^D(\mathbf{R}_1, \dots, \mathbf{R}_N) = \sum_{i=1}^N H(\mathbf{R}_i) = N \log V, \quad (2.17)$$

$$H^{ID}(\mathbf{R}_1, \dots, \mathbf{R}_N) = \log \frac{V^N}{N!}. \quad (2.18)$$

Dove per le particelle *indistinguibili* abbiamo inserito la correlazione precedentemente studiata.

Siamo ora in grado di rappresentare l'entropia di un numero  $N$  di particelle descritte dal loro spazio delle configurazioni.

## 2.3 La MI dell'impulso

La descrizione in meccanica classica delle particelle fa uso di uno spazio delle fasi in cui ogni corpo può avere una singola posizione e una singola velocità. Se ci concentriamo direttamente su un sistema isolato di  $N$  particelle non interagenti possiamo utilizzare la distribuzione di Maxwell-Boltzmann per valutare la distribuzione di probabilità delle velocità delle particelle. La distribuzione di probabilità infatti è l'unica cosa di cui abbiamo bisogno per fare una valutazione dell'entropia da associarle. Shannon stesso dimostrò che tra tutte le distribuzioni possibili con una deviazione standard fissata  $\sigma$  la distribuzione gaussiana è quella che massimizza  $H$ .

Sappiamo che per un gas perfetto in equilibrio la varianza della distribuzione della velocità lungo una direzione è correlata alla Temperatura media del sistema e alla massa delle particelle come  $\sigma_{v_x}^2 = \frac{T}{m}$ , dove  $T$  è la temperatura misurata in unità di energia (equivalente quindi a  $k_B T$  con la temperatura in Kelvin). Possiamo quindi cercare tramite il metodo dei moltiplicatori di Lagrange la distribuzione che massimizzi  $H$  pur rispettando la condizione di normalizzazione e il vincolo sulla varianza:

$$\int_{-\infty}^{+\infty} p(x) dx = 1, \quad (2.19)$$

$$\int_{-\infty}^{+\infty} x^2 p(x) dx = \sigma^2. \quad (2.20)$$

La funzione risultante è:

$$p(x) = \frac{\exp\left[-\frac{x^2}{2\sigma^2}\right]}{\sqrt{2\pi\sigma^2}}. \quad (2.21)$$

Da cui troviamo l'entropia  $H$  come:

$$H = \int_{-\infty}^{+\infty} p(x) \log p(x) dx = \frac{1}{2} \log(2\pi e \sigma^2). \quad (2.22)$$

Aprondo una parentesi nel discorso è interessante notare la maniera con cui abbiamo ritrovato il risultato che ci aspettavamo. Abbiamo massimizzato l'entropia vincolandola con le sole proprietà che volevamo imporre alla distribuzione ricercata. In altre parole la distribuzione di M-B è la distribuzione che massimizza l'entropia. La ricerca delle distribuzioni di probabilità tramite questo processo di massimizzazione dell'entropia ha rivestito un ruolo fondamentale nello studio dei sistemi termodinamici, tanto da dover essere promosso a principio della termodinamica secondo alcuni [8].

In seguito a questi ragionamenti siamo pronti per scrivere l'entropia della velocità e dell'impulso di un sistema di  $N$  particelle distinguibili non interagenti all'equilibrio:

$$H^D(\mathbf{v}_1, \dots, \mathbf{v}_N) = \frac{3N}{2} \log(2\pi eT/m), \quad (2.23)$$

$$H^D(\mathbf{p}_1, \dots, \mathbf{p}_N) = \frac{3N}{2} \log(2\pi emT), \quad (2.24)$$

che abbiamo ottenuto sostituendo l'espressione di  $\sigma^2$  e moltiplicando il risultato ottenuto per un singola dimensione per  $N$ , il numero di particelle e per 3, il numero di dimensioni dello spazio degli impulsi.

## 2.4 La relazione tra le due MI

Per concludere la ricerca dell'entropia del nostro sistema termodinamico di particelle non interagenti non possiamo semplicemente sommare le due entropie provenienti dalla distribuzione delle posizioni e degli impulsi. Gli effetti del principio di indeterminazione di Heisenberg ci costringono a suddividere lo spazio delle fasi in piccoli volumi la cui dimensione è data dalla costante di Planck  $h$  e ad abbandonare la possibilità di una precisione arbitraria. Per cui alla somma delle due entropie dobbiamo sottrarre il fattore  $\log h$  moltiplicato per il numero di assi e per il numero di particelle  $N$ :

$$H^D(\mathbf{R}^N, \mathbf{p}^N) = H^D(\mathbf{R}^N) + H^D(\mathbf{p}^N) - 3N \log h, \quad (2.25)$$

dove abbiamo indicato con  $\mathbf{R}^N$  e  $\mathbf{p}^N$  i vettori  $\mathbf{R}_1, \dots, \mathbf{R}_N$  e  $\mathbf{p}_1, \dots, \mathbf{p}_N$ .

Se compiamo invece le stesse identiche considerazioni per un sistema di particelle *indistinguibili* dobbiamo inoltre sottrarre al risultato finale la correlazione già nota:

$$\begin{aligned} H^{ID}(\mathbf{R}^N, \mathbf{p}^N) &= H^D(\mathbf{R}^N) + H^D(\mathbf{p}^N) - 3N \log h - \log N! \\ &= N \ln V + \frac{3N}{2} \ln(2\pi emT) - 3N \ln h - \ln N! \\ &= N \ln \left[ \frac{V}{N} \left( \frac{2\pi emT}{h^2} \right)^{\frac{3}{2}} \right] + \frac{5N}{2}. \end{aligned} \quad (2.26)$$

Notiamo che il risultato che abbiamo ottenuto è perfettamente compatibile con l'equazione di Sackur-Tetrode, che descrive l'entropia di un gas perfetto monoatomico, ed è

frutto di considerazioni esclusivamente di tipo informazionale. L'unica differenza che riscontriamo è l'assenza della costante di Boltzmann a moltiplicare l'intera espressione, ma ciò è dovuto all'aver indicato con  $T$  la temperatura in unità di energia. Abbiamo inoltre cambiato la base del logaritmo per una maggiore aderenza con il risultato atteso. L'espressione che abbiamo ottenuto quindi è frutto di due diverse informazioni e due diverse correlazioni (quest'ultime di natura quantistica). Ricordiamo che tale identificazione tra l'entropia informazionale  $H$  e l'entropia termodinamica (solitamente indicata con)  $S$ , vale esclusivamente per i sistemi macroscopici all'equilibrio.

## 2.5 Confronto con le altre definizioni di Entropia

Possiamo dire ora di aver concluso il percorso per la definizione dell'entropia basata sulle relazioni informazionali di un sistema termodinamico. L'entropia che abbiamo così trovato ha un'origine diversa da quella delle prime definizioni individuate per esempio da Boltzmann [2] e Gibbs [7] nel 19-esimo secolo, che fondano la loro definizione su considerazioni riguardanti lo spazio delle fasi.

L'entropia di Gibbs viene definita, all'*equilibrio* come:

$$S_G = k_B \ln \left[ \frac{1}{h^{3N}} \frac{1}{N!} \int dp \int dq \Theta(E - H(q, p)) \right], \quad (2.27)$$

cioè come un'integrale sulle  $3N$  posizioni  $q$  e sui  $3N$  impulsi  $p$  nello spazio delle fasi del sistema, dove  $\Theta(\cdot)$  è la così detta *funzione scalino*,  $E$  è l'energia del sistema e  $H(q, p)$  è l'hamiltoniana del sistema. In questa definizione influiscono sull'entropia non solo i punti dello spazio delle fasi ma tutto il volume racchiuso dalla superficie equienergetica caratterizzante il sistema. Per questo motivo si fa riferimento a questa forma di entropia come *entropia di volume*.

L'entropia di Boltzmann invece, all'*equilibrio* risulta essere:

$$S_B = k_B \ln \left[ \frac{1}{h^{3N}} \frac{1}{N!} \int dp \int dq \delta(E - H(q, p)) \right], \quad (2.28)$$

cioè un integrale nello spazio delle fasi in corrispondenza della superficie equienergetica di energia  $E$ . I risultati di quest'ultima definizione, all'*equilibrio* coincidono con quelli dell'entropia di origini informazionali per come l'abbiamo definita noi nella (2.26). Il modo per valutare le possibili configurazioni di un gas perfetto può essere quello di valutare tutti i possibili stati accessibili nello spazio delle fasi data una certa energia, che all'equilibrio risulta essere un'espressione del tutto analoga a quella utilizzata nella definizione della (2.28):

$$W^{ID} = \frac{1}{h^{3N}} \frac{1}{N!} \int dp \int dq \delta(E - H(q, p)). \quad (2.29)$$

Il dibattito su quale sia la giusta definizione di entropia è ancora piuttosto acceso, poichè ognuna tra queste definizioni ha dei vantaggi e degli svantaggi in termini di qualità delle approssimazioni che è in grado di compiere.

È possibile trovare ulteriori definizioni di entropia a seconda di alcune proprietà del sistema considerato. Se l'energia del sistema non è costante ma può variare esso viene detto sistema *canonico*. Se è anche il numero di particelle a poter essere modificato diciamo che il sistema è di tipo gran canonico ed, in entrambi questi casi, è possibile dare due ulteriori definizioni di entropia che differiscono leggermente dalla definizione che stiamo seguendo in questa tesi.

Le differenze principali hanno principalmente due origini. La prima sta nell'utilizzo più o meno raffinato di alcune approssimazioni, in particolare l'approssimazione di Stirling, che permette di valutare il fattore  $\ln N!$ . La seconda differenza importante, come sottolinea Swendsen in [12], si basa sulle deviazioni statistiche dalle grandezze descritte e il confronto tra esse e le fluttuazioni termiche del sistema. Se le fluttuazioni termiche del sistema sono di un certo ordine di grandezza, tipicamente proporzionali ad  $\frac{1}{\sqrt{N}}$  allora le deviazioni statistiche delle grandezze devono essere più piccole, in modo da poter essere trascurabili e non misurabili, tipicamente dell'ordine di  $\frac{1}{N}$ . Per una trattazione più completa dell'argomento rimandiamo ad un articolo di Swendsen [14].

Diverse definizioni dell'entropia portano a proprietà di additività e di estensività diverse e a dipendenze dall'energia diverse. La relazione tra energia ed entropia è infatti la chiave per poter definire la temperatura:

$$\left(\frac{\partial S}{\partial E}\right)_{V,N} = k_B \beta = \frac{1}{T}, \quad (2.30)$$

a seconda che ci sia una relazione di monotonia o meno tra queste due grandezze è possibile definire o meno temperature assolute negative. Tramite la nostra definizione di entropia non abbiamo nessun vincolo sulla monotonia della dipendenza tra entropia ed energia, per cui siamo liberi di definire temperature negative qualora le circostanze lo permettano.

Un'ultima differenza che vale la pena di sottolineare è che l'entropia definita come in (2.26) è in grado di rappresentare il sistema attraverso il rilascio di un suo vincolo interno, cioè a seguito di una trasformazione irreversibile. Le (2.27) e (2.28) invece, essendo definite sullo spazio delle fasi, richiedono una riformulazione dopo il passaggio attraverso stati di non equilibrio. Come conseguenza del teorema di Liouville sappiamo che la densità degli stati nello spazio delle fasi è indipendente dal tempo [7], per cui (2.27) e (2.28) non percepirebbero nessuna variazione.

# Capitolo 3

## Termodinamica

Le strade per costruire una teoria della termodinamica tramite la meccanica statistica sono principalmente due.

La strada classica è quella seguita da Boltzmann [2], basata sulle equazioni del moto della meccanica hamiltoniana e la teoria dell'ergodicità. Quello che si compie è una discretizzazione degli stati energetici permessi al sistema così da poter calcolare la *molteplicità* di un certo stato del sistema. In seguito si definisce l'entropia tramite la celebre equazione:

$$S = k_B \log W, \quad (3.1)$$

dove  $W$  indica la molteplicità degli stati permessi al sistema. Dopo aver definito l'entropia si è in grado di ripercorrere tutte le derivazioni necessarie per identificare le grandezze termodinamiche con le loro descrizioni statistiche.

La seconda strada è quella che percorreremo noi. Grazie alla teoria dell'informazione formuliamo l'entropia di un sistema e poi applichiamo in principio di massimizzazione dell'entropia, che ci permette di trovare gli stessi risultati ottenuti seguendo la via classica. In questo capitolo ci concentreremo proprio sull'applicazione e sull'importanza di questo principio che appare come una nuova chiave di lettura dei problemi di natura statistica. Il principio di massimizzazione dell'entropia risulta essere infatti il miglior modo di gestire una quantità limitata di informazione riguardo un certo sistema; se questa quantità di informazione risulta sufficiente a descrivere il comportamento del sistema otteniamo una descrizione coerente con le previsioni e con le conferme sperimentali.

Il capitolo si conclude riportando quelli che Swendsen chiama *postulati della termodinamica* in [13], che propongono una base formale a questo diverso modo di procedere.

### 3.1 Termodinamica di un gas perfetto

Nel primo capitolo abbiamo capito come calcolare l'entropia di una distribuzione di probabilità data, senza ambiguità. Nel capitolo precedente invece abbiamo visto come



sia possibile ritrovare dei risultati tipici del gas perfetto applicando il concetto di entropia basato sulla sua definizione in teoria dell'informazione. Vediamo ora come procedere con lo sviluppo dell'entropia per arrivare alla distribuzione di probabilità di un gas perfetto.

Individuiamo con  $x_i$  i possibili stati del sistema ed esprimiamo con  $\epsilon(x_i)$  l'energia di tale stato e con  $p_i$  la probabilità di occupare tale stato. Vogliamo che le probabilità siano in relazione tale che

$$\sum_{i=1}^n p_i = 1, \quad (3.2)$$

e supponiamo di essere a conoscenza del valore di aspettazione dell'energia

$$\langle \epsilon(x) \rangle = \sum_{i=1}^n p_i \epsilon(x_i). \quad (3.3)$$

Quello che bisogna fare è cercare il massimo dell'entropia vincolato dalle due proprietà che vogliamo imporre sulla distribuzione tramite il metodo dei moltiplicatori di Lagrange. Quello che otteniamo è:

$$p_i = e^{-\lambda - \beta \epsilon(x_i)}, \quad (3.4)$$

dove le costanti  $\lambda$  e  $\beta$  sono determinate sostituendo la (3.4) rispettivamente in (3.2) e (3.3). Riconosciamo quindi  $\lambda$  come:

$$\lambda = \ln Z(\beta), \quad (3.5)$$

dove abbiamo indicato con  $Z(\beta)$  la *funzione di partizione*, definita come:

$$Z(\beta) = \sum_{i=1}^n e^{\beta \epsilon(x_i)}, \quad (3.6)$$

permettendoci di scrivere la distribuzione di probabilità risultante come:

$$p_i = \frac{e^{-\beta \epsilon(x_i)}}{Z(\beta)}. \quad (3.7)$$

Questo sviluppo ci permette di tracciare tutte le identificazioni con le grandezze tipiche della termodinamica, come la temperatura e l'energia libera:

$$\beta = \frac{1}{k_B T}, \quad (3.8)$$

$$\Phi = U - TS = -k_B T \ln Z, \quad (3.9)$$

e quindi ritrovare la relazione che lega l'entropia alla variazione di energia libera sulla variazione di temperatura:

$$S = -\frac{\partial \Phi}{\partial T}. \quad (3.10)$$

Si conclude in questa maniera il procedimento di ricerca della descrizione termodinamica di un gas perfetto partendo dalla descrizione informazionale dell'entropia.

## 3.2 Il principio di massimizzazione dell'entropia

Questo processo di massimizzazione dell'entropia si propone come uno strumento esplorativo molto forte, capace di riprodurre risultati noti con una grande facilità. Non si tratta di un principio fondamentale della meccanica statistica ma più di una metodologia investigativa che ci permette di sfruttare al meglio la nostra conoscenza di un sistema e di non distorcere il ragionamento con eventuali supposizioni arbitrarie inconsapevoli. Nel suo articolo [8] Jaynes analizza tale principio come strumento per costruire la meccanica statistica alternativo alla teoria degli ensemble e senza l'ausilio di teorie dell'ergodicità.

Quello che colpisce è la facilità e la naturalezza con cui si ottengono e si riconfermano risultati noti solo applicando questo principio. Data una qualsiasi distribuzione di probabilità, che rispecchi essa una distribuzione di energia, volume, pressione o altro, si costruisce la corrispondente entropia di Shannon. Imponendo su questa entropia i vincoli noti tramite il metodo dei moltiplicatori di Lagrange otteniamo l'espressione delle diverse probabilità associate ai possibili valori. I vincoli che imponiamo sono solitamente imposizioni di valori medi della distribuzione. Tipicamente siamo infatti in grado di misurare solo queste grandezze, che rimangono costanti nel tempo e invece raramente riusciamo a fornire dei valori particolari di alcune specifiche probabilità. Jaynes inoltre spiega come sia necessario utilizzare l'entropia di Shannon e non una delle altre misure dell'informazione proposte da altri in quel periodo [6], [5]. Essa è infatti l'unica che attribuisca sempre probabilità non nulle ai diversi eventi a meno che non siano esplicitamente vietati e soprattutto è l'unica formulazione che gode della proprietà dell'additività o di *consistenza* come l'abbiamo chiamata nel primo capitolo.

### 3.2.1 Indipendenza dall'energia

La potenza e la larga applicabilità di questo principio suggeriscono che l'entropia, come misura dell'incertezza su un sistema, possa svolgere il ruolo fondamentale nella ricerca delle distribuzioni di probabilità in meccanica statistica come in termodinamica statistica. Il principio di massimizzazione dell'entropia ha un'origine del tutto indipendente da altre speculazioni fisiche, come la conservazione dell'energia. Per cui se operiamo in questo ordine di idee tutte le grandezze misurabili devono ricoprire un ruolo equivalente nello sviluppo dei risultati statistici. Possiamo quindi vedere una conferma esplicita di questa equivalenza.

Supponiamo di avere un sistema descritto da un certo insieme di funzioni  $f_k(x_i)$ , di cui abbiamo i valori di aspettazione  $\langle f_k(x_i) \rangle$ . Pensiamo di variare queste funzioni di una quantità  $\delta f_k(x_i)$  in maniera indipendente per ogni  $k$  e  $i$ . Supponiamo inoltre di modificare il valore di aspettazione delle  $f_k$  in maniera indipendente dagli  $f_k$  affinché non ci sia nessuna relazione tra le  $\delta \langle f_k \rangle$  e le  $\langle \delta f_k \rangle$ . Quella che otteniamo è una distribuzione di probabilità leggermente diversa dalla precedente, in cui le diverse probabilità degli stati  $i$ -esimi vengono modificate di una quantità  $\delta p_i$  e allo stesso modo i valori dei moltiplicatori

di Lagrange  $\lambda_k$  si modificano di  $\delta\lambda_k$ . Quali sono quindi le differenze che otteniamo in termini di entropia? La funzione di partizione viene modificata di una quantità:

$$\delta\lambda = \delta \log Z = - \sum_k [\delta\lambda_k \langle f_k \rangle + \lambda_k \langle \delta f_k \rangle], \quad (3.11)$$

adoperando la relazione tra entropia e funzione di partizione:

$$\begin{aligned} \delta S &= \sum_k \lambda_k [\delta \langle f_k \rangle - \langle \delta f_k \rangle] \\ &= \sum_k \lambda_k \delta Q_k. \end{aligned} \quad (3.12)$$

La quantità

$$\delta Q_k = \delta \langle f_k \rangle - \langle \delta f_k \rangle, \quad (3.13)$$

fornisce una generalizzazione della nozione di calore infinitesimo fornito al sistema e può essere interpretato come "calore del  $k$ -esimo tipo". Nel caso in cui la funzione  $f_k$  fosse l'energia la quantità  $\delta Q_k$  coincide con il calore nel suo senso ordinario. Vediamo come il moltiplicatore di Lagrange  $\lambda_k$  sia il fattore di integrazione del  $k$ -esimo tipo di calore  $Q_k$  e che quindi sia possibile parlare di  $k$ -esimo tipo di *temperatura*, poiché

$$\lambda_k = \frac{\partial S}{\partial f_k} = \frac{1}{T_k}. \quad (3.14)$$

Esattamente come nel caso consueto una misura di temperatura è sufficiente per decretare la distribuzione di probabilità in energia di un sistema termodinamico è possibile compiere un ragionamento analogo con una qualsiasi altra quantità che goda di una conservazione. Definendo i concetti di *serbatoio di volume*, di *particelle* o di *momento* invece che di *serbatoio di calore* siamo in grado di concepire uno strumento corrispondente al *termometro*, che si basi sullo scambio di questa quantità sotto una regola di conservazione del totale, che misuri la corrispondente *temperatura*.

Quello di cui necessitiamo per sviluppare la distribuzione è infatti la misura di un valore di aspettazione  $\langle \delta f_k \rangle$  o della sua quantità statistica coniugata  $\lambda_k$ .

Ciò che ne consegue è che in linea di principio si può cercare la distribuzione di probabilità di un sistema attraverso un processo matematico del tutto analogo a quello utilizzato di solito sull'energia ma invece basato su una diversa quantità conservata.

### 3.3 I postulati della termodinamica

Attraverso quelli che chiamiamo *postulati termodinamici* riusciamo a formulare le quattro leggi della termodinamica e a definire tutte le grandezze caratterizzanti di questa

disciplina in maniera coerente ai risultati tipici. In [13] Swendsen propone una rivisitazione dei postulati originariamente formulati da H. B. Callen in [4] che chiama i postulati *essenziali* della termodinamica.

I postulati *essenziali* sono i seguenti:

1. STATI DI EQUILIBRIO

*"Esistono stati di equilibrio di un sistema macroscopico che sono caratterizzati unicamente da un piccolo numero di variabili estensive."*

Useremo quindi il termine *funzione di stato* per indicare una qualsiasi funzione che abbia dipendenza esclusivamente dai parametri estensivi. Questi parametri estensivi sono tipicamente l'energia  $E$ , il volume  $V$  e il numero di particelle  $N$ . Le grandezze intensive possono essere derivate dai parametri estensivi e, attraverso trasformazioni di Legendre potranno essere utilizzate per caratterizzare lo stato termodinamico del sistema.

2. MASSIMIZZAZIONE DELL'ENTROPIA

*"I parametri estensivi di un sistema isolato in assenza di vincoli interni assumono quei valori che, tra tutti quelli permessi, massimizzano l'entropia."*

Qui ritroviamo il principio di massimizzazione dell'entropia calato a postulato della termodinamica. Riconosciamo in questo postulato in effetti una riformulazione della seconda legge della termodinamica.

3. ADDITIVITÀ

*"L'entropia di un sistema composto è additiva sui suoi sottosistemi costituenti."*

Questa è l'unica relazione che si instaura tra le sottoparti di un sistema termodinamico ed è alla base della definizione di equilibrio termico e della costruzione di un termometro.

Con additività si intende che se sono dati due sistemi  $j$  e  $k$ :

$$S_{j,k}(E_j, V_j, N_j; E_k, V_k, N_k) = S_j(E_j, V_j, N_j) + S_k(E_k, V_k, N_k), \quad (3.15)$$

dove  $S_j$  e  $S_k$  sono le entropie dei sistemi  $j$  e  $k$ .

4. CONTINUITÀ E DIFFERENZIABILITÀ

*"L'entropia è una funzione continua e differenziabile dei parametri estensivi del sistema."*

Postulato necessario per poter costruire le relazioni analitiche tra i diversi parametri intensivi ed estensivi di un sistema termodinamico.

Dopo aver definito quelli che secondo lui sono i postulati minimi per poter descrivere un qualsiasi sistema termodinamico, Swendsen propone quelli che chiama i postulati *opzionali*, cioè quelli che può convenire tenere in considerazione in alcuni casi più specifici.

## 5. ESTENSIVITÀ

*"L'entropia è una funzione estensiva dei parametri estensivi."*

L'estensività di una funzione è intesa come un'omogeneità di primo grado:

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N). \quad (3.16)$$

Non è una proprietà vera in assoluto, ma è molto utile per studiare le proprietà dei materiali tralasciando eventuali effetti di superficie. È una proprietà vera però nel caso della descrizione di un gas perfetto e ci permette la validità dell'equazione di Eulero:

$$U = TS - PV + \mu N. \quad (3.17)$$

## 6. MONOTONIA

*"L'entropia è una funzione monotona crescente dell'energia per valori di equilibrio dell'energia."*

Questo postulato permettere di compiere l'inversione di  $S = S(E, N, V)$  per ottenere l'energia  $U = U(S, N, V)$  e di conseguenza permettere di formulare alcuni potenziali termodinamici nella maniera familiare. Nella maggiorparte dei casi è un'assunzione conveniente ma vedremo nel prossimo capitolo come questo postulato non risulterebbe compatibile con la descrizione di un sistema termodinamico tramite l'utilizzo di temperature assolute negative.

## 7. POSTULATO DI NERST

*"L'entropia di un qualsiasi sistema è non negativa."*

Si preferisce questa forma del postulato di Nerst, in alternativa alla sua formulazione secondo il limite  $(\frac{\partial S}{\partial U})_{V,N} \rightarrow \infty$ , poichè si è a conoscenza di alcuni modelli teorici, tra cui i vetri di spin, che prevedono un'entropia non nulla per  $T \rightarrow 0$ . Questo postulato oltre ad essere in accordo con la definizione di entropia data nella (1.1) è conosciuto come terza legge della termodinamica. Trova la sua vera applicazione nei sistemi quantistici, ma poichè tutti i sistemi reali hanno basi nella meccanica quantistica, si può pensare di estendere la sua validità in generale.

### 3.3.1 Derivazione delle leggi della termodinamica

Seguendo l'idea di Swendsen la base formale e concettuale della termodinamica è gettata dei postulati del paragrafo precedente e dalle corrette proprietà dell'entropia. Possiamo quindi derivare facilmente le familiari leggi della termodinamica da questa struttura.

La prima legge della termodinamica è vista come una riformulazione della conservazione dell'energia. Poiché sembra essere un'assunzione imprescindibile in qualsiasi branca

della fisica non mettiamo in dubbio la sua validità. Bisogna sottolineare però che è necessario riconoscere il *calore* come quantità di energia, come il lavoro, per poter scrivere la legge:

$$dU = \delta Q - \delta W, \quad (3.18)$$

dove  $dU$  è la variazione infinitesima di energia interna di un sistema,  $\delta Q$  è il calore scambiato con il sistema e  $\delta W$  è il lavoro compiuto da o sul sistema (l'uso del simbolo  $\delta$  indica il concetto di differenziale non esatto).

La seconda legge della termodinamica invece deriva direttamente dal secondo postulato della termodinamica, che possiamo riscrivere in maniera concisa come:

$$\Delta S \geq 0, \quad (3.19)$$

dove  $\Delta S$  indica il cambiamento di entropia in seguito al rilascio di un qualsiasi vincolo interno al sistema.

La terza legge della termodinamica invece, come discusso precedentemente, non è altro che il postulato di Nerst sulla non negatività dell'entropia di un sistema termodinamico.

Per ritrovare la legge zero della termodinamica dobbiamo riconoscere l'equilibrio termico tra sistemi come una relazione di equivalenza, cioè una relazione simmetrica, transitiva e riflessiva. Riusciamo in questo compito grazie ai postulati numero due e tre per cui due sistemi  $j$  e  $k$  in equilibrio termico debbano soddisfare la relazione:

$$\left(\frac{\partial S_j}{\partial U_j}\right)_{V_j, N_j} = \left(\frac{\partial S_k}{\partial U_k}\right)_{V_k, N_k}, \quad (3.20)$$

ricordando la (2.30) riconosciamo nella (3.20) un'uguaglianza tra le temperature dei due sistemi.

Reputiamo così concluso il percorso alternativo di costruzione della termodinamica partendo dal concetto di entropia basato sulla teoria dell'informazione.

# Capitolo 4

## Temperature Negative

Quando non imponiamo che l'entropia debba essere monotona crescente nei confronti dell'energia del sistema apriamo le porte alla possibilità di descrivere tale sistema tramite valori di temperatura negativi.

Ricorrendo la relazione:

$$\frac{1}{T} = \left( \frac{\partial S}{\partial E} \right)_{V,N} \quad (4.1)$$

risulta evidente che per alcuni sistemi particolari in cui all'aumentare dell'energia l'entropia diminuisca ci si imbatte in derivate di segno negativo e quindi in temperature negative.

### 4.1 Sistemi ad energia vincolata

La condizione fondamentale affinché ciò avvenga è che il sistema abbia un limite superiore all'energia accessibile. Possiamo per esempio considerare un semplice sistema le cui parti possono assumere solo due valori di energia  $E_-$  ed  $E_+$  con  $E_- < E_+$ . Nel momento in cui sappiamo che l'energia del sistema è al suo minimo possibile sappiamo che tutte le parti che lo costituiscono hanno assunto l'energia minima  $E_-$  per cui la  $MI$  del sistema è nulla e ciò conduce ad un valore di entropia  $S = 0$ . Viceversa se sappiamo che il sistema risiede nella sua configurazione di energia massima sappiamo che tutte le sue parti hanno l'energia massima  $E_+$  e anche in questo caso la  $MI$  è nulla, per cui  $S = 0$ . La configurazione con la massima incertezza è quella in cui l'energia del sistema è a metà del suo valore massimo per cui abbiamo il maggior valore di *missing Information*. In Figura 4.1 possiamo vedere tutte le possibili configurazioni di un esempio di questo tipo di sistemi con un numero di particelle  $N = 4$ .

Questo tipo di sistemi porta a risultati apparentemente incoerenti che vale la pena di studiare con un po' più di rigore. Prima di vedere più nel dettaglio questo tipo di sistemi

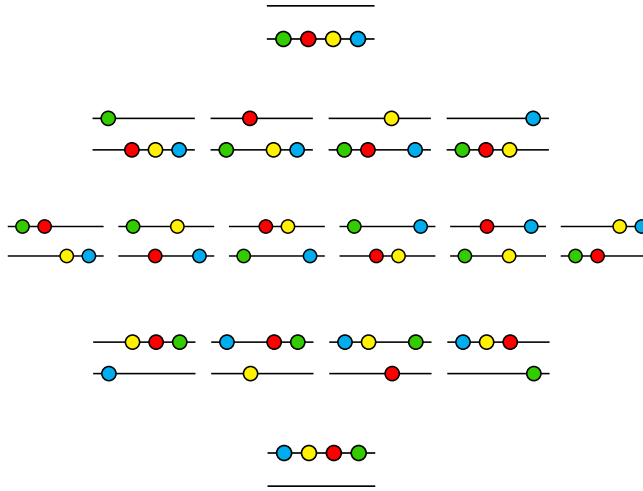


Figura 4.1: Tutte le possibili configurazioni tra due livelli energetici di quattro particelle.

citiamo l'articolo di Ramsey [9] in cui si conduce una buona analisi e giustificazione dell'introduzione delle temperature negative.

Ipotizziamo di avere un sistema composto da  $N$  particelle distinguibili con uno spettro energetico vincolato del tipo descritto precedentemente. Siano le due energie  $E_-$  ed  $E_+$  rispettivamente i valori  $-\epsilon$  e  $+\epsilon$ . Denotiamo con  $n_-$  e con  $n_+$  il numero di particelle che in un certo istante risiedono nei rispettivi livelli energetici. Se consideriamo di lavorare con un sistema in cui l'energia e il numero di particelle totali rimangano invariate nel tempo possiamo esprimere queste due quantità come:

$$E = \epsilon n_+ - \epsilon n_-, \quad (4.2)$$

$$N = n_+ + n_-. \quad (4.3)$$

Il numero delle possibili configurazioni  $\Omega(E)$  di questo sistema lo si può facilmente ricavare considerando il numero di modi di scegliere  $n_+$  oggetti tra  $N$ :

$$\Omega(E) = \binom{N}{n_+} = \frac{N!}{n_+!(N-n_+)!} = \frac{N!}{n_+!(n_-)!}. \quad (4.4)$$

Grazie a quest'ultima quantità siamo in grado di valutare la quantità di informazione nascosta nel sistema:

$$\begin{aligned} \frac{S(E)}{k_B} &= \ln \Omega(E) = \ln N! - \ln n_+! - \ln n_-! \\ &= N \ln N - N - n_+ \ln n_+ + n_+ - n_- \ln n_- + n_- \\ &= (n_+ + n_-) \ln N - n_+ \ln n_+ - n_- \ln n_- \\ &= n_+ \ln \frac{N}{n_+} + n_- \ln \frac{N}{n_-}, \end{aligned} \quad (4.5)$$



dove abbiamo utilizzato tre volte l'approssimazione di Stirling per svolgere i termini del tipo  $\ln N!$ .

Dalle (4.2) e (4.3) abbiamo inoltre che:

$$n_+ = \frac{N}{2} + \frac{E}{2\epsilon} \quad , \quad n_- = \frac{N}{2} - \frac{E}{2\epsilon}, \quad (4.6)$$

indicando quindi con  $x = \frac{E}{2\epsilon N}$  nel dominio  $-1/2 \leq x \leq 1/2$  si ottiene:

$$S = -k_B N \left[ \left( \frac{1}{2} + x \right) \ln \left( \frac{1}{2} + x \right) + \left( \frac{1}{2} - x \right) \ln \left( \frac{1}{2} - x \right) \right], \quad (4.7)$$

di cui troviamo una rappresentazione in Figura 4.2.

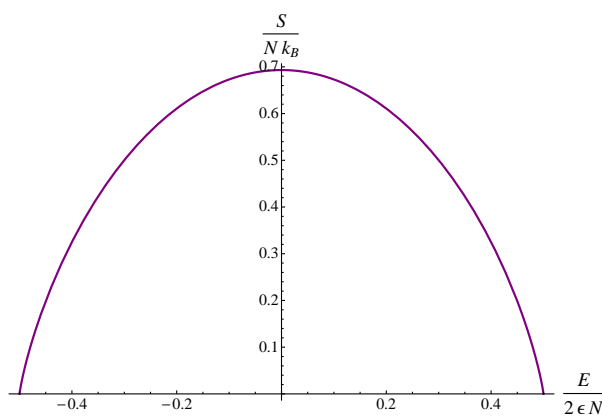


Figura 4.2: Grafico dell'entropia in funzione dell'energia del sistema.

Per trovare il massimo dell'entropia è sufficiente compiere la derivata rispetto ad  $x$ :

$$0 = \frac{dS}{dx} = -k_B N \ln \frac{\frac{1}{2} + x}{\frac{1}{2} - x}, \quad (4.8)$$

da cui individuiamo un punto di massimo per  $x = 0$  in corrispondenza ad un valore di energia  $E = 0$ . Il valore massimo di entropia ottenuto risulta essere  $S_{MAX} = k_B N \ln 2$ .

Siamo inoltre in grado di individuare la relazione che sussiste tra la temperatura e l'energia, ricordando che:

$$T^{-1} = \frac{\partial S}{\partial x} \frac{dx}{dE}, \quad (4.9)$$

otteniamo

$$T = \frac{2\epsilon}{k_B} \left( \ln \frac{\frac{1}{2} - x}{\frac{1}{2} + x} \right)^{-1}, \quad (4.10)$$

rappresentata in Figura 4.3.

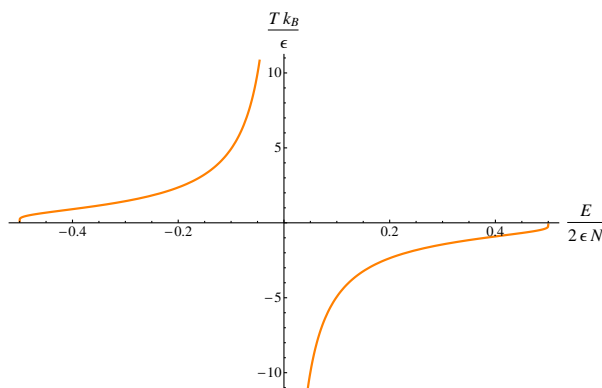


Figura 4.3: *Grafico della temperatura del sistema in funzione della sua energia.*

Vediamo che la relazione che abbiamo individuato presenta una discontinuità di secondo tipo nel punto di energia  $E = 0$ , in corrispondenza del massimo di entropia. Essendo la temperatura sostanzialmente la derivata dell'entropia ritroviamo che la temperatura assume valori negativi per valori  $0 < E < \epsilon N$ .

In questo tipo di descrizione capiamo che qualsiasi temperatura negativa risulti essere più *calda* di una qualsiasi temperatura positiva, nel senso che se i due corpi fossero messi in contatto termico il calore fluirebbe da quello con la temperatura negativa al corpo a temperatura positiva. Questa apparente contraddizione notazionale è dovuta all'uso in termodinamica della corrente scala di temperatura. Essendo una scelta arbitraria, se avessimo invece utilizzato una relazione per cui  $T = -\frac{1}{T}$  ritroveremmo una scala delle temperature più *intuitiva*.

Nella termodinamica a temperature negative termini fondamentali come *calore* e *lavoro* mantengono il loro significato, mentre certe formulazioni di alcune leggi della termodinamica richiedono qualche precisazione. Anche i termini *caldo* e *freddo*, come accennato precedentemente, non fanno più riferimento al valore algebrico della temperatura di un corpo ma invece devono far riferimento all'ipotetico contatto termico tra due sistemi e all'eventuale flusso di calore tra un corpo e l'altro. La scala delle temperature che otteniamo dalla più fredda alla più calda è rappresentata in Figura 4.4.

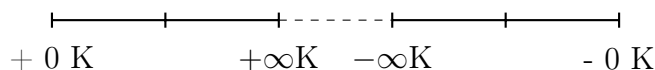


Figura 4.4: *Scala delle temperature in un sistema che ammette temperature negative secondo la definizione (4.1).*

Questo tipo di sistemi è in grado di assumere temperature di segni diversi ed è un sistema in grado di compiere cicli di trasformazioni termodinamiche. Sappiamo che a temperature positive e negative l'efficienza di un ciclo di Carnot si può valutare in base

ai calori scambiati e alle temperature coinvolte nel ciclo come:

$$\eta = 1 - \frac{Q_1}{Q_2} = 1 - \frac{T_1}{T_2}. \quad (4.11)$$

Verrebbe da pensare di poter compiere un ciclo tra due temperature di segno discorde e così ottenere un'efficienza  $\eta > 1$ . Questo è in realtà un processo impossibile poiché le due regioni di temperature sono separate da una regione di discontinuità, rendendo impossibile il passaggio da una zona all'altra tramite una trasformazione reversibile.

È possibile invece che, operando un ciclo tra due temperature negative, il rapporto  $\frac{T_2}{T_1} > 1$  rendendo l'efficienza  $\eta < 0$ . Significa che in un tale ciclo bisogna fornire lavoro al sistema per fargli assorbire calore dal serbatoio a temperatura maggiore. Ipotizzando un suo funzionamento inverso significa che in tale ciclo si produce lavoro per trasferire calore da un serbatoio freddo ad uno più caldo (Figura 4.5). Se tale calore trasferito è in grado di tornare al serbatoio a temperatura minore esisterebbe un apparato in grado di operare in un ciclo chiuso con il solo effetto di estrarre calore e di compiere una quantità di lavoro equivalente. L'esistenza di una tale macchina termica è in contraddizione con una delle convenzionali formulazioni della seconda legge della termodinamica, ma non in contraddizione con la sua corretta riformulazione nel campo delle temperature negative.

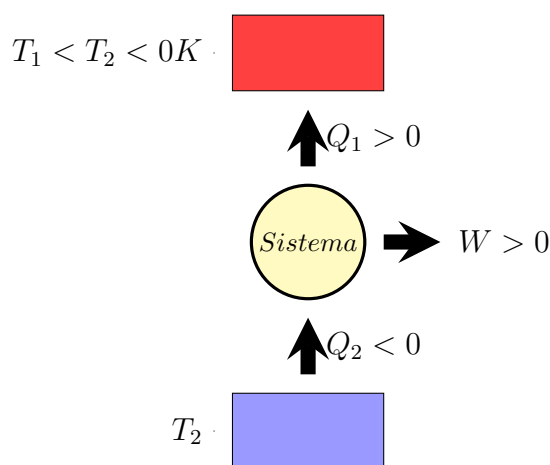


Figura 4.5: *Schema del funzionamento inverso di una ipotetica macchina termica che agisce tra due temperature negative tale che  $\eta < 0$ .*

Tale ipotetica macchina non sarebbe infatti in contraddizione con il principio della crescita dell'entropia. Come vediamo dalla Figura 4.2 estrarre energia da un corpo a temperatura negativa comporta una crescita di entropia. Dobbiamo notare infatti che in una tale macchina che assorbe calore da un serbatoio a temperatura minore e cede calore ad un serbatoio a temperatura maggiore l'efficienza (4.11) non solo sarebbe positiva ma anche minore dell'unità. Questo ciclo, come il suo corrispettivo a temperature negative,

quindi produce lavoro in una quantità minore del calore che ha assorbito. Quello che occorre fare è quindi chiarire la formulazione della seconda legge data da Kelvin-Planck: "*È impossibile creare un dispositivo, che operi in un ciclo chiuso, che non produca altro effetto che trasferire calore da un corpo più freddo ad uno più caldo*". Sarebbe più corretto dire che "*È impossibile creare un dispositivo, operante in un ciclo chiuso, che non produca altro effetto che: 1) estrarre calore da un serbatoio di calore a temperatura positiva e compiere una quantità equivalente di lavoro, 2) assorbire calore da un serbatoio a temperatura negativa in seguito all'applicazione di una quantità di lavoro corrispondente*". Capiamo così che la difficoltà di raffreddare un corpo a temperatura positiva è analoga alla difficoltà di riscaldare un corpo a temperatura negativa.

Analogamente è giusto modificare la formulazione del terzo principio della dinamica in "*È impossibile con qualsiasi procedimento finito, non importa quanto idealizzato, ridurre la temperatura di un sistema allo zero positivo della temperatura o aumentarla fino allo zero negativo*".

Dopo aver commentato quali siano le conseguenze del definire temperatura negative vediamo quali siano le richieste essenziali affinché un sistema ammetta questo tipo di descrizione. Ramsey ne individua tre in [9]:

1. Gli elementi del sistema termodinamico devono essere in equilibrio termodinamico tra di loro, in modo che il sistema possa in generale essere descritto tramite una temperatura.
2. Deve esserci un limite superiore alle possibili energie permesse al sistema.
3. Il sistema deve essere termicamente isolato da qualsiasi altro sistema che non soddisfi le due richieste precedenti. Il tempo di equilibrio termico del sistema deve essere piccolo rispetto al tempo in cui i due sistemi si scambiano energia.

La seconda condizione deve essere soddisfatta per permettere di raggiungere temperature negative con un'energia finita. Se infatti consideriamo  $W_m$  l'energia del  $m$ -esimo stato di un elemento del sistema il numero di elementi in quello stato è dato dal fattore di Boltzmann  $e^{-W_m/kT}$ . Per temperature negative questo fattore cresce esponenzialmente con il crescere di  $W_m$  rendendo gli stati a maggiore energia maggiormente occupati rispetto a quelli a minore energia. Senza un limite superiore all'energia non è possibile quindi raggiungere temperature negative. La grande maggioranza dei sistemi non gode di questa proprietà, per esempio non c'è limite superiore all'energia cinetica di un gas di molecole. Per questa ragione i sistemi a temperatura negativa si incontrano raramente.

Possiamo trovare alcune recenti evidenze sperimentali di sistemi termodinamici in stati equilibrati a temperature negative. Possiamo citare per esempio [3], in cui si descrive un sistema di bosoni ultrafreddi.

# Conclusioni

L'obiettivo di questa tesi è esplorare una differente definizione di entropia e verificare come si possa costruire una teoria della termodinamica coerente con i risultati classici su questa definizione alternativa.

Partendo dalla definizione di entropia di Shannon siamo prima riusciti a riprodurre i risultati riguardo ad un sistema noto come il gas perfetto e poi, grazie al principio di massimizzazione dell'entropia, abbiamo trovato un insieme di postulati che gettano le basi della termodinamica e permettono di derivare le leggi con cui siamo familiari. Ci siamo soffermati su questo principio, sulla sua potenza e la sua versatilità e abbiamo visto come possa essere conveniente.

Come ultima cosa ci siamo avvicinati al concetto di temperatura negativa e alle precisazioni che è necessario compiere per renderle uno strumento utile nei casi in cui è possibile definirle.

Abbiamo visto quindi come la discussione sui temi fondanti della termodinamica statistica, che saremmo portati a pensare che siano consolidati o risolti, sia tutt'altro che sopita e scontata.

# Bibliografia

- [1] A. Ben-Naim, *A Farewell to Entropy: Statistical Thermodynamics Based on Information*, World Scientific, 2008.
- [2] L. Boltzmann (1866). *Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie*. Wiener Berichte. 53: 195–220.
- [3] S. Braun, J. P. Ronzheimer, M. Schreiber, S. S. Hodgman, T. Rom, I. Bloch, U. Schneider, *Negative Absolute Temperature for Motional Degrees of Freedom*, Sci. 04/01/2013.
- [4] H.B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed.; Wiley: New York, NY, USA, 1985.
- [5] J. L. Doob, *Statistical Estimation*, Trans. Amer. Math. Soc. 39 (1936).
- [6] R. A. Fisher, *Theory of Statistical Estimation*, Proc. Cambridge Phil. Soc. 22, 700, 1925.
- [7] J. W. Gibbs, *Elementary Principles of Statistical Mechanics* (Yale University Press, New Haven, 1902) reprinted by Dover, New York, 1960.
- [8] E. T. Jaynes, *Information Theory and Statistical Mechanics*, *Phys. Rev.* Vol. 106, No. 4, 1957.
- [9] N. T. Ramsey, *Thermodynamics and Statistical Mechanics at Negative Absolute Temperature*, *Phys. Rev.* Vol. 103, No. 1, 1956.
- [10] C.E. Shannon, *A Mathematical Theory of Communication*, Bell Syst. Tech. J., vol. 27, luglio-ottobre 1948.
- [11] C.E. Shannon, *Sci. Am.*, vol. 225, (p. 180), 1971.
- [12] R.H. Swendsen, *An Introduction to Statistical Mechanics and Thermodynamics*, Oxford University Press, 2012.

- [13] R.H. Swendsen, *Thermodynamics, Statistical Mechanics and Entropy*. Entropy (2017), 19, 603.
- [14] R.H. Swendsen, *Resolving the debate about proposed expressions for the classical entropy*, (2017) arXiv:1702.05810v5